

# Multiscale Modeling and Multirate Time-Integration of Field/Circuit Coupled Problems

## Dissertation

zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften (Dr. rer. nat.)

dem Fachbereich C - Mathematik und Naturwissenschaften -  
der Bergischen Universität Wuppertal vorgelegt

und/and

submitted in fulfillment of the requirements for the degree of a

Doctor in de Wetenschappen: Fysica

at the Faculteit Wetenschappen (Fysica)  
of the Katholieke Universiteit Leuven

von/by

Sebastian Schöps, M.Sc.

born 1980-11-24 in Hilden

Supervisor      Prof. Dr. Michael Günther  
Co-Supervisor   Prof. Dr. Herbert De Gerssem  
Co-Supervisor   Dr. Andreas Bartel

May, 2011

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20110718-110524-8

[<http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20110718-110524-8>]

---

# Preface

This treatise was written during my employment at the Chair of Angewandte Mathematik und Numerische Analysis of the Bergische Universität Wuppertal and my post-graduate research at the Wave Propagation and Signal Processing Research Group of the Katholieke Universiteit Leuven funded by a Jahresstipendium für Doktoranden from the DAAD (German Academic Exchange Service). At the Bergische Universität Wuppertal I have been a member of the Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM). Correspondingly this thesis consists of parts devoted to modeling, analysis and computational mathematics.

In concluding this treatise, I would like to thank all those without whose help this thesis would not have been possible.

First of all, I want to thank Prof. Dr. Michael Günther for his supervision and the opportunity to work at the Chair of Applied Mathematics in Wuppertal. I thank Prof. Dr.-Ir. Herbert De Gersem for his supervision, scientific guidance who was a source of motivation and many fruitful discussions. Herbert De Gersem introduced me to the electrical engineering community and it was only due to his support that I was able to understand and tackle the engineering problems of this treatise. Finally I thank Dr. Andreas Bartel for his scientific supervision, his support of my research and his rigorous mathematical review of all results.

I would like to thank Prof. Dr. Markus Clemens, Prof. Dr. Karl Meerbergen and Prof. Dr. Caren Tischendorf who agreed to be members of my doctoral committee.

During my doctoral studies I collaborated with external researchers. The topics ranged from general discussions of mathematical ideas to the solution of concrete problems. In this context I am particularly grateful for the cooperations with Sascha Baumanns, Dr. Markus Brunk and Prof. Dr. Markus Clemens.

I thank my colleagues from Wuppertal and Leuven, present and former, for many mathematical discussions but also for being friends who it was always a pleasure to work with. I would like to name just a few. In particular I want to thank Dr. Markus Brunk and Dr. Michael Striebel from Wuppertal and Bart Van Damme and Stijn Rebry from Leuven/Kortrijk. Further thanks go to my Italian colleagues Dr. Giuseppe Alí, Dr. Massimiliano Cuplo and Dr. Carmelo Scordia.

I would also like to thank Alun Davies for proof reading this thesis.

Finally, I would like to thank my parents Brigitte and Fred Schöps, my girlfriend Julia Heuer and my friends for giving me all their love and support.



---

# Contents

<b>List of Figures</b>	<b>VII</b>
<b>List of Symbols</b>	<b>IX</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Multiscale and Multirate . . . . .	1
1.2 Related Works . . . . .	2
1.3 Overview . . . . .	3
<b>2 Electromagnetism</b>	<b>5</b>
2.1 Maxwell's Equations . . . . .	5
2.1.1 Boundary and Initial Conditions . . . . .	6
2.1.2 Partitioning into Regions and Materials . . . . .	8
2.2 Space Discretization . . . . .	9
2.2.1 Maxwell's Grid Equations . . . . .	10
2.2.2 Semi-discrete Problem Formulation . . . . .	12
2.2.3 Gauging . . . . .	13
2.3 Conclusions . . . . .	16
<b>3 Multiscale Device Models</b>	<b>17</b>
3.1 Electric Network . . . . .	18
3.2 Magnetoquasistatic Conductor Model . . . . .	20
3.2.1 More Conductor Models . . . . .	24
3.3 Semiconductor Device Model . . . . .	26
3.4 Conclusions . . . . .	28
<b>4 DAE-Index Analysis</b>	<b>29</b>
4.1 Tractability Index . . . . .	29
4.2 Index Analysis for the Field/Circuit Problem . . . . .	30
4.2.1 Index-1 Conditions . . . . .	32
4.2.2 Index-2 Conditions . . . . .	34
4.3 Conclusions . . . . .	38
<b>5 Multirate Methods</b>	<b>39</b>
5.1 Single Rate DAE Time-integration . . . . .	40
5.2 Multirate Bypassing of MQS Schur Complements . . . . .	41
5.2.1 Bypassing as Multirate Time-Integration . . . . .	44
5.2.2 Conclusions . . . . .	45
5.3 Multirate Cosimulation . . . . .	45
5.3.1 Abstract DAE-DAE Coupling . . . . .	47

5.3.2	Dynamic Iteration . . . . .	49
5.3.3	Recursion Estimate . . . . .	52
5.3.4	Convergence and Stability . . . . .	58
5.3.5	Application to Field/Circuit Coupling . . . . .	63
5.3.6	Application to Semiconductor/Circuit Coupling . . . . .	72
5.4	Domain Substructuring in MQS Devices . . . . .	78
5.4.1	Linear/Nonlinear Decomposition . . . . .	79
5.4.2	Convergence and Eigenvalues . . . . .	80
5.4.3	Inner Solvers . . . . .	82
5.4.4	Conclusions . . . . .	83
<b>6</b>	<b>Numerical Examples</b>	<b>84</b>
6.1	DAE-Index in Applications . . . . .	85
6.2	Multirate Bypassing . . . . .	87
6.2.1	Conclusion . . . . .	89
6.3	Cosimulation . . . . .	90
6.3.1	Field/Circuit Problem . . . . .	91
6.3.2	Semiconductor/Circuit Problem . . . . .	95
6.4	Domain Substructuring of a Transformer . . . . .	99
6.4.1	2D Test Problem . . . . .	100
6.4.2	3D Test Problem . . . . .	100
6.4.3	Conclusions . . . . .	102
<b>7</b>	<b>Conclusions</b>	<b>103</b>
7.1	Outlook . . . . .	104
<b>A</b>	<b>Discretization Properties</b>	<b>105</b>
A.1	Discrete Operators . . . . .	105
A.1.1	FIT Numbering Scheme . . . . .	105
A.1.2	Phantom Objects on Finite Grids . . . . .	107
A.1.3	Degrees of Freedom in FIT . . . . .	110
A.2	Material Matrices . . . . .	111
A.2.1	Element Matrices . . . . .	112
A.2.2	Global Matrix Assembly . . . . .	113
A.2.3	Nonlinearity . . . . .	113
A.3	Differential Curl-Curl Matrix . . . . .	114
	<b>Bibliography</b>	<b>116</b>

---

# List of Figures

1.1	Introduction to multirate . . . . .	2
1.2	Introduction to cosimulation . . . . .	3
2.1	Boundary conditions . . . . .	7
2.2	Maxwell's house . . . . .	9
2.3	Primary and dual cells . . . . .	10
2.4	Regularized curl-curl matrix . . . . .	13
2.5	Shift of eigenvalues . . . . .	14
3.1	Multiscale modeling . . . . .	17
3.2	MQS device types . . . . .	20
3.3	MQS device coupling . . . . .	21
3.4	Full and sparse coupling . . . . .	22
3.5	Smearing . . . . .	24
5.1	Multirate behavior. . . . .	39
5.2	Field/circuit time scales . . . . .	44
5.3	Bypassing . . . . .	45
5.4	Dynamic iteration . . . . .	47
5.5	Gauß-Seidel scheme . . . . .	51
5.6	Contraction of waveforms . . . . .	54
5.7	Lady Windermere's Fan . . . . .	59
5.8	Field/circuit coupling interfaces . . . . .	65
5.9	Festoon-like behavior of sources. . . . .	66
5.10	PWM frequency spectrum . . . . .	68
5.11	Spline approximation of waveforms . . . . .	69
5.12	Semiconductor/circuit interfaces . . . . .	73
5.13	Domain substructuring . . . . .	79
5.14	Eigenvalues and PCG convergence . . . . .	82
6.1	Flow chart of software packages . . . . .	84
6.2	DAE-index example circuits . . . . .	85
6.3	Inductor example . . . . .	86
6.4	Index-1 vs. index-2 errors . . . . .	87
6.5	Bypassing example . . . . .	88
6.6	Field/circuit cosimulation example . . . . .	90
6.7	Convergence of field/circuit cosimulation . . . . .	91
6.8	Contraction of field/circuit cosimulation . . . . .	92
6.9	Multirate cosimulation example . . . . .	93
6.10	Convergence of multirate cosimulation . . . . .	94

6.11	Semiconductor/circuit cosimulation example . . . . .	95
6.12	Contraction of semiconductor/circuit cosimulation. . . . .	96
6.13	Contraction and Lipschitz constants . . . . .	97
6.14	Convergence of semiconductor/circuit cosimulation . . . . .	98
6.15	Domain substructuring example . . . . .	99
6.16	Eigenvalues in substructuring example . . . . .	100
A.1	Cells and primary curl operator. . . . .	106
A.2	Primary FIT grid of dimensions $2 \times 2 \times 2$ . . . . .	107
A.3	The number of total and non-phantom objects in Maxwell's house . . . . .	109
A.4	Local elements and numbering in FIT case. . . . .	112

---

# List of Symbols

## Electromagnetism

$\vec{A}$	magnetic vector potential . . . . .	5
$\vec{B}$	magnetic flux density . . . . .	5
$\vec{D}$	electric flux density . . . . .	5
$\vec{E}$	electric field strength . . . . .	5
$\varepsilon$	electric permittivity . . . . .	5
$\Gamma$	boundary of the computational domain $\Omega$ . . . . .	8
$\vec{H}$	magnetic field strength . . . . .	5
$\vec{J}$	electric current density . . . . .	5
$\vec{n}$	outward normal unit vector (at boundary) . . . . .	7
$\nu$	magnetic reluctivity (inverse permeability) . . . . .	5
$\Omega$	computational domain . . . . .	8
$\Omega^{(r)}$	region, i.e., $r$ -th subdomain of $\Omega$ . . . . .	8
$\varphi$	electric scalar potential . . . . .	5
$\rho$	electric charge density . . . . .	5
$\sigma$	electric conductivity . . . . .	5

## Discrete Electromagnetism

$\hat{\mathbf{a}}$	line-integrated magnetic vector potential . . . . .	11
$\hat{\mathbf{b}}$	area-integrated magnetic flux density . . . . .	10
$\mathbf{C}$	curl matrix ( $\tilde{\mathbf{C}} = \mathbf{C}^\top$ is the dual curl matrix) . . . . .	10
$\hat{\mathbf{d}}$	area-integrated electric flux density . . . . .	10
$\mathbf{S}$	divergence matrix ( $\tilde{\mathbf{S}}$ is the dual divergence matrix) . . . . .	10
$\hat{\mathbf{e}}$	line-integrated electric field strength . . . . .	10
$\hat{\mathbf{h}}$	line-integrated magnetic field strength . . . . .	10
$\hat{\mathbf{j}}$	area-integrated electric current density . . . . .	10
$\mathbf{k}_\nu$	curl-curl matrix (chord) . . . . .	14
$\mathbf{K}_\nu$	differential curl-curl matrix . . . . .	14
$\mathbf{M}_\varepsilon$	electric permittivity matrix . . . . .	11
$\mathbf{M}_\nu$	(nonlinear) magnetic reluctivity matrix . . . . .	11
$\mathbf{M}_\sigma$	electric conductivity matrix . . . . .	11
$\Omega_e$	element for discretization (subdomain of $\Omega$ ) . . . . .	10
$\Phi$	discrete electric scalar potential . . . . .	11
$\mathbf{Q}_\sigma$	constant projector onto $\text{Ker } \mathbf{M}_\sigma$ . . . . .	15
$\mathbf{P}_\sigma$	complementary projector $\mathbf{P}_\sigma = \mathbf{I} - \mathbf{Q}_\sigma$ . . . . .	15
$\mathbf{q}$	volume-integrated electric charge density . . . . .	10

$\mathbf{Z}_\sigma$	regularization matrix (grad-div gauge) . . . . .	14
<b>Electric Network</b>		
$\mathbf{A}_\star$	(reduced) circuit incidence matrix . . . . .	18
$\mathbf{C}$	Jacobian of the capacitance relation (matrix of lumped capacitances) . .	19
$\mathbf{g}_R$	constitutive relation for resistances . . . . .	18
$\mathbf{i}_D$	currents through semiconductor devices . . . . .	19
$\mathbf{i}_L$	current through inductances . . . . .	18
$\mathbf{i}_M$	currents through MQS field devices . . . . .	19
$\mathbf{i}_s$	constitutive relation for current sources . . . . .	18
$\mathbf{i}_V$	current through voltage sources . . . . .	18
$\mathbf{L}$	Jacobian of the inductance relation (matrix of lumped inductances) . . .	19
$\phi$	magnetic flux . . . . .	18
$\phi_L$	constitutive relation for inductances . . . . .	18
$\mathbf{q}$	electric charge . . . . .	18
$\mathbf{q}_C$	constitutive relation for capacitances . . . . .	18
$\mathbf{G}$	Jacobian of the resistance relation (matrix of lumped conductances) . .	19
$\mathbf{u}$	nodal potentials . . . . .	18
$\mathbf{v}_s$	constitutive relation for voltage sources . . . . .	18
<b>MQS device</b>		
$\gamma$	arbitrary path through the conductor part of the MQS device . . . . .	21
$\mathbf{i}_M$	currents through the MQS devices . . . . .	24
$\mathbf{L}_{\text{str}}$	nonlinear inductance matrix (for stranded conductors) . . . . .	25
$\mathbf{M}_{\bar{\sigma}}$	modified electric conductivity matrix (for stranded conductors) . . . . .	25
$\Omega^{(m)}$	conductor region $\Omega^{(m)} \subset \Omega^{(M)}$ . . . . .	20
$\mathbf{X}_M$	full coupling matrix . . . . .	23
$\bar{\mathbf{X}}_M$	sparse coupling matrix . . . . .	22
$\mathbf{R}_{\text{sol}}$	diagonal matrix of dc resistances (for solid conductors) . . . . .	25
$\mathbf{R}_{\text{str}}$	diagonal matrix of dc resistances (for stranded conductors) . . . . .	25
$\mathbf{v}_M$	voltage drops at the MQS devices . . . . .	24
<b>Semiconductor</b>		
$C$	doping concentration . . . . .	27
$\mathbf{C}_D$	extracted capacitance modeling the displacement current . . . . .	27
$\mathbf{i}_D$	total current through semiconductor . . . . .	27
$\mathbf{i}_{SD}$	current through semiconductor without displacement current . . . . .	27
$\mathbf{M}_\star$	drift diffusion mass matrices (where $\star \in \{n, p\}$ ) . . . . .	27
$\mathbf{K}_\star$	drift diffusion stiffness matrices (where $\star \in \{n, p\}$ ) . . . . .	27
$\mu_\star$	mobility parameters (where $\star \in \{n, p\}$ ) . . . . .	27
$n$	electron density . . . . .	27
$\mathbf{n}$	discrete electron density . . . . .	27
$\Omega^{(D)}$	semiconductor domain . . . . .	26

$p$	hole density . . . . .	27
$\mathbf{p}$	discrete hole density . . . . .	27
$q$	elementary charge . . . . .	27
$\mathbf{r}_\star$	discrete recombination term (where $\star \in \{n, p\}$ ) . . . . .	27
$R$	generation-recombination term . . . . .	27
$U_T$	thermal voltage . . . . .	27

### Time-integration and Cosimulation

$\alpha_n$	contraction factor for the $n$ -th time window . . . . .	53
$d$	distance from the solution . . . . .	52
$\Delta_{\star,n}^k$	difference of two waveforms, $n$ -th window, $k$ -th iteration . . . . .	53
$\delta_{\star,n}^k$	maximum of $\delta_{\star,n}^k$ on the $n$ -th time window . . . . .	53
$\mathbf{F}$	splitting function (differential) . . . . .	50
$\mathbf{f}$	right-hand-side function (differential) . . . . .	47
$\mathbf{G}$	splitting function (algebraic) . . . . .	50
$\mathbf{g}$	right-hand-side function (algebraic) . . . . .	47
$H_n$	time window size ( $n$ -th window) . . . . .	49
$h_n$	time step size ( $n$ -th step) . . . . .	40
$\mathbf{J}_n^{(i)}$	system matrix in the $i$ -th Newton iteration at time step $n$ . . . . .	41
$\Phi_n$	extrapolation operator ( $n$ -th window) . . . . .	49
$\Psi_n$	solution operator operator ( $n$ -th window) . . . . .	50
$T_n$	$n$ -th synchronization time point . . . . .	49
$t_n$	$n$ -th time point on the time-discretization grid . . . . .	40
$\mathcal{U}_{d,n}$	neighborhood of the solution on $n$ -th time window . . . . .	52
$\mathbf{y}$	differential variables . . . . .	47
$\tilde{\mathbf{y}}$	approximation of differential variables . . . . .	49
$\mathbf{z}$	algebraic variables . . . . .	47
$\tilde{\mathbf{z}}$	approximation of algebraic variables . . . . .	49
$\zeta$	function defined implicitly by algebraic constraint . . . . .	54

### Multirate Applications

$\mathbf{E}_n$	estimated energy level . . . . .	44
$\mathbf{g}_D$	nonlinear semiconductor resistance . . . . .	73
$\mathbf{L}_h^{(i)}$	generalized inductance matrix (step size $h$ , iteration $i$ ) . . . . .	43
$L_C$	maximum Lipschitz constant (w.r.t. $\mathbf{z}_D$ ) . . . . .	76
$L_D$	maximum Lipschitz constant (w.r.t. $\mathbf{y}_C$ ) . . . . .	76
$\mathbf{L}_M$	extracted nonlinear inductance from field model . . . . .	65
$\phi_M$	magnetic flux through field model . . . . .	65
$\psi$	time integrated voltage (energy estimation) . . . . .	69
$\mathbf{g}_M$	dc resistance of field model . . . . .	65

### Index Analysis

$\mathbf{A}$	mass matrix . . . . .	30
--------------	-----------------------	----

<b>b</b>	coefficient function (stiffness) . . . . .	30
<b>b'<sub>x</sub></b>	linearized stiffness matrix . . . . .	32
<b>d</b>	coefficient function (mass) . . . . .	30
<b>G<sub>0</sub></b>	linearized mass matrix . . . . .	31
<b>H<sub>11</sub></b>	Schur complement matrix . . . . .	34
<b>R</b>	projector for properly stated leading term . . . . .	31
<b>x</b>	vector of unknowns (degrees of freedom) . . . . .	30

---

# 1 Introduction

Faster innovation cycles and the increasing complexity of electrical devices necessitate their simulation. Nowadays engineers design new devices on desktop computers thus reducing the construction of prototypes to a minimum.

Such virtual prototyping consists of several steps: the engineer draws the geometry of a particular device (e.g. a transformer) using a computer aided design tool. He assigns material data and boundary conditions. Then he carries out a simulation to verify that his device is adequate for its purpose. For example the engineer might be interested in the behavior of a transformer after connecting it to the electric power network. The power network supplies energy and this excites the device. Now, to predict the performance of it, the physical quantities must be computed by solving differential equations for the given excitation, e.g., the time-varying currents and voltages. Typically, the corresponding simulation in time domain ('time-integration') is computationally very expensive because millions of equations must be solved repeatedly at various time points.

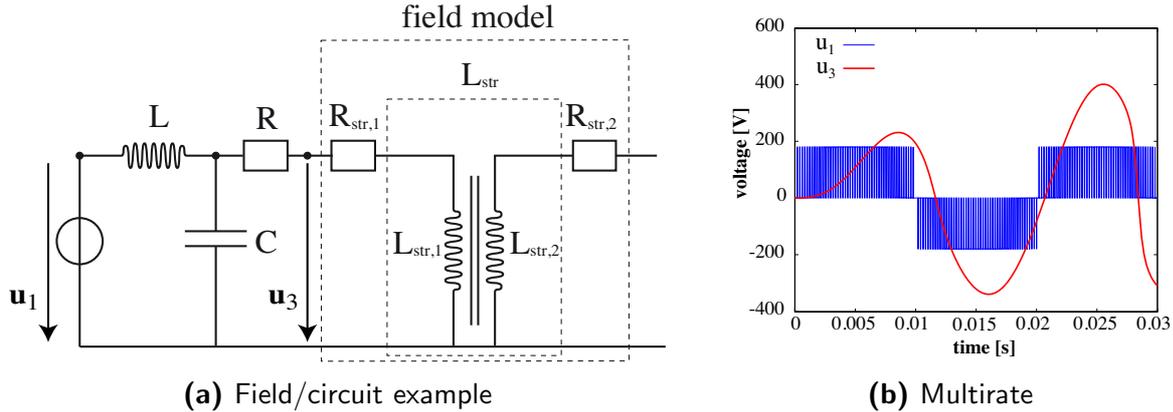
The main goal of this treatise is to develop methods speeding up those simulations. The idea behind it is trivial: different parts of a given problem are not equally important. For example the simulation of a power plant is not necessary to predict the behavior of a transformer plugged into an electric socket. Thus some effects can be neglected or simulated by simpler models, while other problem parts are crucial and thus need models of higher quality. Obviously, higher quality automatically implies higher computational costs.

This treatise deals with the coupling and simulation of device models of different quality and scales, in other words multirate methods are applied to a hierarchy of multiscale models.

## 1.1 Multiscale and Multirate

Coupled multiscale problems are a challenging task for time integrators: the intrinsic time rates of subsystems can differ by several orders of magnitude. Some effects are very fast, while other are very slow. Furthermore the type of the equations can be different, i.e., their symmetry and definiteness. For example Figure 1.1a shows an electric circuit modeled by differential-algebraic equations (DAEs), where a part is refined by a magnetoquasistatic field model using partial differential equations (PDEs). The simulation of the resulting PDAE model is often denoted as mixed mode simulation. In the example here, the voltages are only pulsed in the circuit part but slow and sinusoidal in the field part, Figure 1.1b. Due to the different approaches, i.e., network and space discretization, the underlying equations are different in shape, the circuit system is non-symmetric and typically tackled by direct solvers, while the field is symmetric and solved by iterative methods.

Standard time-integration is inefficient for those problems and thus tailor-made approaches are presented in this thesis: bypassing and cosimulation. Cosimulation in partic-



**Figure 1.1:** Introduction to multirate. Example configuration (a) and multirate behavior in example (b), see [118].

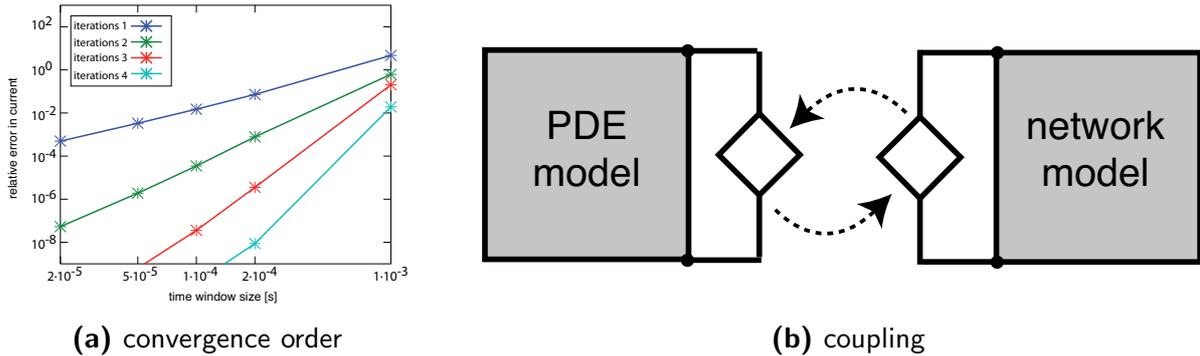
ular allows for multirate and multimethod, but the presence of algebraic constraints can handicap higher-order integration and may cause divergence. In this treatise it will be shown that the stability and convergence of a cosimulation can be verified by analyzing the coupling interface. The design of the coupling interface is crucial: it determines the convergence speed and may or may not support multirate time-integration.

In problems with different time-scales the representation of the slow subsystem by a reduced order model for as many time steps as possible is advantageous, because the weaker the coupling, the more efficient the cosimulation. For example in most magnetoquasistatic applications a lumped linear inductance model extracted from the field equations is sufficient to reflect the device characteristics for a reasonable time span. The nonlinearity of the inductance due to magnetic saturation can be slowly resolved in time because of the inertia of energy transport. Thus this treatise proposes to solve a network problem consisting of entirely lumped models and fit the parameters on the fly by an outer iteration with PDE device models, Fig. 1.2b.

For that the cosimulation approach is adapted and sufficient conditions for stability and convergence are derived and the convergence rate is well understood. The parameter coupling approach is rather general and can be applied to any DAE-DAE problem and thus a second example from electrical engineering is given: an electric circuit coupled to a semiconductor model.

## 1.2 Related Works

Since [70] the strongly coupled time-transient simulation of the field/circuit problem has become an established methodology, [82, 132]. Different approaches have been combined: the field subsystems were discretized e.g. by finite elements or the finite integration technique and the circuit has been represented by different formulations, e.g. loop or nodal analysis, [137, 46, 15]. For the coupling different conductor models were proposed, e.g. the stranded and solid conductors, [14, 49, 40]. Some of those results are already available in text books, e.g. the 2D coupling for stranded conductors is discussed in [111]. Furthermore the properties of the coupled system have been studied, [131]. Higher order time-integration, [101] with adaptive step size control, [28] and the different fixed point



**Figure 1.2:** Introduction to cosimulation. Order in dependence of window size (a) and cosimulation (b).

iterations to solve the nonlinear system (Newton-Raphson, successive substitution etc.), are state of the art, [102], [100]. There has been extensive research on linear solvers, either direct or iterative, in particular conjugate gradients, [110] and multigrid, [95], [107].

Domain decomposition and in particular substructuring methods were applied to both the field and circuit problems to reduce their complexities, e.g. [3, 130]. Similarly the elimination of the circuit equations by Schur complements was proposed in [84] and vice versa the elimination of the field equations in [134].

At the same time cosimulations based on source coupling were developed, [14]. The more advanced approaches [80, 143] deployed reduced inductances for the inductive coupling and this was further enhanced in [85, 86]. They applied several time-steps in the circuit subsystem, while only one time step in the field part. This was a first multirate technique for the field/circuit problem, but it suffers from bad stability properties, because each subsystem was only solved using once without error control. Dynamic iteration schemes are superior when they solve the subproblems iteratively and exchange coupling information in each sweep, [88, 141]. This improves the overall convergence order of the method, see Fig. 1.2a. The convergence of these schemes is well understood in the case of ordinary differential equations, [97, 27]. Additional algebraic constraints are known to endanger convergence, [76, 5]. Dynamic iteration has been traditionally applied in circuit simulation, but was also applied to multiscale problems, e.g., for the coupling of semiconductors and circuits, [62, 7, 51].

## 1.3 Overview

The focus of this thesis is the efficient time-transient simulation of device/circuit coupled problems by exploiting properties of the multiscale models (time scales, symmetries etc). Several methods are proposed in this thesis, but the cosimulation is the most natural choice. One has the opportunity to freely choose different methods for each subproblem. On the other hand cosimulation requires more analysis of the problem formulation and the interfaces.

The structure of the work is as follows: the next chapter introduces the electromagnetic field problem. The quasistatic approximations to Maxwell's equations are recapitulated and the space discretization is briefly discussed. Maxwell's equations are the foundation

of the multiscale models that are derived in Chapter 3. It is subdivided into Section 3.1 on electric network modeling, Section 3.2 on magnetoquasistatic devices and finally Section 3.3 that recapitulates the drift-diffusion model of a semiconductor. The results in those Sections are standard, although Section 3.2 contains some new aspects from [118, 8].

Chapter 5 is the core of this thesis. It starts with a brief summary of numerical time-integration (Section 5.1). In the following Section 4 the time-discretization properties of the field/circuit coupled problem are analyzed using the differential-algebraic index concept. The section is based on a recent publication compiled in cooperation with Sascha Baumanns and Andreas Bartel, [8], but also includes results from material that was presented at SCEE 2008, [117]. Section 5.2 discusses a new multirate technique based on standard time-integration: the field subsystem is reduced by Schur complements and updated according to its energy-level. This approach was presented at SCEE 2010 and was devised in collaboration with Andreas Bartel and Herbert De Gersem, [116]. The cosimulation stability and convergence theory for general index-1 DAEs is presented in Section 5.3. It utilizes results from [7], but features the multi-window propagation of splitting errors in dynamic iteration schemes, similarly to [5]. Furthermore important special cases are studied with more rigor. The dynamic iteration schemes are applied to the field/circuit (Section 5.3.5) and semiconductor/circuit coupled problems (Section 5.3.6). The convergence theory and its application are the result of a collaboration with Andreas Bartel, Markus Brunk and Michael Günther, [9]. The multirate coupling in Section 5.3.5 is developed together with Herbert De Gersem and was presented at Compumag 2009 and EPNC 2010, [118]. Furthermore some of the convergence studies for the semiconductor case in Section 5.3.6 were presented at SCEE 2010, [1]. In Section 5.4 a domain substructuring approach is discussed that exploits the particular structure, i.e., linear/nonlinear and conductive/nonconductive domains, of the eddy current problem. The method was presented at IGTE 2010 in collaboration with Markus Clemens, Andreas Bartel and Herbert De Gersem, [36].

Chapter 6 shows the significance and practicability of the various results and methods: the index-2 time-integration of a magnetostatic device model (Section 6.1), the multirate bypassing of a magnetoquasistatic transformer model (Section 6.2), the cosimulation of both MQS and semiconductor devices coupled to circuits (Section 6.3.1 and 6.3.2), and finally the domain substructuring approach is applied to 2D and 3D transformer models (Section 6.4). With the exception of the semiconductor/circuit cosimulation all the experiments were performed using the software FIDES (within Octave/CoMSON DP). The numerical results for the semiconductor problem were obtained by Markus Brunk using Matlab, [9]. The thesis closes with conclusions in Chapter 7.

---

## 2 Electromagnetism

### 2.1 Maxwell's Equations

Electromagnetic phenomena are described on the macroscopic level by Maxwell's partial differential equations, [93]

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad \nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{J}, \quad \nabla \cdot \vec{D} = \rho, \quad \nabla \cdot \vec{B} = 0, \quad (2.1)$$

where  $\vec{E} = \vec{E}(\vec{r}, t)$  is the electric field strength,  $\vec{B} = \vec{B}(\vec{r}, t)$  the magnetic flux density,  $\vec{H} = \vec{H}(\vec{r}, t)$  the magnetic field strength,  $\vec{D} = \vec{D}(\vec{r}, t)$  the electric flux density,  $\rho = \rho(\vec{r}, t)$  the electric charge density and  $\vec{J} = \vec{J}(\vec{r}, t)$  is the electric current density. All quantities depend on space  $\vec{r} \in \Omega$  and time  $t \in \mathcal{I}$ . These equations are related to each other by the additional material relations

$$\vec{D} = \vec{\epsilon} \vec{E}, \quad \vec{J} = \vec{\sigma} \vec{E}, \quad \vec{H} = \vec{\nu} \vec{B}, \quad (2.2)$$

where the permittivity  $\vec{\epsilon} = \vec{\epsilon}(\vec{r})$ , conductivity  $\vec{\sigma} = \vec{\sigma}(\vec{r})$  and reluctivity (inverse permeability)  $\vec{\nu} = \vec{\nu}(\vec{r}, \|\vec{B}\|^2)$  are rank-2 tensors. All field quantities can be expressed by the magnetic vector potential (MVP)  $\vec{A} : \mathcal{I} \times \Omega \rightarrow \mathbb{R}^3$  and its integration constant, the electric scalar potential  $\varphi : \mathcal{I} \times \Omega \rightarrow \mathbb{R}$

$$\vec{B} = \nabla \times \vec{A} \quad \text{and} \quad \vec{E} = -\frac{\partial \vec{A}}{\partial t} - \nabla \varphi. \quad (2.3)$$

The magnetic flux density  $\vec{B}$  defines the magnetic vector potential  $\vec{A}$  only up to a gradient field. For a unique solution an additional gauging condition is required, e.g. Coulomb's gauge  $\nabla \cdot \vec{A} = 0$ .

The engineering applications in this treatise are either in the high or low frequency regime. This allows us to consider approximations to Maxwell's full set of equations for special frequency ranges, [71].

**Assumption 2.1** (Quasistatics). *We assume that the problems considered in this treatise are either*

- i) magnetoquasistatic (MQS), i.e.,  $\frac{\partial \vec{D}}{\partial t} = 0$*
- ii) electroquasistatic (EQS), i.e.,  $\frac{\partial \vec{B}}{\partial t} = 0$*
- iii) static, i.e., both assumptions i) and ii)*

Assumption 2.1 i) and inserting Maxwell's equations into each other, starting from Ampère's law, i.e., the second equation of (2.1), gives the *magnetoquasistatic* approximation, [71]

$$\vec{\sigma} \frac{\partial \vec{A}}{\partial t} + \nabla \times \left( \vec{\nu} \nabla \times \vec{A} \right) = -\vec{\sigma} \nabla \varphi. \quad (2.4)$$

This curl-curl equation is of parabolic type, but if we further disregard changes in the magnetic field (Ass. 2.1 iii), we end up with the *magnetostatic* approximation

$$\nabla \times \left( \vec{\nu} \nabla \times \vec{A} \right) = -\vec{\sigma} \nabla \varphi \quad (2.5)$$

which is an elliptic problem. From the electric permittivity material law and Ohm's Law, i.e., the first two material relations of (2.2) and using the assumption of statics (Ass. 2.1 iii) one deduces the electrostatic approximations

$$-\nabla \cdot \left( \vec{\varepsilon} \nabla \varphi \right) = \rho \quad \text{and} \quad -\nabla \cdot \left( \vec{\sigma} \nabla \varphi \right) = 0, \quad (2.6)$$

that are elliptic equations of Poisson's type, [71]. The formulations above are not the only possible choice. There are other formulations (e.g. [18]) based on various quantities. Those are not used in the following.

### 2.1.1 Boundary and Initial Conditions

Maxwell's Equations (2.1)-(2.2) are defined on an infinite space. In our device simulations, we are only interested in the solution on a finite domain whose center is given by a device:

**Assumption 2.2** (Computational domain). *The computational domain  $\Omega$  is a simply connected polyhedral domain with the boundary  $\Gamma = \partial\bar{\Omega}$ .*

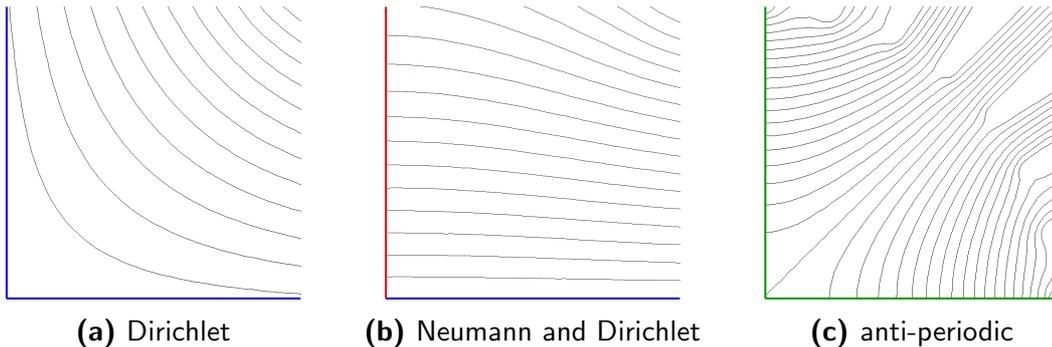
This truncation is reasonable for our parabolic and elliptic PDEs if the boundary is sufficiently far from the center, i.e., the device. This can be enforced by extending the computational domain using a sufficiently large region of air around the device. In MQS problems, the magnitude of the magnetic flux decays rapidly outside of the device towards the boundary. The rule of thumb commonly given in the literature recommends a distance from the center to the outer boundary to be at least five times the radius of the device, [30].

The boundary conditions (BC) are given at the geometrical borders  $\Gamma$  of our truncated problem and they model the behavior in the outside world. Furthermore, they can be used to model symmetries.

**Assumption 2.3** (Boundary conditions). *The boundary is either of Dirichlet (dir), Neumann (neu) or anti-periodic (anti) type:*

$$\Gamma = \Gamma_{\text{dir}} \dot{\cup} \Gamma_{\text{neu}} \dot{\cup} \Gamma_{\text{anti}}.$$

The choice of the BCs depends on several restrictions: on one hand at least one boundary must be defined as a Dirichlet type to yield a uniquely solvable problem, [105]. On the



**Figure 2.1:** Boundary conditions and flux lines. (a) Dirichlet condition imposed on left and bottom boundary, (b) Neumann condition on left and Dirichlet on bottom boundary, (c) anti-periodic conditions on both boundaries.

other hand the circuit coupling, as discussed in Section 3.2, adds further constraints. In the case of excitation by source terms, the boundaries are commonly of Dirichlet type (unless symmetries are exploited), while a coupling via boundary conditions requires more sophisticated assumptions, [75]. In the magnetic vector potential formulation (2.4) the conditions read thus:

**Definition 2.1** (Dirichlet boundary). Let  $\Gamma_{\text{dir}}$  denote the Dirichlet boundary and  $\vec{n}$  its outward normal unit vector then

$$\vec{A}(\vec{r}) \times \vec{n} = \vec{A}_{\text{dir}} \quad \text{with} \quad \vec{r} \in \Gamma_{\text{dir}}$$

sets the tangent components of the magnetic vector potential at the boundary to the prescribed value  $\vec{A}_{\text{dir}}$ . In the case of  $\vec{A}_{\text{dir}} = 0$  the condition is called homogeneous Dirichlet BC.

The condition is called *electric BC*, often also described as “flux wall” or “current gate”. It does not allow the magnetic flux to pass through the border (the flux lines stay parallel to the boundary). On the other hand it corresponds to perfectly conducting borders, i.e., with infinite electric conductance, see Fig. 2.1a.

**Definition 2.2** (Neumann boundary). Let  $\Gamma_{\text{neu}}$  denote the Neumann boundary and  $\vec{n}$  its outward normal unit vector then

$$\vec{\nu} \left( \nabla \times \vec{A}(\vec{r}) \right) \times \vec{n} = \vec{H}_{\text{neu}} \quad \text{with} \quad \vec{r} \in \Gamma_{\text{neu}}$$

sets the tangent components of the magnetic field strength at the boundary to the prescribed value  $\vec{H}_{\text{neu}}$ . In the case of  $\vec{H}_{\text{neu}} = 0$  the condition is called a homogeneous Neumann BC (for the magnetic vector potential).

The condition is called *magnetic BC*, often also described as “flux gate” or “current wall”. It forces the magnetic flux to leave the computational domain perpendicular to the border and corresponds on the other hand to perfectly resistive borders, see Fig. 2.1b.

**Definition 2.3** (Anti-periodic boundary). Anti-periodic boundaries are used to exploit symmetry and thus they are defined in pairs. Let  $\Gamma_{\text{anti}} = \Gamma_{\text{anti,+}} \cup \Gamma_{\text{anti,-}}$  denote two

boundaries connected by a mapping  $\vec{s}(\vec{r}) : \Gamma_{\text{anti},+} \rightarrow \Gamma_{\text{anti},-}$  and let  $\vec{n}_+$ ,  $\vec{n}_-$  describe their outer normal unit vectors, then

$$\nabla \times \vec{A}(\vec{r}) \cdot \vec{n}_+ = -\nabla \times \vec{A}(\vec{s}) \cdot \vec{n}_-$$

forces the same magnetic flux entering/leaving through one boundary  $\Gamma_{\text{anti},+}$  to equally enter/leave through the other one  $\Gamma_{\text{anti},-}$  (for an equal spatial distribution), see Fig. 2.1c.

Finally the problem description is complete with the specification of the time interval of interest, i.e.,  $\mathcal{I} := [t_0, t_e]$  and an initial value for the parabolic setting.

**Assumption 2.4** (Initial value). *The initial conditions are given for  $t_0$*

$$\vec{A}(t_0, \vec{r}) = \vec{A}_0(\vec{r}) \quad \text{for all} \quad \vec{r} \in \Omega. \quad (2.7)$$

### 2.1.2 Partitioning into Regions and Materials

In engineering applications the devices can be decomposed into several regions. For example a single-phase transformer model consists of two coils wound around a common core and it is surrounded by air. The air region is important, because it ensures that the computational domain is sufficiently large, see Ass. 2.2. Each region is made of a material. For example the coils are made of copper strands and the core is made of iron. Mathematically speaking the regions define a partitioning of the computational domain.

**Assumption 2.5** (Regions). *We assume that the computational domain is partitioned into  $n_\Omega$  subdomains  $\Omega^{(r)}$ , called regions*

$$\bar{\Omega} = \bigcup_r \bar{\Omega}^{(r)} \quad \text{with} \quad \Omega^{(r)} \cap \Omega^{(s)} = \emptyset \quad \text{for } r \neq s \quad (2.8)$$

with interfaces  $\Gamma^{(r,s)} = \bar{\Omega}^{(r)} \cap \bar{\Omega}^{(s)}$  that are piecewise  $C^2$  and Lipschitz-continuous, [72].

We assume that the materials properties are constant within each region.

**Assumption 2.6** (Permittivity). *We assume an isotropic permittivity that is given by a positive constant in each region, i.e.,*

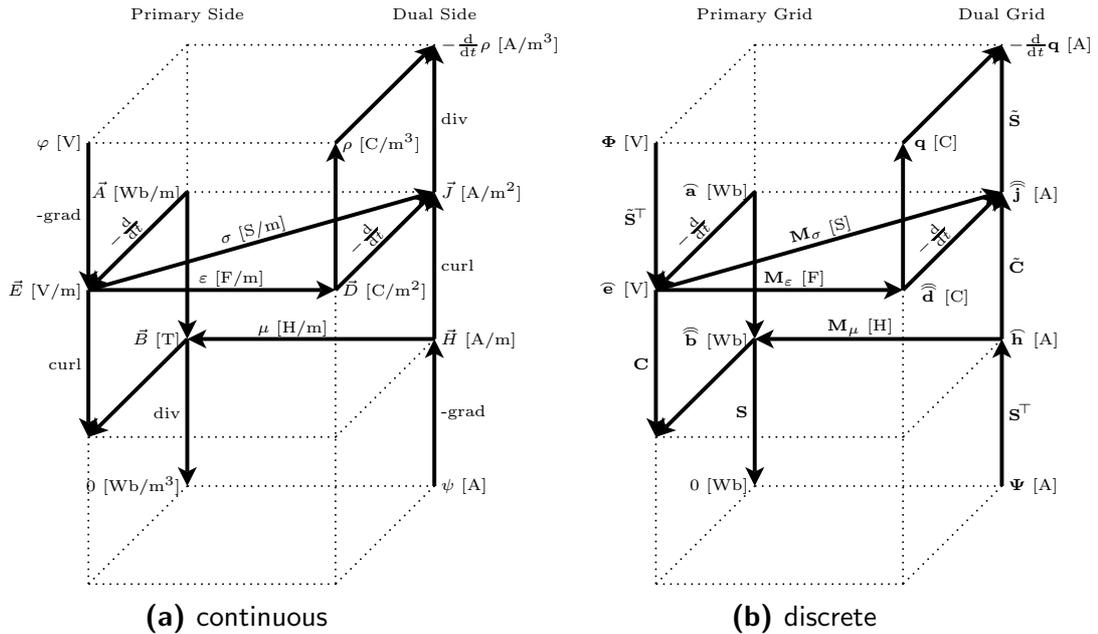
$$\frac{\partial \vec{D}}{\partial \vec{E}} = \vec{\varepsilon}(\vec{r}) = \varepsilon^{(r)} \mathbf{I} \quad \text{with} \quad \varepsilon^{(r)} > 0 \quad \text{for } \vec{r} \in \Omega^{(r)}. \quad (2.9)$$

**Assumption 2.7** (Conductivity). *We assume a region-wise constant non-negative conductivity, whose axes of anisotropy are aligned with the coordinate system, i.e.,*

$$\frac{\partial \vec{J}}{\partial \vec{E}} = \vec{\sigma}(\vec{r}) = \vec{\sigma}^{(r)} = \text{diag}(\sigma_x^{(r)}, \sigma_y^{(r)}, \sigma_z^{(r)}) \quad \text{with} \quad \sigma_{\{x,y,z\}}^{(r)} \geq 0 \quad \text{for } \vec{r} \in \Omega^{(r)} \quad (2.10)$$

**Assumption 2.8** (Reluctivity). *We assume that the reluctivity tensor is given by Brauer's model, [24]. It is defined as the superposition of an isotropic and anisotropic tensor, i.e.,*

$$\vec{\nu}(\|\vec{B}\|^2) = \nu_{iso}(\vec{r}, \|\vec{B}\|^2) \cdot \mathbf{I} + \vec{\nu}_{aniso}(\vec{r}) \quad (2.11)$$



**Figure 2.2:** Maxwell's house. Maxwell's equations and differential forms, i.e., Tonti's diagram for the discrete and continuous formulations, [129, 34].

with the region-wise constant differential reluctivity

$$\frac{\partial \vec{H}}{\partial \vec{B}} = \frac{\partial (\vec{H}_{iso} + \vec{H}_{aniso})}{\partial \vec{B}} = \vec{\nu}_d(\vec{r}, \|\vec{B}\|^2) = \vec{\nu}_d^{(r)}(\|\vec{B}\|^2) \quad \text{for } \vec{r} \in \Omega^{(r)}.$$

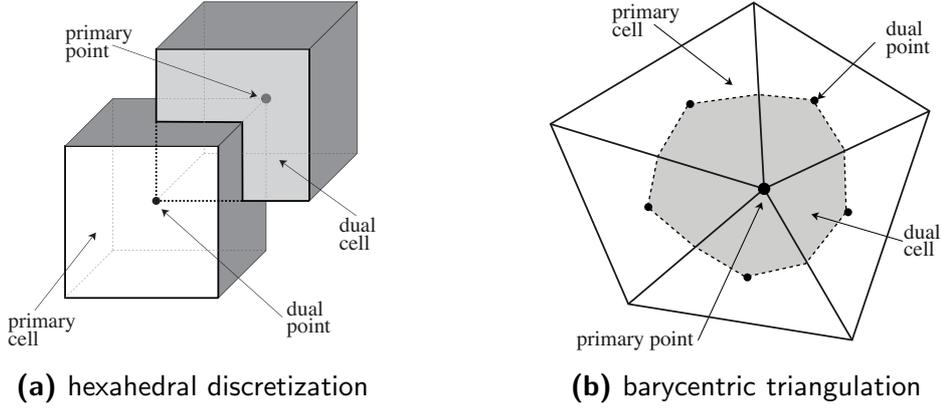
Only the isotropic material scalar  $\nu_{iso}$  depends nonlinearly on the magnetic flux density  $\|\vec{B}\|^2$  and it is assumed that this tensor and all derivatives are positive definite and bounded

$$\nu_2 > \nu_{iso}^{(r)}(\|\vec{B}\|^2) > \nu_1 \quad , \quad \nu_{d,iso}^{(r)} := \frac{\partial \vec{H}_{iso}}{\partial \vec{B}} = \frac{\partial \nu_{iso}^{(r)}(\|\vec{B}\|^2) \vec{B}}{\partial \vec{B}} \quad \text{and} \quad \vec{\nu}_{aniso}^{(r)} := \frac{\partial \vec{H}_{aniso}}{\partial \vec{B}}$$

by constants  $\nu_2 > \nu_1 > 0$ .

## 2.2 Space Discretization

Following a discrete approach to electromagnetics, see Fig. 2.2, or discretizing the equations (2.1) and (2.2) in space by the finite element method (FEM) with Whitney elements, [19, 22] or the finite integration technique (FIT), [138], one subdivides the computational domain into elements. We consider here FEM for 2D problems with triangular elements and a hexahedral FIT discretization for 3D problems. The latter uses explicitly staggered grids, while the FEM discretization can be interpreted accordingly, [19]. Fig. 2.3 shows a sketch of a primary and dual grid for hexahedral 3D discretization and a 2D triangulation. We assume that the discretization (primary grid) respects the regions defined in Section 2.1.2; they are subdivided into disjunct elements:



**Figure 2.3:** Primary and dual grid cells

**Assumption 2.9** (Discretization). *The primary grid discretizes the regions, i.e.,*

$$\bar{\Omega}^{(r)} = \bigcup_{\Omega_e \subset \Omega^{(r)}} \bar{\Omega}_e \quad \text{with} \quad \Omega_e \cap \Omega_j = \emptyset \text{ for } e \neq j, \quad (2.12)$$

where  $\bar{\Omega}_e$  denotes (the closure of) elements (triangle, hexahedron etc.) of the primary grid.

This treatise deals with polytopal elements only and thus Ass. 2.9 implies that the computational domain  $\bar{\Omega}$  and all regions  $\bar{\Omega}^{(r)}$  are polytopes. Again, this assumption is only made for simplicity of notation and can be overcome, e.g. by curved elements, [33].

**Definition 2.4** (Primary objects). The pairwise intersection of all elements of the primary grid yields **vertices** (0D polytope) denoted by  $P_i$  with  $i = 1, \dots, n_0$ , **edges** (1D polytope) denoted by  $L_i$  with  $i = 1, \dots, n_1$ , **facets** (2D polytope) denoted by  $A_i$  with  $i = 1, \dots, n_2$ , and **cells** (3D polytope) denoted by  $V_i$  with  $i = 1, \dots, n_3$ .

Similarly, one defines objects for the dual grid, see Section A.1.1.

### 2.2.1 Maxwell's Grid Equations

On both, primary and dual objects, one formulates Maxwell's grid equations, [138, 140]. Their definition and relation to differential forms using the Whitney complex was studied in [21, 19]. The equations can be read from Tonti's diagram, Fig. 2.2, independently of the element definitions above

$$\mathbf{C}\widehat{\mathbf{e}} = -\frac{d}{dt}\widehat{\mathbf{b}}, \quad \tilde{\mathbf{C}}\widehat{\mathbf{h}} = \frac{d}{dt}\widehat{\mathbf{d}} + \widehat{\mathbf{j}}, \quad \tilde{\mathbf{S}}\widehat{\mathbf{d}} = \mathbf{q}, \quad \mathbf{S}\widehat{\mathbf{b}} = 0, \quad (2.13)$$

with discrete curl operators  $\mathbf{C}$  and  $\tilde{\mathbf{C}}$ , divergence operators  $\mathbf{S}$  and  $\tilde{\mathbf{S}}$  on the primal and dual grid, respectively. The variables are time dependent quantities, whose diacritical bows correspond to the dimension of the underlying object, i.e., edges, facets and cells, [21]. On the one hand there are the line-integrals of electric and magnetic field strength  $\widehat{\mathbf{e}}(t)$  and  $\widehat{\mathbf{h}}(t)$ . They are located on primary and dual edges. For example the line-integrated electric field strength and magnetic vector potential are time dependent functions  $\mathcal{I} \rightarrow \mathbb{R}^{n_1}$

measured in Volt (V) and Weber (Wb), respectively, see Fig. 2.2b. They read

$$\widehat{\mathbf{e}}_i := \int_{L_i} \vec{E} \cdot d\mathbf{s} \quad \text{and} \quad \widehat{\mathbf{a}}_i := \int_{L_i} \vec{A} \cdot d\mathbf{s} \quad i = 1, \dots, n_1, \quad (2.14)$$

and are defined on the primary edges  $L_i \subset \Omega$ . The scalar  $n_1$  denotes the number of all primary edges. The other quantities, i.e., the discrete magnetic flux density, the discrete current density and the displacement field  $\widehat{\mathbf{b}}(t)$ ,  $\widehat{\mathbf{j}}(t)$  and  $\widehat{\mathbf{d}}(t)$ , respectively, are defined as surface integrals. The first is located on primary facets and the latter ones are defined on dual facets. Finally the electric flux density  $\mathbf{q}$  is defined on dual volumes.<sup>1</sup>

The constitutive material laws relate the primary and dual quantities to each other, see Fig. 2.2b, such that Maxwell's Grid Equations are completed by

$$\widehat{\mathbf{d}} = \mathbf{M}_\varepsilon \widehat{\mathbf{e}}, \quad \widehat{\mathbf{j}} = \mathbf{M}_\sigma \widehat{\mathbf{e}}, \quad \widehat{\mathbf{h}} = \mathbf{M}_\nu \widehat{\mathbf{b}}, \quad (2.15)$$

where the matrices  $\mathbf{M}_\varepsilon$ ,  $\mathbf{M}_\sigma$  and  $\mathbf{M}_\nu = \mathbf{M}_\nu(\widehat{\mathbf{b}})$  represent the permittivities, conductivities and (nonlinear) reluctivities, respectively. The matrices of permittivities and reluctivities are always positive definite, while the matrix of conductivities is typically positive semi-definite due to non-conducting parts in the computational domain, see Section 2.2.3.

In the FEM context the material matrices are typically constructed by assembling local material matrices defined per element. On the other hand in the classical FIT there is no need for the construction of local matrices due to the sophisticated index mappings on the underlying (structured) grid. Nonetheless for unity of notation the element-wise approach can be used for both methods, [42, 41]. This approach is generalized in Section A.2.2 to apply to anisotropic materials.

For both discretization the material properties were assumed region-wise constant, Ass. 2.6-2.8 and this property is inherited by the elements. It holds for the element in the  $r$ -th region ( $\vec{r}_e \in \Omega_e \subset \Omega^{(r)}$ ):

$$\vec{\varepsilon}(\vec{r}_e) = \vec{\varepsilon}^{(r)}, \quad \vec{\sigma}(\vec{r}_e) = \vec{\sigma}^{(r)} \quad \text{and} \quad \vec{\nu}(\vec{r}_e, \|\mathbf{B}_e\|^2) = \vec{\nu}^{(r)}(\|\mathbf{B}_e\|^2)$$

with constant permittivity  $\vec{\varepsilon}^{(r)}$ , conductivity  $\vec{\sigma}^{(r)}$  and reluctivity  $\vec{\nu}^{(r)}$ . The nonlinear reluctivity is evaluated using an element-wise averaged magnetic flux density  $\|\mathbf{B}_e\|^2$ .

The rough material approximation above implies that curved material boundaries are discretized by staircase approximations. This simplifies the notation because no material parameter averaging is necessary, but in practice this limitation is overcome by subgridding or other more elaborate schemes, [126, 83]. In either case the global material matrices

$$\mathbf{M}_\varepsilon := \sum_r \mathbf{M}_{\varepsilon,r}, \quad \mathbf{M}_\sigma := \sum_r \mathbf{M}_{\sigma,r}, \quad \mathbf{M}_\nu(\widehat{\mathbf{b}}) := \sum_r \mathbf{M}_{\nu,r}(\widehat{\mathbf{b}}), \quad (2.16)$$

can be written as the sum of region-wise matrices, assembled from the element contributions, see Section A.2.2.

**Definition 2.5** (Regional index sets). The indices of all primary objects belonging to a region can be expressed by index sets. For example the edges for the (closed) region  $r$  are

<sup>1</sup>For simplicity of notation the electric flux density is denoted without bows, although located on volumes.

given by

$$\mathcal{L}(\bar{\Omega}^{(r)}) := \{i \in \mathbb{N} \mid L_i \in \bar{\Omega}^{(r)}\}.$$

Note that an object may be included in one or more index sets because of intersections of closures (for example the region-wise matrices overlap only at region boundaries).

**Lemma 2.10** (Conductive regions). *On a (conductive) region  $r$  with  $\bar{\sigma}^{(r)} > 0$  the region-wise material matrix is positive definite, i.e.,*

$$\bar{\mathbf{e}}^\top \mathbf{M}_{\sigma,r} \bar{\mathbf{e}} > 0 \quad \text{if } i \in \mathcal{L}(\bar{\Omega}^{(r)}) \quad \text{such that } \bar{\mathbf{e}}_i \neq 0.$$

*Proof.* Direct consequence of the element assembly, see for example the FIT case in Section A.2.2.  $\square$

Analogous results hold for the region-wise permittivities and reluctivities.

## 2.2.2 Semi-discrete Problem Formulation

In analogy to the continuous formulation, Maxwell's grid equations (2.13) can be rewritten to form a single equation. Its unknown is the line-integrated magnetic vector potential, [17]. It is accompanied by the discrete electric scalar potential  $\Phi : \mathcal{I} \rightarrow \mathbb{R}^{n_0}$  located on primary points with  $n_0$  denoting their number. We find analogously to the continuous case

$$\hat{\mathbf{b}} = \mathbf{C}\hat{\mathbf{a}} \quad \text{and} \quad \bar{\mathbf{e}} = -\frac{d}{dt}\hat{\mathbf{a}} + \tilde{\mathbf{S}}^\top \Phi. \quad (2.17)$$

For 3D discretizations the matrix  $\mathbf{C}$  is singular, such that the vector potential is not uniquely defined (as in in the continuous case). Section 2.2.3 discusses possible regularization techniques. In further correspondence with the continuous formulation, i.e., (2.4) and (2.5), one obtains the semi-discrete curl-curl equations for the magnetoquasistatic (Ass. 2.1 i) and magnetostatic regimes (Ass. 2.1 iii) as

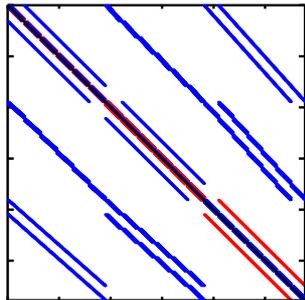
$$\mathbf{M}_\sigma \frac{d}{dt}\hat{\mathbf{a}} + \tilde{\mathbf{C}}\mathbf{M}_\nu(\mathbf{C}\hat{\mathbf{a}})\mathbf{C}\hat{\mathbf{a}} = \hat{\mathbf{j}}_s \quad \text{and} \quad \tilde{\mathbf{C}}\mathbf{M}_\nu(\mathbf{C}\hat{\mathbf{a}})\mathbf{C}\hat{\mathbf{a}} = \hat{\mathbf{j}}_s \quad (2.18)$$

with a given (facet-integrated) source current density  $\hat{\mathbf{j}}_s$ . The degrees of freedom (DoF) of this problem are the line-integrated vector potentials  $\hat{\mathbf{a}}$ . The mass matrix  $\mathbf{M}_\sigma$  is the algebraic origin for eddy currents and thus magnetoquasistatic problems are called *eddy-current problems*. The stiffness matrix is the derivative of the curl-curl term with respect to  $\hat{\mathbf{a}}$ . It is called the (differential) *curl-curl matrix*

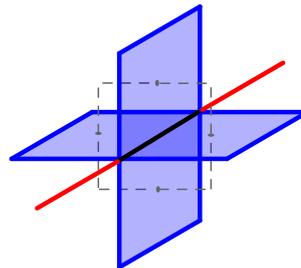
$$\frac{\partial}{\partial \hat{\mathbf{a}}} \left( \tilde{\mathbf{C}}\mathbf{M}_\nu(\mathbf{C}\hat{\mathbf{a}})\mathbf{C}\hat{\mathbf{a}} \right) = \tilde{\mathbf{C}}\mathbf{M}_{\nu,d}(\mathbf{C}\hat{\mathbf{a}})\mathbf{C}, \quad (2.19)$$

whose derivation is given in more detail for the FIT case in Section A.3. Finally the semi-discrete Poisson problems in the electrostatic approximation (Ass. 2.1 iii) read

$$\tilde{\mathbf{S}}\mathbf{M}_\epsilon \tilde{\mathbf{S}}^\top \Phi = \mathbf{q} \quad \text{and} \quad \tilde{\mathbf{S}}\mathbf{M}_\sigma \tilde{\mathbf{S}}^\top \Phi = 0. \quad (2.20)$$



(a) spy plot of (regularized) curl-curl matrix



(b) curl-curl discretization stencil

**Figure 2.4:** The regularized curl-curl matrix consists of entries from the original matrix and additions from the regularization, see [36].

similarly to (2.6). The equations (2.18) and (2.20) yield well-posed problems when boundary conditions and for the initial value problems the corresponding initial values are prescribed. For the magneto(quasi)static problems the discretization of the boundary conditions, Section 2.1.1 yields restrictions for MVP; for example in the homogeneous Dirichlet case the tangential components vanish

$$\widehat{\mathbf{a}}_i = 0 \quad \text{if } i \in \mathcal{L}(\Gamma_{\text{dir}}),$$

and for the (anti-)periodic condition we identify

$$\widehat{\mathbf{a}}_r = \vartheta \widehat{\mathbf{a}}_s \quad \text{if } r \in \mathcal{L}(\Gamma_{\text{anti},+}) \quad \text{and } L_s = \vec{s}(L_r)$$

where  $\vartheta = \pm 1$  depends on the orientation of the edges and the type of periodicity. The homogeneous Neumann condition is naturally fulfilled by the discretization. In any case the conditions are built into the system matrices, such that superfluous components are removed, e.g. [103, Section 8.4.5].

### 2.2.3 Gauging

On a 3D domain the curl-operator inherits the non-uniqueness from its continuous counterpart, i.e.,

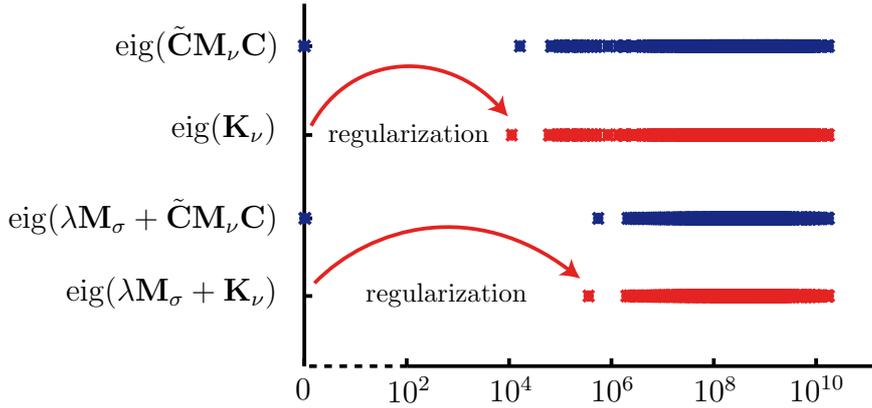
$$\mathbf{C}\tilde{\mathbf{S}}^\top \Phi = 0 \quad \text{for all } \Phi. \quad (2.21)$$

whereby the gradient operator  $-\tilde{\mathbf{S}}^\top$  has full column rank, Section A.1.1. Thus the curl-curl matrix (2.19) has zero eigenvalues and it is not invertible. This is inconvenient for the structural analysis and renders direct matrix factorizations impossible. This can be overcome by regularizations in a similar way to the continuous Coulomb gauge, [20, 37, 81]. Those regularizations remove the nullspace by addition

**Definition 2.6** (Regularization). The regularized curl-curl term reads

$$\mathbf{k}_\nu(\widehat{\mathbf{a}}) := \tilde{\mathbf{C}}\mathbf{M}_\nu(\mathbf{C}\widehat{\mathbf{a}})\mathbf{C} + \mathbf{Z}_\sigma, \quad \text{where } \Phi^\top \tilde{\mathbf{S}}\mathbf{Z}_\sigma \tilde{\mathbf{S}}^\top \Phi > 0 \quad \text{for all } \Phi \neq 0, \quad (2.22)$$

such that  $\mathbf{k}_\nu(\widehat{\mathbf{a}})$  is positive definite (and  $\mathbf{Z}_\sigma$  is positive semi-definite).



**Figure 2.5:** Shift of eigenvalues due to grad-div regularization, parameter  $\lambda = 10^3$ , cf. [36].

This approach transfers directly to the differential curl-curl matrix.

**Corollary 2.11** (Regularization of differential curl-curl matrix). *The differential curl-curl matrix is positive definite and it reads*

$$\mathbf{K}_\nu(\bar{\mathbf{a}}) := \frac{\partial}{\partial \bar{\mathbf{a}}} \left( \tilde{\mathbf{C}}\mathbf{M}_\nu(\mathbf{C}\bar{\mathbf{a}})\mathbf{C}\bar{\mathbf{a}} \right) + \mathbf{Z}_\sigma. \quad (2.23)$$

*Proof.* See Section A.3. □

The regularization  $\mathbf{Z}_\sigma$  is supposed to shift the zero eigenvalues to the positive real axes, but should not affect nonzero eigenvalues. Otherwise the corresponding flux  $\widehat{\mathbf{b}} = \mathbf{C}\bar{\mathbf{a}}$  would be altered. A common regularization approach is the following, [20, 37]:

**Definition 2.7** (Grad-div regularization). The particular regularization

$$\mathbf{Z}_\sigma := \mathbf{M}_1 \tilde{\mathbf{S}}^\top \mathbf{M}_2 \tilde{\mathbf{S}} \mathbf{M}_1 \quad (2.24)$$

is called the Grad-Div regularization. It utilizes the gradient and divergence operators  $\tilde{\mathbf{S}}$  from above and suitable (artificial) material matrices:  $\mathbf{M}_1$  maps primary edges to dual facets and the norm matrix  $\mathbf{M}_2$  maps dual points to primary volumes, [37].

For a homogenous material distribution and an equidistant grid, the Grad-Div regularized curl-curl matrix (see Fig. 2.4) corresponds to the discrete vector Laplacian  $(-\nabla^2)$ , [37]. Suitable choices for the material matrices  $\mathbf{M}_1$  and  $\mathbf{M}_2$  are discussed in [36].

When considering a magnetoquasistatic problem (2.18), we are interested in the regularity of the matrix pencil (c.f. Fig. 2.5)

$$[\mathbf{M}_\sigma, \mathbf{K}_\nu(\bar{\mathbf{a}})] := \lambda \mathbf{M}_\sigma + \mathbf{K}_\nu(\bar{\mathbf{a}}), \quad \text{with } \lambda > 0 \quad (2.25)$$

which occurs naturally when solving problems in the time domain, e.g.  $\lambda = 1/h$  for the implicit Euler, see Section 5.1. The conductivity matrix  $\mathbf{M}_\sigma$  already shifts some of the zero eigenvalues of the curl-curl matrix onto the positive real axis, see Section 5.2. The regularization (2.22) would affect those eigenvalues superfluously. Hence a gauging in the nonconductive domain is sufficient, [37, 36]. This domain can be addressed formally by the following projector:

**Definition 2.8** (Projector onto non-conductive regions). Let  $\mathbf{Q}_\sigma$  be a constant projector onto  $\text{Ker } \mathbf{M}_\sigma$ , such that  $\mathbf{Q}_\sigma = \mathbf{Q}_\sigma^\top$ . The complementary projector  $\mathbf{P}_\sigma = \mathbf{I} - \mathbf{Q}_\sigma$  is given by

$$\mathbf{P}_\sigma = \mathbf{M}_\sigma^+ \mathbf{M}_\sigma \quad (2.26)$$

where  $\mathbf{M}_\sigma^+$  is the (Moore-Penrose) pseudo-inverse of  $\mathbf{M}_\sigma$ , [61].

In the MQS case it is sufficient to regularize only in the non-conductive parts, e.g. by restricting the regularization to  $\mathbf{Z}_\sigma = \mathbf{Q}_\sigma^\top \mathbf{Z}_\sigma \mathbf{Q}_\sigma$ . The restriction can be incorporated into the artificial material matrices in Definition 2.7. This is the ‘Grad-Div Gauging Type II’ in [36]. In either case the following result holds (in particular for the Grad-Div Gauging)

**Corollary 2.12.** *If a positive semi-definite regularization  $\mathbf{Z}_\sigma$  fulfills*

$$\widehat{\mathbf{a}}^\top \mathbf{Z}_\sigma \widehat{\mathbf{a}} > 0 \quad \text{for all } \widehat{\mathbf{a}} \in \text{Ker } \mathbf{K}_\nu \cap \text{Ker } \mathbf{M}_\sigma \quad (2.27)$$

*then the matrix pencil  $[\mathbf{M}_\sigma, \mathbf{K}_\nu(\widehat{\mathbf{a}})]$  is positive definite for all  $\widehat{\mathbf{a}}$  and  $\lambda > 0$ .*

*Proof.* We find that

$$\Phi^\top \tilde{\mathbf{S}} (\lambda \mathbf{M}_\sigma + \mathbf{Z}_\sigma) \tilde{\mathbf{S}}^\top \Phi > 0 \quad \text{for all } \Phi \neq 0.$$

is positive in both of the two possible cases:

1. if  $\tilde{\mathbf{S}}^\top \Phi \in \text{Ker } \mathbf{M}_\sigma$  then there is an  $\widehat{\mathbf{a}}$  such that  $\tilde{\mathbf{S}}^\top \Phi = \mathbf{Q}_\sigma \widehat{\mathbf{a}}$ . The first summand vanishes but the second summand is positive because the element is not in  $\text{Ker } \mathbf{Z}_\sigma$ .
2. else  $\mathbf{M}_\sigma \tilde{\mathbf{S}}^\top \Phi \neq 0$  and thus the first summand is positive and the second summand is non-negative.  $\square$

Independently of the particular choice for the regularization, the following assumption is made:

**Assumption 2.13** (Gauge). *In the magnetostatic case the curl-curl matrix is fully regularized (Definition 2.6) with a positive semi-definite matrix  $\mathbf{Z}_\sigma$  such that  $\mathbf{k}_\nu$  and  $\mathbf{K}_\nu$  are positive definite. In the MQS case the same regularization is applied but only for elements in  $\text{Ker } \mathbf{M}_\sigma$ .*

As a consequence of this assumption, in the following the regularized curl-curl equations are used for the analysis

$$\mathbf{M}_\sigma \frac{d}{dt} \widehat{\mathbf{a}} + \mathbf{k}_\nu(\widehat{\mathbf{a}}) \widehat{\mathbf{a}} = \widehat{\widehat{\mathbf{j}}}_s \quad \text{and} \quad \mathbf{k}_\nu(\widehat{\mathbf{a}}) \widehat{\mathbf{a}} = \widehat{\widehat{\mathbf{j}}}_s, \quad (2.28)$$

although in practice the gauging is not applied if an iterative solver is available:

**Remark 2.1** (Weak Gauging). No explicit regularization is necessary when solving linear systems with system matrices of the form (2.25) with Krylov subspace methods. They exhibit a weak gauging property, [37], i.e., they will not alter the initial guess in the nullspace of the system matrix, [78].

## 2.3 Conclusions

This chapter has introduced briefly the continuous and discrete formulation of Maxwell's equations and their common quasistatic approximations. General properties of the discrete material and differential operators for spatial discretizations based on edge-elements have been presented. The notation was based on FIT, but without loss of generality.

Special emphasis was put on a uniquely solvable discrete problem formulation, because the following chapters will discuss time-integration and the corresponding theory and methods rely on uniqueness.

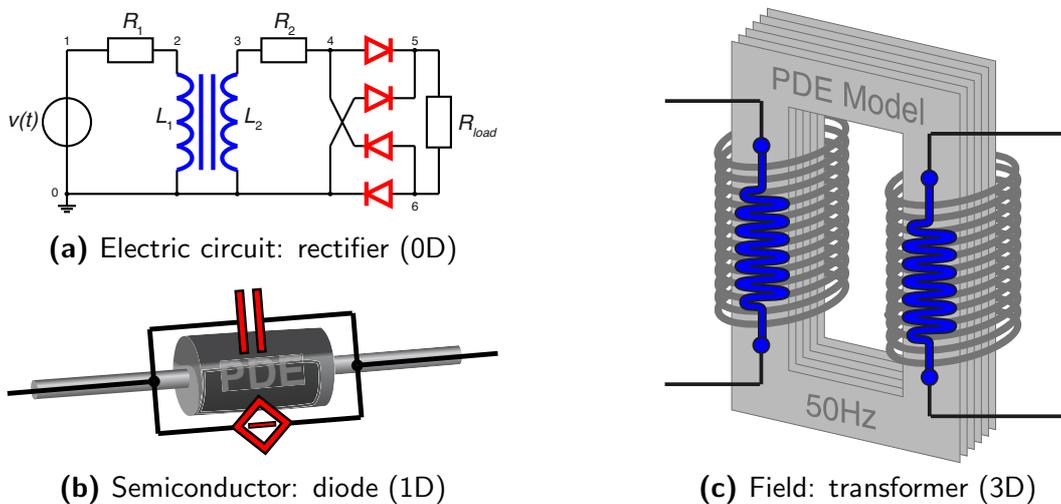
This chapter has established the electromagnetic framework into which the following chapter embeds (multiscale) models for electromagnetic networks and distributed device models, i.e., the magnetoquasistatic and semiconductor devices.

### 3 Multiscale Device Models

In the previous chapter the electromagnetic phenomena of a single problem ('device') were described on the macroscopic scale. In electric circuit simulation, e.g. [32], the behavior of the combination of a large number of such devices is in focus. This chapter focusses on effects in circuit simulation that must be resolved on different spatial scales. This will be clarified by the description of a transformer and a semiconductor, whose spatial dimensions differ by many orders of magnitude.

A multiscale simulation with PDE models of all devices, is obviously computationally inappropriate. Instead the complexity of some devices is reduced, e.g., by disregarding the spatial distribution (lumping). Then a network of idealized basic elements is considered, where each element describes only one effect, e.g. resistances, inductances, capacitances and sources. A single physical device is approximated by several basic elements. They are called equivalent, compact or companion models. Most often, the network equations for all devices are set up element-wise according to the modified nodal analysis (MNA), which is introduced in Section 3.1. Mathematically speaking the result is a set of differential algebraic equations (DAEs).

However, many devices cannot be described sufficiently accurately in terms of a few idealized lumped elements. The corresponding equivalent circuits become too complex and contain hundreds of parameters, most of them with no physical interpretation, [44]. Especially if one is interested in the effects on all scales at the same time, a hierarchical modeling of PDE devices and the electric network DAEs is the only viable strategy (*mixed mode simulation*, [62]). The result is a system of partial differential algebraic equations (PDAEs), which models the distributed effects by computationally expensive PDE models only where necessary.



**Figure 3.1:** Multiscale modeling: 0D network coupled to 1D and 3D devices.

In the next sections we present different models (see Fig. 3.1) and couplings. The first model in Section 3.1 is a recapitulation of the electric network description in terms of the modified nodal analysis (MNA). This formulation is the standard approach in commercial tools for circuit analysis (e.g. SPICE-like simulators as Titan by Infineon), [108]. In Section 3.2 the magnetoquasistatic field device and the corresponding field/circuit coupling is established. The device model yields a current/voltage relation and this makes the circuit coupling straightforward, independently of the type of circuit analysis. This treatise follows [113, 8] and features the MNA. Other approaches use the standard nodal analysis [46, 137] or study the coupling and numerical treatment for a coupled system using loop/branch analysis, e.g. [15]. Finally in Section 3.3 the standard drift-diffusion model of a semiconductor is reproduced, [123, 119, 62]. This well-known model features a microscopic correction of the electrostatic Maxwell's equations in terms of holes and electrons.

### 3.1 Electric Network

The electric network model, [32, 74], is derived from Maxwell's equations (2.1). The spatial distributions are disregarded and instead a 0D network is considered. Its basic elements, i.e., capacitors, inductors, resistors correspond to the constitutive relations (2.2), extended by voltage and current sources. The unknowns are lumped quantities obtained from Maxwell's distributed quantities, e.g., the currents  $\mathbf{i}$  through the elements are defined as surface integrals of the current densities  $\vec{J}$  at the contacts of the elements. The network topology is maintained by (reduced) incidence matrices:  $\mathbf{A}_C$ ,  $\mathbf{A}_R$ ,  $\mathbf{A}_L$ ,  $\mathbf{A}_V$  and  $\mathbf{A}_I$  where each matrix corresponds to an element type. They state the node-branch relations for each element type for the underlying digraph:

$$(\mathbf{A}_\star)_{ij} = \begin{cases} 1, & \text{if branch } j \text{ leaves node } i \\ -1, & \text{if branch } j \text{ enters node } i \\ 0, & \text{if branch } j \text{ is not incident with node } i \end{cases}$$

where each row of  $\mathbf{A}_\star$  refers to a network node. One node is identified as the mass node ('ground') and its row is skipped in the reduced matrix  $\mathbf{A}_\star$ . The flux/charge oriented modified nodal analysis (MNA) yields equations of the form, [64, 57, 54]

$$\mathbf{A}_C \frac{d}{dt} \mathbf{q} + \mathbf{A}_R \mathbf{g}_R(\mathbf{A}_R^\top \mathbf{u}, t) + \mathbf{A}_L \dot{\mathbf{i}}_L + \mathbf{A}_V \mathbf{i}_V + \mathbf{A}_I \mathbf{i}_s(t) = 0, \quad (3.1a)$$

$$\mathbf{q} - \mathbf{q}_C(\mathbf{A}_C^\top \mathbf{u}, t) = 0, \quad (3.1b)$$

$$\frac{d}{dt} \phi - \mathbf{A}_L^\top \mathbf{u} = 0, \quad (3.1c)$$

$$\phi - \phi_L(\dot{\mathbf{i}}_L, t) = 0, \quad (3.1d)$$

$$\mathbf{A}_V^\top \mathbf{u} - \mathbf{v}_s(t) = 0, \quad (3.1e)$$

where the time is denoted by  $t \in \mathcal{I} := [t_0, t_e]$  defined on an interval of interest. The given functions  $\mathbf{q}_C(\mathbf{v}, t)$ ,  $\mathbf{g}_R(\mathbf{v}, t)$ ,  $\phi_L(\dot{\mathbf{i}}, t)$ ,  $\mathbf{v}_s(t)$  and  $\mathbf{i}_s(t)$  describe the constitutive relations for the circuit elements, i.e., capacitances, resistances, inductances, voltage and current sources, respectively. The sources are assumed here to be independent, the extension

to classes of controlled sources is straightforward, [54]. The unknowns of the system are most importantly the node potentials  $\mathbf{u} : \mathcal{I} \rightarrow \mathbb{R}^{n_u}$  (without ground), they correspond to the voltage drop between the node and ground. Further unknowns are the currents  $\mathbf{i}_L : \mathcal{I} \rightarrow \mathbb{R}^{n_L}$ ,  $\mathbf{i}_V : \mathcal{I} \rightarrow \mathbb{R}^{n_V}$  through inductors and voltage sources and the charges and fluxes  $\mathbf{q} : \mathcal{I} \rightarrow \mathbb{R}^{n_C}$  and  $\phi : \mathcal{I} \rightarrow \mathbb{R}^{n_L}$ , respectively (where  $n_L$ ,  $n_V$  and  $n_C$  denote their respective number), see [57, 65]. The problem is completed with (consistent) initial conditions at time  $t = t_0$ , [8].

The flux/charge-oriented MNA above is reduced to the traditional MNA, when the unknowns for the fluxes and charges are eliminated. This yields a smaller system, but does not guarantee the conservation of energy, [65]. Structurally both approaches are nearly equivalent, [54] and in either case a mathematically consistent description must fulfill some topological conditions. The constitutive relations should be passive, i.e.,

**Assumption 3.1** (Local passivity). *The functions  $\mathbf{q}_C(\mathbf{v}, t)$ ,  $\phi_L(\mathbf{i}, t)$  and  $\mathbf{g}_R(\mathbf{v}, t)$  are continuous differentiable with positive definite Jacobians:*

$$\mathbf{C}(\mathbf{v}, t) := \frac{\partial \mathbf{q}_C(\mathbf{v}, t)}{\partial \mathbf{v}}, \quad \mathbf{L}(\mathbf{i}, t) := \frac{\partial \phi_L(\mathbf{i}, t)}{\partial \mathbf{i}}, \quad \mathbf{G}(\mathbf{v}, t) := \frac{\partial \mathbf{g}_R(\mathbf{v}, t)}{\partial \mathbf{v}}.$$

In the following sections PDE devices are added to the network. That is, we enlarge our list of basic elements by magnetoquasistatic field and semiconductor devices, see Section 3.2 and Section 3.3, respectively. The coupling yields an extended circuit problem (this is the spatially discretized PDAE system). In the MNA framework we simply add the unknown currents  $\mathbf{i}_M \in \mathbb{R}^{n_M}$  and  $\mathbf{i}_D \in \mathbb{R}^{n_D}$  to the current balance equation (3.1a) using the corresponding incidence matrices  $\mathbf{A}_M$  and  $\mathbf{A}_D$ . Then Kirchhoff's current law (3.1a) for the coupled problem reads

$$\mathbf{A}_C \frac{d}{dt} \mathbf{q}_C(\mathbf{A}_C^\top \mathbf{u}, t) + \mathbf{A}_R \mathbf{g}_R(\mathbf{A}_R^\top \mathbf{u}, t) + \mathbf{A}_L \mathbf{i}_L + \mathbf{A}_V \mathbf{i}_V + \mathbf{A}_{I_s}(t) + \mathbf{A}_M \mathbf{i}_M + \mathbf{A}_D \mathbf{i}_D = 0. \quad (3.2)$$

To obtain a uniquely solvable system we need further equations for the devices which describe the unknown currents in terms of the applied voltage drops. These voltage drops are given by  $\mathbf{v}_M = \mathbf{A}_M^\top \mathbf{u}$  and  $\mathbf{v}_D = \mathbf{A}_D^\top \mathbf{u}$  and the corresponding model equations are the topic of Section 3.2 and Section 3.3. Furthermore the extended circuit must fulfill the following condition to prevent short circuits

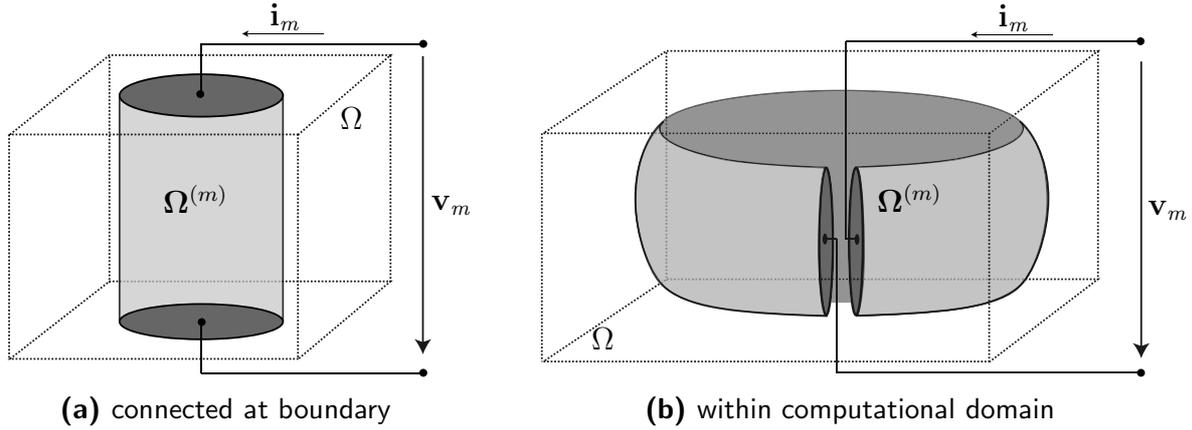
**Assumption 3.2** (Soundness of extended circuit). *The circuit shall be connected and the matrices*

$$\mathbf{A}_V \quad \text{and} \quad [\mathbf{A}_C \ \mathbf{A}_R \ \mathbf{A}_L \ \mathbf{A}_V \ \mathbf{A}_M \ \mathbf{A}_D]^\top$$

*have full column rank, i.e., there is neither a loop containing only voltage sources nor a cutset containing only current sources.*

If Ass. 3.2 is violated, the circuit equations (with initial conditions) would have either no solution or infinitely many solutions due to Kirchhoff's laws.

**Remark 3.1** (Incidence matrices). A (reduced) incidence matrix  $\mathbf{A}$  has full column rank if and only if there are no loops in the graph. It has full row rank if the graph is connected, i.e., there is a spanning tree. Let  $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2]$  describe a connected graph, where also  $\mathbf{A}_1$  has a spanning tree. Then  $\mathbf{A}_2$  has full row rank if and only if the subgraph contains no cutset of elements from the other subgraph  $\mathbf{A}_1$ .



**Figure 3.2:** MQS device types. Solid conductor connected at boundary and forming a loop within the domain.

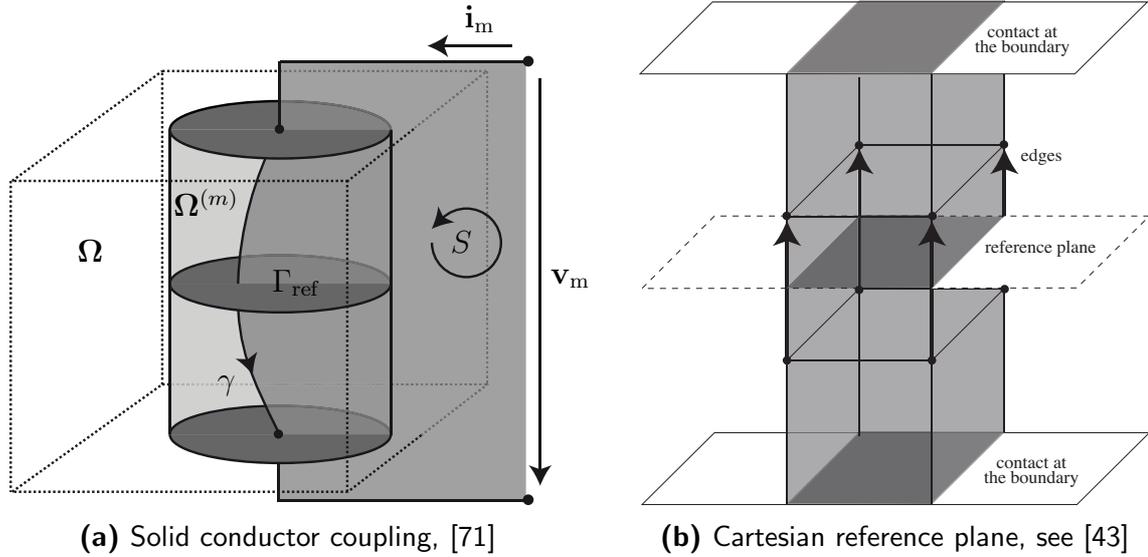
The network is now soundly defined. It is the coupling framework for all devices. The elements (multiscale models from 0D to 3D) may communicate only via lumped network quantities. Thus the following PDE models must have corresponding boundary conditions and source terms, such that physical correctness of the overall model is ensured. This is the topic of the following sections.

### 3.2 Magnetoquasistatic Conductor Model

The coupling of the MQS device to the circuit is established by identifying a device’s regions as circuit branches. Those regions must either be connected to the boundary, Fig. 3.2a, or form a loop within the computational domain, Fig. 3.2b. Otherwise the divergence-freeness of the current density could be violated. The coupling for the first configuration is straightforward, only appropriate boundary conditions must be set, [75]. In the second configuration, which is an important model for a coil, an (artificial) cut is introduced in the coil model (the *reference layer*) such that the circuit can be attached. We do not distinguish between the two configurations in the following: the procedure is analogous if we consider an infinitesimal reference layer, [49, 73]. In both cases the electric current is imposed from the coupled circuit by assigning the integrated current density to a branch current of the electric circuit and computing an electric field from the applied voltage drop. To guarantee a current flow, it is assumed

**Assumption 3.3** (Soundness of conductor region). *We assume the conductor region  $\Omega^{(m)} \subset \Omega$  to be conductive, i.e., a region where the material tensor  $\vec{\sigma}^{(M)}$  is positive definite.*

From the engineering point of view, it is necessary to distinguish between different conductor types. The most important types are the solid and stranded model, [133, 14]. We will only discuss the solid conductor in detail. It is a single massive bar connected with an electrode at each end. The contacts exhibit an equipotential voltage. On the other hand the stranded conductor models windings are made of thin strands. This model is based on the assumption that the current at a cross section is constant, [49, 48, 43]. The derivation of both models is analogous and it can be shown that both are structurally equivalent



**Figure 3.3:** MQS device coupling. Source term coupling at the reference plane.

(albeit different from the engineering point of view), [117]. The important differences are discussed in Section 3.2.1.

Let us consider in the following a single (solid) conductor (see Fig. 3.3a). Its region shall be denoted by  $\Omega^{(m)}$ . It corresponds to a circuit branch, which is connected by two perfect conducting contacts. The 0D-voltage drop  $\mathbf{v}_m$  must be distributed onto the 3D-domain; this defines an electric field.

For that we denote by  $\gamma$  an arbitrary path through the conductor domain and  $S$  is the area inside the loop described by the path  $\gamma$ , see Fig. 3.3a. The wires connecting the conductor to the circuit are regarded as 0D objects and therefore they are not considered in the field model, see [71, Chapter 8]. The voltage drop is given by Faraday's Law

$$\int_{\gamma} \vec{E} \cdot d\vec{s} - \mathbf{v}_m = -\frac{d}{dt} \int_S \vec{B} \cdot d\vec{A}.$$

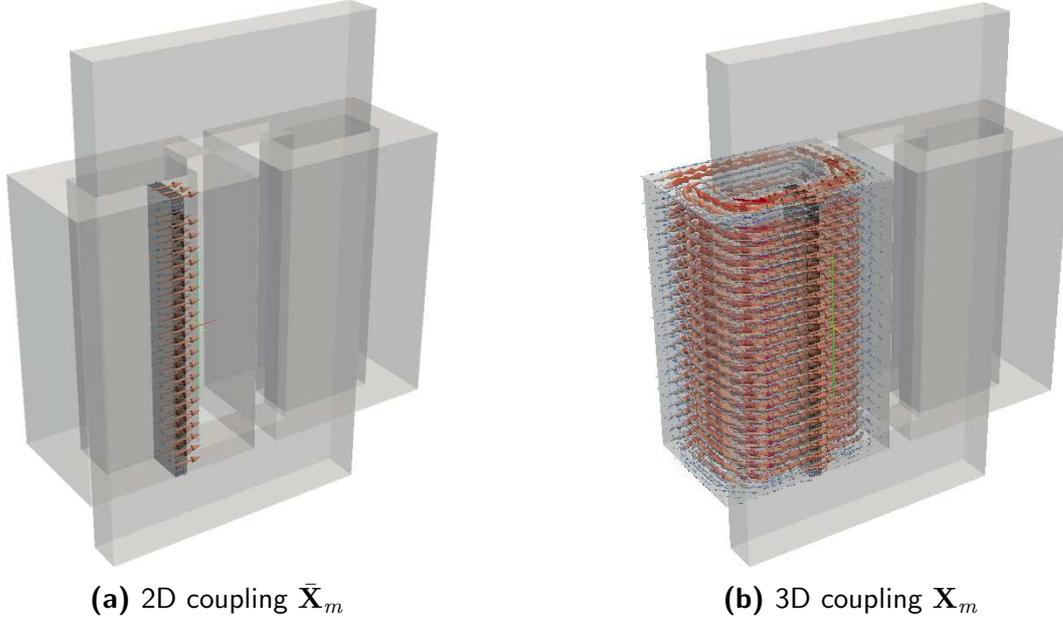
Replacing  $E = -\frac{d}{dt}\vec{A} - \nabla\varphi$ ,  $\vec{B} = \nabla \times \vec{A}$  and applying Stokes theorem yields

$$-\int_{\gamma} \frac{d}{dt} \vec{A} \cdot d\vec{s} - \int_{\gamma} \nabla\varphi \cdot d\vec{s} - \mathbf{v}_m = -\frac{d}{dt} \int_{\partial S} \vec{A} \cdot d\vec{s}.$$

The connecting wires are not considered in  $\Omega$ , i.e.,  $\partial S = \gamma$  and thus

$$\mathbf{v}_m = -\int_{\gamma} \nabla\varphi \cdot d\vec{s}. \quad (3.3)$$

Finally the electric field is only defined via its integral. This can be exploited in the construction of a discrete distribution. The discrete electric field  $\mathbf{e}_M$  is located on primary edges. Since we are only interested in line integrals on those edges, we look at an arbitrary



**Figure 3.4:** Full and sparse coupling. Stranded conductor coupling vectors and for the first coil of a transformer model discretized by FIT. This corresponds to a 1A excitation.

discrete path  $\bar{\gamma} \in \{-1, 0, 1\}^{n_1}$  from one contact to the other (within  $\Omega^{(m)}$ ). Due to the linearity of Ohm's law, it is sufficient to consider an applied voltage  $\mathbf{v}_m = 1\text{V}$  and define a corresponding distribution matrix  $\bar{\mathbf{X}}_m \in \mathbb{R}^{n_1}$ , such that  $\bar{\mathbf{X}}_m^\top \bar{\gamma} = 1$  for all  $\bar{\gamma}$ .

In [43, 8] a construction is proposed that imposes the voltages only onto edges that cross a reference plane  $\Gamma_{\text{ref}}$  (see Fig. 3.3b). This implies that the reference plane is a set of dual facets. The definition of the coupling reads for a Cartesian grid with an orthogonal aligned reference plane

$$(\bar{\mathbf{X}}_m)_i = \begin{cases} \pm 1 & \text{if edge } L_i \subset \Omega^{(m)} \text{ crosses the reference plane,} \\ 0 & \text{else,} \end{cases} \quad (3.4)$$

where the sign depends on the directions of the edges. This coupling vector  $\bar{\mathbf{X}}_m$  is sparse; it features only nonzero entries at the reference plain („2D coupling”). This is computationally less costly than a full coupling that exhibits entries in the whole 3D domain  $\Omega^{(m)}$ , [43]. Fig. 3.4 shows the coupling pattern of both approaches.

Nonetheless the full coupling is favored in the following, because it can be constructed in a way that the source current density is divergence free. The divergence property of the sparse coupling  $\bar{\mathbf{X}}_m$  is corrected by previously solving the following Poisson problem

$$\tilde{\mathbf{S}}\mathbf{M}_{\sigma,m}\tilde{\mathbf{S}}^\top \Phi_m = \tilde{\mathbf{S}}\mathbf{M}_{\sigma,m}\bar{\mathbf{X}}_m\mathbf{v}_m \quad \text{with} \quad \mathbf{Q}_{\sigma,m}\tilde{\mathbf{S}}^\top \Phi_m = 0. \quad (3.5)$$

with unit excitation  $\mathbf{v}_m = 1\text{V}$  and projector  $\mathbf{Q}_{\sigma,m}$  onto  $\text{Ker } \mathbf{M}_{\sigma,m}$ , see Definition 2.8.

**Lemma 3.4** (Excitation). *Let a sparse coupling vector  $\bar{\mathbf{X}}_m$  be given as defined in (3.4) and let the potential  $\Phi_m$  denote the solution of (3.5). Then the full coupling vector reads*

$$\mathbf{X}_m := \bar{\mathbf{X}}_m - \tilde{\mathbf{S}}^\top \Phi_m.$$

The current excitation is divergence free and the voltage drops are only applied in conductive regions, i.e.,  $\mathbf{X}_m = \mathbf{P}_\sigma \mathbf{X}_m$ .

*Proof.* From Lemma 2.10 follows  $\bar{\mathbf{X}}_m \subset \text{im } \mathbf{M}_{\sigma,m}$  and thus  $\mathbf{Q}_\sigma \bar{\mathbf{X}}_m = 0$ . The rest is clear from equation (3.5).  $\square$

By linearity, the coupling matrix  $\mathbf{X}_m$  allows the application of an arbitrary voltage  $\mathbf{v}_m = \mathbf{A}_m^\top \mathbf{u}$  by multiplication. The applied source current density is given by

$$\hat{\mathbf{j}}_m = \mathbf{M}_{\sigma,m} \tilde{\mathbf{S}} \Phi = \mathbf{M}_{\sigma,m} \mathbf{X}_m \mathbf{v}_m . \quad (3.6)$$

In the sparse coupling the total current through the conductor is given by integrating over the reference cross section. In the full coupling approach, the coupling vector  $\mathbf{X}_m$  averages over all cross section integrals. We find by using Ohm's Law in the solid conductor region

$$\mathbf{i}_m = \mathbf{X}_m^\top \hat{\mathbf{j}} = \mathbf{X}_m^\top \mathbf{M}_{\sigma,m} \mathbf{e} = \mathbf{X}_m^\top \mathbf{M}_{\sigma,m} \mathbf{X}_m \mathbf{v}_m - \mathbf{X}_m^\top \mathbf{M}_{\sigma,m} \frac{d}{dt} \hat{\mathbf{a}} . \quad (3.7)$$

The excitation is easily generalized to the case of an arbitrary number of conductors

$$\hat{\mathbf{j}}_M = \sum_{m \in M} \mathbf{M}_{\sigma,m} \mathbf{X}_m \mathbf{v}_m . \quad (3.8)$$

where  $M$  is the index set of all conductor regions (their total number is denoted by  $n_M = |M|$ ). Thus the field device has  $n_M$  terminals. In the case of the transformer shown in Fig. 3.4, the set consists of two regions that correspond to the primary and secondary coil.

In practice the region-wise construction of multiple neighbored conductors can suffer from smearing effects, [117]. The conductivities and thus the coupling vectors are not clearly separated at material boundaries. For a nonzero voltage drop  $\mathbf{v}_m$  the following currents are not the same

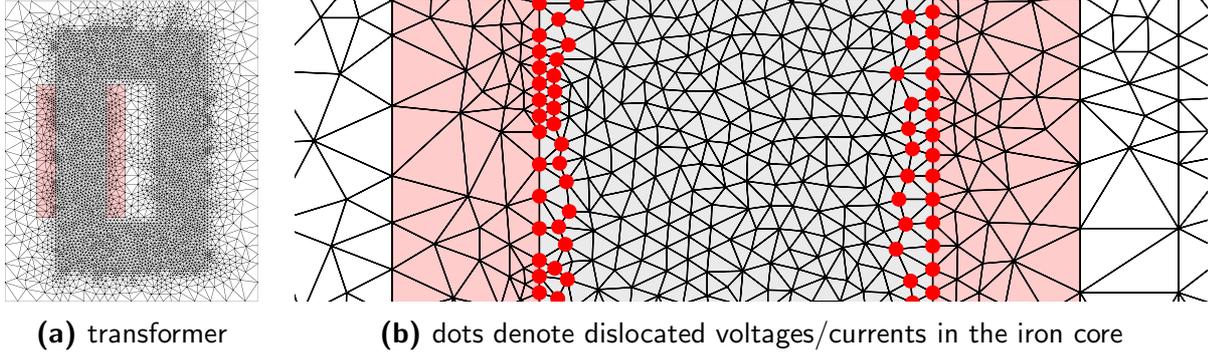
$$\mathbf{M}_\sigma \mathbf{X}_m \mathbf{v}_m \neq \mathbf{M}_{\sigma,m} \mathbf{X}_m \mathbf{v}_m ,$$

if a conductive region is adjacent to the coupling region  $\Omega^{(m)}$ . This may cause the model to behave unexpectedly. Let us consider a transformer discretized by 2D FEM, see Fig. 3.5. A voltage excitation in the (stranded) conductor region  $\Omega^{(m)}$  causes no current flow in the region itself because stranded conductors are commonly modeled as nonconductive material to disable eddy currents, see Section 3.2.1. On the other hand there is a current within the neighboring iron core due to smearing at the boundary.

This smearing is a typical discretization error. It can be prevented by using adequate material matrices, a sufficiently large insulation region around each conductor or the error can be made arbitrarily small by mesh refinements. The problem with the latter approach is that the insulation layers may be very thin but have to be resolved by the mesh. In the following it is assumed that there is no smearing.

**Assumption 3.5** (No smearing). *There is no smearing, i.e., the images of the region-wise conductivity matrices are distinct*

$$\text{im } \mathbf{M}_{\sigma,r} \cap \text{im } \mathbf{M}_{\sigma,m} = \{0\} \quad \text{for } r \neq m . \quad (3.9)$$



**Figure 3.5:** Smearing. Current in the iron core while excitation only applied to the coil  $\Omega^{(m)}$ .

This allows for the following simplified notation of the device equations using the curl-curl equation (2.28) and the voltage drop  $\mathbf{v}_M = \mathbf{A}_M^\top \mathbf{u}$

$$\mathbf{M}_\sigma \frac{d}{dt} \bar{\mathbf{a}} + \mathbf{k}_\nu(\bar{\mathbf{a}}) \bar{\mathbf{a}} = \mathbf{M}_\sigma \mathbf{X}_M \mathbf{A}_M^\top \mathbf{u}, \quad \text{with} \quad \bar{\mathbf{a}}(t_0) = \bar{\mathbf{a}}_0, \quad (3.10a)$$

where the total current (3.7) is equivalently given by

$$\mathbf{i}_M = \mathbf{X}_M^\top \mathbf{k}_\nu(\bar{\mathbf{a}}) \bar{\mathbf{a}}. \quad (3.10b)$$

Each column of the matrix  $\mathbf{X}_M$  is a coupling vector  $\mathbf{X}_m$  that corresponds to the region  $\Omega^{(m)}$ . The branch currents are gathered in the MQS device's current vector  $\mathbf{i}_M$ .

### 3.2.1 More Conductor Models

The solid conductor was characterized by the equipotential voltage at the reference plane. This property followed immediately from Maxwell's Equations. Now, if the conductor consists of very thin strands, one would be forced to resolve each strand in the grid. Obviously this is computationally not efficient. Thus other conductor models were introduced to overcome this impasse, e.g. the stranded and foil model, [39]. We discuss here briefly the stranded conductor, whose region is denoted by  $\Omega_{\text{str}}$ .

**Assumption 3.6** (Stranded conductor). *The current at the reference plain in a stranded conductor is homogeneously distributed.*

This assumption implies that we neglect the insulator around each strand and that there are no eddy currents in the strands, i.e., their diameter is below the skin-depth. The model accounts for that by using a modified conductivity tensor that vanishes in the stranded conductor region. The homogeneous current distribution allows to impose the current by a coupling vector similarly to the voltage drops for the solid model  $\mathbf{X}_m$ . The source current density is given for stranded conductors by

$$\widehat{\mathbf{j}}_{\text{str}} = \mathbf{X}_{\text{str}} \mathbf{i}_{\text{str}} \quad (3.11)$$

with a divergence-free stranded conductor coupling vector  $\mathbf{X}_{\text{str}}$ . The construction is similar to the solid conductor coupling vector, starting from a sparse coupling vector  $\bar{\mathbf{X}}_{\text{str}}$ . The only difference is the scaling to account for the area of the reference plane and the number

of windings, [49, 48, 43]. Using the no-smearing Ass. 3.5, the coupled system for solid and stranded conductors reads

$$\mathbf{M}_{\bar{\sigma}} \frac{d}{dt} \hat{\mathbf{a}} + \mathbf{k}_{\nu} \hat{\mathbf{a}} = \mathbf{M}_{\sigma} \mathbf{X}_{\text{sol}} \mathbf{v}_{\text{sol}} + \mathbf{X}_{\text{str}} \mathbf{i}_{\text{str}} \quad (3.12a)$$

with gauging, initial and boundary conditions for  $\hat{\mathbf{a}}$  and the coupling equations

$$-\frac{d}{dt} \mathbf{X}_{\text{sol}}^{\top} \mathbf{M}_{\sigma} \hat{\mathbf{a}} + \mathbf{R}_{\text{sol}}^{-1} \mathbf{v}_{\text{sol}} = \mathbf{i}_{\text{sol}}, \quad \mathbf{R}_{\text{sol}}^{-1} := \mathbf{X}_{\text{sol}}^{\top} \mathbf{M}_{\sigma} \mathbf{X}_{\text{sol}}, \quad (3.12b)$$

$$-\frac{d}{dt} \mathbf{X}_{\text{str}}^{\top} \hat{\mathbf{a}} + \mathbf{R}_{\text{str}} \mathbf{i}_{\text{str}} = \mathbf{v}_{\text{str}}, \quad \mathbf{R}_{\text{str}} := \mathbf{X}_{\text{str}} \mathbf{M}_{\sigma}^{+} \mathbf{X}_{\text{str}}, \quad (3.12c)$$

where the superscript ‘+’ denotes a Moore-Penrose pseudo-inverse [61]. The matrices  $\mathbf{R}_{\star}$  denote extracted dc resistances, cf. (3.7). The modified conductivity matrix

$$\mathbf{M}_{\bar{\sigma}} := \sum_{\Omega_r \neq \Omega_{\text{str}}} \mathbf{M}_{\sigma, r} \quad (3.13)$$

corresponds to the global conductivity matrix  $\mathbf{M}_{\sigma}$  defined in (2.16), but without the contributions from stranded conductors. This has consequences for the gauging, Section 2.2.3 because modified conductivity matrix  $\mathbf{M}_{\bar{\sigma}}$  has a larger nullspace

**Remark 3.2** (Gauging in presence of stranded conductors). If stranded conductors are present, a new projector  $\mathbf{Q}_{\bar{\sigma}}$  onto  $\text{Ker } \mathbf{M}_{\bar{\sigma}}$  must be defined analogously to  $\mathbf{Q}_{\sigma}$  in Definition 2.8.

Similarly to the definition of the lumped dc resistances in (3.12b) and (3.12c), an inductance matrix can be extracted. The approach corresponds to an 1 Ampere excitation  $\hat{\mathbf{j}}_{\text{str}} = \mathbf{X}_{\text{str}}$  of the curl-curl matrix, [111].

**Lemma 3.7** (Inductance extraction). *If the stranded conductor coupling vector is divergence free, i.e.,  $\mathbf{S} \mathbf{X}_{\text{str}} = 0$  and the curl-curl matrix  $\mathbf{k}_{\nu}$  is symmetric positive semi-definite, then the extracted inductance*

$$\mathbf{L}_{\text{str}}(\hat{\mathbf{a}}) := \mathbf{X}_{\text{str}}^{\top} \mathbf{k}_{\nu}^{+}(\hat{\mathbf{a}}) \mathbf{X}_{\text{str}} \quad (3.14)$$

*is symmetric positive definite.*

*Proof.* We can add a grad-div regularization, without altering the inductance matrix, i.e.,

$$\mathbf{L}_{\text{str}}(\cdot) = \mathbf{X}_{\text{str}}^{\top} \mathbf{k}_{\nu}^{+}(\cdot) \mathbf{X}_{\text{str}} = \mathbf{X}_{\text{str}}^{\top} \mathbf{k}_{\nu}^{+}(\cdot) \mathbf{X}_{\text{str}} + \mathbf{X}_{\text{str}}^{\top} \mathbf{S}^{\top} \mathbf{S} \mathbf{X}_{\text{str}} = \mathbf{X}_{\text{str}}^{\top} (\mathbf{k}_{\nu}^{+}(\cdot) + \mathbf{S}^{\top} \mathbf{S}) \mathbf{X}_{\text{str}}.$$

The inner matrix is symmetric positive definite, see Remark 4.1 and thus the nullspace of this matrix is given by  $\mathbf{X}_{\text{str}}$ . This matrix has full column rank, thus we conclude symmetric positive definiteness.  $\square$

This reveals that a stranded conductor model corresponds to the series connection of a lumped (nonlinear) inductance  $\mathbf{L}_{\text{str}}$  and a resistance  $\mathbf{R}_{\text{str}}$ .

**Lemma 3.8** (Structural equivalence of conductor models, [117]). *The solid and stranded conductor models are structural equivalent, i.e., the stranded conductor can be written as a solid conductor with a particular conductivity matrix.*

*Proof.* To show the structural equivalence of the solid and stranded model, [117], we left-multiply (3.12c) by  $\mathbf{X}_{\text{str}}\mathbf{R}_{\text{str}}^{-1}$  and add the result to (3.12a)

$$\left(\mathbf{M}_{\tilde{\sigma}} + \underbrace{\mathbf{X}_{\text{str}}\mathbf{R}_{\text{str}}^{-1}\mathbf{X}_{\text{str}}^{\top}}_{=:\mathbf{M}_{\text{equiv}}}\right)\frac{d}{dt}\hat{\mathbf{a}} + \mathbf{k}_{\nu}(\hat{\mathbf{a}}) = \mathbf{M}_{\sigma}\mathbf{X}_{\text{sol}}\mathbf{v}_{\text{sol}} + \underbrace{\mathbf{X}_{\text{str}}\mathbf{R}_{\text{str}}^{-1}\mathbf{X}_{\text{str}}^{\top}}_{=:\mathbf{M}_{\text{equiv}}}\underbrace{\mathbf{M}_{\sigma,\text{str}}^+}_{=:\mathbf{X}_{\text{equiv}}}\mathbf{X}_{\text{str}}\mathbf{R}_{\text{str}}^{-1}\mathbf{v}_{\text{str}}. \quad (3.15a)$$

Then left-multiplying (3.15a) by  $\mathbf{R}_{\text{str}}^{-1}\mathbf{X}_{\text{str}}^{\top}\mathbf{M}_{\sigma,\text{str}}^+$  and adding to (3.12c) gives

$$\underbrace{\mathbf{R}_{\text{str}}^{-1}\mathbf{X}_{\text{str}}^{\top}\mathbf{M}_{\sigma,\text{str}}^+}_{=:\mathbf{X}_{\text{equiv}}^{\top}}\mathbf{k}_{\nu}(\hat{\mathbf{a}}) = \mathbf{i}_{\text{str}}. \quad (3.15b)$$

This is the same structure as for solid conductors only, i.e., (3.10a)-(3.10b). Note that the new (artificial) conductivity matrix  $\mathbf{M}_{\text{equiv}}$  contributes with only one (positive) eigenvalue for each stranded conductor to the global conductivity matrix.  $\square$

Similarly other conductor models can be rewritten in the form of a solid conductor with special conductivity matrix, e.g. the foil model, [39]. In the following chapters we will chose the formulation that is more convenient for the corresponding section. Due to the equivalence above the analysis and methods remain applicable for all conductor models. For example the following field/circuit coupled system Section 3.4 is given in the solid conductor formulation.

### 3.3 Semiconductor Device Model

In contrast to the MQS device, the semiconductor model as a whole is identified as a circuit branch. In the coupled device/circuit setting the coupling is carried out by the prescription of the voltage drop  $\mathbf{v}_{\text{D}}$  at the contacts of the device, [26, 1]. In the following the drift-diffusion model is recapitulated in the particular case of a pn-diode, [123, 119]. It is crucial for the physical correctness of the coupled problem that the semiconductor and MQS device domains are separate problems. In practice the spatial discretization of both problems is done independently to resolve the effects on the different scales.

The computational domain is denoted by  $\Omega^{(\text{D})}$ . In this simple model it is not further partitioned into regions, only the boundary is divided into Dirichlet and Neumann type

$$\Gamma^{(\text{D})} = \Gamma_{\text{dir}}^{(\text{D})} \cup \Gamma_{\text{neu}}^{(\text{D})} \quad \text{where} \quad \Gamma^{(\text{D})} := \partial\Omega^{(\text{D})}.$$

The drift-diffusion model is a microscopic correction of the macroscopic Poisson equation, i.e., the electrostatic approximation (3.5). It reads

$$-\nabla \cdot \left( \vec{\varepsilon} \nabla \varphi \right) = \rho \quad \text{where} \quad \rho = q(n - p - C(\vec{r})) \quad (3.16)$$

where the refined electric flux density  $\rho$  is given in dependence of the elementary charge  $q$ , the electron and hole densities  $n, p$  and the doping concentration  $C(\vec{r})$ . The densities are determined by the additional conservation laws

$$\frac{\partial n}{\partial t} - q^{-1} \nabla \cdot \vec{J}_n = -R, \quad \vec{J}_n = \mu_n (U_T \nabla n - n \nabla \varphi), \quad (3.17a)$$

$$\frac{\partial p}{\partial t} + q^{-1} \nabla \cdot \vec{J}_p = -R, \quad \vec{J}_p = -\mu_p (U_T \nabla p + p \nabla \varphi), \quad (3.17b)$$

where  $R = R(n, p)$  denotes the generation-recombination term,  $\mu_n, \mu_p$  are the mobility parameters and  $U_T$  is the thermal voltage. The total current leaving the device at terminal  $\Gamma_k^{(D)} \subset \Gamma_{\text{dir}}^{(D)}$  ( $k = 1, 2$ ) is then given by

$$\mathbf{i}_D = \int_{\Gamma_k^{(D)}} \left( \vec{J}_{\text{disp}} - (\vec{J}_n + \vec{J}_p) \right) \cdot d\vec{A} \quad \text{where} \quad \vec{J}_{\text{disp}} = \vec{\varepsilon} \frac{\partial}{\partial t} \nabla \varphi \quad (3.17c)$$

denotes the displacement current.

Due to charge conservation the absolute current  $\mathbf{i}_D$  is the same. The integrals for  $k = 1, 2$  differ only by sign. The model is supplemented with initial conditions for  $n, p$  at  $t = t_0$  and boundary conditions for  $\varphi, n, p$  on the Dirichlet boundary  $\Gamma_{\text{dir}}^{(D)}$  and for  $\vec{J}_n, \vec{J}_p, \nabla \varphi$  on the Neumann boundary  $\Gamma_{\text{neu}}^{(D)}$ . Space discretization, e.g. by exponentially fitted mixed finite elements [90, 26], yields the semi-discrete problem

$$\mathbf{M}_n(\Phi) \frac{d}{dt} \mathbf{n} + \mathbf{K}_n(\Phi) \mathbf{n} = \mathbf{r}_n(\mathbf{p}, \Phi), \quad \tilde{\mathbf{S}} \mathbf{M}_\varepsilon \tilde{\mathbf{S}}^\top \Phi = \mathbf{q}(\mathbf{n}, \mathbf{p}, \mathbf{v}_D), \quad (3.18a)$$

$$\mathbf{M}_p(\Phi) \frac{d}{dt} \mathbf{p} + \mathbf{K}_p(\Phi) \mathbf{p} = \mathbf{r}_p(\mathbf{n}, \Phi), \quad \mathbf{i}_D = \mathbf{j}_D(\mathbf{n}, \mathbf{p}, \Phi, \frac{d}{dt} \Phi), \quad (3.18b)$$

with (regular) matrix functions  $\mathbf{M}_n, \mathbf{M}_p, \mathbf{K}_n, \mathbf{K}_p$  and the Laplacian  $\tilde{\mathbf{S}} \mathbf{M}_\varepsilon \tilde{\mathbf{S}}^\top$ . The bold symbols represent the vectors containing the discrete approximations of the corresponding continuous quantities in (3.17). The lumped current  $\mathbf{i}_D$  is the discrete approximation of the total current and it can be obtained in a post-processing step, i.e., it is not necessary to solve a differential equation for  $\Phi$ . Finally  $\mathbf{v}_D$  denotes the the voltage drop applied to the device. It is determined by the surrounding circuit. The boundary conditions are incorporated in the functions  $\mathbf{r}_n, \mathbf{r}_p$  and  $\mathbf{q}$ .

The displacement current in (3.18b) can be expressed in terms of the time derivative of the applied voltage drop  $\mathbf{v}_D$ , [2]. This *capacitance extraction* gives the following equivalent definition of the coupling current

$$\mathbf{i}_D = \mathbf{C}_D \frac{d}{dt} \mathbf{v}_D - \mathbf{i}_{\text{SD}} \quad \text{with} \quad \mathbf{i}_{\text{SD}} := \mathbf{j}_{\text{SD}}(\mathbf{n}, \mathbf{p}, \Phi). \quad (3.18c)$$

The capacitance  $\mathbf{C}_D$  may either be extracted from the discrete model or computed analytically, e.g.  $\mathbf{C}_D = \varepsilon A_D / l_D$  for a cubic diode with isotropic permittivity  $\varepsilon$ , length  $l_D$  and cross-section  $A_D$ .

This reveals that the semiconductor model corresponds to the parallel connection of a lumped capacitance and a lumped (nonlinear) resistance.

**Remark 3.3** (Singular mass matrices). The spatial discretization (e.g. using mixed finite elements) of the semiconductor (3.18) might yield singular mass matrices  $\mathbf{M}_n, \mathbf{M}_p$ . This turns  $\mathbf{n}, \mathbf{p}$  in some discretization nodes into algebraic variables, [26]. However, this does not affect the following analysis; projectors are defined similarly to Definition 2.8. This particular case is disregarded in the following.

### 3.4 Conclusions

At the conclusion of this chapter we assemble the equations of the spatial discrete models into one system of equations: the extended circuit model (3.2), (3.1b)-(3.1e), the MQS device model (3.10a)-(3.10b) and the drift-diffusion model (3.18). The coupled PDAE system is given by the extended network equations (for simplicity of notation in the notation of traditional MNA)

$$\mathbf{A}_C \frac{d}{dt} \mathbf{q}_C(\mathbf{A}_C^\top \mathbf{u}, t) + \mathbf{A}_R \mathbf{g}_R(\mathbf{A}_R^\top \mathbf{u}, t) + \mathbf{A}_L \mathbf{i}_L + \mathbf{A}_V \mathbf{i}_V + \mathbf{A}_I \mathbf{i}_s(t) \quad (3.19a)$$

$$+ \mathbf{A}_M \mathbf{X}_M^\top \mathbf{k}_\nu(\widehat{\mathbf{a}}) \widehat{\mathbf{a}} + \mathbf{A}_D \mathbf{j}_D(\mathbf{n}, \mathbf{p}, \Phi, \frac{d}{dt} \Phi) = 0$$

$$\frac{d}{dt} \phi_L(\mathbf{i}_L, t) - \mathbf{A}_L^\top \mathbf{u} = 0 \quad (3.19b)$$

$$\mathbf{A}_V^\top \mathbf{u} - \mathbf{v}_s(t) = 0 \quad (3.19c)$$

coupled to the MQS device problem

$$\mathbf{M}_\sigma \frac{d}{dt} \widehat{\mathbf{a}} + \mathbf{k}_\nu(\widehat{\mathbf{a}}) \widehat{\mathbf{a}} - \mathbf{M}_\sigma \mathbf{X}_M \mathbf{A}_M^\top \mathbf{u} = 0 \quad (3.19d)$$

and the semiconductor device problem

$$\mathbf{M}_n(\Phi) \frac{d}{dt} \mathbf{n} + \mathbf{K}_n(\Phi) \mathbf{n} - \mathbf{r}_n(\mathbf{p}, \Phi) = 0 \quad (3.19e)$$

$$\mathbf{M}_p(\Phi) \frac{d}{dt} \mathbf{p} + \mathbf{K}_p(\Phi) \mathbf{p} - \mathbf{r}_p(\mathbf{n}, \Phi) = 0 \quad (3.19f)$$

$$\tilde{\mathbf{S}} \mathbf{M}_\epsilon \tilde{\mathbf{S}}^\top \Phi - \mathbf{q}(\mathbf{n}, \mathbf{p}, \mathbf{A}_D^\top \mathbf{u}) = 0 \quad (3.19g)$$

with the unknown node potentials  $\mathbf{u} : \mathcal{I} \rightarrow \mathbb{R}^n$  without ground and the currents  $\mathbf{i}_L : \mathcal{I} \rightarrow \mathbb{R}^{n_L}$ ,  $\mathbf{i}_V : \mathcal{I} \rightarrow \mathbb{R}^{n_V}$  through inductors and voltage sources (where  $n_L$  and  $n_V$  denote their number), the magnetic vector potential  $\widehat{\mathbf{a}} : \mathcal{I} \rightarrow \mathbb{R}^{n_1}$  for the MQS device and the electric scalar potential  $\Phi : \mathcal{I} \rightarrow \mathbb{R}^{n_0}$ , the electron and hole densities  $\mathbf{n}, \mathbf{p} : \mathcal{I} \rightarrow \mathbb{R}^{n_0}$  for semiconductors. In the spirit of traditional MNA, the device currents  $\mathbf{i}_M$  and  $\mathbf{i}_D$  in Kirchhoff's current law (3.2) are replaced by the corresponding current assignments (3.10b) and (3.18b).

The following chapter discusses the expected numerical difficulties during the time-integration of problem (3.19a)-(3.19d) with  $\mathbf{A}_D = \square$  in terms of the DAE-index. Later on, in Section 5 efficient multirate strategies for the solution in the time domain are discussed.

---

## 4 DAE-Index Analysis

Differential/algebraic equations are structurally different from ordinary differential equations. The additional algebraic constraints make those problems ‘infinitely’ stiff, so only implicit time-integration methods can be applied. Furthermore the errors of numerical approximations are more critical. To classify this criticality, the DAE-index was introduced, e.g. [25, 68]. There are different concepts, but roughly speaking the idea is that the index corresponds to the highest derivative of an input function that enters the problem. This highest derivative affects the numerical approximation most severely. In the index-0 case, which corresponds to an ordinary differential equation, there is no such derivative. On the other hand systems of index larger than 1 are called higher index problems and suffer from higher derivatives. They may require special numerical treatment.

In this chapter we will analyze the DAE-index of the coupled system (3.19). It consists of contributions from three different subproblems. The index results for circuits containing semiconductor devices based on the drift-diffusion equation are well-known, e.g. [13, 120, 127]. Thus this section focuses on the index-analysis of the field/circuit coupled problem. The extension to all three problems is straightforward when using topological assumptions that keep the circuit’s branches of both PDE problems separate.

In the following we employ the tractability index. This concept is often used in the circuit analysis community. It gives a detailed view of the structure of the equations by using projectors and especially for the MNA the tractability index is only determined by the circuit’s topology, [54]. However, we expect for the field/circuit problem the same results when employing other index concepts.

### 4.1 Tractability Index

The tractability index is a projector-based approach, where a projector  $\mathbf{Q} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is an idempotent operator such that  $\mathbf{Q}^2 = \mathbf{Q}$ . The tractability concept provides an index characterization in terms of the original problem’s unknowns; it also leads to a precise solution description and requires low smoothness of the involved functions [63].

We deal here with a special class of DAEs that can be written with a properly stated leading term. In particular DAEs stemming from MNA (see Section 4) with only basic elements can always be given in this formulation [91] and the DAE-index does not exceed two under passivity assumptions (Ass. 3.1), [54].

Let be given an abstract DAE

$$\mathbf{A} \frac{d}{dt} \mathbf{d}(\mathbf{x}, t) + \mathbf{b}(\mathbf{x}, t) = 0 \quad (4.1)$$

with a matrix  $\mathbf{A} \in \mathbb{R}^{l \times m}$ , coefficient functions  $\mathbf{d}(\mathbf{x}, t) \in \mathbb{R}^m$  and  $\mathbf{b}(\mathbf{x}, t) \in \mathbb{R}^l$  that are

assumed continuous in their arguments and smooth. Their partial derivatives are

$$\mathbf{d}'_{\mathbf{x}}(\mathbf{x}, t) := \frac{\partial}{\partial \mathbf{x}} \mathbf{d}(\mathbf{x}, t) \quad \text{and} \quad \mathbf{b}'_{\mathbf{x}}(\mathbf{x}, t) := \frac{\partial}{\partial \mathbf{x}} \mathbf{b}(\mathbf{x}, t).$$

The unknown solution is  $\mathbf{x} = \mathbf{x}(t) \in \mathcal{D} \subset \mathbb{R}^l$ ,  $t \in \mathcal{I} = [t_0, t_e]$ .

**Definition 4.1** (Properly stated leading term, [92]). The DAE (4.1) has a properly stated leading term if and only if

$$\text{Ker } \mathbf{A} \oplus \text{Im } \mathbf{d}'_{\mathbf{x}}(\mathbf{x}, t) = \mathbb{R}^n \quad \text{for all } \mathbf{x} \in \mathcal{D}, t \in \mathcal{I},$$

and if there is a representing projector  $\mathbf{R} \in C^1(\mathcal{I}, \mathbb{R}^n)$ ,  $\text{Ker } \mathbf{A} = \text{Ker } \mathbf{R}(t)$ ,  $\text{Im } \mathbf{d}'_{\mathbf{x}}(\mathbf{x}, t) = \text{Im } \mathbf{R}(t)$  and  $\mathbf{d}(\mathbf{x}, t) = \mathbf{R}(t) \mathbf{d}(\mathbf{x}, t)$  for all  $\mathbf{x} \in \mathcal{D}$  and  $t \in \mathcal{I}$ .

The following definitions are used to discuss the index, see [91].

**Definition 4.2** (Matrix chain and subspaces). Given the DAE (4.1), we define recursively the following objects:

$$\begin{aligned} \mathbf{G}_0(\mathbf{x}, t) &:= \mathbf{A} \mathbf{d}'_{\mathbf{x}}(\mathbf{x}, t), \\ \mathbf{N}_0(\mathbf{x}, t) &:= \text{Ker } \mathbf{G}_0(\mathbf{x}, t), \\ \mathbf{P}_0(\mathbf{x}, t) &:= \mathbf{I} - \mathbf{Q}_0(\mathbf{x}, t), \quad \mathbf{Q}_0(\mathbf{x}, t) \text{ projector onto } \mathbf{N}_0(\mathbf{x}, t), \\ \mathbf{S}_0(\mathbf{x}, t) &:= \{\mathbf{z} \in \mathbb{R}^m \mid \mathbf{b}'_{\mathbf{x}}(\mathbf{x}, t) \mathbf{z} \in \text{Im } \mathbf{G}_0(\mathbf{x}, t)\}, \\ \mathbf{G}_1(\mathbf{x}, t) &:= \mathbf{G}_0(\mathbf{x}, t) + \mathbf{b}'_{\mathbf{x}}(\mathbf{x}, t) \mathbf{Q}_0(\mathbf{x}, t), \\ \mathbf{N}_1(\mathbf{x}, t) &:= \text{Ker } \mathbf{G}_1(\mathbf{x}, t) \\ \mathbf{S}_1(\mathbf{x}, t) &:= \{\mathbf{z} \in \mathbb{R}^m \mid \mathbf{b}'_{\mathbf{x}}(\mathbf{x}, t) \mathbf{P}_0(\mathbf{x}, t) \mathbf{z} \in \text{Im } \mathbf{G}_1(\mathbf{x}, t)\}. \end{aligned}$$

The definitions above allow to characterize the tractability-index.

**Definition 4.3** (Tractability index, [92]). The DAE (4.1) with a properly stated leading term is called DAE of (tractability) index-0 if and only if

$$\mathbf{N}_0(\mathbf{x}, t) = \{0\} \quad \text{for all } \mathbf{x} \in \mathcal{D}, t \in \mathcal{I}$$

or it is of index-1 if and only if

$$(\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t) = \{0\} \quad \text{for all } \mathbf{x} \in \mathcal{D}, t \in \mathcal{I}$$

or it is of index-2 if and only if

$$(\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t) = \text{const.} \quad \text{and} \quad (\mathbf{N}_1 \cap \mathbf{S}_1)(\mathbf{x}, t) = \{0\} \quad \text{for all } \mathbf{x} \in \mathcal{D}, t \in \mathcal{I}.$$

## 4.2 Index Analysis for the Field/Circuit Problem

The Kronecker-index, [68], for the eddy-current problem was studied first by [101]. Later a simple source coupling, i.e., a linear 2D MQS device without an external network, was analyzed by Tsukerman, [131]. In [117], the differential-index-1 was investigated for a

linear 3D model with an attached network. Now following [8], the nonlinear field/circuit problem in 3D is studied using the tractability index concept, [63]. This allows for a deeper understanding of the index-2 case, for example compared to [117]. For the tractability analysis the coupled system without semiconductors ( $\mathbf{A}_D = \mathbf{[]}$ ), i.e., equations (3.19a-3.19d), is given in the notation of traditional MNA with a properly stated leading term, i.e., in the form of (4.1)

$$\begin{aligned} \begin{bmatrix} \mathbf{A}_C & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbf{M}_\sigma \end{bmatrix} \frac{d}{dt} \begin{bmatrix} \mathbf{A}_C^+ \mathbf{A}_C \mathbf{q}_C(\cdot) \\ \phi_L(\cdot) \\ \mathbf{M}_\sigma^+ \mathbf{M}_\sigma \hat{\mathbf{a}} \end{bmatrix} \\ + \begin{bmatrix} \mathbf{A}_R \mathbf{g}_R(\cdot) + \mathbf{A}_L \mathbf{i}_L + \mathbf{A}_V \mathbf{i}_V + \mathbf{A}_M \mathbf{X}^\top \mathbf{k}_\nu(\cdot) \hat{\mathbf{a}} + \mathbf{A}_I \mathbf{i}_s(t) \\ -\mathbf{A}_L^\top \mathbf{u} \\ \mathbf{A}_V^\top \mathbf{u} - \mathbf{v}_s(t) \\ \mathbf{k}_\nu(\cdot) \hat{\mathbf{a}} - \mathbf{M}_\sigma \mathbf{X} \mathbf{A}_M^\top \mathbf{u} \end{bmatrix} = 0. \end{aligned} \quad (4.2)$$

where the unknown vector is given by  $\mathbf{x}^\top = [\mathbf{u}^\top \ \mathbf{i}_L^\top \ \mathbf{i}_V^\top \ \hat{\mathbf{a}}^\top]$ . Let  $\mathbf{Q}_\sigma$  and  $\mathbf{P}_\sigma = \mathbf{I} - \mathbf{Q}_\sigma$  be defined as in Definition 2.8, i.e.,  $\mathbf{Q}_\sigma$  is the constant projector onto  $\text{Ker } \mathbf{M}_\sigma$ , such that Ass. 2.13 ensures

$$\text{Ker}(\mathbf{M}_\sigma + \mathbf{Q}_\sigma^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma) = \{0\}$$

which is a necessary condition for a uniquely solvable problem. Let further  $\mathbf{P}_C = \mathbf{I} - \mathbf{Q}_C$  be defined using the projector  $\mathbf{Q}_C$  onto  $\text{Ker } \mathbf{A}_C^\top$  as in the classical MNA case, [54]. Then (4.2) has a properly stated leading term with projector

$$\mathbf{R} := \begin{bmatrix} \mathbf{P}_C & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{P}_\sigma \end{bmatrix} = \begin{bmatrix} \mathbf{A}_C^+ \mathbf{A}_C & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{M}_\sigma^+ \mathbf{M}_\sigma \end{bmatrix}.$$

Compared with the circuit only case,  $\mathbf{R}$  features an additional row and column for  $\mathbf{P}_\sigma$ , because of the additional curl-curl equation. The curl-curl equation contributes also with the conductivity matrix  $\mathbf{M}_\sigma$  to the overall mass matrix

$$\mathbf{G}_0(\mathbf{x}, t) := \mathbf{A} \mathbf{d}'_{\mathbf{x}}(\mathbf{x}, t) = \begin{bmatrix} \mathbf{A}_C \mathbf{C}(\cdot) \mathbf{A}_C^\top & 0 & 0 & 0 \\ 0 & \mathbf{L}(\cdot) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{M}_\sigma \end{bmatrix}. \quad (4.3)$$

If the mass matrix above is regular, all equations are differential equations, such that the problem is an ODE. This is the case for the following class of circuits, [8].

**Theorem 4.1** (Index-0). *Let Ass. 2.13, 3.1 and 3.2 be fulfilled. Then the DAE (4.2) has index-0 if and only if there is no voltage source, a tree containing capacitors only and*

- (a) *the circuit does not contains MQS devices, or*
- (b) *the MQS devices only have conductive regions.*

*Proof.* The matrix  $\mathbf{G}_0$  is nonsingular iff all blocks on the diagonal have full rank. The matrices  $\mathbf{C}$  and  $\mathbf{L}$  are positive definite by assumption and  $\text{Ker } \mathbf{A}_C^\top = \{0\}$  is trivial iff the circuit has a tree containing capacitors only, see Remark 3.1. The third row/column vanishes iff there are no voltage sources. Finally,  $\text{Ker } \mathbf{M}_\sigma = \{0\}$  iff the circuit does not include MQS devices (this is obviously the classical MNA case, [54]) or the domains of all MQS device are conductive ( $\mathbf{M}_\sigma$  has full rank).  $\square$

## 4.2.1 Index-1 Conditions

The index-0 (ODE) case above is very restrictive. For example flux/charge oriented MNA or the introduction of the currents  $\mathbf{i}_M$  as variables turns the problem into a DAE, [54]. When generalizing the problem to the index-1 case, additional elements, i.e., voltage sources, can be considered. This requires the following definitions, [8]

$$\mathbf{Q}_0 := \begin{bmatrix} \mathbf{Q}_C & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{Q}_\sigma \end{bmatrix}, \quad \mathbf{b}'_x(\mathbf{x}, t) = \begin{bmatrix} \mathbf{A}_R \mathbf{G}(\cdot) \mathbf{A}_R^\top & \mathbf{A}_L & \mathbf{A}_V & \mathbf{A}_M \mathbf{X}^\top \mathbf{K}_\nu(\cdot) \\ -\mathbf{A}_L^\top & 0 & 0 & 0 \\ \mathbf{A}_V^\top & 0 & 0 & 0 \\ -\mathbf{M}_\sigma \mathbf{X} \mathbf{A}_M^\top & 0 & 0 & \mathbf{K}_\nu(\cdot) \end{bmatrix},$$

and

$$\mathbf{b}'_x(\mathbf{x}, t) \mathbf{Q}_0 = \begin{bmatrix} \mathbf{A}_R \mathbf{G}(\cdot) \mathbf{A}_R^\top \mathbf{Q}_C & 0 & \mathbf{A}_V & \mathbf{A}_M \mathbf{X}^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \\ -\mathbf{A}_L^\top \mathbf{Q}_C & 0 & 0 & 0 \\ \mathbf{A}_V^\top \mathbf{Q}_C & 0 & 0 & 0 \\ -\mathbf{M}_\sigma \mathbf{X} \mathbf{A}_M^\top \mathbf{Q}_C & 0 & 0 & \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \end{bmatrix},$$

where  $\mathbf{Q}_0$  is a constant projector onto  $\text{Ker } \mathbf{G}_0(\mathbf{x}, t)$ . Again, the matrices  $\mathbf{Q}_0$  and  $\mathbf{b}'_x(\mathbf{x}, t) \mathbf{Q}_0$  are similar to the well-known  $3 \times 3$  block structure in MNA index analysis. Thus the index-1 proof as given in [8] is a straightforward extension of [54].

**Theorem 4.2** (Index-1). *Let Ass. 2.13, 3.1 and 3.2 be fulfilled and the circuit contains at least an MQS device, a voltage source or there is no tree containing capacitors only. Then the DAE (4.2) has index-1 if and only if there is neither*

- (a) a LIM-cutset, i.e., a cutset consisting of inductances, current sources and MQS devices only, nor
- (b) a CV-loop, i.e., a loop consisting of capacitances and at least a voltage source only.

*Proof.* The intersection  $(\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t)$  is analyzed in the following, see Definition 4.2. Let  $\mathbf{W}_0(\mathbf{x}, t)$  denote a projector along  $\text{Im } \mathbf{G}_0(\mathbf{x}, t)$ . Consequently  $\mathbf{W}_0^\top(\mathbf{x}, t)$  is a projector onto  $\text{Ker } \mathbf{G}_0^\top(\mathbf{x}, t)$ . Symmetry gives  $\text{Ker } \mathbf{G}_0^\top(\mathbf{x}, t) = \text{Ker } \mathbf{G}_0(\mathbf{x}, t)$  and thus we can choose  $\mathbf{W}_0^\top(\mathbf{x}, t) = \mathbf{Q}_0$ . We find starting from Definition 4.2

$$(\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t) = \text{Ker } \mathbf{G}_0(\mathbf{x}, t) \cap \text{Ker } (\mathbf{W}_0 \mathbf{b}'_x)(\mathbf{x}, t) \quad (4.4)$$

with

$$\mathbf{S}_0(\mathbf{x}, t) = \{\mathbf{z} \in \mathbb{R}^m \mid \mathbf{b}'_x(\mathbf{x}, t) \mathbf{z} \in \text{Im } \mathbf{G}_0(\mathbf{x}, t)\}$$

$$\begin{aligned}
 &= \{ \mathbf{z} \in \mathbb{R}^m \mid \mathbf{W}_0(\mathbf{x}, t) \mathbf{b}'_x(\mathbf{x}, t) \mathbf{z} = 0 \} \\
 &= (\mathbf{W}_0 \mathbf{b}'_x)(\mathbf{x}, t)
 \end{aligned}$$

and consequently

$$(\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t) = \text{Im } \mathbf{Q}_0 \cap \text{Ker } \mathbf{W}_0 \mathbf{b}'_x(\mathbf{x}, t) \mathbf{Q}_0.$$

Exploiting the identity  $\mathbf{W}_0 = \mathbf{Q}_0^\top$  yields

$$\mathbf{W}_0 \mathbf{b}'_x(\mathbf{x}, t) \mathbf{Q}_0 = \begin{bmatrix} \mathbf{Q}_C^\top \mathbf{A}_R \mathbf{G}(\cdot) \mathbf{A}_R^\top \mathbf{Q}_C & 0 & \mathbf{Q}_C^\top \mathbf{A}_V & \mathbf{Q}_C^\top \mathbf{A}_M \mathbf{X}^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \\ 0 & 0 & 0 & 0 \\ \mathbf{A}_V^\top \mathbf{Q}_C & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{Q}_\sigma^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \end{bmatrix}.$$

We show that  $\mathbf{z} = 0$  when  $\mathbf{z} \in (\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t)$ . Let  $\mathbf{z}^\top = [\mathbf{z}_1^\top \ \mathbf{z}_2^\top \ \mathbf{z}_3^\top \ \mathbf{z}_4^\top]$ , then follows from  $\mathbf{Q}_0 \mathbf{z} = \mathbf{z}$

$$\mathbf{Q}_C \mathbf{z}_1 = \mathbf{z}_1, \quad (4.5)$$

$$\mathbf{Q}_\sigma \mathbf{z}_4 = \mathbf{z}_4, \quad (4.6)$$

$$\mathbf{z}_2 = 0, \quad (4.7)$$

and  $\mathbf{W}_0 \mathbf{b}'_x(\mathbf{x}, t) \mathbf{Q}_0 \mathbf{z} = 0$  implies

$$\mathbf{Q}_C^\top \mathbf{A}_R \mathbf{G}(\cdot) \mathbf{A}_R^\top \mathbf{Q}_C \mathbf{z}_1 + \mathbf{Q}_C^\top \mathbf{A}_V \mathbf{z}_3 + \mathbf{Q}_C^\top \mathbf{A}_M \mathbf{X}^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \mathbf{z}_4 = 0, \quad (4.8)$$

$$\mathbf{A}_V^\top \mathbf{Q}_C \mathbf{z}_1 = 0, \quad (4.9)$$

$$\mathbf{Q}_\sigma^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \mathbf{z}_4 = 0. \quad (4.10)$$

From (4.10) and  $\mathbf{Q}_\sigma^\top \mathbf{M}_\sigma = 0$  it follows immediately that

$$\mathbf{Q}_\sigma^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \mathbf{z}_4 = \mathbf{Q}_\sigma^\top (\mathbf{M}_\sigma + \mathbf{K}_\nu(\cdot)) \mathbf{Q}_\sigma \mathbf{z}_4 = 0,$$

where  $\mathbf{M}_\sigma + \mathbf{K}_\nu(\cdot)$  is positive definite by Ass. 2.13 and thus we achieve  $\mathbf{z}_4 = 0$  using (4.6).

Now we are back to the ‘classical’ case, i.e., without MQS device. From (4.5-4.10) it follows

$$\mathbf{Q}_C^\top \mathbf{A}_R \mathbf{G}(\cdot) \mathbf{A}_R^\top \mathbf{Q}_C \mathbf{z}_1 + \mathbf{Q}_C^\top \mathbf{A}_V \mathbf{z}_3 = 0, \quad (4.11)$$

$$\mathbf{A}_V^\top \mathbf{Q}_C \mathbf{z}_1 = 0, \quad (4.12)$$

$$\mathbf{Q}_C \mathbf{z}_1 = \mathbf{z}_1, \quad (4.13)$$

$$\mathbf{z}_2 = 0, \quad \mathbf{z}_4 = 0. \quad (4.14)$$

Left-multiplication of (4.11) by  $\mathbf{z}_1^\top$  and using (4.12) gives  $\mathbf{z}_1 \in \text{Ker } \mathbf{A}_R^\top \mathbf{Q}_C$  and thus  $\mathbf{Q}_C^\top \mathbf{A}_V \mathbf{z}_3 = 0$ . Now, from (4.13) follows

$$\mathbf{z}_1 \in \text{Ker } [\mathbf{A}_C \ \mathbf{A}_R \ \mathbf{A}_V]^\top \quad \text{and} \quad \mathbf{Q}_C^\top \mathbf{A}_V \mathbf{z}_3 = 0.$$

Then  $(\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t) = \{0\}$  holds iff there are neither *LIM*-cutsets nor *CV*-loops with at

least one voltage source.  $\square$

The intersection (4.4) can be given elegantly in terms of additional projectors from the classical MNA index-analysis, [54]: the constant projector  $\mathbf{Q}_{CRV} = \mathbf{Q}_C \mathbf{Q}_{V-C} \mathbf{Q}_{R-CV}$  onto  $\text{Ker}[\mathbf{A}_C \ \mathbf{A}_R \ \mathbf{A}_V]^\top$ , where  $\mathbf{Q}_{C-V}$ ,  $\mathbf{Q}_{V-C}$  and  $\mathbf{Q}_{R-CV}$  are constant projectors onto  $\text{Ker} \mathbf{Q}_C^\top \mathbf{A}_V$ ,  $\text{Ker} \mathbf{A}_V^\top \mathbf{Q}_C$  and  $\text{Ker} \mathbf{A}_R^\top \mathbf{Q}_C \mathbf{Q}_{C-V}$ . From those projectors this result follows immediately:

**Lemma 4.3.** *The dimension of the intersection*

$$(\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t) = \{z \in \mathbb{R}^n \mid \mathbf{z}_1 \in \text{Im } \mathbf{Q}_{CRV}, \mathbf{z}_3 \in \text{Im } \mathbf{Q}_{C-V}, [\mathbf{z}_2 \ \mathbf{z}_4] = 0\} \quad (4.15)$$

is constant.

This result will be exploited in the next section in Theorem 4.5, where we prove that the DAE-index is at most 2.

## 4.2.2 Index-2 Conditions

In the index-2 case, the main difficulty is to verify that the excitation of the field model is consistent and that the coupling vectors do not extract currents from non-conductive edges, [117]. To prove this (without additional assumptions), we need the following auxiliary result for pseudo-inverses and projectors:

**Remark 4.1.** Let  $\mathbf{K}$ ,  $\mathbf{K}^+$  and  $\mathbf{Q}$  denote an arbitrary matrix, its Moore-Penrose pseudo-inverse and a projector, respectively. If  $\mathbf{K} = \mathbf{Q}^\top \mathbf{K} \mathbf{Q}$  then it follows that  $\mathbf{K}^+ = \mathbf{Q} \mathbf{K}^+ \mathbf{Q}^\top$ .

*Proof.* One verifies easily that  $\mathbf{Q} \mathbf{K}^+ \mathbf{Q}^\top$  is a Moore-Penrose pseudoinverse of  $\mathbf{K} = \mathbf{Q}^\top \mathbf{K} \mathbf{Q}$  if  $\mathbf{K}^+$  is a pseudoinverse. The uniqueness of the Moore-Penrose pseudoinverse proves the equality of both inverses.  $\square$

With Remark 4.1 the following lemma from [8] can be proved, which is structurally similar to the extraction of a lumped inductance from the PDE model, see Lemma 3.7.

**Lemma 4.4** (Consistent excitation). *Let Ass. 2.13 and 3.4 be fulfilled, then*

$$\mathbf{X}^\top \mathbf{H}_k(\cdot) \mathbf{X} \quad \text{with} \quad \mathbf{H}_k(\cdot) = \mathbf{K}_\nu(\cdot) \left( \mathbf{K}_\nu(\cdot)^+ - (\mathbf{Q}_\sigma^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma)^+ \right) \mathbf{K}_\nu(\cdot)$$

is positive definite.

*Proof.* A straightforward computation using properties of the projector and pseudo-inverse (Remark 4.1), the soundness of the excitation  $\mathbf{X} = \mathbf{P}_\sigma \mathbf{X}$  and its divergence-freeness  $\mathbf{S} \mathbf{M} \mathbf{X} = 0$  (Lemma 3.4) yield

$$\mathbf{X}^\top \mathbf{H}_k(\cdot) \mathbf{X} = \mathbf{X}^\top \mathbf{T} \mathbf{Z}_k(\cdot) \mathbf{T}^\top \mathbf{X}$$

with a curl-curl matrix  $\mathbf{Z}_k(\cdot) := \mathbf{K}_\nu(\cdot) + \mathbf{M}^\top \tilde{\mathbf{S}}^\top \tilde{\mathbf{S}} \mathbf{M}$  that is fully regularized and the block-elimination  $\mathbf{T}(\cdot) := \mathbf{I} - (\mathbf{P}_\sigma^\top \mathbf{Z}_k(\cdot) \mathbf{Q}_\sigma)(\mathbf{Q}_\sigma^\top \mathbf{Z}_k(\cdot) \mathbf{Q}_\sigma)^+$ .

The matrix  $\mathbf{T}(\cdot)$  is regular with  $\mathbf{T}^{-1}(\cdot) = \mathbf{I} + (\mathbf{P}_\sigma^\top \mathbf{Z}_k(\cdot) \mathbf{Q}_\sigma)(\mathbf{Q}_\sigma^\top \mathbf{Z}_k(\cdot) \mathbf{Q}_\sigma)^+$  and  $\mathbf{X}$  has full column rank by construction. Thus the definiteness of  $\mathbf{Z}_k(\cdot)$  must be shown, where the only interesting elements are from  $\text{Ker } \mathbf{C} = \text{im } \tilde{\mathbf{S}}^\top$ , i.e.,

$$\mathbf{x}^\top \tilde{\mathbf{S}} \left( \mathbf{Z}_\sigma + \mathbf{M}^\top \tilde{\mathbf{S}}^\top \tilde{\mathbf{S}} \mathbf{M} \right) \tilde{\mathbf{S}}^\top \mathbf{x} > 0 \quad \text{for all } \mathbf{x} \neq 0$$

with  $\mathbf{Z}_\sigma$  as defined in (2.22). Positive definiteness follows in both cases

1. if  $\tilde{\mathbf{S}}^\top \mathbf{x} \in \text{Ker } \mathbf{M}$  then there is an  $\mathbf{y}$  such that  $\tilde{\mathbf{S}}^\top \mathbf{x} = \mathbf{Q}_\sigma \mathbf{y}$ . Thus the second summand vanishes and the first summand is positive because of Ass. 2.13.
2. else  $\mathbf{M} \tilde{\mathbf{S}}^\top \mathbf{x} \neq 0$  and thus the second summand is positive (because its kernel is  $\text{Ker } \mathbf{M} \tilde{\mathbf{S}}^\top$ ) and the first summand is non-negative.  $\square$

**Remark 4.2** (Schur complement). The matrix  $\mathbf{H}_k$  in Lemma 4.4 corresponds to a Schur-complement. Let us assume a convenient partitioning of the curl-curl-equations into equations for conducting and nonconductive domains  $\mathbf{M}_\sigma = \text{diag}(\mathbf{M}_{11}, \mathbf{0})$ . Then the projector onto  $\text{Ker } \mathbf{M}_\sigma$  can be given as  $\mathbf{Q}_\sigma = \text{diag}(\mathbf{0}, \mathbf{I})$ . With a corresponding block partitioning of the curl-curl matrix

$$\mathbf{K}_\nu(\cdot) = \begin{bmatrix} \mathbf{K}_{11}(\cdot) & \mathbf{K}_{12}(\cdot) \\ \mathbf{K}_{12}^\top(\cdot) & \mathbf{K}_{22}(\cdot) \end{bmatrix} \quad \text{follows} \quad \mathbf{H}_k(\cdot) = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (4.16)$$

with  $\mathbf{H}_{11} := \mathbf{K}_{12}(\cdot) \mathbf{K}_{22}^+(\cdot) \mathbf{K}_{12}^\top(\cdot)$ . The Schur complement is the simplest form of domain substructuring, [104]. It yields a reduced system because only the first block must be solved. This can be exploited when solving linear systems, see Section 5.4.

Now, having computed the important ingredient Lemma 4.4, the next element of the matrix chain Definition 4.2, i.e.,  $\mathbf{G}_1(\mathbf{x}, t) = \mathbf{G}_0(\mathbf{x}, t) + \mathbf{b}'_x(\mathbf{x}, t) \mathbf{Q}_0$  is analyzed to extract the topological conditions for the index-2 case.

$$\mathbf{G}_1(\mathbf{x}, t) = \begin{bmatrix} \mathbf{A}_C \mathbf{C}(\cdot) \mathbf{A}_C^\top + \mathbf{A}_R \mathbf{G}(\cdot) \mathbf{A}_R^\top \mathbf{Q}_C & \mathbf{0} & \mathbf{A}_V & \mathbf{A}_M \mathbf{X}^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \\ -\mathbf{A}_L^\top \mathbf{Q}_C & \mathbf{L}(\cdot) & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_V^\top \mathbf{Q}_C & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{M}_\sigma \mathbf{X} \mathbf{A}_M^\top \mathbf{Q}_C & \mathbf{0} & \mathbf{0} & \mathbf{M}_\sigma + \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \end{bmatrix}.$$

Following [8], the next proof will not be based on the classical procedure as given in [54], where the projector  $\mathbf{Q}_1$  onto  $\text{Ker } \mathbf{G}_1(\mathbf{x}, t)$  is computed. Instead it will be shown that the intersection  $\mathbf{N}_1 \cap \mathbf{S}_1$  is trivial, see Definition 4.2 in Section 4.1.

**Theorem 4.5** (Index-2). *Let Ass. 2.13, 3.1, 3.2, and 3.4 be fulfilled and let the circuit contain at least one MQS device or one voltage source or there is no tree containing capacitors only. Then the DAE (4.2) has index-2 if and only if there is either*

- (a) a LIM-cutset, i.e., a cutset consisting of inductances, current sources and MQS devices only, or
- (b) a CV-loop, i.e., a loop consisting of capacitances and at least a voltage source only.

*Proof.* Using Remark 4.1 the projector

$$\mathbf{W}_1(\mathbf{x}, t) = \begin{bmatrix} \mathbf{Q}_{\text{CRV}}^\top & 0 & 0 & -\mathbf{Q}_{\text{CRV}}^\top \mathbf{A}_M \mathbf{X}^\top \mathbf{K}_\nu(\cdot) (\mathbf{Q}_\sigma^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma)^\dagger \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{Q}_{\text{C-V}}^\top & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is defined and it holds true  $\text{Im } \mathbf{G}_1(\mathbf{x}, t) \subset \text{Ker } \mathbf{W}_1(\mathbf{x}, t)$ . Now, we can reformulate the object

$$\begin{aligned} \mathbf{S}_1(\mathbf{x}, t) &= \{z \in \mathbb{R}^n \mid \mathbf{b}'_x(\mathbf{x}, t) \mathbf{P}_0 z \in \text{Im } \mathbf{G}_1(\mathbf{x}, t)\} \\ &\subset \{z \in \mathbb{R}^n \mid \mathbf{W}_1(\mathbf{x}, t) \mathbf{b}'_x(\mathbf{x}, t) \mathbf{P}_0 z = 0\} =: \tilde{\mathbf{S}}_1(\mathbf{x}, t). \end{aligned}$$

Due to the inclusion  $\mathbf{S}_1(\mathbf{x}, t) \subset \tilde{\mathbf{S}}_1(\mathbf{x}, t)$  it is sufficient to show  $(\mathbf{N}_1 \cap \tilde{\mathbf{S}}_1)(\mathbf{x}, t) = \{0\}$ .

The new set  $\tilde{\mathbf{S}}_1$  defines the matrix

$$\mathbf{W}_1(\mathbf{x}, t) \mathbf{b}'_x(\mathbf{x}, t) \mathbf{P}_0 = \begin{bmatrix} 0 & \mathbf{Q}_{\text{CRV}}^\top \mathbf{A}_L & 0 & \mathbf{Q}_{\text{CRV}}^\top \mathbf{A}_M \mathbf{X}^\top \mathbf{H}_k(\cdot) \mathbf{P}_\sigma \\ 0 & 0 & 0 & 0 \\ \mathbf{Q}_{\text{C-V}}^\top \mathbf{A}_V^\top \mathbf{P}_C & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

with  $\mathbf{H}_k$  as defined in Lemma 4.4. The matrix imposes the following conditions on  $\mathbf{z}$  (this is now almost the same procedure as in the proof of Theorem 4.2):

$$\mathbf{Q}_{\text{CRV}}^\top \mathbf{A}_L \mathbf{z}_2 + \mathbf{Q}_{\text{CRV}}^\top \mathbf{A}_M \mathbf{X}^\top \mathbf{H}_k(\cdot) \mathbf{P}_\sigma \mathbf{z}_4 = 0, \quad (4.17)$$

$$\mathbf{Q}_{\text{C-V}}^\top \mathbf{A}_V^\top \mathbf{P}_C \mathbf{z}_1 = 0, \quad (4.18)$$

and  $\mathbf{G}_1(\mathbf{x}, t) z = 0$  gives

$$(\mathbf{A}_C \mathbf{C}(\cdot) \mathbf{A}_C^\top + \mathbf{A}_R \mathbf{G}(\cdot) \mathbf{A}_R^\top \mathbf{Q}_C) \mathbf{z}_1 + \mathbf{A}_V \mathbf{z}_3 + \mathbf{A}_M \mathbf{X}^\top \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma \mathbf{z}_4 = 0, \quad (4.19)$$

$$\mathbf{z}_2 - \mathbf{L}^{-1}(\cdot) \mathbf{A}_L^\top \mathbf{Q}_C \mathbf{z}_1 = 0, \quad (4.20)$$

$$\mathbf{A}_V^\top \mathbf{Q}_C \mathbf{z}_1 = 0, \quad (4.21)$$

$$-\mathbf{M}_\sigma \mathbf{X} \mathbf{A}_M^\top \mathbf{Q}_C \mathbf{z}_1 + (\mathbf{M}_\sigma + \mathbf{K}_\nu(\cdot) \mathbf{Q}_\sigma) \mathbf{z}_4 = 0. \quad (4.22)$$

The left-multiplication of equation (4.22) by  $(\mathbf{Q}_\sigma \mathbf{z}_4)^\top$  together with the Gauging Ass. 2.13 ( $\mathbf{M}_\sigma + \mathbf{K}_\nu(\cdot)$  is positive definite) gives

$$\mathbf{Q}_\sigma \mathbf{z}_4 = 0, \quad \text{i.e.,} \quad \mathbf{z}_4 = \mathbf{P}_\sigma \mathbf{z}_4. \quad (4.23)$$

From Equation (4.22) with  $\mathbf{P}_\sigma = \mathbf{M}_\sigma^+ \mathbf{M}_\sigma$ ,  $\mathbf{z}_4 = \mathbf{P}_\sigma \mathbf{z}_4$  and Lemma 4.4 ( $\mathbf{X} = \mathbf{P}_\sigma \mathbf{X}$ ) follows

$$\mathbf{z}_4 = \mathbf{X} \mathbf{A}_M^\top \mathbf{Q}_C \mathbf{z}_1. \quad (4.24)$$

A ‘classical’ procedure of MNA index-analysis is the left-multiplication of (4.19) by  $(\mathbf{Q}_C \mathbf{z}_1)^\top$  and using both (4.21) and (4.23) to arrive at

$$\mathbf{Q}_C \mathbf{z}_1 \in \text{Ker } \mathbf{A}_R^\top. \quad (4.25)$$

Still using standard techniques, equations (4.20), (4.25) and the definition of  $\mathbf{Q}_C$  yield

$$\mathbf{Q}_C \mathbf{z}_1 \in \text{Ker} [\mathbf{A}_C \mathbf{A}_R \mathbf{A}_V]^\top = \text{Im } \mathbf{Q}_{\text{CRV}},$$

and thus  $\mathbf{Q}_C \mathbf{z}_1 = \mathbf{Q}_{\text{CRV}} \mathbf{Q}_C \mathbf{z}_1$ . Now, substituting (4.20) and (4.24) in (4.17) gives

$$\mathbf{Q}_{\text{CRV}}^\top \mathbf{A}_L \mathbf{L}^{-1}(\cdot) \mathbf{A}_L^\top \mathbf{Q}_C \mathbf{z}_1 + \mathbf{Q}_{\text{CRV}}^\top \mathbf{A}_M \mathbf{X}^\top \mathbf{H}_k(\cdot) \mathbf{X} \mathbf{A}_M^\top \mathbf{Q}_C \mathbf{z}_1 = 0$$

where the important non-standard term  $\mathbf{X}^\top \mathbf{H}_k(\cdot) \mathbf{X}$  is positive definite as shown in Lemma 4.4. Therefore it follows that  $\mathbf{A}_L^\top \mathbf{Q}_C \mathbf{z}_1 = 0$  and  $\mathbf{A}_M^\top \mathbf{Q}_C \mathbf{z}_1 = 0$  (using  $\mathbf{Q}_C \mathbf{z}_1 = \mathbf{Q}_{\text{CRV}} \mathbf{Q}_C \mathbf{z}_1$ ). Hence  $\mathbf{Q}_C \mathbf{z}_1 \in \text{Ker} [\mathbf{A}_C \mathbf{A}_R \mathbf{A}_L \mathbf{A}_V \mathbf{A}_M]^\top$ , which is trivial due to Ass. 3.2 (without semiconductors  $\mathbf{A}_D = []$ ).  $\mathbf{Q}_C \mathbf{z}_1 = 0$  or equivalently  $\mathbf{P}_C \mathbf{z}_1 = \mathbf{z}_1$ .

The rest of the proof only uses standard arguments: it follows from (4.20) immediately  $\mathbf{z}_2 = 0$  and (4.19) can be restated as (using a ‘regularization’ of the capacitive term)

$$\mathbf{H}_C(\cdot) \mathbf{P}_C \mathbf{z}_1 = -\mathbf{A}_V \mathbf{z}_3 \quad \text{and thus} \quad \mathbf{z}_1 = -\mathbf{H}_C(\cdot)^{-1} \mathbf{A}_V \mathbf{z}_3$$

because  $\mathbf{H}_C(\cdot) = \mathbf{A}_C \mathbf{C}(\cdot) \mathbf{A}_C^\top + \mathbf{Q}_C^\top \mathbf{Q}_C$  is positive definite. Multiplying of (4.19) from the left by  $\mathbf{Q}_C^\top$  leads to  $\mathbf{Q}_C^\top \mathbf{A}_V \mathbf{z}_3 = 0$  and  $\mathbf{z}_3 \in \text{Im } \mathbf{Q}_{\text{C-V}}$  respectively. Finally (4.18) and  $\mathbf{z}_3 \in \text{Im } \mathbf{Q}_{\text{C-V}}$  give

$$\mathbf{Q}_{\text{C-V}}^\top \mathbf{A}_V^\top \mathbf{H}_C(\cdot)^{-1} \mathbf{A}_V \mathbf{Q}_{\text{C-V}} \mathbf{z}_3 = 0.$$

Hence  $\mathbf{A}_V \mathbf{z}_3 = 0$  and  $\mathbf{z}_3 = 0$  because  $\mathbf{A}_V$  has full column rank. Now  $\mathbf{H}_C(\cdot) \mathbf{z}_1 = 0$  implies  $\mathbf{z}_1 = 0$ . Thus it has been shown that  $(\mathbf{N}_1 \cap \tilde{\mathbf{S}}_1)(\mathbf{x}, t)$  and consequently  $(\mathbf{N}_1 \cap \mathbf{S}_1)(\mathbf{x}, t)$  is trivial iff there is a *LIM*-cutset or a *CV*-loop with at least one voltage source.  $\square$

**Remark 4.3** (Flux/charge-oriented MNA). Using the flux/charge-oriented formulation instead of the traditional MNA for system (4.2) does not change the index results for the index-1 and index-2 cases, [54].

While the time-integration of index-1 DAEs behaves numerically almost like integrating a stiff ODE, it can be shown using perturbation-index analysis that index-2 problems are more severe, [67]. On the other hand it is known for index-2 Hessenberg systems with linear index-2 variables, [25, 6], and for index-2 circuits, [128], that the numerical difficulties in time-integration are moderate. The derivative of the perturbation does not affect the differential (index-0) variables and thus they do not propagate and accumulate. The field/circuit coupled problem belongs to this class, because the MQS device does not contribute with additional index-2 other than the voltages  $\mathbf{v}_M$ , [8]. This is shown numerically in the example in Section 6.1 and mathematically by the following result which matches the classical MNA result extend by a trivial column and row, [128].

**Remark 4.4** (Linear index-2 variables). The index-2 variables are components that depend on the first derivatives of the input functions. Following [53] they are described by the constant projector

$$\mathbf{T} := \begin{bmatrix} \mathbf{Q}_{\text{CRV}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{Q}_{\text{C-V}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

onto  $(\mathbf{N}_0 \cap \mathbf{S}_0)(\mathbf{x}, t)$  and the complimentary projector  $\mathbf{U} := \mathbf{I} - \mathbf{T}$ . The stiffness term can be split accordingly  $\mathbf{b}(\mathbf{x}, t) = \mathbf{b}(\mathbf{U}\mathbf{x}, t) + \mathbf{B}\mathbf{T}\mathbf{x}$  using a problem-specific matrix  $\mathbf{B}$  and  $\mathbf{d}(\mathbf{x}, t) = \mathbf{d}(\mathbf{U}\mathbf{x}, t)$ . Thus it can be concluded that the index-2 variables enter our system linearly.

### 4.3 Conclusions

In this section the structural properties of magnetoquasistatic devices in electrical circuits modeled by MNA were discussed. The coupled multiscale system of lumped devices (resistors, inductors, capacitors, independent current and voltage sources) and MQS devices was formulated with a proper leading term and analyzed by the tractability index concept.

The field/circuit coupled problem was proved to be numerically harmless, i.e., it is index-2 (with linear index-2 components) at most and index-1 under rather mild conditions (analogously to the classical network case). The MQS devices were plugged into the circuit as controlled current sources, but the analysis shows that they behave topologically as inductances. This corresponds to the physical effects covered by the eddy current problem.

The index results are numerically verified in Section 6.1 by applying Euler's method as a time-integrator to a PDE inductance model. More efficient (multirate) time-integration methods are discussed in the following section.

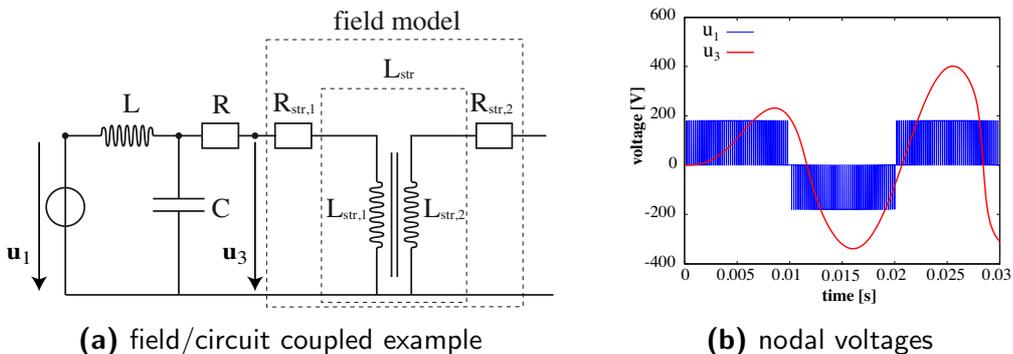
## 5 Multirate Methods

As we mentioned in the introduction, the coupled multiscale problems of Section 3 are a challenging task for time-integrators: the intrinsic time rates of subsystems differ by several orders of magnitude and the type of the subsystems is different, i.e., concerning their symmetry and definiteness. Unfortunately the structural properties can also change. The coupling of DAE index-1 subsystems can yield an arbitrarily high index problem, e.g. [51].

The introductory example Figure 5.1a shows an electric circuit configuration, where a part is refined by a magnetoquasistatic field model using partial differential equations (PDEs). Due to the different modeling techniques, i.e., network and space discretization, the underlying equations are different in shape, the circuit system is non-symmetric and typically tackled by direct solvers, while the field is symmetric and solved by iterative methods. Furthermore the voltages, Figure 5.1b, are only pulsed in the circuit part but slow and sinusoidal in the field part. The discretization has to resolve the dynamics of the coupled system as a whole and thus it produces a series of time steps that matches the dynamics of the most active component (i.e., the one working at the highest frequency). Due to switches, filters or high integration there may only be a small number of devices active at any given moment, while the others remain latent. The time-integrator will resolve those parts with an unnecessarily high time resolution causing an avoidable high computational cost.

Standard single-rate time-integration (described in Section 5.1) is inefficient for those problems. A possible work-around is presented in Section 5.2: a Schur complement approach for the MQS device is introduced, cf. [134, 58]. It allows us to use different linear solvers for the subsystems, i.e. a direct solver for the circuit and an iterative one for the field. Furthermore the Schur complement is used for several time steps (bypassing), such that multiple time scales are exploited to some extent.

A more general way is to treat every subsystem independently within a cosimulation scheme, Section 5.3. This permits the use of multirate and multimethod techniques and



**Figure 5.1:** Multirate behavior. Example is described in Section 6.3.1.

simplifies the coupling of simulator packages. The drawback is the decline in stability: the presence of algebraic constraints can handicap higher-order integration, [136] and may cause divergence, [1, 5].

In this thesis stability and convergence are proved for an iterative cosimulation scheme (‘dynamic iteration’) by analyzing the coupling interface, see Section 5.3. The analysis includes particularly the application in field/circuit and semiconductor/circuit problems as discussed in Section 5.3.5 and Section 5.3.6, respectively. In Section 5.3.5 the cosimulation approach is adapted for efficient multirate time-integration of field/circuit coupled problems by using reduced order models: ‘fitting parameters on the fly’, [118]. The conditions for stability and convergence are obtained and the convergence order is well understood.

## 5.1 Single Rate DAE Time-integration

In Section 4 we have shown, that the coupled field/circuit problem as derived in Section 3 can be given as a DAE with properly stated leading term

$$\mathbf{A} \frac{d}{dt} \mathbf{d}(\mathbf{x}, t) + \mathbf{b}(\mathbf{x}, t) = 0$$

with a matrix  $\mathbf{A} \in \mathbb{R}^{l \times m}$ , coefficient functions  $\mathbf{d}(\mathbf{x}, t) \in \mathbb{R}^m$  and  $\mathbf{b}(\mathbf{x}, t) \in \mathbb{R}^l$ . The numerical difficulties in the time-integration of those problems depend strongly on their DAE-index, see Section 4. The difficulty of numerical time-integration increases with a higher index, [25]. In the present case of problems that do not exceed index-2 standard time-integration methods for stiff problems can be applied without difficulty, e.g. implicit Runge-Kutta methods (IRK), linearly implicit schemes or Backward Differentiation Formulas (BDF), [68]. All methods discretize the time interval  $\mathcal{I}$  into a series of time points

$$t_0 < t_1 < \dots < t_n < \dots < t_e.$$

The solution of all components of  $\mathbf{x}$  is computed at each time point (‘single rate’), independently of their dynamics. For simplicity of notation a constant step size is assumed, i.e.,  $t_{n+1} - t_n = h$ . In practice an error control estimator together with a step size predictor adjusts the step size adaptively, [68].

The various time-integration methods differ in the construction of higher-order approximations  $\mathbf{x}_n$  to  $\mathbf{x}(t_n)$ . We focus here on BDF because they are traditionally used in circuit simulators. The other methods are analogously applicable and for example in Section 6 some examples are discretized by a Runge-Kutta method, i.e., RADAU5, [68]. The BDF scheme turns the time-continuous problem into a series of nonlinear time-discrete problems, where the higher order approximation is obtained by constructing a polynomial of degree  $k$  using previously computed data. The time-discrete problems reads for  $t_n$  in the notation above with properly stated leading term, [91]:

$$\mathbf{F}(\mathbf{x}_{n-k}, \dots, \mathbf{x}_n, t_{n-k}, \dots, t_n) := \mathbf{A} \frac{1}{h} \sum_{i=0}^k \alpha_i \mathbf{d}(\mathbf{x}_{n-i}, t_{n-i}) + \mathbf{b}(\mathbf{x}_n, t_n) = 0 \quad (5.1)$$

using coefficients  $\alpha_i$  ( $k$ -th order BDF). For the special choice  $k = 1$  with  $\alpha_0 = -\alpha_1 = 1$ , the method is the implicit Euler scheme. The nonlinear time-discrete problems are usually

solved by Newton-Raphson

$$\mathbf{J}_n^{(i)} \mathbf{x}_n^{(i+1)} = \underbrace{-\mathbf{F}_n^{(i)} + \mathbf{J}_n^{(i)} \mathbf{x}_n^{(i)}}_{=: \mathbf{r}} \quad \text{with} \quad \begin{aligned} \mathbf{J}_n^{(i)} &:= \frac{\alpha_0}{h} \mathbf{G}_0(\mathbf{x}_n^{(i)}, t_n) + \mathbf{b}'_x(\mathbf{x}_n^{(i)}, t_n) \\ \mathbf{F}_n^{(i)} &:= \mathbf{F}(\mathbf{x}_{n-k}, \dots, \mathbf{x}_{n-1}, \mathbf{x}_n^{(i)}, t_{n-k}, \dots, t_n) \end{aligned} \quad (5.2)$$

where  $\mathbf{G}_0 := \mathbf{A} \mathbf{d}'_{\mathbf{x}}(\mathbf{x}, t)$  and  $\mathbf{b}'_x$  are the (differential) mass and stiffness matrices, respectively (see Definition 4.2). The problems in this treatise exhibit only constant mass matrices, such that  $\mathbf{G}_0 = \text{const}$  and thus the time-derivative is discretized directly.

In Section 2.2.3 the regularization of the curl-curl operator was discussed, such that the matrix pencil  $\mathbf{J}_n^{(i)}$  above is invertible, cf. equation (2.25). This guarantees well-posed linear problems for direct solvers in the Newton-Raphson iteration (5.2). Alternatively iterative solvers could be used, which would benefit from their weak gauging property as noted in Remark 2.1, [37].

The single rate approach is not efficient for problems with different time-scales. Some components of the unknown vector may behave slowly and almost linearly in time, while other are fast and nonlinear. This limitation can be mitigated by using bypassing techniques, Section 5.2 or completely overcome by cosimulation Section 5.3.

## 5.2 Multirate Bypassing of MQS Schur Complements

In this section a bypassing algorithm based on Schur complements for the field/circuit coupled problem is developed, see [116]. It exploits the problem specific multirate behavior in the field/circuit coupled problem. It is important to note that classical multirate techniques (for example [60, 10, 125, 112] to name only a few) cannot exploit the multirate behavior as shown in Fig. 5.2, because the problem exhibits a special mixed signal ('multi-tone'): the nonlinearity of the saturation curve varies on a slow scale, while the MQS device behaves rather linearly on the fast scale. There are approaches that can cope with such signals, e.g., multirate partial differential algebraic equations (MPDAEs) split the time axis into a fast and slow scale and solve a PDAE instead of a DAE, [23], but those approaches come with a high computational overhead and cannot simply be implemented into existing simulation environments.

The first part of the section introduces the basics of matrix assembly and Schur complements in circuit simulation and Section 5.2.1 explains which time scales are relevant.

Circuit simulators assemble the time-discretized MNA equations (3.19) element-wise.<sup>1</sup> The time-integrator is typically a BDF method, see above. Each circuit element is represented by an current/voltage relation possibly with some additional equations and internal unknowns. Those relations are algebraic functions  $\mathbf{F}_{e,n}$  that finally form the overall system  $\mathbf{F}_n = \mathbf{0}$  at the  $n$ -th time step. Similarly each element contributes with  $\mathbf{J}_{e,n}$  to the overall Jacobian. Suppressing the subscript  $n$  (for simplicity of notation), we read

$$\mathbf{J}^{(i)} := \sum_e \mathbf{Q}_e \mathbf{J}_e^{(i)} \mathbf{Q}_e^\top \quad \text{and} \quad \mathbf{F}^{(i)} := \sum_e \mathbf{Q}_e \mathbf{F}_e^{(i)} \quad (5.3)$$

with element-wise assembly matrices  $\mathbf{Q}_e$  that organize the contributions such that the

<sup>1</sup>Element refers here to circuit elements and they are not related to elements used for space discretization.

MNA equations (3.19) are obtained, [116]. In particular they incorporate the incidence matrices  $\mathbf{A}_e$ . One calls the tuple  $\mathbf{J}_e^{(i)}$ ,  $\mathbf{F}_e^{(i)}$  the element stamp. It consists of internal and external variables, i.e., variables used only inside the particular element and variables that are related to other elements by the simulator, [56, 58].

Following [116, 134] we want to speed up the solution of the Newton system in field/circuit applications by eliminating the magnetic vector potential  $\widehat{\mathbf{a}}$ . We focus here on the MQS device's contribution to the time-discretization of system (3.19). This is in the notation above a particular contribution to the function  $\mathbf{F}$  in (5.3). It reads

$$\mathbf{F}_M^{(i)} := \begin{bmatrix} \mathbf{I} & 0 & 0 \\ \mathbf{R}_M & -\mathbf{I} & 0 \\ -\mathbf{X}_M & 0 & \mathbf{k}_\nu(\widehat{\mathbf{a}}^{(i)}) \end{bmatrix} \mathbf{x}_M^{(i)} + \frac{1}{h} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \mathbf{X}_M^\top \\ 0 & 0 & \mathbf{M}_{\bar{\sigma}} \end{bmatrix} \rho \mathbf{x}_M^{(i)},$$

with the variables

$$\mathbf{x}_M := \mathbf{x}_{M,n} := \begin{bmatrix} \mathbf{i}_{M,n} \\ \mathbf{v}_{M,n} \\ \widehat{\mathbf{a}}_n \end{bmatrix} \quad \text{and} \quad \rho \mathbf{x}_M^{(i)} := \alpha_0 \mathbf{x}_{M,n}^{(i)} + \sum_{i=1}^k \alpha_i \mathbf{x}_{M,n-i}.$$

The first row contains the contribution to the current balance equation, the second line is the coupling equation and finally the last row represents the MQS curl-curl equation. Here the model is excited by stranded conductors (the solid conductor case is analogous). We obtain the following Jacobian contribution:

$$\mathbf{J}_M^{(i)} := \begin{bmatrix} \mathbf{I} & 0 & 0 \\ \mathbf{R}_M & -\mathbf{I} & \frac{\alpha_0}{h} \mathbf{X}_M^\top \\ -\mathbf{X}_M & 0 & \mathbf{K}_h^{(i)} \end{bmatrix} \quad \text{with} \quad \mathbf{K}_h^{(i)} := \mathbf{K}_\nu(\widehat{\mathbf{a}}^{(i)}) + \frac{\alpha_0}{h} \mathbf{M}_{\bar{\sigma}} \quad (5.4)$$

with the differential reluctivity matrix  $\mathbf{K}_\nu(\widehat{\mathbf{a}}^{(i)})$ , see (2.23). The contribution to the right-hand side is given by, cf. (5.2):

$$\mathbf{r}_M^{(i)} := -\mathbf{F}_M^{(i)} + \mathbf{J}_M^{(i)} \mathbf{x}_M^{(i)} = \frac{1}{h} \begin{bmatrix} 0 \\ \mathbf{X}_M^\top \\ \mathbf{M}_{\bar{\sigma}} \end{bmatrix} (\alpha_0 \widehat{\mathbf{a}}^{(i)} - \rho \widehat{\mathbf{a}}^{(i)}) + \begin{bmatrix} 0 \\ 0 \\ \mathbf{K}_\nu(\widehat{\mathbf{a}}^{(i)}) - \mathbf{k}_\nu(\widehat{\mathbf{a}}^{(i)}) \end{bmatrix} \widehat{\mathbf{a}}^{(i)}, \quad (5.5)$$

where the only unknown  $\widehat{\mathbf{a}}$  is the magnetic vector potential, which is internal, i.e., it is not used outside the MQS stamp. Only the current/voltage relation of the series connection of a (nonlinear) inductor and a resistor needs to be revealed to the host circuit simulator. This is the case if  $\widehat{\mathbf{a}}$  is eliminated from the Newton system by the Schur complement.

This is beneficial for all kinds of large elements, e.g. it allows cache-optimized stamping of semiconductor models, [58], or field devices, [134, 116]. In either case more compact stamps are obtained, which fit better into the overall MNA framework.

The unknown  $\widehat{\mathbf{a}}$  is removed and one ends up with a reduced stamp in terms of  $\tilde{\mathbf{x}}_M^\top = (\tilde{\mathbf{i}}_M^\top, \tilde{\mathbf{v}}_M^\top)$ . The corresponding reduced Jacobian reads

$$\tilde{\mathbf{J}}_M^{(i)} := \begin{bmatrix} \mathbf{I} & 0 & 0 \\ 0 & \mathbf{I} & -\frac{\alpha_0}{h} \mathbf{X}_M^\top \left( \mathbf{K}_h^{(i)} \right)^{-1} \end{bmatrix} \mathbf{J}_M^{(i)} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 \\ \mathbf{R}_M + \frac{\alpha_0}{h} \mathbf{L}_h^{(i)} & -\mathbf{I} \end{bmatrix} \quad (5.6)$$

with generalized inductance matrix

$$\mathbf{L}_h^{(i)} := \mathbf{X}_M^\top \left( \mathbf{K}_h^{(i)} \right)^{-1} \mathbf{X}_M \quad (5.7)$$

using  $\mathbf{K}_h^{(i)}$  from (5.4). For static models, equation (5.7) extracts an inductance by one Ampère excitation, [118], but here  $\mathbf{L}_h^{(i)}$  also takes eddy current effects into account (due to the presence of the conductance matrix  $\mathbf{M}_{\bar{\sigma}}$ ). Thus the inductance depends on the frequency (via the step size  $h$ ) and therefore the matrix must be recomputed or interpolated for any change of  $h$ .

The reduced right-hand side contribution is

$$\tilde{\mathbf{r}}_M^{(i)} = \begin{bmatrix} 0 \\ \tilde{\mathbf{r}}_{M,v}^{(i)} \end{bmatrix} \quad \text{where} \quad \tilde{\mathbf{r}}_{M,v}^{(i)} = \frac{1}{h} \mathbf{X}_M^\top \left( \mathbf{I} - \frac{\alpha_0}{h} \left( \mathbf{K}_h^{(i)} \right)^{-1} \mathbf{M}_{\bar{\sigma}} \right) (\alpha_0 \bar{\mathbf{a}}^{(i)} - \rho \bar{\mathbf{a}}^{(i)}) \\ - \frac{\alpha_0}{h} \mathbf{X}_M^\top \left( \mathbf{K}_h^{(i)} \right)^{-1} (\mathbf{K}_\nu(\bar{\mathbf{a}}^{(i)}) - \mathbf{k}_\nu(\bar{\mathbf{a}}^{(i)})) \bar{\mathbf{a}}^{(i)}.$$

The magnetic vector potential is still needed to evaluate the reduced right-hand sides (and thus the nonlinear material curve). This evaluation and the recomputation of the Schur complements come with computational costs.

In the following the computational costs are discussed in terms of direct solvers (matrix factorizations). Their application is beneficial for our approach, because the factorization can be stored. For the Schur complement in the Newton iteration  $i+1$ , we need to compute  $\mathbf{L}_h^{(i)}$ . When applying a direct solver, the matrix  $\mathbf{K}_h^{(i)}$  is factorized (one LU decomposition) and afterwards forward/backward substitutions are carried out for each circuit branch ( $m = 1, \dots, n_M$ ):

$$\mathbf{K}_h^{(i)} \bar{\mathbf{a}}_m^{(i)} = \mathbf{X}_m \quad \text{s.t.} \quad \mathbf{L}_h^{(i)} = \mathbf{X}_M^\top \bar{\mathbf{a}}_M^{(i)} \quad \text{with} \quad \bar{\mathbf{a}}_M^{(i)} := [\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_{n_M}] \quad (5.8)$$

by sparse inner products. Also, the magnetic vector potential for the right-hand-side voltage must be computed. To this end, we solve the following equation inside each Newton iteration (derived from Jacobian (5.4) and right-hand side (5.5)):

$$\mathbf{K}_h^{(i)} \bar{\mathbf{a}}^{(i+1)} = \mathbf{r}_{M,a}^{(i)} + \mathbf{X}_M \mathbf{i}_M^{(i+1)}, \quad \mathbf{r}_{M,a}^{(i)} := \frac{1}{h} \mathbf{M}_{\bar{\sigma}} (\alpha_0 \bar{\mathbf{a}}^{(i)} - \rho \bar{\mathbf{a}}^{(i)}) + (\mathbf{K}_\nu(\bar{\mathbf{a}}^{(i)}) - \mathbf{k}_\nu(\bar{\mathbf{a}}^{(i)})) \bar{\mathbf{a}}^{(i)}.$$

Thus we determine the remaining term  $\bar{\mathbf{a}}_V^{(i)}$  by forward/backward substitutions from:

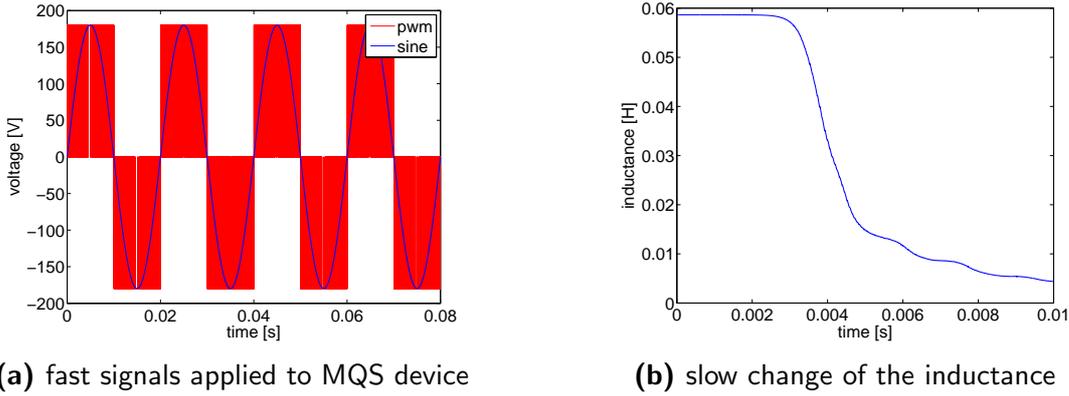
$$\mathbf{K}_h^{(i)} \bar{\mathbf{a}}_V^{(i)} = \mathbf{r}_{M,a}^{(i)}, \quad (5.9)$$

and finally one obtains for the magnetic vector potential

$$\bar{\mathbf{a}}^{(i+1)} = \bar{\mathbf{a}}_V^{(i)} + \bar{\mathbf{a}}_M^{(i)} \mathbf{i}_M^{(i+1)}.$$

Moreover, we find for the reduced right-hand side the simplification:

$$\tilde{\mathbf{r}}_{M,v}^{(i)} = \frac{1}{h} \mathbf{X}_M^\top \left( \alpha_0 \bar{\mathbf{a}}^{(i)} - \rho \bar{\mathbf{a}}^{(i)} - \bar{\mathbf{a}}_V^{(i)} \right).$$



**Figure 5.2:** Time scales in a field/circuit coupled problem: the inductivity changes according to the slow sine wave, which is the low frequency part of the pulsed signal.

This improves the result of [134] for a similar setting. It has been shown that only one factorization and  $n_M + 1$  forward/backward substitutions, i.e., equations (5.8) and (5.9), are necessary for the Schur complement in each Newton iteration. The choice of the linear solver for the Schur complement is independent of the solver used in the circuit host simulator (typically a direct solver). So, for example an iterative method such as (block) PCG could be used for the Schur complement, [110]. This allows us to solve 3D problems within the circuit simulator without additional gauging, because of the *weak gauging* property, Remark 2.1. This highly improves the efficiency of the linear solver, especially if multiple right-hand-sides are supported to solve (5.8-5.9), e.g., [142].

### 5.2.1 Bypassing as Multirate Time-Integration

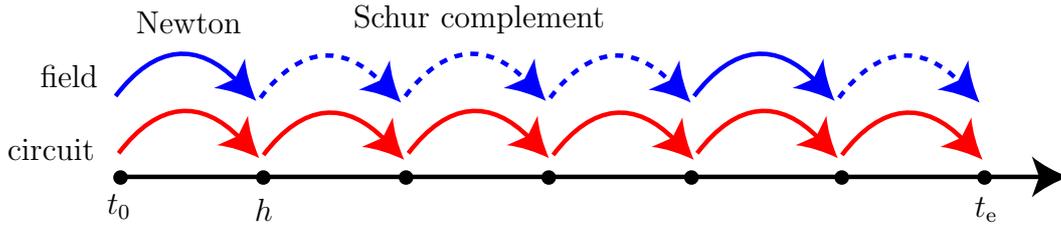
The (generalized) inductance matrix  $\mathbf{L}_h$  is determined by the magnetic saturation due to the underlying BH-curve (2.15). The saturation is driven by the supplied energy and thus it is rather slow compared to other time rates of the electric circuit, e.g., the switching frequency of transistors, see Fig. 5.2. Thus the relevant time rate of the nonlinearity corresponds to the dynamics of the following integral, even if the applied voltage is a much faster switching signal

$$\mathbf{E}(t_n) = \mathbf{E}(t_0) + \int_0^{t_n} \mathbf{i}_M(s) \mathbf{v}_M(s) ds \quad (5.10)$$

with the initial energy level of the device  $\mathbf{E}_0$ . In the  $i$ -th Newton iteration of time step  $n$  the energy is approximated by

$$\mathbf{E}_n \approx \mathbf{E}_0 + h \left( \sum_{j=0}^{n-1} \mathbf{i}_{M,j} \mathbf{v}_{M,j} + \mathbf{i}_{M,j}^{(i)} \mathbf{v}_{M,j}^{(i)} \right)$$

and this can easily be compared to the initial energy level  $\mathbf{E}_0$  using a (relative) norm. Updates of the nonlinearity are only necessary if the energy level changes significantly. Consequently, the computation of the material matrices can often be bypassed, Fig. 5.3. Then the model behaves (nearly) linearly and only one forward/backward substitution



**Figure 5.3:** Bypassing. The field and circuit problems are decoupled: the solid arrows denote time steps where the system is solved by an ordinary Newton-Raphson iteration. Time steps with the dashed arrows indicate the usage of the Schur complement.

for the right-hand-side per iteration is necessary (5.9). This defines a simplified Newton algorithm, where the Jacobian (5.6) is frozen for several iterations and possibly several time steps if the (relative) change of energy does not exceed a threshold and thus the reluctivity is (nearly) constant.

Furthermore, if the problem is rather latent the right-hand-side evaluation can be bypassed as well. The vector potential needs no update and thus the field problem is decoupled from the circuit, where it is represented by an inductance matrix. This is obviously not free of risks, because the bypassing of the right-hand-side changes the fixed point of the Newton scheme. The algorithm is given in Listing 1 where the energy level and the reluctivity  $\nu$  are monitored by relative norms. This algorithm relieves the host simulator of the burden of solving unnecessarily large system of equations, especially if the nonlinearity is weak. On the other hand the host-simulator still can use a Jacobian for its Newton iteration. The drawback are the additional iterations due to the inferior convergence of simplified Newton, albeit solving a sequence of reduced systems.

## 5.2.2 Conclusions

The Schur complement approach yields significantly smaller element stamps that are equivalent to a series connection of an inductance and a resistance (this is in line with DAE-index result, Section 4). The additional costs of the complement computation can be disregarded if linear solvers with multiple right-hand side techniques are available. In particular iterative solvers will improve the efficiency of this solution process. Due to bypassing, the field/circuit coupling is weakened and thus the time-integration of the circuit is cheapened because only basic elements are evaluated. This decoupling exploits the multirate time behavior of the coupled system if present, see Section 6.2.

Although the bypassing allows for a decoupling, the circuit simulator still controls the simulation: the same time-integrator is used for all subproblems, the circuit simulator decides about step sizes and Newton iterations etc., see Fig. 5.3. If the level of decoupling is to be further increased the use of waveform relaxation approaches is appropriate. This is the topic of the next section.

## 5.3 Multirate Cosimulation

Coupled multiscale problems like (3.19) arise naturally in electromagnetic device simulation and in the previous section on Schur complement bypassing, we have introduced a method

---

**Listing 1** Algorithm: Bypassing

---

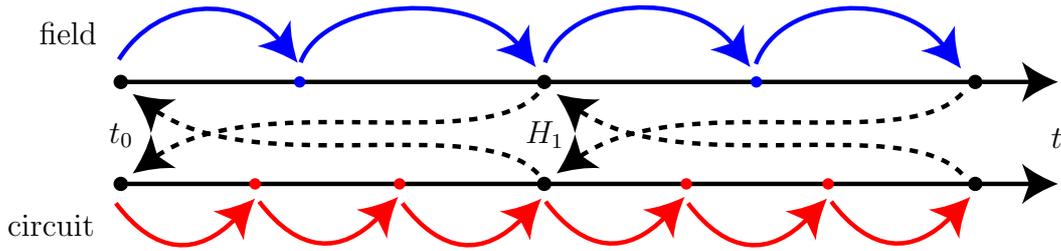
- 1) compute  $\widehat{\mathbf{a}}^{(i)}$  and  $\mathbf{E}_n^{(i)}$
  - 2) **if**  $\text{norm}(\mathbf{E}_n^{(i)} - \mathbf{E}_0) > \text{tol}$ 
    - then** evaluate material curve  $\nu^{(i)} := \nu(\widehat{\mathbf{a}}^{(i)})$
    - 2a) **if**  $\text{norm}(\nu^{(i)} - \nu^{(i-1)}) > \text{tol}$ 
      - then** compute  $\mathbf{L}_h^{(i)}$  and  $\mathbf{v}_M^{(i)}$
      - else** bypass matrix update  $\mathbf{L}_h^{(i)} := \mathbf{L}_h^{(i-1)}$  and  $\mathbf{v}_M^{(i)} := \mathbf{v}_M^{(i-1)}$
    - else** bypass material update  $\nu^{(i)} := \nu^{(i-1)}$  and  $\mathbf{L}_h^{(i)} := \mathbf{L}_h^{(i-1)}$ ,  $\mathbf{v}_M^{(i)} := \mathbf{v}_M^{(i-1)}$
  - 3) return to host simulator.
- 

that is able to mitigate some of the most severe problems (multirate and the efficiency of linear solvers). On the other hand, problems often come along with their own specific simulation packages and thus the coupled system cannot be solved monolithically (as a whole system) in the time domain. This renders standard time-integration impossible and even the bypassing approach is only feasible if the simulation packages supply the corresponding interfaces, e.g., such that we can access the magnetic vector potential to ensure correct initial values.

One way to overcome this impasse is to apply cosimulation methods, i.e. methods for the coupling of simulator packages. Each model is simulated separately by its package, e.g. network and PDE models. Thus all problems may be solved on their own time scale, with tailor-made methods (multimethod). Information on how the models interact are exchanged only at synchronization points (on ‘time windows’).

In the following we will revisit the *waveform relaxation* or *dynamic iteration* schemes, which solve the subproblems iteratively and exchange coupling information in each sweep, [88, 141]. Fig. 5.4 depicts the general idea of separate time-stepping, windowing and iteration. Those schemes are known to be unconditionally stable for coupled ODEs, [97, 27] and their convergence is well understood for ODEs that stem from circuit simulation and space discretized PDEs, [135, 77]. On the other hand instabilities are known when those schemes are applied to DAEs. The introduction of a windowing technique has been proposed to speed-up convergence and avoid instabilities. Nonetheless the contraction of the underlying fixed point iteration is only guaranteed if a stability constraint is fulfilled. This constraint forces the algebraic coupling to be weak, but then convergence on a single window is enforced, [76, 7].

The error transport for multiple windows has been analyzed so far for a special class of DAEs, where the coupling is established by Lagrangian multipliers, [5]. In [51] the error transport of the general case was considered using a differential equation for the error propagation, but the DAE system was reduced to its underlying ODE. This has simplified several aspects, that are discussed in [9], which is the basis of the following analysis. It generalizes the approach of [5] to the most general form of index-1 DAEs, for which a similar stability constraint is derived. Global convergence and stability (with error propagation) will be guaranteed if the splitting error remains in a neighborhood of the analytical solution. Sections 5.3.5 and 5.3.6 will apply the theory to the electromagnetic



**Figure 5.4:** Dynamic iteration. The subsystems (e.g. field and circuit) are discretized on their own time scales using different time steps (solid arrows). Several time steps are a time window and they are computed iteratively (the ‘sweeps’ are denoted by the dashed arrows).

models in electric networks of the previous sections and show that dynamic iteration is indeed a strategy for exploiting multirate behavior using reduced order models similar to those illustrated in [106].

### 5.3.1 Abstract DAE-DAE Coupling

In the index analysis Section 4, it has been shown that the coupled problem is addressed by the abstract equation (4.1). The projectors in Section 4 can be used to decompose the problem down into dynamic and algebraic parts, [92]. The resulting parts can be reorganized such that we obtain a semi-explicit initial value problem of differential algebraic equations (DAE-IVP)

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, \mathbf{z}), \quad \text{with } \mathbf{y}(t_0) = \mathbf{y}_0, \quad (5.11a)$$

$$0 = \mathbf{g}(\mathbf{y}, \mathbf{z}), \quad \text{with } \mathbf{z}(t_0) = \mathbf{z}_0, \quad (5.11b)$$

where the variables are renamed to  $\mathbf{x}^\top = [\mathbf{y}^\top, \mathbf{z}^\top]$  with the vector functions  $\mathbf{f}$  and  $\mathbf{g}$ . This formulation addresses the whole problem abstractly and thus allows us to extract the underlying principles without the dispensable details. The following analysis of system (5.11) is limited to the index-1 case. In the field/circuit case this can be assured by Theorem 4.2 that demands loop/cutset conditions for the circuit. This prerequisite is mathematically formalized by the following index-1 assumption:

**Assumption 5.1** (Monolithic index-1). *The differential algebraic initial value problem (5.11)*

(a) *has a unique solution on  $\mathcal{I} = [t_0, t_e]$*

$$\mathbf{y} : [t_0, t_e] \rightarrow \mathbb{R}^{n_y} \quad \text{and} \quad \mathbf{z} : [t_0, t_e] \rightarrow \mathbb{R}^{n_z} \quad (5.12)$$

*and the initial values are particularly consistent, i.e.,  $\mathbf{y}_0$  and  $\mathbf{z}_0$  solve the algebraic equation (5.11b),*

(b) *the right-hand-side functions  $\mathbf{f}$  and  $\mathbf{g}$  are supposed to be sufficiently often differentiable in the neighborhood of the solution,*

(c) *the Jacobian  $\partial \mathbf{g} / \partial \mathbf{z}$  is non-singular in the neighborhood of the solution (the coupled problem is index-1).*

**Remark 5.1** (Index concepts for Hessenberg systems). Please note that we do not distinguish between different DAE index-concepts for semi-explicit systems. Equation (5.11) describes the special case of an index-1 Hessenberg system and for Hessenberg systems of index-1, index-2 and index-3 it has been shown in [67, Page 13] that the important index-concepts coincide, [29].

In the index-1 case above the vector  $\mathbf{y}$  contains only the *differential variables*, i.e., the variables that are defined by first derivatives with respect to time, while  $\mathbf{z}$  contains the *algebraic variables* that are not described by any derivatives.

Let the coupled problem (5.11) consist of  $r$  subsystems. For example a network containing an MQS and a semiconductor device consists of  $r = 3$  subsystems. Let us further assume that the equations are partitioned accordingly, then the  $i$ -th subsystem is given by (for  $i = 1, \dots, r$ )

$$\dot{\mathbf{y}}_i = \mathbf{f}_i(\mathbf{y}, \mathbf{z}), \quad \text{with } \mathbf{y}^\top = [\mathbf{y}_1^\top, \dots, \mathbf{y}_r^\top] \quad (5.13a)$$

$$0 = \mathbf{g}_i(\mathbf{y}, \mathbf{z}), \quad \text{with } \mathbf{z}^\top = [\mathbf{z}_1^\top, \dots, \mathbf{z}_r^\top] \quad (5.13b)$$

where the global right-hand-sides are assembled in the obvious way  $\mathbf{f}^\top = [\mathbf{f}_1^\top, \dots, \mathbf{f}_r^\top]$  and  $\mathbf{g}^\top = [\mathbf{g}_1^\top, \dots, \mathbf{g}_r^\top]$ . Similar to the monolithic Ass. 5.1 we require for each subsystem

**Assumption 5.2** (Subsystem index-1). *The Jacobian*

$$\partial \mathbf{g}_i / \partial \mathbf{z}_i \text{ is not singular for all } i = 1, \dots, r, \quad (5.14)$$

*in the neighborhood of the solution (the subsystems are index-1).*

This guarantees by the Implicit Function Theorem that each subsystem  $\mathbf{g}_i(\mathbf{y}, \mathbf{z}) = 0$  is (locally) uniquely solvable for  $\mathbf{z}_i$ . It is important to note that Assumptions 5.1 and 5.2 do not imply each other, as shown by the following examples.

**Example 5.1** (DAE-index of subsystems and monolithic DAE-index). The following examples show that coupling two index-1 subsystems can result in a (monolithic) index-2 system and vice-versa.

- Let us consider the following scalar subsystems separately, [50]

$$\begin{aligned} \dot{\mathbf{y}}_1 = \mathbf{f}_1(\mathbf{y}, \mathbf{z}) &:= \mathbf{z}_1 & \text{and} & & 0 = \mathbf{g}_1(\mathbf{y}, \mathbf{z}) &:= \mathbf{y}_1 + \mathbf{z}_1 + \mathbf{w}_1 \\ \dot{\mathbf{y}}_2 = \mathbf{f}_2(\mathbf{y}, \mathbf{z}) &:= \mathbf{z}_2 & \text{and} & & 0 = \mathbf{g}_2(\mathbf{y}, \mathbf{z}) &:= \mathbf{y}_2 + \mathbf{z}_2 + \mathbf{w}_2, \end{aligned}$$

where  $\mathbf{w}_1(t)$  and  $\mathbf{w}_2(t)$  are given inputs and each subsystem is index-1

$$\partial \mathbf{g}_1 / \partial \mathbf{z}_1 = \partial \mathbf{g}_2 / \partial \mathbf{z}_2 = \mathbf{I}, \quad \text{but} \quad \mathbf{w}_1 = \mathbf{z}_2, \mathbf{w}_2 = \mathbf{z}_1 \quad \text{yields} \quad \partial \mathbf{g} / \partial \mathbf{z} = \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} \end{bmatrix}$$

such that the monolithically coupled system is index-2.

- Let us consider the following scalar subsystems (Hessenberg form) separately

$$\begin{aligned} \dot{\mathbf{y}}_1 = \mathbf{f}_1(\mathbf{y}, \mathbf{z}) &:= \mathbf{y}_1 + \mathbf{z}_1 & \text{and} & & 0 = \mathbf{g}_1(\mathbf{y}, \mathbf{z}) &:= \mathbf{y}_1 + \mathbf{w}_1 \\ \dot{\mathbf{y}}_2 = \mathbf{f}_2(\mathbf{y}, \mathbf{z}) &:= \mathbf{y}_2 + \mathbf{z}_2 & \text{and} & & 0 = \mathbf{g}_2(\mathbf{y}, \mathbf{z}) &:= \mathbf{y}_2 + \mathbf{w}_2, \end{aligned}$$

where  $\mathbf{w}_1(t)$  and  $\mathbf{w}_2(t)$  are given inputs and each subsystem is index-2

$$\partial \mathbf{g}_1 / \partial \mathbf{z}_1 = \partial \mathbf{g}_2 / \partial \mathbf{z}_2 = 0, \quad \text{but} \quad \mathbf{w}_1 = \mathbf{z}_2, \mathbf{w}_2 = \mathbf{z}_1 \quad \text{yields} \quad \partial \mathbf{g} / \partial \mathbf{z} = \begin{bmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{bmatrix}$$

such that the monolithically coupled system is index-1.

Now, having ensured that the problem is well-posed, the next section will introduce the iteration schemes using the concept of splitting functions.

### 5.3.2 Dynamic Iteration

This section defines the fundamental ingredients for a dynamic iteration scheme: the partitioning into time windows, the extrapolation and iteration steps. The iteration step is applied repeatedly, but it is stopped after a finite number of iterations. Thus the dynamic iteration returns an approximation to problem (5.11) that is affected by a splitting error. The approximation is denoted by, cf. [9]

$$\tilde{\mathbf{x}} := \begin{bmatrix} \tilde{\mathbf{y}} \\ \tilde{\mathbf{z}} \end{bmatrix} \quad \text{with} \quad \tilde{\mathbf{y}} : [t_0, t_e] \rightarrow \mathbb{R}^{n_y} \quad \text{and} \quad \tilde{\mathbf{z}} : [t_0, t_e] \rightarrow \mathbb{R}^{n_z} \quad (5.15)$$

being close to the exact solution (5.12). As indicated before, the dynamic iteration scheme operates on time windows  $[T_n, T_{n+1}]$ , such that

$$t_0 = T_0 < T_1 < T_2 < \dots < T_N = t_e$$

with window size  $H_n := T_{n+1} - T_n$ . In the multirate context those windows are called macro steps in contrast to the micro steps  $h$  of the numerical time-integration, see Fig. 5.4. Assuming a numerical approximation is computed on the window  $[T_{n-1}, T_n]$ , the dynamic iteration defines a new approximation on the consecutive window

$$\tilde{\mathbf{x}}|_{[T_n, T_{n+1}]} \in C_n^{1,0} \quad \text{with} \quad C_n^{1,0} := C^1([T_n, T_{n+1}], \mathbb{R}^{n_y}) \times C([T_n, T_{n+1}], \mathbb{R}^{n_z})$$

using an extrapolation step followed by (one or more) iteration steps. The corresponding steps are defined by operators as in [5].

**Extrapolation step.** Let the operator  $\Phi_n : C_{n-1}^{1,0} \rightarrow C_n^{1,0}$  denote a continuous extrapolation from the old window  $[T_{n-1}, T_n]$  to the new window  $[T_n, T_{n+1}]$ . This defines an initial guess of the new approximation

$$\begin{bmatrix} \tilde{\mathbf{y}}_n^{(0)} \\ \tilde{\mathbf{z}}_n^{(0)} \end{bmatrix} := \Phi_n \begin{bmatrix} \tilde{\mathbf{y}}|_{[T_{n-1}, T_n]} \\ \tilde{\mathbf{z}}|_{[T_{n-1}, T_n]} \end{bmatrix} \quad \text{with} \quad \Phi_n = \begin{bmatrix} \Phi_{\mathbf{y},n} \\ \Phi_{\mathbf{z},n} \end{bmatrix}. \quad (5.16)$$

Actually the initial value  $\tilde{\mathbf{x}}_n^{(0)}(T_n)$  is fixed from the previous window. Its constant extrapolation on the new window is the most common choice for an initial guess:

$$\begin{bmatrix} \tilde{\mathbf{y}}_n^{(0)}(t) \\ \tilde{\mathbf{z}}_n^{(0)}(t) \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{y}}_n(T_n) \\ \tilde{\mathbf{z}}_n(T_n) \end{bmatrix} \quad \text{for all} \quad t \in [T_n, T_{n+1}].$$

This operator introduces an error in  $\mathcal{O}(H_n)$ , which can be improved by linear or higher order polynomial extrapolation. It is uniformly Lipschitz-continuous, independently of  $H_n$ , see [5].

**Iteration step** The extrapolation has defined an initial guess. It is followed by an iteration step defined by the mapping  $\Psi_n : C_n^{1,0} \rightarrow C_n^{1,0}$

$$\begin{bmatrix} \tilde{\mathbf{y}}_n^{(k-1)} \\ \tilde{\mathbf{z}}_n^{(k-1)} \end{bmatrix} \rightarrow \begin{bmatrix} \tilde{\mathbf{y}}_n^{(k)} \\ \tilde{\mathbf{z}}_n^{(k)} \end{bmatrix} := \Psi_n \begin{bmatrix} \tilde{\mathbf{y}}_n^{(k-1)} \\ \tilde{\mathbf{z}}_n^{(k-1)} \end{bmatrix} \quad \text{with } \Psi_n = \begin{bmatrix} \Psi_{\mathbf{y},n} \\ \Psi_{\mathbf{z},n} \end{bmatrix} \quad (5.17)$$

where  $k = 1, \dots, k_n$  and  $k_n$  is the (finite) number of iterations. The mapping  $\Psi_n$  defines the solution operator for the initial value problem

$$\dot{\tilde{\mathbf{y}}}_n^{(k)} = \mathbf{F}(\tilde{\mathbf{y}}_n^{(k)}, \tilde{\mathbf{y}}_n^{(k-1)}, \tilde{\mathbf{z}}_n^{(k)}, \tilde{\mathbf{z}}_n^{(k-1)}), \quad \text{with } \tilde{\mathbf{y}}_n^{(k)}(T_n) = \tilde{\mathbf{y}}_n^{(k-1)}(T_n), \quad (5.18a)$$

$$0 = \mathbf{G}(\tilde{\mathbf{y}}_n^{(k)}, \tilde{\mathbf{y}}_n^{(k-1)}, \tilde{\mathbf{z}}_n^{(k)}, \tilde{\mathbf{z}}_n^{(k-1)}). \quad (5.18b)$$

where the differential and algebraic splitting functions  $\mathbf{F}$  and  $\mathbf{G}$ , respectively, are consistent, i.e., they fulfill the following compatibility condition:

**Definition 5.1** (Consistent splitting functions). Splitting functions  $\mathbf{F}$  and  $\mathbf{G}$  are called consistent if they are sufficiently differentiable and fulfill the compatibility condition

$$\mathbf{F}(\mathbf{y}, \mathbf{y}, \mathbf{z}, \mathbf{z}) = \mathbf{f}(\mathbf{y}, \mathbf{z}) \quad \text{and} \quad \mathbf{G}(\mathbf{y}, \mathbf{y}, \mathbf{z}, \mathbf{z}) = \mathbf{g}(\mathbf{y}, \mathbf{z}). \quad (5.19)$$

**Remark 5.2** (Fixed-point). The following statements follow immediately from the definitions above:

- the compatibility condition implies the exact solution  $\mathbf{x}$  is a fixed-point of the iteration operator  $\Psi_n$ .
- the solution on  $[T_n, T_{n+1}]$  after extrapolation and  $k_n$  iterations reads

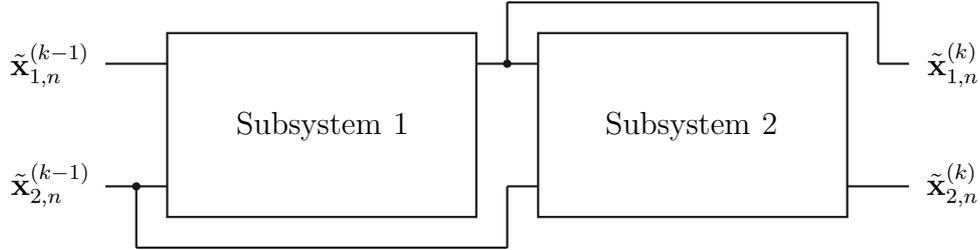
$$\begin{bmatrix} \tilde{\mathbf{Y}}|_{[T_n, T_{n+1}]} \\ \tilde{\mathbf{Z}}|_{[T_n, T_{n+1}]} \end{bmatrix} := (\Psi_n^{k_n} \circ \Phi_n) \begin{bmatrix} \tilde{\mathbf{Y}}|_{[T_{n-1}, T_n]} \\ \tilde{\mathbf{Z}}|_{[T_{n-1}, T_n]} \end{bmatrix}. \quad (5.20)$$

- for the partitioned system (5.13) with corresponding unknowns  $\tilde{\mathbf{y}}_n^\top = [\tilde{\mathbf{y}}_{1,n}^\top \dots \tilde{\mathbf{y}}_{r,n}^\top]$ ,  $\tilde{\mathbf{z}}_n^\top = [\tilde{\mathbf{z}}_{1,n}^\top \dots \tilde{\mathbf{z}}_{r,n}^\top]$ , the iteration operator  $\Psi_n$  is defined by  $r$  initial-value problems:

$$\begin{aligned} \dot{\tilde{\mathbf{y}}}_{i,n}^{(k)} &= \mathbf{F}_i(\tilde{\mathbf{y}}_n^{(k)}, \tilde{\mathbf{y}}_n^{(k-1)}, \tilde{\mathbf{z}}_n^{(k)}, \tilde{\mathbf{z}}_n^{(k-1)}), \quad \text{with } \tilde{\mathbf{y}}_{i,n}^{(k)}(T_n) = \tilde{\mathbf{y}}_{i,n}^{(k-1)}(T_n), \\ 0 &= \mathbf{G}_i(\tilde{\mathbf{y}}_n^{(k)}, \tilde{\mathbf{y}}_n^{(k-1)}, \tilde{\mathbf{z}}_n^{(k)}, \tilde{\mathbf{z}}_n^{(k-1)}) \end{aligned} \quad (5.21)$$

where  $i = 1, \dots, r$  and splitting functions  $\mathbf{F}^\top = [\mathbf{F}_1^\top \dots \mathbf{F}_r^\top]$  and  $\mathbf{G}^\top = [\mathbf{G}_1^\top \dots \mathbf{G}_r^\top]$ .

All common iteration schemes can be encoded by the splitting functions above, i.e., Picard, Jacobi or Gauß-Seidel-type schemes, [9]. For example the Gauß-Seidel-type scheme is visualized in Fig. 5.5 for the case of two subsystems. It computes sequentially the solutions of all  $r$  subsystems. It iterates such that the  $i$ -th subsystem in the  $k$ -th iteration solves for the new solution  $\mathbf{x}_{i,n}^{(k)}$  where it utilizes the latest available data for the variables



**Figure 5.5:** Schematic representation of the  $k$ -th Gauß-Seidel iteration on the  $n$ -th time window for  $r = 2$  subsystems; schematic as given in [51].

of the previous subsystems  $(1, \dots, i - 1)$  and old data (from the previous iteration) for the variables of following subsystems  $(i + 1, \dots, r)$ .

**Definition 5.2** (Gauß-Seidel). The splitting functions ( $i = 1, \dots, r$ )

$$\mathbf{F}_i(\tilde{\mathbf{y}}_n^{(k)}, \tilde{\mathbf{y}}_n^{(k-1)}, \tilde{\mathbf{z}}_n^{(k)}, \tilde{\mathbf{z}}_n^{(k-1)}) := \mathbf{f}_i(\tilde{\mathbf{Y}}_{i,n}^{(k)}, \tilde{\mathbf{Z}}_{i,n}^{(k)}), \quad (5.22a)$$

$$\mathbf{G}_i(\tilde{\mathbf{y}}_n^{(k)}, \tilde{\mathbf{y}}_n^{(k-1)}, \tilde{\mathbf{z}}_n^{(k)}, \tilde{\mathbf{z}}_n^{(k-1)}) := \mathbf{g}_i(\tilde{\mathbf{Y}}_{i,n}^{(k)}, \tilde{\mathbf{Z}}_{i,n}^{(k)}) \quad (5.22b)$$

define a Gauß-Seidel-type iteration scheme with

$$\tilde{\mathbf{X}}_{i,n}^{(k)} := \begin{bmatrix} \tilde{\mathbf{Y}}_{i,n}^{(k)} \\ \tilde{\mathbf{Z}}_{i,n}^{(k)} \end{bmatrix} \quad \text{with} \quad \tilde{\mathbf{Y}}_{i,n}^{(k)}(t) := \begin{bmatrix} \tilde{\mathbf{y}}_{1,n}^{(k)}(t) \\ \vdots \\ \tilde{\mathbf{y}}_{i,n}^{(k)}(t) \\ \tilde{\mathbf{y}}_{i+1,n}^{(k-1)}(t) \\ \vdots \\ \tilde{\mathbf{y}}_{r,n}^{(k-1)}(t) \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{Z}}_{i,n}^{(k)}(t) := \begin{bmatrix} \tilde{\mathbf{z}}_{1,n}^{(k)}(t) \\ \vdots \\ \tilde{\mathbf{z}}_{i,n}^{(k)}(t) \\ \tilde{\mathbf{z}}_{i+1,n}^{(k-1)}(t) \\ \vdots \\ \tilde{\mathbf{z}}_{r,n}^{(k-1)}(t) \end{bmatrix}.$$

**Remark 5.3.** The Gauß-Seidel-type iteration scheme of Definition 5.2 defines consistent splitting functions  $\mathbf{F}$  and  $\mathbf{G}$ , i.e., they fulfill the compatibility condition and inherit their smoothness and differentiability from  $\mathbf{f}$  and  $\mathbf{g}$ , see Definition 5.1. Please note that the splitting functions reflect changes in the computational sequence. In other words: it matters in which sequence the subsystems (5.21) are solved.

For coupled ODEs (short ‘ODE-ODE’) various splitting schemes can be shown to be convergent, [27] and similar results are known for the special case of an ODE coupled to an algebraic equation (short ‘ODE-AE’), [136].

**Example 5.2** (Gauß-Seidel examples). Let us consider the following examples

- (i) The Gauß-Seidel type scheme for  $r = 2$  ODE subsystems has the following form

$$\begin{aligned} \dot{\tilde{\mathbf{y}}}_1^{(k)} &= \mathbf{f}_1(\tilde{\mathbf{y}}_1^{(k)}, \tilde{\mathbf{y}}_2^{(k-1)}), \\ \dot{\tilde{\mathbf{y}}}_2^{(k)} &= \mathbf{f}_2(\tilde{\mathbf{y}}_1^{(k)}, \tilde{\mathbf{y}}_2^{(k)}), \end{aligned} \quad \text{with} \quad \mathbf{F} = \begin{bmatrix} \mathbf{f}_1(\tilde{\mathbf{y}}_1^{(k)}, \tilde{\mathbf{y}}_2^{(k-1)}) \\ \mathbf{f}_2(\tilde{\mathbf{y}}_1^{(k)}, \tilde{\mathbf{y}}_2^{(k)}) \end{bmatrix}.$$

- (ii) The Gauß-Seidel type scheme for an ODE and AE subsystem, i.e., the ‘fractional

step method for index-1 DAEs', has the following form

$$\begin{aligned} \dot{\tilde{\mathbf{y}}}^{(k)} &= \mathbf{f}(\tilde{\mathbf{y}}^{(k)}, \tilde{\mathbf{z}}^{(k-1)}), & \mathbf{F} &= \mathbf{f}(\tilde{\mathbf{y}}^{(k)}, \tilde{\mathbf{z}}^{(k-1)}), \\ 0 &= \mathbf{g}(\tilde{\mathbf{y}}^{(k)}, \tilde{\mathbf{z}}^{(k)}), & \mathbf{G} &= \mathbf{g}(\tilde{\mathbf{y}}^{(k)}, \tilde{\mathbf{z}}^{(k)}). \end{aligned} \quad \text{with}$$

Please note that the old waveforms  $\tilde{\mathbf{y}}_2^{(k-1)}$  and  $\tilde{\mathbf{z}}^{(k-1)}$  enter in both examples only in the first differential equation. The importance of that fact becomes clear by Cor. 5.6.

In contrast to the special cases above, the general DAE case is more involved: the algebraic constraints (5.22b) may depend on old algebraic variables, i.e., those of the previous iteration  $\tilde{\mathbf{z}}_n^{(k-1)}$ . They can cause divergence and thus the scheme is carefully analyzed in the following: Section 5.3.3 gives a fixed point analysis in function space of the iteration scheme and Section 5.3.4 carries out an error analysis (using the fixed point argument) to prove stability and convergence.

The reader who is not interested in mathematical details and proofs may skip to Section 5.3.5 or Section 5.3.6 for applications in field/circuit and semiconductor/circuit coupling, respectively.

### 5.3.3 Recursion Estimate

The  $k$ -th dynamic iteration starts from a given waveform. It stems either from a previous iteration or from an extrapolation step. In any case it must already be sufficiently close to the exact solution to guarantee contraction. This is not clear because the number of iterations ( $k_n$ ) is finite and thus a splitting error occurs on each window (and this error is even propagated to subsequent windows). These errors and their propagation are discussed in Section 5.3.4. For now the quantification of 'close' is given by the following neighborhood in function space:

**Definition 5.3** (Neighborhood of the solution). For each time window  $T_n$  the distance  $d > 0$  defines a neighborhood of the exact solution

$$\mathcal{U}_{d,n} = \left\{ \mathbf{X} := \begin{bmatrix} \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} \in C_n^{1,0} : \|\mathbf{Y} - \mathbf{y}|_{[T_n, T_{n+1}]}\|_{2,\infty}, \|\mathbf{Z} - \mathbf{z}|_{[T_n, T_{n+1}]}\|_{2,\infty} \leq d \right\}, \quad (5.23)$$

where  $\|\mathbf{v}(t)\|_{2,\infty} := \max_t \|\mathbf{v}(t)\|_2$  utilizes the maximum-norm in time and the Euclidean norm in space.

The splitting functions must fulfill the following smoothness properties on that neighborhood, which are typically inherited from the original right-hand-side functions  $\mathbf{f}$  and  $\mathbf{g}$ :

**Assumption 5.3** (Smoothness in the neighborhood of the solution). *Let the problem (5.11) with consistent splitting functions  $\mathbf{F}$ ,  $\mathbf{G}$  be given, then it is assumed that there is a  $d_0 > 0$  with*

- (a) *the (differential) splitting Function  $\mathbf{F}$  is Lipschitz-continuous on  $\mathcal{U}_{d_0,n}$  with constant  $L_{\mathbf{F}} > 0$*
- (b) *the (algebraic) splitting Function  $\mathbf{G}$  is totally differentiable with Lipschitz-continuous derivatives on  $\mathcal{U}_{d_0,n}$*

(c) the partial derivative  $\mathbf{G}_{\mathbf{z}^{(k)}}$  is invertible on  $\mathcal{U}_{d_0,n}$ .

Ass. 5.3 ensures that the split problems (5.18) are index-1 and have a well-defined solution. For the Gauß-Seidel iteration the smoothness of the neighborhood in Ass. 5.3 is implied by the smoothness of the right-hand-sides, see Ass. 5.2 and Definition 5.2.

Now, let the functions  $\mathbf{X}, \tilde{\mathbf{X}} \in \mathcal{U}_{d_0,n}$  be given and further  $k$  denote the number of iterations on the  $n$ -th time window. This allows for the following abbreviations, [9]

$$\mathbf{Y}_n^k := \Psi_{\mathbf{y},n}^k \mathbf{X}, \quad \mathbf{Z}_n^k := \Psi_{\mathbf{z},n}^k \mathbf{X}, \quad \text{and} \quad \tilde{\mathbf{Y}}_n^k := \Psi_{\mathbf{y},n}^k \tilde{\mathbf{X}}, \quad \tilde{\mathbf{Z}}_n^k := \Psi_{\mathbf{z},n}^k \tilde{\mathbf{X}}, \quad (5.24)$$

that are introduced to measure the distance of a approximation from the exact solution after  $k$  iterations, i.e., the following abbreviation for the differences

$$\begin{aligned} \Delta_{\mathbf{y},n}^k(t) &:= \mathbf{Y}_n^k(t) - \tilde{\mathbf{Y}}_n^k(t), & \delta_{\mathbf{y},n}^k &:= \|\Delta_{\mathbf{y},n}^k\|_{2,\infty}, \\ \Delta_{\mathbf{z},n}^k(t) &:= \mathbf{Z}_n^k(t) - \tilde{\mathbf{Z}}_n^k(t), & \delta_{\mathbf{z},n}^k &:= \|\Delta_{\mathbf{z},n}^k\|_{2,\infty}. \end{aligned} \quad (5.25)$$

Based on the result in [7] for particular waveforms, the following general result yields an estimate for the dynamic iteration of two arbitrary waveforms  $\mathbf{X}$  and  $\tilde{\mathbf{X}}$  on  $\mathcal{U}_{d,n}$

**Lemma 5.4** (Recursion estimate). *Let Ass. 5.1 and Ass. 5.3 be fulfilled for the consistent splitting functions  $\mathbf{F}, \mathbf{G}$ . Then there is a constant  $C > 1$ , such that for a distance*

$$d < \min\{d_0/C, 1/(4C)\}$$

and a window size  $H_n < H_{max}$ , two waveforms

$$\mathbf{X}_n^{k-1}, \tilde{\mathbf{X}}_n^{k-1} \in \mathcal{U}_{d,n}. \quad (5.26)$$

fulfill the recursion estimate

$$\begin{bmatrix} \delta_{\mathbf{y},n}^k \\ \delta_{\mathbf{z},n}^k \end{bmatrix} \leq \mathbf{K} \begin{bmatrix} \delta_{\mathbf{y},n}^{k-1} \\ \delta_{\mathbf{z},n}^{k-1} \end{bmatrix} + \begin{bmatrix} 1 + CH_n \\ C \end{bmatrix} \|\Delta_{\mathbf{y},n}^0(T_n)\|_2 \quad (5.27)$$

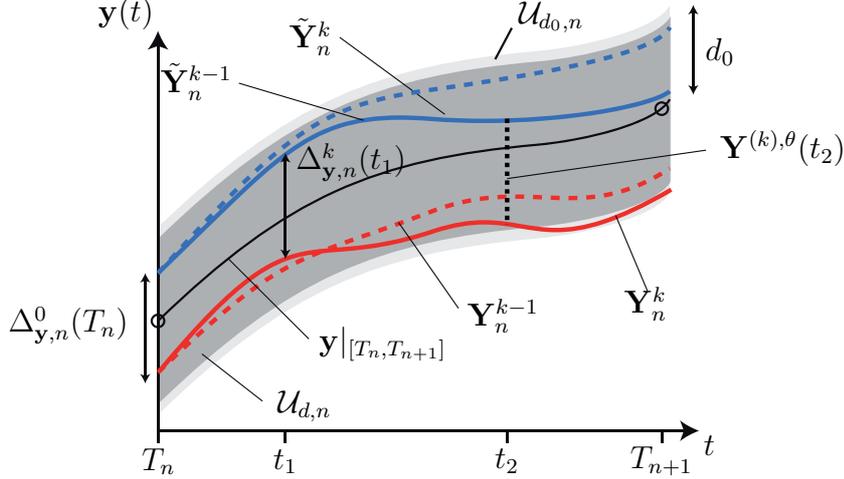
with recursion matrix and contraction factor

$$\mathbf{K} := \begin{bmatrix} CH_n & CH_n \\ C & CH_n + \alpha_n \end{bmatrix} \quad \text{and} \quad \alpha_n := (1 + Cd) \|\mathbf{G}_{\mathbf{z}^{(k)}}^{-1} \mathbf{G}_{\mathbf{z}^{(k-1)}}\|_{2,\infty} + Cd, \quad (5.28)$$

respectively.

*Proof.* The proof follows [9] and generalizes [5, 7]. It is split into two parts: in the first part the recursion estimate (5.27) for the differential unknowns  $\Delta_{\mathbf{y},n}^k$  is shown (similarly to the classical proof of the Picard-Lindelöf iteration). The second part proves the estimate for the algebraic unknowns  $\Delta_{\mathbf{z},n}^k$  using a homotopy between the two waveforms (5.26), see Fig. 5.6.

*Estimate for the differential components.* Inserting the two waveforms (5.26) into (5.18a) yields two differential equations. Subtracting one from the other and then integrating over



**Figure 5.6:** Contraction of waveforms. A differential component is shown for the  $k$ -th iteration of the time window  $T_n$ , see the proof of Lemma 5.4.

the interval  $[T_n, \tau]$  with the limits  $T_n < \tau \leq T_{n+1}$  yields

$$\|\Delta_{\mathbf{y},n}^k(\tau)\|_2 \leq \|\Delta_{\mathbf{y},n}^k(T_n)\|_2 + L_{\mathbf{F}} \int_{T_n}^{\tau} (\|\Delta_{\mathbf{y},n}^k\|_2 + \|\Delta_{\mathbf{y},n}^{k-1}\|_2 + \|\Delta_{\mathbf{z},n}^k\|_2 + \|\Delta_{\mathbf{z},n}^{k-1}\|_2) dt,$$

using the consistency of  $\mathbf{F}$  and its Lipschitz-continuity on  $\mathcal{U}_{d_0,n}$  (Ass. 5.3). To obtain the new waveforms at iteration  $k$  the solvability of the ODE (5.18a) is required, which can be assured for a sufficiently small time window  $[T_n, \tau]$  by standard ODE theory. The smallness is concretized in (5.33) by the constant  $C$ . From the fact that the initial offset is given by error propagation and thus cannot be improved by iterations, follows

$$\leq \|\Delta_{\mathbf{y},n}^0(T_n)\|_2 + L_{\mathbf{F}} \int_{T_n}^{\tau} (\|\Delta_{\mathbf{y},n}^k\|_2 + \|\Delta_{\mathbf{y},n}^{k-1}\|_2 + \|\Delta_{\mathbf{z},n}^k\|_2 + \|\Delta_{\mathbf{z},n}^{k-1}\|_2) dt. \quad (5.29)$$

The Implicit Function Theorem (with Ass. 5.3) allows us to solve the algebraic equation (5.18b) for  $\mathbf{Z}^{(k)} = \hat{\zeta}(\mathbf{Y}^{(k)}, \mathbf{Y}^{(k-1)}, \mathbf{Z}^{(k-1)})$  and analogously for  $\tilde{\mathbf{Z}}^{(k)}$ . This yields

$$\begin{aligned} \|\Delta_{\mathbf{z},n}^k\|_2 &= \|\hat{\zeta}(\mathbf{Y}_n^{(k)}, \mathbf{Y}_n^{(k-1)}, \mathbf{Z}_n^{(k-1)}) - \hat{\zeta}(\tilde{\mathbf{Y}}_n^{(k)}, \tilde{\mathbf{Y}}_n^{(k-1)}, \tilde{\mathbf{Z}}_n^{(k-1)})\|_2 \\ &\leq L_{\hat{\zeta}} \left( \|\Delta_{\mathbf{y},n}^k\|_2 + \|\Delta_{\mathbf{y},n}^{k-1}\|_2 + \|\Delta_{\mathbf{z},n}^{k-1}\|_2 \right) \end{aligned} \quad (5.30)$$

with the Lipschitz constant  $L_{\hat{\zeta}} > 0$ . Inserting this result into (5.29) yields the maximum

$$\Delta_{\mathbf{y},n}^k \leq \|\Delta_{\mathbf{y},n}^0(T_n)\|_2 + L_0 H_n (\Delta_{\mathbf{y},n}^k + \Delta_{\mathbf{y},n}^{k-1} + \Delta_{\mathbf{z},n}^{k-1})$$

with  $L_0 := L_{\mathbf{F}}(1 + L_{\hat{\zeta}})$ , then reordering yields

$$\leq \left( 1 + \frac{L_0}{1 - L_0 H_n} H_n \right) \|\Delta_{\mathbf{y},n}^0(T_n)\|_2 + \frac{L_0}{1 - L_0 H_n} H_n (\delta_{\mathbf{y},n}^{k-1} + \delta_{\mathbf{z},n}^{k-1}). \quad (5.31)$$

Let  $H_n < H_{\max} := 1/(2L_0)$  then  $L_0 H_n < 1/2$  implies

$$\delta_{\mathbf{y},n}^k \leq (1 + 2L_0 H) \|\Delta_{\mathbf{y},n}^{k-1}(T_n)\|_2 + 2L_0 H (\delta_{\mathbf{y},n}^{k-1} + \delta_{\mathbf{z},n}^{k-1}). \quad (5.32)$$

This proves the differential part of the estimate (5.27). To verify that the new waveforms are in the neighborhood, the distance from the exact solution is measured. This can be done using the estimates above but for the special cases, that one of the waveforms is identified as the exact solution, i.e., the fixed point. We set

$$\begin{aligned} \mathbf{X}_n^k &= \mathbf{x}|_{[T_n, T_{n+1}]} \quad \text{such that} \quad \Delta_{\mathbf{x},n}^k = \tilde{\mathbf{X}}_n^k - \mathbf{x}|_{[T_n, T_{n+1}]} \\ \text{and } \tilde{\mathbf{X}}_n^k &= \mathbf{x}|_{[T_n, T_{n+1}]} \quad \text{such that} \quad \Delta_{\mathbf{x},n}^k = \mathbf{X}_n^k - \mathbf{x}|_{[T_n, T_{n+1}]} \end{aligned}$$

Then the following a priori estimates are found from the estimate (5.32)

$$\begin{aligned} |\mathbf{Y}_n^k - \mathbf{y}|_{[T_n, T_{n+1}]}, |\tilde{\mathbf{Y}}_n^k - \mathbf{y}|_{[T_n, T_{n+1}]} &\leq 7d, \\ |\mathbf{Z}_n^k - \mathbf{z}|_{[T_n, T_{n+1}]}, |\tilde{\mathbf{Z}}_n^k - \mathbf{z}|_{[T_n, T_{n+1}]} &\leq 9L_\zeta d, \end{aligned} \quad (5.33)$$

and thus  $d < d_0/C$  implies  $\mathbf{X}_n^k, \tilde{\mathbf{X}}_n^k \in \mathcal{U}_{d_0,n}$  for a constant

$$C > \max \left\{ 2L_0, 7, 9L_\zeta \right\}.$$

*Estimate for the algebraic components.* In the second part of the proof, the inequality for the algebraic component is shown. The key is the following homotopy for  $\theta \in [0, 1]$ :

$$\mathbf{Y}^{(k),\theta}(t) := \theta \tilde{\mathbf{Y}}_n^k(t) + (1 - \theta) \mathbf{Y}_n^k(t), \quad \text{and} \quad \mathbf{Z}^{(k),\theta}(t) := \theta \tilde{\mathbf{Z}}_n^k(t) + (1 - \theta) \mathbf{Z}_n^k(t).$$

Insertion of the homotopies into the splitting function  $\mathbf{G}$  defines implicitly an ‘overloaded’ version of  $\mathbf{G}$  in the only parameter  $\theta$

$$\mathbf{G}(\theta) := \mathbf{G}(\mathbf{Y}^{(k),\theta}, \mathbf{Y}^{(k-1),\theta}, \mathbf{Z}^{(k),\theta}, \mathbf{Z}^{(k-1),\theta}) \quad \text{and} \quad \mathbf{G}_{\mathbf{u}}(\theta) := \frac{\partial \mathbf{G}}{\partial \mathbf{u}}(\theta).$$

where  $\mathbf{u}$  denotes an arbitrary argument of the splitting function. The waveforms fulfill the algebraic constraint and thus it holds that  $\mathbf{G}(0) = \mathbf{G}(1) = 0$ . Therefore the integral

$$\begin{aligned} 0 &= \mathbf{G}(1) - \mathbf{G}(0) \\ &= \int_0^1 \left( \mathbf{G}_{\mathbf{y}^{(k)}}(\theta) \Delta_{\mathbf{y},n}^k + \mathbf{G}_{\mathbf{y}^{(k-1)}}(\theta) \Delta_{\mathbf{y},n}^{k-1} + \mathbf{G}_{\mathbf{z}^{(k)}}(\theta) \Delta_{\mathbf{z},n}^k + \mathbf{G}_{\mathbf{z}^{(k-1)}}(\theta) \Delta_{\mathbf{z},n}^{k-1} \right) d\theta \end{aligned} \quad (5.34)$$

is obtained, where all partial derivatives are identified as the error abbreviations (5.25), e.g.,  $\frac{\partial}{\partial \theta} \mathbf{Y}^{(k),\theta} = \tilde{\mathbf{Y}}_n^k - \mathbf{Y}_n^k = \Delta_{\mathbf{y},n}^k$ . The Estimate (5.33) and Ass. 5.3 guarantee that all arguments of  $\mathbf{G}$  are in the neighborhood  $\mathcal{U}_{d,n}$ . Thus the Lipschitz continuity of  $\mathbf{G}$  on  $\mathcal{U}_{d,n}$  ( $Cd \leq d_0$ ) with the constant  $L_{\mathbf{G}'}$  yields

$$\begin{aligned} \|\mathbf{G}_{\mathbf{u}}(\theta) - \mathbf{G}_{\mathbf{u}}(0)\|_2 &\leq L_{\mathbf{G}'} \left( \|\theta \tilde{\mathbf{Y}}_n^k + (1 - \theta) \mathbf{Y}_n^k - \mathbf{Y}_n^k\|_2 + \dots \right. \\ &\quad \left. \dots + \|\theta \tilde{\mathbf{Z}}_n^{k-1} + (1 - \theta) \mathbf{Z}_n^{k-1} - \mathbf{Z}_n^{k-1}\|_2 \right) \end{aligned}$$

$$\begin{aligned}
 &= L_{\mathbf{G}'} \theta \left( \|\Delta_{\mathbf{y},n}^k\|_2 + \|\Delta_{\mathbf{y},n}^{k-1}\|_2 + \|\Delta_{\mathbf{z},n}^k\|_2 + \|\Delta_{\mathbf{z},n}^{k-1}\|_2 \right) \\
 &\leq 18(1 + L_{\hat{\zeta}}) L_{\mathbf{G}'} d.
 \end{aligned} \tag{5.35}$$

Now, Ass. 5.3 guarantees the regularity of  $\mathbf{G}_{\mathbf{z}^{(k)}}(0)$  and thus left-multiplication of (5.34) yields

$$\begin{aligned}
 0 = \int_0^1 \mathbf{G}_{\mathbf{z}^{(k)}}^{-1}(0) &\left( \left( \mathbf{G}_{\mathbf{z}^{(k)}}(0) + [\mathbf{G}_{\mathbf{z}^{(k)}}(\theta) - \mathbf{G}_{\mathbf{z}^{(k)}}(0)] \right) \Delta_{\mathbf{z},n}^k \right. \\
 &+ \left( \mathbf{G}_{\mathbf{z}^{(k-1)}}(0) + [\mathbf{G}_{\mathbf{z}^{(k-1)}}(\theta) - \mathbf{G}_{\mathbf{z}^{(k-1)}}(0)] \right) \Delta_{\mathbf{z},n}^{k-1} \\
 &+ \mathbf{G}_{\mathbf{y}^{(k)}}(\theta) \Delta_{\mathbf{y},n}^k \\
 &\left. + \mathbf{G}_{\mathbf{y}^{(k-1)}}(\theta) \Delta_{\mathbf{y},n}^{k-1} \right) d\theta.
 \end{aligned} \tag{5.36}$$

Furthermore Ass. 5.3 (smoothness) guarantees that the Jacobians  $\mathbf{G}_{\mathbf{z}^{(k)}}^{-1}$ ,  $\mathbf{G}_{\mathbf{z}^{(k-1)}}$ ,  $\mathbf{G}_{\mathbf{y}^{(k)}}$  and  $\mathbf{G}_{\mathbf{y}^{(k-1)}}$  are uniformly bounded on  $\mathcal{U}_{d_0,n}$ . Let the corresponding constant be denoted by  $c_g$ . Then solving equation (5.36) for  $\mathbf{G}_{\mathbf{z}^{(k)}}^{-1}(0) \mathbf{G}_{\mathbf{z}^{(k)}}(0) \Delta_{\mathbf{z},n}^k = \Delta_{\mathbf{z},n}^k$  and application of the maximum norm in conjunction with estimate (5.35) yields

$$\delta_{\mathbf{z},n}^k \leq \left( \|\mathbf{G}_{\mathbf{z}^{(k)}}^{-1} \mathbf{G}_{\mathbf{z}^{(k-1)}}\|_{2,\infty} + \frac{\tilde{c}}{2} d \right) \delta_{\mathbf{z},n}^{k-1} + \frac{\tilde{c}}{2} d \delta_{\mathbf{z},n}^k + c_g^2 (\delta_{\mathbf{y},n}^k + \delta_{\mathbf{y},n}^{k-1}),$$

with the constant  $\tilde{c} := 36(1 + L_{\hat{\zeta}}) L_{\mathbf{G}'} c_g$  and

$$\|\mathbf{G}_{\mathbf{z}^{(k)}}^{-1} \mathbf{G}_{\mathbf{z}^{(k-1)}}\|_{2,\infty} = \max_{t \in [T_n, T_{n+1}]} \|\mathbf{G}_{\mathbf{z}^{(k)}}^{-1} \mathbf{G}_{\mathbf{z}^{(k-1)}}\|_2(\mathbf{Y}_n^k(t), \mathbf{Y}_n^k(t), \mathbf{Z}_n^k(t), \mathbf{Z}_n^k(t)).$$

Finally the estimate (5.32) with  $H_n < H_{\max}$  and  $d < \min \{d_0/C, \frac{1}{2\tilde{c}}\}$  give

$$\begin{aligned}
 \delta_{\mathbf{z},n}^k &\leq 3(1 + \tilde{c}d) c_g^2 \left( \|\Delta_{\mathbf{y},n}^{k-1}(t_n)\|_2 + \delta_{\mathbf{y},n}^{k-1} \right) \\
 &\quad + (1 + \tilde{c}d) \left( 2c_g^2 L_0 H + \|\mathbf{G}_{\mathbf{z}^{(k)}}^{-1} \mathbf{G}_{\mathbf{z}^{(k-1)}}\|_{2,\infty} + \frac{\tilde{c}}{2} d \right) \delta_{\mathbf{z},n}^{k-1} \\
 &\leq C \left( \|\Delta_{\mathbf{y},n}^{k-1}(T_n)\|_2 + \delta_{\mathbf{y},n}^{k-1} \right) + (CH_n + \alpha_n) \delta_{\mathbf{z},n}^{k-1}.
 \end{aligned} \tag{5.37}$$

and this proves the algebraic part of the estimate (5.27). Using  $d < d_0$  the global constant

$$C > \max \left\{ 2L_0, 7, 9L_{\hat{\zeta}}, 3(1 + \tilde{c}d_0) c_g^2, (1 + \tilde{c}d_0) c_g^2 L_0, \frac{\tilde{c}}{2} \right\}$$

is large enough to deduce all estimates, i.e., (5.32), (5.33) and (5.37).  $\square$

Lemma 5.4 and especially the proof above distinguish between the different (Lipschitz) constants and their respective origin. On the other hand the estimates in [5] are rougher. Naturally the recursion estimate (5.27) can be brought to the form of Lemma 3.1 as it is given in [5].

**Remark 5.4** (Rougher recursion estimate). In the infinity norm the initial offset is bounded by the maximum error on time window, i.e.,  $|\Delta_{\mathbf{y},n}^{k-1}(T_n)| \leq \delta_{\mathbf{y},n}^{k-1}$ . Thus (5.27) implies the following (rougher) estimate

$$\begin{pmatrix} \delta_{\mathbf{y},n}^k \\ \delta_{\mathbf{z},n}^k \end{pmatrix} \leq \mathbf{K} \begin{pmatrix} \delta_{\mathbf{y},n}^0 \\ \delta_{\mathbf{z},n}^{k-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \|\Delta_{\mathbf{y},n}^{k-1}(T_n)\|_2, \tag{5.38}$$

with a (possibly) larger constant  $C$ . The rougher estimate matches the structure in [5, 7].

The consequences of the iterative application of estimates of the structure (5.38) are discussed in [5]. The following result is easily transferred to the general index-1 setting, it holds that

**Proposition 5.5** (Iteration estimate). *Let the same assumptions as for Lemma 5.4 be fulfilled with a constant  $C > 1 > \alpha_n$ . Then there is a new constant  $C_0 > C$  such that for all  $k \geq 1$  and  $H_n \leq H_{max}$  it holds:*

$$\begin{pmatrix} \delta_{\mathbf{y},n}^k \\ \delta_{\mathbf{z},n}^k \end{pmatrix} \leq \begin{pmatrix} C(4C+1)H_n\mu_n^{\max(0,k-2)} & 4CH_n\mu_n^{k-2} \\ 4C\mu_n^{k-1} & \mu_n^k + (\mu_n - \alpha_n)^k \end{pmatrix} \begin{pmatrix} \delta_{\mathbf{y},n}^0 \\ \delta_{\mathbf{z},n}^0 \end{pmatrix} + \begin{pmatrix} 1+C_0H_n \\ C_0 \end{pmatrix} \delta_{\mathbf{y},n}^0(T_n) \quad (5.39)$$

where

$$\mu_n = \mu(\alpha_n, H_n) := \alpha_n + \frac{2CH_n}{\frac{\alpha_n}{2C} + \sqrt{H_n}}. \quad (5.40)$$

*Proof.* The total error after  $k$  iterations corresponds to the iterative application of estimate (5.38). This is the multiplication by the  $k$ -th power of the matrix  $\mathbf{K}$ , see (5.28). Thus the claim is deduced using the eigenvalues of  $\mathbf{K}$ , i.e.,

$$\lambda_{1,2}(\mathbf{K}) = \frac{1}{2} \left( \alpha_n + 2CH_n \pm \sqrt{\alpha_n^2 + 4C^2H_n} \right), \quad (5.41)$$

the details of the proof can be found in [5].  $\square$

Without discussing the contraction of the recursion above (i.e., spectral radius  $\rho(\mathbf{K}) < 1$ ) and shifting the analysis of the contraction factor  $\alpha_n$  to Theorem 5.7, the following corollary is immanent

**Corollary 5.6** (Simple coupling and convergence Rate). *The eigenvalues (5.41) determine the rate of contraction of the recursion above for the limit  $H_n \rightarrow 0$  and  $\alpha < 1$ .*

(i) *Given a general splitting with*

$$4C^2H_n < \alpha_n^2,$$

*then the convergence rate is  $\alpha_n + \mathcal{O}(H_n)$ .*

(ii) *Given a splitting ('simple coupling'), where no algebraic constraint depends on old algebraic iterates, i.e.,*

$$\mathbf{G}_{\mathbf{z}^{(k-1)}} = 0,$$

*then  $\alpha_n = 0$  and the convergence rate is  $\mathcal{O}(\sqrt{H_n})$ .*

(iii) *Given a splitting, where no algebraic constraint depends on old iterates, i.e.,*

$$\mathbf{G}_{\mathbf{z}^{(k-1)}} = 0 \quad \text{and} \quad \mathbf{G}_{\mathbf{y}^{(k-1)}} = 0,$$

*then the convergence rate is  $\mathcal{O}(H_n)$ .*

*Proof.* The first claim (i) is shown by Taylor expansion of (5.41). It holds

$$\sqrt{\alpha_n^2 + 4C^2H_n} = \alpha_n(1 + 2C^2H_n/\alpha_n^2) + \mathcal{O}(H_n^2).$$

and this, together with the assumptions, concludes the proof. The other claims (i) and (ii) are shown in the same way as Lemma 5.4. One exploits that some Jacobians vanish, namely  $\mathbf{G}_{\mathbf{z}^{(k-1)}}$  and  $\mathbf{G}_{\mathbf{y}^{(k-1)}}$ .  $\square$

The different convergence rates imply that it is beneficial to design the coupling interface in such a way that the contraction factor  $\alpha_n$  vanishes. This is for example obtained if algebraic couplings are avoided. This guarantees immediately a higher order convergence (cf., ‘simple coupling’ in [7]). Whether a differential coupling can be achieved or not, depends obviously on the particular problem, the splitting scheme and especially on the computational sequence of the subsystems, [5]. This will be discussed in more detail in Section 5.3.5 and Section 5.3.6 for examples in field/circuit and semiconductor/circuit applications, respectively.

**Remark 5.5** (Convergence rate of the fractional step method). Cor. 5.6 includes two important special cases: the fractional step method, see Example 5.2, exhibits a convergence rate in the order of  $\mathcal{O}(H_n)$ . This special case is discussed in [136]. The same rate applies to the ODE-ODE Gauß-Seidel splitting, but the theory in this treatise is obviously not tailored for a deeper understanding of the ODE case, see [27].

Now, having established a recursion estimate for a finite number of iterations, it will be used in the next section to obtain convergence and stability of the scheme.

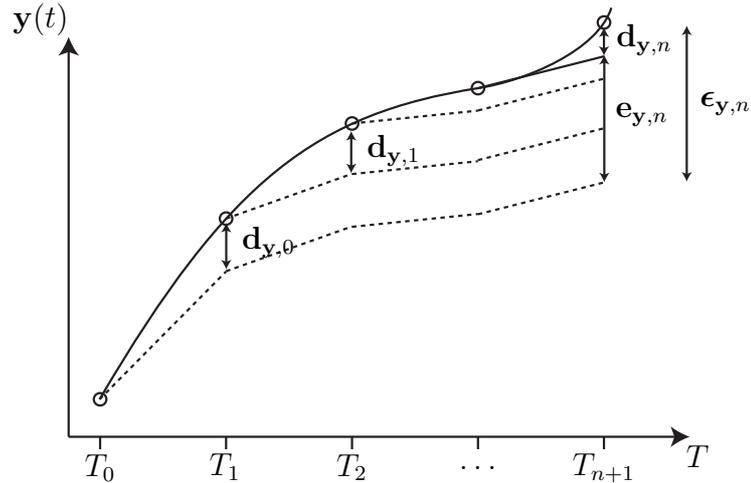
### 5.3.4 Convergence and Stability

For convergence and stability it will be shown that the global splitting error can be controlled only by the window size and a finite number of iterations. The analysis consists of two parts, firstly the local splitting error (due to stopping after a finite number of iterations) and secondly the propagation error (transport of local errors from window to window). Discretization errors originating from numerical time integration are disregarded, i.e., it is assumed that problems (5.21) are solved analytically. The next sections follow the approach in [9], where the results of [5, 7] are transferred to the new (general) setting (5.11). In the following the local and global splitting errors are defined, similar as in the analysis of standard time-integration methods, [69]. The errors and their propagation are visualized by Lady Windermere’s Fan, Fig. 5.7.

#### Definition of the Splitting Errors

Following the standard technique in the error analysis of time-integration methods, e.g. [68], the global error is split into contributions as depicted by Lady Windermere’s Fan, Fig. 5.7. Each error definition corresponds to a particular difference of two waveforms as defined in (5.25). Let us start with the global error:

**Definition 5.4** (Global splitting error). The global error for  $k$  iterations on the  $n$ -th time



**Figure 5.7:** Lady Windermere's Fan. Error propagation for the differential component. The solid line depicts the exact solution, the dotted lines are approximations, [69].

window ( $t \in [T_n, T_{n+1}]$ ) is defined by

$$\boldsymbol{\epsilon}_{\mathbf{x},n}^k := \begin{bmatrix} \boldsymbol{\epsilon}_{\mathbf{y},n}^k \\ \boldsymbol{\epsilon}_{\mathbf{z},n}^k \end{bmatrix} := \begin{bmatrix} (\tilde{\mathbf{y}} - \mathbf{y})|_{[T_n, T_{n+1}]} \\ (\tilde{\mathbf{z}} - \mathbf{z})|_{[T_n, T_{n+1}]} \end{bmatrix} = (\Psi_n^k \circ \Phi_n) \begin{bmatrix} \tilde{\mathbf{y}}|_{[T_{n-1}, T_n]} \\ \tilde{\mathbf{z}}|_{[T_{n-1}, T_n]} \end{bmatrix} - \begin{bmatrix} \mathbf{y}|_{[T_n, T_{n+1}]} \\ \mathbf{z}|_{[T_n, T_{n+1}]} \end{bmatrix},$$

where  $\tilde{\mathbf{y}}(t)$  and  $\tilde{\mathbf{z}}(t)$  denote the numerical approximations of the exact solution given by  $\mathbf{y}(t)$  and  $\mathbf{z}(t)$ , see (5.20).

The global error consists of the local splitting error of the current window and of the splitting errors of previous windows, that are propagated by the initial values of each window.

**Definition 5.5** (Local splitting error). The local splitting error

$$\mathbf{d}_{\mathbf{x},n}^k := \begin{bmatrix} \mathbf{d}_{\mathbf{y},n}^k \\ \mathbf{d}_{\mathbf{z},n}^k \end{bmatrix} := (\Psi_n^k \circ \Phi_n) \begin{bmatrix} \mathbf{y}|_{[T_{n-1}, T_n]} \\ \mathbf{z}|_{[T_{n-1}, T_n]} \end{bmatrix} - \Psi_n^k \begin{bmatrix} \mathbf{y}|_{[T_n, T_{n+1}]} \\ \mathbf{z}|_{[T_n, T_{n+1}]} \end{bmatrix} \quad (5.42)$$

describes the difference between the exact solution  $\mathbf{x}|_{[T_n, T_{n+1}]}$  and the approximation after  $k$  iterations that is obtained when starting from the exact data  $\mathbf{x}|_{[T_{n-1}, T_n]}$  on the  $n$ -th window.

The definition above exploits the fact that the exact solution  $\mathbf{x}$  is a fixed point of  $\Psi_n$ , i.e.,  $\mathbf{x} = \Psi_n^k \mathbf{x}$  and thus the definition above gives indeed the local error.

**Definition 5.6** (Propagated splitting error). The propagated splitting error

$$\mathbf{e}_{\mathbf{x},n}^k := \begin{bmatrix} \mathbf{e}_{\mathbf{y},n}^k \\ \mathbf{e}_{\mathbf{z},n}^k \end{bmatrix} := (\Psi_n^k \circ \Phi_n) \begin{bmatrix} \tilde{\mathbf{y}}|_{[T_{n-1}, T_n]} \\ \tilde{\mathbf{z}}|_{[T_{n-1}, T_n]} \end{bmatrix} - (\Psi_n^k \circ \Phi_n) \begin{bmatrix} \mathbf{y}|_{[T_{n-1}, T_n]} \\ \mathbf{z}|_{[T_{n-1}, T_n]} \end{bmatrix} \quad (5.43)$$

describes the difference between the solutions that are obtained by  $k$  iterations on window  $n$  when starting from the exact solution  $\mathbf{x}|_{[T_{n-1}, T_n]}$  and the approximation  $\tilde{\mathbf{x}}|_{[T_{n-1}, T_n]}$ .

The splitting error definitions imply the additive error decomposition in the style of Lady Windermere's Fan, Fig. 5.7,

$$\begin{bmatrix} \boldsymbol{\epsilon}_{\mathbf{y},n} \\ \boldsymbol{\epsilon}_{\mathbf{z},n} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_{\mathbf{y},n} \\ \mathbf{d}_{\mathbf{z},n} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{\mathbf{y},n} \\ \mathbf{e}_{\mathbf{z},n} \end{bmatrix} \quad (5.44)$$

where the superscript for the iteration number of the error on window  $n$  is disregarded if a fixed number  $k_n$  of iteration is applied, e.g.,  $\boldsymbol{\epsilon}_{\mathbf{y},n} := \boldsymbol{\epsilon}_{\mathbf{y},n}^{k_n}$ .

### Splitting Error Analysis

Let us start with the analysis of the local error, [5, 7]. The exact solution  $\mathbf{x}$  is compared to the splitting error on the  $n$ -th window after extrapolation and  $k$  iterations. The following theorem from [9] holds

**Theorem 5.7** (Contraction). *Let an index-1 DAE (5.11) with Ass. 5.1, the constant extrapolation operator and consistent splitting functions (Ass. 5.3) be given. Then for  $d$  and  $H < H_0$  small enough, the hypothesis*

$$\|\mathbf{G}_{\mathbf{z}^{(k)}}^{-1} \mathbf{G}_{\mathbf{z}^{(k-1)}}\|_{2,\infty} < 1 \quad (5.45)$$

*implies that the local error mapping decreases strictly for all  $k > 1$ , i.e.,*

$$\|\mathbf{d}_{\mathbf{x},n}^{k-1}\|_{2,\infty} > \|\mathbf{d}_{\mathbf{x},n}^k\|_{2,\infty}, \quad (5.46)$$

*such that  $\Psi$  is strongly contractive.*

*Proof.* The proof assumes a constant extrapolation with accuracy  $\mathcal{O}(H_n)$ , see (5.16). The generalization to higher order schemes is straightforward. Contraction is shown by induction on  $k$ .

*Induction Basis.* For the differential equation at iteration  $k = 0$  follows

$$\|\mathbf{y}(\tau) - \mathbf{y}^0(\tau)\|_2 = \left\| \int_{T_n}^{\tau} \mathbf{f}(\mathbf{y}, \mathbf{z}) dt \right\|_2 \leq c_f H_n, \quad \tau \in [T_n, T_{n+1}], \quad (5.47)$$

with  $c_f := \|\mathbf{f}(\mathbf{y}|_{[T_n, T_{n+1}]}, \mathbf{z}|_{[T_n, T_{n+1}]})\|_{2,\infty}$ . The index-1 Ass. 5.1 together with the implicit function theorem guarantees the solvability for the algebraic component, i.e.,  $\mathbf{z} = \boldsymbol{\phi}(\mathbf{y})$  and thus

$$\|\mathbf{z}(\tau) - \mathbf{z}^0(\tau)\|_2 = \|\boldsymbol{\phi}(\mathbf{y}(\tau)) - \boldsymbol{\phi}(\mathbf{y}^0(\tau))\|_2 \leq L_\phi c_f H_n \quad (5.48)$$

where  $L_\phi$  denotes the corresponding Lipschitz constant. Thus for  $H_n$  sufficiently small, i.e.,  $H_n < H_0 := \frac{1}{c_f(L_\phi+1)}$ , the extrapolated waveform is in the neighborhood, i.e.,

$$\Phi_n \mathbf{x}|_{[T_{n-1}, T_n]} \in \mathcal{U}_{d,n}. \quad (5.49)$$

*Induction step.* The error decreases if the matrix  $\mathbf{K}$  in the recursion estimate (5.27) has a spectral radius  $\rho(\mathbf{K}) < 1$ , see Prop. 5.5. From the eigenvalues (5.41) it follows that  $\alpha_n < 1$

is sufficient for contraction (given  $d$  and  $H_0$  small enough). By the definition of  $\alpha_n$  in (5.28), this gives us our hypothesis (5.45). The theorem is concluded by Lemma 5.4 where  $\Delta_{\mathbf{y},n}^{k-1}(T_n) = 0$ . There is no initial offset due to the Definition 5.5 of the local error (the exact solution is extrapolated from the previous window).  $\square$

Now that contraction on a single window is guaranteed by hypothesis (5.45) in Theorem 5.7, an upper bound for the local splitting error will be derived. This will eventually give stability and convergence for a single time window.

**Proposition 5.8** (Local error estimate). *Let Ass. 5.1 and Ass. 5.3 be fulfilled for the consistent splitting functions  $\mathbf{F}$ ,  $\mathbf{G}$ . Then for hypothesis (5.45) and  $H$  sufficiently small ( $H < H_0$ ), there is a constant  $C_{\mathbf{d}}^*$  independent of window size  $H_n$ , contraction factor  $\alpha_n$  and iteration number  $k_n$ , such that the local error is bounded in terms of the step size*

$$\|\mathbf{d}_{\mathbf{y},n}\|_2 + H_n \|\mathbf{d}_{\mathbf{z},n}\|_2 \leq C_{\mathbf{d}}^* H_n \delta_n^0, \quad (5.50)$$

with the extrapolation error (5.40)

$$\delta_n^0 := \mu_n^{\max(0, k_n - 2)} \delta_{\mathbf{y},n}^0 + \mu_n^{k_n - 1} \delta_{\mathbf{z},n}^0. \quad (5.51)$$

with  $\mu_n$  as defined in (5.40).

*Proof.* Theorem 5.7 states  $\Phi_n \mathbf{x}|_{[T_{n-1}, T_n]} \in \mathcal{U}_{d,n}$  given  $H$  small enough, see (5.49). Thus the recursion estimate (5.39) in Prop. 5.5 is applicable ( $\alpha_n < 1$ ) for the two particular waveforms  $\mathbf{X} := \mathbf{x}|_{[T_n, T_{n+1}]}$  (exact solution) and  $\tilde{\mathbf{X}} := \Phi_n \mathbf{x}|_{[T_{n-1}, T_n]}$  (extrapolation of exact data). There is no initial offset:  $\delta_{\mathbf{y},n}^0(T_n) = 0$ . Finally summation of both the differential and algebraic estimates conclude the proof using a sufficiently large constant  $C_{\mathbf{d}}^*$ .  $\square$

For a sufficiently small window size  $H_n$  and assuming hypothesis (5.45), it follows  $\mu_n < 1$ . Thus the dynamic iteration on a single window converges to a fixed point as  $k \rightarrow \infty$ . After stopping this fixed point iteration after a finite number of iterations  $k_n$  a local splitting error remains. The propagation of this error to subsequent windows is analyzed in the following section.

## Error Propagation

Due to the windowing technique errors on previous windows accumulate and propagate to the current window, Fig. 5.7. Stability and convergence require that this error is controllable. For this the differential and algebraic propagation errors  $\mathbf{e}_{\mathbf{y},n}$  and  $\mathbf{e}_{\mathbf{z},n}$  are analyzed. Following [9] the Prop. 5.5 is utilized to obtain (cf. [5]):

**Proposition 5.9** (Propagation error). *Let an index-1 DAE (5.11) with Ass. 5.1, the constant extrapolation operator and consistent splitting functions (Ass. 5.3) be given. If  $\mu_n < 1$ , then there is a constant  $C_{\mathbf{e}}^* > 0$ , such that the propagation error is bounded*

$$\begin{bmatrix} \|\mathbf{e}_{\mathbf{y},n}\|_2 \\ \|\mathbf{e}_{\mathbf{z},n}\|_2 \end{bmatrix} \leq \begin{bmatrix} 1 + C_{\mathbf{e}}^* H_n & C_{\mathbf{e}}^* H_n \\ C_{\mathbf{e}}^* & \alpha_n^* \end{bmatrix} \cdot \begin{bmatrix} \|\boldsymbol{\epsilon}_{\mathbf{y},n-1}\|_2 \\ \|\boldsymbol{\epsilon}_{\mathbf{z},n-1}\|_2 \end{bmatrix} \quad (5.52)$$

with  $\alpha_n^*$  depending on the Lipschitz constant  $L_\Phi$  of the extrapolation operator

$$\alpha_n^* := L_\Phi (\mu_n^{k_n} + (\mu_n - \alpha_n)^{k_n}). \quad (5.53)$$

*Proof.* The application of Prop. 5.5 with the waveforms  $\mathbf{X} := \Phi_n \mathbf{x}|_{(T_{n-1}, T_n]}$  (extrapolation of exact data) and  $\tilde{\mathbf{X}} := \Phi_n \tilde{\mathbf{x}}|_{(T_{n-1}, T_n]}$  (extrapolation of erroneous data) yields an initial offset at  $T_n$ , which is bounded by the total error on the previous time window, i.e.,

$$\|\Delta_{\mathbf{y},n}^0(T_n)\|_2 \leq \|\mathbf{y}|_{(T_{n-1}, T_n]} - \tilde{\mathbf{y}}|_{(T_{n-1}, T_n]}\|_2.$$

The extrapolation operator is a uniformly Lipschitz continuous mapping, see (5.16). Let  $L_\Phi$  denote the corresponding constant. It follows

$$\begin{bmatrix} \delta_{\mathbf{y},n}^0 \\ \delta_{\mathbf{z},n}^0 \end{bmatrix} \leq L_\Phi \begin{bmatrix} \|\mathbf{y}|_{(T_{n-1}, T_n]} - \tilde{\mathbf{y}}|_{(T_{n-1}, T_n]}\|_2 \\ \|\mathbf{z}|_{(T_{n-1}, T_n]} - \tilde{\mathbf{z}}|_{(T_{n-1}, T_n]}\|_2 \end{bmatrix} = L_\Phi \begin{bmatrix} \|\mathbf{e}_{\mathbf{y},n-1}\|_2 \\ \|\mathbf{e}_{\mathbf{z},n-1}\|_2 \end{bmatrix},$$

such that the proof is completed by the application of Prop. 5.5 for the particular choice of waveforms above.  $\square$

Finally, combining all the previous results, the following theorem guarantees stability and from that the global convergence result is deduced by iterative application, [5].

**Theorem 5.10** (Stability). *Let an index-1 DAE (5.11) with Ass. 5.1, the constant extrapolation operator and consistent splitting functions (Ass. 5.3) be given. If the contractivity constant is bounded*

$$\alpha_m \leq \bar{\alpha} < 1 \quad \text{and} \quad L_\Phi \alpha_m^{k_m} \leq \bar{\alpha} \quad \text{for } 0 \leq m \leq n,$$

and the approximation is close to the solution, i.e.,

$$\|\boldsymbol{\epsilon}_{\mathbf{y},m}\|_2 + \|\boldsymbol{\epsilon}_{\mathbf{z},m}\|_2 \leq d \quad \text{for } 0 \leq m < n,$$

then there is a constant  $C^* > 0$  (independent of the window number  $n$  and of the window sizes  $H_m$ ) such that the global error on the  $n$ -th time window satisfies

$$\|\boldsymbol{\epsilon}_{\mathbf{y},n}\|_2 + \|\boldsymbol{\epsilon}_{\mathbf{z},n}\|_2 \leq C^* \max_{0 \leq m < n} \delta_m^0 \leq d \quad (5.54)$$

for all window sizes  $0 < H_m < H_0$  small enough.

*Proof.* From equation (5.40) the following estimate is derived

$$\mu_m = \alpha_m + \frac{2CH_m}{\frac{\alpha_m}{2C} + \sqrt{H_m}} < \alpha_m + C_{\alpha,m} \sqrt{H_m} \quad (5.55)$$

where  $C_{\alpha,m}$  is a sufficiently large constant. Then the assumption  $L_\Phi \alpha_m^{k_m} \leq \bar{\alpha} < 1$  yields

$$\alpha_m^* = L_\Phi (\mu_m^{k_m} + (\mu_m - \alpha_m)^{k_m}) < L_\Phi \left( (\alpha_m + C_{\alpha,m} \sqrt{H_m})^{k_m} + (C_{\alpha,m} \sqrt{H_m})^{k_m} \right) < 1,$$

for  $H_m$  small enough and finally  $\alpha^* := \max_{0 \leq m \leq n} \alpha_m^* < 1$ . The Prop. 5.8 and Prop. 5.9 imply

$$\begin{pmatrix} \|\boldsymbol{\epsilon}_{\mathbf{y},n}\|_2 \\ \|\boldsymbol{\epsilon}_{\mathbf{z},n}\|_2 \end{pmatrix} \leq \begin{pmatrix} 1 + C_e^* H & C_e^* H \\ C_e^* & \alpha^* \end{pmatrix} \cdot \begin{pmatrix} \|\boldsymbol{\epsilon}_{\mathbf{y},n-1}\|_2 \\ \|\boldsymbol{\epsilon}_{\mathbf{z},n-1}\|_2 \end{pmatrix} + \begin{pmatrix} C_d^* H \delta_n^0 \\ C_d^* \delta_n^0 \end{pmatrix}.$$

This proves the first half of the inequality (5.54), while the second half is clear by the definition of the extrapolation operator, i.e., for the constant extrapolation  $\delta_m^0 = \mathcal{O}(H_m)$ . Thus the extrapolation errors can be made arbitrarily small by the window size  $H_m$ .  $\square$

Finally Theorem 5.10 above is applied iteratively, [9, 5]. This guarantees that the approximating waveform remains in the neighborhood of the exact solution, which depends only on the sizes of the time windows  $H_n$ . This is the desired global convergence and stability result.

**Corollary 5.11** (Convergence and stability). *Let the same assumptions as for Theorem 5.10 be fulfilled. Then there is a constant  $C^*$ , such that*

$$\|\tilde{\mathbf{y}}|_{[0,t_e]} - \mathbf{y}|_{[0,t_e]}\|_2 + \|\tilde{\mathbf{z}}|_{[0,t_e]} - \mathbf{z}|_{[0,t_e]}\|_2 \leq C^* \cdot \max_{0 \leq m < N} \delta_m^0,$$

where  $\delta_m^0$  is the extrapolation error on the  $m$ -th window.

Let us conclude that the only important constraint for stability and convergence is the hypothesis (5.45), i.e., the algebraic-to-algebraic coupling, [5, 7]

$$\|\mathbf{G}_{\mathbf{z}^{(k)}}^{-1} \mathbf{G}_{\mathbf{z}^{(k-1)}}\|_{2,\infty} < 1,$$

while the dependence on the window size is a natural condition that cannot be circumvented. In the following Sections 5.3.5 and 5.3.6 cosimulations of applications from electrical engineering are discussed; the focus is especially on the hypothesis above and how this coupling can be avoided in practice.

### 5.3.5 Application to Field/Circuit Coupling

This section deals with the application of dynamic iteration methods to the coupled problem of circuits described in terms of the MNA and field problems given by MQS devices, (3.19). The time-transient monolithic simulation of this coupled problem has been an established technique for several years, [133] and higher order methods have been studied, e.g. [15]. On the other hand the weak coupling (or cosimulation) where the circuit and field subproblems are kept separately is also well-known: for example the coupling via current and voltage sources is proposed in [14, 113]. A more advanced cosimulation strategy is for example studied by [80, 143], where inductances are extracted from the finite element model and inserted into the circuit subsystem. This idea is further enhanced in [85, 86] by using a multirate approach, i.e., different time steps in each subproblem. Nonetheless all the approaches have in common that they solve the subsystems only once per time window (i.e., one iteration  $k_n = 1$ ) and then proceed to the next window. This makes those cosimulation approaches very dependent on the particular choice of the window size. Furthermore

there is no error control and convergence analysis. Thus this treatise features the dynamic iteration approach as introduced in Section 5.3.2 and analyzed in Sections 5.3.3-5.3.4.

The DAE-index of the subsystems and the coupling interface are the crucial points for the convergence of a dynamic iteration scheme. The hypothesis (5.45) identified the exchange of algebraic variables as the mathematical reason for divergence of the iteration scheme. Thus an interface is required that circumvents those problems. Let us recapitulate the coupled field/circuit system in order to derive an adequate interface. It reads for the flux/charge oriented MNA and excited by stranded conductors, see Section 3.2

$$\begin{aligned}
 \mathbf{A}_C \frac{d}{dt} \mathbf{q} + \mathbf{A}_R \mathbf{g}_R(\mathbf{A}_R^\top \mathbf{u}, t) + \mathbf{A}_L \mathbf{i}_L + \mathbf{A}_V \mathbf{i}_V + \mathbf{A}_I \mathbf{i}_s(t) + \mathbf{A}_M \mathbf{i}_M &= 0, \\
 \mathbf{q} - \mathbf{q}_C(\mathbf{A}_C^\top \mathbf{u}, t) &= 0, \\
 \frac{d}{dt} \boldsymbol{\phi} - \mathbf{A}_L^\top \mathbf{u} &= 0, \\
 \boldsymbol{\phi} - \boldsymbol{\phi}_L(\mathbf{i}_L, t) &= 0, \\
 \mathbf{A}_V^\top \mathbf{u} - \mathbf{v}_s(t) &= 0
 \end{aligned} \tag{5.56a}$$

with the coupling interface

$$\frac{d}{dt} \boldsymbol{\phi}_M + \mathbf{R}_M \mathbf{i}_M - \mathbf{v}_M = 0 \tag{5.56b}$$

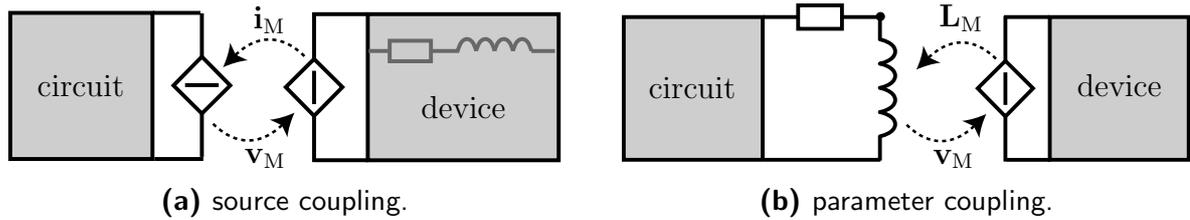
and the curl-curl equation

$$\mathbf{M}_\sigma \frac{d}{dt} \widehat{\mathbf{a}} + \mathbf{k}_\nu(\widehat{\mathbf{a}}) \widehat{\mathbf{a}} - \mathbf{X}_M \mathbf{i}_M = 0 \tag{5.56c}$$

where  $\widehat{\mathbf{a}}$  denotes the line-integrated magnetic vector potentials,  $\boldsymbol{\phi}_M := \mathbf{X}_M^\top \widehat{\mathbf{a}}$  is introduced only for the matter of notation and  $\mathbf{v}_M = \mathbf{A}_M^\top \mathbf{u}$  are the voltage drops applied to the coils, cf. equation (3.7). For the circuit unknowns we refer to Section 3.1. When the field and circuit subproblems are solved separately, the field system (5.56c) above must be fed by an input signal from the circuit system (5.56a) and vice versa. This may either be the current  $\mathbf{i}_M$  or the voltage drop  $\mathbf{v}_M$ . These inputs can be interpreted as surrogate models (i.e., time-dependent sources) of the circuit subproblem for the duration of a time window. Analogously an output of the field system (5.56) that enters the circuit subsystem must be defined.

The dynamic of the MQS device (especially its nonlinearity) varies rather slowly varying in time (cf. Section 5.2.1) and thus typically the circuit is discretized using smaller time steps than the MQS device. Consequently it is advantageous to use a high quality surrogate model as a representation of the MQS device, because it will be the object of multiple evaluations in the time-integration of the circuit. In fact, the better the quality of this model, the larger the time windows can be chosen, lesser communication is needed and the more efficient is the method.

**Definition 5.7** (Field/circuit coupling interfaces). Let the MQS device be excited by a given voltage drop. Depending on the splitting scheme, this is for example  $\mathbf{v}_M(t) = \mathbf{A}_M^\top \mathbf{u}^{(k-1)}(t)$  with  $t \in [T_n, T_{n+1}]$ , where  $\mathbf{u}^{(k-1)}$  are the waveforms of node potentials computed in the previous iteration ( $k-1$ ) by solving the circuit subproblem. Then the following surrogate models of the field device are defined:



**Figure 5.8:** Field/circuit coupling interfaces. In interface (a) the coupling to the circuit is given by the current through the device and in model (b) by an extracted inductance, [9].

- (a) **Source coupling.** For the given voltage  $\mathbf{v}_M$  the system (5.56) is used to compute a new waveform for the current  $\mathbf{i}_M^{(k)}$ . Then it is inserted as a time-dependent current source into the network equations (5.56a), see Fig. 5.8a. This interface corresponds to the model in [14]. The circuit input on the  $n$ -th time window is given by

$$\mathbf{i}_s(t) := \mathbf{i}_M^{(k)}(t), \quad \text{for } t \in [T_n, T_{n+1}]$$

and  $\mathbf{A}_I = \mathbf{A}_M$ , where it is assumed for simplicity of notation that the current through the MQS device is the only current source in the circuit.

- (b) **Parameter coupling.** For the given voltage  $\mathbf{v}_M$  the system (5.56) is used to compute the saturation level, i.e., the vector potential  $\bar{\mathbf{a}}^{(k)}$ . Then a lumped parameter model, e.g., a time-dependent inductance matrix and a constant resistance are extracted, see, Fig. 5.8b. For the  $n$ -th time window, i.e.,  $t \in [T_n, T_{n+1}]$ , it holds

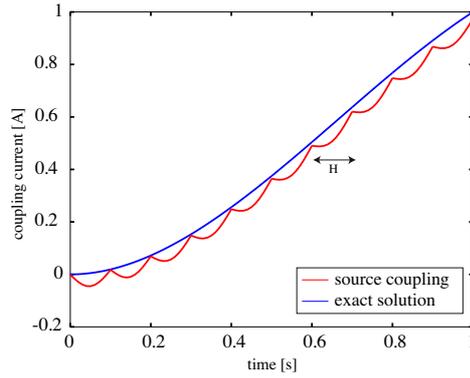
$$\begin{aligned} \mathbf{g}_M &:= \mathbf{R}_M^{-1} \mathbf{A}_M \mathbf{u} & \text{with } \mathbf{R}_M &:= \mathbf{X}_M^\top \mathbf{M}_\sigma^{-1} \mathbf{X}_M, \\ \boldsymbol{\phi}_M &:= \mathbf{L}_M^{(k)}(t) \mathbf{i}_L + \boldsymbol{\phi}_{\text{eddy}}^{(k)}(t) & \text{with } \mathbf{L}_M^{(k)}(t) &:= \mathbf{X}_M^\top \mathbf{k}_\nu^{-1}(\bar{\mathbf{a}}^{(k)}(t)) \mathbf{X}_M \\ & & \text{and } \boldsymbol{\phi}_{\text{eddy}}^{(k)}(t) &:= \mathbf{L}_M^{(k)}(t) \mathbf{i}_M^{(k)}(t) - \mathbf{X}_M^\top \bar{\mathbf{a}}^{(k)}(t), \end{aligned} \quad (5.57)$$

with the incidence matrix  $\mathbf{A}_R = \mathbf{A}_L = \mathbf{A}_M$ , where it is assumed for simplicity of notation that the MQS device is the only inductive  $\boldsymbol{\phi}_L = \boldsymbol{\phi}_M$  and resistive element  $\mathbf{g}_R = \mathbf{g}_M$  in the circuit. The new characteristic equations above replace the corresponding definitions in the network equations (5.56a).

This model covers the inductive and eddy current effects due to the inductance  $\mathbf{L}_M$  and the flux  $\boldsymbol{\phi}_M$ . It generalizes the approaches in [143, 85].

- (c) **MOR coupling.** This is a generalization of (b), where  $\mathbf{L}_M$  may describe an arbitrary (linear) system obtained by Model Order Reduction (MOR): for the given voltage  $\mathbf{v}_M$  the system (5.56) is used to compute the MVP  $\bar{\mathbf{a}}^{(k)}(t)$ . Then a standard model order reduction can be applied to the system (5.56), e.g. proper orthogonal decomposition, [106]. It may utilize the waveform of the MVP on  $t \in [T_n, T_{n+1}]$ , for example such that the curl-curl term loses its nonlinear character, i.e.,  $\mathbf{k}_\nu^{-1}(\bar{\mathbf{a}}^{(k)}(t))$  depends only on time.

In practice the subproblems (field and circuit) are solved numerically by a single rate time-integrator, see Section 5.1. Hence the solution is not a continuous waveform, but a series of discrete solutions at time steps  $t_i$ . A (continuous) waveform can be reconstructed from them for example using dense-output or spline interpolation, [69]. Obviously, the



**Figure 5.9:** Qualitative behavior (‘festoos-like’) of the waveforms when using the source coupling (a) in Definition 5.7. Window size  $H = 10^{-2}$ .

interpolation must be of sufficient quality (order), so that after exchanging the waveforms the other time-integrator can benefit from it.

In a monolithic simulation, the interfaces (a) and (b) in Definition 5.7 are equivalent: one can interpret (5.57) as the Schur complement of the system (5.56). In fact, this is another way of looking at the bypassing approach in Section 5.2. On the contrary approach (c) will introduce a modeling error depending on the reduction technique, [106]. For the weak coupling by a dynamic iteration scheme, the models behave differently: the source approach (a) is only a black-box coupling. Additional evaluations of the source model at time points  $t_i \in (T_n, T_{n+1})$  do not reflect the physical behavior, see Fig. 5.9. On the other hand the lumped parameter approach (b) still models the inductive effect (Faraday’s Law) correctly. The drawback is the additional computational cost: the inductance matrices  $\mathbf{L}_M$  must be computed for each time step. This drawback applies even more to general MOR approaches. Nonetheless these approaches pay off in practice due to the better decoupling.

**Remark 5.6** (Interface and DAE-index of the MQS device, [117]). It has been shown in Example 5.1 that the design of the coupling interface may change the DAE-index. For the field/circuit case the models in Definition 5.7 guarantee that the MQS device-subsystem remains an index-1 problem. On the other hand the current-driven case, e.g. where equation (5.56c) is fed by a given current and the voltage  $\mathbf{v}_M$  is computed and supplied to the network equations (3.1e) corresponds to an index-2 system (Theorem 4.5) and thus the dynamic iteration theory of Section 5.3.2 is not applicable, [117].

### Abstract Coupling Analysis

For the circuit system the standard loop and cutset conditions from Theorem 4.2 are assumed and thus the circuit system (5.56a) is index-1. Moreover, the coupled system is index-1 (another application of Theorem 4.2). Then the following problem description is derived, where the MQS device subsystem is abstractly addressed by the functions with subscript M and the circuit subsystem by subscript C (the structure below is the same for all interfaces in Definition 5.7)

$$\begin{aligned} \dot{\mathbf{y}}_M &= \mathbf{f}_M(\mathbf{z}_M, \mathbf{z}_C), & \dot{\mathbf{y}}_C &= \mathbf{f}_C(\mathbf{y}_C, \mathbf{z}_C), \\ \mathbf{0} &= \mathbf{g}_M(\mathbf{y}_M, \mathbf{z}_M), & \mathbf{0} &= \mathbf{g}_C(\mathbf{y}_C, \mathbf{z}_C, \mathbf{z}_M), \end{aligned} \quad (5.58)$$

with regular  $\partial \mathbf{g}_M / \partial \mathbf{z}_M$  and  $\partial \mathbf{g}_C / \partial \mathbf{z}_C$ . The variables of the field and circuit equations are

$$\mathbf{y}_M := \begin{bmatrix} \mathbf{P}_\sigma \hat{\mathbf{a}} \\ \phi_M \end{bmatrix}, \quad \mathbf{z}_M := \begin{bmatrix} \mathbf{Q}_\sigma \hat{\mathbf{a}} \\ \mathbf{i}_M \\ \mathbf{v}_M \\ \mathbf{L}_M \end{bmatrix} \quad \text{and} \quad \mathbf{y}_C := \begin{bmatrix} \mathbf{q} \\ \phi \end{bmatrix}, \quad \mathbf{z}_C := \begin{bmatrix} \mathbf{u} \\ \mathbf{i}_L \\ \mathbf{i}_V \end{bmatrix}$$

where  $\mathbf{y}$  denotes differential (i.e., defined by differential equations) and  $\mathbf{z}$  algebraic components (i.e., only defined by algebraic constraints). Similar to flux/charge oriented MNA, the magnetic  $\phi_M := \mathbf{X}_M \hat{\mathbf{a}}$  is typically not computed as an explicit unknown of the system. Projector  $\mathbf{P}_\sigma$  picks out the differential part of the magnetic vector potential, i.e., the components defined in conductive materials, see Definition 2.8. A detailed derivation of structure (5.58) is given in [117].

Now, having defined the subproblems and identified the coupling variables, a splitting scheme must be chosen that defines the computational sequence of the waveforms. From its structure a convergence guarantee can be deduced (depending on hypothesis (5.45)). Following [118, 9] a Gauß-Seidel-type dynamic iteration scheme is applied to (5.58), so the newest data available is always exploited. Let us start with the computation of the field subproblem. The splitting scheme reads

$$\mathbf{F}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) := \begin{bmatrix} \mathbf{f}_M(\mathbf{z}_M^{(k)}, \mathbf{z}_C^{(k-1)}) \\ \mathbf{f}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}) \end{bmatrix}, \quad (5.59a)$$

$$\mathbf{G}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) := \begin{bmatrix} \mathbf{g}_M(\mathbf{y}_M^{(k)}, \mathbf{z}_M^{(k)}) \\ \mathbf{g}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}, \mathbf{z}_M^{(k)}) \end{bmatrix}, \quad (5.59b)$$

with  $\mathbf{y}^\top := [\mathbf{y}_M^\top \ \mathbf{y}_C^\top]$  and  $\mathbf{z}^\top := [\mathbf{z}_M^\top \ \mathbf{z}_C^\top]$ . The only old iterate  $\mathbf{z}_C^{(k-1)}$ , i.e., the voltage drop defined by the circuit, enters a differential equation via the function  $\mathbf{f}_M$ . Thus the contraction factor  $\alpha$  vanishes, see Cor. 5.6 (iii).

**Theorem 5.12** (Convergence of field/circuit cosimulation). *Let the assumptions of Theorem 5.10 be given. Then the Gauß-Seidel-type dynamic iteration of the field (5.56) and circuit subsystems (3.1) coupled by one of the interfaces from Definition 5.7 is unconditionally stable and convergent with a window-wise convergence rate  $\mathcal{O}(H_n)$ , if the iteration starts with the computation of the MQS device.*

*Proof.* Application of Cor. 5.6 to the splitting functions (5.59). □

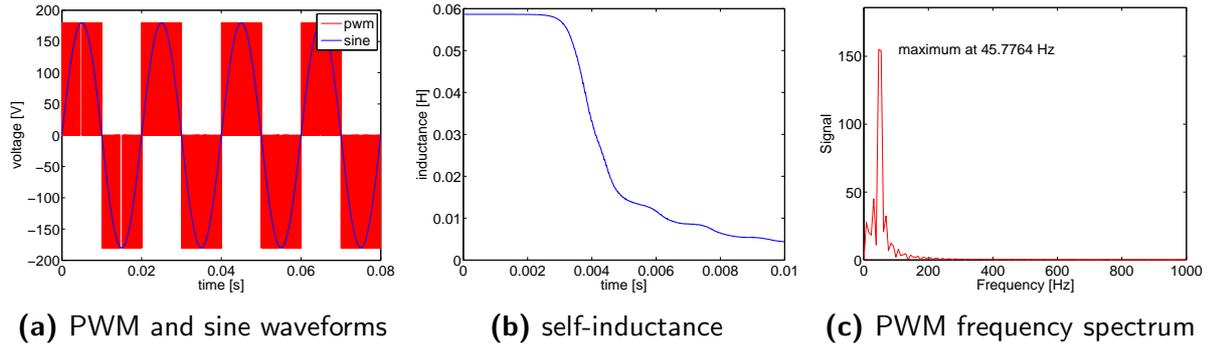
A reordering of the computational sequence creates a mutual algebraic dependence. In contrast to the previous scheme (5.59) the Gauß-Seidel scheme

$$\tilde{\mathbf{F}}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) := \begin{bmatrix} \mathbf{f}_M(\mathbf{z}_M^{(k)}, \mathbf{z}_C^{(k)}) \\ \mathbf{f}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}) \end{bmatrix}, \quad (5.60a)$$

$$\tilde{\mathbf{G}}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) := \begin{bmatrix} \mathbf{g}_M(\mathbf{y}_M^{(k)}, \mathbf{z}_M^{(k)}) \\ \mathbf{g}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}, \mathbf{z}_M^{(k-1)}) \end{bmatrix}, \quad (5.60b)$$

does not fulfill the hypothesis (5.45) trivially, i.e.,

$$\left\| \mathbf{G}_{\mathbf{z}^{(k)}}^{-1} \mathbf{G}_{\mathbf{z}^{(k-1)}} \right\|_{2, \infty} > 0.$$



**Figure 5.10:** Applied PWM voltage and its frequency spectrum.

Thus the smallness of the contraction factor  $\alpha$  is not automatically fulfilled, see equation (5.28). Consequently convergence is not guaranteed by the theory described above and divergence might occur. Even if  $0 < \alpha < 1$  holds true in (5.60) the convergence properties of (5.58) would be better, see Cor. 5.6. Thus scheme (5.59) is employed in the following, see Listing 2 on Page 71.

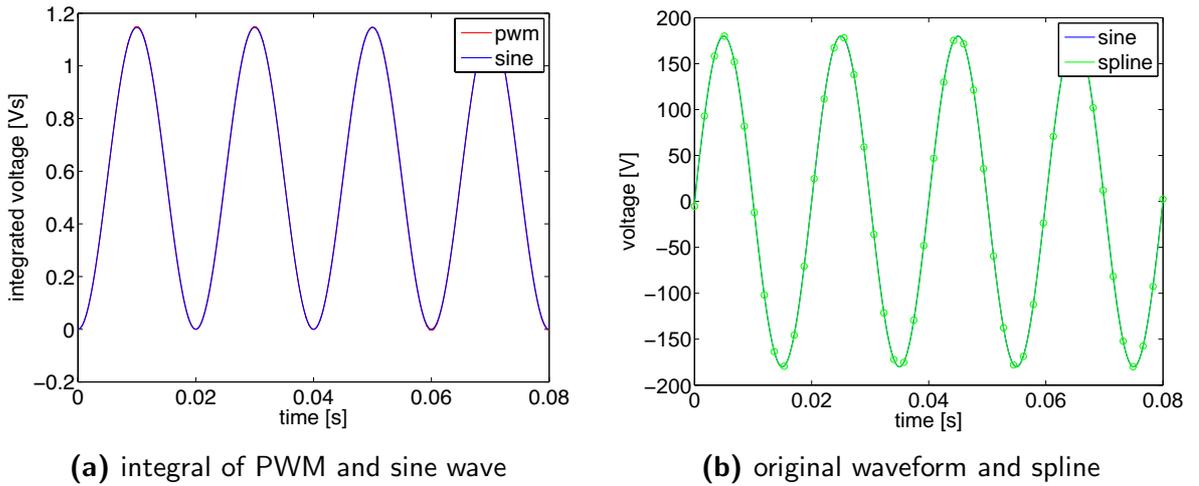
As previously mentioned, cosimulation is especially efficient if the adaptive time-integration of the subproblems can exploit different time scales.

### Multirate Phenomenon in Field/Circuit Coupling

The drawback of single rate time-integration is that it resolves the dynamics of a system as a whole, cf., Sections 5.1 and 5.2.1. Thus it yields a series of time steps that matches the dynamics of the most active component, i.e., the one working at the highest frequency. In coupled, multiphysical systems (e.g. electromagnetic problems with heating effects) one can split the equations corresponding to their time constants on the basis of physical reasoning. Let us consider a 2D model of an induction motor, [40]. Its rotor position can be updated using an additional ODE and, furthermore, changes in the conductivity due to the MQS device heating up can be modeled by another PDE. Here the separation into subproblems allows immediately for an efficient cosimulation exploiting the inherent time-constant of the subproblems. This approach is particularly convenient when different physical effects are simulated by different software tools.

In contrast to this, the field and circuit subproblems describe the same physical phenomena and one speaks of ‘refined modeling’, [11]. Hence the subproblems feature similar time constants. Nevertheless, due to switches or filters in the circuit there may only a subset of the devices active at any given time, while the others remain latent. Then cosimulation is more efficient than the single-rate approach, see the example in Section 6.3.1 where a low-pass filter causes dynamics at different rates.

If the circuit topology does not provide such a splitting, the time rates are not well separated in the coupling variables. Let us consider another example, where the voltage that is applied to the MQS device is a pulse-width-modulated (PWM) sine-wave switching at 20 kHz, see Fig. 5.10. The saturation, i.e., the nonlinearity of the inductance matrix, is characterized by the underlying, much slower sine wave, see Fig. 5.10b. Mathematically speaking the saturation of nonlinear materials in the PDE model depends on the energy supplied, see (5.10) in the section on bypassing Section 5.2. This is related to the time



**Figure 5.11:** Spline approximation of the integral waveform and its derivative.

integral of the applied voltage  $\mathbf{v}_M(t)$

$$\psi(t) = \int_{T_n}^t \mathbf{v}_M(s) ds \quad \text{with } t \in [T_n, T_{n+1}].$$

Consequently the relevant time rates of the nonlinear behavior is given by the time rates of the integral above, even if the voltage applied is a much faster switching signal. The waveform in Fig. 5.10a shows the PWM voltage while Fig. 5.11a depicts its integral with respect to time. The integral is a step function approximating the cosine at a frequency of 50Hz, see Fig. 5.10c. For high frequencies the approximation is very accurate (below the accuracy of the nonlinear curve), so one could use the smooth signal instead of the step function.

The same approach is valid if the signal (voltage) is composed of a fast and a slow part

$$\mathbf{v}_M(t) = \mathbf{v}_{\text{fast}}(t) + \mathbf{v}_{\text{slow}}(t).$$

Problems often exhibit a time rate of interest that is given by  $\mathbf{v}_{\text{slow}}$ , and either the amplitude of the fast signal  $\max |\psi_{\text{fast}}| \ll \max |\psi_{\text{slow}}|$  is negligible or the energy of the fast switching voltage  $\mathbf{v}_{\text{fast}}(t)$  evolves at a slower time rate. The impact of the fast signal on the nonlinear effects can be disregarded, similar to thermal coupling that takes effect only at a slow time rate due to energy transport, [47, 7].

### Signal Separation and Smoothing

As explained above, feeding the field model by the low frequency part of the signal (the sine wave) will have the same effect on the nonlinearity as the fast waveform (PWM), but time-integration will need fewer time steps. On the other hand the low-frequency part of the waveform, i.e., the sinusoidal voltage in the example above, is typically not explicitly given because other circuit elements will have an impact on the signal too. Various methods can be used to extract the low-frequency part of the waveform, we propose here a simple method based on spline interpolation of the integrated voltage.

The known waveform  $\mathbf{v}_{\text{fast}}(t)$  is integrated in time on the interval  $\mathcal{I}_n := [T_n, T_{n+1}]$ , which corresponds to a summation of the discrete solutions  $\mathbf{v}_M(t_i)$  ( $t_i \in \mathcal{I}_n$ ) multiplied by the corresponding step sizes  $h_i$

$$\psi_n := \sum_{i=1}^n \mathbf{v}_M(t_i) h_i. \quad (5.61)$$

Now we shall define a cubic spline interpolation  $\tilde{\psi}(t)$  of  $\psi(t)$  that gives a smooth time-integrated voltage. The interpolation knots should be chosen accordingly to the dynamics of the low-frequency part of the waveform. For example the number of knots can easily be estimated by a Fourier analysis or might be known beforehand. In the example Fig. 5.10b) 20 uniformly distributed knots per period were chosen, and these yield a satisfactory approximation of the sinusoidal waveform. It is crucial that the spline is a good approximation of (5.61), because otherwise the energy balance will be violated.

Finally the cubic polynomials of the spline interpolation are differentiated piecewise with respect to time yielding a slowly varying spline approximation  $\tilde{\mathbf{v}}_{\text{slow}}(t)$  to the low-frequency part of the waveform. The adaptive time-integrator will require only a few time steps for the smooth slowly-varying signal compared to the hundreds of steps that are necessary to sample the original, non-smooth and fast switching PWM signal.

This smoothing requires only a small change in Listing 2: the integration, resampling and derivation of the spline for the voltage excitation must be done in Step 2d. This approach can reduce the computational effort significantly, but it introduces a model error that relates to the frequency of the pulsed signal and the accuracy of the spline interpolation.

## Conclusions

In this section the stability and convergence of the field/circuit cosimulation were mathematically analyzed, following [118, 9]. It was shown that different coupling interfaces are feasible, but that the parameter coupling is superior because it reflects the underlying laws of physics and exploits multirate behavior (located in the nonlinearity of the MQS device). This is documented by numerical examples in Sections 6.3 and 6.3.1.

---

**Listing 2** Multirate Dynamic Iteration Algorithm (Gauß-Seidel Type), cf. [118]

---

0) **Initialization.** Set first time window size  $H_0$ , window counter  $n := 0$ , sweep counter  $k := 0$  and initial values  $\phi_0$ ,  $\mathbf{q}_0$  and  $\widehat{\mathbf{a}}_0$ . Go to Step 1).

1) **Guess.** Set window  $T_{n+1} = T_n + H_n$  and extrapolate the circuit solution

$$\mathbf{v}_M^{(0)}(t) := \mathbf{A}_M^\top \mathbf{u}^{(k)}(T_n) \quad \text{for } t \in [T_n, T_{n+1}],$$

where  $\mathbf{u}^{(k)}$  is the waveform of the node potentials from the last iteration on the previous window. Set  $k := 1$  and start dynamic iteration, i.e., go to Step 2).

2) **Solve the DAE initial value problems.**

a) Adaptive time-integration of the MQS device IVP on time window  $t \in [T_n, T_{n+1}]$  with initial value  $\widehat{\mathbf{a}}_0$

$$\begin{aligned} \mathbf{M}_{\bar{\sigma}} \frac{d}{dt} \widehat{\mathbf{a}}^{(k)} + \mathbf{k}_\nu(\widehat{\mathbf{a}}^{(k)}) \widehat{\mathbf{a}}^{(k)} &= \mathbf{X}_M \mathbf{i}_M^{(k)}, \\ \frac{d}{dt} \mathbf{X}_M^\top \widehat{\mathbf{a}}^{(k)} + \mathbf{R}_M \mathbf{i}_M^{(k)} &= \boxed{\mathbf{v}_M^{(k-1)}}. \end{aligned}$$

b) Computing the reduced order model, see equation (5.57)

$$\phi_M(\mathbf{i}_L, t) := \mathbf{L}_M^{(k)}(t) \mathbf{i}_L + \phi_{\text{eddy}}^{(k)}(t).$$

c) Adaptive time-integration of the circuit IVP on time window  $t \in [T_n, T_{n+1}]$  with initial values  $\phi_0$  and  $\mathbf{q}_0$

$$\begin{aligned} \mathbf{A}_C \frac{d}{dt} \mathbf{q}^{(k)} + \mathbf{A}_M \mathbf{R}_M^{-1} \mathbf{A}_M^\top \mathbf{u}^{(k)} + \mathbf{A}_V \mathbf{i}_V^{(k)} + \mathbf{A}_I \mathbf{i}_s^{(k)}(t) + \mathbf{A}_M \mathbf{i}_L^{(k)} &= 0, \\ \frac{d}{dt} \phi^{(k)} - \mathbf{A}_M^\top \mathbf{u}^{(k)} = 0, \quad \phi^{(k)} - \boxed{\phi_M(\mathbf{i}_L^{(k)}, t)} &= 0, \\ \mathbf{q} - \mathbf{q}_C(\mathbf{A}_C^\top \mathbf{u}^{(k)}, t) = 0, \quad \mathbf{A}_V^\top \mathbf{u}^{(k)} - \mathbf{v}_s(t) &= 0. \end{aligned}$$

d) Compute the field excitation from the node-potentials and go to Step 3)

$$\mathbf{v}_M^{(k)}(t) := \mathbf{A}_M^\top \mathbf{u}^{(k)}(t).$$

3) **Sweep control.** If e.g.  $\text{norm}(\mathbf{i}_L^{(k)} - \mathbf{i}_M^{(k)}) > \text{tol}$ , then repeat step, i.e., set  $k := k + 1$  and go to Step 2), else goto Step 4)

4) **Next window.** If  $T_{n+1} \geq t_e$  then go to Step 5), else set new initial values

$$\phi_0 := \phi^{(k)}(T_{n+1}), \quad \mathbf{q}_0 := \mathbf{q}^{(k)}(T_{n+1}) \quad \text{and} \quad \widehat{\mathbf{a}}_0 := \widehat{\mathbf{a}}^{(k)}(T_{n+1}).$$

Determine the new window size  $H_{n+1}$ , e.g. from the step size predictor of the time-integrators in Step 2a) and 2b). Go to Step 1) with  $n := n + 1$ .

5) **Stop.**

---

### 5.3.6 Application to Semiconductor/Circuit Coupling

Similarly to the previous section, this one deals with the application of dynamic iteration methods. Here the coupling of the semiconductor problem (mathematical model from Section 3.3) and electric circuits (network model from Section 3.1) is described. The idea is based on [1, 9] using the previous works of [26] and [7]. Especially in the last mentioned treatise [7] the dynamic iteration of semiconductors and circuits was studied in a similar way, but without considering the error propagation from window to window, see Section 5.3.4.

In the following the focus is the derivation of a natural coupling interface ('parameter coupling'), such that the algebraic-to-algebraic coupling term vanishes, [9]. This will be compared with the classical source coupling, cf. Section 5.3.5. For this, the semiconductor/circuit problem is briefly recapitulated. It reads for the flux/charge oriented MNA coupled to a single semiconductor device, see (3.19)

$$\begin{aligned}
 \mathbf{A}_C \frac{d}{dt} \mathbf{q} + \mathbf{A}_R \mathbf{g}_R(\mathbf{A}_R^\top \mathbf{u}, t) + \mathbf{A}_L \mathbf{i}_L + \mathbf{A}_V \mathbf{i}_V + \mathbf{A}_I \mathbf{i}_s(t) + \mathbf{A}_D \mathbf{i}_D &= 0, \\
 \mathbf{q} - \mathbf{q}_C(\mathbf{A}_C^\top \mathbf{u}, t) &= 0, \\
 \frac{d}{dt} \boldsymbol{\phi} - \mathbf{A}_L^\top \mathbf{u} &= 0, \\
 \boldsymbol{\phi} - \boldsymbol{\phi}_L(\mathbf{i}_L, t) &= 0, \\
 \mathbf{A}_V^\top \mathbf{u} - \mathbf{v}_s(t) &= 0,
 \end{aligned} \tag{5.62a}$$

with the coupling interface

$$\mathbf{i}_D = \mathbf{C}_D \frac{d}{dt} \mathbf{v}_D - \mathbf{i}_{SD} \quad \text{with} \quad \mathbf{i}_{SD} := \mathbf{j}_{SD}(\mathbf{n}, \mathbf{p}, \boldsymbol{\Phi}) \tag{5.62b}$$

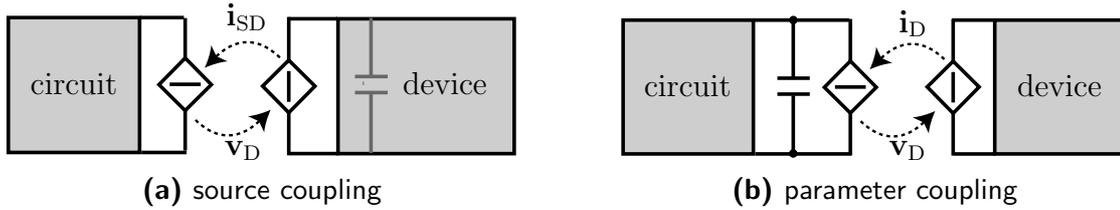
and the drift-diffusion equations

$$\begin{aligned}
 \mathbf{M}_n \frac{d}{dt} \mathbf{n} + \mathbf{K}_n \mathbf{n} - \mathbf{r}_n(\mathbf{p}, \boldsymbol{\Phi}) &= 0, \\
 \tilde{\mathbf{S}} \mathbf{M}_e \tilde{\mathbf{S}}^\top \boldsymbol{\Phi} - \mathbf{q}(\mathbf{n}, \mathbf{p}, \mathbf{v}_D) &= 0, \\
 \mathbf{M}_p \frac{d}{dt} \mathbf{p} + \mathbf{K}_p \mathbf{p} - \mathbf{r}_p(\mathbf{n}, \boldsymbol{\Phi}) &= 0,
 \end{aligned} \tag{5.62c}$$

where the unknowns  $\boldsymbol{\Phi}, \mathbf{n}, \mathbf{p}$  denote the electric scalar potential, the electron and hole densities, see (5.62c);  $\mathbf{i}_D$  is the discrete current through the device and  $\mathbf{v}_D = \mathbf{A}_D^\top \mathbf{u}$  denotes the applied voltage drop. For the circuit unknowns we refer to Section 3.1. This structure allows for different interfaces, similarly to the field/circuit case in Definition 5.7:

**Definition 5.8** (Semiconductor/circuit coupling interfaces). Let the semiconductor be excited by a given voltage drop. Depending on the splitting scheme, this is for example  $\mathbf{v}_D(t) = \mathbf{A}_D^\top \mathbf{u}^{(k-1)}(t)$  with  $t \in [T_n, T_{n+1}]$ , where  $\mathbf{u}^{(k-1)}$  are the waveforms of node potentials computed in the previous iteration ( $k-1$ ) by solving the circuit subproblem. Then the following surrogate models of the semiconductor device are defined

- (a) **Source coupling.** For the given voltage  $\mathbf{v}_D$  the system (5.62c) is used to compute a new waveform for the current  $\mathbf{i}_D^{(k)}$ . Then it is inserted as a time-dependent current source into the network equations (5.62a), see Fig. 5.12a. This interface corresponds



**Figure 5.12:** Semiconductor/circuit interfaces. (a) the displacement current is modeled in the PDE device (semiconductor) and (b) given by a parallel lumped capacitance in the network model (circuit), see [1, 9].

to the model by [7]. The waveform on the  $n$ -th time window is given by

$$\mathbf{i}_s(t) := \mathbf{i}_D^{(k)}(t), \quad \text{for } t \in [T_n, T_{n+1}]$$

and  $\mathbf{A}_I = \mathbf{A}_D$ , where it is assumed for simplicity of notation that the current through the semiconductor is the only current source in the circuit.

- (b) **Parameter coupling.** The capacitance  $\mathbf{C}_D$  can be extracted from the semiconductor model beforehand. From the given voltage  $\mathbf{v}_D$  the current  $\mathbf{i}_{SD}$  is computed, i.e., the one without the displacement current, see equation (5.62b) and Fig. 5.12b. For the  $n$ -th time window, i.e.,  $t \in [T_n, T_{n+1}]$ , it holds

$$\mathbf{A}_C \frac{d}{dt} \mathbf{q} := \mathbf{A}_D \mathbf{C}_D \mathbf{A}_D^\top \frac{d}{dt} \mathbf{u} \quad \text{and} \quad \mathbf{i}_s(t) := \mathbf{i}_{SD}^{(k)}(t) \quad (5.63)$$

with the incidence matrix  $\mathbf{A}_I = \mathbf{A}_C = \mathbf{A}_D$ , where it is assumed for simplicity of notation that the semiconductor is the only capacitance and current source in the circuit. The new characteristic equations above replace the corresponding definitions in the network equations (5.62a). Please note that (5.63) uses the traditional MNA for the constant capacitance  $\mathbf{C}_D$ . This does not compromise the conservation laws because the charge is given here by a linear relation, [57].

**Remark 5.7** (Parameter coupling for semiconductors). The parameter coupling for semiconductors in (5.63) uses additional information on the model for the capacitive effect, but it is still a source coupling (it directly utilizes the current  $\mathbf{i}_{SD}$ ). The parameter approach can be further improved by using the nonlinear lumped conductance

$$\mathbf{g}_D(\mathbf{A}_D^\top \mathbf{u}, t) := \frac{\mathbf{i}_{SD}^{(k)}(t)}{\mathbf{v}_D^{(k)}(t)} \mathbf{A}_D^\top \mathbf{u} \quad \text{for } t \in [T_n, T_{n+1}]$$

instead of the plain current  $\mathbf{i}_{SD}^{(k)}$ . The equation above is given for the special case of a diode, i.e., scalar voltage and current.

In the case of a monolithic coupling, the interfaces (a)-(b) are equivalent, [9]. On the other hand for a weak coupling by a dynamic iteration scheme, they behave differently, cf. Definition 5.7. The additional lumped device in the interface (b), i.e., the capacitance  $\mathbf{C}_D$ , is a simple compact model for the displacement current of the diode. It reflects the physical law correctly, even though in lumped form. This idea is similarly used in [51],

where a capacitance and an exponential resistance model are exploited as a preconditioner for dynamic iteration.

Alternatively the parallel capacitance  $\mathbf{C}_D$  can be interpreted as a special case of an overlapping technique, e.g., [59]: the capacitive effect is computed in both, the device and circuit subsystems, but it is removed in the coupling interface. If the coupling current in interface (b) is defined as  $\mathbf{i}_{SD} := \mathbf{i}_D - \mathbf{C}_D \frac{d}{dt} \mathbf{v}_D$ . One can use here an arbitrary capacitance  $\mathbf{C}_D$ , as long as it is correctly added to the circuit subsystem, see equation (5.63).

In the following section the two coupling interfaces are analyzed in the framework of dynamic iteration, i.e., Section 5.3.4.

### Abstract Coupling Analysis

It is assumed that the circuit system fulfills the standard loop and cutset conditions from Theorem 4.2 for both coupling interfaces, i.e., neither the current sources  $\mathbf{i}_D$  and  $\mathbf{i}_{SD}$  introduce  $LI$ -cutsets nor the capacitance  $\mathbf{C}_D$  creates a  $CV$ -loop, see Section 4. For the semiconductor subsystem (5.62c) a prescribed voltage drop  $\mathbf{v}_D = \mathbf{A}_D^T \mathbf{u}$  also yields an subsystem index-1 DAE, [26]. Moreover, it can be assured that the coupled system is monolithic index-1, [120]. This allows the application of the dynamic iteration theory above and the following analysis is eligible, [1, 9].

**Source coupling.** The semiconductor/circuit subsystems, i.e., (5.62a) and (5.62c), with interface (b) in Definition 5.8 have the following semi-explicit form

$$\begin{aligned} \dot{\mathbf{y}}_D &= \mathbf{f}_D(\mathbf{y}_D, \mathbf{z}_D), & \dot{\mathbf{y}}_C &= \mathbf{f}_C(\mathbf{y}_C, \mathbf{z}_C, \mathbf{z}_D), \\ \mathbf{0} &= \mathbf{g}_D(\mathbf{y}_D, \mathbf{z}_D, \mathbf{z}_C), & \mathbf{0} &= \mathbf{g}_C(\mathbf{y}_C, \mathbf{z}_C, \mathbf{z}_D) \end{aligned} \quad (5.64)$$

where the circuit subsystem is denoted by subscript C and the semiconductor subsystem by D. The partial derivatives  $\partial \mathbf{g}_D / \partial \mathbf{z}_D$  and  $\partial \mathbf{g}_C / \partial \mathbf{z}_C$  are regular due to the index-1 assumptions above. The differential and algebraic components of the subsystems are given by

$$\mathbf{y}_D := \begin{bmatrix} \mathbf{n} \\ \mathbf{p} \end{bmatrix}, \quad \mathbf{z}_D := \begin{bmatrix} \Phi \\ \mathbf{i}_D \end{bmatrix} \quad \text{and} \quad \mathbf{y}_C := \begin{bmatrix} \mathbf{q} \\ \phi \end{bmatrix}, \quad \mathbf{z}_C := \begin{bmatrix} \mathbf{u} \\ \mathbf{i}_L \\ \mathbf{i}_V \end{bmatrix},$$

where the space discrete electric scalar potential is denoted by the vector  $\Phi$  and the device current by  $\mathbf{i}_D$ , see the definitions in (5.62c).

The circuit is given in flux/charge oriented MNA and thus the node potentials  $\mathbf{u}$  and consequently the voltage drop  $\mathbf{v}_D = \mathbf{A}_D^T \mathbf{u}$  are algebraic variables of the circuit. The device subsystem depends only on the algebraic circuit variables  $\mathbf{z}_C$  in the algebraic equation  $\mathbf{g}_D$ . In turn the algebraic variable  $\mathbf{z}_C$  (that contains the device current) may enter the differential  $\mathbf{f}_C$  and/or the algebraic equations  $\mathbf{g}_C$  of the circuit subsystem (5.64) (that depends on the circuit topology), [1, 9]. In a nutshell: the algebraic equations of both subsystems depend on the algebraic variable of the other subsystem. As a consequence the contraction factor does not vanish, i.e.,  $\alpha > 0$ , independently of the computational sequence of the subsystems in the dynamic iteration scheme. For example the splitting

functions for the device-first Gauß-Seidel iteration are

$$\begin{aligned}\tilde{\mathbf{F}}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) &:= \begin{bmatrix} \mathbf{f}_D(\mathbf{y}_D^{(k)}, \mathbf{z}_D^{(k)}) \\ \mathbf{f}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}, \mathbf{z}_D^{(k)}) \end{bmatrix}, \\ \tilde{\mathbf{G}}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) &:= \begin{bmatrix} \mathbf{g}_D(\mathbf{y}_D^{(k)}, \mathbf{z}_D^{(k)}, \mathbf{z}_C^{(k-1)}) \\ \mathbf{g}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}, \mathbf{z}_D^{(k)}) \end{bmatrix},\end{aligned}\tag{5.65}$$

with  $\mathbf{y}^\top := [\mathbf{y}_D^\top \ \mathbf{y}_C^\top]$  and  $\mathbf{z}^\top := [\mathbf{z}_D^\top \ \mathbf{z}_C^\top]$ . The superscript  $(k)$  denotes the iteration number. It encodes the computational order. Due to the dependence of  $\mathbf{G}$  on an old algebraic iterate convergence cannot be guaranteed by structural analysis. The parameters of all devices and the circuit topology will have a serious influence, see Section 6.3.2. Thus the following result is obtained, [1]

**Lemma 5.13** (Semiconductor/circuit cosimulation via sources). *Let the assumptions of Theorem 5.10 be given. Then the Gauß-Seidel-type dynamic iteration of the semiconductor (5.62) and circuit subsystems (5.62a) coupled by the interface (a) from Definition 5.8 is not convergent in general.*

*Proof.* From the splitting functions (5.65) it follows that the hypothesis (5.45) is not trivially fulfilled, i.e., convergence and stability are only guaranteed for cases where  $\alpha < 1$ .  $\square$

**Parameter coupling.** The interface (b) of Definition 5.8 replaces the current  $\mathbf{i}_D$  in the balance equation (3.1a) by the current  $\mathbf{i}_{SD}$  with a parallel capacitance  $\mathbf{C}_D$  in traditional MNA notation, i.e., by introducing some node potentials as differential unknowns:

$$\begin{aligned}\dot{\mathbf{y}}_D &= \mathbf{f}_D(\mathbf{y}_D, \mathbf{z}_D), & \dot{\mathbf{y}}_C &= \mathbf{f}_C(\mathbf{y}_C, \mathbf{z}_C, \mathbf{z}_D), \\ \mathbf{0} &= \mathbf{g}_D(\mathbf{y}_D, \mathbf{z}_D, \mathbf{y}_C), & \mathbf{0} &= \mathbf{g}_C(\mathbf{y}_C, \mathbf{z}_C),\end{aligned}\tag{5.66}$$

with the following variables for both subsystems

$$\mathbf{y}_D := \begin{bmatrix} \mathbf{n} \\ \mathbf{p} \end{bmatrix}, \quad \mathbf{z}_D := \begin{bmatrix} \Phi \\ \mathbf{i}_{SD} \end{bmatrix} \quad \text{and} \quad \mathbf{y}_C := \begin{bmatrix} \mathbf{q} \\ \phi \\ \mathbf{P}_D \mathbf{u} \end{bmatrix}, \quad \mathbf{z}_C := \begin{bmatrix} \mathbf{Q}_D \mathbf{u} \\ \mathbf{i}_L \mathbf{i}_V \end{bmatrix},$$

where  $\mathbf{Q}_D$  is a projector onto the kernel of  $\mathbf{A}_D^\top$  and  $\mathbf{P}_D = \mathbf{I} - \mathbf{Q}_D$  its complement. The projectors separate the differential from the algebraic components, i.e.,  $\mathbf{P}_D$  picks out the difference of the node potentials at the capacitance  $\mathbf{C}_D$  and  $\mathbf{Q}_D$  addresses the other potentials. Consequently the current  $\mathbf{i}_{SD}$  enters only in  $\mathbf{f}_C$  via  $\mathbf{z}_D$ , because it is related to a differential equation.

When starting, as before, with the semiconductor device computation, the splitting functions read (device first)

$$\begin{aligned}\mathbf{F}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) &:= \begin{bmatrix} \mathbf{f}_D(\mathbf{y}_D^{(k)}, \mathbf{z}_D^{(k)}) \\ \mathbf{f}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}, \mathbf{z}_D^{(k)}) \end{bmatrix}, \\ \mathbf{G}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) &:= \begin{bmatrix} \mathbf{g}_D(\mathbf{y}_D^{(k)}, \mathbf{z}_D^{(k)}, \mathbf{y}_C^{(k-1)}) \\ \mathbf{g}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}) \end{bmatrix}.\end{aligned}\tag{5.67}$$

The capacitive path between the coupling nodes ensures that the voltage drop  $\mathbf{v}_D$  is part of the differential variables  $\mathbf{y}_C$  and thus the only old iterate used is differential. Consequently the contraction factor  $\alpha$  vanishes for the splitting functions (5.67) and Cor. 5.6 (ii) guarantees a convergence rate of  $\mathcal{O}(\sqrt{H})$ .

On the other hand for the reversed computational order (circuit first)

$$\begin{aligned}\hat{\mathbf{F}}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) &:= \begin{bmatrix} \mathbf{f}_D(\mathbf{y}_D^{(k)}, \mathbf{z}_D^{(k)}) \\ \mathbf{f}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}, \mathbf{z}_D^{(k-1)}) \end{bmatrix}, \\ \hat{\mathbf{G}}(\mathbf{y}^{(k)}, \mathbf{y}^{(k-1)}, \mathbf{z}^{(k)}, \mathbf{z}^{(k-1)}) &:= \begin{bmatrix} \mathbf{g}_D(\mathbf{y}_D^{(k)}, \mathbf{z}_D^{(k)}, \mathbf{y}_C^{(k)}) \\ \mathbf{g}_C(\mathbf{y}_C^{(k)}, \mathbf{z}_C^{(k)}) \end{bmatrix},\end{aligned}\tag{5.68}$$

there is no dependence on old iterates in any algebraic equation, and thus Cor. 5.6 (iii) promises a higher convergence rate, i.e.,  $\mathcal{O}(H)$ , [9]. The alleged difference in the convergence rates is analyzed in the next section.

### Analysis of the Computational Sequence

In the previous section the Cor. 5.6 was used to prove the convergence of the dynamic iteration schemes (5.67) and (5.68). The criteria for convergence are easily verified if the splitting functions are known. But they are not necessary conditions: especially for subsystems that have less mutual dependencies than the general case, problem-specific estimates can improve the expected convergence results, [9]. This is demonstrated in the following for the splitting functions (5.67).

**Definition 5.9** (Lipschitz constants). Let Assumptions 5.1 and 5.2 be fulfilled and let the Lipschitz-continuous functions  $\zeta_C$  and  $\zeta_D$  define  $\mathbf{z}_C$  and  $\mathbf{z}_D$  by the implicit function theorem applied to  $\mathbf{g}_C$  and  $\mathbf{g}_D$  from (5.66), respectively. Then the following Lipschitz constants are defined

- let  $L$  denote the maximum of the Lipschitz constants of  $\mathbf{f}_\star$  and  $\zeta_\star$  w.r.t.  $\mathbf{y}_\star$  and  $\mathbf{z}_\star$  for  $\star \in \{D, C\}$
- let  $L_C$  denote the maximum of  $L$  and the Lipschitz constant of  $\mathbf{f}_C$  w.r.t.  $\mathbf{z}_D$
- let  $L_D$  denote the maximum of  $L$  and the Lipschitz constant of  $\zeta_D$  w.r.t.  $\mathbf{y}_C$ .

The latter Lipschitz constants  $L_C$  and  $L_D$  are a measure of the strength of the mutual coupling between the semiconductor and circuit subsystems. Following [9], a problem-specific version of the recursion estimate Lemma 5.4 is found

**Lemma 5.14** (Refined recursion estimate). *Let the assumptions of Lemma 5.4 be fulfilled. Then for the Gauß-Seidel-type dynamic iteration scheme (5.67) the recursion estimate*

$$\begin{bmatrix} \delta_{\mathbf{y},n}^k \\ \delta_{\mathbf{z},n}^k \end{bmatrix} \leq \underbrace{\begin{bmatrix} C_D H_n & 0 \\ C & 0 \end{bmatrix}}_{=: \mathbf{K}_D} \begin{bmatrix} \delta_{\mathbf{y},n}^{k-1} \\ \delta_{\mathbf{z},n}^{k-1} \end{bmatrix} + \begin{bmatrix} 1 + C H_n \\ C \end{bmatrix} \|\Delta_{\mathbf{y},n}^0(T_n)\|_2\tag{5.69}$$

holds true and similarly for the reversed order scheme (5.68)

$$\begin{bmatrix} \delta_{\mathbf{y},n}^k \\ \delta_{\mathbf{z},n}^k \end{bmatrix} \leq \underbrace{\begin{bmatrix} 0 & CH_n \\ 0 & C_C H_n \end{bmatrix}}_{=: \mathbf{K}_C} \begin{bmatrix} \delta_{\mathbf{y},n}^{k-1} \\ \delta_{\mathbf{z},n}^{k-1} \end{bmatrix} + \begin{bmatrix} 1 + CH_n \\ C \end{bmatrix} \|\Delta_{\mathbf{y},n}^0(T_n)\|_2 \quad (5.70)$$

with the constants

$$C_D := \left| \frac{L_C L_D}{1 - L(1 + L_C)H_{\max}} \right|, \quad \text{and} \quad C_C := \left| \frac{L_C L_D}{1 - L(1 + L_D)H_{\max}} \right| \quad (5.71)$$

where  $H_{\max} > H_n$  denotes the maximum time window size and  $C > 1$  is a sufficiently large constant.

*Proof.* The proof is basically the same as for Lemma 5.4, but here the Lipschitz constants from Definition 5.9 are analyzed separately. Instead of equation (5.30) one finds for the particular splitting scheme (5.67)

$$\|\Delta_{\mathbf{z},n}^k\|_2 \leq L \|\Delta_{\mathbf{y},n}^k\|_2 + L_D \|\Delta_{\mathbf{y},n}^{k-1}\|_2 \quad (5.72)$$

and analogously to (5.29) integration gives for  $T_n < \tau \leq T_{n+1}$

$$\|\Delta_{\mathbf{y},n}^k(\tau)\|_2 \leq \|\Delta_{\mathbf{y},n}^0(T_n)\|_2 + L \int_{T_n}^{\tau} \|\Delta_{\mathbf{y},n}^k\|_2 dt + L_C \int_{T_n}^{\tau} \|\Delta_{\mathbf{z},n}^k\|_2 dt \quad (5.73)$$

with Lipschitz constants  $L$ ,  $L_C$  and  $L_D$  as defined in Definition 5.9. Now the insertion of (5.72) into (5.73) and solving for the differential difference  $\delta_{\mathbf{y},n}^k$  yields

$$\delta_{\mathbf{y},n}^k \leq (1 + CH_n) \|\Delta_{\mathbf{y},n}^0(T_n)\|_2 + C_D H_n \delta_{\mathbf{y},n}^{k-1}, \quad (5.74)$$

with the constant  $C_D$  as defined above,  $C$  as defined in (5.76) and using the upper bound of the time window size  $H < H_{\max}$ . Then (5.72) and (5.74) imply

$$\delta_{\mathbf{z},n}^k \leq C \|\Delta_{\mathbf{y},n}^0(T_n)\|_2 + C \delta_{\mathbf{y},n}^{k-1}. \quad (5.75)$$

with the overall constant

$$C > \left\{ \frac{L(1 + L_C)}{1 - L(1 + L_C)H_{\max}}, (L_C + L_D) \right\}. \quad (5.76)$$

This concludes the proof of (5.69); the other estimate (5.70) is shown analogously.  $\square$

Now the following convergence result is immanent, [9]:

**Theorem 5.15** (Convergence of semiconductor/circuit cosimulation). *Let the assumptions of Theorem 5.10 be given. Then the Gauß-Seidel-type dynamic iteration of the semiconductor (5.62c) and circuit subsystems (5.62a) coupled by the interfaces (b) from Definition 5.8 is unconditionally stable and convergent with a window-wise convergence rate  $\mathcal{O}(H)$ , independently of the computational sequence.*

*Proof.* Stability and convergence follow from Cor. 5.6. Thus only the convergence rate  $\mathcal{O}(H)$  must be shown. The spectral radii of the iteration matrices  $\mathbf{K}_C$  and  $\mathbf{K}_D$  are given by  $\rho(\mathbf{K}_C) = C_C H_n$  and  $\rho(\mathbf{K}_D) = C_D H_n$ , respectively. Thus for  $H \geq \max H_n$  follows convergence rate  $\mathcal{O}(H)$ , where  $C_C$  and  $C_D$  are estimates for the leading coefficients.  $\square$

Although both computational sequences have the same order, the proof shows that the speed of the iteration scheme depends on different leading order coefficients, i.e.,  $C_C$  and  $C_D$ . Again, their estimates differ only by the Lipschitz constants  $L_C$  and  $L_D$ , reflecting the strength of the coupling via differential and algebraic equations, see (5.71). This difference can be observed in numerical simulations, see Section 6.3.2.

## Conclusions

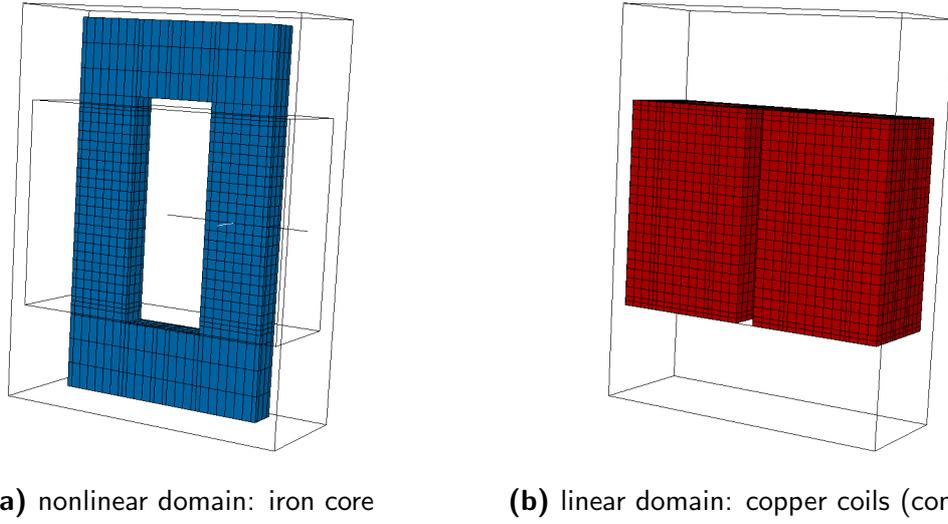
In this section the stability and convergence of the semiconductor/circuit cosimulation was mathematically analyzed, following [1, 9]. It was shown that only parameter coupling guarantees convergence for both computational sequences. This is documented by a numerical example using the different coupling interfaces and sequences in Section 6.3.2.

## 5.4 Domain Substructuring in MQS Devices

The key idea of the multirate approaches in the previous Sections 5.2 and Section 5.3 in particular was that the field subproblem was reduced within the field/circuit coupled problem, e.g., by Schur complements. The multirate behavior was exploited by the fact that the reduced model was only updated on demand. In this sense the dynamic iteration method is a domain decomposition technique in time: the different domains (i.e., subproblems) are solved separately and solutions are exchanged. In the special case of domain decomposition where the different domains do not overlap, the algebraic counterpart of the Steklov-Poincaré operator is in fact a Schur complement (cf. Section 5.2). Its application within the field problem already occurred naturally in the DAE-index analysis, cf. equation (4.16). This special case of domain decomposition is referred to as *domain substructuring*, see e.g. [104, Chapter 2.3].

Although the ideas are similar, the focus in this chapter is not on multirate time-integration for coupled problems but on finding a beneficial decomposition of the field subproblem itself, i.e., system (3.10). The chapter follows [36], where the solution of the transient nonlinear magnetoquasistatic field computation is accelerated by the Schur complement method. This approach is well-known and similar to eliminations in saddle point problems, [16], or the reduction of subdomains in simulations with different discretization techniques, e.g. finite elements coupled to boundary elements, [87], or finite elements coupled to spectral elements, [38]. Furthermore, cascaded conjugate gradient methods have been used before to remove material jumps, [89].

The Schur complement method bases on the important observation that eddy-current problems consist of large subdomains that are only described by linear and static equations (e.g. air), Fig. 5.13. For the corresponding degrees of freedom it is sufficient to solve a linear system once on beforehand and use this information through out the whole simulation, i.e., a factorization of the linear parts is precomputed and reused in every solve of a linear system.



**Figure 5.13:** Domain substructuring of a transformer. Iron core exhibits eddy currents and a nonlinear permeability. The surrounding air and the coils are modeled by static and linear equations (no eddy currents, strands below skin-depth), cf. Fig. 3.4.

### 5.4.1 Linear/Nonlinear Decomposition

Implicit time-integrators solve a series of nonlinear systems for each time step, Section 5.1. Typically this is done by the Newton-Raphson method solving linear systems that feature the matrix pencil, cf. equation (2.25):

$$\mathbf{J}(\bar{\mathbf{a}}) := \lambda \mathbf{M}_{\bar{\sigma}} + \mathbf{K}_{\nu}(\bar{\mathbf{a}}) \quad (5.77)$$

where  $\lambda$  is related to the time step size. For the implicit Euler method with constant time step size  $h$ , it holds:  $\lambda = 1/h$ , Section 5.1. The differential curl-curl matrix  $\mathbf{K}_{\nu}(\bar{\mathbf{a}})$ , see Definition 2.6 and the conductivity matrix  $\mathbf{M}_{\bar{\sigma}}$  have typically a common nullspace in 3D formulations. This is removed by a regularization technique, e.g. Grad-Div, as given in Section 2.2.3.

In relevant technical applications, typically only few materials are modeled in such a way that (non-)linear or even hysteretic behavior is taken into account. Furthermore many regions are non-conductive, either because the material is nonconductive (air) or because it is modeled non-conductively to prevent eddy currents, e.g. without resolving laminations or windings in the discretization (e.g. as for copper coils, see Section 3.2). This turns the corresponding degrees of freedom into algebraic variables and increases the nullspace of the conductivity matrix  $\mathbf{M}_{\bar{\sigma}}$ .

Consequently the models contain large regions described only by linear algebraic relations. The entries of the system matrix  $\mathbf{J}$  related to edges that are strictly linear model parts do not change during the overall simulation. This is commonly exploited in professional implementations, e.g., the assembly of the corresponding elements is bypassed. But this does not fully exploit the structure: the linear model parts are still solved unnecessarily in every step of the nonlinear iteration in the time-integration, [36].

The degrees of freedom of the eddy-current problem (2.18) can be separated by projec-

tors into edges belonging to conducting  $\mathbf{P}_\sigma \widehat{\mathbf{a}}$  and non-conducting regions  $\mathbf{Q}_\sigma \widehat{\mathbf{a}}$ , see Definition 2.8. For simplicity of notation it is assumed in the following that the DoFs are partitioned accordingly into  $\widehat{\mathbf{a}}_1$  and  $\widehat{\mathbf{a}}_2$ , cf. Remark 4.2. Then the eddy current problem becomes a coupled magnetoquasistatic/magnetostatic system

$$\mathbf{M}_{11} d_t \widehat{\mathbf{a}}_1 + \mathbf{K}_{11}(\widehat{\mathbf{a}}) \widehat{\mathbf{a}}_1 + \mathbf{K}_{12} \widehat{\mathbf{a}}_2 = \widehat{\mathbf{j}}_{s,1} \quad (5.78a)$$

$$\mathbf{K}_{12}^\top \widehat{\mathbf{a}}_1 + \mathbf{K}_{22} \widehat{\mathbf{a}}_2 = \widehat{\mathbf{j}}_{s,2} \quad (5.78b)$$

where  $\mathbf{K}_{11}$  and  $\mathbf{K}_{22}$  are the curl-curl matrices of the respective subdomains and  $\mathbf{K}_{12}$  is a interface matrix. The regularity of  $\mathbf{K}_{22}$  follows from the Gauging Assumption 2.13. Thus (5.78) fulfills the standard criterion for a differential-index-1 equation, Section 4, [101].

The first equation (5.78a) is an ordinary differential equation with a positive definite conductivity matrix  $\mathbf{M}_{11}$  and the second equation (5.78b) is an algebraic constraint. However time-discretization turns (5.78) into a nonlinear algebraic problem that has to be solved by Newton-Raphson.

For a given  $\widehat{\mathbf{a}}_1$  we can solve the algebraic equation for the non-conductive domain and reinsert the resulting  $\widehat{\mathbf{a}}_2$  into the differential equation (5.78a). This is in terms of discrete domain substructuring the Schur complement

$$\Sigma_{11}(\widehat{\mathbf{a}}_1) := \mathbf{K}_{11}(\widehat{\mathbf{a}}_1) - \mathbf{H}_{11} \quad \text{with} \quad \mathbf{H}_{11} := \mathbf{K}_{12}(\mathbf{K}_{22})^{-1} \mathbf{K}_{12}^\top$$

that was already used in the index-analysis, see equation (4.16). This approach is known in DAE theory as index-reduction, because we have transformed the index-1 DAE (5.78) into an ordinary differential equation (index-0)

$$\mathbf{M}_{11} d_t \widehat{\mathbf{a}}_1 + \Sigma_{11}(\widehat{\mathbf{a}}_1) \widehat{\mathbf{a}}_1 = \widehat{\mathbf{j}}_{11} \quad (5.79)$$

with the reduced right-hand-side  $\widehat{\mathbf{j}}_{11} := \widehat{\mathbf{j}}_{s,1} - \mathbf{K}_{12}(\mathbf{K}_{22})^{-1} \widehat{\mathbf{j}}_{s,2}$ . If the application of the inverse  $\mathbf{K}_{22}$  is feasible, the Schur complement system is iteratively solved by Listing 3, that is basically the preconditioned conjugate gradients method with some additional linear solves for the complement, [110].

In Listing 3, the matrix  $\mathbf{J}_{11} := \lambda \mathbf{M}_{11} + \mathbf{K}_{11}(\cdot)$  denotes the upper left block of the matrix pencil (5.77) and  $\widehat{\mathbf{a}}_1(t_0)$  is the initial value for time-integration and  $\mathbf{P}$  is a matrix for preconditioning. So far, this method only differs from the Schur complement equipped by standard CG by the fact that, here, the Schur complement is not explicitly formed, [36].

## 5.4.2 Convergence and Eigenvalues

The convergence speed of conjugate gradients is determined by the effective condition number  $\lambda_{max}/\lambda_{min}$ , which is the quotient of the largest  $\lambda_{max}$  and smallest  $\lambda_{min}$  nonzero eigenvalue, [122, 78]. For the 3D eddy-current problem this is especially important, because a gauging shifts zero eigenvalues to the positive axes and may impair the condition number, [20, 37]. The same analysis must be carried out for the Schur method above, since it is closely related to gauging (i.e., equation (5.79) is regularized).

The spectrum of the Schur complement consists of two kinds of eigenvalues: those from conductive and those from non-conductive regions. In conductive regions the positive eigenvalues of the conductivity matrix are further increased by the inverse of the time step

---

**Listing 3** Schur complement PCG method, [36]

---

```

solve  $\mathbf{K}_{22}\mathbf{p}_2^{(0)} = \widehat{\mathbf{j}}_{s,2}$ 

 $\mathbf{r}_1^{(0)} := \widehat{\mathbf{j}}_{s,1} - \mathbf{K}_{12}\mathbf{p}_2^{(0)} - \mathbf{J}_{11}\widehat{\mathbf{a}}_1^{(0)} + \mathbf{J}_{11}\widehat{\mathbf{a}}_1(t_0)$ 

solve  $\mathbf{P}\mathbf{z}_1^{(0)} := \mathbf{r}_1^{(0)}$ 

 $\mathbf{p}_1^{(0)} := \mathbf{z}_1^{(0)}$ 

for  $i := 0, 1, \dots$ , until convergence
  solve  $\mathbf{K}_{22}\mathbf{p}_2^{(i+1)} = \widehat{\mathbf{j}}_{s,2} - \mathbf{K}_{12}^\top\mathbf{p}_1^{(i)}$ 

   $\mathbf{x}_1^{(i)} := \mathbf{J}_{11}\mathbf{p}_1^{(i)} + \mathbf{K}_{12}\mathbf{p}_2^{(i+1)}$ 

   $\alpha^{(i)} := (\mathbf{r}_1^{(i)}, \mathbf{r}_1^{(i)}) / (\mathbf{x}_1^{(i)}, \mathbf{p}_1^{(i)})$ 

   $\widehat{\mathbf{a}}_1^{(i+1)} := \widehat{\mathbf{a}}_1^{(i)} + \alpha^{(i)}\mathbf{p}_1^{(i)}$ 

   $\mathbf{r}_1^{(i+1)} := \mathbf{r}_1^{(i)} - \alpha^{(i)}\mathbf{x}_1^{(i)}$ 

  solve  $\mathbf{P}\mathbf{z}_1^{(i+1)} := \mathbf{r}_1^{(i+1)}$ 

   $\beta^{(i)} := (\mathbf{r}_1^{(i+1)}, \mathbf{z}_1^{(i+1)}) / (\mathbf{r}_1^{(i)}, \mathbf{z}_1^{(i)})$ 

   $\mathbf{p}_1^{(i+1)} := \mathbf{z}_1^{(i+1)} + \beta^{(i)}\mathbf{p}_1^{(i)}$ 

end

```

---

$\lambda$ . This shifts the corresponding eigenvalues of the curl-curl matrix (further) to the right on the positive axis, while the eigenvalues of the non-conductive domain remain unaltered, Fig. 5.14a.

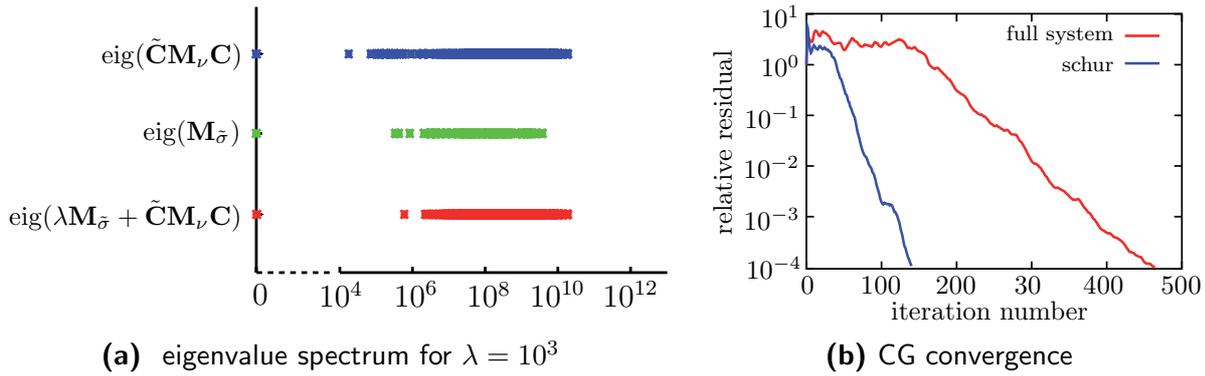
By construction, all eigenvalues of the Schur system (5.79) are affected by conductivities ( $\mathbf{M}_{11}$  has full rank). This diminishes the influence of the material jump from the system's spectrum and improves the speed of convergence of the CG method. Additionally, the (maximum) number of iterations decreases because the system has less DoFs than before.<sup>2</sup> Fig. 5.14b shows the improvement in the convergence rate using the 2D example of Section 6.4. This plot does not take the additional computational costs into account: the higher convergence speed comes at the price of a more expensive iteration since in each step an inner solver computes the Schur system, see Listing 3

$$\mathbf{K}_{22}\mathbf{p}_2^{(i+1)} = \widehat{\mathbf{j}}_{s,2} - \mathbf{K}_{12}^\top\mathbf{p}_1^{(i)}.$$

Furthermore an additional matrix factorization must be created on beforehand, once for the whole time-integration procedure. Consequently it is important to make the inner solving procedure as computationally cheap as possible. If this is feasible the reduced number of iterations will compensate for the increased costs.

---

<sup>2</sup>In exact arithmetics CG needs fewer iterations than the number of DoF, [110].



**Figure 5.14:** Eigenvalues and PCG convergence. The eigenvalues in (a) are given for the FIT discretization of a 3D transformer model: 882 of 4752 eigenvalues of the curl-curl matrix are zero, 2457 eigenvalues of the conductivity matrix are zero and finally their sum has 673 zeros. The plot (b) shows the convergence for the full and Schur system (5.79) for a 2D model (the models are discussed in Section 6.4, see [36]).

### 5.4.3 Inner Solvers

Instead of computing the explicit inverse for the Schur complement an inner solving step is introduced into PCG in Listing 3. There are several possibilities for solving or modeling this block, [36]

- an iterative method, preferably equipped with deflation, in order to benefit from the repeated solving with the same system matrix, e.g. [52, 35]
- a sparsity preserving factorization (e.g., Cholmod [31]) in the preprocessing phase together with forward/backward substitution steps during the outer CG iterations;
- a sparse, approximate inverse, e.g., using hierarchical matrices, [66]
- MOR (e.g. using proper orthogonal decomposition, Krylov techniques, etc.) constructed from the first solutions of the time-stepping process;
- a model of the non-conductive subdomain based on another discretization technique that is better suited for models with homogeneous material, e.g. boundary elements, [87], spectral elements, [38] or nodal formulations for the curl-curl equation, [79].

The first three ‘solvers’ compute (approximations of) the solution that belong to the original problem, whereas the last two alternatives solve a modified problem (in the static and linear regions).

The second approach is discussed in Section 6.2 for the eddy-current problem of a transformer model in 2D and 3D formulations. It features a sparse Cholesky factorization of the matrix  $\mathbf{K}_{22}$ . In 2D this approach works very well, while in 3D the additional burden of the forward/backward substitutions in each iterations may become dominant, [36].

As we have said before, in 3D the block  $\mathbf{K}_{22}$  has a non-trivial nullspace (the gradient fields, see (2.21)). Thus standard factorizations fail and a regularization must be applied, Section 2.2.3. For the Schur complement method an additional constraint for the regularization is crucial to preserve: the separation between conductive and non-conductive model

regions. This ensures that the Schur complement removes the (conductivity) material jump completely and only then can a significant gain in PCG's convergence be expected, [36].

#### 5.4.4 Conclusions

In this section we have proposed a variant of the Schur complement method for the eddy-current problem. This adapted version has been shown to exploit the material structure, i.e., the static and linear part. It reduced the differential-algebraic problem to an ordinary differential equation and the resulting system matrix has an advantageous eigenvalue spectrum. This speeds up the convergence of the preconditioned conjugate gradient algorithm. Computational examples are given in Section 6.2.

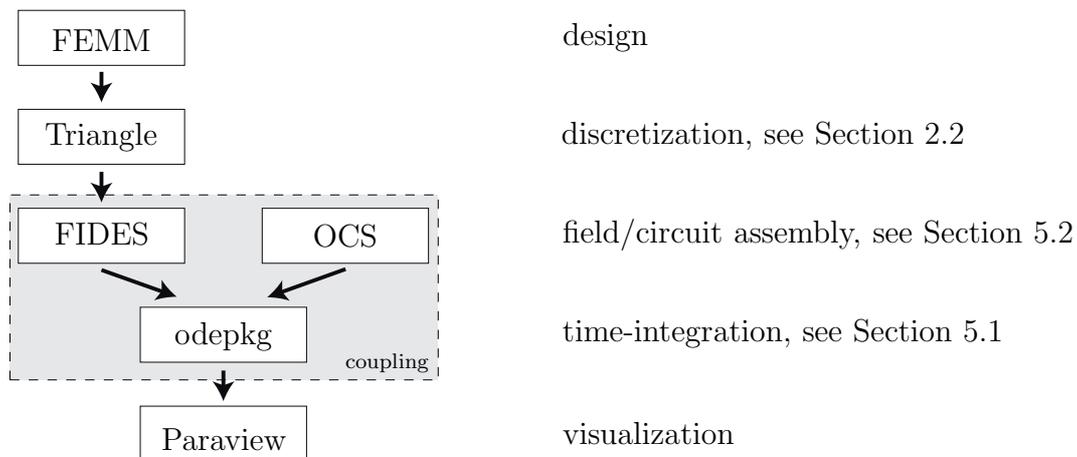
## 6 Numerical Examples

In this chapter the multirate methods and analysis given in Chapter 5 are numerically verified by examples from electrical engineering.

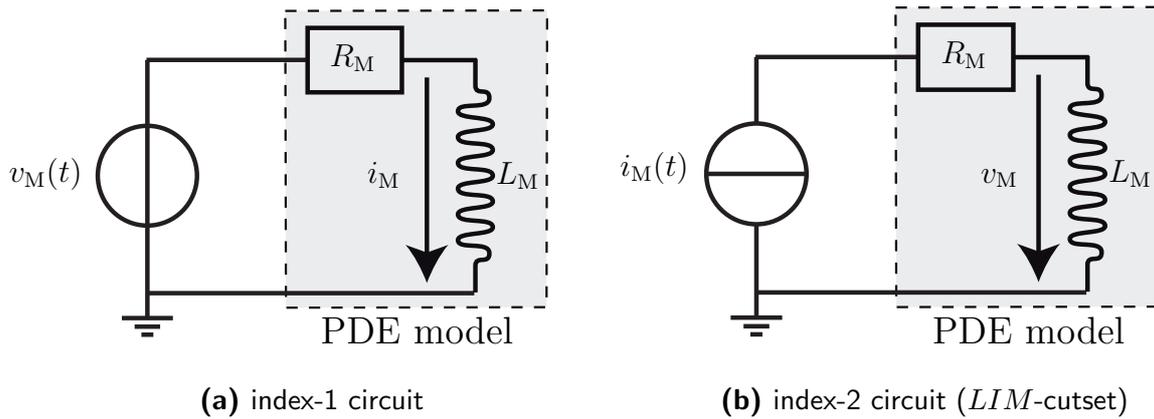
The software is written in Octave within the framework of the demonstrator platform of the CoMSON project (Coupled Multiscale Simulation and Optimization in Nanoelectronics). For the circuit simulation part the OCS package (Octave's Circuit Simulator) is used. It is coupled to the new package FIDES (Field Device Simulator) for magnetoquasistatic field device simulations, [113]. The 3D examples are obtained from handmade FIT discretizations, while 2D examples are designed in FEMM and discretized by Triangle [94, 121]. Visualizations are obtained by Paraview, [98]. The workflow is shown in Fig. 6.1.

The MQS devices may consist of several conductor models (stranded or solid), connected to the electric circuit as one multiport device, Section 3.2. The connection to OCS is established by calling a corresponding device file in the circuit netlist, that is an IFF-file, [55]. For the strong field/circuit coupling (*monolithic coupling*) the device file `Mfidesmono` defines the (full or reduced) element stamp, such that the field equations are solved along with the circuit equations by the same time stepping scheme, Section 5.2. Consequently the device file adds additional unknowns (external and internal variables) to the circuit problem: each conductor inside the model is excited by a voltage drop and hence it is represented in the circuit by two pins (2 external variables), Section 3.2.

The internal variables are the magnetic vector potential  $\hat{\mathbf{a}}$  and the currents through each conductor model  $\mathbf{i}_M$ . They require several input parameters: the filename of the model, followed by the specification of external variables. These specifications must meet the topology of the field model, e.g., the number of unknown currents must match the number of columns of the coupling matrix  $\mathbf{X}_M$ , [114].



**Figure 6.1:** Flow chart of software packages. The focus in this treatise is on the efficient coupling of simulator packages, i.e., the dashed box. The analysis and methods from Section 5 belong there.



**Figure 6.2:** DAE-index example circuits. (a) voltage-driven and (b) current-driven devices have a different DAE-index, [8].

The procedure for the cosimulation (*dynamic iteration*) follows a different approach: the resistances are defined separately and the fluxes are given by a special nonlinear inductance device `Mfidesinduct`. This device file receives the extracted inductances from the field subproblem using an outer iteration, see interface (5.57) and interpolates if necessary, [114].

The following examples are discussed in their corresponding sections

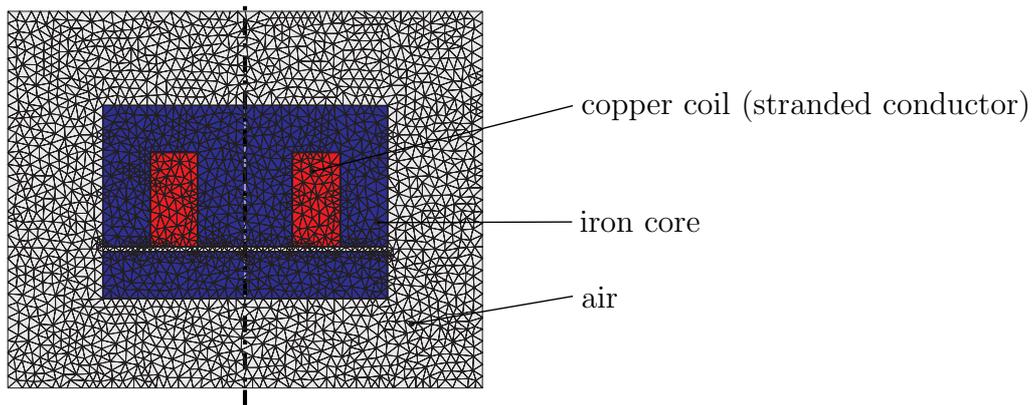
- 5.1 **DAE-index:** the increased error when solving a DAE-index-2 system monolithically for the MQS device is analyzed using a linear axisymmetric inductance example,
- 5.2 **bypassing:** the multirate bypassing for reduced field stamps is applied to a pulsed circuit coupled to a nonlinear 2D model of a transformer,
- 5.3 **cosimulation:** examples of the Gauß-Seidel-like dynamic iteration of MQS and semiconductor devices with circuits; the results for the semiconductor example are taken from [9],
- 5.4 **substructuring:** the domain substructuring method is applied to a transformer in 2D and 3D formulations.

## 6.1 DAE-Index in Applications

In this section the simplest possible field/circuit configuration is analyzed: a linear magnetostatic field model is connected to an independent voltage or current source, see Fig. 6.2 from [8]. This setting is already sufficient to demonstrate the numerical consequences of the DAE-index results stated in Section 4. One solves the system

$$\begin{aligned} \mathbf{K}_\nu \hat{\mathbf{a}} &= \mathbf{X}_M i_M \\ -\frac{d}{dt} \mathbf{X}_M^\top \hat{\mathbf{a}} + \mathbf{R}_M i_M &= v_M \end{aligned}$$

either for given voltages  $v(t)$  or currents  $i(t)$ . The magnetostatic system only exhibits inductive effects (no eddy currents) and the coil is modeled by a stranded conductor, see Section 3.2.1. The corresponding axisymmetric PDE model was discretized by FEMM,



**Figure 6.3:** Inductor example. Axisymmetric inductor model from FEMM, discretized by Triangle, [94]. The coupling is established via stranded conductor models, see [9].

[94]. The device model, Fig. 6.3, is taken from the FEMM examples section in the online tutorial (‘Inductance Calculation Example’, file: `induct1a.fem1`). In this simple case the tractability concept matches the Kronecker and differentiation indices and thus it corresponds to the special cases in [117, 131].

Solely the circuit topology determines the DAE-index of the coupled problem, see Theorem 4.2 and Theorem 4.5. The index-1 and index-2 cases are obtained for the different choices (see Fig. 6.2)

- (a) voltage source connected to the device, i.e., incidence matrices  $\mathbf{A}_V = [1]$  and  $\mathbf{A}_M = [-1]$ , states an index-1 problem,
- (b) current source connected to the device, i.e., incidence matrices  $\mathbf{A}_I = [1]$  and  $\mathbf{A}_M = [-1]$ , states an index-2 problem (*LIM*-cutset).

As a consequence of linearity and statics a lumped analytical solution is given by a resistance and an inductance: the magnetic vector potential is determined by

$$\hat{\mathbf{a}} := \mathbf{K}_\nu^{-1} \mathbf{X}_M i_M$$

and thus the following lumped quantities are easily extracted from the PDE model

$$R_M := \mathbf{X}_M^\top \mathbf{M}_\sigma^+ \mathbf{X}_M = 0.070197 \Omega \quad \text{and} \quad L_M := \mathbf{X}_M^\top \mathbf{K}_\nu^{-1} \mathbf{X}_M = 1.73 \cdot 10^{-3} \text{H}$$

and thus for the sinusoidal current source

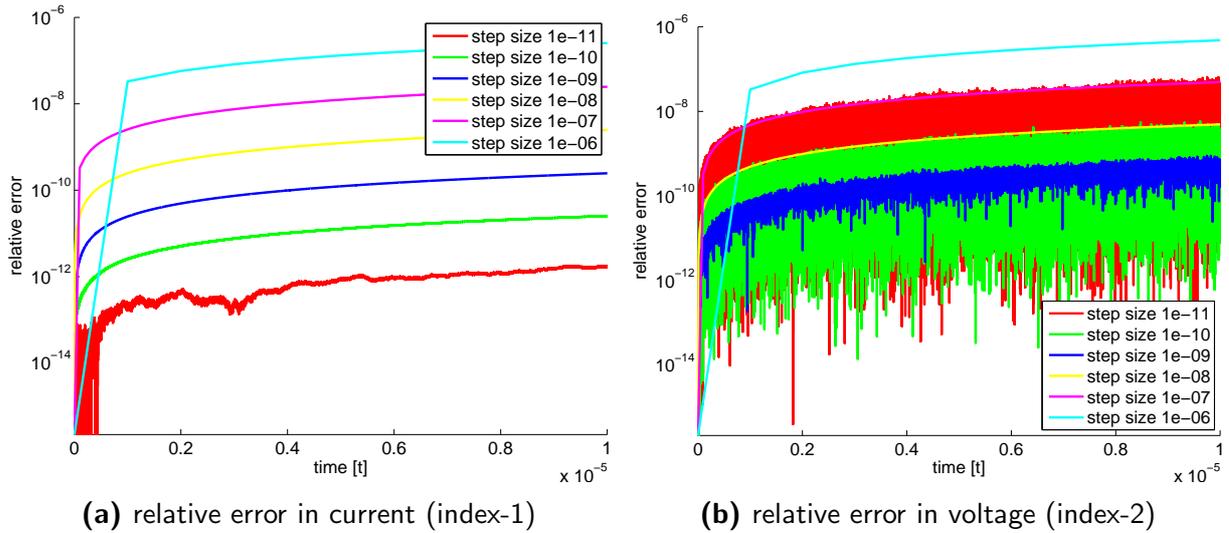
$$i_M(t) = \sin(2\pi ft) \quad \text{with a frequency} \quad f = 50\text{Hz} \quad (6.1)$$

the voltage drop (at the RL-element) is given by

$$v_M(t) = -2\pi f L_M \cos(2\pi ft) - R_M \sin(2\pi ft). \quad (6.2)$$

For the time-discretization the implicit Euler scheme was used with fixed step sizes  $h = 10^{-11}\text{s}, 10^{-10}\text{s}, \dots, 10^{-6}\text{s}$ , Section 5.1. The application of higher-order methods is

<sup>1</sup>see <http://www.femm.info/wiki/InductanceExample>



**Figure 6.4:** Index-1 vs. index-2 errors. Errors for (a) the index-1 setting (driven by voltage source) and (b) the index-2 setting (driven by current source), [9].

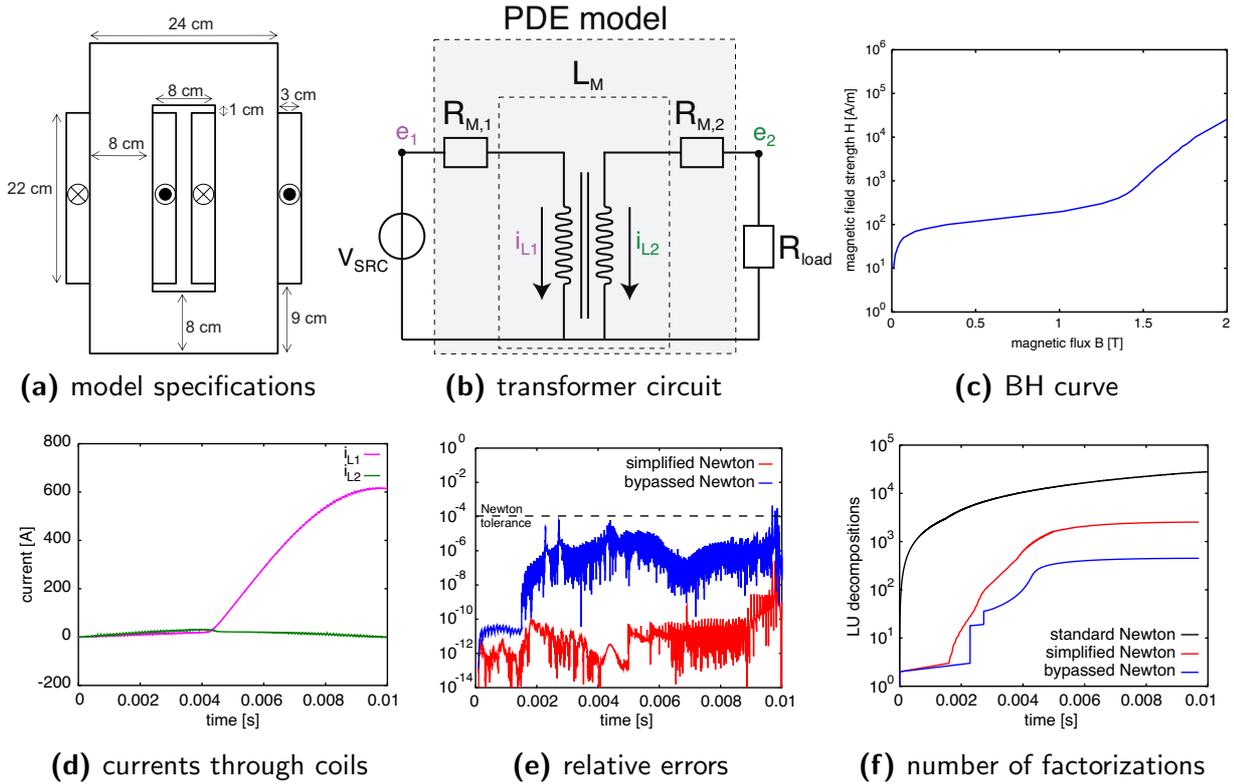
straightforward but it will not give further insight (e.g. using FIDES and a RADAU scheme supplied by odepkg, see Fig. 6.1).

The numerical solutions of the index-1 and index-2 case are compared to the analytical reference solution as given above (either the current (6.1) or voltage (6.2)). Fig. 6.4 shows the numerical errors due to time-integration for both cases. The relative errors of both problems behave very differently: in the index-1 setting, Fig. 6.4a, the relative error of the current decreases with the step size. Small oscillations (near the machine precision) occur at the smallest step size  $h = 10^{-11}$ s. On the other hand, in the index-2 setting Fig. 6.4b, the error oscillates at a high amplitude for step sizes below  $h = 10^{-8}$ s. This is a typical index-2 phenomenon: the error increases while the step size decreases. These numerical results underline the difference that is mathematically described by Theorem 4.2 and Theorem 4.5.

Although index-2 problems are rather ill-conditioned, Fig. 6.4b shows clearly that the index-2 errors are not propagated in time. Remark 4.4 explains this behavior: the index-2 components enter the system only linearly and cannot affect subsequent time-steps, [6, 128]. Nonetheless one must not use the index-2 variables (the voltages) for step size control, because the (numerical) oscillations might be detected by the control and hence the predictor would suggest unreasonably small step sizes.

## 6.2 Multirate Bypassing

This section discussed the various Schur complement strategies introduced in Section 5.2 and [116]. The field part is a 2D model of a transformer. The secondary side is connected to a load resistance. Fig. 6.5a shows the specifications of the model. The first coil consists of 358 and the second one of 206 copper winding turns with a conductivity of  $57.7 \cdot 10^6$ S/m. The coils are connected as stranded conductors. They are wound around an iron core (with conductivity  $0.5 \cdot 10^6$ S/m); its reluctivity is given by a nonlinear material curve (interpolated by splines), Fig. 6.5c.



**Figure 6.5:** Bypassing example. Device and circuit problem description, reference solution, errors and decompositions, cf. [116].

The model has been drawn and discretized by FEMM and simulated by the software packages OCS, FIDES and odepkg, see Fig. 6.1. The simulation is carried out only for the startup phase, i.e., until the saturation phase is reached, Fig. 6.5d. Afterwards the computation can be continued by a linear model without difficulty.

Fig. 6.5b shows the field/circuit coupling, where a pulse width modulated (PWM) voltage source is connected to the primary side of a transformer. The PWM voltage is switching at a frequency of 20kHz. The secondary side is connected to a load resistance  $R_{\text{load}} = 10\Omega$ .

For the simulation the different strategies of Section 5.2 were implemented in `Mfidesschur` for OCS, [114]: this is (a) the standard Newton-Raphson scheme that serves as a reference, (b) a simplified Newton iteration with Schur complements and (c) a simplified Newton with bypassing of right-hand-side evaluations.

The circuit device element `Mfidesschur` uses the bypassing heuristics defined in Listing 1. This algorithm allows the element to decide independently whether a new factorization of (5.8) or a new right hand-side evaluation (5.9) is necessary or not. The advantage is that this implementation only requires changes to the MQS device element. No other parts of the host circuit simulator need adjustment. The other (basic) elements remain responsible for their respective contributions and the outer (standard) Newton-Raphson scheme is still available for them.

On the other hand, if the step size control of the host simulator is accessible, it should be configured so as to be as conservative as possible, i.e., the step size  $h$  should be kept constant as long as possible. In fact, step size changes require a recomputation of the

**Table 6.1:** Computational costs for the different Newton strategies. The ‘full Newton’ solves the full system of equations without bypassing. The ‘simplified Newton’ solves the reduced system and bypasses some Jacobian updates. The ‘bypassed Newton’ solves the reduced system and bypasses Jacobian and right-hand-side updates.

	decompositions	forward/backward substitutions	stamp evaluations	time
full Newton	23371	27936	27936	20h
simplified Newton	2531	36460	31398	1h
bypassed Newton	450	3171	20449	25min

Schur complement  $\mathbf{L}_h$  but alternatively one may continue with a ‘wrong’ Jacobian, cf. equation (5.7). For the following example from [116] this is not implemented: only backward Euler with fixed step sizes is used, because the examples are determined by pulsed inputs and step size prediction would yield high amounts of rejected steps. In principle adaptive higher-order methods are available by odepkg.

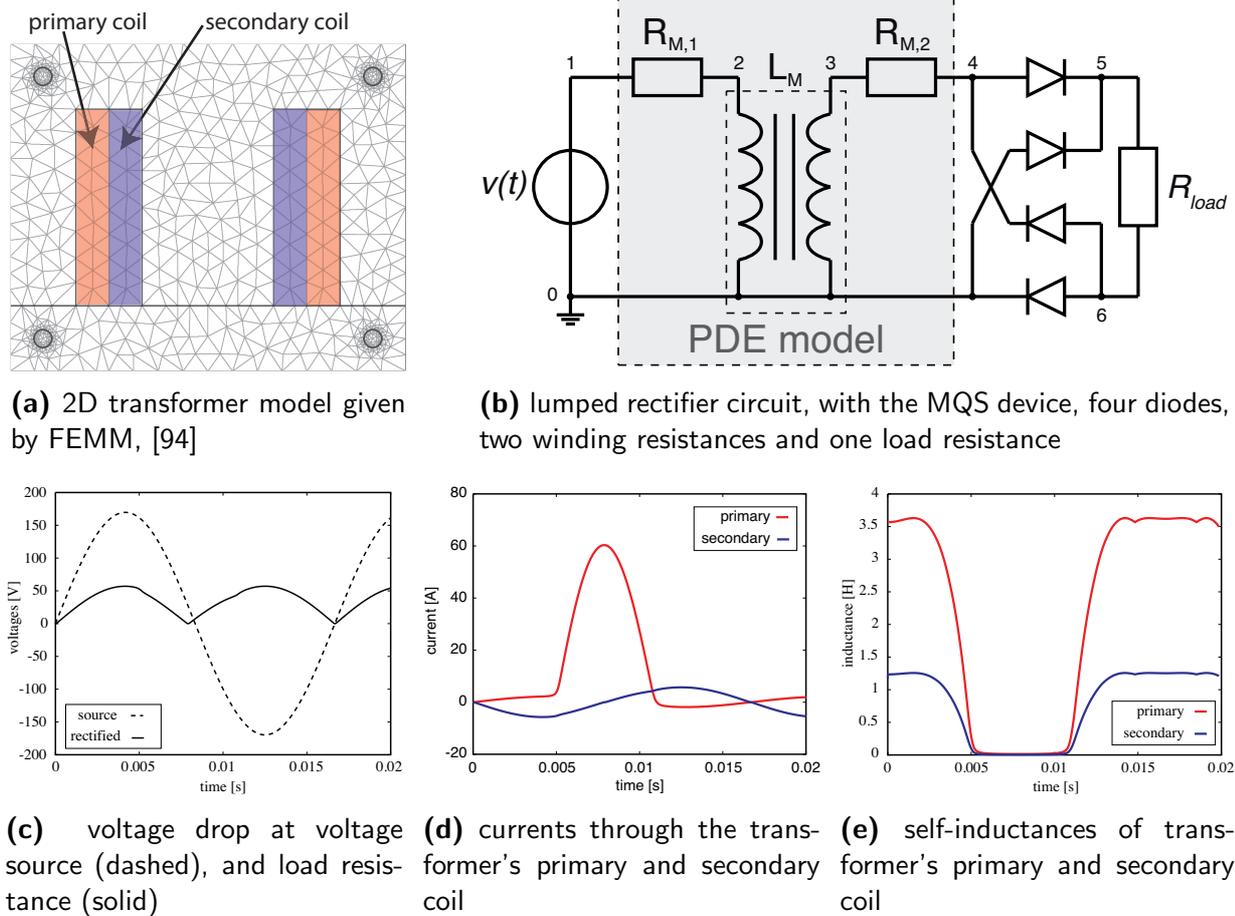
Fig. 6.5f shows the different bypassing strategies in comparison with the reference approach. The reference method is a non-optimized Newton algorithm; it evaluates the material curve (unnecessarily) in every iteration, see Table 6.1. On the other hand both bypassing strategies (simplified and bypassed Newton) detect the linearity in the material curve (for  $t \leq 0.003s$ ) and skip the superfluous evaluations and matrix factorizations. Depending on the error, the nonlinear effect is not important here and the reduced (lumped) models are sufficient, see Fig. 6.5e. Those models preserve correctly the characteristics on the fast scale, e.g., they resemble the switching in the current due to PWM, Fig. 6.5d, while other multirate techniques would probably fail here, especially when waveforms are interpolated and not models (‘the multirate behavior is in the nonlinearity’).

In the highly nonlinear saturation phase,  $0.003s < t \leq 0.007s$ , all approaches require approximately the same (high) number of Jacobian updates. Without those updates the Newton iteration might diverge or even converge to a wrong solution if the bypassing of right-hand-side evaluations is not controlled. Each bypassing of the right-hand-side assumes linearity and as a consequence the Newton iteration requires less Jacobian updates but the error increases, see Fig. 6.5. Finally after the saturation level is reached,  $t > 0.007$ , the field problem behaves again linearly and the updates of the simplified and bypassing Newton are clearly reduced, [116].

### 6.2.1 Conclusion

The bypassing approach was shown to exploit reliably multirate potential in the nonlinear field part. Of course the efficiency depends on the particular choice of error norms, tolerances and the device’s characteristics. Nonetheless, the numerical experiments indicated that the heuristic is rather insensitive to changes in those parameters and that the computational costs can be significantly reduced even when using conservative parameters (i.e., small tolerances).

Obviously, the changes in the saturation cause the high computational costs. For a transformer example this typically occurs only during the start-up phase, but in an induction machine, where the saturation follows the rotation, one is forced to recompute the Schur complement in every turn. Nonetheless the rotation is still determined by the



**Figure 6.6:** Field/circuit cosimulation example. Nonlinear 2D transformer (a) coupled to a nonlinear rectifier (b). The plots show voltages (c), currents (d) and self inductances (e), [9].

energy (5.10) and not by the fast frequency of the pulsed inputs and thus solving on the slow scale should be sufficient for most applications. Furthermore an optimized version of the bypassing could save previously computed inductances, e.g. in dependence of the rotor angle, and reuse them in the subsequent simulation.

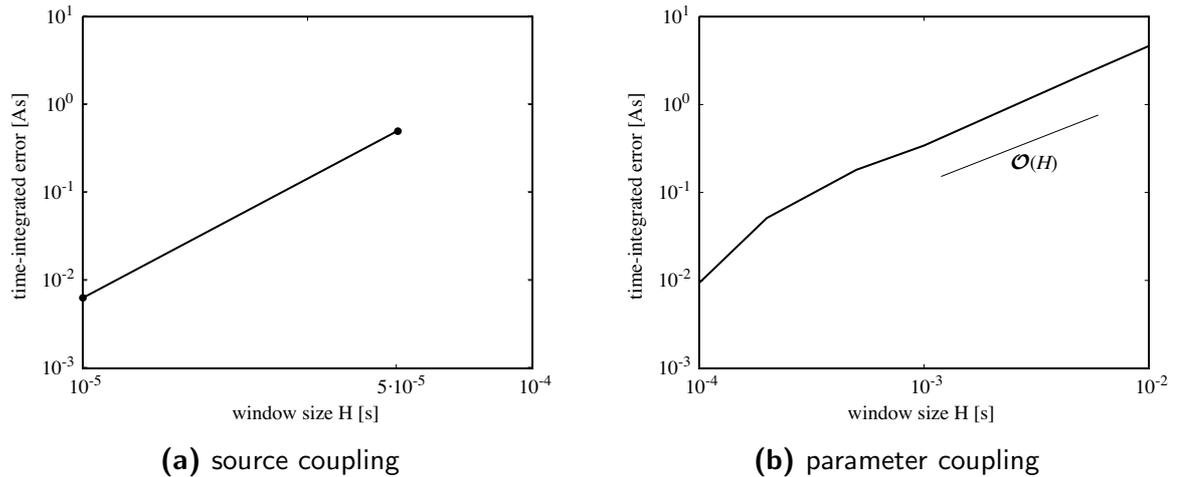
### 6.3 Cosimulation

In this section cosimulations of exemplary field/circuit and semiconductor/circuit problems are performed using Gauß-Seidel-like dynamic iteration schemes, see Sections 5.3.5 and 5.3.6. For both settings the different coupling interfaces

- (a) source coupling
- (b) parameter coupling (using extracted, lumped models)

are implemented and numerically compared. Furthermore it will shown that the convergence speed predicted in Section 5.3 is verified by the simulations.

The focus in the field/circuit case is primarily on the multirate potential that can be exploited in cosimulation, while the main topic in semiconductor/coupling is the compu-



**Figure 6.7:** Convergence of field/circuit cosimulation. Figures show the splitting error in the primary current on the full time interval in dependence of the window size  $H$ .

tational sequence and the convergence properties. Due to similar time scales for semi-conductors and integrated circuits there is typically no multirate potential in this kind of simulations (with the obvious exception of devices in latent branches, [124]).

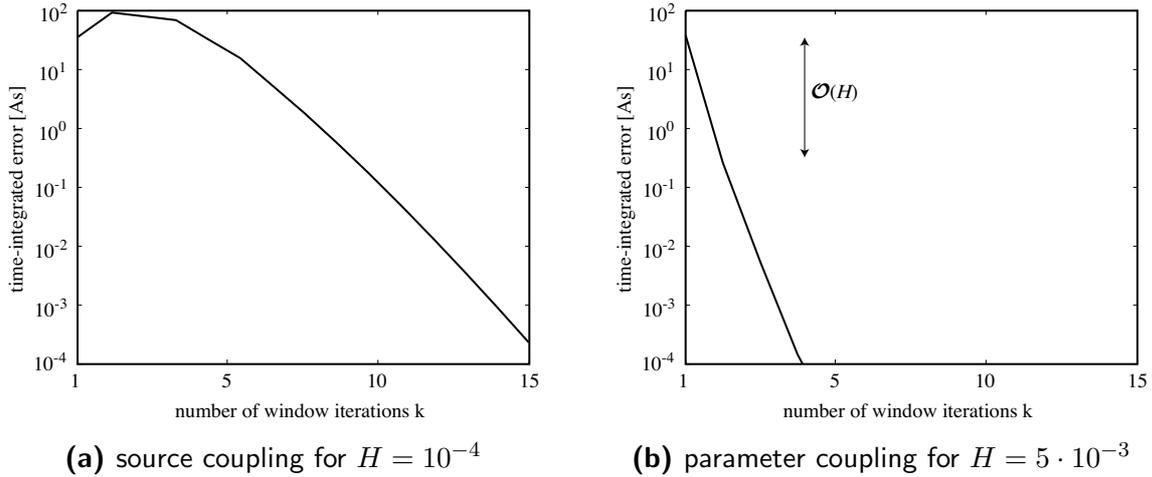
### 6.3.1 Field/Circuit Problem

The dynamic iteration given in Section 5.3 is applied here to the field/circuit coupled example from [9]. This is the rectifier circuit shown in Fig. 6.6b connected via the transformer (see Fig. 6.6a) to a sinusoidal voltage source with an effective voltage  $160 \sin(2\pi ft)$  V at low frequency ( $f = 60$  Hz). The transformer shown in Fig. 6.6a is a single-phase isolation transformer, whose coils are described by the stranded conductor model, see Section 3.2.1. The discretization was obtained by FEMM, [94] and model specifications are also taken from the FEMM examples section in the online tutorial (‘Determination of Transformer Operating Point’, file: `mytransformer.fem2`). The primary coil has 260 turns and the secondary coil 90 turns. The reluctivity is given by the same nonlinear BH-curve  $\nu(\hat{\mathbf{b}})$  that was used in the previous examples, see Fig. 6.5c.

The circuit model consists of four diodes described by the Shockley’s lumped exponential model (as it is available in OCS). The reverse bias saturation current of those devices is  $\mathbf{I}_s = 10^{-9}$  A. The constant load resistance is given by  $\mathbf{R}_{\text{load}} = 10 \Omega$  and the resistances of the coil windings are extracted from the FEMM model, i.e.,  $\mathbf{R}_1 = 0.44937 \Omega$  and  $\mathbf{R}_2 = 0.061526 \Omega$ , see (3.12c).

The problem is simulated on the time interval  $\mathcal{I} := [0, 10^{-2}]$  s, this is approximately one period of the applied voltage. This time span is sufficient to cover the nonlinear start-up behavior due to inrush currents. The figures 6.6c-6.6e show voltages, currents and inductances, respectively. The voltage plot shows the applied voltage versus the transformed and rectified voltage drop at the load resistance. The inrush currents are clearly visible in the current plot and they cause the very high saturation in the transformer such that the corresponding inductances decrease rapidly.

<sup>2</sup>see <http://www.femm.info/wiki/MyTransformer>

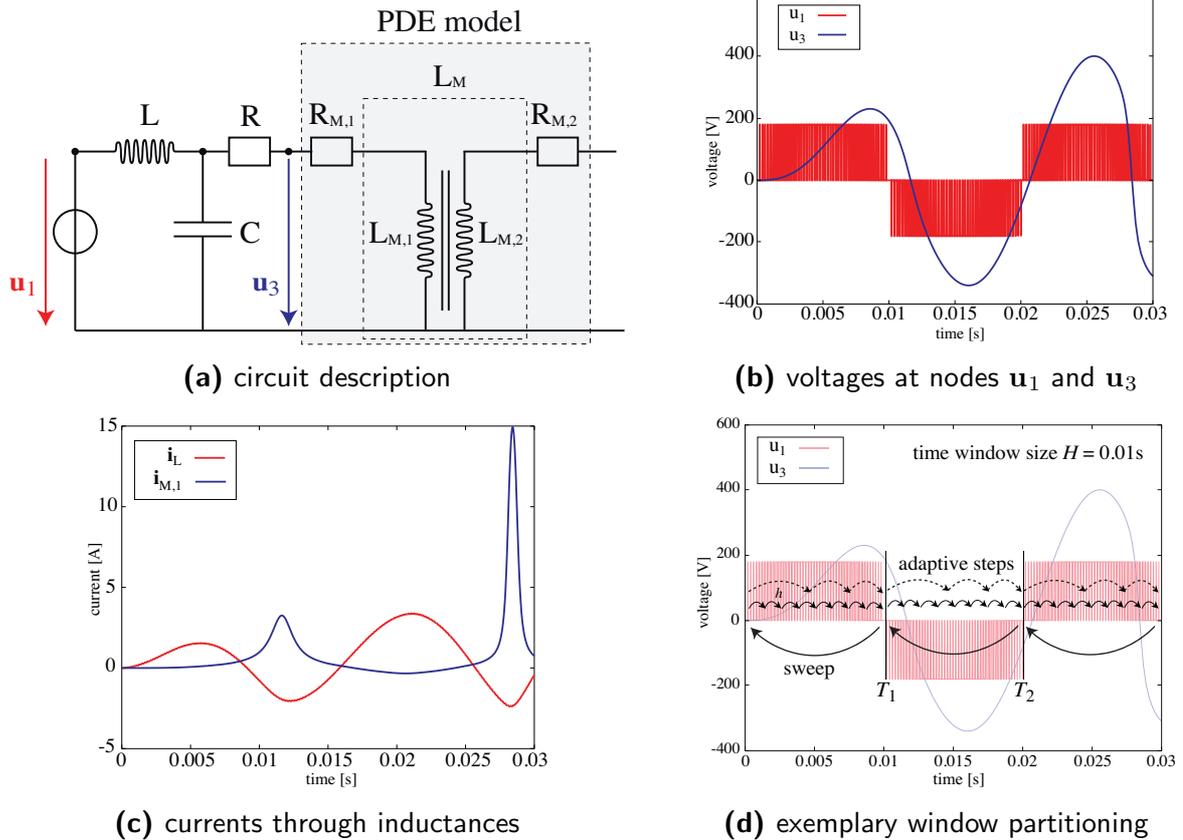


**Figure 6.8:** Contraction of field/circuit cosimulation. Splitting errors in the primary current on the first time window  $[0, H]$  in dependence of the number of iterations.

For the numerical experiments, the time-integrator has been the implicit Euler scheme with fixed time step  $h = 10^{-5}$ s, see Section 5.1. The dynamic-iteration scheme (5.58) has been performed using various window sizes and the two interfaces **(a) source coupling** and **(b) parameter coupling**, see Definition 5.7. The interfaces are implemented in the functions `Mfidescurrent` and `Mfidescurrent` within FIDES, [114]. The cosimulation results have been compared with a monolithic simulation (implicit Euler with  $h = 10^{-5}$ s). Thus the comparison neglects the time-discretization error (due to the implicit Euler) and visualizes only on the splitting error (due to the iteration scheme).

### Contraction

The splitting errors of the source coupling interface are shown in Fig. 6.7a. The simulations were performed with constant extrapolation and up to 5 iterations per window. No interpolation was necessary to recover the waveforms from the discrete currents, due to the same step size  $h$  in both subproblems. Convergence was only obtained for very small time window sizes near the time step size, i.e.,  $H = 2 \cdot 10^{-5}$ s and  $5 \cdot 10^{-5}$ s. The parameter coupling interface comes with additional costs: two additional linear systems must be solved afterwards to extract the  $2 \times 2$  inductance matrix  $\mathbf{L}_D(t)$ , but the effort pays out as shown in Fig. 6.7b. The dynamic iteration with the parameter interface converges faster and for the much larger window sizes  $H = 5 \cdot 10^{-5}$ s,  $\dots$ ,  $10^{-2}$ s as the source coupling (up to 5 iterations per window, constant extrapolation and spline interpolation). Both results confirm neatly the theoretic results of Section 5.3: the error decreases with the time window size if the the window size is sufficiently small. The second convergence study, Fig. 6.8 shows the time-integrated error on the first time window versus the number of iterations. The source coupling converges badly on the first time window ( $H = 10^{-4}$ s) and it divergences on a subsequent window (i.e., the corresponding window size  $H_n$  is not small enough). For the parameter coupling interface, the convergence is much better even for the larger window size  $H = 5 \cdot 10^{-3}$ s. The iteration order is here approximately linear in  $H$ , which matches the expected order in Theorem 5.12. For more than 4 iterations the splitting error is in the order of the time-integration error.



**Figure 6.9:** Multirate cosimulation example. Nonlinear field/circuit configuration (a) exhibiting different time constants in the voltages  $u_1$  and  $u_3$  due to a fast switching PWM voltage source (b); currents (c) and partitioning into time windows  $H$  and time steps  $h$  (d), [118].

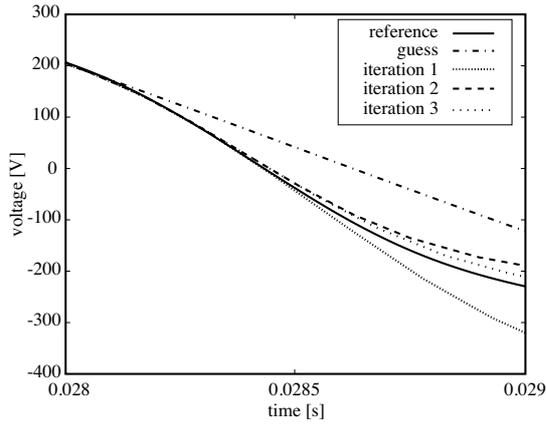
For small window sizes near the step size of the time discretization ( $H \approx h$ ) the discretization error may dominate the asymptotic behavior; then additional iterations are superfluous. In the present case of the implicit Euler method, its accuracy is first order, thus depending on the quality of the initial guess, one or two iterations are typically sufficient. This changes for higher order methods.

The rectifier example could not benefit from multirate time-integration: the time stepping was fixed to simplify the convergence analysis. This will be generalized in the following multirate example.

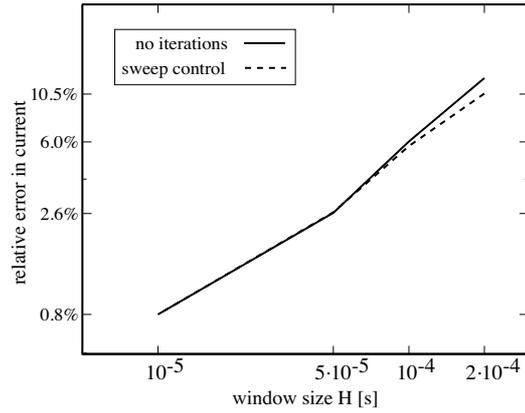
### Multirate in the Field/Circuit Problem

To prove the feasibility of multirate time-integration by dynamic iteration, the following example was given in [118]. The circuit in Fig. 6.9a acts as a low pass filter feeding only the important frequencies to an MQS device. This is a typical engineering approach for signal separation, in contrast to the numerical approach in Section 5.3.5.

The voltage source applies a pulse-width modulated (PWM) signal, while only the low frequency part arrives at the transformer. The MQS device is described here by the same nonlinear model as in Section 6.2, see Fig. 6.5a. The application of an adaptive time-integrator yields time step sizes in the order of  $h_C = 10^{-6}$ s, although step sizes of



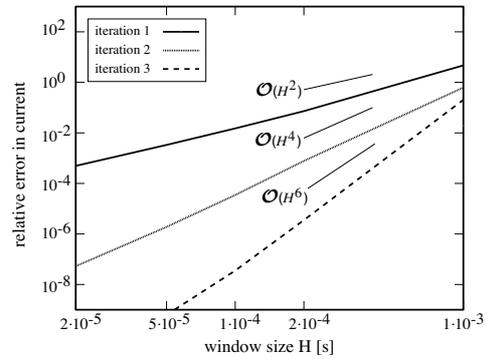
(a) waveforms are improved by additional iterations ( $H = 10^{-2}s$ )



(b) relative error in the current in dependence of the window size

coupling method (time stepping)	linear solves	error
strong (step size $h = 10^{-5}s$ )	6574	5.8%
weak (no sweeps, $H = 5 \cdot 10^{-5}s$ )	3942 + 600	2.6%
weak (sweep control, $H = 10^{-4}s$ )	4282 + 560	5.7%
weak (no sweeps, $H = 10^{-4}s$ )	2196 + 300	6.0%
weak (sweep control, $H = 2 \cdot 10^{-4}s$ )	2726 + 310	10.5%
weak (no sweeps, $H = 2 \cdot 10^{-4}s$ )	1274 + 150	12.7%

(c) computational effort



(d) higher order time integration

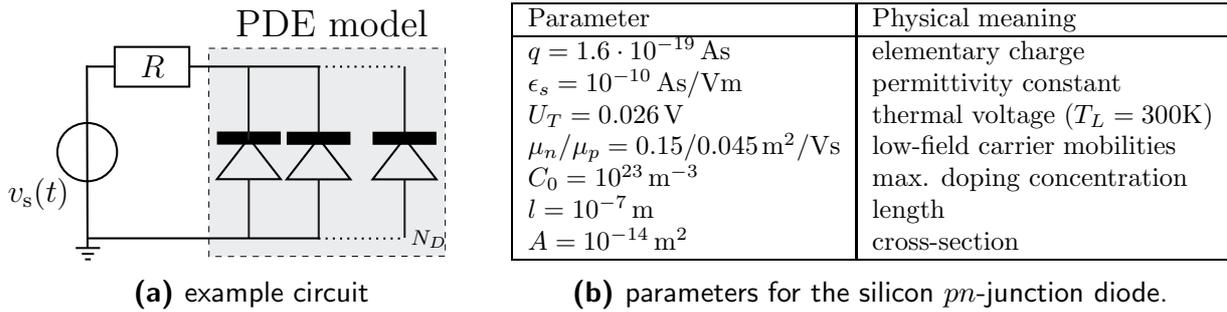
**Figure 6.10:** Convergence of multirate cosimulation. dynamic iterations (a), convergence (b)-(c) and convergence for a higher order Runge-Kutta time-integrator (d), [118].

$h_L = 10^{-4}s$  would be sufficient to render the dynamics of the field model (Fig. 6.9b).

The interface couples via inductance parameters as defined in equation (5.57). All time-integrations have been performed by the implicit Euler method, Section 5.1. The circuit subproblem is discretized by a fixed step size of  $h = 10^{-6}s$ , which is reasonable for fast switching PWM signal, while the field subsystem is solved adaptively. The low order method was chosen to allow easy comparison, although high order adaptive multi-method time-integration is straightforward in FIDES.

The dynamic iteration is performed with linear extrapolation, spline interpolation, fixed time window sizes from  $H = 10^{-5}s$  to  $2 \cdot 10^{-4}s$ , either with ( $k \leq 3$  iterations) or without sweep control (1 sweep), [118]. Table 6.10C shows the computational effort expressed in linear systems solved. The first summand relates to the time-integration and the second relates to the inductance extraction. The costs of the monolithic integration (fixed step size  $h_{ref} = 10^{-6}s$ ) are included for comparison. The relative errors shown in Table 6.10C and Fig. 6.10b are always given with respect to the monolithic reference solution (scaled by the maximal current 15.3 A). Due different time-steppings for the field subsystem, the depicted errors consists of both splitting and time-discretization errors.

The dynamic iteration ('weak' coupling) gives for the window size  $H = 10^{-4}s$  the same level of accuracy as the monolithic approach ('strong'), but requires less than half of the



**Figure 6.11:** Semiconductor/circuit cosimulation example. Model and parameters from [9].

computational effort. The additional iterations improve the accuracy but also increase the computational costs, Fig. 6.10a. The iterations cannot improve the order of the method Fig. 5.14b. This is mainly because the order of the time-discretization dominates the plot. If one uses higher order methods additional iterations are important to conserve the quality of approximation. Fig. 6.10d shows the convergence of the cosimulation when using a higher order Runge-Kutta method (RADAU5). The convergence order in terms of the window size increases clearly with the number of dynamic iterations (for this plot the same example was used, but the applied input signal was smoothed to benefit from the higher order of RADAU5).

Larger time windows ( $H = 2 \cdot 10^{-4} \text{ s}$ ) cause larger errors ( $> 12\%$ ) but at reduced computational costs. On the other hand simulations with  $H \leq 5 \cdot 10^{-5} \text{ s}$  require less than 70% of the computational effort of the monolithic coupling ( $h = 10^{-5} \text{ s}$ ) while being significantly more accurate, see Fig. 5.14b. For small window sizes the sweep control does not require iterations and thus the curves of both methods coincide in Fig. 5.14b, [118].

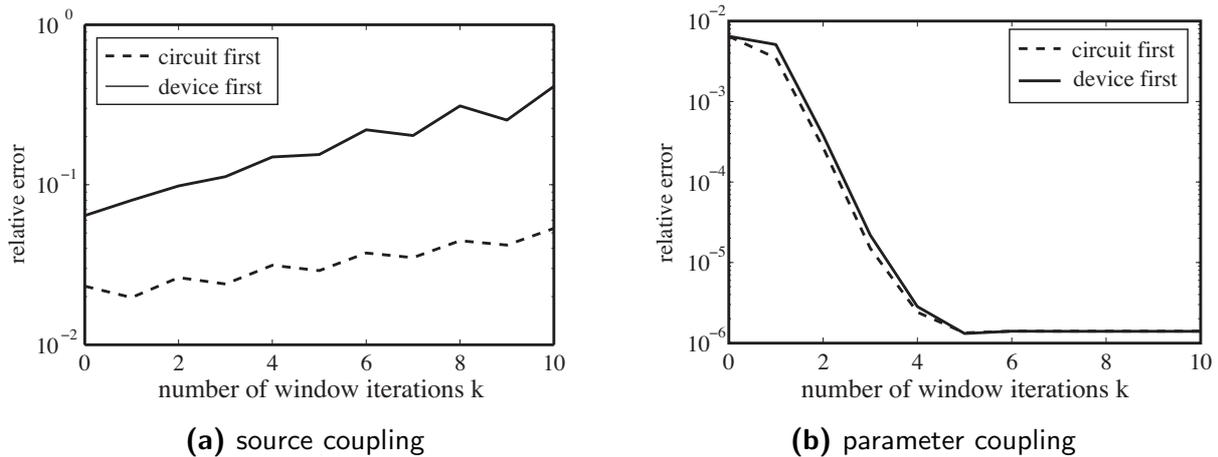
## Conclusions

Two examples underlined the convergence and stability of the dynamic iteration approach for the field/circuit coupled problem. The convergence rate was discussed and it has been shown that the scheme automatically exploits multirate behavior due to the decoupling of subproblems. The source and parameter interface were implemented. The parameter approach together with outer iterations allows for enlarged window sizes. For optimal results higher order time-integration is needed and a combined window size and sweep control will further improve the efficiency.

### 6.3.2 Semiconductor/Circuit Problem

In [9] the dynamic iteration technique is applied to the series connection of a voltage source, resistor and a block of  $N_D$  silicon  $pn$ -diodes connected in parallel (using the 1D-model described in Section 3.3). The resistance is given by a linear model with  $R = 1\Omega$ , the voltage source  $v(t) = \sin(2\pi ft)\text{V}$  is a fast oscillating sinus wave at a frequency  $f = 10^{11}$  Hz. Each diode consists of a 50nm  $n$ -region doped with doping concentration  $C_0$  and a 50nm  $p$ -region doped with  $-C_0$ , see Table 6.11B.

The example is used to underline the theoretic results of Section 5.3.6 by numerical computations. The convergence rates of both interfaces (‘source coupling’ and ‘parameter coupling’) are compared and the impact of computational sequence is discussed.



**Figure 6.12:** Contraction of semiconductor/circuit cosimulation. Splitting error of the network components in dependence of the iterations  $k$  on the time window  $[2.2, 2.3] \cdot 10^{-12}$ s, see [9].

### Contraction

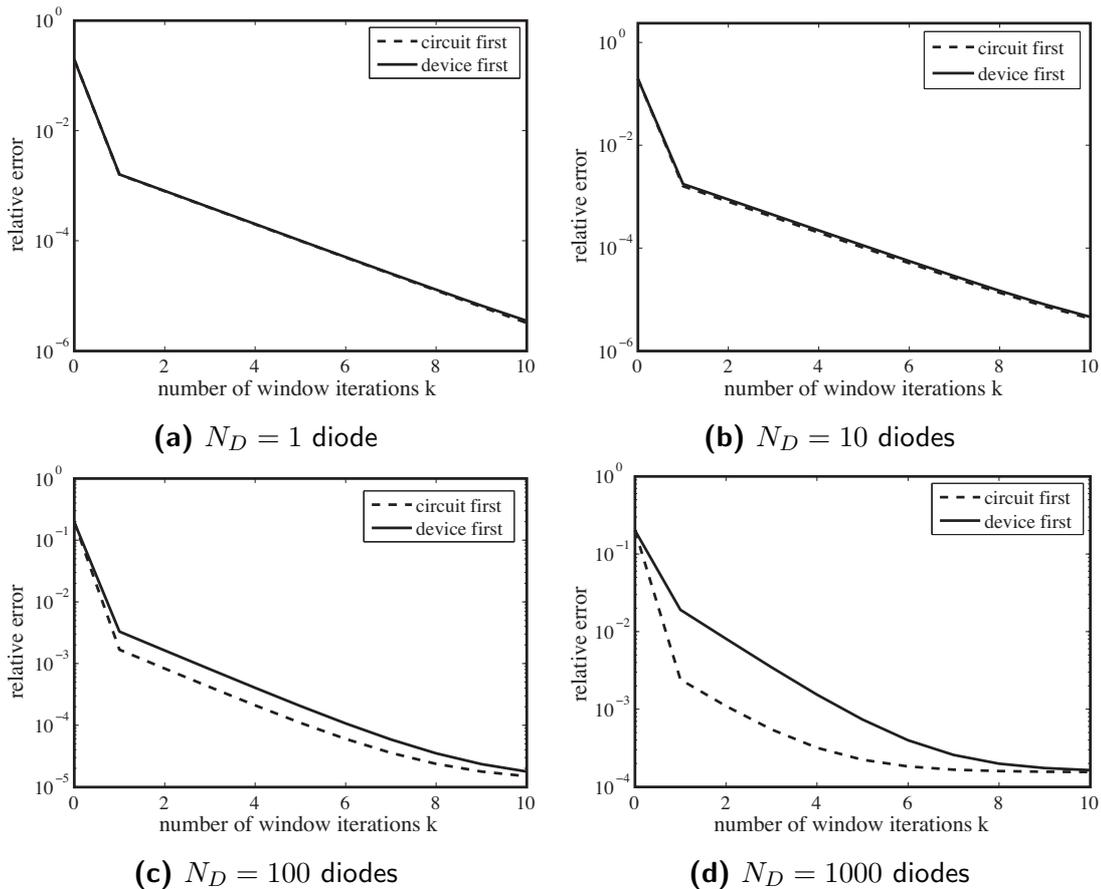
For the convergence study we analyzed the academic test example shown in Fig. 6.11a. The diode-block consists of  $N_D = 1500$  diodes. Due to their parallel connection it is sufficient to simulate a single PDE device and multiplying the output current by the number of devices.

The dynamic iteration is applied with constant extrapolation and 10 iterations per window on the time interval  $\mathcal{I} = [0, 10] \cdot 10^{-12}$ s. The underlying time-discretization was performed by the implicit Euler method. The time windows and time steps are chosen to be the same  $H = h = 0.1 \cdot 10^{-12}$ s, i.e., after each time step both subproblems are synchronized. The detailed algorithm is similar to Listing 2; it is described in more detail in [1].

Now the contraction analysis follows. The solutions of the cosimulation after 10 iterations is compared with the monolithic reference. The reference allows the verification of convergence of the dynamic iteration scheme. It is performed with same method and step size, such that the comparison shows the splitting error only (time-integration errors are neglected).

**(a) Source Coupling.** As predicted in Lemma 5.13 the dynamic iteration scheme does not converge reliably for the source coupling. The convergence depends on the contraction parameter  $\alpha$  in (5.28). In the present numerical example the amplification of the diode's current by the factor 1500 directly affects the corresponding Lipschitz constants and thus causes divergence. This is shown in Fig. 6.12a where the relative error of the network components (with respect to the monolithic reference solution) is plotted in dependence of the number of iterations for the time window  $[2.2, 2.3] \cdot 10^{-12}$ s. The dynamic iteration schemes converges (slowly) on the previous windows until it diverges on the window depicted. The same problem occurs independently of the computational sequence, i.e., the iteration for both device-first and circuit-first do not converge. The different starting values for the sequences are due to different errors on previous time windows (error propagation).

**(b) Parameter Coupling.** In the second approach the displacement current of the diode is extracted and modeled by a lumped parallel capacitance, see Definition 5.8. It

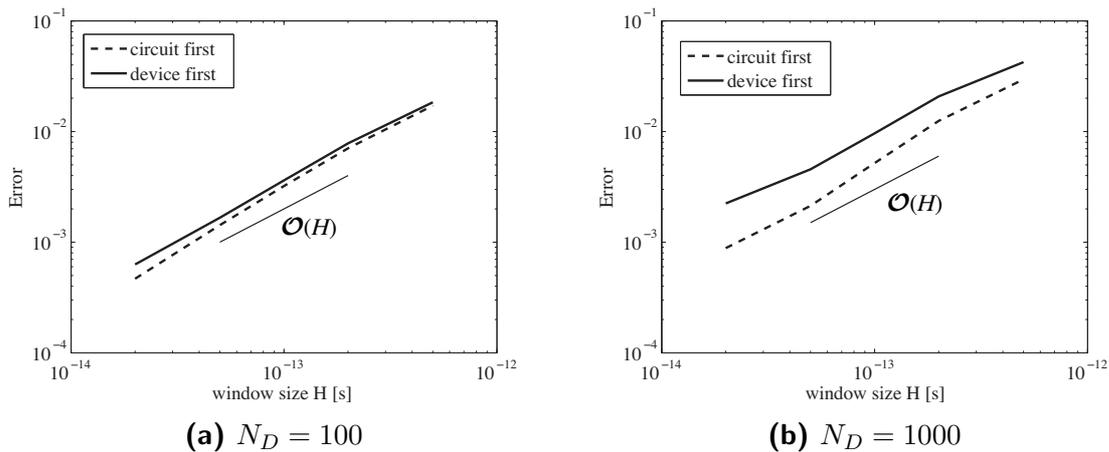


**Figure 6.13:** Contraction and Lipschitz constants. Splitting errors on the time window  $[0.4, 0.5] \cdot 10^{-12}$ s for different numbers of diodes, see [9].

is for the given example  $C_D = 10^{-17}$ F for each diode and this accumulates to a single capacitance of  $1.5 \cdot 10^{-14}$ F. Consequently the interface couples via the current  $\mathbf{i}_{SD}$ . In contrast to the source coupling approach above the parallel capacitance aids the dynamic iteration scheme. One obtains a robust algorithm that yields a sufficiently accurate solution after a few iterations, see Fig. 6.12b. The convergence plot depicts the relative error (w.r.t. the monolithic reference solution) of the network components in dependence of the number of iterations for the same interval as above, i.e.,  $[2.2, 2.3] \cdot 10^{-12}$ s. Moreover, due to significantly better convergence on the previous time windows the initial error is reduced.

### Computational Sequence

The convergence study above revealed that for the parameter coupling, i.e., interface (b), both computational sequences yield a convergent cosimulation, but at slightly different speeds, see Fig. 6.12. According to Theorem 5.15 the speed is determined by the Lipschitz constants and they can be influenced by the number of diodes  $N_D$ . More precisely the number of diodes affects the Lipschitz constant  $L_F$  from Definition 5.9. Figure 6.13 shows the corresponding convergence plots of the dynamic iteration scheme as above applied to the same problems but for varying numbers  $N_D = 1, 10, 100$  and 1000. As before the splitting error is computed with respect to the reference solution and again the time step size  $h = 0.1 \cdot 10^{-12}$ s was applied.



**Figure 6.14:** Convergence of semiconductor/circuit cosimulation. Splitting errors for different values of  $L_C$  after  $2 \cdot 10^{-12}$ s for different time step sizes  $H$  and for the circuit first and device first approach with one iteration per time window.

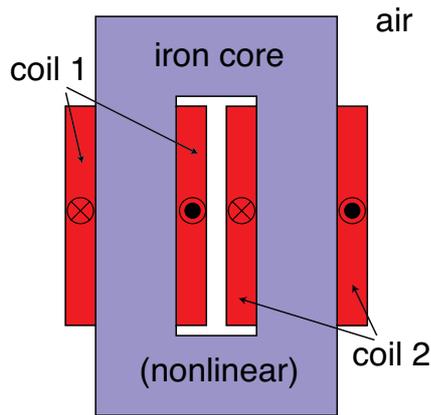
Figures 6.13a and 6.13b show that for a small number of diodes ( $N_D \leq 10$ ), i.e., with a small Lipschitz constant  $L_C$ , the speed of convergence is nearly the same for both computational sequences. However, for an increasing number of diodes ( $N_D > 10$ ) the convergence speed is clearly superior when solving the circuit subproblem first: for  $N_D = 100$  diodes the same level of accuracy is obtained with one iteration less and for  $N_D = 1000$  diodes one saves nearly two iterations. Those results reflect the effect of the leading order coefficients  $C_D$  and  $C_C$  as predicted in (5.71), cf. [9].

For further increasing values of  $N_D$  (and thus  $L_C$ ) the advantage of the device-first approach is expected to improve further on and it will require fewer iterations than the reversed order approach. This shows that a deeper knowledge on the strength of the coupling, i.e., a good estimation of the Lipschitz constants ( $L_D, L_C$ ) helps to determine an optimal sequence for solving the subsystems. Similarly it is known that reordering the computational sequence can turn a splitting scheme with contraction factor  $\alpha > 1$  into a convergent setting, see [5].

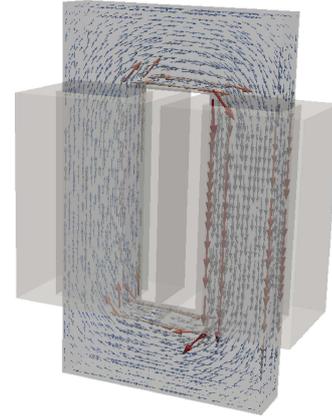
## Conclusions

Finally global convergence and stability of the parameter coupling scheme is numerically verified on the time interval  $\mathcal{I} = [0, 2] \cdot 10^{-12}$ s by performing simulations for decreasing time window sizes, i.e.,  $H = 10^{-12}, 10^{-13}, 10^{-14}$ . Simulation are computed for both sequences (circuit-first and device-first), where only one Gauß-Seidel iteration ( $k = 1$ ) is applied. Time-integration is still obtained by the implicit Euler method. On one hand the additional iterations decrease the splitting error by  $\mathcal{O}(H)$ , but on the other hand the implicit Euler method is globally first order accurate  $\mathcal{O}(h) = \mathcal{O}(H)$  (step size equals window size) and thus more iterations will not improve the total order of the scheme (when considering splitting and time-discretization errors).

The relative errors (with respect to the monolithic reference solution using the same time-stepping) are depicted in Fig. 6.14 for different window sizes. This is an order plot for the convergence of the splitting scheme. For both sequences the convergence order is approximately linear in the time window size  $H$ . Again, the larger Lipschitz-constants  $L_C$



(a) nonlinear, conductive region (iron), linear, non-conductive regions: coils and air



(b) transformer model in 3D, eddy currents at the surface of the solid iron core

material region	number of triangles	number of hexahedra
iron	619878	21120
air	61771	22997
copper	23459	14144

(c) Number of elements per material in the 2D and 3D examples

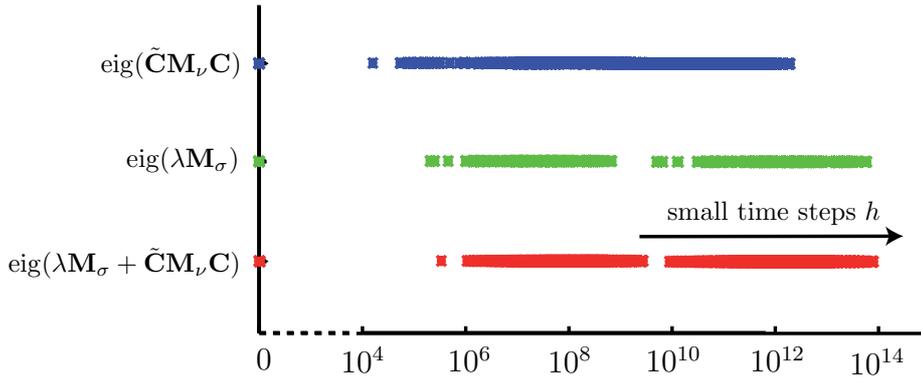
**Figure 6.15:** Domain substructuring example: (a) material regions, (b) field distribution in 3D, (c) region-wise degrees of freedom, see [36].

(for  $N_D = 1000$ ) causes the circuit-first approach to perform better than the device-first approach (although both sequences are still of the same order). This underlines the results of Theorem 5.15.

## 6.4 Domain Substructuring of a Transformer

This section discusses the numerical application of the Schur complement method given in Section 5.4. The transformer model and the analysis are taken from [36]. The example is  $30\text{cm} \times 40\text{cm} \times 6\text{cm}$  in size with 358 primary and 206 secondary copper strands wound around a (nonlinear) iron core modeled by the nonlinear material curve shown in Fig. 6.5c. The coils do not exhibit eddy currents due to strands with cross-sectional diameters below the skin depth. Consequently they are modeled by stranded conductors that are excited here by a sinusoidal currents at 20A, see Section 3.2.1. Their DoFs belong to the non-conductive domain and they are treated together with the other DoFs in the air region by the Schur complement method. Fig. 6.15a shows the linear nonconductive/nonlinear conductive partitioning. The transformer is described using two different formulations: 2D FEM and 3D FIT discretization are applied.

The 2D example is discretized by FEMM and Triangle using two meshes, with 333732 and 352990 DoFs (i.e., nodes), respectively. For the 3D example the finite integration technique was applied, it uses 174783 DoFs (edges). For simplicity all problems have been discretized by the implicit Euler method, see Section 5.1. Higher-order time-integration



**Figure 6.16:** Eigenvalues in substructuring example. The eigenvalues of the scaled conductivity matrix dominate the eigenvalues of the curl-curl matrix, time step  $h = 10^{-5}$ s, see also Fig. 5.14a and [36].

is straightforward and it is available within FIDES (using `odepkg`, see Fig. 6.1). The embedded Newton-Raphson scheme solves the linear problems directly by SuiteSparse (Cholmod, [31]) or by PCG with a Jacobi preconditioner for the original, the regularized and the Schur complement system (5.79).

The 2D and 3D problem formulations are structurally different: the block  $\mathbf{K}_{22}$  is only invertible in the 2D case where it corresponds to the discretization of the Laplacian. In the 3D FIT discretization the same block must be regularized to allow for direct solvers. A grad-div term was added in the non-conductive region as described in [36]. Furthermore it will be shown that the sparsity patterns in 2D and 3D are different and thus direct solvers behave differently. Let us start with the 2D problem.

### 6.4.1 2D Test Problem

The 2D problem has been discretized by two different meshes to illustrate the dependence on the number of DoFs. The refinement was only performed in the air region, i.e., the part that is removed by the Schur complement. Table 6.2 shows for the different methods the total simulation time, the total number of all PCG iterations and the condition number as approximated by the PCG method (averaged over all system matrices). The methods are in detail: a) Cholmod and b) PCG applied to the full system and c) the iterative solver using the Schur complement method (Listing 3) with embedded Cholesky factorization.

The results indicate that for both meshes the number of PCG iterations can be significantly reduced by using the Schur complement. For the given example this comes almost without additional costs. The forward/backward substitutions per iteration are cheap, because the factorization can preserve the sparsity (the averaged time per iteration increase slightly from 0.09s to 0.1s). On the other hand the direct approach to the whole problem is very time consuming and scales badly when increasing the number of unknowns in the air region.

### 6.4.2 3D Test Problem

The analysis of the 3D example focuses on the dependence of the time step size  $h$ . When it is decreased the gap between the largest and smallest eigenvalue increases, see Fig. 6.16.

**Table 6.2:** Substructuring of the 2D model. Time transient simulation, 10 time steps, 48 Newton iterations, fixed step-size  $h = 10^{-3}$ . ‘iterations’ denotes the total number of all CG iterations and ‘condition’ refers to the averaged condition number as approximated by PCG. ‘Cholmod’ and ‘PCG’ refer to direct/iterative solves of the full system, while ‘Schur PCG’ denotes the proposed algorithm from Listing 3 with Jacobi preconditioning, [36].

	time	iterations	condition
Cholmod	37.0min	-	-
PCG	18.7min	13186	$9.4 \cdot 10^3$
Schur PCG	6.5min	4083	$7.4 \cdot 10^2$

(a) 2D transformer with 333732 DoFs.

	time	iterations	condition
Cholmod	55.3min	-	-
PCG	21.3min	14057	$1.0 \cdot 10^4$
Schur PCG	7.1min	4163	$7.7 \cdot 10^2$

(b) 2D transformer with 352990 DoFs (refined).

**Table 6.3:** Substructuring of the 3D model. Time transient simulation, 48 steps, step-sizes  $h = 10^{-5}$  and  $10^{-4}$ ; ‘iterations’ denotes the total number of all PCG iterations and ‘condition’ refers to the averaged condition number as approximated by PCG. PCG was used to solve the curl-curl equation (‘PCG original’), its regularized version (‘PCG gauged’) and the Schur system (5.79) (‘Schur method’); regularization is carried out by an edge-wise grad-div term, [36].

	time	iterations	condition
PCG original	12.3min	13423	$1.4 \cdot 10^4$
PCG gauged	49.8min	38425	$1.2 \cdot 10^5$
Schur method	7.2min	815	$1.9 \cdot 10^1$

(a) 3D example with step size  $h = 10^{-5}$

	time	iterations	condition
PCG original	17.8min	19019	$3.0 \cdot 10^4$
PCG gauged	76.1min	56935	$1.5 \cdot 10^5$
Schur method	20.6min	2456	$1.4 \cdot 10^2$

(b) 3D example with step size  $h = 10^{-4}$

The time step pushes the eigenvalues of the conductivity matrix further along the positive axis, while the non-conductive ones remain unaltered. Thus especially in the case of tiny step sizes, the eigenvalues of the conductivity matrix dominate the spectrum.

Table 6.3 shows the total simulation time, the number of PCG iterations and the estimated condition number for (i) the original system, (ii) the system with applied grad-div regularization and finally (iii) the Schur system Listing 3. Again, the Schur complement method reduces the number of PCG iterations substantially, but for large step sizes this does not compensate for the higher computational costs due to the additional forward/backward substitutions. The sparsity pattern is not well preserved by Cholmod and thus the matrix/vector operations increase the averaged time for a single PCG iteration from 0.05s to 0.5s (other direct solvers using other ordering strategies might perform better<sup>3</sup>).

Only for small step sizes  $h \leq 10^{-5}$  the gap between the eigenvalues becomes dominant and the fewer number of iterations compensates for the additional costs per iteration. In 3D, the direct solver fails for the singular matrix pencil and when applied to the regularized system, the memory requirements are so huge that the simulation takes an unreasonable time of ten hours.

<sup>3</sup>Cholmod uses METIS for ordering, <http://www-users.cs.umn.edu/~karypis/metis>

### 6.4.3 Conclusions

The 2D and 3D computational examples have demonstrated the general feasibility of the Schur complement method. The direct linear system solvers perform very well for (small scale) 2D problems, but they become too expensive for large 3D problems. The usage of low-rank approximations, model order reduction techniques and inner iterative solvers with deflation is expected to further improve the proposed method. Combinations of the method with other more sophisticated preconditioners or multigrid solvers is the subject of current research, [115].

---

## 7 Conclusions

In this thesis multiscale models from electrical engineering, i.e., the lumped electric network, the magnetoquasistatic field device and the drift-diffusion model of a semiconductor device were derived, analyzed and mutually coupled. This culminates in a system of partial differential algebraic equations (PDAEs) that may contain device models of any dimension (from 0D to 3D) interconnected by the surrounding electric network.

For the spatial discretized PDAE system new differential-algebraic-index results in terms of the tractability concept were obtained: in particular it was shown that the field/circuit coupled problem is at most index-2 and that the index-2 variables enter the system only linearly. Thus the error propagation during time-integration is not seriously affected. Numerical test examples have underlined this result and show that the index-2 case is rather harmless but nonetheless the time-integration error can be decreased when modeling the circuit topology according to the index-1 results.

Two different classes of multirate time-integration methods for the given PDAE problem were discussed. They both follow a hierarchical multiscale approach where lumped models (for example generalized inductances) are fitted on the fly by spatially distributed PDE models (for example magnetoquasistatic Maxwell's Equations).

This is in particular accomplished by the multirate Schur complement approach in which the PDE device equations are eliminated from the circuit equations such that only a lumped model remains. The lumped model is updated according to an energy-based scheme such that the problem-specific multirate potential is exploited. Furthermore the complement makes multimethod approaches feasible, i.e., different linear solvers can be used for the subproblems. The acceleration of the time-integration is significant: a numerical example shows a  $40\times$  speed-up compared to a standard approach.

Secondly the multirate cosimulation (using dynamic iteration) was adapted for this hierarchical approach. For that purpose we introduced a parameter coupling interface that uses lumped surrogate models. This interface was shown to be superior to the common source coupling approach. Again, this unlocks multirate potential and allows even the application of different time-integrators for each subproblem – with the drawback of an increased computational overhead due to additional iterations. Nonetheless numerical examples from field/circuit coupling show a clear reduction in the computational costs compared to a monolithic single rate approach.

Furthermore, it was mathematically proved that the dynamic iteration applied to both field/circuit and semiconductor/circuit coupled problems is always convergent if the coupling interface is modeled by the parameter approach. To this end we carried out a fixed point analysis in function space and a propagation error analysis. This allowed us to compare the different interface models and we discussed their advantages and disadvantages. The underlying principles were given within an abstract framework such that this approach can be easily applied to other fields of application.

## 7.1 Outlook

This treatise raises new interesting questions and problems: the DAE-index analysis and multirate cosimulation of a coupled system consisting of all three subproblems, i.e., field, circuit and semiconductor seems to be the inevitable next step. A tentative analysis reveals that the interaction of coupling interface, number of subsystems and convergence rate offers interesting new aspects. Along with the analysis, numerical experiments should be carried out for this enlarged system.

Another interesting aspect is the relation of the order of the time-integrator and the overall cosimulation. The results in Section 6.3.1 imply that higher order cosimulation requires additional iteration. This is clear for simulations where time window and step size coincide, but what happens in the general (multirate) case? If this is well understood and one knows the approximation order of each iterate, this will allow us to predict the splitting error and consequently to derive a very efficient control for the window size and the number of iterations.

# A Discretization Properties

## A.1 Discrete Operators

### A.1.1 FIT Numbering Scheme

This section discusses some aspects of the finite integration technique (FIT) for rectilinear grids, [138]. The primary grid is given by a finite number of volumes

$$\mathcal{G} := \{V(n) := V(i_x, i_y, i_z) \mid V(i_x, i_y, i_z) = [x_{i_x}, x_{i_x+1}] \times [y_{i_y}, y_{i_y+1}] \times [z_{i_z}, z_{i_z+1}]; \\ i_x = 1, \dots, n_x - 1; i_y = 1, \dots, n_y - 1; i_z = 1, \dots, n_z - 1\},$$

The three indices  $i_x$ ,  $i_y$  and  $i_z$ , given  $n_x$ ,  $n_y$  and  $n_z \in \mathbb{N}$ , are combined into one space index, which allows us to number the objects consecutively:

$$n = n(i_x, i_y, i_z) = i_x \cdot k_x + (i_y - 1) \cdot k_y + (i_z - 1) \cdot k_z, \quad (\text{A.1})$$

with  $k_x := 1$ ,  $k_y := n_x$  and  $k_z := n_x \cdot n_y$ , such that  $n \leq n_0 := n_x \cdot n_y \cdot n_z$ .

The primary grid  $\mathcal{G}$  is accompanied by the dual grid  $\tilde{\mathcal{G}}$ , which is constructed by connecting the center points of the primary volumes, i.e., the dual points. The intersection of two primary volumes is by construction either empty, a point  $\mathbf{P}(n)$ , an edge  $\mathbf{L}_w(n)$ , a facet  $\mathbf{A}_w(n)$  or a volume  $\mathbf{V}(n)$ , where  $w \in \{x, y, z\}$  denotes the direction of the object, see Definition 2.4. The intersections in the dual grid are  $\tilde{\mathbf{P}}(n)$ ,  $\tilde{\mathbf{L}}_w(n)$ ,  $\tilde{\mathbf{A}}_w(n)$  and  $\tilde{\mathbf{V}}_w(n)$ , respectively.

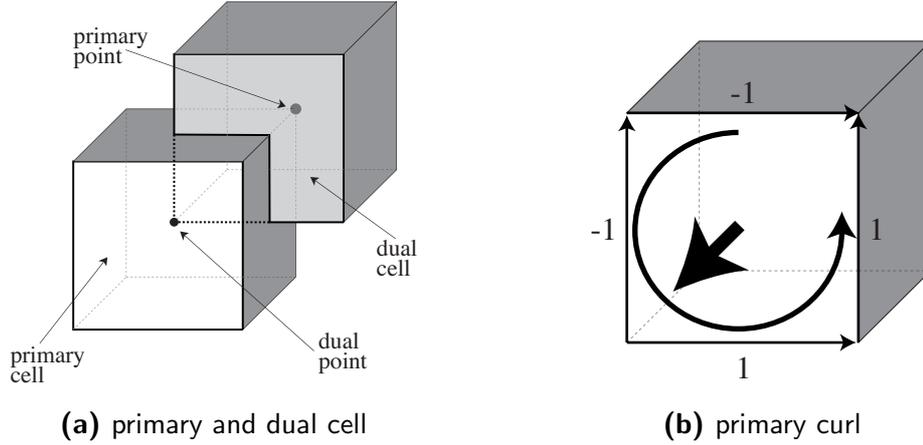
Every object (edge, facet, volume) is associated with its smallest numbered connected point  $\mathbf{P}(n)$ . Thus three edges, three facets and one volume are connected to each point. An edge  $\mathbf{L}_w(n)$  connects two in  $w$ -direction neighbored points  $\mathbf{P}(n)$  and  $\mathbf{P}(n + k_w)$  and is always directed from  $\mathbf{P}(n)$  towards  $\mathbf{P}(n + k_w)$ . A facet  $\mathbf{A}_w(n)$  is defined by  $\mathbf{P}(n)$  and the direction  $w$ , in which its normal vector points. The dual objects are numbered slightly different, e.g. a dual edge  $\tilde{\mathbf{L}}_w(n)$  points from  $\tilde{\mathbf{P}}(n - k_w)$  to  $\tilde{\mathbf{P}}(n)$ .<sup>1</sup>

In Section 2.2 each discrete field quantity was defined on an geometrical object. In FIT those objects are numbered in  $x$ ,  $y$  and  $z$  order. For example the discrete electric field strength was given as

$$\hat{\mathbf{e}}_i := \int_{L_i} \vec{E} \cdot d\vec{s} \quad \text{such that} \quad L_i = \begin{cases} \mathbf{L}_x(i) & \text{if } i \leq n_0 \\ \mathbf{L}_y(i - n_0) & \text{if } n_0 < i \leq 2n_0 \\ \mathbf{L}_z(i - 2n_0) & \text{else} \end{cases}$$

corresponds to the classical FIT numbering scheme.

<sup>1</sup>The numbering scheme as presented in this paragraph introduces phantom objects at the grid's boundary, see Section A.1.2.



**Figure A.1:** Cells and primary curl operator.

### Discrete Operators

The curl and divergence operators are discretized according to the primary and dual grid. This results in structured  $3n_0 \times 3n_0$ -matrices that reflect the continuous operators. When neglecting the effects at the boundary of the grid, the discrete curl operators read

$$\mathbf{C}_\infty = \begin{bmatrix} 0 & -\mathbf{P}_z & \mathbf{P}_y \\ \mathbf{P}_z & 0 & -\mathbf{P}_x \\ -\mathbf{P}_y & \mathbf{P}_x & 0 \end{bmatrix}, \quad \text{and} \quad \tilde{\mathbf{C}}_\infty = \mathbf{C}_\infty^\top.$$

The divergence operators are

$$\mathbf{S}_\infty = [\mathbf{P}_x \quad \mathbf{P}_y \quad \mathbf{P}_z], \quad \text{and} \quad \tilde{\mathbf{S}}_\infty = [-\mathbf{P}_x^\top \quad -\mathbf{P}_y^\top \quad -\mathbf{P}_z^\top],$$

where the partial differential operators  $\mathbf{P}_w \in \{-1, 0, 1\}^{n_0 \times n_0}$  for each spatial direction  $w \in \{x, y, z\}$  are defined as

$$(\mathbf{P}_w)_{p,q} := \delta_{p+k_w,q} - \delta_{p,q} = \begin{cases} -1 & \text{for } q = p, \\ 1 & \text{for } q = p + k_w, \\ 0 & \text{else;} \end{cases} \quad (\text{A.2})$$

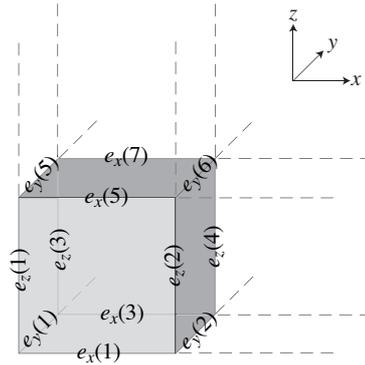
where  $\mathbf{I}$  denotes the identity matrix and  $k_w$  is defined as in (A.1). For inner edges each partial differential operator exhibits exactly two non-zeros (1 and  $-1$ ) per row. Consequently the primary curl matrix has four entries ( $-1, -1, 1, 1$ ) in each row which pick out the corresponding line-integrals from a vector that lives on the primary edge, see Fig. A.1b.

**Lemma A.1.** *Two discrete partial differential operators  $\mathbf{P}_v$  and  $\mathbf{P}_w$  with  $v, w \in \{x, y, z\}$  commute*

$$\mathbf{P}_v \mathbf{P}_w = \mathbf{P}_w \mathbf{P}_v. \quad (\text{A.3})$$

*Proof.* Let  $\mathbf{D}_{p,q} := \delta_{p+1,q}$  denote a shift matrix then  $\mathbf{P}_w = \mathbf{D}^{k_w} - \mathbf{I}$  and

$$\begin{aligned} \mathbf{P}_v \mathbf{P}_w &= (\mathbf{D}^{k_v} - \mathbf{I})(\mathbf{D}^{k_w} - \mathbf{I}) = \mathbf{D}^{k_v+k_w} - \mathbf{D}^{k_v} - \mathbf{D}^{k_w} + \mathbf{I} = (\mathbf{D}^{k_w} - \mathbf{I})(\mathbf{D}^{k_v} - \mathbf{I}) \\ &= \mathbf{P}_w \mathbf{P}_v, \end{aligned}$$



**Figure A.2:** Primary FIT grid of dimensions  $2 \times 2 \times 2$ .

which is obviously true for all  $k_v, k_w \in \{k_x, k_y, k_z\}$ .  $\square$

The result of Lemma A.1 corresponds to the interchange of partial derivatives in the continuous case. A similar proof is given for example in [139, 12].

**Theorem A.2.** *The product of the discrete divergence and the discrete curl matrices is zero*

$$\mathbf{S}_\infty \mathbf{C}_\infty = 0 \quad (\text{A.4})$$

$$\tilde{\mathbf{S}}_\infty \tilde{\mathbf{C}}_\infty = 0. \quad (\text{A.5})$$

*Proof.* Both relations are simple applications of Lemma A.1. For example we find for the operators on the primary grid

$$\mathbf{S}_\infty \mathbf{C}_\infty = [\mathbf{P}_y \mathbf{P}_z - \mathbf{P}_z \mathbf{P}_y \quad \mathbf{P}_z \mathbf{P}_x - \mathbf{P}_x \mathbf{P}_z \quad \mathbf{P}_x \mathbf{P}_y - \mathbf{P}_y \mathbf{P}_x] = 0$$

which is zero because the  $\mathbf{P}_w$  commute. The dual case is analogous. This proof and its consequences are discussed in several publications, e.g. [12].  $\square$

### A.1.2 Phantom Objects on Finite Grids

The numbering scheme introduces superfluous *phantom* objects at the boundary of the grid  $\mathcal{G}$ . Those edges, facets and volumes must be disregarded in the computations.

**Example A.1.** Let us examine the (primary) FIT grid of dimensions  $2 \times 2 \times 2$  as shown in Fig. A.2. It describes only one complete volume with 6 facets and 12 edges, but another 7 volumes, 18 facets and 12 edges are superfluously included in the numbering scheme.

#### Projectors for Partial Differential Operators

The phantom objects are always attached to those points on the boundary that are addressed by the multi-index  $n(i_x, i_y, i_z)$ , where one direction index reaches its maximum  $i_w = n_w$ , where  $w \in \{x, y, z\}$ . The following sets contain the indices of all primary points with an attached phantom edge in direction  $w$ :

$$\mathcal{H}_w = \{1 \leq n(i_x, i_y, i_z) \leq n_0 \mid i_w = n_w\}$$

We denote the set of all points connected to at least one phantom edge by  $\mathcal{H}_{xyz} := \mathcal{H}_x \cup \mathcal{H}_y \cup \mathcal{H}_z$  and define  $\mathcal{H}_{xy}$ ,  $\mathcal{H}_{xz}$ ,  $\mathcal{H}_{yz}$  accordingly. Each index set  $\mathcal{H}_w$  with  $w \in \{x, y, z\}$  gives rise to a diagonal matrix  $\mathbf{L}_w \in \{0, 1\}^{n_0 \times n_0}$  with

$$(\mathbf{L}_w)_{pq} = \begin{cases} 1 & \text{for } p = q \text{ and } p \notin \mathcal{H}_w, \\ 0 & \text{else.} \end{cases} \quad (\text{A.6})$$

**Lemma A.3.** *The matrices  $\mathbf{L}_w$  are projectors, i.e.,  $\mathbf{L}_w^2 = \mathbf{L}_w$  and have the following properties for two directions  $v, w \in \{x, y, z\}$  with  $v \neq w$*

$$\mathbf{L}_w \mathbf{L}_v = \mathbf{L}_v \mathbf{L}_w, \quad (\text{A.7})$$

$$\mathbf{L}_w \mathbf{L}_v \mathbf{P}_v = \mathbf{L}_v \mathbf{P}_v \mathbf{L}_w, \quad (\text{A.8})$$

where  $\mathbf{P}_w$  is defined as in equation (A.2) and  $\mathbf{L}_w$  as in equation (A.6).

*Proof.* The idem-potency of the projectors and their commuting is trivial, since they are diagonal matrices containing only ones and zeros. The left-hand-side of the third equation (A.8) reads

$$(\mathbf{L}_w \mathbf{L}_v \mathbf{P}_v)_{pq} = \begin{cases} -1 & \text{for } p = q \wedge p \notin \mathcal{H}_w \cup \mathcal{H}_v, \\ 1 & \text{for } q = p + k_v \wedge p \notin \mathcal{H}_w \cup \mathcal{H}_v, \\ 0 & \text{else.} \end{cases}$$

which is equivalent to the right-hand-side

$$(\mathbf{L}_w \mathbf{P}_v \mathbf{L}_v)_{pq} = \begin{cases} -1 & \text{for } p = q \wedge q \notin \mathcal{H}_w \cup \mathcal{H}_v, \\ 1 & \text{for } q = p + k_v \wedge p \notin \mathcal{H}_v \wedge q \notin \mathcal{H}_w, \\ 0 & \text{else.} \end{cases}$$

because  $p \notin \mathcal{H}_v$  implies that

$$p = i_x + i_y k_y + i_z k_z \quad \text{with} \quad i_v < n_v$$

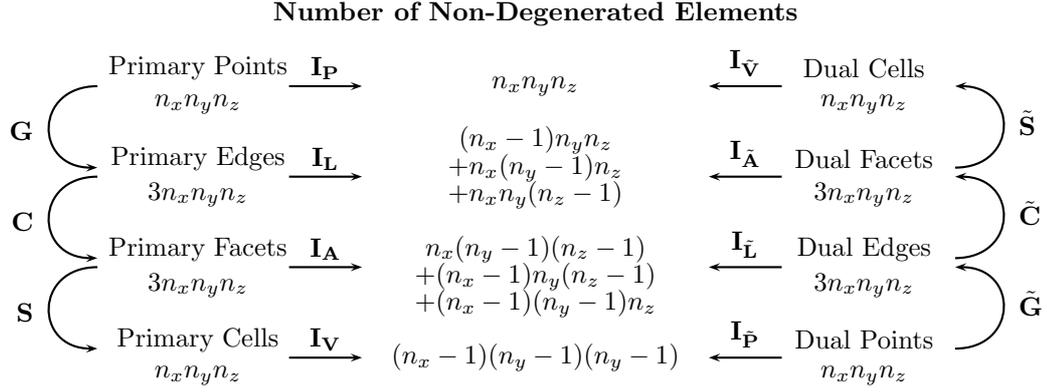
and thus  $q = p + k_v$  gives

$$q = i'_x + i'_y k_y + i'_z k_z \quad \text{with} \quad i'_v = i_v + 1 \leq n_v$$

but  $i_w = i'_w$  (for  $w \neq v$ ) and thus the condition  $p \notin \mathcal{H}_w$  is equivalent to the condition  $q = p + k_v \notin \mathcal{H}_w$  for  $w \neq v$ .  $\square$

Only the edges in the set  $\mathcal{H}_w$  exist in the grid, i.e., are degrees of freedom. We define analogously that facets and volumes exist if none of their edges are phantoms. We assemble the projectors corresponding to the index sets in  $x, y, z$ -order, such that

$$\begin{aligned} \mathbf{I}_\mathbf{P} &= \mathbf{I}_{\hat{\mathbf{V}}} := \mathbf{I} \\ \mathbf{I}_\mathbf{L} &= \mathbf{I}_{\hat{\mathbf{A}}} := \text{blkdiag}(\mathbf{L}_x, \mathbf{L}_y, \mathbf{L}_z) \\ \mathbf{I}_\mathbf{A} &= \mathbf{I}_{\hat{\mathbf{L}}} := \text{blkdiag}(\mathbf{L}_y \mathbf{L}_z, \mathbf{L}_x \mathbf{L}_z, \mathbf{L}_x \mathbf{L}_y) \\ \mathbf{I}_\mathbf{V} &= \mathbf{I}_{\hat{\mathbf{P}}} := \mathbf{L}_x \mathbf{L}_y \mathbf{L}_z \end{aligned}$$



**Figure A.3:** The number of total and non-phantom objects in Maxwell's house

where  $\mathbf{I}_P$ ,  $\mathbf{I}_L$ ,  $\mathbf{I}_A$  and  $\mathbf{I}_V$  denote the projectors for all points, edges, facets and volumes in the primary grid; the ones with a tilde are the corresponding counterparts on the dual grid, see Fig. A.3. These definitions match the matrices in [45, Section 2.2.2].

**Corollary A.4.** *The matrices  $\mathbf{I}_P$ ,  $\mathbf{I}_L$ ,  $\mathbf{I}_A$  and  $\mathbf{I}_V$  are projectors.*

### Projected Operators

The curl operator relates the edges to the flux through their facet and therefore we ignore contributions from phantom edges and facets. Using the matrices from above we can apply the curl operator on finite domains by defining (cf. [45, Section 2.2.2])

$$\mathbf{C} := \mathbf{I}_A \mathbf{C}_\infty \mathbf{I}_L \quad \text{and} \quad \tilde{\mathbf{C}} := \mathbf{I}_{\tilde{A}} \tilde{\mathbf{C}}_\infty \mathbf{I}_{\tilde{L}}. \quad (\text{A.9})$$

The divergence operators are a mapping between facets and volumes; the primary and dual operator read using the projectors

$$\mathbf{S} := \mathbf{I}_V \mathbf{S}_\infty \mathbf{I}_A \quad \text{and} \quad \tilde{\mathbf{S}} := \mathbf{I}_{\tilde{V}} \tilde{\mathbf{S}}_\infty \mathbf{I}_{\tilde{A}}. \quad (\text{A.10})$$

**Remark A.1.** The curl operator  $\mathbf{C}_\infty$  was constructed from only three partial differential matrices  $\mathbf{P}_w$  with  $w \in \{x, y, z\}$ , but the finite operator  $\mathbf{C}$  in equation (A.9) is constructed from 6 pairwise distinct blocks, e.g.

$$\mathbf{L}_y \mathbf{L}_z \mathbf{P}_y \mathbf{L}_z \neq \mathbf{L}_x \mathbf{L}_y \mathbf{P}_y \mathbf{L}_x \quad (\text{A.11})$$

and those blocks differ again from the blocks in both, the primary and dual divergence operators  $\mathbf{S}$  and  $\tilde{\mathbf{S}}$

$$\mathbf{L}_x \mathbf{L}_y \mathbf{L}_z \mathbf{P}_y \mathbf{L}_x \mathbf{L}_z \neq \mathbf{P}_y \mathbf{L}_x \mathbf{L}_y \mathbf{L}_z \quad (\text{A.12})$$

Nonetheless the blocks have a redundancy in the projectors, which is revealed by the following corollary:

**Corollary A.5.** *Let  $\mathbf{S}$  denote the primary divergence operator as defined in equation (A.10) and  $\mathbf{C}$  the primary curl operator from equation (A.9), then*

$$\mathbf{C} := \mathbf{I}_A \mathbf{C}_\infty \mathbf{I}_L = \mathbf{I}_A \mathbf{C}_\infty$$

$$\mathbf{S} := \mathbf{I}_V \mathbf{S}_\infty \mathbf{I}_A = \mathbf{I}_V \mathbf{S}_\infty$$

*Proof.* This is just a consequence of Lemma A.3. We show here only the second equation; it holds

$$\mathbf{S} = \mathbf{I}_V \mathbf{S}_\infty \mathbf{I}_A = \mathbf{L}_x \mathbf{L}_y \mathbf{L}_z [\mathbf{P}_x \mathbf{L}_y \mathbf{L}_z \quad \mathbf{P}_y \mathbf{L}_x \mathbf{L}_z \quad \mathbf{P}_z \mathbf{L}_x \mathbf{L}_y] = \mathbf{L}_x \mathbf{L}_y \mathbf{L}_z [\mathbf{P}_x \quad \mathbf{P}_y \quad \mathbf{P}_z] = \mathbf{I}_V \mathbf{S}_\infty$$

because the projectors on the left commute, so they can be reordered. Then the operators commute with the projectors on their right and finally the projectors are idem-potent.  $\square$

**Remark A.2.** Sometimes in the literature, the partial differential matrices are directly constructed for finite domains, for example

$$(\bar{\mathbf{P}}_y)_{pq} = \begin{cases} -1 & \text{for } q = p \text{ and } j \leq J - 1, \\ +1 & \text{for } q = p + k_y \text{ and } j \leq J - 1, \\ 0 & \text{else;} \end{cases}$$

with row index  $1 \leq p = p(i, j, k) \leq n_0$  and column index  $1 \leq q \leq n_0$ . This corresponds to

$$\bar{\mathbf{P}}_y = \mathbf{L}_y \mathbf{P}_y$$

in our notation and it unveils that the boundary constraints of  $\bar{\mathbf{P}}_y$  are not sufficient, compare equations (A.11) and (A.12).

### A.1.3 Degrees of Freedom in FIT

#### Properties from Vector Calculus

The validity of the fundamental operator properties Theorem A.2 on finite domains can be easily verified using the projectors, e.g. the fact that the divergence of a rotation vanishes is assured by the following theorem

**Theorem A.6.** *Let  $\mathbf{S}$  denote the primary divergence operator as defined in equation (A.10) and  $\mathbf{C}$  the primary curl operator, equation (A.9), then*

$$\begin{aligned} \mathbf{S}\mathbf{C} &= 0, \\ \tilde{\mathbf{S}}\tilde{\mathbf{C}} &= 0. \end{aligned}$$

*Proof.* We show the first equation using the results from Corollary A.5 for  $\mathbf{S}$  and  $\mathbf{C}$ , we obtain directly

$$\mathbf{S}\mathbf{C} = \mathbf{I}_V \mathbf{S}_\infty \mathbf{C}_\infty \mathbf{I}_L = 0$$

because  $\mathbf{S}_\infty \mathbf{C}_\infty = 0$  holds as a stated in Theorem A.2. The dual case can be proved analogously.  $\square$

#### Curl-Curl Matrix

The projectors apply also naturally to material matrices. In this case the consistency of the discrete problem is ensured, since material properties might be assigned to phantom

objects, e.g. if all diagonal entries of the reluctivity matrix are positive. Let us discuss in the following the curl-curl matrix as it is used in the vector-potential formulations.

**Corollary A.7.** *For the curl-curl matrix it is equivalent to either impose the boundary conditions on the material matrix only or on the curl matrices only*

$$\tilde{\mathbf{C}}\mathbf{M}_\nu\mathbf{C} = \tilde{\mathbf{C}}_\infty\mathbf{M}_\nu\mathbf{C}_\infty = \tilde{\mathbf{C}}\mathbf{M}_{\nu,\infty}\mathbf{C}.$$

*Proof.* Starting from the curl-curl matrix expressed in the infinite operators with the material matrix  $\mathbf{M}_\nu = \mathbf{I}_{\tilde{\mathbf{L}}}\mathbf{M}_{\nu,\infty}\mathbf{I}_{\mathbf{A}}$  mapping from primary facets to dual edges

$$\begin{aligned}\tilde{\mathbf{C}}\mathbf{M}_\nu\mathbf{C} &= \mathbf{I}_{\tilde{\mathbf{A}}}\tilde{\mathbf{C}}_\infty\mathbf{I}_{\tilde{\mathbf{L}}}\mathbf{I}_{\tilde{\mathbf{L}}}\mathbf{M}_{\nu,\infty}\mathbf{I}_{\mathbf{A}}\mathbf{I}_{\mathbf{A}}\mathbf{C}_\infty\mathbf{I}_{\tilde{\mathbf{L}}} = \tilde{\mathbf{C}}\mathbf{M}_{\nu,\infty}\mathbf{C} \\ &= \tilde{\mathbf{C}}_\infty\mathbf{I}_{\tilde{\mathbf{L}}}\mathbf{M}_{\nu,\infty}\mathbf{I}_{\mathbf{A}}\mathbf{C}_\infty = \tilde{\mathbf{C}}_\infty\mathbf{M}_\nu\mathbf{C}_\infty,\end{aligned}$$

which shows the equivalence in imposing the boundary condition in both the curl matrices and material matrix  $\mathbf{M}_\nu$ , or just in one of them.  $\square$

## Conclusion

The phantom edges do not belong to the degrees of freedom. It was shown that they can be disregarded in the construction of the differential operators, if the material matrices are assembled correctly.

## A.2 Material Matrices

The global material matrices are formed by looping over the contributions from each element  $\Omega_e$ , see Ass. 2.9. Mathematically speaking this assembly is carried out with the help of incidence matrices, which map the local element-wise indices to global ones. We assume for simplicity and without loss of generality that

**Assumption A.8** (Orientation). *The local orientation of objects (edges and facets) corresponds to the global one.*

Consequently the indices are mapped by the matrices

$$\mathbf{Q}_{p,e} = \left( \mathbb{1}_{\pi_{p,e}(1)} \quad \mathbb{1}_{\pi_{p,e}(2)} \quad \dots \quad \mathbb{1}_{\pi_{p,e}(m_p)} \right), \quad p \in \{0, 1, 2\}$$

where  $p$  refers to the dimension of the object,  $e$  denotes the element and  $\mathbb{1}_i$  a column vector with a 1 at the  $i$ -th position and 0 otherwise (the non-negativity follows from Ass. A.8). The length of the vectors is given by the overall number of either points ( $n_0$ ), edges ( $n_1$ ) or facets ( $n_2$ ). The local numbers are denoted by  $m_p$ , see Table A.4a. The index mappings  $\pi_{p,e}$  are injective embeddings identifying the local to its global number

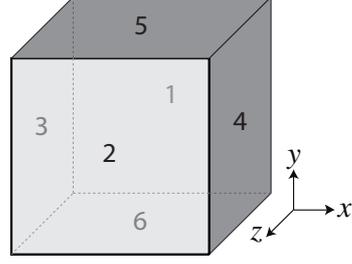
$$\pi_{p,e} : \{1, \dots, m_p\} \rightarrow \{1, \dots, n_p\} \quad \text{thus} \quad \text{rank } \mathbf{Q}_{p,e} = m_p,$$

i.e., full column rank. For example in the 3D FIT case the incidence matrix  $\mathbf{Q}_{2,e}$  for the facets (Fig. A.4b), is constructed by the index mappings

$$\pi_{2,e}(1) = e, \quad \pi_{2,e}(3) = n_1 + e, \quad \pi_{2,e}(5) = 2n_1 + e$$

Method (elements)	points	edges	facets
FIT 3D (hexahedron)	8	12	6
FEM 3D (tetrahedron)	4	6	4
FEM 2D (triangle)	3	3	1

(a) number of points, edges and facets of primary elements for first order methods



(b) facets numbering in FIT

**Figure A.4:** Local elements and numbering in FIT case.

$$\pi_{2,e}(2) = e + k_x, \quad \pi_{2,e}(4) = n_1 + e + k_y, \quad \pi_{2,e}(6) = 2n_1 + e + k_z,$$

where  $k_x$ ,  $k_y$  and  $k_z$  are defined as in (A.1) and  $n_1$  is the global number of edges.

**Assumption A.9** (No phantom objects). *It is assumed that degrees of freedom are not located on phantom objects, i.e., that each  $p$ -dimensional object belongs to one or more grid elements, see Section A.1.3*

$$\forall i_p \exists(e, j_p) \pi_{p,e}(j_p) = i_p, \quad (\text{A.13})$$

where  $e$  refers to an element,  $i_p$  and  $j_p$  are the global and local number of the same  $p$ -dimensional object.

## A.2.1 Element Matrices

In this section we focus on the construction of the anisotropic reluctivity matrix in a 3D FIT discretization. The nonlinearity is taken into account in Section A.2.3. The constructions for the other material properties and discretizations is analogous, [41].

The reluctivity relates the magnetic flux  $\widehat{\mathbf{b}}$  (2-form) to the field strength  $\widehat{\mathbf{h}}$  (1-form). In FIT both are connected at the centers of primary facets. Each element, i.e., hexahedron, in the  $i$ -th region, i.e.,  $\Omega_e \subset \Omega^{(r)}$ , contributes to the global material matrix with the  $6 \times 6$  element matrix

$$\mathbf{M}_{\nu,e}(\widehat{\mathbf{b}}_e) := \frac{1}{4} \left( \mathbf{D}_{1,e} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \nu_{\text{FIT}}^{(r)}(\widehat{\mathbf{b}}_e) \left( \mathbf{D}_{2,e}^{-1} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \quad (\text{A.14})$$

with the  $6 \times 6$  facet-wise reluctivity tensor

$$\begin{aligned} \nu_{\text{FIT}}^{(r)}(\widehat{\mathbf{b}}_e) := & \text{diag}(\nu^{(r)}(\|\mathbf{B}_e\|^2)) \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \nu^{(r)}(\|\mathbf{B}_e\|^2) \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ & - \text{diag}(\nu^{(r)}(\|\mathbf{B}_e\|^2)) \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \end{aligned}$$

where the diagonal  $3 \times 3$ -matrices  $\mathbf{D}_{1,e}$  and  $\mathbf{D}_{2,e}$  contain the primary edge lengths and primary facets areas of the element  $e$ . The local area-integrated magnetic flux densities (i.e., magnetic fluxes) are given by  $\widehat{\mathbf{b}}_e := \mathbf{Q}_{2,e}^T \widehat{\mathbf{b}}$ . Allocation of the fluxes at the facet centers and averaging the opposing ones yields the local flux density  $\mathbf{B}_e$  in the element center

$$\mathbf{B}_e := \frac{1}{2} (\mathbf{D}_{2,e}^{-1} \otimes [1 \quad 1]) \widehat{\mathbf{b}}_e \quad \text{and} \quad \|\mathbf{B}_e\|^2 = (\mathbf{B}_e)_1^2 + (\mathbf{B}_e)_2^2 + (\mathbf{B}_e)_3^2$$

is the corresponding magnitude of the magnetic flux. The definition (A.14) yields for diagonal material tensors, i.e.,  $\nu_{\text{aniso}}^{(r)} = 0$ , diagonal and positive definite matrices. In this case it coincides with the classical FIT approach. On the other hand the definition above preserves symmetry even for anisotropic materials, which is in contrast to other approaches, [99]:

**Lemma A.10.** *The matrix  $\mathbf{M}_{\nu,e}$  is symmetric positive definite if the material tensor  $\nu$  is so.*

*Proof.* For the hexahedral FIT-discretization the volume matrix of each element is given by  $\mathbf{D}_{3,e} = \mathbf{D}_{1,e}\mathbf{D}_{2,e} = |\Omega_e| \cdot \mathbf{I}$ ; rewriting  $\mathbf{M}_{\nu,e}$  in the form  $\mathbf{T}^\top \nu^{(r)} \mathbf{T}$  yields symmetry by the assumption above (with a suitable matrix  $\mathbf{T}$ ). On the other hand positive semi-definiteness follows from the positive semi-definiteness of each Kronecker product and finally the full rank of the matrix sum is proved by Sylvester's inequality (the product of both summands is zero).  $\square$

## A.2.2 Global Matrix Assembly

The material matrices for the regions are given by the sums

$$\mathbf{M}_\varepsilon^{(r)} := \sum_{\Omega_e \subset \Omega^{(r)}} \mathbf{Q}_{1,e} \mathbf{M}_{\varepsilon,e} \mathbf{Q}_{1,e}^\top, \quad \mathbf{M}_\sigma^{(r)} := \sum_{\Omega_e \subset \Omega^{(r)}} \mathbf{Q}_{1,e} \mathbf{M}_{\sigma,e} \mathbf{Q}_{1,e}^\top, \quad \mathbf{M}_\nu^{(r)}(\widehat{\mathbf{b}}) := \sum_{\Omega_e \subset \Omega^{(r)}} \mathbf{Q}_{2,e} \mathbf{M}_{\nu,e}(\widehat{\mathbf{b}}_e) \mathbf{Q}_{2,e}^\top,$$

and finally the global matrices are defined in the obvious way

$$\mathbf{M}_\varepsilon := \sum_i \mathbf{M}_\varepsilon^{(r)}, \quad \mathbf{M}_\sigma := \sum_i \mathbf{M}_\sigma^{(r)}, \quad \mathbf{M}_\nu(\widehat{\mathbf{b}}) := \sum_i \mathbf{M}_\nu^{(r)}(\widehat{\mathbf{b}}), \quad (\text{A.15})$$

with symmetric positive definite local matrices  $\mathbf{M}_{\varepsilon,e}$  and  $\mathbf{M}_{\nu,e}$  (as shown above for the reluctivity matrix). The matrix  $\mathbf{M}_{\sigma,e}$  is in general only symmetric positive semi-definite due to non-conducting areas (e.g. due to air). In the special case of FIT the construction yields diagonal global material matrices if only isotropic materials or anisotropic materials with principal directions coinciding with the grid directions are present.

In all cases the symmetry (diagonality) of the local matrices obviously carries over to the global material matrix. Even the definiteness is ensured since each object (edge, facet) is assigned at least one material parameter, mathematically speaking:

**Lemma A.11.** *The global matrix  $\mathbf{M}_\xi$  with  $\xi \in \{\varepsilon, \sigma, \nu\}$  is positive definite if all element contributions  $\mathbf{M}_{\xi,e}$  are so; it is positive semi-definite if  $\mathbf{M}_{\xi,e}$  are only semi-definite.*

*Proof.* Let  $p$  denote edges or facets depending on the material property  $\xi$  and let there be a vector  $\mathbf{x} \neq 0 \in \mathbb{R}^{N_p}$  and  $i_p$  denote the index of a non-zero degree of freedom. Then follows from (A.13) that there is an element  $e$  such that  $\mathbf{y} := \mathbf{Q}_{p,e} \mathbf{x} \neq 0$  and hence the corresponding summand is positive  $\mathbf{y}^\top \mathbf{M}_{\xi,e} \mathbf{y} > 0$ . Therefore the whole sum is positive, since all other summands are non-negative. The semi-definite case is trivial.  $\square$

## A.2.3 Nonlinearity

In real world applications the material properties depend on additional conditions for example due to nonlinear material relations as in the case of the reluctivity matrix in

(A.15). In this case the additional dependence has to be introduced into the global material and local matrices. This yields matrix-valued functions instead of the constant matrices above. The following analysis applies to nonlinear reluctivities, but also to temperature-dependent conductivities: let be given a positive scalar valued function  $f$  (e.g.,  $f = \nu_{\text{iso}}$ )

$$f : \mathbb{R}^{n_p} \rightarrow \mathbb{R}_{>0}$$

whose argument  $\alpha_e$  can be any quantity located at the points, edges, facets or centers of primary element (for example the averaged flux magnitude  $\|\mathbf{B}_e\|^2$  is located at the center). Although the function itself depends nonlinearly on the parameter, its value affects the local element matrix only affine linearly

$$\mathbf{M}_\xi(\alpha_e) := \sum_e \mathbf{Q}_{p,e} f(\alpha_e) \mathbf{M}_{\xi,e} \mathbf{Q}_{p,e}^\top. \quad (\text{A.16})$$

where  $p = \{1, 2\}$  is chosen accordingly to the material, as defined above. The parameter-dependent material matrix has the same properties as before: it is still symmetric (diagonal if applicable) and positive (semi-)definite, with a kernel that does not depend on the parameter (the image is  $\mathbb{R}_{>0}$ ).

**Lemma A.12.** *The kernel of  $\mathbf{M}_\xi(\alpha_e)$  as defined in (A.16) is constant, i.e., independently of the parameter  $\alpha_e$ .*

*Proof.* Let  $\alpha_e$  and  $\alpha_e^*$  denote two distinct parameters and  $\mathbf{x} \in \text{Ker } \mathbf{M}_\xi(\alpha_e)$ , then follows  $\text{Ker } \mathbf{M}_\xi(\alpha_e) \subset \text{Ker } \mathbf{M}_\xi(\alpha_e^*)$  from the scalar nature of  $f_e$

$$\begin{aligned} 0 &= \mathbf{x}^\top \mathbf{M}_\xi(\alpha_e) \mathbf{x} = \sum_e f_e(\alpha_e) \mathbf{x}^\top \mathbf{Q}_{p,e} \mathbf{M}_{\xi,e} \mathbf{Q}_{p,e}^\top \mathbf{x} \\ &= \sum_e f_e(\alpha_e^*) \mathbf{x}^\top \mathbf{Q}_{p,e} \mathbf{M}_{\xi,e} \mathbf{Q}_{p,e}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{M}_\xi(\alpha_e^*) \mathbf{x}, \end{aligned}$$

because all summands are necessarily zero, independently of  $f_e > 0$ . Now using the same arguments for  $\mathbf{x}^* \in \text{Ker } \mathbf{M}_\xi(\alpha_e^*)$  yields  $\text{Ker } \mathbf{M}_\xi(\alpha_e) = \text{Ker } \mathbf{M}_\xi(\alpha_e^*)$ .  $\square$

The arguments above are only valid for effects that have an isotropic (scalar valued) impact on the material and do not force the material property to vanish (e.g. switches). Typical applications are in the modeling of temperature dependent conductivities or saturation effects of the reluctivities.

### A.3 Differential Curl-Curl Matrix

The derivative of the curl-curl equation with respect to the magnetic vector potential reads

$$\begin{aligned} \mathbf{K}_\nu &= \frac{d\mathbf{k}_\nu(\widehat{\mathbf{a}})}{d\widehat{\mathbf{a}}} = \frac{d}{d\widehat{\mathbf{a}}} \left( \widetilde{\mathbf{C}} \mathbf{M}_\nu(\mathbf{C}\widehat{\mathbf{a}}) \mathbf{C}\widehat{\mathbf{a}} \right) + \mathbf{Z}_\sigma \\ &= \frac{d}{d\widehat{\mathbf{a}}} \left( \widetilde{\mathbf{C}} \mathbf{M}_\nu(\widehat{\mathbf{b}}) \widehat{\mathbf{b}} \right) + \mathbf{Z}_\sigma = \frac{d}{d\widehat{\mathbf{b}}} \left( \widetilde{\mathbf{C}} \mathbf{M}_\nu(\widehat{\mathbf{b}}) \widehat{\mathbf{b}} \right) \frac{d\widehat{\mathbf{b}}}{d\widehat{\mathbf{a}}} + \mathbf{Z}_\sigma \\ &= \widetilde{\mathbf{C}} \frac{d}{d\widehat{\mathbf{b}}} \left( \mathbf{M}_\nu(\widehat{\mathbf{b}}) \widehat{\mathbf{b}} \right) \mathbf{C} + \mathbf{Z}_\sigma \end{aligned}$$

and due to the linearity the derivative with respect to  $\widehat{\mathbf{b}}$ , it can be passed through the sums in equation (A.15) right to the nonlinear isotropic material tensor and thus

$$\nu_{\text{FIT,d}}^{(r)}(\widehat{\mathbf{b}}_e) := \nu_{\text{FIT}}^{(r)}(\widehat{\mathbf{b}}_e) + \frac{d\nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2)}{d\|\mathbf{B}_e\|^2} \left( \mathbf{D}_\nu^{(r)} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \left( \mathbf{D}_{2,e}^{-1} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \widehat{\mathbf{b}}_e \widehat{\mathbf{b}}_{e,av}^\top,$$

with the averaged flux density

$$\widehat{\mathbf{b}}_{e,av} := \left( \mathbf{D}_{2,e} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \frac{d\|\mathbf{B}_e\|^2}{d\widehat{\mathbf{b}}_e} = \left( \mathbf{I} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \mathbf{B}_e = \frac{1}{2} \left( \mathbf{D}_{2,e}^{-1} \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right) \widehat{\mathbf{b}}_e.$$

In conclusion, the differential reluctivity (and thus the differential curl-curl matrix) are assembled by using the differential reluctivity tensor  $\nu_{\text{FIT,d}}^{(r)}$  instead of the chord reluctivity tensor  $\nu_{\text{FIT}}$  in (A.14).

$$\mathbf{M}_{\nu,d,e}(\widehat{\mathbf{b}}_e) := \frac{1}{4} \left( \mathbf{D}_{1,e} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \nu_{\text{FIT,d}}^{(r)}(\widehat{\mathbf{b}}_e) \left( \mathbf{D}_{2,e}^{-1} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \quad (\text{A.17})$$

The resulting matrix can be nonsymmetric depending on the construction, [41]. However, for physical correct material curves positive definiteness can be shown.

**Corollary A.13.** *For Brauer's model the differential reluctivity matrix is positive definite.*

*Proof.* Assumption 2.8 assures a positive isotropic differential reluctivity

$$0 < \frac{\partial \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2)}{\partial \|\mathbf{B}_e\|^2} = \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2) \mathbf{I} + 2 \frac{\partial \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2)}{\partial \|\mathbf{B}_e\|^2} \|\mathbf{B}_e\|^2$$

which is the only non-trivial eigenvalue of

$$= \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2) + \frac{\partial \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2)}{\partial \|\mathbf{B}_e\|^2} \text{eig} \left( \left( \mathbf{D}_{2,e}^{-1} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \widehat{\mathbf{b}}_e \widehat{\mathbf{b}}_{e,av}^\top \right).$$

The isotropic scalar, i.e.,  $\nu_{\text{iso}}^{(r)} > 0$ , is positive by Ass. 2.8. We conclude for all eigenvalues

$$0 < \text{eig} \left( \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2) \mathbf{I} + \frac{\partial \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2)}{\partial \|\mathbf{B}_e\|^2} \left( \mathbf{D}_{2,e}^{-1} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \widehat{\mathbf{b}}_e \widehat{\mathbf{b}}_{e,av}^\top \right)$$

and the anisotropic tensor is positive definite by Ass. 2.8

$$\begin{aligned} &= \text{eig} \left( \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2) \mathbf{I} + \frac{\partial \nu_{\text{iso}}^{(r)}(\|\mathbf{B}_e\|^2)}{\partial \|\mathbf{B}_e\|^2} \left( \mathbf{D}_{2,e}^{-1} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \widehat{\mathbf{b}}_e \widehat{\mathbf{b}}_{e,av}^\top + \nu_{\text{aniso}}^{(r)} \right) \\ &= \text{eig} \frac{\partial \nu^{(r)}(\|\mathbf{B}_e\|^2)}{\partial \|\mathbf{B}_e\|^2}. \end{aligned}$$

This proves that the differential reluctivity material tensor is positive definite and thus the material matrix is positive definite by Lemma A.11.  $\square$

---

# Bibliography

- [1] Giuseppe Alì, Andreas Bartel, Markus Brunk, and Sebastian Schöps. „A convergent iteration scheme for semiconductor/circuit coupled problems“. In: *Scientific Computing in Electrical Engineering SCEE 2010*. Ed. by Bas Michielsen. Mathematics in Industry. To be published. Berlin: Springer, 2011. ISBN: 978-3-540-71979-3.
- [2] Giuseppe Alì, Andreas Bartel, and Michael Günther. „Parabolic differential-algebraic models in electrical network design“. In: *SIAM Journal on Multiscale Modeling and Simulation* 4.3 (2005), pp. 813–838. DOI: 10.1137/040610696.
- [3] Ana Alonso and Alberto Valli. „A domain decomposition approach for heterogeneous time-harmonic Maxwell equations“. In: *Computer Methods in Applied Mechanics and Engineering* 143.1-2 (1997), pp. 97–112. ISSN: 0045-7825. DOI: 10.1016/S0045-7825(96)01144-9.
- [4] Angelo Marcello Anile, Giuseppe Alì, and Giovanni Mascali, eds. *Scientific Computing in Electrical Engineering SCEE 2004*. Mathematics in Industry 9. Berlin: Springer, 2006.
- [5] Martin Arnold and Michael Günther. „Preconditioned Dynamic Iteration for Coupled Differential-Algebraic Systems“. In: *BIT Numerical Mathematics* 41.1 (2001), pp. 1–25. DOI: 10.1023/A:1021909032551.
- [6] Martin Arnold, Karl Strehmel, and Rüdiger Weiner. *Errors in the numerical solution of nonlinear differential-algebraic systems of index 2*. Tech. rep. Martin-Luther-University Halle, 1995.
- [7] Andreas Bartel. *Partial Differential-Algebraic Models in Chip Design - Thermal and Semiconductor Problems*. Fortschritt-Berichte VDI, Reihe 20. Dissertation. Düsseldorf: VDI Verlag, 2004.
- [8] Andreas Bartel, Sascha Baumanns, and Sebastian Schöps. „Structural analysis of electrical circuits including magnetoquasistatic devices“. In: *Applied Numerical Mathematics* (2010). Submitted.
- [9] Andreas Bartel, Markus Brunk, Michael Günther, and Sebastian Schöps. „Dynamic iteration schemes for coupled problems in circuit/semiconductor simulation“. In preparation. 2011.
- [10] Andreas Bartel, Michael Günther, and Anne Kværnø. „Multirate Methods in Electrical Circuit Simulation“. In: *Progress in Industrial Mathematics at ECMI 2000*. Ed. by Angelo Marcello Anile, Vincenzo Capasso, and Antonio M. Greco. Berlin: Springer, 2002, pp. 258–265.

- 
- [11] Andreas Bartel and Roland Pulch. „A Concept for Classification of Partial Differential Algebraic Equations in Nanoelectronics“. In: *Progress in Industrial Mathematics at ECMI 2006*. Ed. by Luis L. Bonilla, Jose M. Vega, Miguel Moscoso, and Gloria Platero. Berlin: Springer, 2006. ISBN: 978-3-540-71991-5. DOI: 10.1007/978-3-540-71992-2\_79.
- [12] Michael Bartsch et al. „Solution of Maxwell’s Equations“. In: *Computer Physics Communications* 73.1-3 (1992), pp. 22–39. DOI: 10.1016/0010-4655(92)90026-U.
- [13] Sascha Baumanns, Monica Selva Soto, and Caren Tischendorf. „Consistent initialization for coupled circuit-device simulation“. In: *Scientific Computing in Electrical Engineering SCEE 2008*. Ed. by Janne Roos and Luis R. J. Costa. Mathematics in Industry 14. Springer, 2010, pp. 297–304.
- [14] Gary Bedrosian. „A New Method for Coupling Finite Element Field Solutions with External Circuits and Kinematics“. In: *IEEE Transactions on Magnetics* 29.2 (1993), pp. 1664–1668. DOI: 10.1109/20.250726.
- [15] Galina Benderskaya. „Numerical Methods for Transient Field-Circuit Coupled Simulations Based on the Finite Integration Technique and a Mixed Circuit Formulation“. PhD thesis. TU Darmstadt, 2006.
- [16] Michele Benzi, Gene H. Golub, and Jörg Liesen. „Numerical solution of saddle point problems“. In: *Acta Numerica* 14.1 (2005), pp. 1–137. DOI: 10.1017/S0962492904000212.
- [17] Oszkár Bíró and Kurt Preis. „On the use of the magnetic vector potential in the finite-element analysis of three-dimensional eddy currents“. In: *IEEE Transactions on Magnetics* 25.4 (July 1989), pp. 3145–3159. DOI: 10.1109/20.34388.
- [18] Oszkár Bíró, Kurt Preis, and Kurt R. Richter. „Various FEM formulations for the calculation of transient 3D eddy currents in nonlinear media“. In: *IEEE Transactions on Magnetics* 31.3 (May 1995), pp. 1307–1312. DOI: 10.1109/20.376269.
- [19] Alain Bossavit. *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements*. San Diego: Academic Press, 1998. ISBN: 0121187101.
- [20] Alain Bossavit. „Stiff problems in eddy-current theory and the regularization of Maxwell’s equations“. In: *IEEE Transactions on Magnetics* 37.5 (Sept. 2001), pp. 3542–3545. ISSN: 0018-9464. DOI: 00189464/01\$10.00.
- [21] Alain Bossavit. „Whitney forms: a class of finite elements for three-dimensional computations in electromagnetism“. In: *IEE Proceedings* 135.8 (1988), pp. 493–500. DOI: 10.1049/ip-a-1:19880077.
- [22] Alain Bossavit and Lauri Kettunen. „Yee-like schemes on staggered cellular grids: a synthesis between FIT and FEM approaches“. In: *IEEE Transactions on Magnetics* 36.4 (July 2000), pp. 861–867. DOI: 10.1109/20.877580.
- [23] Hans Georg Brachtendorf, G. Welsch, R. Laur, and A. Bunse-Gerstner. „Numerical steady state analysis of electronic circuits driven by multi-tone signals“. In: *Electrical Engineering (Archiv für Elektrotechnik)* 79.2 (1996), pp. 103–112. DOI: 10.1007/BF01232919.

- [24] John R. Brauer and Franz Hirtenfelder. „Anisotropic materials in electromagnetic finite element analysis“. In: *International Conference on Computation in Electromagnetics, 1991* (Nov. 1991), pp. 144–147.
- [25] Kathryn Eleda Brenan, Stephen L. Campbell, and Linda Ruth Petzold. *Numerical solution of initial-value problems in differential-algebraic equations*. Philadelphia: SIAM, 1995.
- [26] Markus Brunk and Anne Kværnø. „Positivity preserving discretization of time dependent semiconductor drift-diffusion equations“. In: *Applied Numerical Mathematics* (2010). To be published.
- [27] Kevin Burrage. *Parallel and sequential methods for ordinary differential equations*. Oxford: Oxford University Press, 1995. ISBN: 0-19-853432-9.
- [28] F. Cameron, R. Piche, and K. Forsman. „Variable step size time-integration methods for transient eddy current problems“. In: *IEEE Transactions on Magnetics* 34.5 (Sept. 1998), pp. 3319–3322. ISSN: 0018-9464. DOI: 10.1109/20.717780.
- [29] Stephen L. Campbell and Charles William Gear. „The index of general nonlinear DAEs“. In: *Numerische Mathematik* 72.2 (Dec. 1995), pp. 173–196. DOI: 10.1007/s002110050165.
- [30] Qiushi Chen and A. Konrad. „A review of finite element open boundary techniques for static and quasi-static electromagnetic field problems“. In: *IEEE TransMagn* 33.1 (Jan. 1997), pp. 663–676. ISSN: 0018-9464. DOI: 10.1109/20.560095.
- [31] Yanqing Chen, Timothy A. Davis, William W. Hager, and Sivasankaran Rajamanickam. „Algorithm 887: CHOLMOD, Supernodal Sparse Cholesky Factorization and Update/Downdate“. In: *ACM Trans. Math. Softw.* 35.3 (Oct. 2008), pp. 1–14. DOI: 10.1145/1391989.1391995.
- [32] Leon O. Chua and Paul Y. Lin. *Computer aided analysis of electronic circuits*. 1st ed. Prentice-Hall series in electrical and computer engineering. Englewood Cliffs, NJ: Prentice-Hall, 1975. ISBN: 0-13-165415-2.
- [33] Philippe G. Ciarlet. *The Finite Element Method for Elliptic Problems*. 2nd. Vol. 40. Classics in applied mathematics. Society for Industrial Mathematics, Apr. 2002.
- [34] Markus Clemens. „Large systems of equations in a discrete electromagnetism: formulations and numerical algorithms“. In: *IEE Proceedings - Science, Measurement and Technology* 152.2 (Mar. 2005), pp. 50–72. DOI: 10.1049/ip-smt:20050849.
- [35] Markus Clemens, Moritz Helias, Thorsten Steinmetz, and Georg Wimmer. „Multiple right-hand side techniques for the numerical simulation of quasistatic electric and magnetic fields“. In: *Journal of Computational and Applied Mathematics* 215.2 (2008), pp. 328–338. ISSN: 0377-0427. DOI: 10.1016/j.cam.2006.04.072.
- [36] Markus Clemens, Sebastian Schöps, Herbert De Gerssem, and Andreas Bartel. „Decomposition and Regularization of Nonlinear Anisotropic Curl-Curl DAEs“. In: *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering* 30.6 (2011). To be published.
- [37] Markus Clemens and Thomas Weiland. „Regularization of eddy-current formulations using discrete grad-div operators“. In: *IEEE Transactions on Magnetics* 38.2 (Mar. 2002), pp. 569–572. DOI: 10.1109/20.996149.

- 
- [38] Herbert De Gersem, Markus Clemens, and Thomas Weiland. „Coupled finite-element, spectral-element discretisation for models with circular inclusions and far-field domains“. In: *IET Science, Measurement & Technology* 149.5 (2002), pp. 237–241. DOI: 10.1049/ip-smt:20020620.
- [39] Herbert De Gersem and Kay Hameyer. „A Finite Element Model for Foil Winding Simulation“. In: *IEEE Transactions on Magnetics* 37.5 (Sept. 2001), pp. 3472–3432. DOI: 10.1109/20.952629.
- [40] Herbert De Gersem, Kay Hameyer, and Thomas Weiland. „Field-Circuit Coupled Models in Electromagnetic Simulation“. In: *Journal of Computational and Applied Mathematics* 168.1-2 (2004), pp. 125–133. DOI: 10.1016/j.cam.2003.05.008.
- [41] Herbert De Gersem, Irina Munteanu, and Thomas Weiland. „Construction of Differential Material Matrices for the Orthogonal Finite-Integration Technique With Non-linear Materials“. In: *IEEE Transactions on Magnetics* 44.6 (June 2008), pp. 710–713. DOI: 10.1109/TMAG.2007.915819.
- [42] Herbert De Gersem, Irina Munteanu, and Thomas Weiland. „Differential material matrices for the Finite Integration Technique“. In: *The European Physical Journal-Applied Physics* 39 (2007), pp. 165–169. DOI: 10.1051/epjap:2007119.
- [43] Herbert De Gersem and Thomas Weiland. „Field-Circuit Coupling for Time-Harmonic Models Discretized by the Finite Integration Technique“. In: *IEEE Transactions on Magnetics* 40.2 (Mar. 2004), pp. 1334–1337. DOI: 10.1109/TMAG.2004.824536.
- [44] Georg Denk. „Circuit Simulation for Nanoelectronics“. In: *Scientific Computing in Electrical Engineering SCEE 2004*. Ed. by Angelo Marcello Anile, Giuseppe Ali, and Giovanni Mascali. Mathematics in Industry 9. Berlin: Springer, 2006. DOI: 10.1007/978-3-540-32862-9\_2.
- [45] Martin Dohlus. „Ein Beitrag zur numerischen Berechnung elektromagnetischer Felder im Zeitbereich“. German. Dissertation. TU Darmstadt, 1992.
- [46] Thomas Dreher and Gérard Meunier. „3D Line Current Model of Coils and External Circuits“. In: *IEEE Transactions on Magnetics* 31.3 (May 1995), pp. 1853–1856. DOI: 10.1109/20.376398.
- [47] Johan Driesen, Ronnie J. M. Belmans, and Kay Hameyer. „Methodologies for coupled transient electromagnetic-thermal finite-element modeling of electrical energy transducers“. In: *IEEE Transactions on Industry Applications* 38.5 (Sept. 2002), pp. 1244–1250. DOI: 10.1109/TIA.2002.803024.
- [48] Patrick Dular and Johan Gyselinck. „Modeling of 3-D stranded inductors with the magnetic vector potential formulation and spatially dependent turn voltages of reduced support“. In: *IEEE Transactions on Magnetics* 40.2 (Mar. 2004), pp. 1298–1301. DOI: 10.1109/TMAG.2004.825153.
- [49] Patrick Dular, François Henrotte, and Willy Legros. „A general and natural method to define circuit relations associated with magnetic vector potential formulations“. In: *IEEE Transactions on Magnetics* 35.3 (May 1999), pp. 1630–1633. DOI: 10.1109/20.767310.

- [50] Falk Ebert. *Convergence of relaxation methods for coupled systems of ODEs and DAEs*. Tech. rep. 200. MATHEON, 2004.
- [51] Falk Ebert. „On Partitioned Simulation of Electrical Circuits using Dynamic Iteration Methods“. PhD thesis. Berlin: TU Berlin, 2008.
- [52] Jocelyne Erhel and Frederic Guyomarc’h. „An Augmented Conjugate Gradient Method for Solving Consecutive Symmetric Positive Definite Linear Systems“. In: *SIAM Journal on Matrix Analysis and Applications* 21.4 (2000), pp. 1279–1299. DOI: 10.1137/S0895479897330194.
- [53] Diana Estévez Schwarz. „Consistent initialization for index-2 differential algebraic equations and its application to circuit simulation“. PhD thesis. Berlin: Humboldt-Universität, 2000.
- [54] Diana Estévez Schwarz and Caren Tischendorf. „Structural analysis of electric circuits and consequences for MNA“. In: *International Journal of Circuit Theory and Applications* 28.2 (2000), pp. 131–162. DOI: 10.1002/(SICI)1097-007X(200003/04)28:2<131::AID-CTA100>3.0.CO;2-W.
- [55] Carlo de Falco. *The IFF File Format Specification*. CoMSON Project. Wuppertal, 2008.
- [56] Carlo de Falco, Georg Denk, and Reinhart Schultz. „A Demonstrator Platform for Coupled Multiscale Simulation“. In: *Scientific Computing in Electrical Engineering SCEE 2006*. Ed. by Gabriela Ciuprina and Daniel Ioan. Mathematics in Industry 11. Berlin: Springer, 2007, pp. 63–71. ISBN: 978-3-540-71979-3. DOI: 10.1007/978-3-540-71980-9\_5.
- [57] Uwe Feldmann and Michael Günther. „CAD-based electric-circuit modeling in industry I: mathematical structure and index of network equations“. In: *Surveys on Mathematics for Industry* 8.2 (1999), pp. 97–129.
- [58] Uwe Feldmann, Masataka Miyake, Takahiro Kajiwara, and Mitiko Miura-Mattausch. „On Local Handling of Inner Equations in Compact Models“. In: *Scientific Computing in Electrical Engineering SCEE 2008*. Ed. by Janne Roos and Luis R. J. Costa. Mathematics in Industry 14. Springer, 2010, pp. 143–150.
- [59] Martin J. Gander and Albert E. Ruehli. „Optimized waveform relaxation methods for RC type circuits“. In: *IEEE Transactions on Circuits and Systems* 51.4 (Apr. 2004), pp. 755–768. ISSN: 1549-8328. DOI: 10.1109/TCSI.2004.826193.
- [60] Charles William Gear and D. R. Wells. „Multirate Linear Multistep Methods“. In: *BIT Numerical Mathematics* 24.4 (1984), pp. 484–502. DOI: 10.1007/BF01934907.
- [61] Gene H. Golub and Charles F. van Loan. *Matrix Computations*. 3rd ed. The Johns Hopkins University Press, 1996.
- [62] Klaus-Tibor Grasser. „Mixed-Mode Device Simulation“. PhD thesis. TU Wien, 1999.
- [63] Eberhard Griepentrog and Roswitha März. *Differential-Algebraic Equations and Their Numerical Treatment*. Teubner, Leipzig, 1986.
- [64] Michael Günther and Uwe Feldmann. „The DAE-Index in Electric Circuit Simulation“. In: *Mathematics and Computers in Simulation* 39 (Nov. 1995), pp. 573–582.

- 
- [65] Michael Günther, Uwe Feldmann, and Jan ter Maten. In: *Numerical Methods in Electromagnetics*. Ed. by Wil Schilders, Philippe Ciarlet, and Jan ter Maten. Vol. 13. Handbook of Numerical Analysis. North-Holland, 2005. Chap. Modelling and Discretization of Circuit Problems. ISBN: 978-0-444-51375-5.
- [66] Wolfgang Hackbusch. „Direct Domain Decomposition using the Hierarchical Matrix Technique“. In: *Fourteenth International Conference on Domain Decomposition Methods*. Ed. by Ismael Herrera, David E. Keyes, Olof B. Widlund, and Robert Yates. 2003.
- [67] Ernst Hairer, Christian Lubich, and Michel Roche. *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*. Lecture Notes in Mathematics. Berlin: Springer, 1989. ISBN: ISBN 3-540-51860-6. DOI: 10.1007/BFb0093947.
- [68] Ernst Hairer, Syvert P. Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. 2nd ed. Springer Series in Computational Mathematics. Berlin: Springer, 2002. ISBN: 978-3-540-604525-0.
- [69] Ernst Hairer, Syvert P. Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. 2nd ed. Springer Series in Computational Mathematics. Berlin: Springer, 2000. ISBN: 978-3-540-56670-0.
- [70] A.Y. Hannalla and D.C. MacDonald. „Numerical analysis of transient field problems in electrical machines“. In: *Proceedings of the Institution of Electrical Engineers* 123.9 (Sept. 1976), pp. 893–898. ISSN: 0020-3270. DOI: 10.1049/piee.1976.0190.
- [71] Hermann A. Haus and James R. Melcher. *Electromagnetic Fields and Energy*. Prentice Hall, 1989. ISBN: 013249020X.
- [72] Bodo Heise. „Analysis of a Fully Discrete Finite Element Method for a Nonlinear Magnetic Field Problem“. In: *SIAM Journal on Numerical Analysis* 31.3 (1994), pp. 745–759.
- [73] Ralf Hiptmair and Oliver Sterz. „Current and voltage excitations for the eddy current model“. In: *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* 18.1 (2005), pp. 1–21. DOI: 10.1002/jnm.555.
- [74] Chung-Wen Ho, Albert E. Ruehli, and Pierce A. Brennan. „The modified nodal approach to network analysis“. In: *IEEE Transactions on Circuits and Systems* 22.6 (June 1975), pp. 504–509. ISSN: 0098-4094. DOI: 10.1109/TCS.1975.1084079.
- [75] Daniel Ioan and Irina Munteanu. „Missing link rediscovered: the electromagnetic circuit element concept“. In: *The Japan Society of Applied Electromagnetics and Mechanics* 8 (1998), pp. 302–320.
- [76] Zdzislaw Jackiewicz and Marian Kwapisz. „Convergence of Waveform Relaxation Methods for Differential-Algebraic Systems“. In: *SIAM Journal on Numerical Analysis* 33.6 (Dec. 1996), pp. 2303–2317. DOI: 10.1137/S0036142992233098.
- [77] Jan Janssen and Stefan Vandewalle. „On SOR Waveform Relaxation Methods“. In: *SIAM Journal on Numerical Analysis* 34 (1997), pp. 2456–2481.
- [78] Enrique Francisco Kaasschieter. „Preconditioned conjugate gradients for solving singular systems“. In: *Journal of Computational and Applied Mathematics* 24.1-2 (1988), pp. 265–275. DOI: 10.1016/0377-0427(88)90358-5.

- [79] Manfred Kaltenbacher. *Numerical Simulation of Mechatronic Sensors and Actuators*. 2nd ed. Berlin: Springer, 2007. ISBN: 978-3-540-71359-3.
- [80] Sami Kanerva. „Data transfer methodology between a FEM program and a system simulator“. In: *Proceedings of the Fifth International Conference on Electrical Machines and Systems, 2001. ICEMS 2001*. Vol. 2. Aug. 2001, pp. 1121–1124.
- [81] Lauri Kettunen, Kimmo Forsman, and Alain Bossavit. „Gauging in Whitney spaces“. In: *IEEE Transactions on Magnetics* 35.3 (May 1999), pp. 1466–1469. DOI: 10.1109/20.767243.
- [82] Adalbert Konrad. „Integrodifferential finite element formulation of two-dimensional steady-state skin effect problems“. In: *IEEE Transactions on Magnetics* 18.1 (Jan. 1982), pp. 284–292. DOI: 10.1109/TMAG.1982.1061775.
- [83] Bernd Krietenstein, Rolf Schuhmann, Peter Thoma, and Thomas Weiland. „The Perfect Boundary Approximation Technique Facing the Big Challenge of High Precision Field Computation“. In: *Proceedings of the XIX International Linear Accelerator Conference*. 1998, pp. 860–862.
- [84] Stefan Kurz and Volker Rischmüller. „Field-circuit coupling by means of the Woodbury formula“. In: *Scientific Computing in Electrical Engineering SCEE 2004*. Ed. by Angelo Marcello Anile, Giuseppe Ali, and Giovanni Mascali. Mathematics in Industry 9. Berlin: Springer, 2004, pp. 281–288.
- [85] Enno Lange, François Henrotte, and Kay Hameyer. „A Circuit Coupling Method Based on a Temporary Linearization of the Energy Balance of the Finite Element Model“. In: *IEEE Transactions on Magnetics* 44.6 (2008), pp. 838–841. DOI: 10.1109/TMAG.2007.916138.
- [86] Enno Lange, François Henrotte, and Kay Hameyer. „An Efficient Field-Circuit Coupling Based on a Temporary Linearization of FE Electrical Machine Models“. In: *IEEE Transactions on Magnetics* 45.3 (2009), pp. 1258–1261. ISSN: 0018-9464. DOI: 10.1109/TMAG.2009.2012585.
- [87] Ulrich Langer. *Boundary and Finite Element Domain Decomposition Methods*. Ed. by Karl Kunisch, Günther Of, and Olaf Steinbach. Springer Berlin Heidelberg, 2008. ISBN: 978-3-540-69777-0. DOI: 10.1007/978-3-540-69777-0\_2.
- [88] Ekachai Lelarasme. „The Waveform Relaxation Method for Time Domain Analysis of Large Scale Integrated Circuits: Theory and Applications“. PhD thesis. EECS Department, University of California, Berkeley, 1982.
- [89] Konstantin Lipnikov and Peter Deuffhard. „Domain Decomposition with Subdomain CCG for Material Jump Elliptic Problems“. In: *East-West J. Numer. Anal* 6 (1997), pp. 81–100.
- [90] Luisa Donatella Marini and Paola Pietra. „New mixed finite element schemes for current continuity equations“. In: *COMPEL* 9 (1990), pp. 257–268. DOI: 10.1108/eb010080.
- [91] Roswitha März. „Differential Algebraic Systems with Properly Stated Leading Term and MNA Equations“. In: *Modelling, Simulation and Optimization of Integrated Circuits*. Ed. by Kurt Anstreich, Roland Bulirsch, Albert Gilg, and Peter Rentrop. Berlin: Birkhäuser, 2003, pp. 135–151.

- 
- [92] Roswitha März. „Solvability of linear differential algebraic equations with properly stated leading terms“. In: *Results in Mathematics* 45.1-2 (2004). Preprint, pp. 88–105.
- [93] James Clerk Maxwell. „A Dynamical Theory of the Electromagnetic Field“. In: *Royal Society Transactions CLV* (1864), pp. 459–512.
- [94] David Meeker. *Finite Element Method Magnetics User’s Manual*. Version 4.2 (09Nov2010 Build). 2010.
- [95] Ronny Mertens et al. „An algebraic multigrid method for solving very large electromagnetic systems“. In: *IEEE Transactions on Magnetics* 34.5 (Sept. 1998), pp. 3327–3330. DOI: 10.1109/20.717782.
- [96] Bas Michielsen, ed. *Scientific Computing in Electrical Engineering SCEE 2010*. Mathematics in Industry. To be published. Berlin: Springer, 2011. ISBN: 978-3-540-71979-3.
- [97] Ulla Miekka and Olavi Nevanlinna. „Convergence of dynamic iteration methods for initial value problems“. In: *SIAM Journal on Scientific and Statistical Computing* 8 (July 1987), pp. 459–482. DOI: 10.1137/0908046.
- [98] Kenneth Moreland. *The ParaView Tutorial*. 3.8. Sandia National Laboratories. 2010.
- [99] Victor C. Motrescu and Ursula van Rienen. „Computation of electrostatic fields in anisotropic human tissues using the Finite Integration Technique (FIT)“. In: *Advances in Radio Science* 2 (2004), pp. 309–313. DOI: 10.5194/ars-2-309-2004.
- [100] Takayoshi Nakata et al. „Improvements of convergence characteristics of Newton-Raphson method for nonlinear magnetic field analysis“. In: *IEEE Transactions on Magnetics* 28.2 (Mar. 1992), pp. 1048–1051. ISSN: 0018-9464. DOI: 10.1109/20.123861.
- [101] André Nicolet and François Delincé. „Implicit Runge-Kutta Methods for Transient Magnetic Field Computation“. In: *IEEE Transactions on Magnetics* 32.3 (May 1996), pp. 1405–1408. DOI: 10.1109/20.497510.
- [102] Kurt Preis, H. Stogner, and Kurt R. Richter. „Finite element analysis of anisotropic and nonlinear magnetic circuits“. In: *IEEE Transactions on Magnetics* 17.6 (Nov. 1981), pp. 3396–3398. DOI: 10.1109/TMAG.1981.1061469.
- [103] Alfio Quarteroni. *Numerical Models for Differential Problems*. Modeling, Simulation and Applications. Heidelberg: Springer, 2009. ISBN: 978-88-470-1070-3. DOI: 10.1007/978-88-470-1071-0.
- [104] Alfio Quarteroni and Alberto Valli. *Domain Decomposition Methods for Partial Differential Equations*. Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press, 1999.
- [105] Alfio Quarteroni and Alberto Valli. *Numerical Approximation of Partial Differential Equations*. Vol. 23. Springer Series in Computational Mathematics. Berlin: Springer, 2008.
- [106] Muruhan Rathinam, Linda, and Linda R. Petzold. „Dynamic Iteration Using Reduced Order Models: A Method For Simulation Of Large Scale Modular Systems“. In: *SIAM Journal on Numerical Analysis* 40 (2002), pp. 1446–1474.

- [107] Stefan Reitzinger and Joachim Schöberl. „An algebraic multigrid method for finite element discretizations with edge elements“. In: *Numerical Linear Algebra with Applications* 9.3 (2002), pp. 223–238. DOI: 10.1002/nla.271.
- [108] Ricardo Riaza. *Differential-Algebraic Systems: Analytical Aspects and Circuit Applications*. World Scientific, 2008.
- [109] Janne Roos and Luis R. J. Costa, eds. *Scientific Computing in Electrical Engineering SCEE 2008*. Mathematics in Industry 14. Springer, 2010.
- [110] Yousef Saad. *Iterative methods for sparse linear systems*. 2nd ed. Boston: SIAM, 2003. ISBN: 978-0898715347.
- [111] Sheppard J. Salon. *Finite Element Analysis of Electrical Machines*. Kluwer, 1995.
- [112] Valeriu Savcenko, Willem Hundsdorfer, and Jan G. Verwer. *A multirate time stepping strategy for parabolic PDE*. Tech. rep. MAS-E0516. Amsterdam: Centrum voor Wiskunde en Informatica, 2005.
- [113] Sebastian Schöps. „Coupling and Simulation of Lumped Electric Circuits Refined by 3-D Magnetoquasistatic Conductor Models Using MNA and FIT“. MA thesis. Wuppertal: Bergische Universität, 2008.
- [114] Sebastian Schöps. *Field Device Simulator User’s Manual*. 2009-08-31. 2009.
- [115] Sebastian Schöps, Andreas Bartel, and Markus Clemens. „Efficient Solution of Transient Eddy Current Problems Using Linear/Nonlinear Domain Substructuring“. In preparation. 2011.
- [116] Sebastian Schöps, Andreas Bartel, and Herbert De Gersem. „Multirate Time-Integration of Field/Circuit Coupled Problems by Schur Complements“. In: *Scientific Computing in Electrical Engineering SCEE 2010*. Ed. by Bas Michielsen. Mathematics in Industry. To be published. Berlin: Springer, 2011. ISBN: 978-3-540-71979-3.
- [117] Sebastian Schöps, Andreas Bartel, Herbert De Gersem, and Michael Günther. „DAE-Index and Convergence Analysis of Lumped Electric Circuits Refined by 3-D MQS Conductor Models“. In: *Scientific Computing in Electrical Engineering SCEE 2008*. Ed. by Janne Roos and Luis R. J. Costa. Mathematics in Industry 14. Springer, 2010, pp. 341–350. DOI: 10.1007/978-3-642-12294-1.
- [118] Sebastian Schöps, Herbert De Gersem, and Andreas Bartel. „A Cosimulation Framework for Multirate Time-Integration of Field/Circuit Coupled Problems“. In: *IEEE Transactions on Magnetics* 46.8 (2010), pp. 3233–3236. DOI: 10.1109/TMAG.2010.2045156.
- [119] Siegfried Selberherr. *Analysis and Simulation of Semiconductor Devices*. Springer, 1984.
- [120] Monica Selva Soto and Caren Tischendorf. „Numerical Analysis of DAEs from Coupled Circuit and Semiconductor Simulation“. In: *Applied Numerical Mathematics* 53.2–4 (2005), pp. 471–488. DOI: 10.1016/j.apnum.2004.08.009.

- 
- [121] Jonathan Richard Shewchuk. „Triangle: Engineering a 2D Quality Mesh Generator and Delaunay triangulator“. In: *Applied Computational Geometry: Towards Geometric Engineering*. Ed. by Ming C. Lin and Dinesh Manocha. Vol. 1148. Lecture Notes in Computer Science. From the First ACM Workshop on Applied Computational Geometry. Springer-Verlag, May 1996, pp. 203–222.
- [122] Abraham van der Sluis and Henk A. van der Vorst. „The rate of convergence of Conjugate Gradients“. In: *Numerische Mathematik* 48.5 (1986), pp. 543–560. DOI: 10.1007/BF01389450.
- [123] R. Stratton. „Diffusion of hot and cold electrons in semiconductor barriers“. In: *Physical Review* 126.6 (1962), pp. 2002–2014. DOI: 10.1103/PhysRev.126.2002.
- [124] Michael Striebel. *Hierarchical Mixed Multirating for Distributed Integration of DAE Network Equations in Chip Design*. Fortschritt-Berichte VDI Reihe 20. Dissertation. Düsseldorf: VDI Verlag, 2006.
- [125] Michael Striebel and Michael Günther. *A Charge Oriented Mixed Multirate Method for a special Class of Index-1 Network Equations in Chip Design*. Preprint BUW-AMNA 04/01. Bergische Universität Wuppertal, 2004.
- [126] Peter Thoma and Thomas Weiland. „A Consistent Subgridding Scheme for the Finite Difference Time Domain Method“. In: *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* 9.5 (1996), pp. 359–374. DOI: 10.1002/(SICI)1099-1204(199609)9:5<359::AID-JNM245>3.0.CO;2-A.
- [127] Caren Tischendorf. „Coupled Systems of Differential Algebraic and Partial Differential Equations in Circuit and Device Simulation“. Habil. Berlin: Humboldt-Universität, June 2003.
- [128] Caren Tischendorf. „Model Design Criteria for Integrated Circuits to Have a Unique Solution and Good Numerical Properties“. In: *Scientific Computing in Electrical Engineering SCEE 2000*. Ed. by Ursula van Rienen, Michael Günther, and Dirk Hecht. Mathematics in Industry. Berlin: Springer, 2001, pp. 179–198.
- [129] Enzo Tonti. „On the Geometrical Structure of Electromagnetism“. In: *Gravitation, Electromagnetism and Geometrical Structures*. Ed. by G. Ferrarese. Bologna: Pitagora Editrice, 1995, pp. 281–308.
- [130] Andrea Toselli, Olof B. Widlund, and Barbara I. Wohlmuth. „An Iterative Substructuring Method for Maxwell’s Equations in Two Dimensions“. In: *Mathematics of Computation* 70.235 (2001), pp. 935–949. ISSN: 00255718.
- [131] Igor A. Tsukerman. „Finite Element Differential-Algebraic Systems for Eddy Current Problems“. In: *Numerical Algorithms* 31.1 (Dec. 2002), pp. 319–335. DOI: 10.1023/A:1021112107163.
- [132] Igor A. Tsukerman, Adalbert Konrad, Gary Bedrosian, and M. V. K. Chari. „A Survey of Numerical Methods for Transient Eddy Current Problems“. In: *IEEE Transactions on Magnetics* 29.2 (Mar. 1993), pp. 1711–1716. DOI: 10.1109/20.250735.
- [133] Igor A. Tsukerman, Adalbert Konrad, Gérard Meunier, and Jean Claude Sabonnadière. „Coupled Field-Circuit Problems: Trends and Accomplishments“. In: *IEEE Transactions on Magnetics* 29.2 (Mar. 1993), pp. 1701–1704. DOI: 10.1109/20.250733.

- [134] Janne Vaananen. „Circuit Theoretical Approach to Couple Two-Dimensional Finite Element with External Circuit Equations“. In: *IEEE Transactions on Magnetics* 32.2 (1996), pp. 400–410. DOI: 10.1109/20.486524.
- [135] Stefan Vandewalle. *Parallel multigrid waveform relaxation for parabolic problems*. Teubner Skripten zur Numerik. Stuttgart: Teubner, 1993.
- [136] Prashanth K. Vijalapura, John Strain, and Sanjay Govindjee. „Fractional step methods for index-1 differential-algebraic equations“. In: *Journal of Computational Physics* 203.1 (2005), pp. 305–320. DOI: 10.1016/j.jcp.2004.08.015.
- [137] Jian-She Wang. „A Nodal Analysis Approach for 2D and 3D Magnetic-Circuit Coupled Problems“. In: *IEEE Transactions on Magnetics* 32.3 (May 1996), pp. 1074–1077. DOI: 10.1109/20.497427.
- [138] Thomas Weiland. „A discretization model for the solution of Maxwell’s equations for six-component fields“. In: *International Journal of Electronics and Communications* 31 (Mar. 1977), pp. 116–120.
- [139] Thomas Weiland. „On the unique numerical solution of Maxwellian eigenvalue problems in three dimensions“. In: *Particle Accelerators* 17.227–242 (1985).
- [140] Thomas Weiland. „Time Domain Electromagnetic Field Computation with Finite Difference Methods“. In: *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* 9.4 (1996), pp. 295–319. DOI: 10.1002/(SICI)1099-1204(199607)9:4<295::AID-JNM240>3.0.CO;2-8.
- [141] Jacob K. White, F. Odeh, Alberto L. Sangiovanni-Vincentelli, and Albert E. Ruehli. *Waveform Relaxation: Theory and Practice*. Tech. rep. UCB/ERL M85/65. Berkley: EECS Department, University of California, 1985.
- [142] Georg Wimmer, Thorsten Steinmetz, and Markus Clemens. „Reuse, Recycle, Reduce (3R) – strategies for the calculation of transient magnetic fields“. In: *Applied Numerical Mathematics* 59.3-4 (2008), pp. 830–844. DOI: 10.1016/j.apnum.2008.03.015.
- [143] Ping Zhou et al. „A General Co-Simulation Approach for Coupled Field-Circuit Problems“. In: *IEEE Transactions on Magnetics* 42.4 (Apr. 2006), pp. 1051–1054. DOI: 10.1109/TMAG.2006.871374.

# Index

- algebraic variable, 47
- assembly
  - circuit simulator, 41
  - finite elements, 113
- backward differentiation formulas, **41**
- BDF, *see* backward differentiation formulas
- boundary conditions, 6, 13
- Brauer's model, **8**, 115
- bypassing, **44**
- capacitance extraction, 27, 73
- circuit analysis
  - cutset, **19**
  - CV-loop, 32, 74
  - LIM-cutset, 32
  - loop, **19**
- conductivity, **5**, 9, 11
  - matrix, 11, 25
  - tensor, 9
- conductor model
  - foil, 26
  - solid, **20**, 25
  - stranded, **25**
- cosimulation, **46**, 72, 90, 91, *see also* dynamic iteration
  - field/circuit, 91
  - semiconductor/circuit, 72
- Coulomb's gauge, *see* regularization
- coupling
  - DAEs, 47
  - field/circuit, 20
  - semiconductor/circuit, 26
- coupling interface
  - field/circuit, 64, 92
  - semiconductor/circuit, 72, 96
- coupling matrix, **22**, 25
  - full, **22**
  - sparse, **22**
- curl matrix, 10, **106**
- curl-curl equation, **6**, 12
  - discrete, 12, 64
- curl-curl matrix, 12
  - differential, 114
- cutset, *see* circuit analysis
- CV-loop, *see* circuit analysis
- differential algebraic equations
  - index, **29**, 48, *see also* index
  - index reduction, 80
  - monolithic index, 47
  - MQS index, **30**
- tractability index, 29
- differential index, *see* index
- differential variable, 47
- divergence matrix, 10, **106**
- domain decomposition, 78
  - linear/nonlinear, 79
  - Schur complement, 35, 78, **80**
  - substructuring, **78**
- doping concentration, 27
- drift-diffusion model, 26
- dynamic iteration, 49
  - extrapolation step, 49
  - field/circuit, **63**, 91
  - Gauß-Seidel, **51**, 58, 67, 74
  - iteration step, 50
  - recursion, 53
  - semiconductor/circuit, 72
  - simple coupling, 57
- eigenvalues
  - dynamic iteration, 57
  - gauging, 13
  - PCG, 81
  - substructuring, 100
- electric circuit, *see* network

- electric scalar potential, **5**, 12, 21, 26
  - discrete, **12**, 22, 27, 72
- electron density, 27
  - discrete, 27, 72
- electrostatic, *see* static
- elementary charge, 27
- energy estimation, 44, 69
- EQS, *see* quasistatic
- extrapolation step, *see* dynamic iteration
- field device
  - conductor, *see* conductor model
  - modeling, 20
- finite elements, **9**, 112
  - element definition, 9
  - material matrices, 112
- finite integration, **9**, 105
  - grids, 105
  - material matrices, 112
  - nonlinear materials, 113
  - operators, 106
  - phantom objects, 107
- fractional step method, 58
- Gauß-Seidel, *see* dynamic iteration
- grad-div term, *see* regularization
- Hessenberg system, 48
- hole density, 27
  - discrete, 27, 72
- implicit Euler, **41**, 79, 87, 92, 96, 100
- incidence matrix, 18, **19**
- index, *see* differential algebraic equations
  - differential, 30
  - Kronecker, 30
  - linear index-2, 37
  - monolithic, 48
  - perturbation, 37
- inductance
  - circuit simulation, 18
  - extraction, 43
  - nonlinear, 45, 69
- interpolation step, *see* dynamic iteration
- Kronecker index, *see* index
- Lady Windermere, 58
- LIM-cutset, *see* circuit analysis
- linear solvers, 82
  - bypassing, 43, 89
  - PCG, 44, 80
  - substructuring, 99
- Lipschitz, 53, 76, 97
  - extrapolation, 50
  - neighborhood, 53
  - regions, 8
  - semiconductor, 76
  - semiconductor example, 97
- loop, *see* circuit analysis
- magnetic vector potential, **5**, 11
  - line-integrated, **11**, 12, 24, 64
- magnetostatic, *see* static
- mass matrix
  - magnetoquasistatic, 12
  - semiconductor, 27
- conductivity matrix, 12
- material matrices, 113
- matrix pencil, 14, 41, 79, 80
- MNA, *see* modified nodal analysis
- modified nodal analysis, **18**
- MQS, *see* quasistatic
- multirate, **39**, 44, 87, 93
  - bypassing, 44, 87
  - cosimulation, 46
  - field/circuit, 93
- multiscale, 17, 20
  - 0D network model, 18
  - 1D semiconductor device, 26
  - 3D field device, 20
- MVP, *see* magnetic vector potential
- network model, **18**
- Newton-Raphson, **40**
- permittivity, **5**, 9, 11
  - tensor, 9
- perturbation index, *see* index
- projectors, **14**
- properly stated leading term, 31, **40**
- quasistatic, 5
  - electroquasistatic, 5
  - magnetoquasistatic, **5**, 12
- regions, **8**, 113

- 
- regularization, 13
    - Coulomb's gauge, 5, 13
    - grad-div, 13
  - reluctivity, 5, 11
    - nonlinear, 9, 11, 43, 44, 113
    - tensor, 9
  - Runge-Kutta methods, 40, 95
  - Schur complement, *see* domain decomposition
    - inductance, 43
  - semiconductor
    - modeling, 26
  - simple coupling, *see* dynamic iteration
  - single rate, 40
  - smearing, 23
  - solid conductor, *see* conductor model
  - space discretization, 9, 105
  - splitting functions, 51
    - field/circuit, 67
    - semiconductor/circuit, 74–76
  - stamping, 41
  - static, 5
    - electrostatic, 6, 12, 26
    - magnetostatic, 6, 12
  - step size, 40, 95
  - stranded conductor, *see* conductor model
  - substructuring, *see* domain decomposition
  - time discretization, 40, 42, 58
  - time step size, *see* step size
  - time window size, *see* window size
  - tractability index, *see* index
  - window size, 49, 95
    - field example, 91
    - semiconductor example, 98