

Numerical homogenization of time-dependent Maxwell's equations with dispersion effects

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Jan Philip Freese

Tag der mündlichen Prüfung: 25.11.2020

Referent: Prof. Dr. Christian Wieners

Korreferent: Prof. Dr. Dietmar Gallistl

Acknowledgement

I would like to express my deep gratitude to Prof. Dr. Christian Wieners and Prof. Dr. Dietmar Gallistl for the supervision of this work. Furthermore, I would like to thank all members of the research groups IANM 1 and 3 of the Faculty of Mathematics who accompanied me on the way to my doctorate. I am deeply grateful for my friends and family, without whom all this would not have been possible.

Moreover, I thankfully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173 “Wave phenomena: analysis and numerics” and by the Baden-Württemberg Stiftung through the project “Mehrskalmethoden für Wellenausbreitung in heterogenen Materialien und Metamaterialien.” During an early stage of my doctoral studies I enjoyed the participation in the trimester program “Multiscale Problems: Algorithms, Numerical Analysis and Computation” at the Hausdorff Research Institute for Mathematics (HIM) where this work was initiated.

1	Introduction	1
1.1	Motivation	1
1.2	Literature review	4
1.3	Main results	5
1.4	Outline	6
2	Preliminaries	7
3	Mathematical modeling of electromagnetic waves	11
3.1	Maxwell's equations	11
3.1.1	Constitutive laws	12
3.1.2	The Maxwell system with conductivity	14
3.1.3	The Maxwell–Debye system as a model problem	15
3.1.4	A class of dispersive materials	16
3.1.5	Boundary conditions	18
3.2	Wellposedness	19
3.2.1	Maximal monotone operator theory and semigroups	19
3.2.2	Application to Maxwell's equations	23
3.2.3	Energy and stability	27
4	Homogenization of Maxwell's equations in locally periodic media	29
4.1	Heterogeneous materials and their restrictions in simulation	29
4.2	Homogenization results for the Maxwell system with conductivity	31
4.2.1	Asymptotic expansion for the Maxwell system with conductivity	32
4.3	Long time dispersive effects in homogenization of general wave equations	34
4.4	Homogenization of the Maxwell system	35
4.4.1	Heterogeneous Maxwell system	35
4.4.2	Periodic unfolding as generalization of two-scale convergence	35
4.4.3	The homogeneous Maxwell system and the delay effect	38

4.4.4	Reformulation of parameters and cell correctors	40
4.4.5	Wellposedness of cell problems	46
4.4.6	Symmetric representation of the convolution kernel and the extra source	52
4.4.7	On the exponential structure of the convolution kernel	53
4.4.8	Wellposedness of the integro-differential homogeneous system	54
4.5	Homogenization beyond periodicity	59
4.6	Application to the Maxwell–Debye system	59
5	The Finite Element Heterogeneous Multiscale Method	63
5.1	The finite element method	63
5.1.1	Nédélec finite elements	66
5.1.2	Lagrange finite elements	66
5.1.3	Quadrature	67
5.2	The HMM framework and the application to Maxwell’s equations	67
5.2.1	Approximating the effective parameters on the microscale	69
5.2.2	Results for the classical Maxwell system	74
5.2.3	Difficulties in the approximation of time-dependent parameters	76
5.3	A FE-HMM algorithm for the dispersive Maxwell system	76
5.3.1	Error analysis of the damping parameter	80
5.3.2	Error estimates for the Sobolev equation and the time-dependent parameter	83
5.3.3	Wellposedness of the semi-discrete system	91
5.3.4	Semi-discrete a priori error analysis	92
6	Time discretization and approximation of the convolution	105
6.1	Time integration of first-order ordinary differential equations	106
6.1.1	Explicit schemes	107
6.1.2	Implicit schemes	107
6.1.3	Problems for integro-differential equations	108
6.2	Approximating the convolution	109
6.2.1	Numerical quadrature	109
6.2.2	Recursive convolution	109
6.2.3	Memory variables	112
6.3	The fully discrete scheme and its error analysis	113
6.3.1	Recursive FE-HMM scheme for approximation of the effective solution	113
6.3.2	Fully discrete error analysis	115
7	Implementation and numerical experiments	119
7.1	Algorithmic approach	119
7.1.1	Efficient parallel computation of the HMM parameters	120
7.1.2	Time integration and convolution approximation	121
7.2	Numerical examples on the microscopic level	123

7.2.1	Microstructures and cell correctors in the conductivity setting	123
7.2.2	Microstructure in the Debye setting	137
7.3	Numerical examples for the HMM	137
7.3.1	Dimensionless Maxwell system	137
7.3.2	A test case with exact solution	139
7.3.3	Micro error analysis and locally periodic setting	142
8	Conclusion and Outlook	147
A	Appendix	149
A.1	The Debye model	149
A.1.1	Mechanisms of polarization	150
A.1.2	The model	152
A.2	Transformation of cell problems and parameters	154
A.3	Effective parameters of isotropically layered material	155
A.4	Exact solution for macroscopic Maxwell with layer material	157
	List of symbols	159
	List of function spaces	163
	Bibliography	165

CHAPTER 1

Introduction

1.1 Motivation

When light travels through a prism it separates in the different colors of the visible spectrum as shown in Figure 1.1. The reason for this effect is the difference in the frequencies, or equivalently the wavelengths, of the colors, causing them to propagate at different speeds. In reality not only a prism but every natural material shows these dispersive effects at least for certain parameters.

Materials that recently attracted great interest are so-called *metamaterials*. They are artificially constructed as composite materials. These consist of so-called *unit cells* whose size is significantly smaller than the wavelength of an incident wave. Examples for the unit cells are fishnet structures or split-ring resonators (SRR). An example of the latter one is shown in Figure 1.2. The reason for the great attention are extraordinary effects such as negative refractive indices, perfect lensing, electromagnetic cloaking or subwavelength imaging. All the aforementioned effects can be described by dispersive models, which links the topic of dispersion with the field of metamaterials. More precisely, the heterogeneity of the material leads to new dispersion effects which are the cause for the unusual properties. In 2000 [Smith et al. \(2000\)](#) were the first to construct those materials and in [Shelby et al. \(2001\)](#) the astonishing effect of negative

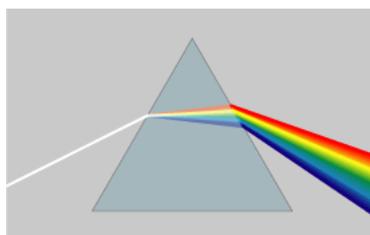


Figure 1.1: Dispersion in a prism [https://commons.wikimedia.org/wiki/File:Prism_rainbow_schema.png].

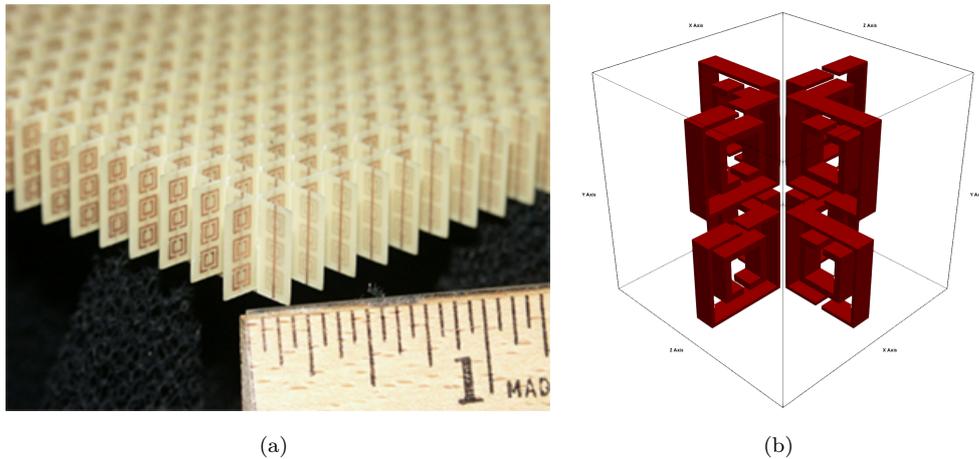


Figure 1.2: Metamaterial slab consisting of split ring resonators [https://commons.wikimedia.org/wiki/File:Split-ring_resonator_array_10K_sq_nm.jpg] (a) and a unit cell with eight SRRs (b).

refraction was first observed. Since then the interest in these artificial materials and their properties has increased and thus mathematical questions have arisen that are strongly related to dispersion models. In general the propagation of light, or other electromagnetic waves, such as radio or microwaves, is modeled using the famous Maxwell's equations. Heterogeneous materials that exhibit properties at different scales such as metamaterials enter these equations by rapidly varying parameters. Thus, although experiments showing the unexpected behavior of these materials are possible, the mathematical analysis of such structures is difficult. As we are interested in the effective behavior of the material, every fine scale must be resolved to cover the effects of the heterogeneous structure.

Likewise, if not even more, the numerical simulation has to cover the different scales. On the one hand, the microscopic structure has to be resolved to display all effects but on the other hand the macroscopic material slab itself has to be modeled, which is not possible simultaneously even with modern computer capacities. Thus, the modeling of wave propagation in heterogeneous media is a multiscale problem. In principle, however, we are only interested in effective properties of the macroscopic material.

For periodic structures such as the metamaterial consisting of split ring resonators a method to derive these effective properties is *homogenization*. The idea is to actually model a different material that is homogeneous but reflects the effective behavior of its heterogeneous counterpart. Of course, the challenge is how to derive this homogeneous material model. Here the assumption on periodicity on the one hand and the length of the waves compared to the size of the unit cells on the other hand are crucial. Under these assumptions it is reasonable to consider the limit when the size of the unit cells tends to zero. Eventually, this yields a Maxwell system with parameters that model the effective wave propagation through the heterogeneous material. Consequently, we call these parameters *effective*, since they model the behavior we are interested in. Except from special cases these parameters are not known analytically. Instead, they are represented as averages over the unit cells and include the solutions of partial differential equations, the so-called *micro problems*, posed on the unit cells. In this way microscopic effects enter the homogeneous system.

Similar to the astonishing effects occurring in the experiments above, this is also the case in homoge-

nization. These theoretical results are proposed in [Sánchez-Palencia \(1980\)](#) and especially in [Wellander \(2001\)](#). Here the authors consider a classical set of Maxwell's equations including Ohm's law with highly oscillatory periodic parameters. Thus, only instantaneous effects are taken into account. However, the derived effective system shows delay effects represented as convolutions in the time-dependent equations. The transfer of these equations to the frequency domain reveals the frequency dependence of the electric permittivity and implies that the effective material is dispersive. These results coincide with physically derived models for metamaterials such as the Drude ([Pendry et al., 1996](#)) or the Lorentz model ([Smith and Kroll, 2000](#)), which include dispersion as well.

As pointed out, dispersion does not occur solely in metamaterials but is present in every material. This is even the so-called *causality principle*. Therefore, in this thesis we consider a general class of Maxwell's equations that include dispersion already in the heterogeneous parameters. Recently, the effective system for this type of heterogeneous structures has been derived in [Bokil et al. \(2018\)](#). Logically, this system involves even more complicated dispersive models than the several composites.

From an analytical point of view, these results are simultaneously good and bad. On the one hand, representations of the effective parameters are derived and these do not depend on the microscopic scale anymore. On the other hand, the effective dispersion models yield a different type of system. Namely, we changed from a (partial) differential to an integro-differential equation, which causes new problems. Generally speaking, we exchange the heterogeneity in space with a non-local dependence in time, and we already mentioned that this yields frequency dependent parameters. In the classical Maxwell system these dispersion effects are contained as the polarization and magnetization. Consequently, such systems have been studied extensively in the past, including the time discretization for various polarization and magnetization models, such as the Debye, Drude or Lorentz models. Nevertheless, in general these results are not applicable to the effective system since the structure of the convolution kernel depends on the heterogeneity. More precisely, since the effective parameters are not given explicitly but as solutions of micro problems, the convolution may be arbitrarily complicated. In principle, however, the effective parameters only vary on the macroscopic scale.

Thus, at first glance, classical space discretizations such as finite differences or finite elements are applicable. But since we do not have an explicit representation of the effective parameters, there is no way to assemble the required matrices. Here multiscale methods apply that either try to overcome the above problem or directly tackle the heterogeneous system. In this thesis we focus on (locally) periodic structures and use the Heterogeneous Multiscale Method (HMM), which belongs to the first category. In general the HMM implements a framework that splits the multiscale problem in a macroscopic and a microscopic part with solvers for both. The method used in this thesis is also called Finite Element Heterogeneous Multiscale Method (FE-HMM) since we only consider the solvers on the macroscopic and microscopic level to be finite element methods.

In this sense the representation of the effective parameters that stem from homogenization is used as microscopic problem in the HMM. The solution of these micro problems may then be used to approximate the effective parameters on the macroscopic level. Eventually, this yields a Maxwell system that can be solved by a classical finite element method in space. Unfortunately, the time discretization of the HMM system involves a convolution that covers the dispersive effects but is again a challenge for the numerical approximation. The idea to overcome these difficulties is directly related to classical dispersion models. At least, for various of these models the convolution kernels are exponential functions, which allow to

compute a convolution recursively. This observation is taken into account for the time discretization of the HMM Maxwell system by an exponential fitting of the effective convolution kernel to an exponential function.

1.2 Literature review

We continue this introduction with a review on the literature concerning the above mentioned topics. The theory of Maxwell's equations is classical and may be found in [Banks et al. \(2000\)](#); [Jackson \(1999\)](#); [Landau and Lifshitz \(1960\)](#). The wellposedness of these equations may be shown using semigroup theory. The central references for this concept are [Engel and Nagel \(2000\)](#); [Pazy \(1983\)](#). The application of this theory to the Maxwell system is considered for example in [Hochbruck et al. \(2015a\)](#); [Sturm \(2017\)](#).

The propagation of electromagnetic waves in heterogeneous media and their unnatural effects have been studied extensively throughout the past. In [Veselago \(1968\)](#) the first theoretical result concerning negative refractive indices is provided. Finally in 2000, the construction of a metamaterial was the starting point of an extensively growing interest in these materials [Pendry \(2000\)](#); [Shelby et al. \(2001\)](#); [Smith et al. \(2000\)](#). For general overviews on the field of metamaterials and their discretization we refer to [Li and Huang \(2013\)](#); [Solyman and Shamonina \(2009\)](#) and the references therein.

In this thesis we follow the approach of homogenization to derive the effective material behavior. An overview of this method is found in [Cioranescu and Donato \(1999\)](#); [Jikov et al. \(1994\)](#). More specific for Maxwell's equations we refer to first results in [Sánchez-Palencia \(1980\)](#) and most importantly to [Wellander \(2001\)](#) for the time-domain homogenization of linear material laws. Non-linearity is taken into account in [Wellander \(2002\)](#), whereas in [Wellander and Kristensson \(2003\)](#) the frequency domain homogenization is considered. Wellander uses the concept of two-scale convergence ([Allaire, 1992](#); [Nguetseng, 1989](#)) to derive the effective system. The approach within this thesis is based on results presented in [Bokil et al. \(2018\)](#) that include linear dispersive effects. Even more complicated material laws, including time dependent parameters, are considered in [Bossavit et al. \(2005\)](#). The latter references use the method of periodic unfolding ([Cioranescu et al., 2002, 2008](#)) to derive the homogeneous system, which is a generalization of the two-scale convergence approach.

For the numerical approximation of the homogeneous solution we use the method of finite elements. Introductory reading may be found in [Brenner and Scott \(2008\)](#); [Ciarlet \(2002\)](#); [Ern and Guermond \(2004\)](#). The finite element method (FEM) applied to the Maxwell system is covered in [Monk \(2003\)](#); [Nédélec \(1980\)](#). Important within this thesis is the multiscale character of the wave propagation in heterogeneous media. Thus, we briefly discuss some multiscale methods. The method used here is the Heterogeneous Multiscale Method (HMM) introduced by [E and Engquist \(2003\)](#). See also [Abdulle \(2009\)](#); [Abdulle et al. \(2012\)](#); [E and Engquist \(2005\)](#) for survey articles. The HMM provides a general framework for the solution of multiscale problems. On the macroscopic scale we are concerned with a problem that inherits microscopic information. Thus, we choose a macroscopic solver, for e.g. a finite element method, and have to estimate the missing information from a suitable micro scale model. This problem has again to be solved with a suitable scheme. The HMM applied to the Maxwell system in time harmonic formulation is presented in [Ciarlet et al. \(2017\)](#) as well as [Henning et al. \(2016\)](#). An application to the time-domain Maxwell system is considered in [Hochbruck et al. \(2019\)](#) but without dispersive effects.

As alternative methods to cover the multiscale character we mention the Multiscale Finite Element

Method (MsFEM) (Efendiev and Hou, 2009; Hou and Wu, 1997) and the multiscale hybrid-mixed (MHM) finite element method from Harder et al. (2013). The latter one has recently been applied to an instantaneous heterogeneous Maxwell system in Lanteri et al. (2018). Moreover, multiscale methods for the Maxwell's equations are presented in Zhang et al. (2010). Finally, we mention the Localized Orthogonal Decomposition (LOD) (Henning et al., 2014; Målqvist and Peterseim, 2014; Peterseim, 2016). First introduced for the Laplace operator it was shown in Gallistl et al. (2018) that this technique is also applicable to the (stationary) Maxwell system. The aforementioned multiscale methods are especially useful for unstructured heterogeneities. In this thesis, however, we are interested in (locally) periodic structures and thus, the HMM seems to be the right choice.

Since we consider time dependent Maxwell's equations, also its discretization with respect to time is a concern. For an overview on the topic of time discretization see Hairer and Wanner (1996); Hairer et al. (1993, 2006). Time integration schemes for Maxwell systems without dispersive effects are studied in Hochbruck et al. (2015a, 2019). The approximation of dispersion in time-domain is either achieved using auxiliary differential equations (ADE) (Lanteri and Scheid, 2013; Li, 2011; Li and Zhang, 2010) or via recursive convolution as in Li (2007); Li and Chen (2008). The latter one is a technique introduced in Luebbers et al. (1990) to approximate convolutions with exponential kernel.

In many of the above references numerical experiments were carried out. Concerning the time dependent cell problems arising in the homogenization of Maxwell's equations we highlight the results in Banks et al. (2006). Here the authors perform numerical experiments on a unit cell and demonstrate the exponential decay of the effective convolution kernel for a heterogeneous Debye medium. An implementation of a HMM to the Maxwell's equations is found in Hochbruck et al. (2019), where a classical system without dispersion effects is considered.

With the above literature we are now in the position to highlight the results obtained in this thesis.

1.3 Main results

Our analysis and the numerical approach start with a reformulation of the effective system proposed in Bokil et al. (2018) and earlier in Bossavit et al. (2005). As it turns out, the Sobolev equation plays an important role in this new system. Thus, the analysis of these equations including wellposedness, stability and regularity is one part of our study. In Theorem 4.4.17, we show an H^2 -norm estimate for the solution of the Sobolev equation. We use these results to analyze the effective parameters. In Lemma 4.4.19, we prove the boundedness of the parameters, especially of the convolution kernel. Most importantly this parameter is uniformly bounded with respect to time. Moreover, in Section 4.4.7 we show that under additional assumptions this kernel is actually exponentially decaying. The wellposedness Theorem 4.4.21 for the effective macroscopic Maxwell system includes a refined stability estimate compared to the original result from Bokil et al. (2018).

The macroscopic Maxwell system including the effective parameters is then solved using the Heterogeneous Multiscale Method (HMM) in space. To our knowledge we are the first to apply this method to an effective Maxwell system with dispersive effects. Therefore, we derive estimates on the micro errors of all parameters, especially for the convolution kernel. The corresponding new results, that are crucial for the error analysis, are found in Lemmas 5.3.2, 5.3.7 and 5.3.9. Here we are again concerned with the Sobolev equation for which we propose a separate semi-discrete error estimate in Theorem 5.3.6. Subsequently,

we show the wellposedness of the HMM system and analyze the semi-discrete error of the Heterogeneous Multiscale Method applied to the effective Maxwell system including dispersive effects. The result, which bounds the L^2 -error in terms of the macroscopic discretization size H and the microscopic mesh size h , is found in Theorem 5.3.23.

The rigorous error analysis in space is followed by a rather standard time discretization at the end of which an efficient, fully discrete method is proposed. This method uses a recursive approximation of the convolution that relies on the assumption that the convolution kernel is an exponential function. In general this assumption is of course not true. However, we highlight that there are cases, as shown in Section 4.4.7 where this assumption is satisfied. Hence, for the general convolution kernels arising in homogenization we apply an exponential fit, which is then used in the recursive approach. Although we do not analyze the fully discrete error here, we provide approaches to this fully discrete analysis and eventually show numerical experiments. Within these we again observe the exponential decay of the convolution kernel for all our examples of heterogeneous structures.

Finally, we point out that all the numerical experiments in Chapter 7 are three-dimensional simulations. To our knowledge we are the first to provide such an algorithm including dispersion effects resulting from heterogeneous media.

1.4 Outline

The structure of this thesis is as follows. In the next chapter we start with some preliminaries concerning notation and basic results. Chapter 3 is concerned with the modeling of electromagnetic waves. We present a class of Maxwell's equations that include dispersive effects. Moreover, we show classic results from semi group theory, which yield the wellposedness of these equations. In Chapter 4, we investigate the theory of homogenization and the method of periodic unfolding. Moreover, we present two results on the homogenization of dispersive Maxwell systems. Following this, we use those results to derive a new effective Maxwell equation in differential form and show the wellposedness of this system. This is the starting point for the numerical approximation. Therefore, in Chapter 5 we introduce the method of finite elements. In addition, the Heterogeneous Multiscale Method (HMM) is presented as one possibility to handle the structure of the effective system. The wellposedness of the HMM-system and the semi discrete error estimate for the discretization using the HMM, is found in Theorem 5.3.23. The time discretization is then content of Chapter 6, where we present a recursive convolution approach for the effective integro-differential system. Moreover, we present the fully discrete recursive finite element heterogeneous multiscale method that we use for the numerical experiments. Finally, in Chapter 7, we illustrate how the proposed scheme is efficiently implemented and furthermore, numerical experiments both, on the microscopic and macroscopic level support our theoretical findings.

CHAPTER 2

Preliminaries

Throughout this thesis we use the following notation and basic results.

Miscellaneous

First, we make use of the concept of a generic constant $C > 0$, which does not depend on the mesh sizes or the gradient of parameters. It may, however, depend on the contrast of parameters and have different values for various cases. We consider $d, n \in \mathbb{N}$ to be integers. Moreover, let $\Omega, Y \subseteq \mathbb{R}^3$ be given domains, i.e., open, bounded and simply connected. The quantity id denotes the identity operator for an infinite dimensional space.

Vectors and Matrices

For vectors $a, b \in \mathbb{R}^d$, we denote the Euclidean scalar product as

$$a \cdot b = a^T b = \sum_{i=1}^d a_i b_i.$$

The Euclidean norm is denoted as $|a|$. Moreover, with $\mathbf{e}_\ell \in \mathbb{R}^d$, for $\ell \in \{1, \dots, d\}$ we denote the ℓ -th canonical basis vector of \mathbb{R}^d , whereas $\mathbf{0}_d \in \mathbb{R}^d$ denotes the vector consisting only of zeros. Similarly, $\mathbf{I}_{d \times d} \in \mathbb{R}^{d \times d}$ is the identity matrix while the matrix $\mathbf{0}_{d \times n} \in \mathbb{R}^{d \times n}$ has just zero entries. We may omit the indices if there are no ambiguities.

For a matrix $\mathbf{G} \in \mathbb{R}^{d \times n}$ we define the Frobenius norm as

$$\|\mathbf{G}\|_F := \sqrt{\sum_{i=1}^d \sum_{j=1}^n |\mathbf{G}_{ij}|^2}.$$

Differential operators

The derivative with respect to time is denoted with ∂_t . Now let $\Psi : \Omega \rightarrow \mathbb{R}$ and $\Phi : \Omega \rightarrow \mathbb{R}^3$ be sufficiently smooth functions. We define the ∇ -, curl- and div-operator as

$$\nabla \Psi(x) = \begin{pmatrix} \partial_{x_1} \Psi(x) \\ \partial_{x_2} \Psi(x) \\ \partial_{x_3} \Psi(x) \end{pmatrix}, \quad \text{curl } \Phi(x) = \begin{pmatrix} \partial_{x_2} \Phi_3(x) - \partial_{x_3} \Phi_2(x) \\ \partial_{x_3} \Phi_1(x) - \partial_{x_1} \Phi_3(x) \\ \partial_{x_1} \Phi_2(x) - \partial_{x_2} \Phi_1(x) \end{pmatrix}, \quad \text{div } \Phi(x) = \sum_{i=1}^3 \partial_{x_i} \Phi_i(x).$$

If a function $\Theta : \Omega \times Y \rightarrow \mathbb{R}^3$ is defined on a product space, we denote for $x \in \Omega$ and $y \in Y$ the derivatives with respect to the corresponding space as ∂_x and ∂_y . This is used in the same way for ∇_x , ∇_y , curl_x , curl_y and div_x , div_y .

Function spaces

Throughout this thesis we denote by $C^k(\Omega)$ the space of k times continuously differentiable functions on Ω . Moreover, the space of test functions is given as $C_0^\infty(\Omega)$, which is the space of smooth functions with compact support in Ω .

For $1 \leq p < \infty$ we denote with $L^p(\Omega; \mathbb{R})$ the standard space of measurable functions ϕ on Ω with values in \mathbb{R} that satisfy

$$\int_{\Omega} |\phi(x)|^p dx < \infty.$$

The norm on these spaces is denoted by $\|\cdot\|_{L^p(\Omega; \mathbb{R})}$. Most importantly, for $p = 2$ the space $L^2(\Omega; \mathbb{R})$ equipped with the scalar product

$$(\phi, \psi) := (\phi, \psi)_{L^2(\Omega)} = \int_{\Omega} \phi(x)\psi(x) dx \quad \text{for all } \phi, \psi \in L^2(\Omega; \mathbb{R}),$$

is a Hilbert space. With $L^\infty(\Omega; \mathbb{R})$ we denote the space of all essentially bounded measurable functions on Ω . The standard Sobolev spaces (Adams and Fournier, 2003) of k -times weakly differentiable functions in $L^p(\Omega; \mathbb{R})$ are denoted by $W^{k,p}(\Omega; \mathbb{R})$. The corresponding norms and seminorms are denoted as $\|\cdot\|_{W^{k,p}(\Omega; \mathbb{R})}$ and $|\cdot|_{W^{k,p}(\Omega; \mathbb{R})}$. As usual, we use the identification

$$H^k(\Omega; \mathbb{R}) = W^{k,2}(\Omega; \mathbb{R}).$$

The spaces $H^k(\Omega; \mathbb{R})$ are Hilbert spaces with respect to the standard inner product. As we are interested in vector valued problems, we generalize the above definitions by the use of the Euclidean inner product and use the notation $L^p(\Omega; \mathbb{R}^d)$, $W^{k,p}(\Omega; \mathbb{R}^d)$ for the corresponding functions with values in \mathbb{R}^d . The space of three-dimensional functions possessing a weak curl is denoted by

$$H(\text{curl}, \Omega) = \{f \in L^2(\Omega; \mathbb{R}^3) : \text{curl } f \in L^2(\Omega; \mathbb{R}^3)\},$$

which is equipped with the norm

$$\|\phi\|_{H(\text{curl}, \Omega)}^2 = \|\phi\|_{L^2(\Omega; \mathbb{R}^3)}^2 + \|\text{curl } \phi\|_{L^2(\Omega; \mathbb{R}^3)}^2 \quad \text{for all } \phi \in H(\text{curl}, \Omega).$$

The closure of $C_0^\infty(\Omega; \mathbb{R}^3)$ with respect to the $H(\text{curl}, \Omega)$ -norm is denoted as

$$H_0(\text{curl}, \Omega) = \overline{C_0^\infty(\Omega; \mathbb{R}^3)}^{\|\cdot\|_{H(\text{curl}, \Omega)}}.$$

Furthermore, for a Banach space $(X, \|\cdot\|_X)$ and a measurable space (S, Σ, ζ) we introduce the Bochner–Lebesgue spaces $L^p(S; X)$ (Hytönen et al., 2016). The measurable space in our case is either an interval $[0, T]$ with one dimensional Lebesgue measure or the domain Ω with corresponding Lebesgue measure. For $1 \leq p < \infty$, $\phi \in L^p(S; X)$ and $\psi \in L^\infty(S; X)$ these spaces are Banach spaces equipped with the norms

$$\|\phi\|_{L^p(S; X)} = \left(\int_S \|\phi(s)\|_X^p d\zeta(s) \right)^{\frac{1}{p}}, \quad \|\psi\|_{L^\infty(S; X)} = \text{ess sup}_{s \in S} \|\psi(s)\|_X.$$

Similarly to the above definitions we define the Banach space valued Sobolev spaces $W^{k,p}(S; X)$. Additionally, we define the mean of ϕ over Y with finite Lebesgue measure $|Y| < \infty$ as

$$\int_Y \phi(y) dy = \frac{1}{|Y|} \int_Y \phi(y) dy.$$

Analytical tools

We introduce the Fourier–Laplace transform \mathcal{F} of an integrable function u as

$$\widehat{u}(\omega) = \mathcal{F}(u)(\omega) := \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-i\omega t} u(t) dt.$$

The central property of this transform is

$$\widehat{\partial_t u}(\omega) = i\omega \widehat{u}(\omega) - \frac{1}{\sqrt{2\pi}} u(0).$$

Moreover, the Fourier-Laplace transform of a convolution yields the product of the individual transforms

$$\mathcal{F} \left(\int_0^t G(t-s)u(s) ds \right) (\omega) = \sqrt{2\pi} \widehat{G}(\omega) \widehat{u}(\omega).$$

Finally, we present a result concerning the derivative of a convolution. The proof follows from arguments using difference quotients.

Lemma 2.0.1. *Let $0 \leq t < \infty$. For any $G \in C^1([0, t])$ and $u \in L^1(0, t)$ we find*

$$\partial_t \int_0^t G(t-s)u(s) ds = \int_0^t (\partial_t G)(t-s)u(s) ds + G(0)u(t).$$

The preceding lemma directly extends to the case of Banach-space valued functions.

Notation

For a sum over a_i , $i = 0, \dots, N$, $N \in \mathbb{N}$ we abbreviate $\bar{a} = a_i$ and the sum is written short as

$$\sum_{i=0}^N a_i = \sum_i \bar{a}.$$

Finally, we use the symbol \approx to indicate that some kind of approximation is taking place. This can either be a quadrature formula, a finite element approximation or an exponential fitting.

CHAPTER 3

 Mathematical modeling of electromagnetic waves

In this first chapter we derive the governing equations that we intend to solve. Since we are interested in modeling electromagnetic waves, these equations are going to be the Maxwell's equations. We divide this chapter into two parts. In the next section we derive the mathematical models for electromagnetic waves. After that, in Section 3.1.1, we concentrate on the constitutive laws, which represent the material properties, and derive an abstract Maxwell system. We show two main examples, which fit in our class of Maxwell's equations in Sections 3.1.2 and 3.1.3. In the second part of this chapter, in Section 3.2, we show wellposedness for the class of Maxwell's equations we introduced in Section 3.1 and over that in Section 3.2.3 show some stability and energy estimates.

3.1 Maxwell's equations

Let us start with Maxwell's equations in time domain and differential formulation. The results of this section are well-known and may be found in Landau and Lifshitz (1960) or Jackson (1999). These equations couple the quantities \mathbf{E} and \mathbf{H} of the electric and magnetic field intensities with \mathbf{D} and \mathbf{B} , the electric and magnetic displacement fields or flux densities. All four fields depend on time and space and therefore we choose a time interval $[0, T]$ with a finite end time $T \in \mathbb{R}_{>0}$ and a bounded domain $\Omega \subseteq \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$ in three-dimensional space. With this notation we can give the formulation of Maxwell's equations in differential form. Including the electric current density \mathbf{J} and the electric charge density ρ we get

$$\partial_t \mathbf{D}(t, x) - \operatorname{curl} \mathbf{H}(t, x) = -\mathbf{J}(t, x), \quad (3.1a)$$

$$\partial_t \mathbf{B}(t, x) + \operatorname{curl} \mathbf{E}(t, x) = \mathbf{0}, \quad (3.1b)$$

$$\operatorname{div} \mathbf{D}(t, x) = \rho(t, x), \quad (3.1c)$$

$$\operatorname{div} \mathbf{B}(t, x) = 0. \quad (3.1d)$$

Equation (3.1a) is Ampère's circuital law with Maxwell's extension, which states how magnetic fields are generated. Equation (3.1b) is Faraday's law of induction, and it states that a time varying magnetic field induces an electric field. The additional equations (3.1c) and (3.1d) are given by Gauss's electric and magnetic law. Gauss's electric law relates electric charges to the electric field whereas Gauss's magnetic law states that there are no magnetic charges.

Observe that from Ampère's law (3.1a) and Gauss's electric law (3.1c) we find the continuity equation

$$\partial_t \rho(t, x) + \operatorname{div} \mathbf{J}(t, x) = 0,$$

which relates the electric current density with the electric charge density.

The Maxwell system is endowed with initial and boundary conditions on the fields that will be specified later on. The number of unknowns in the Maxwell system (3.1) is 12, but the system only involves eight equations. Thus, in order to get a wellposed problem we need additional conditions: the constitutive laws that describe how the displacement fields \mathbf{D} and \mathbf{B} depend on the intensities \mathbf{E} and \mathbf{H} . The most general form of this dependence would be

$$\mathbf{D} = \mathbf{D}(\mathbf{E}, \mathbf{H}), \quad \mathbf{B} = \mathbf{B}(\mathbf{E}, \mathbf{H}),$$

where the relation might be arbitrarily complicated, for example non-linear. In the next section we concentrate on the constitutive laws and derive an abstract Maxwell system that represents a class of linear but possibly non-local materials.

3.1.1 Constitutive laws

As pointed out in the previous section, we have to take care of the constitutive laws that relate the fields occurring in the Maxwell system (3.1). In order to do so we follow the classical approach, which introduces two auxiliary fields, the polarization \mathbf{P} and magnetization $\overline{\mathbf{M}}$, which characterize how the material responds to an electromagnetic wave. They are defined to satisfy

$$\mathbf{D}(t, x) = \varepsilon_0 \mathbf{E}(t, x) + \mathbf{P}(t, x), \quad (3.2a)$$

$$\mathbf{H}(t, x) = \frac{1}{\mu_0} \mathbf{B}(t, x) - \overline{\mathbf{M}}(t, x). \quad (3.2b)$$

The constants ε_0 and μ_0 are the electric permittivity and magnetic permeability of vacuum given by

$$\varepsilon_0 = 8.854 \times 10^{-12} \text{ A s V}^{-1} \text{ m}^{-1},$$

$$\mu_0 = 4\pi \times 10^{-7} \text{ V s A}^{-1} \text{ m}^{-1}.$$

In vacuum, there is no dielectric or magnetic material present. Hence, we have the simplest constitutive relation because here the polarization and magnetization both vanish. For, so-called, linear and local in space materials the polarization and magnetization are given as

$$\mathbf{P}(t, x) = \varepsilon_0 \chi_e(x) \mathbf{E}(t, x), \quad (3.3a)$$

$$\overline{\mathbf{M}}(t, x) = \chi_m(x) \mathbf{H}(t, x), \quad (3.3b)$$

where χ_e and χ_m are the electric and magnetic susceptibility, respectively. Here the constitutive relations become the most common ones

$$\mathbf{D}(t, x) = \varepsilon_0 \varepsilon_r(x) \mathbf{E}(t, x), \quad (3.4a)$$

$$\mathbf{B}(t, x) = \mu_0 \mu_r(x) \mathbf{H}(t, x). \quad (3.4b)$$

We point out that the parameters ε_r and μ_r , the relative electric permittivity tensor and the relative magnetic permeability tensor, respectively, are matrix-valued functions unless the material is isotropic. A direct consequence of (3.2), (3.3) and (3.4) are the relations

$$\varepsilon_r(x) = 1 + \chi_e(x), \quad \mu_r(x) = 1 + \chi_m(x).$$

Although the constitutive relation (3.4) yields good approximations, it violates some physical principles as shown in Landau and Lifshitz (1960). Especially dispersion effects, which are always present in matter, are not reflected by such material laws. Thus, we consider a broader class of constitutive relations where the dependence is still local in space but may be non-local in time. To be more precise the electric and magnetic susceptibilities χ_e and χ_m are not only space but also time dependent. With these parameters the polarization and magnetization are defined as convolutions

$$\mathbf{P}(t, x) = \varepsilon_0 \int_0^t \chi_e(t-s, x) \mathbf{E}(s, x) ds, \quad (3.5a)$$

$$\overline{\mathbf{M}}(t, x) = \int_0^t \chi_m(t-s, x) \mathbf{H}(s, x) ds. \quad (3.5b)$$

Here we observe that the polarization and magnetization depend on the history of the electric and magnetic field. By definition, the polarization and magnetization are zero at $t = 0$ but we allow instantaneous effects as well. Thus, we introduce the instantaneous polarization \mathbf{P}^{in} , which we assume to be proportional to the electric field as in (3.3)

$$\mathbf{P}^{\text{in}}(t, x) = \varepsilon_0 \chi_e^{\text{in}}(x) \mathbf{E}(t, x).$$

The total polarization is the sum of the instantaneous one and the already known polarization field in (3.5a). We insert this in the definition (3.2) of the polarization to get

$$\begin{aligned} \mathbf{D}(t, x) &= \varepsilon_0 \mathbf{E}(t, x) + \mathbf{P}^{\text{tot}}(t, x) = \varepsilon_0 \mathbf{E}(t, x) + \mathbf{P}^{\text{in}}(t, x) + \mathbf{P}(t, x) \\ &= \varepsilon_0 \mathbf{E}(t, x) + \varepsilon_0 \chi_e^{\text{in}}(x) \mathbf{E}(t, x) + \mathbf{P}(t, x) = \varepsilon_0 (1 + \chi_e^{\text{in}}(x)) \mathbf{E}(t, x) + \mathbf{P}(t, x). \end{aligned}$$

Thus, the new definition of the polarization reads

$$\mathbf{D}(t, x) = \varepsilon_0 \varepsilon_{\text{in}}(x) \mathbf{E}(t, x) + \mathbf{P}(t, x), \quad (3.6)$$

where $\varepsilon_{\text{in}}(x) = 1 + \chi_e^{\text{in}}(x)$ is a permittivity that also accounts instantaneous effects and the polarization field is still given as in (3.5a). With the same approach we get with $\mu_{\text{in}} = 1 + \chi_m^{\text{in}}$ for the magnetization

$$\mathbf{B}(t, x) = \mu_0 \mu_{\text{in}}(x) \mathbf{H}(t, x) + \mu_0 \overline{\mathbf{M}}(t, x). \quad (3.7)$$

The relations in (3.6) and (3.7) combined with (3.5) are the most general ones we consider in this thesis. We specify the possible convolution kernels in Section 3.1.4.

The end of this section is dedicated to the transfer of the relation (3.5) into frequency domain. For this purpose we apply the Fourier–Laplace transform to (3.5) resulting in

$$\hat{\mathbf{P}}(\omega, x) = \varepsilon_0 \hat{\chi}_e(\omega, x) \hat{\mathbf{E}}(\omega, x), \quad \hat{\overline{\mathbf{M}}}(\omega, x) = \hat{\chi}_m(\omega, x) \hat{\mathbf{H}}(\omega, x). \quad (3.8)$$

This representation highlights that the susceptibilities are frequency dependent. Thus, the non-local material law in time domain corresponds to frequency dependent laws in frequency domain. This dependence on the frequency is exactly the effect of dispersion, i.e, the propagation of a wave depends on its frequency. The constitutive relation in frequency domain including instantaneous effects becomes

$$\hat{\mathbf{D}}(\omega, x) = \varepsilon_0 (\varepsilon_{\text{in}}(x) + \hat{\chi}_e(\omega, x)) \hat{\mathbf{E}}(\omega, x), \quad (3.9a)$$

$$\hat{\mathbf{B}}(\omega, x) = \mu_0 (\mu_{\text{in}}(x) + \hat{\chi}_m(\omega, x)) \hat{\mathbf{H}}(\omega, x). \quad (3.9b)$$

Hence, the frequency dependent permittivity and permeability are given as

$$\hat{\varepsilon}(\omega, x) = \varepsilon_0 (\varepsilon_{\text{in}}(x) + \hat{\chi}_e(\omega, x)), \quad \hat{\mu}(\omega, x) = \mu_0 (\mu_{\text{in}}(x) + \hat{\chi}_m(\omega, x)).$$

As already mentioned, the frequency dependence allows incorporating dispersive effects into the Maxwell system. Thus, the constitutive relations (3.5) or equivalently (3.9) are the right choice for modeling dispersive effects. Before we state the most general Maxwell system in Section 3.1.4, we show two examples that occur frequently throughout this thesis. In the next section we present the classical Maxwell system with Ohm's law.

3.1.2 The Maxwell system with conductivity

In this setting we consider a classic Maxwell system where we have neither polarization nor magnetization, i.e., $\mathbf{P} \equiv \mathbf{0}$ as well as $\overline{\mathbf{M}} \equiv \mathbf{0}$. To be more precise we only consider instantaneous effects, i.e., the constitutive relation is given as in (3.4). Moreover, we include Ohm's law that states that the current density \mathbf{J} can be split into the external current \mathbf{J}_0 and a part that is proportional to the electric field, i.e., for $x \in \Omega$ we find

$$\mathbf{J}(t, x) = \sigma(x)\mathbf{E}(t, x) + \mathbf{J}_0(t, x).$$

The Maxwell system in this case is

$$\varepsilon_0 \varepsilon_r(x) \partial_t \mathbf{E}(t, x) + \sigma(x) \mathbf{E}(t, x) - \text{curl } \mathbf{H}(t, x) = -\mathbf{J}_0(t, x), \quad (3.10a)$$

$$\mu_0 \mu_r(x) \partial_t \mathbf{H}(t, x) + \text{curl } \mathbf{E}(t, x) = \mathbf{0}. \quad (3.10b)$$

We introduce some short notation. For that purpose collect the solution components as

$$\mathbf{u}(t, x) = \begin{pmatrix} \mathbf{E}(t, x) \\ \mathbf{H}(t, x) \end{pmatrix},$$

and define the matrices

$$\mathbf{M}(x) = \begin{pmatrix} \varepsilon_0 \varepsilon_r(x) & \\ & \mu_0 \mu_r(x) \end{pmatrix}, \quad \mathbf{R}(x) = \begin{pmatrix} \sigma(x) & \\ & \mathbf{0}_{3 \times 3} \end{pmatrix}.$$

Thus, the Maxwell conductivity system (3.10) can be written as

$$\mathbf{M}(x) \partial_t \mathbf{u}(t, x) + \mathbf{R}(x) \mathbf{u}(t, x) + \begin{pmatrix} -\text{curl} \\ \text{curl} \end{pmatrix} \mathbf{u}(t, x) = \mathbf{g}(t, x), \quad (3.11)$$

$$\text{where } \mathbf{g}(t, x) = \begin{pmatrix} -\mathbf{J}_0(t, x) \\ \mathbf{0}_3 \end{pmatrix}.$$

Besides this classic Maxwell system we give another example, which is the Maxwell–Debye system presented in the next section.

3.1.3 The Maxwell–Debye system as a model problem

The Maxwell system with conductivity from the previous section does not cover dispersive effects. In this section we present the Debye model for orientation polarization, which actually includes dispersion. For this example we assume that there is no magnetization, i.e., $\overline{\mathbf{M}} \equiv \mathbf{0}$. The polarization is given by the Debye model presented in Debye (1929), which is also discussed in the Appendix A.1. In the Debye model the polarization is given as

$$\mathbf{P}(t, x) = \int_0^t \exp\left(\frac{-(t-s)}{\tau(x)}\right) \frac{\varepsilon_0(\varepsilon_s(x) - \varepsilon_\infty(x))}{\tau(x)} \mathbf{E}(s, x) ds. \quad (3.12)$$

Here ε_s denotes the permittivity at zero frequency and ε_∞ the permittivity at maximum frequency of the material under consideration. The latter one also describes the instantaneous polarization, i.e., in (3.6) we choose $\varepsilon_{\text{in}} = \varepsilon_\infty$. Moreover, τ is the relaxation time, which describes the time electric dipoles need to react on the application of an electromagnetic field. We deduce from (3.5a) that

$$\chi_e(t, x) = \exp\left(\frac{-(t-s)}{\tau(x)}\right) \frac{\varepsilon_s(x) - \varepsilon_\infty(x)}{\tau(x)}.$$

A direct consequence of the definition (3.12) is that the polarization satisfies the ordinary differential equation (ODE)

$$\tau(x)\partial_t \mathbf{P}(t, x) + \mathbf{P}(t, x) = \varepsilon_0(\varepsilon_s(x) - \varepsilon_\infty(x)) \mathbf{E}(t, x), \quad \mathbf{P}(0, x) = \mathbf{0}. \quad (3.13)$$

We now abbreviate $\Delta\varepsilon := \varepsilon_s - \varepsilon_\infty$. From (3.6) together with the definition (3.13) of the polarization the Maxwell system (3.1) reads

$$\begin{aligned} \varepsilon_0 \varepsilon_\infty(x) \partial_t \mathbf{E}(t, x) + \varepsilon_0 \tau(x)^{-1} \Delta\varepsilon(x) \mathbf{E}(t, x) - \tau(x)^{-1} \mathbf{P}(t, x) - \text{curl } \mathbf{H}(t, x) &= -\mathbf{J}(t, x), \\ \tau(x) \partial_t \mathbf{P}(t, x) + \mathbf{P}(t, x) - \varepsilon_0 \Delta\varepsilon(x) \mathbf{E}(t, x) &= \mathbf{0}, \\ \mu_0 \partial_t \mathbf{H}(t, x) + \text{curl } \mathbf{E}(t, x) &= \mathbf{0}. \end{aligned}$$

This system can also be represented in a structure similar to (3.11), where now the solution is given as

$$\mathbf{u}(t, x) = \begin{pmatrix} \mathbf{E}(t, x) \\ \mathbf{P}(t, x) \\ \mathbf{H}(t, x) \end{pmatrix},$$

and the matrices are

$$\mathbf{M}(x) = \begin{pmatrix} \varepsilon_0 \varepsilon_\infty(x) & & \\ & \tau(x) & \\ & & \mu_0 \end{pmatrix}, \quad \mathbf{R}(x) = \begin{pmatrix} \tau(x)^{-1} \varepsilon_0 \Delta\varepsilon(x) & -\tau(x)^{-1} & \mathbf{0}_{3 \times 3} \\ -\varepsilon_0 \Delta\varepsilon(x) & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} \end{pmatrix}.$$

This yields the Maxwell system

$$\mathbf{M}(x) \partial_t \mathbf{u}(t, x) + \mathbf{R}(x) \mathbf{u}(t, x) + \begin{pmatrix} & -\text{curl} \\ \mathbf{0} & \\ \text{curl} & \end{pmatrix} \mathbf{u}(t, x) = \mathbf{g}(t, x), \quad (3.15)$$

where the right-hand side has the definition $\mathbf{g}(t, x) = \begin{pmatrix} -\mathbf{J}(t, x) \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$.

The Debye model is a prototype for more complicated models that we introduce in the next section. Here the central property is that the polarization and magnetization may be represented as solutions of ODEs.

3.1.4 A class of dispersive materials

Most of the mathematical theory of this section is based on [Cassier et al. \(2017\)](#) and [\(Banks et al., 2000, Section 2.1\)](#). In general the material laws are given in (3.6) and (3.7), where we allow the polarization and magnetization to be given as convolutions with the electric and magnetic field, respectively, as seen in (3.5). Since all the mechanisms for polarization and magnetization work in the same way, we focus on the polarization. In the previous section we introduced the Debye model, which is one possibility to model dispersive effects. The polarization in (3.12) is given as convolution, and it turns out that it is equivalent to define it as solution of the first-order ODE (3.13). More complex dispersion models are often given as higher-order ODEs. As we show in this section this is still equivalent to the definition via convolution but with possibly more complicated kernels. These higher-order models are exactly that kind of polarizations we consider in this thesis. For $N_E \in \mathbb{N}$ we define the polarization as solution of the N_E th-order ODE

$$\sum_{j=0}^{N_E} a_j^E(x) \partial_t^j \mathbf{P}(t, x) = \varepsilon_0 b_0^E(x) \mathbf{E}(t, x), \quad \partial_t^j \mathbf{P}(0, x) = \mathbf{0}_3 \quad \text{for } j = 0, \dots, N_E - 1. \quad (3.16)$$

The parameters are scalars, i.e., $a_j^E, b_0^E : \Omega \rightarrow \mathbb{R}$, $j = 1, \dots, N_E$, which represents an isotropic material. In this section we frequently use the identification $a_j^E = a_j^E \mathbf{I}_{3 \times 3}$, $j = 0, \dots, N_E - 1$ and $b_0^E = b_0^E \mathbf{I}_{3 \times 3}$.

The goal of this section is to derive an abstract Maxwell system, which includes these higher-order polarization and magnetization models. The key to this abstract system is that an n th-order ODE can be written as a system of first-order ordinary differential equations.

Here we distinguish two cases. If $N_E = 1$ we already have the first-order ODE

$$a_1^E(x) \partial_t \mathbf{P}(t, x) + a_0^E(x) \mathbf{P}(t, x) - \varepsilon_0 b_0^E \mathbf{E}(t, x) = \mathbf{0}_3, \quad \mathbf{P}(0, x) = \mathbf{0}_3.$$

The Debye polarization from the previous section fits in this setting.

If $N_E \geq 2$ we introduce the notation $\mathbf{P}^{(0)} = \mathbf{P}$ and $\mathbf{P}^{(\ell+1)} = \partial_t \mathbf{P}^{(\ell)}$ for $\ell = 0, \dots, N_E - 2$. Hence, with

$$\mathbb{P}(t, x) = \left(\mathbf{P}^{(0)}(t, x)^T \quad \mathbf{P}^{(1)}(t, x)^T \quad \dots \quad \mathbf{P}^{(N_E-1)}(t, x)^T \right)^T \in \mathbb{R}^{3N_E},$$

we define a first-order ODE by

$$\mathbf{M}_{\mathbb{P}}(x) \partial_t \mathbb{P}(t, x) + \mathbf{R}_{\mathbb{P}\mathbb{P}}(x) \mathbb{P}(t, x) + \mathbf{R}_{\mathbb{P}\mathbf{E}}(x) \mathbf{E}(t, x) = \mathbf{0}_{3N_E}, \quad \mathbb{P}(0, x) = \mathbf{0}_{3N_E}. \quad (3.17)$$

The matrices are defined by

$$\mathbf{M}_{\mathbb{P}}(x) := \begin{pmatrix} \mathbf{I}_{3(N_E-1) \times 3(N_E-1)} & \mathbf{0}_{3(N_E-1) \times 3} \\ \mathbf{0}_{3 \times 3(N_E-1)} & a_{N_E}^E(x) \end{pmatrix} \in \mathbb{R}^{3N_E \times 3N_E}, \quad \mathbf{R}_{\mathbb{P}\mathbb{E}}(x) := \begin{pmatrix} \mathbf{0}_{3(N_E-1) \times 3} \\ -\varepsilon_0 b_0^E(x) \end{pmatrix} \in \mathbb{R}^{3N_E \times 3},$$

$$\mathbf{R}_{\mathbb{P}\mathbb{P}}(x) := \begin{pmatrix} \mathbf{0}_{3 \times 3} & -\mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \cdots & \mathbf{0}_{3 \times 3} \\ \vdots & \mathbf{0}_{3 \times 3} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \cdots & \mathbf{0}_{3 \times 3} & -\mathbf{I}_{3 \times 3} \\ a_0^E(x) & a_1^E(x) & \cdots & a_{N_E-2}^E(x) & a_{N_E-1}^E(x) \end{pmatrix} \in \mathbb{R}^{3N_E \times 3N_E},$$

where we use the matrix interpretation of the parameters. With the obvious identification

$$\mathbf{M}_{\mathbb{P}} = a_1^E, \quad \mathbf{R}_{\mathbb{P}\mathbb{P}} = a_0^E, \quad \mathbf{R}_{\mathbb{P}\mathbb{E}} = -\varepsilon_0 b_0^E,$$

we collect both cases $N_E = 1$ and $N_E \geq 2$ as the system (3.17). For the magnetization we may derive these auxiliary differential equations as well. We just found that we can rewrite the ODE for \mathbf{P} and $\overline{\mathbf{M}}$ as a system of first-order ODEs, which was our goal. Even more, from (3.17) we get with the variation of constants formula that the constitutive laws (3.6) and (3.7) are given as

$$\mathbf{D}(t, x) = \varepsilon_0 \left(\varepsilon_{\text{in}}(x) \mathbf{E}(t, x) + \int_0^t \chi_e(t-s, x) \mathbf{E}(s, x) \, ds \right),$$

$$\mathbf{B}(t, x) = \mu_0 \left(\mu_{\text{in}}(x) \mathbf{H}(t, x) + \int_0^t \chi_m(t-s, x) \mathbf{H}(s, x) \, ds \right),$$

where χ_e and χ_m are the electric and magnetic susceptibility, respectively.

Thanks to the structure of the polarization and magnetization models we can write the Maxwell system (3.1) in an abstract setting. Therefore we define the solution vector

$$\mathbf{u}(t, x) = \begin{pmatrix} \mathbf{E}(t, x) \\ \mathbb{P}(t, x) \\ \mathbf{H}(t, x) \\ \mathbb{M}(t, x) \end{pmatrix}, \quad (3.19)$$

and the parameter matrices

$$\mathbf{M}(x) = \begin{pmatrix} \varepsilon_0 \varepsilon_{\text{in}}(x) & & & \\ & \mathbf{M}_{\mathbb{P}}(x) & & \\ & & \mu_0 \mu_{\text{in}}(x) & \\ & & & \mathbf{M}_{\mathbb{M}}(x) \end{pmatrix}, \quad \mathbf{R}(x) = \begin{pmatrix} \mathbf{R}_{\mathbb{E}\mathbb{E}}(x) & \mathbf{R}_{\mathbb{E}\mathbb{P}}(x) & & \\ \mathbf{R}_{\mathbb{P}\mathbb{E}}(x) & \mathbf{R}_{\mathbb{P}\mathbb{P}}(x) & & \\ & & \mathbf{R}_{\mathbb{H}\mathbb{H}}(x) & \mathbf{R}_{\mathbb{H}\mathbb{M}}(x) \\ & & \mathbf{R}_{\mathbb{M}\mathbb{H}}(x) & \mathbf{R}_{\mathbb{M}\mathbb{M}}(x) \end{pmatrix}.$$

Here the matrices $\mathbf{R}_{\mathbb{E}\mathbb{E}}, \mathbf{R}_{\mathbb{H}\mathbb{H}} \in \mathbb{R}^{3 \times 3}$, $\mathbf{R}_{\mathbb{E}\mathbb{P}} \in \mathbb{R}^{3 \times 3N_E}$ and $\mathbf{R}_{\mathbb{H}\mathbb{M}} \in \mathbb{R}^{3 \times 3N_H}$ again depend on the order of the polarization and magnetization model. For $N_E, N_H = 1$ they are given as

$$\mathbf{R}_{\mathbb{E}\mathbb{E}}(x) := \frac{\varepsilon_0 b_0^E(x)}{a_1^E(x)}, \quad \mathbf{R}_{\mathbb{E}\mathbb{P}}(x) := -\frac{a_0^E(x)}{a_1^E(x)},$$

$$\mathbf{R}_{\mathbb{H}\mathbb{H}}(x) := \frac{\mu_0 b_0^H(x)}{a_1^H(x)}, \quad \mathbf{R}_{\mathbb{H}\mathbb{M}}(x) := -\frac{a_0^H(x)}{a_1^H(x)},$$

whereas for $N_E \geq 2$ we find

$$\begin{aligned} \mathbf{R}_{EE}(x) &:= \mathbf{0}_{3 \times 3}, & \mathbf{R}_{EP}(x) &:= \begin{pmatrix} \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3(N_E-2)} \end{pmatrix}, \\ \mathbf{R}_{HH}(x) &:= \mathbf{0}_{3 \times 3}, & \mathbf{R}_{HM}(x) &:= \begin{pmatrix} \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3(N_E-2)} \end{pmatrix}. \end{aligned}$$

These parameter matrices result from the ODE for both the polarization and magnetization as well as the PDE for the Maxwell system. Moreover, the curl operator and the right-hand side are represented as \mathbf{A} and \mathbf{g} respectively and are defined as

$$\mathbf{A}\mathbf{u}(t, x) = \begin{pmatrix} -\operatorname{curl} \mathbf{H}(t, x) \\ \mathbf{0}_{3N_E} \\ \operatorname{curl} \mathbf{E}(t, x) \\ \mathbf{0}_{3N_H} \end{pmatrix}, \quad \mathbf{g}(t, x) = \begin{pmatrix} -\mathbf{J}(t, x) \\ \mathbf{0}_{3N_E} \\ \mathbf{0}_3 \\ \mathbf{0}_{3N_H} \end{pmatrix}. \quad (3.20)$$

With this notation we rewrite the Maxwell system (3.1) with polarization and magnetization in abstract form as

$$\mathbf{M}(x)\partial_t \mathbf{u}(t, x) + \mathbf{R}(x)\mathbf{u}(t, x) + \mathbf{A}\mathbf{u}(t, x) = \mathbf{g}(t, x), \quad \text{in } (0, T) \times \Omega, \quad (3.21a)$$

$$\mathbf{u}(0, x) = \mathbf{u}_0(x), \quad \text{in } \Omega. \quad (3.21b)$$

The dimension of the system is $n := 3(2 + N_E + N_H)$. In Section 3.1.2 and 3.1.3, we already saw two examples that fit in this framework. In both examples we have $N_H = 0$ where in the conductivity case also $N_E = 0$ and in the Debye setting we find $N_E = 1$.

At the end of this section we again transfer the formulation above to the frequency domain. Therefore, consider (3.16) and apply the Fourier–Laplace transform. This results in

$$\hat{\mathbf{P}}(\omega, x) = \frac{\varepsilon_0 b_0^E(x)}{\sum_{j=0}^{N_E} a_j^E(x)(i\omega)^j} \hat{\mathbf{E}}(\omega, x).$$

Thus, from (3.8) and (3.9), we find the frequency dependent permittivity as

$$\hat{\varepsilon}(\omega, x) = \varepsilon_0 \left(\varepsilon_{\text{in}}(x) + \frac{b_0^E(x)}{\sum_{j=0}^{N_E} a_j^E(x)(i\omega)^j} \right).$$

The Debye polarization model from Section 3.1.3 results in the permittivity

$$\hat{\varepsilon}(\omega, x) = \varepsilon_0 \left(\varepsilon_{\infty}(x) + \frac{\Delta\varepsilon(x)}{1 + i\omega\tau} \right).$$

Having the abstract Maxwell system including high-order polarization and magnetization models at hand we briefly turn to the topic of boundary conditions in the next section. After that, in Section 3.2 we focus on the wellposedness of the abstract Maxwell system (3.21).

3.1.5 Boundary conditions

Since it is not possible to simulate the Maxwell system on the whole \mathbb{R}^3 , we chose a bounded domain Ω . Thus, we have to take care of the boundary $\partial\Omega$ of this domain. There is a vast theory about boundary conditions for Maxwell's equations, see Monk (2003). However, since in this thesis the focus lies on the heterogeneous structure of the materials we will not go into detail, but consider only the easiest boundary

condition, which is the one of a perfect conductor. Throughout this monograph we denote by \mathbf{n} the unit outward normal to $\partial\Omega$. Then the perfectly conducting boundary conditions read

$$\mathbf{E}(t, x) \times \mathbf{n} = \mathbf{0}, \quad \text{for } (t, x) \in (0, T) \times \partial\Omega.$$

If we use the perfectly conducting boundary conditions for the electric field we see that thanks to the Maxwell system (3.1b) we get

$$\partial_t (\mathbf{n} \cdot \mathbf{B}(t, x)) = \mathbf{n} \cdot \partial_t \mathbf{B}(t, x) = -\mathbf{n} \cdot \text{curl } \mathbf{E}(t, x) = \text{div} (\mathbf{n} \times \mathbf{E}(t, x)) - \mathbf{E}(t, x) \cdot \text{curl } \mathbf{n} = 0,$$

and therefore

$$\mathbf{n} \cdot \mathbf{B}(t, x) = \text{const}, \quad \text{for } (t, x) \in (0, T) \times \partial\Omega. \quad (3.22)$$

Thus, if we make sure that the initial value of the magnetic displacement field satisfies (3.22) for $t = 0$, it is satisfied for all $t \in (0, T)$. The complete Maxwell system including boundary and initial conditions is

$$\varepsilon(x) \partial_t \mathbf{E}(t, x) + \partial_t \mathbf{P}(t, x) - \text{curl } \mathbf{H}(t, x) = -\mathbf{J}(t, x), \quad \text{in } (0, T) \times \Omega, \quad (3.23a)$$

$$\mu(x) \partial_t \mathbf{H}(t, x) + \partial_t \overline{\mathbf{M}}(t, x) + \text{curl } \mathbf{E}(t, x) = \mathbf{0}, \quad \text{in } (0, T) \times \Omega, \quad (3.23b)$$

$$\text{div } \mathbf{D}(t, x) = \rho(t, x), \quad \text{in } (0, T) \times \Omega, \quad (3.23c)$$

$$\text{div } \mathbf{B}(t, x) = 0, \quad \text{in } (0, T) \times \Omega, \quad (3.23d)$$

$$\mathbf{E}(0, x) = \mathbf{E}_0(x), \quad \mathbf{H}(0, x) = \mathbf{H}_0(x), \quad \text{in } \Omega, \quad (3.23e)$$

$$\mathbf{E}(t, x) \times \mathbf{n}(x) = \mathbf{0}, \quad \mathbf{B}(0, x) \cdot \mathbf{n}(x) = 0, \quad \text{on } (0, T) \times \partial\Omega. \quad (3.23f)$$

The system (3.23) has to be complemented by material laws for the polarization and magnetization. In the next section we will study the wellposedness of the system (3.23) for the materials that we introduced in Section 3.1.4.

3.2 Wellposedness

Before we move on to the heterogeneous materials, we have to make sure that the system we consider is wellposed. The main tool for the existence and uniqueness result is the theory of maximal monotone operators and semigroups. First we introduce the theoretical foundation we need and after that apply it to the Maxwell system (3.21) in Section 3.2.2. The main reference is the monograph Engel and Nagel (2000). Moreover, we follow the lines of Sturm (2017) and refer to Monk (2003).

3.2.1 Maximal monotone operator theory and semigroups

In this section we present well known results that will enable us to study the wellposedness of the abstract Cauchy problem

$$\partial_t \mathbf{u}(t) = \mathcal{A} \mathbf{u}(t) + \mathbf{f}(t), \quad \mathbf{u}(0) = \mathbf{u}_0. \quad (3.24)$$

We first give the analytical setting for the analysis of the problem (3.24).

Let $(X, (\cdot, \cdot)_X)$ be a Hilbert space with corresponding norm $\|\cdot\|_X^2 = (\cdot, \cdot)_X$ and $\mathcal{L}(X)$ the space of all bounded

linear operators from X into X with operator norm

$$\|\mathcal{A}\|_{X \leftarrow X} = \sup_{x \in X, x \neq 0} \frac{\|\mathcal{A}x\|_X}{\|x\|_X}.$$

The analysis of the evolution equation (3.24) is closely related to the notion of strongly continuous semigroups.

Definition 3.2.1 (Strongly continuous semigroup). *A one-parameter family $(T(t))_{t \geq 0}$ of bounded linear operators from X to X is called a semigroup of bounded linear operators on X if*

- (i) $T(0) = \text{id}$ and
- (ii) $T(t+s) = T(t)T(s)$ for all $t, s \geq 0$.

A semigroup $(T(t))_{t \geq 0}$ is called a strongly continuous semigroup or C_0 -semigroup if for all $x \in X$ it holds

$$\lim_{t \rightarrow 0^+} \|T(t)x - x\|_X = 0,$$

i.e. $t \mapsto T(t)$ is strongly continuous at 0.

If we replace in Definition 3.2.1 the condition “ $t, s \geq 0$ ” by “ $t, s \in \mathbb{R}$ ” and “ $t \rightarrow 0^+$ ” by “ $t \rightarrow 0$ ” we get the definition of a (strongly continuous) group.

Lemma 3.2.2. *A strongly continuous semigroup $(T(t))_{t \geq 0}$ has the following properties:*

- (i) There exist constants $C \geq 1$ and $\beta \in \mathbb{R}$ such that

$$\|T(t)\|_{X \leftarrow X} \leq Ce^{\beta t}, \quad \text{for all } t \geq 0.$$

- (ii) The mapping $t \mapsto T(t)$ is strongly continuous on $[0, \infty)$, i.e.

$$\lim_{s \rightarrow 0} \|T(t+s)x - T(t)x\|_X = 0, \quad \text{for all } t \geq 0.$$

An important special case of a C_0 -semigroup is the one of a contraction semigroup.

Definition 3.2.3 (Contraction semigroup). *An operator generates a semigroup of contraction $(T(t))_{t \geq 0}$ if*

$$\|T(t)\|_X \leq 1.$$

The relation between the strongly continuous semigroup and the operator \mathcal{A} in (3.24) becomes clear by the definition of the generator of a semigroup.

Definition 3.2.4 (Infinitesimal generator). *Let $(T(t))_{t \geq 0}$ be a C_0 -semigroup. We define the linear operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \rightarrow X$ by*

$$\mathcal{A}x = \lim_{t \rightarrow 0^+} \frac{T(t)x - x}{t}, \tag{3.25}$$

where the domain $\mathcal{D}(\mathcal{A})$ consists of all $x \in X$ for which the limit in (3.25) exists.

We call \mathcal{A} the infinitesimal generator of the strongly continuous semigroup $(T(t))_{t \geq 0}$.

The following lemma gives a statement about the differentiability and yields the relation to the evolution equation.

Lemma 3.2.5. *Let $(T(t))_{t \geq 0}$ be a C_0 -semigroup with infinitesimal generator \mathcal{A} . Then the following holds true:*

(i) *For $x \in \mathcal{D}(\mathcal{A})$ and $t \geq 0$ we have $(T(t))x \in \mathcal{D}(\mathcal{A})$.*

(ii) *For all $x \in \mathcal{D}(\mathcal{A})$ and all $t \geq 0$ we have the relation*

$$\frac{d}{dt}(T(t)x) = \mathcal{A}T(t)x = T(t)\mathcal{A}x.$$

(iii) *The domain of \mathcal{A} is dense in X and \mathcal{A} is a closed operator.*

The next result shows that the relation between the generator and the semigroup is in fact unique.

Corollary 3.2.6. *Let $(T_1(t))_{t \geq 0}$ and $(T_2(t))_{t \geq 0}$ be two C_0 -semigroups with generators \mathcal{A}_1 and \mathcal{A}_2 , respectively. If $\mathcal{A}_1 = \mathcal{A}_2$, then $T_1(t) = T_2(t)$ for all $t \geq 0$.*

Now we can state a main result for the connection between evolution equations and the strongly continuous semigroup, first without a right-hand side.

Theorem 3.2.7 ((Engel and Nagel, 2000, Chapter III, 6.2 Proposition)). *Let \mathcal{A} be the infinitesimal generator of the strongly continuous semigroup $(T(t))_{t \geq 0}$. Then, for every $u_0 \in \mathcal{D}(\mathcal{A})$ the problem*

$$\partial_t \mathbf{u}(t) = \mathcal{A}\mathbf{u}(t), \quad \mathbf{u}(0) = \mathbf{u}_0,$$

has the unique solution $\mathbf{u}(t) = T(t)\mathbf{u}_0 \in C^1(\mathbb{R}_+; X) \cap C(\mathbb{R}_+; \mathcal{D}(\mathcal{A}))$.

Note, that for a bounded operator \mathcal{A} the semigroup $(T(t))_{t \geq 0}$ is given as the exponential function

$$e^{At} = \sum_{k=0}^{\infty} \mathcal{A}^k \frac{t^k}{k!}.$$

For the possibly unbounded operator we still use the notation $(T(t))_{t \geq 0} = (e^{At})_{t \geq 0}$. The next result finally connects the semigroup theory with the Cauchy problem (3.24) and states that if we have a strongly continuous semigroup generated by an operator \mathcal{A} the related evolution problem has a unique solution.

Theorem 3.2.8. *Let \mathcal{A} be the infinitesimal generator of the strongly continuous semigroup $(e^{t\mathcal{A}})_{t \geq 0}$ and $\mathbf{u}_0 \in \mathcal{D}(\mathcal{A})$. Moreover, assume that either $\mathbf{f} \in C^1(0, T; X)$ or $\mathbf{f} \in C(0, T; \mathcal{D}(\mathcal{A}))$. Then, there exists a unique solution $\mathbf{u} \in C^1(0, T; X) \cap C(0, T; \mathcal{D}(\mathcal{A}))$ of (3.24) given by*

$$\mathbf{u}(t) = e^{t\mathcal{A}}\mathbf{u}_0 + \int_0^t e^{(t-s)\mathcal{A}}\mathbf{f}(s) ds.$$

The question remains whether the operator \mathcal{A} is in fact the generator of a strongly continuous semigroup. We give two special cases of operators and corresponding theorems that state the assumptions for the operators to actually generate a C_0 -semigroup. The first concept is the one of a dissipative operator.

Definition 3.2.9 (Dissipative operator). *An operator is called dissipative if*

$$(\mathcal{A}x, x)_X \leq 0. \quad \text{for all } x \in X.$$

We call the operator monotone or accretive if $-\mathcal{A}$ is dissipative, i.e.,

$$(\mathcal{A}x, x)_X \geq 0 \quad \text{for all } x \in X.$$

The Lumer-Phillips theorem by Günter Lumer and Ralph Phillips covers the case where the operator \mathcal{A} is dissipative and satisfies the so-called *range condition*.

Theorem 3.2.10 (Lumer-Phillips. (Engel and Nagel, 2000, Chapter II, Corollary 3.20)). *Let \mathcal{A} be a linear operator with domain $\mathcal{D}(\mathcal{A})$ on a Hilbert space X . Then the following statements are equivalent:*

- (i) *\mathcal{A} is densely defined and generates a contraction semigroup.*
- (ii) *\mathcal{A} is dissipative and $\text{ran}(\lambda - \mathcal{A}) = X$ for some $\lambda > 0$.*

Next to dissipative operators generating C_0 -semigroups the notion of skew-adjoint operators gives a sufficient condition for the operator to generate a C_0 -group. We will see below that the Maxwell operator is skew-adjoint.

Definition 3.2.11 (Adjoint operator). *Let $\mathcal{A} : \mathcal{D}(\mathcal{A}) \rightarrow X$ be a linear operator with $\overline{\mathcal{D}(\mathcal{A})} = X$. The adjoint operator \mathcal{A}^* of \mathcal{A} is defined as follows: The domain $\mathcal{D}(\mathcal{A}^*)$ consists of all $y \in X$ such that there exists $z \in X$ satisfying*

$$(\mathcal{A}x, y)_X = (x, z)_X \quad \text{for all } x \in \mathcal{D}(\mathcal{A}).$$

For $y \in \mathcal{D}(\mathcal{A}^)$, the adjoint is defined as $\mathcal{A}^*y = z$.*

With the adjoint of an operator we now give the definition of a skew-adjoint operator.

Definition 3.2.12 (Skew-symmetric and skew-adjoint operator). *Let $\mathcal{A} : \mathcal{D}(\mathcal{A}) \rightarrow X$ be densely defined. The operator \mathcal{A} is called skew-symmetric if $\mathcal{A}x = -\mathcal{A}^*x$ for all $x \in \mathcal{D}(\mathcal{A})$. It is called skew-adjoint if it is skew-symmetric and $\mathcal{D}(\mathcal{A}) = \mathcal{D}(\mathcal{A}^*)$, which results in $\mathcal{A}^* = -\mathcal{A}$.*

The next lemma gives a criterion to check the skew-adjointness of a skew-symmetric operator.

Lemma 3.2.13. *Let $\mathcal{A} : \mathcal{D}(\mathcal{A}) \rightarrow X$ be skew-symmetric. Then, \mathcal{A} is skew-adjoint if $\text{id} \pm \mathcal{A}$ has dense range, i.e. if*

$$\overline{\text{range}(\text{id} \pm \mathcal{A})} = X.$$

Stone's Theorem 3.2.16 shows that skew-adjoint operators generate unitary C_0 -groups. Therefore, we give the definition of those first.

Definition 3.2.14 (Unitary group). *A C_0 -group $(T(t))_{t \in \mathbb{R}}$ is called a unitary group if*

$$\|T(t)x\|_X = \|x\|_X \quad \text{for all } x \in X, t \in \mathbb{R}.$$

It is worth noting the relation between contraction semigroups and unitary groups.

Remark 3.2.15. *If \mathcal{A} generates a unitary C_0 -group $(T(t))_{t \in \mathbb{R}}$, then $-\mathcal{A}$ generates a contraction semigroup $(\tilde{T}(t))_{t \geq 0}$.*

Now we have all ingredients for the statement of Stone's theorem.

Theorem 3.2.16 (Stone's Theorem (Engel and Nagel, 2000, Chapter II, 3.24 Theorem)). *Let $\mathcal{A} : \mathcal{D}(\mathcal{A}) \rightarrow X$ be a linear operator with dense domain $\overline{\mathcal{D}(\mathcal{A})} = X$. Then the following statements are equivalent:*

- (i) \mathcal{A} generates a unitary C_0 -group $(T(t))_{t \in \mathbb{R}}$ on X .
- (ii) \mathcal{A} is skew-adjoint.

We will see in Section 3.2.2 that, in the right analytical setting, the Maxwell operator \mathcal{A} is a skew-adjoint operator. If we consider the linear Maxwell system, i.e., with constitutive law (3.4) and thus without polarization and magnetization, the results from above would be enough to study the resulting system. But since we analyze Maxwell's equations with polarization and magnetization we need some more results from semigroup theory. The additional operators that occur due to the polarization and magnetization coupling will be bounded and in some special cases even dissipative. For both cases there are results about perturbations of operators that generate contracting semigroups, such as the Maxwell operator. We start with the bounded perturbation.

Theorem 3.2.17 (Bounded perturbation (Engel and Nagel, 2000, Chapter III, 1.3 Bounded Perturbation Theorem)). *Let \mathcal{R} be a bounded operator and \mathcal{A} generate a C_0 -semigroup $(T(t))_{t \geq 0}$ satisfying $\|T(t)\|_{X \leftarrow X} \leq Ce^{\beta t}$ for all $t \geq 0$ and some constants $C \geq 1$ and $\beta \in \mathbb{R}$. Then, the sum $\mathcal{A} + \mathcal{R}$ with $\mathcal{D}(\mathcal{A} + \mathcal{R}) = \mathcal{D}(\mathcal{A})$ generates a C_0 -semigroup $(S(t))_{t \geq 0}$ satisfying*

$$\|S(t)\|_{X \leftarrow X} \leq Ce^{(\beta + C\|\mathcal{R}\|_{X \leftarrow X})t} \quad \text{for all } t \geq 0.$$

In our example problem, the Debye polarization model, we will see that the resulting operator is not just bounded but dissipative. There are more cases in which this is the case and the dissipativity enables us to get a better result for the semigroup generated by the perturbed operator.

Theorem 3.2.18 (Dissipative perturbation (Engel and Nagel, 2000, Chapter III, 2.7 Theorem)). *Let \mathcal{A} generate the contraction semigroup $(T(t))_{t \geq 0}$ and \mathcal{R} be dissipative. Assume that \mathcal{R} is \mathcal{A} -bounded with a constant $a < 1$, i.e., there exists $b \geq 0$ such that*

$$\|\mathcal{R}x\|_X \leq a\|\mathcal{A}x\|_X + b\|x\|_X \quad \text{for all } x \in \mathcal{D}(\mathcal{A}).$$

Then $\mathcal{A} + \mathcal{R}$ with $\mathcal{D}(\mathcal{A} + \mathcal{R}) = \mathcal{D}(\mathcal{A})$ generates a contraction semigroup $(S(t))_{t \geq 0}$.

With this perturbation result we have everything at hand that we need for the analysis of the Maxwell system in the next section.

3.2.2 Application to Maxwell's equations

The purpose of this section is to apply the results from semigroup theory to the general Maxwell system with polarization and magnetization. Thus, we have to give the right functional analytical framework.

Recall that the general Maxwell system (3.21) has size $n = 3(2 + N_E + N_H)$. The state space X is chosen to be $L^2(\Omega; \mathbb{R}^n)$ the space of square integrable functions on Ω with values in \mathbb{R}^n . In order to apply the

results from the previous section we rewrite the general Maxwell system (3.21) as an abstract Cauchy problem (3.24). Thus, we have to define the domain of the corresponding operator \mathcal{A} . Observe, that the solution consists of $2 + N_E + N_H$ components that map from Ω to \mathbb{R}^3 . Therefore, we introduce the weak curl of a function in $L^2(\Omega; \mathbb{R}^3)$.

Definition 3.2.19 (Weak curl). *Let $f \in L^2(\Omega; \mathbb{R}^3)$. If there exists a function $g \in L^2(\Omega; \mathbb{R}^3)$ such that*

$$\int_{\Omega} f(x) \cdot \operatorname{curl} \phi(x) \, dx = \int_{\Omega} g(x) \cdot \phi(x) \, dx \quad \text{for all } \phi \in C_0^\infty(\Omega; \mathbb{R}^3),$$

holds true, we call g the weak curl of f . Moreover, we denote $\operatorname{curl} f = g$.

Now the space of functions that is related to Maxwell's equations is the space $H(\operatorname{curl}, \Omega)$.

Definition 3.2.20 ($H(\operatorname{curl}, \Omega)$, $H_0(\operatorname{curl}, \Omega)$). *The space of all $L^2(\Omega; \mathbb{R}^3)$ functions with existing weak curl in $L^2(\Omega; \mathbb{R}^3)$ is defined as*

$$H(\operatorname{curl}, \Omega) := \{f \in L^2(\Omega; \mathbb{R}^3) : \operatorname{curl} f \in L^2(\Omega; \mathbb{R}^3)\}.$$

Equipped with the scalar product

$$(f, g)_{H(\operatorname{curl}, \Omega)} = (f, g)_{L^2(\Omega; \mathbb{R}^3)} + (\operatorname{curl} f, \operatorname{curl} g)_{L^2(\Omega; \mathbb{R}^3)} \quad \text{for all } f, g \in H(\operatorname{curl}, \Omega),$$

this space is a Hilbert space. Moreover, as usual we define

$$H_0(\operatorname{curl}, \Omega) := \overline{C_0^\infty(\Omega; \mathbb{R}^3)}^{\|\cdot\|_{H(\operatorname{curl}, \Omega)}},$$

which suits the boundary condition for the electric field.

As it turns out in Theorem 3.2.21 below the space $H(\operatorname{curl}, \Omega)$ is the right space for the analysis of the Maxwell's equations. We rewrite the Maxwell system (3.21) in the way we introduced the Cauchy problem (3.24) in the previous section. Then we verify the properties of the resulting operator to get the wellposedness result.

From now on we assume that the parameter \mathbf{M} is bounded and positive definite, i.e. $\mathbf{M} \in L^\infty(\Omega; \mathbb{R}^{n \times n})$ and there exist constants $\alpha, C_{\mathbf{M}} > 0$ such that

$$\alpha|\phi|^2 \leq \mathbf{M}(x)\phi \cdot \phi \leq C_{\mathbf{M}}|\phi|^2, \quad (3.26)$$

for all $\phi \in \mathbb{R}^n$ and almost every $x \in \Omega$. Note that due to the properties of the parameter \mathbf{M} we define a weighted scalar product on $X = L^2(\Omega; \mathbb{R}^n)$ as

$$(\phi, \psi)_X := \int_{\Omega} \mathbf{M}(x)\phi(x) \cdot \psi(x) \, dx \quad \text{for all } \phi, \psi \in X.$$

This scalar product induces a norm $\|\cdot\|_X$, which is equivalent to the standard L^2 -norm. Consider equation (3.21) and rewrite it as

$$\partial_t \mathbf{u}(t, x) = \mathbf{M}(x)^{-1} (-\mathbf{A}\mathbf{u}(t, x) - \mathbf{R}(x)\mathbf{u}(t, x) + \mathbf{g}(t, x)).$$

Thus, with the notation from Section 3.2.1 we have

$$\partial_t \mathbf{u}(t, x) = \mathcal{B}\mathbf{u}(t, x) + \mathbf{f}(t, x), \quad (3.27)$$

where the operator \mathcal{B} and the right-hand side \mathbf{f} are defined as

$$\mathcal{B} = -\mathbf{M}^{-1}(\mathbf{A} + \mathbf{R}), \quad \mathbf{f} = \mathbf{M}^{-1}\mathbf{g}.$$

We split the operator \mathcal{B} into the Maxwell operator and the damping operator

$$\mathcal{B} = \mathcal{A} + \mathcal{R},$$

where, for $x \in \Omega$,

$$\mathcal{A}(x) = -\mathbf{M}(x)^{-1}\mathbf{A}, \tag{3.28}$$

$$\mathcal{R}(x) = -\mathbf{M}(x)^{-1}\mathbf{R}(x). \tag{3.29}$$

We first consider the Maxwell operator \mathcal{A} . It is well known that the following holds true.

Theorem 3.2.21. *The Maxwell operator \mathcal{A} from (3.28) with domain \mathcal{A} given as*

$$\mathcal{D}(\mathcal{A}) = \mathbf{H}_0(\text{curl}, \Omega) \times \mathbf{H}(\text{curl}, \Omega)^{N_E} \times \mathbf{H}(\text{curl}, \Omega) \times \mathbf{H}(\text{curl}, \Omega)^{N_H},$$

or

$$\mathcal{D}(\mathcal{A}) = \mathbf{H}_0(\text{curl}, \Omega) \times \mathbf{L}^2(\Omega; \mathbb{R}^3)^{N_E} \times \mathbf{H}(\text{curl}, \Omega) \times \mathbf{L}^2(\Omega; \mathbb{R}^3)^{N_H},$$

generates a unitary C_0 -group e^{At} , i.e.,

$$\|e^{At}\|_{X \leftarrow X} = 1.$$

Proof. The proof relies on an application of Stone's theorem 3.2.16. Therefore, the skew-adjointness of the Maxwell operator has to be shown, which is done by the use of Lemma 3.2.13. Thus, we have to ensure

$$\overline{\text{range}(\text{id} \pm \mathcal{A})} = X,$$

which is shown using Lax-Milgram, see e.g., (Brenner and Scott, 2008, Theorem 2.7.7). For the details see (Sturm, 2017, Theorem 1.25). Further note, that the Maxwell operator only acts on the components of the solution that represent the electric and the magnetic field. Therefore, for the polarization and magnetization we may either choose $\mathbf{L}^2(\Omega; \mathbb{R}^3)^{N_E}$ and $\mathbf{L}^2(\Omega; \mathbb{R}^3)^{N_H}$ or $\mathbf{H}(\text{curl}, \Omega)^{N_E}$ and $\mathbf{H}(\text{curl}, \Omega)^{N_H}$. \square

From Theorem 3.2.21 together with Theorem 3.2.8 we immediately get the wellposedness of the Maxwell system without polarization or magnetization, i.e., with constitutive relations as in (3.4). We now turn to the damping parameter \mathbf{R} , which is a perturbation of the classic Maxwell system. We will assume that the parameter \mathbf{R} is bounded, i.e., $\mathbf{R} \in \mathbf{L}^\infty(\Omega; \mathbb{R}^{n \times n})$. Now we can apply Theorem 3.2.17 and get a wellposed problem, but in general with an exponential growth in the stability estimates. To be more precise, from Theorem 3.2.17 and Theorem 3.2.21 we see that the semigroup $(S(t))_{t \in \geq 0}$ generated by $\mathcal{B} = \mathcal{A} + \mathcal{R}$ satisfies

$$\|S(t)\|_{X \leftarrow X} \leq e^{\|\mathcal{R}\|_{X \leftarrow X} t} \quad \text{for all } t \geq 0.$$

This leads to an exponential growth in the stability estimate.

Theorem 3.2.22. *Let $\mathbf{M}, \mathbf{R} \in L^\infty(\Omega, \mathbb{R}^{n \times n})$, \mathbf{M} symmetric, uniformly positive definite and let $\mathbf{u}_0(x) \in \mathcal{D}(\mathcal{A})$, $\mathbf{f} \in C([0, T], \mathcal{D}(\mathcal{A}))$ or $\mathbf{f} \in C^1([0, T], \mathbf{X})$.*

Then we find a unique solution $\mathbf{u} \in C^1([0, T], \mathbf{X}) \cap C([0, T], \mathcal{D}(\mathcal{A}))$ of (3.27), which satisfies

$$\|\mathbf{u}(t)\|_{\mathbf{X}} \leq e^{t\|\mathcal{R}\|_{\mathbf{X} \leftarrow \mathbf{X}}} \left(\|\mathbf{u}_0\|_{\mathbf{X}} + t \|\mathbf{f}\|_{L^\infty(0, t; \mathbf{X})} \right). \quad (3.30)$$

Proof. This is a direct consequence of Theorem 3.2.8 and Theorem 3.2.21 combined with the perturbation Theorem 3.2.17 for bounded operators. \square

With more information about the parameter \mathbf{R} we get a better estimate.

If \mathbf{R} is at least positive semi-definite, then the associated operator \mathcal{R} is going to be dissipative and therefore we can apply Theorem 3.2.18, which states that the perturbed operator $\mathcal{B} = \mathcal{A} + \mathcal{R}$ generates a contraction semigroup $(S(t))_{t \geq 0}$, i.e.

$$\|S(t)\|_{\mathbf{X} \leftarrow \mathbf{X}} \leq 1 \quad \text{for all } t \geq 0.$$

Theorem 3.2.23. *Let $\mathbf{M}, \mathbf{R} \in L^\infty(\Omega, \mathbb{R}^{n \times n})$, \mathbf{M} symmetric, uniformly positive definite, \mathbf{R} positive semi-definite, and let $\mathbf{u}_0(x) \in \mathcal{D}(\mathcal{A})$, $\mathbf{f} \in C([0, T], \mathcal{D}(\mathcal{A}))$ or $\mathbf{f} \in C^1([0, T], \mathbf{X})$.*

Then we find a unique solution $\mathbf{u} \in C^1([0, T], \mathbf{X}) \cap C([0, T], \mathcal{D}(\mathcal{A}))$ of (3.27) which satisfies

$$\|\mathbf{u}(t)\|_{\mathbf{X}} \leq \|\mathbf{u}_0\|_{\mathbf{X}} + t \|\mathbf{f}\|_{L^\infty(0, t; \mathbf{X})}. \quad (3.31)$$

Proof. We use the Theorem 3.2.18 about the dissipative perturbation with $a = 0$ and $b = \|\mathcal{R}\|_{\mathbf{X} \leftarrow \mathbf{X}} < \infty$ since \mathbf{R} is bounded. The result is a direct consequence of Theorem 3.2.8 and Theorem 3.2.21. See also with (Hipp et al., 2019, Theorem 2.4). \square

We point out that, although in general the parameter \mathbf{R} does not need to be positive semi-definite, in our examples it is always.

The Maxwell system with conductivity

Let us apply the wellposedness theory to the classic Maxwell system with conductivity (3.11). In addition to the properties of \mathbf{M} in (3.26) we assume that the conductivity σ is positive semi-definite and bounded. This yields that the parameter \mathbf{R} is positive semi-definite and bounded such that we can apply Theorem 3.2.23. This immediately yields the wellposedness of the system (3.11). Moreover, the solution satisfies

$$\|\mathbf{u}(t)\|_{\mathbf{X}} \leq \|\mathbf{u}_0\|_{\mathbf{X}} + t \|\mathbf{f}\|_{L^\infty(0, t; \mathbf{X})}.$$

Note that the space \mathbf{X} in this setting is $L^2(\Omega; \mathbb{R}^6)$ since we only have the electric and magnetic field present.

The Maxwell–Debye system

Now we consider our second example, the Maxwell–Debye system given in (3.15). As shown in Section 3.1.3 we can write it in the form

$$\mathbf{M} \partial_t \mathbf{u}(t) + (\mathbf{A} + \mathbf{R}) \mathbf{u}(t) = \mathbf{g}(t).$$

The parameters \mathbf{M} and \mathbf{R} contain the relaxation time τ , which is strictly positive as well as $\Delta\varepsilon = \varepsilon_s - \varepsilon_\infty > 0$. Before we apply Theorem 3.2.18 we multiply the system by the following matrix

$$\mathbf{Q}(x) := \begin{pmatrix} \mathbf{I}_{3 \times 3} & & \\ & \varepsilon_0^{-1} \tau(x)^{-1} \Delta\varepsilon(x)^{-1} & \\ & & \mathbf{I}_{3 \times 3} \end{pmatrix}, \quad (3.32)$$

which is positive definite. This only affects the parameter matrices \mathbf{M} and \mathbf{R} since the Maxwell operator \mathbf{A} and the right-hand side \mathbf{g} have no entry in the component of the polarization. We thus get

$$\mathbf{Q}(x)\mathbf{M}(x) = \begin{pmatrix} \varepsilon_0 \varepsilon(x) & & \\ & \varepsilon_0^{-1} \Delta\varepsilon(x)^{-1} & \\ & & \mu_0 \end{pmatrix}, \quad \mathbf{Q}(x)\mathbf{R}(x) = \begin{pmatrix} \varepsilon_0 \tau(x)^{-1} \Delta\varepsilon(x) & -\tau(x)^{-1} & \mathbf{0} \\ -\tau(x)^{-1} & \varepsilon_0^{-1} \tau(x)^{-1} \Delta\varepsilon(x)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

The matrix $\mathbf{Q}(x)\mathbf{M}(x)$ still has the same properties as \mathbf{M} , i.e., it is positive definite and bounded. Moreover, the matrix $\mathbf{Q}(x)\mathbf{R}(x)$ is symmetric and positive semi-definite. But this is exactly what we need in order to apply the wellposedness Theorem 3.2.23 with a dissipative perturbation of the Maxwell operator. In this setting we have $X = L^2(\Omega; \mathbb{R}^9)$, and we use the inner product

$$(\phi, \psi)_X = \int_{\Omega} \mathbf{Q}(x)\mathbf{M}(x)\phi(x) \cdot \psi(x) \, dx \quad \text{for all } \phi, \psi \in X.$$

Thus, we immediately find that the Maxwell–Debye system is wellposed and again satisfies the stability bound

$$\|\mathbf{u}(t)\|_X \leq \|\mathbf{u}_0\|_X + t \|\mathbf{M}^{-1}\mathbf{g}\|_{L^\infty(0,t;X)}.$$

In the next section we briefly comment on the effect of dissipative operators to the energy of the Maxwell system.

3.2.3 Energy and stability

The energy of a Maxwell system is given as the quantity

$$\mathcal{E}(t) := \frac{1}{2} \|\mathbf{u}(t)\|_X^2.$$

For a Maxwell system with vanishing right-hand side \mathbf{g} and a positive semi-definite damping parameter \mathbf{R} we find

$$\partial_t \mathcal{E}(t) = (\mathbf{M} \partial_t \mathbf{u}(t), \mathbf{u}(t)) = -((\mathbf{A} + \mathbf{R}) \mathbf{u}(t), \mathbf{u}(t)) \leq 0.$$

Thus, the energy of such a system is bounded by the initial energy

$$\mathcal{E}(T) \leq \mathcal{E}(0).$$

This property is also referred to as physically passive as found in [Cassier et al. \(2017\)](#).

Homogenization of Maxwell's equations in locally periodic media

In the first chapter we introduced the Maxwell's equations for the propagation of electromagnetic waves in general media. Our main interest are materials that exhibit heterogeneities at very small scales. For instance consider a mixture of two different materials that are artificially constructed. Examples of these materials are so-called meta materials, which may have unexpected properties such as negative refraction. This chapter is organized as follows. First we explain what the main difficulties in modeling of heterogeneous media are and why we need to use homogenization to overcome these issues in Section 4.1. Then, in Section 4.2 we consider the linear Maxwell system without polarization and magnetization, which has already been analyzed in many works. Since the dispersion effect will play a crucial role in the analysis we briefly consider such effects occurring in long time homogenization in general wave equations in Section 4.3. The main part of this chapter is collected in Section 4.4 where we finally consider the homogenization of the Maxwell system derived in Chapter 3. In the last Section 4.5 we briefly comment on homogenization techniques that are beyond the periodic setting considered here.

4.1 Heterogeneous materials and their restrictions in simulation

The interest of this work are materials that are periodically varying on small scales. Our understanding of small is the following: Eventually, we apply a finite element method (FEM) to a partial differential equation, which in our setting is a Maxwell system. To capture every aspect of the system we have to ensure that all oscillations are resolved by the finite element discretization. When we say that the parameters are oscillating on small scales, we mean that these scales are too small to be resolved by our finite element scheme. Moreover, we are interested in the setting where the micro-structure is small in comparison to the wavelength of the electromagnetic waves. The idea to overcome this problem stems from the fact that even for the most general Maxwell system we see in Theorem 3.2.22 that the solution \mathbf{u}^δ of the system satisfies the bound

$$\|\mathbf{u}^\delta(t)\|_X \leq e^{t\|\mathcal{R}\|_{X \leftarrow X}} \left(\|\mathbf{u}_0^\delta\|_X + t \|\mathbf{f}^\delta\|_{L^\infty(0,t;X)} \right). \quad (4.1)$$

Now if the initial value and the right-hand side can be bounded independent of δ , there exists a bounded sequence \mathbf{u}^δ . Since the space X is a Hilbert space there is a $\mathbf{u}^{\text{eff}} \in X$ and a subsequence still denoted by \mathbf{u}^δ such that

$$\mathbf{u}^\delta \rightharpoonup \mathbf{u}^{\text{eff}} \quad \text{weakly in } X.$$

This gives rise to the question which limit system the quantity \mathbf{u}^{eff} satisfies. The aim of homogenization is to find this limit system. As it turns out the limit system is characterized by homogeneous parameters that do not depend on the microscopic scale anymore and this is of course favorable for a finite element method. The literature on the field of homogenization is vast. The classical results presented here are found in [Allaire \(1992\)](#); [Jikov et al. \(1994\)](#); [Sánchez-Palencia \(1980\)](#). After this motivation of homogenization, let us now introduce some notation that will be used in the remainder. We begin with the definition of a periodic tensor, which is the basic definition related to the heterogeneous parameters.

Definition 4.1.1 (Periodic function). *Let $d \in \mathbb{N}$ and $\Omega \subseteq \mathbb{R}^d$. A tensor $g: \Omega \rightarrow \mathbb{R}^{n \times n}$ is called 1-periodic, if*

$$g(y + \mathbf{e}_k) = g(y),$$

for almost every $y \in \mathbb{R}^n$ and all $k = 1, \dots, d$. The quantity \mathbf{e}_k is the k -th canonical basis vector of \mathbb{R}^d .

Every 1-periodic tensor is defined through its values on the unit cube $Y := (-\frac{1}{2}, \frac{1}{2})^3$. Thus, we may call the tensor Y -periodic. To get highly oscillatory parameters we use the following notation. For a parameter $\delta > 0$ representing the microscopic scale and a 1-periodic tensor g the tensor $g^\delta(\cdot) := g(\frac{\cdot}{\delta})$ is δ -periodic, i.e.,

$$g^\delta(y + \delta \mathbf{e}_k) = g^\delta(y) \quad \text{for all } y \in Y.$$

The concept of periodicity is important in the homogenization process that follows. Nevertheless, we consider a slightly larger class of material parameters. These parameters may at least be locally periodic, which is equivalent to a separation of the scales as shown in the following definition.

Definition 4.1.2 (Locally periodic parameter). *Let $\delta > 0$. A tensor $\alpha^\delta: \Omega \rightarrow \mathbb{R}^{n \times n}$ is called locally δ -periodic if there exists a tensor $\alpha: \Omega \times \mathbb{R}^3 \rightarrow \mathbb{R}^{n \times n}$, which is Y -periodic in its second argument and $\alpha^\delta(x) = \alpha(x, \frac{x}{\delta})$ holds for almost every $x \in \Omega$. We call the function α the blueprint of α^δ .*

Since the periodicity of the parameters is visible in many quantities in this work we next define notation for spaces that consist of periodic functions. It will turn out that the crucial space we need is going to be a Sobolev type space. For that purpose we define the space of smooth Y -periodic functions as

$$C_\#^\infty(Y) := \{f \in C^\infty(\overline{Y}) : f(y + \mathbf{e}_k) = f(y) \text{ for all } y \in Y\}.$$

Now the usual definition of the Sobolev space as closure of infinitely differentiable functions follows as

$$\overline{H}_\#^1(Y) := \overline{C_\#^\infty(Y)}^{\|\cdot\|_{H^1(Y)}}.$$

Moreover, we define the quotient space consisting of the functions with vanishing zero mean

$$H_\#^1(Y) := \left\{ f \in \overline{H}_\#^1(Y) : \int_Y f(y) \, dy = 0 \right\}.$$

In general for a function space $F(Y)$ we denote a suitable subspace containing periodic functions as $F_{\#}(Y)$.

Next to the function spaces that are related to the periodicity of the parameters we present the concept of two-scale convergence, introduced in [Nguetseng \(1989\)](#), which is adjusted to the scale separation. This is the analytic tool that has been used in the past to rigorously analyze the homogenization especially in [Wellander \(2001\)](#).

Definition 4.1.3 (Two-scale convergence). *A sequence of functions $(u^\delta)_{\delta>0} \subseteq L^2(\Omega)$ two-scale converges to a limit $u^{\text{eff}} \in L^2(\Omega \times Y)$, if for any test function $v \in L^2(\Omega; C_{\#}^0(Y))$ we have*

$$\lim_{\delta \rightarrow 0} \int_{\Omega} u^\delta(x) v\left(x, \frac{x}{\delta}\right) dx = \int_{\Omega} \int_Y u^{\text{eff}}(x, y) v(x, y) dy dx.$$

Let us mention that in Section 4.4.2 we use the technique of periodic unfolding introduced in [Cioranescu et al. \(2002\)](#), which conceptually simplifies the proofs. For this method it was shown in ([Cioranescu et al., 2008](#), Proposition 2.14) that the two-scale convergence is equivalent to a weak convergence of the so-called unfolded functions.

4.2 Homogenization results for the Maxwell system with conductivity

Before we study the general Maxwell models that we introduced in Chapter 3 we briefly give the results for the classic linear material laws as in (3.4). This class of Maxwell's equations has been studied starting with [Sánchez-Palencia \(1980\)](#) where the classical method of asymptotic expansion is used to derive the homogeneous model. In [Wellander \(2001\)](#) the author gives rigorous proofs using the notion of two-scale convergence from Definition 4.1.3. Both of these works already considered Ohm's law with conductivity. This seems to be the right starting point to see the effects that occur when homogenizing Maxwell's equations. Let us start with the heterogeneous Maxwell system with a linear constitutive law, which is in time domain given as

$$\varepsilon^\delta(x) \partial_t \mathbf{E}^\delta(t, x) + \sigma^\delta(x) \mathbf{E}^\delta(t, x) - \text{curl} \mathbf{H}^\delta(t, x) = -\mathbf{J}(t, x), \quad (4.2a)$$

$$\mu^\delta(x) \partial_t \mathbf{H}^\delta(t, x) + \text{curl} \mathbf{E}^\delta(t, x) = \mathbf{0}, \quad (4.2b)$$

$$\text{div} [\varepsilon^\delta(x) \mathbf{E}^\delta(t, x)] = \rho(t, x), \quad (4.2c)$$

$$\text{div} [\mu^\delta(x) \mathbf{H}^\delta(t, x)] = 0. \quad (4.2d)$$

Here the parameters $\varepsilon^\delta, \sigma^\delta$ and μ^δ are locally δ -periodic functions that are assumed to be uniformly elliptic or coercive and uniformly bounded. We constrain the system (4.2) with perfectly conducting boundary conditions and with initial data

$$\mathbf{E}^\delta(0, x) = \mathbf{0}, \quad \mathbf{H}^\delta(0, x) = \mathbf{0}.$$

As explained before we consider the fine-scale parameter δ to be too small to be resolved by a classical finite element scheme. Thus, we follow the idea to derive a limit system for $\delta \rightarrow 0$ whose solution is close (in an appropriate sense) to the original heterogeneous system but where the parameters are no longer highly oscillatory. To get an idea how the effective system looks like, we use the easy tool of asymptotic expansion, which allows to formally derive the desired homogeneous system.

4.2.1 Asymptotic expansion for the Maxwell system with conductivity

We derive the effective system related to (4.2) via the use of asymptotic expansion. Recall that we defined the unit cell $Y = (-\frac{1}{2}, \frac{1}{2})^3$. The idea is to use the following ansatz

$$\mathbf{E}^\delta(t, x) = \mathbf{E}^{\text{eff}}(t, x) + \bar{\mathbf{E}}(t, x, \frac{x}{\delta}), \quad \mathbf{H}^\delta(t, x) = \mathbf{H}^{\text{eff}}(t, x) + \bar{\mathbf{H}}(t, x, \frac{x}{\delta}). \quad (4.3)$$

The quantities $\bar{\mathbf{E}}$ and $\bar{\mathbf{H}}$ are Y -periodic in y and have zero mean, i.e.,

$$\int_Y \bar{\mathbf{E}}(t, x, y) \, dy = 0, \quad \int_Y \bar{\mathbf{H}}(t, x, y) \, dy = 0 \quad \text{for all } (t, x) \in [0, T] \times \Omega. \quad (4.4)$$

Thus, substituting the ansatz (4.3) in the system (4.2) yields

$$\begin{aligned} [\varepsilon(\frac{x}{\delta})\partial_t + \sigma(\frac{x}{\delta})] (\mathbf{E}^{\text{eff}}(t, x) + \bar{\mathbf{E}}(t, x, \frac{x}{\delta})) - \text{curl}_x (\mathbf{H}^{\text{eff}}(t, x) + \bar{\mathbf{H}}(t, x, \frac{x}{\delta})) &= -\mathbf{J}(t, x), \\ \mu(\frac{x}{\delta})\partial_t (\mathbf{H}^{\text{eff}}(t, x) + \bar{\mathbf{H}}(t, x, \frac{x}{\delta})) + \text{curl}_x (\mathbf{E}^{\text{eff}}(t, x) + \bar{\mathbf{E}}(t, x, \frac{x}{\delta})) &= \mathbf{0}, \\ \text{div}_x [\varepsilon(\frac{x}{\delta}) (\mathbf{E}^{\text{eff}}(t, x) + \bar{\mathbf{E}}(t, x, \frac{x}{\delta}))] &= \rho(t, x), \\ \text{div}_x [\mu(\frac{x}{\delta}) (\mathbf{H}^{\text{eff}}(t, x) + \bar{\mathbf{H}}(t, x, \frac{x}{\delta}))] &= 0. \end{aligned}$$

Evaluating the curl with the chain rule yields expressions for different powers of δ . For δ^{-1} and $y = \frac{x}{\delta}$ we find

$$\text{curl}_y \bar{\mathbf{H}}(t, x, y) = \mathbf{0}, \quad \text{curl}_y \bar{\mathbf{E}}(t, x, y) = \mathbf{0},$$

which is equivalent to the existence of fields $\bar{\mathbf{H}}, \bar{\mathbf{E}} : [0, T] \times \Omega \times Y \rightarrow \mathbb{R}$ such that

$$\bar{\mathbf{H}}(t, x, y) = \nabla_y \bar{H}(t, x, y), \quad \bar{\mathbf{E}}(t, x, y) = \nabla_y \bar{E}(t, x, y) \quad \text{for all } (t, x, y) \in [0, T] \times \Omega \times Y. \quad (4.6)$$

On the level of δ^0 we find, using again $y = \frac{x}{\delta}$, the Maxwell system

$$\begin{aligned} [\varepsilon(y)\partial_t + \sigma(y)] (\mathbf{E}^{\text{eff}}(t, x) + \bar{\mathbf{E}}(t, x, y)) - \text{curl}_x \mathbf{H}^{\text{eff}}(t, x) - \text{curl}_x \bar{\mathbf{H}}(t, x, y) &= -\mathbf{J}(t, x), \\ \mu(y)\partial_t (\mathbf{H}^{\text{eff}}(t, x) + \bar{\mathbf{H}}(t, x, y)) + \text{curl}_x \mathbf{E}^{\text{eff}}(t, x) + \text{curl}_x \bar{\mathbf{E}}(t, x, y) &= \mathbf{0}, \\ \text{div}_x [\varepsilon(y) (\mathbf{E}^{\text{eff}}(t, x) + \bar{\mathbf{E}}(t, x, y))] &= \rho(t, x), \\ \text{div}_x [\mu(y) (\mathbf{H}^{\text{eff}}(t, x) + \bar{\mathbf{H}}(t, x, y))] &= 0. \end{aligned}$$

The final step is to take the mean value over Y and to use the vanishing mean values in (4.4) as well as (4.6). Thus, we end up with the two-scale Maxwell system

$$\int_Y [\varepsilon(y)\partial_t + \sigma(y)] (\mathbf{E}^{\text{eff}}(t, x) + \nabla_y \bar{E}(t, x, y)) \, dy - \text{curl}_x \mathbf{H}^{\text{eff}}(t, x) = -\mathbf{J}(t, x), \quad (4.8a)$$

$$\int_Y \mu(y)\partial_t (\mathbf{H}^{\text{eff}}(t, x) + \nabla_y \bar{H}(t, x, y)) \, dy + \text{curl}_x \mathbf{E}^{\text{eff}}(t, x) = \mathbf{0}, \quad (4.8b)$$

$$\text{div}_x \left[\int_Y \varepsilon(y) (\mathbf{E}^{\text{eff}}(t, x) + \nabla_y \bar{E}(t, x, y)) \, dy \right] = \rho(t, x), \quad (4.8c)$$

$$\text{div}_x \left[\int_Y \mu(y) (\mathbf{H}^{\text{eff}}(t, x) + \nabla_y \bar{H}(t, x, y)) \, dy \right] = 0. \quad (4.8d)$$

Next we determine representations for $\bar{\mathbf{E}}$ and $\bar{\mathbf{H}}$. Therefore, we test (4.8a) and (4.8b) with $\nabla_y v$ for $v \in \mathbf{H}_{\#}^1(Y)$. This yields after integration by parts, using $\operatorname{div} \operatorname{curl} = 0$

$$\begin{aligned} \int_Y [\varepsilon(y)\partial_t + \sigma(y)] (\mathbf{E}^{\text{eff}}(t, x) + \nabla_y \bar{\mathbf{E}}(t, x, y)) \cdot \nabla_y v(y) \, dy &= \mathbf{0}, \\ \int_Y \mu(y)\partial_t (\mathbf{H}^{\text{eff}}(t, x) + \nabla_y \bar{\mathbf{H}}(t, x, y)) \cdot \nabla_y v(y) \, dy &= \mathbf{0}. \end{aligned}$$

From these two equations we determine $\bar{\mathbf{E}}$ and $\bar{\mathbf{H}}$ as

$$\bar{\mathbf{E}}(t, x, y) = w^\varepsilon(y) \cdot \mathbf{E}^{\text{eff}}(t, x) + \int_0^t w^\chi(t-s, y) \cdot \mathbf{E}^{\text{eff}}(s, x) \, ds, \quad (4.9a)$$

$$\bar{\mathbf{H}}(t, x, y) = w^\mu(y) \cdot \mathbf{H}^{\text{eff}}(t, x), \quad (4.9b)$$

where the cell correctors $w^\varepsilon, w^\mu, w^\chi(t, \cdot) \in \mathbf{H}_{\#}^1(Y)$ solve the so-called micro problems for all $\ell = 1, 2, 3$ and all $v \in \mathbf{H}_{\#}^1(Y)$

$$\begin{aligned} \int_Y \varepsilon(y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(y)) \cdot \nabla_y v(y) \, dy &= 0, \quad \int_Y \mu(y) (\mathbf{e}_\ell + \nabla_y w_\ell^\mu(y)) \cdot \nabla_y v(y) \, dy = 0, \\ \int_Y (\varepsilon(y)\partial_t \nabla_y w_\ell^\chi(t, y) + \sigma(y)\nabla_y w_\ell^\chi(t, y)) \cdot \nabla_y v(y) \, dy &= 0 \quad \text{for } t \in (0, T], \\ \int_Y (\varepsilon(y)\nabla_y w_\ell^\chi(0, y) + \sigma(y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(y))) \cdot \nabla_y v(y) \, dy &= 0. \end{aligned}$$

Now we use the definition of the correctors in (4.9) and insert it in the two-scale limit system (4.8), which results in

$$\varepsilon^{\text{eff}} \partial_t \mathbf{E}^{\text{eff}}(t, x) + \partial_t \mathbf{P}^{\text{eff}}(t, x) + \sigma^{\text{eff}} \mathbf{E}^{\text{eff}}(t, x) + \int_0^t \zeta^{\text{eff}}(t-s) \mathbf{E}^{\text{eff}}(s, x) \, ds - \operatorname{curl} \mathbf{H}^{\text{eff}}(t, x) = -\mathbf{J}(t, x), \quad (4.10a)$$

$$\mu^{\text{eff}} \partial_t \mathbf{H}^{\text{eff}}(t, x) + \operatorname{curl} \mathbf{E}^{\text{eff}}(t, x) = \mathbf{0}, \quad (4.10b)$$

$$\operatorname{div} [\varepsilon^{\text{eff}} \mathbf{E}^{\text{eff}}(t, x) + \mathbf{P}^{\text{eff}}(t, x)] = \rho(t, x), \quad (4.10c)$$

$$\operatorname{div} [\mu^{\text{eff}} \mathbf{H}^{\text{eff}}(t, x)] = 0. \quad (4.10d)$$

The effective polarization is defined as

$$\mathbf{P}^{\text{eff}}(t, x) = \int_0^t \chi^{\text{eff}}(t-s) \mathbf{E}^{\text{eff}}(s, x) \, ds.$$

The effective parameters are given as mean values over the unit cell as

$$\begin{aligned} \varepsilon_{k,\ell}^{\text{eff}} &= \int_Y \varepsilon(y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(y)) \cdot \mathbf{e}_k \, dy, & \chi^{\text{eff}}(t)_{k,\ell} &= \int_Y \varepsilon(y) \nabla_y w_\ell^\chi(t, y) \cdot \mathbf{e}_k \, dy, \\ \sigma_{k,\ell}^{\text{eff}} &= \int_Y \sigma(y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(y)) \cdot \mathbf{e}_k \, dy, & \zeta^{\text{eff}}(t)_{k,\ell} &= \int_Y \sigma(y) \nabla_y w_\ell^\chi(t, y) \cdot \mathbf{e}_k \, dy, \\ \mu_{k,\ell}^{\text{eff}} &= \int_Y \mu(y) (\mathbf{e}_\ell + \nabla_y w_\ell^\mu(y)) \cdot \mathbf{e}_k \, dy. \end{aligned}$$

What we observe here is that the homogenization process introduces a polarization in our Maxwell system. Moreover, the quantity

$$\sigma^{\text{eff}} \mathbf{E}^{\text{eff}}(t, x) + \int_0^t \zeta^{\text{eff}}(t-s) \mathbf{E}^{\text{eff}}(s, x) ds,$$

could be interpreted as a version of Ohm's law, i.e., in frequency domain this reads

$$(\sigma^{\text{eff}} + \zeta^{\text{eff}}(\omega)) \hat{\mathbf{E}}(\omega, x).$$

Remark 4.2.1. *It is worth noting that there are other possibilities to introduce cell correctors and corresponding effective parameters. For instance, we may introduce the corrector $w^\sigma \in \mathbf{H}_{\#}^1(Y)$ which is directly related to the conductivity and solves for all $\ell = 1, 2, 3$ and all $v \in \mathbf{H}_{\#}^1(Y)$*

$$\int_Y \sigma(y) (\mathbf{e}_\ell + \nabla_y w_\ell^\sigma(y)) \cdot \nabla_y v(y) dy = 0.$$

Using this corrector we define slightly different effective parameters as

$$\hat{\chi}^{\text{eff}}(t)_{k,\ell} = \int_Y \varepsilon(y) \nabla_y w_\ell^\chi(t, y) \cdot (\mathbf{e}_k + \nabla_y w_k^\sigma(y)) dy, \quad \hat{\sigma}_{k,\ell}^{\text{eff}} = \int_Y \sigma(y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(y)) \cdot (\mathbf{e}_k + \nabla_y w_k^\sigma(y)) dy,$$

and a different polarization

$$\hat{\mathbf{P}}^{\text{eff}}(t, x) = \int_0^t \hat{\chi}^{\text{eff}}(t-s) \mathbf{E}^{\text{eff}}(s, x) ds.$$

Thus, an equivalent way to write (4.10a) is

$$\partial_t \left(\varepsilon^{\text{eff}} \mathbf{E}^{\text{eff}}(t, x) + \hat{\mathbf{P}}^{\text{eff}}(t, x) \right) + \hat{\sigma}^{\text{eff}} \mathbf{E}^{\text{eff}}(t, x) - \text{curl} \mathbf{H}^{\text{eff}}(t, x) = -\mathbf{J}(t, x).$$

We emphasize again that this is just a formal derivation without rigorous proofs. In Section 4.4, we give the rigorous theoretical result obtained using results from Bokil et al. (2018); Bossavit et al. (2005) via periodic unfolding, which holds true for a more general Maxwell system. In the next section we briefly comment on other effects of homogenization that may occur.

4.3 Long time dispersive effects in homogenization of general wave equations

In our Maxwell-Debye model the dispersion is present from the beginning but there are also dispersive effects occurring in wave equations when one considers long time intervals. With long, we mean in the context of homogenization times of order $T = \delta^{-2}$ and, since $\delta \ll 1$ is assumed to be very small, this is indeed a long time. For general wave equations these effects have been studied analytically as well as numerically in various works. On the analytic side one observes that, due to the small oscillations in the heterogeneous system over long time, dispersion enters, which is not covered by the effective solution. Thus, in order to overcome this issue new homogeneous equations were derived in Abdulle and Pouchon (2016); Dohnal et al. (2014, 2015) whose solutions are able to reflect the dispersive effects at least for longer times. Moreover, in Abdulle et al. (2014) a numerical approach is presented that also include the heterogeneous multiscale method over long time.

4.4 Homogenization of the Maxwell system

This section is dedicated to the homogenization of the general Maxwell system (3.21). We first show the system in its heterogeneous form in Section 4.4.1 and state some general assumptions. Then we introduce the periodic unfolding method as an analytical tool in Section 4.4.2 and finally present the homogenized Maxwell system in Section 4.4.3. The main result in Theorem 4.4.6 is taken from (Bokil et al., 2018, Theorem 5.2). The principal contribution of this chapter follows in Section 4.4.4 where we derive an equivalent formulation of the effective system, which we think is more suitable in the context of finite element methods. Moreover, in Sections 4.4.6, and 4.4.7 we show symmetry and exponential decay properties of the parameters. These properties are crucial for the error analysis in Chapter 5 and for the time integration in Chapter 6. In addition, we use the structure of the parameters for a refined wellposedness result in Section 4.4.8.

4.4.1 Heterogeneous Maxwell system

We assume that the material in the domain Ω has a locally periodic micro-structure characterized by the parameters \mathbf{M} and \mathbf{R} with periodically varying spatial coefficients. The periodic structure is assumed to be characterized by an elementary micro-structure with size $\delta > 0$. This is expressed by δ dependent parameters \mathbf{M}^δ , \mathbf{R}^δ , initial values \mathbf{u}_0^δ and exterior source \mathbf{g}^δ . Given δ we obtain a family of electromagnetic fields $\mathbf{u}^\delta : [0, T] \rightarrow \mathcal{D}(\mathcal{A})$, which are solutions to the evolution problem

$$\mathbf{M}^\delta(x) \partial_t \mathbf{u}^\delta(t, x) + \mathbf{R}^\delta(x) \mathbf{u}^\delta(t, x) + \mathbf{A} \mathbf{u}^\delta(t, x) = \mathbf{g}^\delta(t, x), \quad \text{in } (0, T) \times \Omega, \quad (4.11a)$$

$$\mathbf{u}^\delta(0, x) = \mathbf{u}_0^\delta(x), \quad \text{in } \Omega, \quad (4.11b)$$

$$\mathbf{n} \times \mathbf{u}_1^\delta(t, x) = \mathbf{0}, \quad \text{on } (0, T) \times \partial\Omega. \quad (4.11c)$$

As explained in Section 4.1 we are interested in the asymptotic behavior of the solution \mathbf{u}^δ when the periodicity length δ tends to zero. In view of the bound (4.1) we make the following assumption on the initial data and the source

$$\begin{aligned} \mathbf{u}_0^\delta &\rightarrow \mathbf{u}_0, \quad \text{strongly in } \mathcal{D}(\mathbf{A}), \\ \mathbf{g}^\delta &\rightarrow \mathbf{g}, \quad \text{strongly in } H^1(0, T; L^2(\Omega; \mathbb{R}^n)). \end{aligned} \quad (4.12)$$

With these assumptions we eventually get the desired uniform bound on the solution and thus, the existence of a subsequence, whose limit is the candidate for the homogenized solution. We already mentioned that we follow the approach of Bokil et al. (2018) where the periodic unfolding method is used to derive the effective system. Thus, we briefly introduce the method used in the next section.

4.4.2 Periodic unfolding as generalization of two-scale convergence

In this section we introduce the periodic unfolding method, which was first introduced in Cioranescu et al. (2002). A collection of many results and proofs can be found in Cioranescu et al. (2008) where the Dirichlet problem as standard homogenization example is considered. In Bossavit et al. (2005) the authors showed additional results for the unfolding operator that are related to Maxwell's equations, which were used in numerical experiments in Banks et al. (2006). For completeness, we repeat these results here.

Recall that we denote by $Y = (-\frac{1}{2}, \frac{1}{2})^3$ the unit/reference cell. For a.e. $z \in \mathbb{R}^3$ let $[z]_Y$ be the unique element belonging to \mathbb{Z}^3 such that $z - [z]_Y \in Y$, so that we may write $z = [z]_Y + \{z\}_Y$ for a.e. $z \in \mathbb{R}^3$. Consequently, for all $\delta > 0$ we get the unique decomposition

$$x = \delta \left(\left[\frac{x}{\delta} \right]_Y + \left\{ \frac{x}{\delta} \right\}_Y \right), \quad \text{for a.e. } x \in \mathbb{R}^3. \quad (4.13)$$

We assume that our parameters \mathbf{M}^δ and \mathbf{R}^δ are locally periodic. More precisely we assume that according to the decomposition in (4.13) there exist two parameters \mathbf{M} and \mathbf{R} such that

$$\mathbf{M}^\delta(x) = \mathbf{M} \left(x, \left\{ \frac{x}{\delta} \right\}_Y \right), \quad \mathbf{R}^\delta(x) = \mathbf{R} \left(x, \left\{ \frac{x}{\delta} \right\}_Y \right), \quad \text{for a.e. } x \in \mathbb{R}^3. \quad (4.14)$$

Compare this definition with the one in Definition 4.1.2, which means that the parameters \mathbf{M} and \mathbf{R} are the blueprints of \mathbf{M}^δ and \mathbf{R}^δ , respectively. See also (Cioranescu et al., 2008, Remark 2.3). As pointed out in the Remark (2) to (Cioranescu et al., 2002, Theorem 2) there is a special structure that is allowed for the parameters, which represents the scale separation. See also (Cioranescu et al., 2008, Remark 5.6). We only consider the case where the domain Ω can be covered by disjunct cuboids. This is only a simplification but no essential restriction. In fact one can do everything that follows also for general domains Ω , see Cioranescu et al. (2008) for the details but this would complicate things unnecessarily. Now we define the unfolding operator and give its key properties.

Definition 4.4.1 (Unfolding operator). *The unfolding operator $\mathcal{T}^\delta : L^2(\Omega; \mathbb{R}^n) \rightarrow L^2(\Omega \times Y; \mathbb{R}^n)$ is defined by*

$$\mathcal{T}^\delta(\Phi)(x, y) = \Phi \left(\delta \left[\frac{x}{\delta} \right] + \delta y \right), \quad \text{for a.e. } (x, y) \in \Omega \times Y, \text{ and all } \Phi \in L^2(\Omega; \mathbb{R}^n).$$

Let us give the central results needed for the homogenization process below. We start with the following result (Bossavit et al., 2005, Theorem 2).

Theorem 4.4.2. (i) *For all $\Phi \in L^2(\Omega)$ we have the strong convergence*

$$\mathcal{T}^\delta(\Phi) \rightarrow \Phi \quad \text{in } L^2(\Omega \times Y).$$

(ii) *Let Φ^δ be a family of functions uniformly bounded in $L^2(\Omega)$. There exists $\Phi \in L^2(\Omega \times Y)$ such that, up to a subsequence, we have the weak convergence*

$$\mathcal{T}^\delta(\Phi^\delta) \rightharpoonup \Phi \quad \text{in } L^2(\Omega \times Y).$$

(iii) *Let Φ^δ be a family of functions uniformly bounded in $H(\text{curl}, \Omega)$. There exists three fields*

$$\Phi \in H(\text{curl}, \Omega), \quad \bar{\Phi} \in L^2(\Omega, H_{\#}^1(Y; \mathbb{R})), \quad \bar{\bar{\Phi}} \in L^2(\Omega, H_{\#}^1(Y; \mathbb{R}^3)), \quad \text{div}_y \bar{\bar{\Phi}} = 0,$$

such that up to a subsequence, we have the weak convergences

$$\begin{aligned} \Phi^\delta &\rightharpoonup \Phi && \text{in } H(\text{curl}, \Omega), \\ \mathcal{T}^\delta(\Phi^\delta) &\rightharpoonup \Phi + \nabla_y \bar{\Phi} && \text{in } L^2(\Omega \times Y; \mathbb{R}^3), \\ \mathcal{T}^\delta(\text{curl } \Phi^\delta) &\rightharpoonup \text{curl}_x \Phi + \text{curl}_y \bar{\bar{\Phi}} && \text{in } L^2(\Omega \times Y; \mathbb{R}^3). \end{aligned}$$

Proof. The first two statements are already given in Cioranescu et al. (2002) and again in Cioranescu et al. (2008). The proof of (iii) can be found in Bossavit et al. (2005). It relies on the decomposition

$$\mathbf{H}(\text{curl}, \Omega) = \nabla \mathbf{H}_0^1(\Omega) \oplus \mathbf{H}(\text{curl}, \text{div } 0, \Omega),$$

where $\mathbf{H}(\text{curl}, \text{div } 0, \Omega) := \{\phi \in \mathbf{H}(\text{curl}, \Omega) : \text{div } \phi \in L^2(\Omega) \text{ and } \text{div } \phi = 0\}$. \square

This result is comparable with the two-scale counterpart given in (Wellander, 2001, Proposition 4.3). We now recall the result (Cioranescu et al., 2008, Proposition 2.14) about the equivalence of two-scale convergence to weak $L^2(\Omega \times Y)$ -convergences of the unfolding sequence.

Proposition 4.4.3. *Let $\{\Phi^\delta\}$ be a bounded sequence in $L^2(\Omega)$. The following assertions are equivalent:*

- (i) $\{\mathcal{T}^\delta(\Phi^\delta)\}$ converges weakly to Φ in $L^2(\Omega \times Y)$,
- (ii) $\{\Phi^\delta\}$ two-scale converges to Φ .

Recall that the solution (3.19) of the abstract Maxwell system (3.21) is split in four components, i.e., the electric and magnetic field and possibly multiple polarization and magnetization fields. The notation we used previously is now extended to higher dimensions fitting to this splitting.

Remark 4.4.4. *The dimension of the abstract Maxwell system (3.21) is $n = 3(2 + N_E + N_H)$ and we set $N := \frac{n}{3}$. For $\mathbf{v} = \left(\mathbf{v}_1^T \quad \mathbf{v}_2^T \quad \mathbf{v}_3^T \quad \mathbf{v}_4^T \right)^T \in \mathbb{R}^{3+3N_E+3+3N_H}$ and $\mathbf{w} = \left(w_1 \quad w_2^T \quad w_3 \quad w_4^T \right)^T \in \mathbb{R}^{1+N_E+1+N_H}$ we define*

$$\begin{aligned} \text{curl } \mathbf{v} &:= (\text{curl } \mathbf{v}_\ell)_{\ell=1}^4, \\ \text{div } \mathbf{v} &:= (\text{div } \mathbf{v}_\ell)_{\ell=1}^4, \\ \mathbf{n} \times \mathbf{v} &:= (\mathbf{n} \times \mathbf{v}_\ell)_{\ell=1}^4, \\ \nabla \mathbf{w} &:= \left(\nabla w_1^T \quad \nabla w_2^T \quad \nabla w_3^T \quad \nabla w_4^T \right)^T = \left(\nabla w_1^T, \quad \nabla w_2^T, \quad \nabla w_3^T, \quad \nabla w_4^T \right) \in \mathbb{R}^n. \end{aligned}$$

Here the curl, divergence and cross product for \mathbf{v}_2 and \mathbf{v}_4 are defined component wise. The same holds true for the gradient of w_2 and w_4 .

Furthermore, we introduce the quantity $\bar{y} \in \mathbb{R}^{N \times n}$, where \bar{y}_ℓ has an entry at position $\frac{\ell - (\ell \bmod 3)}{3} + 1$ and zeros elsewhere. The entry is $y_{\ell \bmod 3 + \frac{3}{2}(1 - \ell \bmod 3)(2 - \ell \bmod 3)}$, which is one of the three coordinate directions of the unit cell Y . This gives us a complicated way to write the canonical basis of \mathbb{R}^n and the $n \times n$ identity matrix as

$$\mathbf{e}_\ell = \nabla_y \bar{y}_\ell, \quad \mathbf{I}_n = \left(\mathbf{e}_1 \quad \dots \quad \mathbf{e}_n \right) = \left(\nabla_y \bar{y}_1 \quad \dots \quad \nabla_y \bar{y}_n \right) = \mathbf{D}_y^T \bar{y}.$$

Next we present the limit system, see (Bokil et al., 2018, Theorem 5.1) where the actual proof is found in (Bossavit et al., 2005, Theorem 3). The parameter \mathbf{R}^δ in the following theorem does not need to be symmetric or positive semi-definite, which is indeed the most general case of a coupling. In the next theorem we use the notation \mathbf{A}_x and \mathbf{A}_y . This has to be understood as the Maxwell operator with respect to the x and y variable, respectively.

Theorem 4.4.5. *Let $\mathbf{M}^\delta, \mathbf{R}^\delta \in L^\infty(\Omega; \mathbb{R}^{n \times n})$ be two matrix sequences given by (4.14), with \mathbf{M}^δ symmetric and uniformly positive definite. Assume for the initial condition $\mathbf{u}_0^\delta \in \mathcal{D}(\mathbf{A})$ and the source*

$\mathbf{g}^\delta \in H^1(0, T; L^\infty(\Omega; \mathbb{R}^n))$. Let \mathbf{u}^δ be the solution to the Maxwell problem (4.11).

Then there exist three fields

$$\begin{aligned} \mathbf{u}^{\text{eff}} &\in W^{1,\infty}(0, T; L^2(\Omega; \mathbb{R}^n)) \cap L^\infty(0, T; \mathcal{D}(\mathbf{A})), \\ \bar{\mathbf{u}} &\in W^{1,\infty}(0, T; L^2(\Omega; H_{\#}^1(Y; \mathbb{R}^N))), \\ \bar{\bar{\mathbf{u}}} &\in L^\infty(0, T; L^2(\Omega; H_{\#}^1(Y; \mathbb{R}^n))), \quad \operatorname{div}_y \bar{\bar{\mathbf{u}}} = 0, \end{aligned}$$

which are limits of the following sequences

$$\begin{aligned} \mathbf{u}^\delta &\rightharpoonup \mathbf{u}^{\text{eff}}, && \text{weakly-}^* \text{ in } L^\infty(0, T; \mathcal{D}(\mathbf{A})), \\ \mathcal{T}^\delta(\mathbf{u}^\delta) &\rightarrow \mathbf{u}^{\text{eff}} + \nabla_y \bar{\mathbf{u}}, && \text{strongly in } H^1(0, T; L^2(\Omega \times Y; \mathbb{R}^n)), \\ \mathcal{T}^\delta(\operatorname{curl}_x \mathbf{u}_j^\delta) &\rightarrow \operatorname{curl}_x \mathbf{u}_j^{\text{eff}} + \operatorname{curl}_y \bar{\bar{\mathbf{u}}}_j, && j = 1, 3 \text{ strongly in } L^2((0, T) \times \Omega \times Y; \mathbb{R}^n). \end{aligned}$$

Note that $j = 1, 3$ corresponds to the electric and magnetic fields as in (3.19). These limit fields solve the Maxwell system

$$\begin{aligned} \mathbf{M}(x, y) \partial_t (\mathbf{u}^{\text{eff}}(t, x) + \nabla_y \bar{\mathbf{u}}(t, x, y)) + \mathbf{R}(x, y) (\mathbf{u}^{\text{eff}}(t, x) + \nabla_y \bar{\mathbf{u}}(t, x, y)) &&& \text{in } (0, T) \times \Omega \times Y, \\ + \mathbf{A}_x \mathbf{u}^{\text{eff}}(t, x) + \mathbf{A}_y \bar{\bar{\mathbf{u}}}(t, x, y) = \mathbf{g}(t, x) &&& \\ \mathbf{u}^{\text{eff}}(0) = \mathbf{u}_0, \quad \bar{\mathbf{u}}(0) = 0 &&& \text{in } \Omega \times Y, \\ \mathbf{n} \times \mathbf{u}_1^{\text{eff}} = \mathbf{0} &&& \text{on } (0, T) \times \Omega. \end{aligned}$$

Proof. The proof of the theorem relies on a slightly different result given in Bossavit et al. (2005) and the fact that the parameters ensure the scale separation. To be more precise we reformulate (4.11a) as

$$\mathbf{M}^\delta \partial_t \mathbf{u}^\delta(t) + \mathbf{R}^\delta \mathbf{u}^\delta(t) + \mathbf{A} \mathbf{u}^\delta(t) = \partial_t \left(\mathbf{M}^\delta \mathbf{u}^\delta(t) + \int_0^t \mathbf{R}^\delta \mathbf{u}^\delta(s) \, ds \right) + \mathbf{A} \mathbf{u}^\delta(t).$$

Thus, we can now apply (Bossavit et al., 2005, Theorem 3) which gives the result. \square

The limit system in Theorem 4.4.5 should be compared with the system in (4.8) where the similarities are obvious. Thus, the next step is to derive the effective parameters and the corresponding cell correctors as in Section 4.2.1, which we do in the next section.

4.4.3 The homogeneous Maxwell system and the delay effect

Computing the effective parameters and corresponding correctors is the final step in the derivation of the homogeneous system. We show that the limit solution \mathbf{u}^{eff} from Theorem 4.4.5 solves a global Maxwell problem posed in $(0, T) \times \Omega$, while the corrector $\bar{\mathbf{u}}$ solves local diffusion problems posed in $(0, T) \times Y$ for $x \in \Omega$. The following result is again from (Bokil et al., 2018, Theorem 5.2).

Theorem 4.4.6. For $\delta > 0$, let $\mathbf{M}^\delta \in L^\infty(\Omega; \mathbb{R}^{n \times n})$, symmetric and uniformly positive definite, and $\mathbf{R}^\delta \in L^\infty(\Omega; \mathbb{R}^{n \times n})$, be two families of parameters given as in (4.14). Assume that the initial condition \mathbf{u}_0^δ and the source \mathbf{g}^δ satisfy (4.12). Then there exists a unique effective field

$$\mathbf{u}^{\text{eff}} = \begin{pmatrix} \mathbf{E}^{\text{eff}} \\ \mathbb{P}^{\text{eff}} \\ \mathbf{H}^{\text{eff}} \\ \mathbb{M}^{\text{eff}} \end{pmatrix} \in W^{1,\infty}(0, T; L^2(\Omega; \mathbb{R}^n)) \cap L^\infty(0, T; \mathcal{D}(\mathbf{A})),$$

which solves the effective Maxwell system

$$\mathbf{M}^{\text{eff}}(x)\partial_t \mathbf{u}^{\text{eff}}(t, x) + \tilde{\mathbf{R}}^{\text{eff}}(x)\mathbf{u}^{\text{eff}}(t, x) + \partial_t \int_0^t \tilde{\mathbf{G}}^{\text{eff}}(t-s, x)\mathbf{u}^{\text{eff}}(s, x) ds \quad \text{in } (0, T) \times \Omega, \quad (4.15a)$$

$$+ \mathbf{A}\mathbf{u}^{\text{eff}}(t, x) = \mathbf{g}(t, x) - \mathbf{J}^{\text{eff}}(t, x)\mathbf{u}_0(x) \quad \text{in } \Omega, \quad (4.15b)$$

$$\mathbf{u}^{\text{eff}}(0) = \mathbf{u}_0 \quad \text{in } \Omega, \quad (4.15b)$$

$$\mathbf{n} \times \mathbf{u}_1^{\text{eff}} = \mathbf{0} \quad \text{on } (0, T) \times \partial\Omega. \quad (4.15c)$$

Here the ℓ -th column of the effective parameters and the extra source are given as

$$(\mathbf{M}^{\text{eff}}(x))_\ell = \int_Y \mathbf{M}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) dy, \quad (4.16)$$

$$(\tilde{\mathbf{R}}^{\text{eff}}(x))_\ell = \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) dy, \quad (4.17)$$

$$(\tilde{\mathbf{G}}^{\text{eff}}(t, x))_\ell = \int_Y \left[\mathbf{M}(x, y) \nabla_y \bar{w}_\ell(t, x, y) + \int_0^t \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(s, x, y) ds \right] dy, \quad (4.18)$$

$$(\mathbf{J}^{\text{eff}}(t, x))_\ell = \partial_t \int_Y \left[\mathbf{M}(x, y) \nabla_y w_\ell^0(t, x, y) + \int_0^t \mathbf{R}(x, y) \nabla_y w_\ell^0(s, x, y) ds \right] dy, \quad (4.19)$$

for $\ell = 1, \dots, n$.

The correctors $w_\ell^{\text{M}}(x, \cdot) \in \mathbf{H}_\#^1(Y; \mathbb{R}^N)$, $\bar{w}_\ell \in \mathbf{W}^{2,1}(0, T; \mathbf{H}_\#^1(Y; \mathbb{R}^N))$, $w_\ell^0 \in \mathbf{W}^{2,1}(0, T; \mathbf{H}_\#^1(Y; \mathbb{R}^N))$ solve the following cell problems.

- The corrector w^{M} is a classic cell corrector associated to the parameter \mathbf{M}

$$\int_Y \mathbf{M}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \nabla_y v(y) dy = 0, \quad \text{for all } v \in \mathbf{H}_\#^1(Y; \mathbb{R}^N). \quad (4.20)$$

- The corrector \bar{w} is a time dependent cell corrector the arises due to the damping parameter \mathbf{R} and solves for a.e. $t \in (0, T)$

$$\begin{aligned} & \int_Y \left[\mathbf{M}(x, y) \nabla_y \bar{w}_\ell(t, x, y) + \int_0^t \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(s, x, y) ds \right] \cdot \nabla_y v(y) dy \\ &= - \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \nabla_y v(y) dy, \quad \text{for all } v \in \mathbf{H}_\#^1(Y; \mathbb{R}^N). \end{aligned} \quad (4.21)$$

- The cell corrector w^0 is related to the initial condition and the damping parameter. It solves for a.e. $t \in (0, T)$

$$\begin{aligned} & \int_Y \left[\mathbf{M}(x, y) \nabla_y w_\ell^0(t, x, y) + \int_0^t \mathbf{R}(x, y) \nabla_y w_\ell^0(s, x, y) ds \right] \cdot \nabla_y v(y) dy \\ &= \int_Y \mathbf{M}(x, y) \mathbf{e}_\ell \cdot \nabla_y v(y) dy, \quad \text{for all } v \in \mathbf{H}_\#^1(Y; \mathbb{R}^N). \end{aligned} \quad (4.22)$$

In this theorem we used the notation given in Remark 4.4.4 for both, the correctors and the test functions.

Remark 4.4.7. Observe the following about the effective parameters from Theorem 4.4.6

1. The effective parameter \mathbf{M}^{eff} is symmetric since \mathbf{M} is symmetric. For details see Lemma 4.4.8 below.
2. In the previous theorem the parameter \mathbf{R} is again not assumed to be symmetric. In this general case, even the effective parameters $\tilde{\mathbf{R}}^{\text{eff}}$ and $\tilde{\mathbf{G}}^{\text{eff}}$ will not necessarily be symmetric.
3. In special cases where the parameter \mathbf{R} is symmetric it is still not obvious that the effective parameters are symmetric. Nevertheless, we show in Lemma 4.4.9 that they are in fact symmetric.

The wellposedness of the cell problems is not obvious, especially for the time-dependent problems. One way to show the unique existence of a solution is to use Fredholm theory as in *Bossavit et al. (2005)*, but we use a different approach, namely methods for differential equations. In Section 4.4.4, we differentiate the time dependent cell problems, which yields equivalent formulations for the cell correctors in a setting of differential equations.

The key in the derivation of the effective system is again the representation of the global corrector $\bar{\mathbf{u}}$, which is given as

$$\bar{\mathbf{u}}(t, x, y) = w_\ell^{\text{M}}(x, y) \mathbf{u}_\ell^{\text{eff}}(t, x) + w_\ell^0(t, x, y) \mathbf{u}_\ell^{\text{eff}}(0, x) + \int_0^t \bar{w}_\ell(t-s, x, y) \mathbf{u}_\ell^{\text{eff}}(s, x) \, ds.$$

This representation is almost the same as in (4.9) with the only difference that in Section 4.2.1 we considered the initial value to be zero. In what follows we slightly reformulate the homogeneous system from Theorem 4.4.6 in a way that we think is more convenient to read. Moreover, the reformulation of the system is useful for the analysis and the subsequent numerical approximation.

4.4.4 Reformulation of parameters and cell correctors

Reformulation of parameters

Let us first show that the effective parameter \mathbf{M}^{eff} is symmetric.

Lemma 4.4.8. Let \mathbf{M}^{eff} be given as in (4.16). Then \mathbf{M}^{eff} is symmetric.

Proof. Note that the component $\mathbf{M}_{k,\ell}^{\text{eff}}$ of the effective parameter is given by

$$(\mathbf{M}^{\text{eff}}(x))_{k,\ell} = \int_Y \mathbf{M}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \mathbf{e}_k \, dy.$$

Now we test the cell problem (4.20) for w_ℓ^{M} with $w_k^{\text{M}} \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$ yielding

$$0 = \int_Y \mathbf{M}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \nabla_y w_k^{\text{M}}(x, y) \, dy.$$

Adding this zero to the effective parameter implies

$$\begin{aligned} (\mathbf{M}^{\text{eff}}(x))_{k,\ell} &= \int_{\mathring{Y}} \mathbf{M}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy \\ &= \int_{\mathring{Y}} (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y))^T \mathbf{M}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \, dy. \end{aligned} \quad (4.23)$$

The symmetry of the parameter \mathbf{M} now immediately leads to the symmetry of the effective one. \square

Next we give a reformulation of the system (4.15a), which relies on the evaluation of the time derivative of the convolution.

Lemma 4.4.9. *Let \mathbf{u}^{eff} be the solution of (4.15a) from Theorem 4.4.6. Then*

$$\begin{aligned} \tilde{\mathbf{R}}^{\text{eff}}(x) \mathbf{u}^{\text{eff}}(t, x) + \partial_t \int_0^t \tilde{\mathbf{G}}^{\text{eff}}(t-s, x) \mathbf{u}^{\text{eff}}(s, x) \, ds \\ = \mathbf{R}^{\text{eff}}(x) \mathbf{u}^{\text{eff}}(t, x) + \int_0^t \mathbf{G}^{\text{eff}}(t-s, x) \mathbf{u}^{\text{eff}}(s, x) \, ds, \end{aligned} \quad (4.24)$$

where

$$(\mathbf{R}^{\text{eff}}(x))_{k,\ell} = \int_{\mathring{Y}} \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy, \quad (4.25)$$

$$(\mathbf{G}^{\text{eff}}(t, x))_{k,\ell} = \int_{\mathring{Y}} \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(t, x, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy. \quad (4.26)$$

Moreover, the extra source in (4.19) can be written as

$$(\mathbf{J}^{\text{eff}}(t, x))_{k,\ell} = \int_{\mathring{Y}} \mathbf{R}(x, y) \nabla_y w_\ell^0(t, x, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy. \quad (4.27)$$

Proof. As in the proof of Lemma 4.4.8 we write the parameters component wise

$$\begin{aligned} (\tilde{\mathbf{R}}^{\text{eff}}(x))_{k,\ell} &= \int_{\mathring{Y}} \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \mathbf{e}_k \, dy, \\ (\tilde{\mathbf{G}}^{\text{eff}}(t, x))_{k,\ell} &= \int_{\mathring{Y}} \left[\mathbf{M}(x, y) \nabla_y \bar{w}_\ell(t, x, y) + \int_0^t \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(s, x, y) \, ds \right] \cdot \mathbf{e}_k \, dy. \end{aligned}$$

Now we use the cell problem (4.21) related to \bar{w}_ℓ and test it again with w_k^{M} . This gives

$$\begin{aligned} 0 &= \int_{\mathring{Y}} \left[\mathbf{M}(x, y) \nabla_y \bar{w}_\ell(t, x, y) + \int_0^t \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(s, x, y) \, ds \right] \cdot \nabla_y w_k^{\text{M}}(x, y) \, dy \\ &\quad + \int_{\mathring{Y}} \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \nabla_y w_k^{\text{M}}(x, y) \, dy, \end{aligned}$$

which we add to the parameter $\tilde{\mathbf{G}}^{\text{eff}}$

$$\begin{aligned} (\tilde{\mathbf{G}}^{\text{eff}}(t, x))_{k, \ell} &= \int_Y \mathbf{M}(x, y) \nabla_y \bar{w}_\ell(t, x, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy \\ &\quad + \int_Y \int_0^t \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(s, x, y) \, ds \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy \\ &\quad + \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \nabla_y w_k^{\text{M}}(x, y) \, dy. \end{aligned} \quad (4.28)$$

Observe that the first expression vanishes due to (4.20) since $\bar{w}_\ell(t, x, \cdot) \in \mathbf{H}_{\#}^1(Y)$ for almost every $(t, x) \in (0, T) \times \Omega$.

We show the statement of the lemma component wise, i.e., we rewrite the left-hand side of (4.24) as

$$\begin{aligned} &\left[\tilde{\mathbf{R}}^{\text{eff}}(x) \mathbf{u}^{\text{eff}}(t, x) + \partial_t \int_0^t \tilde{\mathbf{G}}^{\text{eff}}(t-s, x) \mathbf{u}^{\text{eff}}(s, x) \, ds \right]_k \\ &= \sum_{\ell=1}^d \left[\tilde{\mathbf{R}}_{k, \ell}^{\text{eff}}(x) \mathbf{u}_\ell^{\text{eff}}(t, x) + \partial_t \int_0^t \tilde{\mathbf{G}}_{k, \ell}^{\text{eff}}(t-s, x) \mathbf{u}_\ell^{\text{eff}}(s, x) \, ds \right] \\ &= \sum_{\ell=1}^d \left[\tilde{\mathbf{R}}_{k, \ell}^{\text{eff}}(x) \mathbf{u}_\ell^{\text{eff}}(t, x) + \int_0^t \partial_t \tilde{\mathbf{G}}_{k, \ell}^{\text{eff}}(t-s, x) \mathbf{u}_\ell^{\text{eff}}(s, x) \, ds + \tilde{\mathbf{G}}_{k, \ell}^{\text{eff}}(0, x) \mathbf{u}_\ell^{\text{eff}}(t, x) \right], \end{aligned} \quad (4.29)$$

where we used Lemma 2.0.1 in the last equation. Now we use (4.20) tested with $\bar{w}_\ell(0, x)$ and (4.21) tested with $w_k^{\text{M}}(x)$, which yields

$$\begin{aligned} &(\tilde{\mathbf{R}}_{k, \ell}^{\text{eff}}(x) + \tilde{\mathbf{G}}_{k, \ell}^{\text{eff}}(0, x)) \mathbf{u}_\ell^{\text{eff}}(t, x) \\ &= \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \mathbf{e}_k \, dy \mathbf{u}_\ell^{\text{eff}}(t, x) + \int_Y \mathbf{M}(x, y) \nabla_y \bar{w}_\ell(0, x, y) \cdot \mathbf{e}_k \, dy \mathbf{u}_\ell^{\text{eff}}(t, x) \\ &= \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy \mathbf{u}_\ell^{\text{eff}}(t, x) \\ &= \mathbf{R}_{k, \ell}^{\text{eff}}(x) \mathbf{u}_\ell^{\text{eff}}(t, x), \end{aligned} \quad (4.30)$$

where \mathbf{R}^{eff} is defined as in (4.25). It remains to evaluate the time derivative of the effective parameter. We find from (4.28) recalling that the first expression in (4.28) equals zero

$$\begin{aligned} \partial_t (\tilde{\mathbf{G}}^{\text{eff}}(t, x))_{k, \ell} &= \partial_t \int_Y \int_0^t \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(s, x, y) \, ds \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy \\ &= \int_Y \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(t, x, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy. \end{aligned}$$

Plugging this into the convolution yields

$$\begin{aligned}
& \int_0^t \partial_t \tilde{\mathbf{G}}_{k,\ell}^{\text{eff}}(t-s, x) \mathbf{u}_\ell^{\text{eff}}(s, x) \, ds \\
&= \int_0^t \int_Y \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(t-s, x, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy \mathbf{u}_\ell^{\text{eff}}(s, x) \, ds \\
&= \int_0^t \mathbf{G}_{k,\ell}^{\text{eff}}(t-s, x) \mathbf{u}_\ell^{\text{eff}}(s, x) \, ds,
\end{aligned} \tag{4.31}$$

where \mathbf{G}^{eff} is defined as in (4.26). Inserting (4.30) and (4.31) in (4.29) gives the result in (4.24).

The extra source term reformulation follows from the same ideas. First take the component-wise definition and then test the cell problem (4.22) for w_ℓ^0 with w_k^{M} and add the resulting zero to the definition of the extra source. We get

$$\begin{aligned}
(\mathbf{J}^{\text{eff}}(t, x))_{k,\ell} &= \partial_t \left[\int_Y \mathbf{M}(x, y) \nabla_y w_\ell^0(t, x, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy \right. \\
&\quad + \int_Y \int_0^t \mathbf{R}(x, y) \nabla_y w_\ell^0(s, x, y) \, ds \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy \\
&\quad \left. + \int_Y \mathbf{M}(x, y) \mathbf{e}_\ell \cdot \nabla_y w_k^{\text{M}}(x, y) \, dy \right],
\end{aligned}$$

and see that the first expression again vanishes due to the cell problem (4.20) and the last expression is constant in time and therefore the time derivative is zero. The remaining expression evaluates to the requested form in (4.27). \square

Corollary 4.4.10. *Let \mathbf{R}^{eff} be given as in (4.25). If \mathbf{R} is symmetric, also \mathbf{R}^{eff} is symmetric.*

Proof. The symmetry is an immediate consequence of the reformulation of the parameter \mathbf{R}^{eff} in (4.25). \square

If the parameter \mathbf{R} is symmetric also the convolution kernel is symmetric, but this is not as obvious as for the effective parameter \mathbf{R}^{eff} and we show this later in Lemma 4.4.18. Thanks to Lemma 4.4.9 we can rewrite the system (4.15a) from Theorem 4.4.6 as

$$\begin{aligned}
& \mathbf{M}^{\text{eff}}(x) \partial_t \mathbf{u}^{\text{eff}}(t, x) + \mathbf{R}^{\text{eff}}(x) \mathbf{u}^{\text{eff}}(t, x) + \int_0^t \mathbf{G}^{\text{eff}}(t-s, x) \mathbf{u}^{\text{eff}}(s, x) \, ds + \mathbf{A} \mathbf{u}^{\text{eff}}(t, x) \\
&= \mathbf{g}(t, x) - \mathbf{J}^{\text{eff}}(t, x) \mathbf{u}_0(x)
\end{aligned} \tag{4.32}$$

where the parameters are given in (4.16), (4.25), (4.26) and (4.27). The cell correctors are still the same as is the theorem.

Cell problems

We now give the reformulation of the cell problems as differential equations.

Lemma 4.4.11. *Let $t \in (0, T)$, $x \in \Omega$ and $\ell = 1, \dots, n$. The cell corrector $\bar{w}_\ell(t, x, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$ solves (4.21) if and only if it is the solution of:*

Find $\bar{w}_\ell(t, x, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$ such that

$$\int_Y [\mathbf{M}(x, y) \partial_t \nabla_y \bar{w}_\ell(t, x, y) + \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(t, x, y)] \cdot \nabla_y v(y) \, dy = 0 \quad \text{for } t > 0, \quad (4.33)$$

and

$$\int_Y \mathbf{M}(x, y) \nabla_y \bar{w}_\ell(0, x, y) \cdot \nabla_y v(y) \, dy = - \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^M(x, y)) \cdot \nabla_y v(y) \, dy, \quad (4.34)$$

for all $v \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$.

Proof. Let $t \in (0, T)$, $x \in \Omega$ and $\ell = 1, \dots, n$. If $\bar{w}_\ell(t, x, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$ is solution of (4.21), differentiating the equation with respect to time implies (4.33). The evaluation at $t = 0$ of (4.21) yields (4.34).

If, the other way around, $\bar{w}_\ell(t, x, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$ is solution of (4.33) - (4.34) we integrate (4.33) over $(0, t)$ and use (4.34) which yields (4.21). \square

We get a very similar result for the cell corrector w^0 , which is only different in the initial value.

Lemma 4.4.12. *Let $t \in (0, T)$, $x \in \Omega$ and $\ell = 1, \dots, n$. The cell corrector $w_\ell^0(t, x, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$ is solution of (4.22) if and only if it is solution of:*

Find $w_\ell^0(t, x, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$ such that

$$\int_Y [\mathbf{M}(x, y) \partial_t \nabla_y w_\ell^0(t, x, y) + \mathbf{R}(x, y) \nabla_y w_\ell^0(t, x, y)] \cdot \nabla_y v(y) \, dy = 0 \quad \text{for } t > 0, \quad (4.35)$$

and

$$\int_Y \mathbf{M}(x, y) \nabla_y w_\ell^0(0, x, y) \cdot \nabla_y v(y) \, dy = \int_Y \mathbf{M}(x, y) \mathbf{e}_\ell \cdot \nabla_y v(y) \, dy, \quad (4.36)$$

for all $v \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$.

Proof. The proof is very similar to Lemma 4.4.11. \square

Remark 4.4.13. *Observe that the negative initial value of w^0 coincides with the classic corrector w^M , which is obvious due to (4.36) and (4.20), i.e.,*

$$\int_Y \mathbf{M}(x, y) \nabla_y w_\ell^0(0, x, y) \cdot \nabla_y v(y) \, dy = - \int_Y \mathbf{M}(x, y) \nabla_y w_\ell^M(x, y) \cdot \nabla_y v(y) \, dy \quad \text{for all } v \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N).$$

Reformulated Maxwell system

Now the reformulated Maxwell system, which we discretize in the remainder, reads as follows

$$\mathbf{M}^{\text{eff}}(x) \partial_t \mathbf{u}^{\text{eff}}(t, x) + \mathbf{R}^{\text{eff}}(x) \mathbf{u}^{\text{eff}}(t, x) + \int_0^t \mathbf{G}^{\text{eff}}(t-s, x) \mathbf{u}^{\text{eff}}(s, x) \, ds \quad \text{in } (0, T) \times \Omega, \quad (4.37a)$$

$$+ \mathbf{A} \mathbf{u}^{\text{eff}}(t, x) = \mathbf{g}(t, x) - \mathbf{J}^{\text{eff}}(t, x) \mathbf{u}_0(x)$$

$$\mathbf{u}^{\text{eff}}(0) = \mathbf{u}_0 \quad \text{in } \Omega, \quad (4.37b)$$

$$\mathbf{n} \times \mathbf{u}_1^{\text{eff}}(t, x) = \mathbf{0} \quad \text{on } (0, T) \times \partial\Omega, \quad (4.37c)$$

where the k, ℓ -th component of the effective parameters and the extra source are given as

$$(\mathbf{M}^{\text{eff}}(x))_{k,\ell} = \int_Y \mathbf{M}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy, \quad (4.38a)$$

$$(\mathbf{R}^{\text{eff}}(x))_{k,\ell} = \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy, \quad (4.38b)$$

$$(\mathbf{G}^{\text{eff}}(t, x))_{k,\ell} = \int_Y \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(t, x, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy, \quad (4.38c)$$

$$(\mathbf{J}^{\text{eff}}(t, x))_{k,\ell} = \int_Y \mathbf{R}(x, y) \nabla_y w_\ell^0(t, x, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(x, y)) \, dy, \quad (4.38d)$$

for $k, \ell = 1, \dots, n$.

The correctors $w_\ell^{\text{M}}(x, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$, $\bar{w}_\ell \in \mathbf{W}^{2,1}(0, T; \mathbf{H}_{\#}^1(Y; \mathbb{R}^N))$, $w_\ell^0 \in \mathbf{W}^{2,1}(0, T; \mathbf{H}_{\#}^1(Y; \mathbb{R}^N))$ solve the following cell problems.

- The corrector w^{M} is a classic cell corrector associated to the parameter \mathbf{M}

$$\int_Y \mathbf{M}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \nabla_y v(y) \, dy = 0, \quad \text{for all } v \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N). \quad (4.39)$$

- The corrector \bar{w} is a time dependent cell corrector that arises due to the damping parameter \mathbf{R} and solves for a.e. $t \in (0, T)$ and all $v \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$

$$\int_Y [\mathbf{M}(x, y) \partial_t \nabla_y \bar{w}_\ell(t, x, y) + \mathbf{R}(x, y) \nabla_y \bar{w}_\ell(t, x, y)] \cdot \nabla_y v(y) \, dy = 0, \quad (4.40)$$

$$\int_Y \mathbf{M}(x, y) \nabla_y \bar{w}_\ell(0, x, y) \cdot \nabla_y v(y) \, dy + \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y)) \cdot \nabla_y v(y) \, dy = 0. \quad (4.41)$$

- The cell corrector w^0 is related to the initial condition and the damping parameter. It solves for a.e. $t \in (0, T)$ and all $v \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$

$$\int_Y [\mathbf{M}(x, y) \partial_t \nabla_y w_\ell^0(t, x, y) + \mathbf{R}(x, y) \nabla_y w_\ell^0(t, x, y)] \cdot \nabla_y v(y) \, dy = 0, \quad (4.42)$$

$$\int_Y \mathbf{M}(x, y) (\mathbf{e}_\ell - \nabla_y w_\ell^0(0, x, y)) \cdot \nabla_y v(y) \, dy = 0. \quad (4.43)$$

Let us mention that the system (4.37) is equivalent to the system (4.10). This can be seen by considering the parameters as in Section 3.1.2 and the vanishing initial value.

A consequence of the formulation as a differential equation is that we can apply the classic wellposedness theory to the cell problems.

4.4.5 Wellposedness of cell problems

Wellposedness of stationary problems

The theory in this section is well-known theory for elliptic problems. Note that the wellposedness of the standard, time-independent cell problem (4.20) follows from the Lax-Milgram lemma. In addition, this result shows wellposedness of the initial value problems for the cell correctors \bar{w} and w^0 . In this and the following section we drop the dependence on the x variable, since it is only a parameter in these cell problems. Thus, let $x \in \Omega$ be fixed. Furthermore, we abbreviate the solution space by $\mathbf{V}^{\text{mic}} := \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$, which is equipped with the inner product

$$(\phi, \psi)_{\mathbf{V}^{\text{mic}}} = (\nabla_y \phi, \nabla_y \psi)_{L^2(Y; \mathbb{R}^N)} \quad \text{for all } \phi, \psi \in \mathbf{V}^{\text{mic}}.$$

Let us mention that this is a scalar product due to the Poincaré–Wirtinger inequality. All the three mentioned problems can be written as:

Find $w \in \mathbf{V}^{\text{mic}}$ such that

$$s_m(w, v) = b(v), \quad \text{for all } v \in \mathbf{V}^{\text{mic}}. \quad (4.44)$$

Here the bilinear form $m : \mathbf{V}^{\text{mic}} \times \mathbf{V}^{\text{mic}} \rightarrow \mathbb{R}$ is given as

$$s_m(w, v) = \int_Y \mathbf{M}(y) \nabla_y w(y) \cdot \nabla_y v(y) \, dy, \quad \text{for all } w, v \in \mathbf{V}^{\text{mic}}. \quad (4.45)$$

It is bounded and coercive, since the parameter \mathbf{M} is assumed to be bounded and positive definite, cf. (3.26). Thus, the bilinear form is a scalar product on \mathbf{V}^{mic} , and we denote its induced norm by $\|\cdot\|_{s_m}$. Hence, we have an equivalent scalar product, which satisfies for all $\phi \in \mathbf{V}^{\text{mic}}$

$$\sqrt{\alpha} \|\phi\|_{\mathbf{V}^{\text{mic}}} \leq \|\phi\|_{s_m} \leq \sqrt{C_{\mathbf{M}}} \|\phi\|_{\mathbf{V}^{\text{mic}}}. \quad (4.46)$$

The parameter \mathbf{R} is assumed to be bounded as well, and we denote the constant by

$$C_{\mathbf{R}} = \|\mathbf{R}\|_{L^\infty(\Omega, \mathbb{R}^{n \times n})}. \quad (4.47)$$

The functional on the right-hand side in (4.44) is different in all the three cases mentioned above. For $v \in \mathbf{V}^{\text{mic}}$ we define

$$\begin{aligned} b_\ell^{\mathbf{M}}(v) &:= - \int_Y \mathbf{M}(y) \mathbf{e}_\ell \cdot \nabla_y v(y) \, dy, \\ \bar{b}_\ell(v) &:= - \int_Y \mathbf{R}(y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\mathbf{M}}(y)) \cdot \nabla_y v(y) \, dy, \\ b_\ell^0(v) &:= -b_\ell^{\mathbf{M}}(v), \end{aligned}$$

and all these linear forms are bounded and thus functionals on V^{mic} . Therefore, we can apply Lax-Milgram to the three problems (4.20), (4.34) and (4.36), which are in the new notation equivalent to:

Find $w_\ell^{\text{M}}, \bar{w}_\ell(0), w_\ell^0(0) \in V^{\text{mic}}$ such that

$$s_m(w_\ell^{\text{M}}, v) = b_\ell^{\text{M}}(v), \quad (4.48)$$

$$s_m(\bar{w}_\ell(0), v) = \bar{b}_\ell(v), \quad (4.49)$$

$$s_m(w_\ell^0(0), v) = b_\ell^0(v), \quad (4.50)$$

for all $v \in V^{\text{mic}}$. We collect the results from above in the next lemma and extend them by some bounds on the solutions.

Remark 4.4.14. *In this chapter we always consider the unit cell $Y = (-\frac{1}{2}, \frac{1}{2})^3$ which satisfies $|Y| = 1$. Nevertheless, later in Chapter 5 we also consider other cells, which have non-trivial measures. Thus, in the following results we keep track of the measure $|Y|$ although in this section it is equal to one.*

Lemma 4.4.15. *For every $\ell = 1, \dots, n$ the cell problems (4.20), (4.34) and (4.36) are wellposed. Moreover, the solutions satisfy the bounds*

$$\|w_\ell^{\text{M}}\|_{s_m} \leq \sqrt{C_{\mathbf{M}} |Y|}, \quad \|\bar{w}_\ell(0)\|_{s_m} \leq 2 \frac{C_{\mathbf{R}}}{\alpha} \sqrt{C_{\mathbf{M}} |Y|}, \quad \|w_\ell^0(0)\|_{s_m} \leq \sqrt{C_{\mathbf{M}} |Y|}. \quad (4.51)$$

Proof. We showed the wellposedness in this section. Let us briefly comment on a notation we use. As introduced in Remark 4.4.4 we find

$$\|\bar{y}_\ell\|_{s_m}^2 = \int_Y \mathbf{M}(y) \nabla_y \bar{y}_\ell \cdot \nabla_y \bar{y}_\ell \, dy = \int_Y \mathbf{M}(y) \mathbf{e}_\ell \cdot \mathbf{e}_\ell \, dy.$$

For the bound in the s_m -norm on w_ℓ^{M} we choose in (4.48) $v = w_\ell^{\text{M}}$. From the bound on the parameter \mathbf{M} in (4.46) and the Cauchy–Schwarz inequality we find

$$\begin{aligned} \|w_\ell^{\text{M}}\|_{s_m}^2 &= s_m(w_\ell^{\text{M}}, w_\ell^{\text{M}}) = b_\ell^{\text{M}}(w_\ell^{\text{M}}) = - \int_Y \mathbf{M}(y) \mathbf{e}_\ell \cdot \nabla_y w_\ell^{\text{M}}(y) \, dy \leq \|\bar{y}_\ell\|_{s_m} \|w_\ell^{\text{M}}\|_{s_m} \\ &\leq \sqrt{C_{\mathbf{M}} |Y|} \|w_\ell^{\text{M}}\|_{s_m}. \end{aligned}$$

Similarly we choose $v = \bar{w}_\ell(0)$ in (4.49) which yields

$$\begin{aligned} \|\bar{w}_\ell(0)\|_{s_m}^2 &= s_m(\bar{w}_\ell(0), \bar{w}_\ell(0)) = \bar{b}_\ell(\bar{w}_\ell(0)) = - \int_Y \mathbf{R}(y) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(y)) \cdot \nabla_y \bar{w}_\ell(0, y) \, dy \\ &\leq C_{\mathbf{R}} (\|\bar{y}_\ell\|_{V^{\text{mic}}} + \|w_\ell^{\text{M}}\|_{V^{\text{mic}}}) \|\bar{w}_\ell(0)\|_{V^{\text{mic}}} \leq \frac{C_{\mathbf{R}}}{\alpha} (\|\bar{y}_\ell\|_{s_m} + \|w_\ell^{\text{M}}\|_{s_m}) \|\bar{w}_\ell(0)\|_{s_m} \\ &\leq 2 \frac{C_{\mathbf{R}}}{\alpha} \sqrt{C_{\mathbf{M}} |Y|} \|\bar{w}_\ell(0)\|_{s_m}. \end{aligned}$$

For the last estimate we choose $v = w_\ell^0(0)$ in (4.50) and get with the same arguments as for w_ℓ^{M} the bound

$$\|w_\ell^0(0)\|_{s_m}^2 = s_m(w_\ell^0(0), w_\ell^0(0)) = b_\ell^0(w_\ell^0(0)) = \int_Y \mathbf{M}(y) \mathbf{e}_\ell \cdot \nabla_y w_\ell^0(0) \, dy \leq \sqrt{C_{\mathbf{M}} |Y|} \|w_\ell^0(0)\|_{s_m}.$$

□

Before we show wellposedness of the time-dependent problems let us comment on H^2 -regularity results for the stationary problem. As an example consider (4.48) and assume that its solution is H^2 -regular. If the parameter is regular enough we get with integration by parts

$$s_m(w_\ell^M, v) = \int_Y \operatorname{div}_y (\mathbf{M}(y) \mathbf{e}_\ell) \cdot v(y) \, dy = (f, v)_{L^2(Y; \mathbb{R}^N)}.$$

We can now exploit the usual H^2 -regularity estimate from Grisvard (2011) that states

$$\|w_\ell^M\|_{H^2(Y)} \leq C \|f\|_{L^2(Y)} = C \|\operatorname{div}_y (\mathbf{M} \mathbf{e}_\ell)\|_{L^2(Y)} \leq C \|\mathbf{M}\|_{W^{1,\infty}(Y)} \sqrt{|Y|}.$$

The classical assumption in homogenization on the parameters is that $\|\mathbf{M}\|_{W^{1,\infty}(Y)} \leq C\delta^{-1}$ and this yields

$$\|w_\ell^M\|_{H^2(Y)} \leq C\delta^{-1} \sqrt{|Y|}.$$

We will see these H^2 -regularity assumptions later in Chapter 5. Next, we show that the time-dependent problems (4.33) and (4.35) with initial values (4.34) and (4.36) are also wellposed.

Wellposedness of the evolution problems

The two time dependent cell problems (4.33) and (4.35) are so-called Sobolev equations. Observe that these solutions only differ in the initial value. Therefore, we consider a general Sobolev equation with initial value $w^0 \in H_{\neq}^1(Y; \mathbb{R}^N)$

$$\int_Y \mathbf{M}(y) \nabla_y \partial_t w(t, y) \cdot \nabla_y v(y) \, dy + \int_Y \mathbf{R}(y) \nabla_y w(t, y) \nabla_y v(y) \, dy = 0 \quad \text{for all } v \in V^{\text{mic}}, t \in [0, T], \quad (4.52a)$$

$$w(0) = w^0 \quad \text{in } Y. \quad (4.52b)$$

The results in this section are similar to Bekkouche et al. (2019) and Hipp et al. (2019). As in the previous section we rewrite this problem in a shorter way. For that purpose we introduce another bilinear form similar to (4.45) but related to the parameter \mathbf{R} . Again we suppress the x -dependence.

$$s_r(w, v) := \int_Y \mathbf{R}(y) \nabla_y w(y) \cdot \nabla_y v(y) \, dy, \quad \text{for all } w, v \in V^{\text{mic}}. \quad (4.53)$$

In contrast to the bilinear form $s_m(\cdot, \cdot)$ the new form $s_r(\cdot, \cdot)$ is only bounded but not coercive, since the parameter is not necessarily positive definite. With this notation we rewrite the Sobolev equation (4.52a) as

$$s_m(\partial_t w(t), v) + s_r(w(t), v) = 0 \quad \text{for all } v \in V^{\text{mic}}, t \in [0, T]. \quad (4.54)$$

Due to the Riesz-representation theorem we find an operator $\mathcal{S} : V^{\text{mic}} \rightarrow V^{\text{mic}}$ such that

$$s_r(\phi, \psi) = s_m(\mathcal{S}\phi, \psi) \quad \text{for all } \phi, \psi \in V^{\text{mic}}. \quad (4.55)$$

We thus get the Sobolev equation as

$$s_m(\partial_t w(t), v) + s_m(\mathcal{S}w(t), v) = 0 \quad \text{for all } v \in V^{\text{mic}}. \quad (4.56)$$

We collect some properties of the operator \mathcal{S} . Observe that with respect to the inner product $s_m(\cdot, \cdot)$ the operator \mathcal{S} inherits its properties from the bilinear form $s_r(\cdot, \cdot)$ by definition. Therefore, we immediately get that \mathcal{S} is bounded, i.e., with (4.47) and (4.46)

$$|s_m(\mathcal{S}\phi, \psi)| = |s_r(\phi, \psi)| \leq C_{\mathbf{R}} \|\phi\|_{\mathbf{V}^{\text{mic}}} \|\psi\|_{\mathbf{V}^{\text{mic}}} \leq \frac{C_{\mathbf{R}}}{\alpha} \|\phi\|_{s_m} \|\psi\|_{s_m} .$$

Moreover, we find a constant $C_{\mathcal{S}} \geq 0$ such that

$$s_m(\mathcal{S}\phi, \phi) + C_{\mathcal{S}} \|\phi\|_{s_m}^2 \geq 0 .$$

We call the operator \mathcal{S} quasi-monotone and from (Hipp et al., 2019, Lemma 2.3) we see that $-(\mathcal{S} + C_{\mathcal{S}})$ is dissipative. In this setting the operator \mathcal{S} also satisfies the range condition with respect to \mathbf{V}^{mic} . Choosing $f \in \mathbf{V}^{\text{mic}}$ and $\lambda > C_{\mathcal{S}}$ we have to show the wellposedness of:

Find $u \in \mathbf{V}^{\text{mic}}$ such that

$$s_m((\lambda \text{id} + \mathcal{S})u, v) = s_m(f, v) \quad \text{for all } v \in \mathbf{V}^{\text{mic}} .$$

We denote by $a(\cdot, \cdot)$ the bilinear form given as

$$a(\phi, \psi) = \lambda s_m(\phi, \psi) + s_r(\phi, \psi) = s_m((\lambda \text{id} + \mathcal{S})\phi, \psi) ,$$

which is bounded ($|a(\phi, \psi)| \leq (\lambda C_{\mathbf{M}} + C_{\mathbf{R}}) \|\phi\|_{\mathbf{V}^{\text{mic}}} \|\psi\|_{\mathbf{V}^{\text{mic}}}$) and coercive on \mathbf{V}^{mic} . The Lax-Milgram Lemma applies to:

Find $u \in \mathbf{V}^{\text{mic}}$ such that

$$a(u, v) = s_m(f, v) \quad \text{for all } v \in \mathbf{V}^{\text{mic}} .$$

Therefore, we have a unique solution $u \in \mathbf{V}^{\text{mic}}$ for every $f \in \mathbf{V}^{\text{mic}}$ and the operator \mathcal{S} satisfies the range condition with respect to \mathbf{V}^{mic} .

Thus, we find that $-(\mathcal{S} + C_{\mathcal{S}})$ is dissipative and satisfies the range condition. By the Lumer-Phillips theorem it generates a contraction semigroup $(e^{-(\mathcal{S} + C_{\mathcal{S}})t})_{t \geq 0}$, i.e., in the $\|\cdot\|_{s_m}$ -Norm we get

$$\|e^{-\mathcal{S}t}\|_{s_m \leftarrow s_m} \leq e^{C_{\mathcal{S}}t} . \quad (4.57)$$

Moreover, we find that the Sobolev equation (4.54) has a unique solution

$$w(t) = e^{-\mathcal{S}t}w(0) . \quad (4.58)$$

We are now in the position to show the wellposedness of the time-dependent cell problems.

Lemma 4.4.16. *The cell problems (4.21) and (4.35) are wellposed for every $\ell = 1, \dots, n$. Moreover, the solutions have the regularity $\bar{w}_\ell, w_\ell^0 \in C^\infty(0, T; \mathbf{H}_{\#}^1(Y; \mathbb{R}^N))$ and satisfy the stability bound*

$$\|\bar{w}_\ell(t)\|_{s_m} \leq \|e^{-\mathcal{S}t}\|_{s_m \leftarrow s_m} \|\bar{w}_\ell(0)\|_{s_m} , \quad (4.59)$$

$$\|w_\ell^0(t)\|_{s_m} \leq \|e^{-\mathcal{S}t}\|_{s_m \leftarrow s_m} \|w_\ell^0(0)\|_{s_m} . \quad (4.60)$$

Proof. We showed in Lemma 4.4.11 that (4.21) is equivalent to (4.33) together with (4.34). Moreover, we showed that (4.33) is equivalent to the abstract Cauchy problem (4.56) with $w = \bar{w}$, where the initial value satisfies (4.34). Theorem 3.2.10 yields, that the operator $-(\mathcal{S} + C_{\mathcal{S}})$ generates a contraction

semigroup. Hence, due to Theorem 3.2.7 the abstract Cauchy problem is wellposed and the solution is given by (4.58). The initial value is given as solution of (4.34) or (4.49), which are wellposed thanks to Lemma 4.4.15.

For the regularity of the solution we use the representation of the solution given in (4.58) and the fact that the operator is bounded. Thus, we use the representation of the exponential for bounded operators, which we may differentiate infinitely often. The stability bounds again directly follow from the representation of the solution in (4.58).

Exactly the same argumentation is valid for the cell problem related to w_ℓ^0 . \square

In the case of a positive semi-definite parameter \mathbf{R} the operator \mathcal{S} is itself the generator of a contraction semigroup. This is equivalent to the fact that $C_{\mathcal{S}} = 0$, and we get the improved bound

$$\|e^{-\mathcal{S}t}\|_{s_m \leftarrow s_m} \leq 1. \quad (4.61)$$

If we use the contraction property (4.61) we get from Lemma 4.4.16

$$\|\bar{w}_\ell(t)\|_{s_m} \leq \|\bar{w}_\ell(0)\|_{s_m}, \quad \|w_\ell^0(t)\|_{s_m} \leq \|w_\ell^0(0)\|_{s_m},$$

but for other cases we require the stability result to be in the most general form. As for the stationary case we comment on H^2 -regularity. The following result states, that the H^2 -norm is bounded by the corresponding norm of the initial value.

Theorem 4.4.17. *Let $\mathbf{M} \in W^{1,\infty}(Y)$ be symmetric and uniformly positive definite, $\mathbf{R} \in W^{1,\infty}(Y)$ be positive semi-definite and assume that the initial value is H^2 -regular, i.e., $w^0 \in H^2(Y)$. For a solution $w(t, \cdot) \in H^2(Y)$ of (4.52a) we get the following bound*

$$\|w(t)\|_{H^2(Y)} \leq C \|w(0)\|_{H^2(Y)}.$$

Proof. Let $w^0 \in H^2(Y)$ be given. We rewrite the Sobolev equation (4.52a) in strong formulation as: Find $w : [0, T] \rightarrow H^2(Y)$ such that

$$\Delta_{\mathbf{M}} \partial_t w(t) + \Delta_{\mathbf{R}} w(t) = 0 \quad \text{in } Y.$$

The operators $\Delta_{\mathbf{M}} = \operatorname{div}(\mathbf{M}(y)\nabla \cdot)$ and $\Delta_{\mathbf{R}} = \operatorname{div}(\mathbf{R}(y)\nabla \cdot)$ are weighted Laplace operators. Note, that the operator $\Delta_{\mathbf{M}} : H^2(Y) \rightarrow L^2(Y)$ is invertible due to the properties of \mathbf{M} . The operator \mathcal{S} defined in (4.55) as operator on $H^2(Y)$ may be written as $\mathcal{S} = \Delta_{\mathbf{M}}^{-1} \Delta_{\mathbf{R}}$. Next, we consider the closely related operator $\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} : L^2(Y) \rightarrow L^2(Y)$. We show that this operator is also monotone and bounded with respect to $L^2(Y)$ equipped with the weighted inner product

$$(\Phi, \Psi)_{\Delta_{\mathbf{M}}^{-1}} = (\Phi, -\Delta_{\mathbf{M}}^{-1} \Psi)_{L^2(Y)} \quad \text{for all } \Phi, \Psi \in L^2(Y).$$

Note, that the weighted Laplacian has a negative spectrum, and we thus use the negative operator to get a scalar product. For this inner product we find with integration by parts and due to the positive semi-definiteness of \mathbf{R}

$$\begin{aligned} (\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} \Phi, \Phi)_{\Delta_{\mathbf{M}}^{-1}} &= \int_Y \Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} \Phi (-\Delta_{\mathbf{M}}^{-1} \Phi) \, dy = \int_Y \operatorname{div}(\mathbf{R} \nabla (\Delta_{\mathbf{M}}^{-1} \Phi)) (-\Delta_{\mathbf{M}}^{-1} \Phi) \, dy \\ &= \int_Y \mathbf{R} \nabla (\Delta_{\mathbf{M}}^{-1} \Phi) \cdot \nabla (\Delta_{\mathbf{M}}^{-1} \Phi) \, dy \geq 0. \end{aligned}$$

Additionally, using again integration by parts, the boundedness of \mathbf{R} from (4.47) as well as the positive definiteness of \mathbf{M} from (3.26) and the Cauchy–Schwarz inequality we get

$$\begin{aligned} \left| (\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} \Phi, \Psi)_{\Delta_{\mathbf{M}}^{-1}} \right| &= \left| \int_Y \mathbf{R} \nabla (\Delta_{\mathbf{M}}^{-1} \Phi) \cdot \nabla (\Delta_{\mathbf{M}}^{-1} \Psi) \, dy \right| \leq \frac{C_{\mathbf{R}}}{\alpha} \left| \int_Y \mathbf{M} \nabla (\Delta_{\mathbf{M}}^{-1} \Phi) \cdot \nabla (\Delta_{\mathbf{M}}^{-1} \Psi) \, dy \right| \\ &= \frac{C_{\mathbf{R}}}{\alpha} \left| - \int_Y \operatorname{div} (\mathbf{M} \nabla (\Delta_{\mathbf{M}}^{-1} \Phi)) (\Delta_{\mathbf{M}}^{-1} \Psi) \, dy \right| = \frac{C_{\mathbf{R}}}{\alpha} \left| \int_Y \Phi (-\Delta_{\mathbf{M}}^{-1} \Psi) \, dy \right| \\ &= \frac{C_{\mathbf{R}}}{\alpha} \left| (\Phi, \Psi)_{\Delta_{\mathbf{M}}^{-1}} \right| \leq \frac{C_{\mathbf{R}}}{\alpha} \|\Phi\|_{\Delta_{\mathbf{M}}^{-1}} \|\Psi\|_{\Delta_{\mathbf{M}}^{-1}}. \end{aligned}$$

Thus, the operator $-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1}$ is dissipative and generates a contraction semi-group $e^{-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} t}$, i.e.,

$$\left\| e^{-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} t} \right\|_{\Delta_{\mathbf{M}}^{-1} \leftarrow \Delta_{\mathbf{M}}^{-1}} \leq 1.$$

Note that both, $-\Delta_{\mathbf{M}}^{-1} \Delta_{\mathbf{R}}$ and $-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1}$ are bounded operators. Thus, the idea is to use the series representation for their generated semi-groups, which eventually yields an estimate for the $H^2(Y)$ -norm. For that purpose we define for $n \in \mathbb{N}$ the partial sum

$$s_n := \sum_{k=0}^n (-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1})^k \frac{t^k}{k!} \Delta_{\mathbf{M}} w(0).$$

Since $\Delta_{\mathbf{M}} w(0) \in L^2(Y)$ and due to the properties of $-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1}$ we just showed, we find the convergence in $L^2(Y)$ as

$$s_n \rightarrow s = e^{-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} t} \Delta_{\mathbf{M}} w(0) \quad \text{as } n \rightarrow \infty.$$

Moreover, for $z_n := \sum_{k=0}^n (-\Delta_{\mathbf{M}}^{-1} \Delta_{\mathbf{R}})^k \frac{t^k}{k!} w(0)$ we have

$$z_n \rightarrow z = e^{-\Delta_{\mathbf{M}}^{-1} \Delta_{\mathbf{R}} t} w(0) \quad \text{as } n \rightarrow \infty.$$

Observe, that we get the equality

$$\Delta_{\mathbf{M}} z_n = s_n,$$

which enables us to use the fact that $\Delta_{\mathbf{M}}$ is closed. To be precise, we conclude from

$$\begin{aligned} z_n &\rightarrow z \quad \text{as } n \rightarrow \infty, \\ \Delta_{\mathbf{M}} z_n &= s_n \rightarrow s \quad \text{as } n \rightarrow \infty, \end{aligned}$$

that $z \in \mathcal{D}(\Delta_{\mathbf{M}})$ and

$$\Delta_{\mathbf{M}} \left(e^{-\Delta_{\mathbf{M}}^{-1} \Delta_{\mathbf{R}} t} w(0) \right) = \Delta_{\mathbf{M}} z = s = e^{-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} t} \Delta_{\mathbf{M}} w(0).$$

Eventually, we find

$$\begin{aligned} \|w(t)\|_{H^2(Y)} &\leq C \|\Delta_{\mathbf{M}} w(t)\|_{L^2(Y)} = C \left\| \Delta_{\mathbf{M}} e^{-\Delta_{\mathbf{M}}^{-1} \Delta_{\mathbf{R}} t} w(0) \right\|_{L^2(Y)} = C \left\| e^{-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} t} \Delta_{\mathbf{M}} w(0) \right\|_{L^2(Y)} \\ &\leq C \left\| e^{-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} t} \Delta_{\mathbf{M}} w(0) \right\|_{\Delta_{\mathbf{M}}^{-1}} \leq C \left\| e^{-\Delta_{\mathbf{R}} \Delta_{\mathbf{M}}^{-1} t} \right\|_{\Delta_{\mathbf{M}}^{-1} \leftarrow \Delta_{\mathbf{M}}^{-1}} \|\Delta_{\mathbf{M}} w(0)\|_{\Delta_{\mathbf{M}}^{-1}} \leq C \|w(0)\|_{H^2(Y)}. \end{aligned}$$

□

We found that the cell problems are wellposed and that for a positive semi-definite parameter \mathbf{R} the H^2 norm does not grow in time. Moreover, we have already seen that if the parameter \mathbf{R} is symmetric this transfers to the parameter \mathbf{R}^{eff} . In the next section we use the representation (4.58) of the solution we just found to show that also the convolution kernel is symmetric if the parameter \mathbf{R} is.

4.4.6 Symmetric representation of the convolution kernel and the extra source

The convolution kernel

Remember that the matrix \mathbf{M} is symmetric. Let us now assume that the parameter \mathbf{R} is symmetric as well. For example this is the case in a Debye medium or in the conductivity case. The symmetry of \mathbf{R} immediately yields that the bilinear forms (4.45) and (4.53) are symmetric, i.e., for all $\phi, \psi \in V^{\text{mic}}$

$$s_m(\phi, \psi) = s_m(\psi, \phi), \quad s_r(\phi, \psi) = s_r(\psi, \phi).$$

Moreover, the operator \mathcal{S} defined in (4.55) is self-adjoint, which is clear since it is bounded and due to

$$s_m(\mathcal{S}\phi, \psi) = s_r(\phi, \psi) = s_r(\psi, \phi) = s_m(\mathcal{S}\psi, \phi) = s_m(\phi, \mathcal{S}\psi).$$

Consider the operator exponential that occurs in the solution (4.58). Since the operator \mathcal{S} is bounded we can write the exponential in its power series as

$$\lim_{N \rightarrow \infty} \left(\text{id} - \frac{t}{N} \mathcal{S} \right)^{-N} = e^{-St} = \sum_{j=0}^{\infty} \frac{(\mathcal{S})^j (-t)^j}{j!} : V^{\text{mic}} \rightarrow V^{\text{mic}}.$$

We show that this is also a self-adjoint operator. For $\phi, \psi \in V^{\text{mic}}$, it holds:

$$s_m(e^{-St}\phi, \psi) = s_m \left(\sum_{j=0}^{\infty} \frac{\mathcal{S}^j (-t)^j}{j!} \phi, \psi \right) = \sum_{j=0}^{\infty} \frac{(-t)^j}{j!} s_m(\mathcal{S}^j \phi, \psi).$$

For $j \in \mathbb{N}_0$ we get due to the self-adjointness of the operator

$$s_m(\mathcal{S}^j \phi, \psi) = s_m(\phi, \mathcal{S}^j \psi).$$

Thus, we see that every summand is self-adjoint, which yields the requested self-adjointness

$$s_m(e^{-St}\phi, \psi) = s_m(\phi, e^{-St}\psi).$$

We proceed with the symmetry of the effective convolution kernel.

Lemma 4.4.18. *If the parameter \mathbf{R} is symmetric also the effective convolution kernel \mathbf{G}^{eff} is symmetric.*

Proof. Consider the time-dependent parameter given by (4.26) and use the cell problem for the initial value (4.34) tested with $\bar{w}_\ell(t)$ to get

$$\begin{aligned} (\mathbf{G}^{\text{eff}}(t))_{k,\ell} &= \int_Y \mathbf{R}(y) \nabla_y \bar{w}_\ell(t, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(y)) \, dy \\ &= - \int_Y \mathbf{M}(y) \nabla_y \bar{w}_\ell(t, y) \cdot \nabla_y \bar{w}_k(0, y) \, dy \\ &= -s_m(\bar{w}_\ell(t), \bar{w}_k(0)). \end{aligned}$$

Now we use the representation of the solution given in (4.58) and the self-adjointness of the operator $e^{-\mathcal{S}t}$ and find

$$\begin{aligned} (\mathbf{G}^{\text{eff}}(t))_{k,\ell} &= -s_m(\bar{w}_\ell(t), \bar{w}_k(0)) = -s_m(e^{-\mathcal{S}t}\bar{w}_\ell(0), \bar{w}_k(0)) \\ &= -s_m(\bar{w}_\ell(0), e^{-\mathcal{S}t}\bar{w}_k(0)) = -s_m(\bar{w}_\ell(0), \bar{w}_k(t)). \end{aligned}$$

This shows

$$(\mathbf{G}^{\text{eff}}(t))_{k,\ell} = -\frac{1}{2}(s_m(\bar{w}_\ell(t), \bar{w}_k(0)) + s_m(\bar{w}_\ell(0), \bar{w}_k(t))).$$

Here we immediately see the symmetry of the convolution kernel. \square

The extra source

We note that the procedure from the previous section works very similar for the extra source term given in (4.27) with the cell corrector w^0 being the solution of (4.36) and (4.35). Therefore, also the extra source is symmetric if the parameter \mathbf{R} is.

4.4.7 On the exponential structure of the convolution kernel

We now examine the case where the damping parameter \mathbf{R} has a special structure. This is the case if we consider for example the conductivity Maxwell system from Section 3.1.2. In this setting we find that the evolution cell problem reduces to the component of the electric field. Thus, the general cell problem from (4.33) reduces to

$$\int_Y [\varepsilon(x, y)\partial_t \nabla_y \bar{w}_\ell(t, x, y) + \sigma(x, y)\nabla_y \bar{w}_\ell(t, x, y)] \cdot \nabla_y v(y) \, dy = 0,$$

where $\ell = 1, 2, 3$ and the other components of the corrector \bar{w} vanish. Now if we assume that the conductivity σ satisfies

$$\sigma > 0,$$

and if we redo the computations from the wellposedness analysis in Section 4.4.5, we get an operator \mathcal{S} that is strictly monotone. Thus, $-\mathcal{S}$ is strictly dissipative, i.e., we find $\beta > 0$ such that $-\mathcal{S} + \beta \text{id}$ is dissipative. But this is equivalent to the fact that the semigroup generated by \mathcal{S} is exponentially stable, i.e.,

$$\|e^{-\mathcal{S}t}\|_{s_m \leftarrow s_m} \leq \exp(-\beta t).$$

We use this bound in the stability estimate (4.59) and this shows that the convolution kernel is exponentially decaying. The key property to show the exponential decay is the positive definiteness of the damping parameter \mathbf{R} . The Maxwell–Debye system from 3.1.3 is an example where the parameter \mathbf{R} is only positive semi-definite and thus we do not get the exponential decay in this setting.

With this note on the possibly exponential decay of the convolution kernel we now turn to the wellposedness of the homogeneous Maxwell system (4.37).

4.4.8 Wellposedness of the integro-differential homogeneous system

In this section we show that the Maxwell system (4.37) is wellposed and stable. In contrast to the wellposedness theory from Section 3.2.2 we do not use semigroup theory here. The reason is that the effective Maxwell system has a different structure than the Maxwell systems we considered in Section 3.2.2. The difference is the convolution integral, which occurred in the homogenization process. The proof now relies on the Faedo–Galerkin method and Fredholm theory.

From now on we assume that the damping parameter \mathbf{R} is positive semi-definite. The main reason for this assumption is the structure of the micro problems. In Section 4.4.5, we showed that the solutions are stable if the parameter generates a contraction semigroup. There is also a physical interpretation. In the examples from Section 3.1.2 and 3.1.3 the parameter satisfies this assumption.

We start this section with a reformulation of the effective Maxwell system (4.37) in a variational form. With (\cdot, \cdot) we denote the standard $L^2(\Omega; \mathbb{R}^n)$ scalar product, and we search for $\mathbf{u}^{\text{eff}}(t) \in \mathcal{D}(\mathcal{A})$ such that

$$(\mathbf{M}^{\text{eff}} \partial_t \mathbf{u}^{\text{eff}}(t), \Phi) + (\mathbf{R}^{\text{eff}} \mathbf{u}^{\text{eff}}(t), \Phi) + \int_0^t (\mathbf{G}^{\text{eff}}(t-s) \mathbf{u}^{\text{eff}}(s), \Phi) \, ds + (\mathbf{A} \mathbf{u}^{\text{eff}}(t), \Phi) \quad (4.62)$$

$$= (\mathbf{g}(t), \Phi) - (\mathbf{J}^{\text{eff}}(t) \mathbf{u}_0, \Phi) \quad \text{for all } \Phi \in \mathcal{D}(\mathcal{A}),$$

$$\mathbf{u}^{\text{eff}}(0) = \mathbf{u}_0 \quad \text{in } \Omega. \quad (4.63)$$

Now we introduce bilinear forms $m^{\text{eff}}, r^{\text{eff}}, a : \mathcal{D}(\mathcal{A}) \times \mathcal{D}(\mathcal{A}) \rightarrow \mathbb{R}$ such that for every $\Psi, \Phi \in \mathcal{D}(\mathcal{A})$

$$m^{\text{eff}}(\Psi, \Phi) := (\mathbf{M}^{\text{eff}} \Psi, \Phi), \quad (4.64)$$

$$r^{\text{eff}}(\Psi, \Phi) := (\mathbf{R}^{\text{eff}} \Psi, \Phi), \quad (4.65)$$

$$a(\Psi, \Phi) := (\mathbf{A} \Psi, \Phi). \quad (4.66)$$

Moreover, for $t \in [0, T]$ define $g^{\text{eff}} : [0, T] \times \mathcal{D}(\mathcal{A}) \times \mathcal{D}(\mathcal{A}) \rightarrow \mathbb{R}$ such that

$$g^{\text{eff}}(t; \Psi, \Phi) := (\mathbf{G}^{\text{eff}}(t) \Psi, \Phi) \quad \text{for all } \Psi, \Phi \in \mathcal{D}(\mathcal{A}). \quad (4.67)$$

With these definitions we write the system (4.62) as

$$\begin{aligned} m^{\text{eff}}(\partial_t \mathbf{u}^{\text{eff}}(t), \Phi) + r^{\text{eff}}(\mathbf{u}^{\text{eff}}(t), \Phi) + \int_0^t g^{\text{eff}}(t-s; \mathbf{u}^{\text{eff}}(s), \Phi) \, ds + a(\mathbf{u}^{\text{eff}}(t), \Phi) \\ = m^{\text{eff}}(\mathbf{f}(t), \Phi) - (\mathbf{J}^{\text{eff}}(t) \mathbf{u}_0, \Phi) \quad \text{for all } \Phi \in \mathcal{D}(\mathcal{A}), \end{aligned} \quad (4.68)$$

where

$$m^{\text{eff}}(\mathbf{f}(t), \Phi) = (\mathbf{g}(t), \Phi).$$

Before we continue with the wellposedness result for the Maxwell system, let us give some properties of the bilinear forms related to the effective parameters we introduced.

Lemma 4.4.19. *The effective parameters $\mathbf{M}^{\text{eff}}, \mathbf{R}^{\text{eff}}$ are bounded, i.e., $\mathbf{M}^{\text{eff}}, \mathbf{R}^{\text{eff}} \in L^\infty(\Omega; \mathbb{R}^{n \times n})$. In addition, the parameter \mathbf{M}^{eff} is coercive with the same constant α as \mathbf{M} and the parameter \mathbf{R}^{eff} is positive semi-definite. Moreover, the time-dependent parameters $\mathbf{G}^{\text{eff}}(t)$ and $\mathbf{J}^{\text{eff}}(t)$ are bounded for all $t \in [0, T]$, i.e., $\mathbf{G}^{\text{eff}}(t), \mathbf{J}^{\text{eff}}(t) \in L^\infty(\Omega; \mathbb{R}^{n \times n})$. As a direct consequence the bilinear forms $m^{\text{eff}}, r^{\text{eff}}$ and $g^{\text{eff}}(t)$ are bounded.*

Proof. Let us start with the standard parameter in homogenization \mathbf{M}^{eff} for which the boundedness is classical. We use the symmetric representation of the effective parameter \mathbf{M}^{eff} in (4.23) as in Jikov et al. (1994). For $\xi \in \mathbb{R}^n$ we find using the cell problem (4.20)

$$\begin{aligned} \mathbf{M}^{\text{eff}}(x)\xi \cdot \xi &= s_m(\bar{y} + w^{\text{M}}(x), \bar{y} + w^{\text{M}}(x)) \xi \cdot \xi \\ &= (s_m(\bar{y}, \bar{y}) + s_m(\bar{y}, w^{\text{M}}(x)) + s_m(w^{\text{M}}(x), \bar{y} + w^{\text{M}}(x))) \xi \cdot \xi \\ &= (s_m(\bar{y}, \bar{y}) - 2s_m(w^{\text{M}}(x), w^{\text{M}}(x)) + s_m(w^{\text{M}}(x), w^{\text{M}}(x))) \xi \cdot \xi \\ &= (s_m(\bar{y}, \bar{y}) - s_m(w^{\text{M}}(x), w^{\text{M}}(x))) \xi \cdot \xi \\ &\leq C_{\mathbf{M}} |Y| |\xi|^2. \end{aligned}$$

The coercivity also relies on the cell problems as shown in Bensoussan et al. (2011):

$$\begin{aligned} \mathbf{M}^{\text{eff}}(x)\xi \cdot \xi &= s_m(\bar{y} + w^{\text{M}}(x), \bar{y} + w^{\text{M}}(x)) \xi \cdot \xi \geq \alpha(\bar{y} + w^{\text{M}}(x), \bar{y} + w^{\text{M}}(x))_{\text{Vmic}} \xi \cdot \xi \\ &= \alpha |Y| |\xi|^2 + 2\alpha(\bar{y}, w^{\text{M}}(x))_{\text{Vmic}} \xi \cdot \xi. \end{aligned}$$

Integration by parts shows that the last expression vanishes due to the periodicity of the corrector w^{M} . The positive semi-definiteness of \mathbf{R} implies directly the one of \mathbf{R}^{eff} by

$$\mathbf{R}^{\text{eff}}(x)\xi \cdot \xi = s_r(\bar{y} + w^{\text{M}}(x), \bar{y} + w^{\text{M}}(x)) \xi \cdot \xi \geq 0.$$

For the boundedness of the effective damping parameter we use a different technique. With (4.38b) we find

$$\begin{aligned} |\mathbf{R}^{\text{eff}}(x)_{k,\ell}| &= |s_r(\bar{y}_\ell + w_\ell^{\text{M}}(x), y_k + w_k^{\text{M}}(x))| \\ &\leq \frac{C_{\mathbf{R}}}{\alpha} \left(\|\bar{y}_\ell\|_{s_m} \|y_k\|_{s_m} + \|\bar{y}_\ell\|_{s_m} \|w_k^{\text{M}}(x)\|_{s_m} + \|w_\ell^{\text{M}}(x)\|_{s_m} \|y_k\|_{s_m} + \|w_\ell^{\text{M}}(x)\|_{s_m} \|w_k^{\text{M}}(x)\|_{s_m} \right) \\ &\leq 4 \frac{C_{\mathbf{R}} C_{\mathbf{M}}}{\alpha} |Y|, \end{aligned}$$

where we again used Lemma 4.4.15 and the coercivity of \mathbf{M} .

Let us now consider the time-dependent convolution kernel. We can actually bound this expression independent of x and t . We use (4.38c) and Lemmas 4.4.15 and 4.4.16 which yield together with (4.61)

$$\begin{aligned} |\mathbf{G}^{\text{eff}}(t, x)_{k,\ell}| &= |s_r(\bar{w}_\ell(t, x), y_k + w_k^{\text{M}}(x))| \leq \frac{C_{\mathbf{R}}}{\alpha} \|\bar{w}_\ell(t, x)\|_{s_m} \left(\|y_k\|_{s_m} + \|w_k^{\text{M}}(x)\|_{s_m} \right) \\ &\leq 2 \frac{C_{\mathbf{R}} \sqrt{C_{\mathbf{M}} |Y|}}{\alpha} \|\bar{w}_\ell(0, x)\|_{s_m} \leq 4 \left(\frac{C_{\mathbf{R}}}{\alpha} \right)^2 C_{\mathbf{M}} |Y|. \end{aligned}$$

Finally for the extra source term we get from (4.38d) with the same techniques as above

$$\begin{aligned} |\mathbf{J}^{\text{eff}}(t, x)_{k,\ell}| &= |s_r(w_\ell^0(t, x), y_k + w_k^{\text{M}}(x))| \leq \frac{C_{\mathbf{R}}}{\alpha} \|w_\ell^0(t, x)\|_{s_m} \left(\|y_k\|_{s_m} + \|w_k^{\text{M}}(x)\|_{s_m} \right) \\ &\leq 2 \frac{C_{\mathbf{R}} \sqrt{C_{\mathbf{M}} |Y|}}{\alpha} \|w_\ell^0(0, x)\|_{s_m} \leq 2 \frac{C_{\mathbf{R}} C_{\mathbf{M}}}{\alpha} |Y|. \end{aligned}$$

Thanks to the bounds on the parameters we immediately get the continuity of the bilinear forms since

for all $\phi, \psi \in L^2(\Omega; \mathbb{R}^n)$ and $t \geq 0$ we find

$$m^{\text{eff}}(\phi, \phi) \geq \alpha |Y| \|\phi\|_{L^2(\Omega; \mathbb{R}^n)}^2, \quad (4.69)$$

$$|m^{\text{eff}}(\phi, \psi)| \leq C_{\mathbf{M}} |Y| \|\phi\|_{L^2(\Omega; \mathbb{R}^n)} \|\psi\|_{L^2(\Omega; \mathbb{R}^n)}, \quad (4.70)$$

$$r^{\text{eff}}(\phi, \phi) \geq 0, \quad (4.71)$$

$$|r^{\text{eff}}(\phi, \psi)| \leq 4 \frac{C_{\mathbf{R}} C_{\mathbf{M}}}{\alpha} |Y| \|\phi\|_{L^2(\Omega; \mathbb{R}^n)} \|\psi\|_{L^2(\Omega; \mathbb{R}^n)}, \quad (4.72)$$

$$|g^{\text{eff}}(t; \phi, \psi)| \leq 4 \left(\frac{C_{\mathbf{R}}}{\alpha} \right)^2 C_{\mathbf{M}} |Y| \|\phi\|_{L^2(\Omega; \mathbb{R}^n)} \|\psi\|_{L^2(\Omega; \mathbb{R}^n)}. \quad (4.73)$$

Note that the bound on the time-dependent parameters is uniform. Therefore, we have that $\mathbf{G}^{\text{eff}}, \mathbf{J}^{\text{eff}} \in L^\infty(0, T; L^\infty(\Omega; \mathbb{R}^{n \times n}))$. \square

As pointed out in Lemma 4.4.16, the cell correctors are C^∞ in time. Thus, a direct consequence is that the convolution kernel satisfies $\mathbf{G} \in C^\infty(0, T; L^\infty(\Omega; \mathbb{R}^{n \times n}))$. We see that the effective parameters have the same properties as their heterogeneous counterparts with possibly different constants. Here the most important point for the wellposedness result below is that the parameter \mathbf{M}^{eff} is again bounded and coercive. A main ingredient of the proof is the following Lemma, from (Bossavit et al., 2005, Lemma 1.1) which is based on Fredholm theory for Volterra integral equations.

Lemma 4.4.20. *Let $n \in \mathbb{N}$ and let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be a symmetric and uniformly coercive matrix. Moreover, let $\mathbf{G} \in W^{1,1}(0, T; \mathbb{R}^{n \times n})$ and assume $\mathbf{b} \in W^{1,1}(0, T; \mathbb{R}^n)$. Then, there exists a unique $\mathbf{u} \in W^{1,1}(0, T; \mathbb{R}^n)$ such that*

$$\mathbf{M}\mathbf{u}(t) + \int_0^t \mathbf{G}(t-s)\mathbf{u}(s) \, ds = \mathbf{b}(t) \quad \text{for all } t \in (0, T).$$

We now state the wellposedness result for the effective Maxwell system (4.68).

Theorem 4.4.21. *Let $\mathbf{M}^{\text{eff}}, \mathbf{R}^{\text{eff}} \in L^\infty(\Omega; \mathbb{R}^{n \times n})$ and $\mathbf{G}^{\text{eff}}, \mathbf{J}^{\text{eff}} \in W^{1,1}(0, T; L^\infty(\Omega; \mathbb{R}^{n \times n}))$. Additionally, let the initial value satisfy $\mathbf{u}_0 \in \mathcal{D}(\mathcal{A})$ and the exterior source $\mathbf{g} \in W^{1,1}(0, T; L^2(\Omega; \mathbb{R}^n))$. Then, the problem (4.68) has a unique solution, which satisfies*

$$\begin{aligned} \|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)} &\leq \exp\left(\int_0^t \|\mathbf{G}^{\text{eff}}\|_{L^1(0,s; L^\infty(\Omega; \mathbb{R}^{n \times n}))} \, ds\right) \left[\frac{1}{\alpha} t \|\mathbf{g}\|_{L^\infty(0,t; L^2(\Omega; \mathbb{R}^n))} \right. \\ &\quad \left. + \left(1 + \frac{1}{\alpha} \|\mathbf{J}^{\text{eff}}\|_{L^1(0,t; L^\infty(\Omega; \mathbb{R}^{n \times n}))}\right) \|\mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} \right]. \end{aligned} \quad (4.74)$$

Proof. The proof of the wellposedness relies on the Faedo-Galerkin method. The details can be found in (Bossavit et al., 2005, Proposition 1). Let us comment on the procedure. First one derives a sequence of approximate solutions $\mathbf{u}_m^{\text{eff}}$ of \mathbf{u}^{eff} by the use of Lemma 4.4.20. The next step is to show that this sequence is bounded, which yields the existence of a converging subsequence. Finally, we pass to the limit and show that the limit solves the equation (4.68). Here we only show the bounds on the solution since this is crucial for the error analysis.

Consider the system (4.68) with test function $\Phi = \mathbf{u}^{\text{eff}}(t)$. We already know that

$$a(\mathbf{u}^{\text{eff}}(t), \mathbf{u}^{\text{eff}}(t)) = 0, \quad r^{\text{eff}}(\mathbf{u}^{\text{eff}}(t), \mathbf{u}^{\text{eff}}(t)) \geq 0.$$

From the product rule we get

$$m^{\text{eff}}(\partial_t \mathbf{u}^{\text{eff}}(t), \mathbf{u}^{\text{eff}}(t)) = \frac{1}{2} \frac{d}{dt} m^{\text{eff}}(\mathbf{u}^{\text{eff}}(t), \mathbf{u}^{\text{eff}}(t)) \geq \alpha \left(\frac{d}{dt} \|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)} \right) \|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)}.$$

This yields with (4.68)

$$\begin{aligned} \alpha \left(\frac{d}{dt} \|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)} \right) \|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)} &\leq |(\mathbf{g}(t), \mathbf{u}^{\text{eff}}(t))| + |(\mathbf{J}^{\text{eff}}(t) \mathbf{u}_0, \mathbf{u}^{\text{eff}}(t))| \\ &\quad + \left| \int_0^t g^{\text{eff}}(t-s; \mathbf{u}^{\text{eff}}(s), \mathbf{u}^{\text{eff}}(t)) \, ds \right|. \end{aligned}$$

The application of the Cauchy-Schwarz inequality on the right-hand side in all expressions and division by $\|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)}$ leads to

$$\begin{aligned} \alpha \frac{d}{dt} \|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)} &\leq \|\mathbf{g}(t)\|_{L^2(\Omega; \mathbb{R}^n)} + \|\mathbf{J}^{\text{eff}}(t) \mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} + \int_0^t \|\mathbf{G}^{\text{eff}}(t-s) \mathbf{u}^{\text{eff}}(s)\|_{L^2(\Omega; \mathbb{R}^n)} \, ds \\ &\leq \|\mathbf{g}(t)\|_{L^2(\Omega; \mathbb{R}^n)} + \|\mathbf{J}^{\text{eff}}(t)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \|\mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} \\ &\quad + \int_0^t \|\mathbf{G}^{\text{eff}}(t-s)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \|\mathbf{u}^{\text{eff}}(s)\|_{L^2(\Omega; \mathbb{R}^n)} \, ds \\ &\leq \|\mathbf{g}(t)\|_{L^2(\Omega; \mathbb{R}^n)} + \|\mathbf{J}^{\text{eff}}(t)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \|\mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} \\ &\quad + \int_0^t \|\mathbf{G}^{\text{eff}}(t-s)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \, ds \|\mathbf{u}^{\text{eff}}\|_{L^\infty(0, t; L^2(\Omega; \mathbb{R}^n))}. \end{aligned}$$

Next we integrate over $[0, t]$, which yields

$$\begin{aligned} \alpha \|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)} &\leq \alpha \|\mathbf{u}^{\text{eff}}(0)\|_{L^2(\Omega; \mathbb{R}^n)} + \int_0^t \|\mathbf{g}(s)\|_{L^2(\Omega; \mathbb{R}^n)} \, ds + \int_0^t \|\mathbf{J}^{\text{eff}}(s)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \, ds \|\mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} \\ &\quad + \int_0^t \int_0^s \|\mathbf{G}^{\text{eff}}(s-r)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \, dr \|\mathbf{u}^{\text{eff}}\|_{L^\infty(0, s; L^2(\Omega; \mathbb{R}^n))} \, ds. \end{aligned}$$

The final step is to take the supremum over $[0, t]$

$$\begin{aligned} \|\mathbf{u}^{\text{eff}}\|_{L^\infty(0, t; L^2(\Omega; \mathbb{R}^n))} &\leq \left(1 + \frac{1}{\alpha} \int_0^t \|\mathbf{J}^{\text{eff}}(s)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \, ds \right) \|\mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} + \frac{1}{\alpha} \int_0^t \|\mathbf{g}(s)\|_{L^2(\Omega; \mathbb{R}^n)} \, ds \\ &\quad + \frac{1}{\alpha} \int_0^t \int_0^s \|\mathbf{G}^{\text{eff}}(s-r)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \, dr \|\mathbf{u}^{\text{eff}}\|_{L^\infty(0, s; L^2(\Omega; \mathbb{R}^n))} \, ds \\ &= C_0(t) + \int_0^t C_1(s) \|\mathbf{u}^{\text{eff}}\|_{L^\infty(0, s; L^2(\Omega; \mathbb{R}^n))} \, ds, \end{aligned}$$

where

$$\begin{aligned} C_0(t) &:= \left(1 + \frac{1}{\alpha} \int_0^t \|\mathbf{J}^{\text{eff}}(s)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \, ds \right) \|\mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} + \frac{1}{\alpha} \int_0^t \|\mathbf{g}(s)\|_{L^2(\Omega; \mathbb{R}^n)} \, ds \\ &\leq \left(1 + \frac{1}{\alpha} \|\mathbf{J}^{\text{eff}}\|_{L^1(0,t; L^\infty(\Omega; \mathbb{R}^{n \times n}))} \right) \|\mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} + \frac{1}{\alpha} t \|\mathbf{g}\|_{L^\infty(0,t; L^2(\Omega; \mathbb{R}^n))} \end{aligned} \quad (4.75)$$

$$C_1(t) := \frac{1}{\alpha} \int_0^t \|\mathbf{G}^{\text{eff}}(t-s)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \, ds = \frac{1}{\alpha} \|\mathbf{G}^{\text{eff}}\|_{L^1(0,t; L^\infty(\Omega; \mathbb{R}^{n \times n}))} . \quad (4.76)$$

We now apply Gronwall's inequality, which results in

$$\|\mathbf{u}^{\text{eff}}\|_{L^\infty(0,t; L^2(\Omega; \mathbb{R}^n))} \leq \exp\left(\int_0^t C_1(s) \, ds\right) C_0(t) .$$

Since the supremum is bounded we have the same bound for all $s \in [0, t]$. This is the a priori bound in (4.74). □

Thus, we find a unique solution of the effective Maxwell system. But in the most general case where the material parameters are given as in (3.21) without any further assumptions on \mathbf{R} we can not expect to show a stable solution for all times. Instead, we get an exponential growth in the stability estimate, which is clear since the heterogeneous material already has this growth. Nevertheless, if the heterogeneous material yields a stable solution we would expect that the homogenized solution is stable as well. For this purpose consider the example from Section 4.4.7 where we showed that the convolution kernel is exponentially decaying. We then refine the stability result from the previous theorem.

Corollary 4.4.22. *Under the assumptions of Theorem 4.4.21 and the additional assumption that the damping parameter \mathbf{R} is strictly positive we get the following stability estimate*

$$\|\mathbf{u}^{\text{eff}}(t)\|_{L^2(\Omega; \mathbb{R}^n)} \leq e^{\frac{C}{\alpha\beta}(t + \frac{1}{\beta}e^{-\beta t} - \frac{1}{\beta})} \left[\left(1 + \frac{C}{\alpha} \frac{1}{\beta} (1 - e^{-\beta t}) \right) \|\mathbf{u}_0\|_{L^2(\Omega; \mathbb{R}^n)} + \frac{1}{\alpha} t \|\mathbf{g}\|_{L^\infty(0,t; L^2(\Omega; \mathbb{R}^n))} \right] .$$

Proof. We use the fact that we have an exponentially stable semigroup, i.e.,

$$\|e^{-St}\|_{s_m \leftarrow s_m} \leq \exp(-\beta t) ,$$

and redo the computations from Lemma 4.4.19 to find that

$$\|\mathbf{G}^{\text{eff}}(t)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \leq C \exp(-\beta t) , \quad \|\mathbf{J}^{\text{eff}}(t)\|_{L^\infty(\Omega; \mathbb{R}^{n \times n})} \leq C \exp(-\beta t) ,$$

with a constant $C > 0$. The proof now relies on the evaluation of the expressions (4.75) and (4.76), which occur in Theorem 4.4.21

$$\int_0^t \|\mathbf{G}^{\text{eff}}\|_{L^1(0,s; L^\infty(\Omega; \mathbb{R}^{n \times n}))} \, ds \leq \frac{C}{\beta} \left(t + \frac{1}{\beta} e^{-\beta t} - \frac{1}{\beta} \right) , \quad \|\mathbf{J}^{\text{eff}}\|_{L^1(0,t; L^\infty(\Omega; \mathbb{R}^{n \times n}))} \leq C \frac{1}{\beta} (1 - e^{-\beta t}) .$$

□

In comparison to the stability bound in Theorem 4.4.21 we lose one order of t in the leading exponential factor, which is good but not the result we expected. The heterogeneous medium which we homogenize satisfies the stability bound from (3.31), which is without any exponential growth. Hence, we expect that the homogeneous system satisfies a similar bound without the exponential growth in time. This can be achieved in special situations and with further knowledge on the convolution kernel and thus on the heterogeneous parameters. We will show an example in Section 7.2.

Still, we have a wellposed system, and we can proceed with its discretization in the next chapter. Before doing so, we comment on other approaches to the problem of rapidly varying coefficients. Especially, we address the question of parameters with less structure, i.e., without any periodicity assumption or scale separation. This is the scope of the next section.

4.5 Homogenization beyond periodicity

For general parameters that are not (locally) periodic the method of homogenization is still applicable, i.e., the principle existence of a homogenization limit is provided. The drawback, however, is that in general no representation of the effective parameters is achieved as shown in Cioranescu and Donato (1999); Jikov et al. (1994). Methods that overcome this problem usually directly consider the heterogeneous system without a homogenization attempt.

As already mentioned in the introduction examples are the Multiscale Finite Element Method (MsFEM), the multiscale hybrid-mixed (MHM) finite element method and the Localized Orthogonal Decomposition (LOD). If considered in the periodic setting it turns out that the effective coefficients of the LOD coincide with the ones obtained in classical homogenization as shown in Gallistl and Peterseim (2017).

4.6 Application to the Maxwell–Debye system

At the end of this chapter we demonstrate how the rather abstract effective Maxwell system looks like for the Maxwell–Debye example from Section 3.1.3, which we scaled using the matrix \mathbf{Q} given in (3.32). Thus, the Maxwell–Debye system reads

$$\begin{aligned} \varepsilon_0 \varepsilon_\infty(x) \partial_t \mathbf{E}(t, x) + \varepsilon_0 \tau(x)^{-1} \Delta \varepsilon(x) \mathbf{E}(t, x) - \tau(x)^{-1} \mathbf{P}(t, x) & - \operatorname{curl} \mathbf{H}(t, x) = -\mathbf{J}(t, x), \\ \varepsilon_0^{-1} \Delta \varepsilon(x)^{-1} \partial_t \mathbf{P}(t, x) - \tau(x)^{-1} \Delta \varepsilon(x) \mathbf{E}(t, x) + \varepsilon_0^{-1} \tau(x)^{-1} \Delta \varepsilon(x)^{-1} \mathbf{P}(t, x) & = \mathbf{0}, \\ \mu_0 \partial_t \mathbf{H}(t, x) & + \operatorname{curl} \mathbf{E}(t, x) = \mathbf{0}. \end{aligned}$$

The identification of this system with the abstract Maxwell system (3.21) is clear by using the fact that in the Debye setting we have $N_E = 1$ and $N_H = 0$. Furthermore, the heterogeneous counterpart is given in abstract form in (4.11). The homogenization results and the reformulation from this chapter yield the effective system (4.37) with parameters (4.38). As in the example for the Maxwell system with

conductivity considered in Section 4.2.1, we present the decoupled Maxwell–Debye system

$$\begin{aligned} \varepsilon_0 \varepsilon_\infty^{\text{eff}}(x) \partial_t \mathbf{E}^{\text{eff}}(t, x) + \varepsilon_0 \mathbf{R}_{\text{EE}}^{\text{eff}}(x) \mathbf{E}^{\text{eff}}(t, x) + \mathbf{R}_{\text{EP}}^{\text{eff}}(x) \mathbf{P}^{\text{eff}}(t, x) + \int_0^t \mathbf{G}_{\text{EE}}^{\text{eff}}(t-s, x) \mathbf{E}^{\text{eff}}(s, x) \, ds \\ + \int_0^t \mathbf{G}_{\text{EP}}^{\text{eff}}(t-s, x) \mathbf{P}^{\text{eff}}(s, x) \, ds - \text{curl} \mathbf{H}^{\text{eff}}(t, x) = -\mathbf{J}(t, x) - \mathbf{J}_{\text{EE}}^0(t, x) \mathbf{E}_0(x) - \mathbf{J}_{\text{EP}}^0(t, x) \mathbf{P}_0(x), \end{aligned} \quad (4.77a)$$

$$\begin{aligned} \varepsilon_0^{-1} \mathbf{M}_{\text{P}}^{\text{eff}}(x) \partial_t \mathbf{P}^{\text{eff}}(t, x) + \mathbf{R}_{\text{PE}}^{\text{eff}}(x) \mathbf{E}^{\text{eff}}(t, x) + \varepsilon_0^{-1} \mathbf{R}_{\text{PP}}^{\text{eff}}(x) \mathbf{P}^{\text{eff}}(t, x) + \int_0^t \mathbf{G}_{\text{PE}}^{\text{eff}}(t-s, x) \mathbf{E}^{\text{eff}}(s, x) \, ds \\ + \int_0^t \mathbf{G}_{\text{PP}}^{\text{eff}}(t-s, x) \mathbf{P}(s, x) \, ds = -\mathbf{J}_{\text{PE}}^0(t, x) \mathbf{E}_0(x) - \mathbf{J}_{\text{PP}}^0(t, x) \mathbf{P}_0(x), \end{aligned} \quad (4.77b)$$

$$\mu_0 \partial_t \mathbf{H}^{\text{eff}}(t, x) - \text{curl} \mathbf{E}^{\text{eff}}(t, x) = \mathbf{0}. \quad (4.77c)$$

The notation of the parameters is based on the one in Section 3.1.4. For completeness, we also show the lengthy representation of the effective parameters. This relies on the decomposition of the cell correctors according to the three solution components. As an example consider the time dependent corrector $\bar{w}(t)$, solution of (4.34) together with (4.33) which is decomposed as

$$\bar{w}(t) = \begin{pmatrix} \bar{w}^{\varepsilon, \text{E}}(t) & \bar{w}^{\varepsilon, \text{P}}(t) & \bar{w}^{\varepsilon, \text{H}}(t) \\ \bar{w}^{\Delta \varepsilon, \text{E}}(t) & \bar{w}^{\Delta \varepsilon, \text{P}}(t) & \bar{w}^{\Delta \varepsilon, \text{H}}(t) \\ \bar{w}^{\mu, \text{E}}(t) & \bar{w}^{\mu, \text{P}}(t) & \bar{w}^{\mu, \text{H}}(t) \end{pmatrix} \in \mathbb{R}^{3 \times 9}.$$

Since the permeability in this example is constant the components of \bar{w} that involve μ or H vanish. With a similar decomposition of the cell correctors w^{M} and w^0 the derivation of the cell problems and parameters is possible. Eventually, the stationary parameters are given as

$$\begin{aligned} (\varepsilon_\infty^{\text{eff}}(x))_{k, \ell} &= \int_Y \varepsilon_\infty(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^\varepsilon(x, y)) \, dy, \\ (\mathbf{M}_{\text{P}}^{\text{eff}}(x))_{k, \ell} &= \int_Y \Delta \varepsilon(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\Delta \varepsilon}(x, y)) \, dy, \\ (\mathbf{R}_{\text{EE}}^{\text{eff}}(x))_{k, \ell} &= \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^\varepsilon(x, y)) \, dy, \\ (\mathbf{R}_{\text{EP}}^{\text{eff}}(x))_{k, \ell} &= - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\Delta \varepsilon}(x, y)) \, dy, \\ (\mathbf{R}_{\text{PE}}^{\text{eff}}(x))_{k, \ell} &= - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^\varepsilon(x, y)) \, dy, \\ (\mathbf{R}_{\text{PP}}^{\text{eff}}(x))_{k, \ell} &= \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\Delta \varepsilon}(x, y)) \, dy. \end{aligned}$$

Furthermore, the time-dependent convolution kernels and extra sources are defined as

$$\begin{aligned}
(\mathbf{G}_{\text{EE}}^{\text{eff}}(t, x))_{k, \ell} &= \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot \nabla_y \bar{w}_k^{\varepsilon, \text{E}}(t, x, y) \, dy \\
&\quad - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot \nabla_y \bar{w}_k^{\Delta \varepsilon, \text{E}}(t, x, y) \, dy, \\
(\mathbf{G}_{\text{EP}}^{\text{eff}}(t, x))_{k, \ell} &= \int_Y \tau^{-1}(x, y) \Delta \varepsilon(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot \nabla_y \bar{w}_k^{\varepsilon, \text{P}}(t, x, y) \, dy \\
&\quad - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot \nabla_y \bar{w}_k^{\Delta \varepsilon, \text{P}}(t, x, y) \, dy, \\
(\mathbf{G}_{\text{PE}}^{\text{eff}}(t, x))_{k, \ell} &= - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot \nabla_y \bar{w}_k^{\varepsilon, \text{E}}(t, x, y) \, dy \\
&\quad + \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot \nabla_y \bar{w}_k^{\Delta \varepsilon, \text{E}}(t, x, y) \, dy, \\
(\mathbf{G}_{\text{PP}}^{\text{eff}}(t, x))_{k, \ell} &= - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot \nabla_y \bar{w}_k^{\varepsilon, \text{P}}(t, x, y) \, dy \\
&\quad + \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot \nabla_y \bar{w}_k^{\Delta \varepsilon, \text{P}}(t, x, y) \, dy, \\
(\mathbf{J}_{\text{EE}}^0(t, x))_{k, \ell} &= \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y) (\mathbf{e}_\ell + \nabla_y^T w_\ell^\varepsilon(x, y)) \cdot \nabla_y w_k^{0, \varepsilon, \text{E}}(t, x, y) \, dy \\
&\quad - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot \nabla_y w_k^{0, \Delta \varepsilon, \text{E}}(t, x, y) \, dy, \\
(\mathbf{J}_{\text{EP}}^0(t, x))_{k, \ell} &= \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot \nabla_y w_k^{0, \varepsilon, \text{P}}(t, x, y) \, dy \\
&\quad - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^\varepsilon(x, y)) \cdot \nabla_y w_k^{0, \Delta \varepsilon, \text{P}}(t, x, y) \, dy, \\
(\mathbf{J}_{\text{PE}}^0(t, x))_{k, \ell} &= - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot \nabla_y w_k^{0, \varepsilon, \text{E}}(t, x, y) \, dy \\
&\quad + \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot \nabla_y w_k^{0, \Delta \varepsilon, \text{E}}(t, x, y) \, dy, \\
(\mathbf{J}_{\text{PP}}^0(t, x))_{k, \ell} &= - \int_Y \tau(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot \nabla_y w_k^{0, \varepsilon, \text{P}}(t, x, y) \, dy \\
&\quad + \int_Y \tau(x, y)^{-1} \Delta \varepsilon(x, y)^{-1} (\mathbf{e}_\ell + \nabla_y w_\ell^{\Delta \varepsilon}(x, y)) \cdot \nabla_y w_k^{0, \Delta \varepsilon, \text{P}}(t, x, y) \, dy.
\end{aligned}$$

CHAPTER 5

The Finite Element Heterogeneous Multiscale Method

Our ultimate goal in this thesis is to simulate the propagation of electromagnetic waves in heterogeneous media. In the first chapters we have shown mathematical models for these waves. Then we derived the homogeneous material parameters that allow us to simulate the propagation, as we no longer have to resolve every fine scale. In our case the method of choice for the simulation of partial differential equations is the finite element method, which we introduce briefly in the next section. As an example, we use the heterogeneous Maxwell system (4.11). Within its discretization we observe why we should use the homogeneous system (4.32) instead. One issue still present is that we only have the representation of the effective material parameters as mean values of solutions of partial differential equations. In general, it is not possible to give an exact solution of these equations and therefore we can not assume to get an exact analytic representation of the effective parameters. In Section 5.2, we introduce the Heterogeneous Multiscale Method (HMM) [Abdulle et al. \(2012\)](#); [E and Engquist \(2003\)](#) to overcome this problem. Moreover, we show the wellposedness and analyze the spatial error of this approach in Section 5.3.

5.1 The finite element method

In this section we briefly recapture the finite element method including the finite elements we use. The theory is mainly from [Brenner and Scott \(2008\)](#); [Ciarlet \(2002\)](#); [Ern and Guermond \(2004\)](#); [Monk \(2003\)](#). We start with the heterogeneous Maxwell system from Section 4.4.1

$$\mathbf{M}^\delta(x)\partial_t\mathbf{u}^\delta(t,x) + \mathbf{R}^\delta(x)\mathbf{u}^\delta(t,x) + \mathbf{A}\mathbf{u}^\delta(t,x) = \mathbf{g}^\delta(t,x), \quad \text{in } (0,T) \times \Omega, \quad (5.1)$$

$$\mathbf{u}^\delta(0,x) = \mathbf{u}_0^\delta(x), \quad \text{in } \Omega, \quad (5.2)$$

$$\mathbf{n} \times \mathbf{u}_1^\delta(t,x) = \mathbf{0}, \quad \text{on } (0,T) \times \partial\Omega. \quad (5.3)$$

Next we multiply equation (5.1) by a test function Φ . The space of test functions in our case is the domain of the Maxwell operator, and we give it another name for the sake of clarity, i.e.,

$$V^{\text{mac}} := \mathcal{D}(\mathcal{A}) = H_0(\text{curl}, \Omega) \times H(\text{curl}, \Omega)^{N_E} \times H(\text{curl}, \Omega) \times H(\text{curl}, \Omega)^{N_H}.$$

The superscript mac indicates that this space is used for the macroscopic Maxwell system. We now proceed with the variational formulation of (5.1).

Find $\mathbf{u}^\delta : [0, T] \rightarrow V^{\text{mac}}$ such that

$$(\mathbf{M}^\delta \partial_t \mathbf{u}^\delta(t) + \mathbf{R}^\delta \mathbf{u}^\delta(t) + \mathbf{A} \mathbf{u}^\delta(t), \Phi)_{L^2(\Omega; \mathbb{R}^n)} = (\mathbf{g}^\delta(t), \Phi)_{L^2(\Omega; \mathbb{R}^n)} \quad \text{for all } \Phi \in V^{\text{mac}}. \quad (5.4)$$

Next we choose a finite dimensional subspace $V_H \subseteq V^{\text{mac}}$ with $\dim V_H = N_{V_H}$. We give the definition of the space V_H in the Section 5.2. For now, it is enough that the finite dimensional space has a basis $\{\Phi_i\}_{i=1}^{N_{V_H}}$. Hence, we use it to represent the approximation $\mathbf{u}^{\delta, H} : [0, T] \rightarrow V_H$ of the solution as

$$\mathbf{u}^{\delta, H}(t, x) = \sum_{i=1}^{N_{V_H}} \mathbf{U}_i^\delta(t) \Phi_i(x).$$

Insert this representation in (5.4) to get

$$\begin{aligned} \sum_{i=1}^{N_{V_H}} (\mathbf{M}^\delta \Phi_i, \Phi_j)_{L^2(\Omega; \mathbb{R}^n)} \partial_t \mathbf{U}_i^\delta(t) + \sum_{i=1}^{N_{V_H}} (\mathbf{R}^\delta \Phi_i, \Phi_j)_{L^2(\Omega; \mathbb{R}^n)} \mathbf{U}_i^\delta(t) + \sum_{i=1}^{N_{V_H}} (\mathbf{A} \Phi_i, \Phi_j)_{L^2(\Omega; \mathbb{R}^n)} \mathbf{U}_i^\delta(t) \\ = (\mathbf{g}^\delta(t), \Phi_j)_{L^2(\Omega; \mathbb{R}^n)} \quad \text{for all } j = 1, \dots, N_{V_H}. \end{aligned} \quad (5.5)$$

Here it is enough to test with the basis functions itself. The system of equations (5.5) is equivalent to

$$\mathfrak{M}^\delta \partial_t \mathbf{U}^\delta(t) + \mathfrak{R}^\delta \mathbf{U}^\delta(t) + \mathfrak{A} \mathbf{U}^\delta(t) = \mathfrak{g}^\delta(t),$$

where the matrices and the right-hand side are defined for $i, j = 1, \dots, N_{V_H}$ as

$$\begin{aligned} \mathfrak{M}_{i,j}^\delta &:= (\mathbf{M}^\delta \Phi_j, \Phi_i)_{L^2(\Omega; \mathbb{R}^n)}, & \mathfrak{R}_{i,j}^\delta &:= (\mathbf{R}^\delta \Phi_j, \Phi_i)_{L^2(\Omega; \mathbb{R}^n)}, \\ \mathfrak{A}_{i,j} &:= (\mathbf{A} \Phi_j, \Phi_i)_{L^2(\Omega; \mathbb{R}^n)}, & \mathfrak{g}_i^\delta &:= (\mathbf{g}^\delta(t), \Phi_i)_{L^2(\Omega; \mathbb{R}^n)}. \end{aligned} \quad (5.6)$$

The solution vector is given as $\mathbf{U}^\delta(t) = \left(\mathbf{U}_1^\delta(t) \ \dots \ \mathbf{U}_{N_{V_H}}^\delta(t) \right)^T$. We derived a semi-discrete system that is an ODE and thus can in principle be solved by a time integration scheme. The matrices and the right-hand side in (5.6) have to be computed but still depend on the micro scale δ , which we can not resolve. To solve this resolution problem we introduced the homogeneous system (4.32), which we considered in Section 5.2.

For now, the open question is how to choose the finite dimensional subspace V_H . This is where the finite elements actually occur.

Remark 5.1.1. *Note that the subspace property $V_H \subseteq V^{\text{mac}}$ is one ingredient of a so-called conforming finite element method. We point out that this is not restrictive and that it is possible to use a finite dimensional space that is not a subspace of V^{mac} , e.g, a discontinuous Galerkin approach.*

The method of finite elements follows the idea to choose a subdivision of the computational domain Ω . This division is called triangulation and for its definition we follow (Ciarlet, 2002, Chapter 2).

Definition 5.1.2 (Triangulation). A triangulation \mathcal{T}_H of the domain Ω is a subdivision of Ω in a finite number of subsets K , the cells, such that the following conditions hold

- (i) $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_H} K$,
- (ii) For each $K \in \mathcal{T}_H$, the set K is closed and the interior $\overset{\circ}{K}$ is non-empty,
- (iii) For each distinct $K_i, K_j \in \mathcal{T}_H$, one has $\overset{\circ}{K}_i \cap \overset{\circ}{K}_j = \emptyset$,
- (iv) For each $K \in \mathcal{T}_H$, the boundary ∂K is Lipschitz-continuous.

If every cell $K \in \mathcal{T}_H$ is a polyhedron we call the triangulation adjacent if

- (v) Any face of K_i in the triangulation is either a subset of the boundary ∂K_i , or a face of another K_j in the triangulation.

For every cell we define its diameter $H_K := \text{diam}(K) = \max_{x, y \in K} |x - y|$. Thus, we get the maximum diameter of all elements $K \in \mathcal{T}_H$ as $H = \max_{K \in \mathcal{T}_h} H_K$, which explains the notation as \mathcal{T}_H . Moreover, denote by ϱ_K the diameter of the largest ball that can be inscribed in K .

A family of triangulations $\{\mathcal{T}_H : H > 0\}$ is called shape regular as $H \rightarrow 0$ if there exists a constant C independent of K such that for every H and for each $K \in \mathcal{T}_H$ we find

$$\frac{H_K}{\varrho_K} \leq C.$$

The implementation we strive for uses hexahedral elements only. Therefore, from now on assume that we have an adjacent and shape regular family of triangulations $\{\mathcal{T}_H\}_{H>0}$ of the domain Ω in parallelepipeds. We use the unit cube as reference cell $\hat{K} = [0, 1]^3$. Then, for every cell $K \in \mathcal{T}_H$ there exists an affine linear map F_K such that $F_K(\hat{K}) = K$.

Other choices of triangulations are possible, but we stick to this assumption in the remainder. Let us now define finite elements.

Definition 5.1.3 (Finite element). A finite element is a triple $\{K, \mathcal{P}_K, \mathcal{N}_K\}$ consisting of

- (i) K , a cell of the triangulation, i.e., a parallelepiped,
- (ii) \mathcal{P}_K , a space of functions on K ,
- (iii) \mathcal{N}_K , a set of linear functionals on \mathcal{P}_K , which are called degrees of freedom.

We call a finite element $(K, \mathcal{P}_K, \mathcal{N}_K)$ unisolvent if specifying a value for each degree of freedom in \mathcal{N}_K uniquely determines a function in \mathcal{P}_K .

On $K \in \mathcal{T}_H$ we define the space of polynomials with maximal degree ℓ in the first, m in the second and n in the third component as

$$\mathcal{Q}^{\ell, m, n}(K) = \left\{ p \in C^\infty(K) : \frac{\partial^{\ell+1} p}{\partial x_1^{\ell+1}} \equiv \frac{\partial^{m+1} p}{\partial x_2^{m+1}} \equiv \frac{\partial^{n+1} p}{\partial x_3^{n+1}} \equiv 0 \right\}.$$

A polynomial $p \in \mathcal{Q}^{\ell, m, n}(K)$ is of the form

$$p(x_1, x_2, x_3) = \sum_{r_1 \leq \ell, r_2 \leq m, r_3 \leq n} a_{r_1, r_2, r_3} x_1^{r_1} x_2^{r_2} x_3^{r_3} \quad \text{for } r_1, r_2, r_3 \in \mathbb{N}_0, a_{r_1, r_2, r_3} \in \mathbb{R}.$$

Next we finally define two finite elements. On the macro level we use Nédélec finite elements.

5.1.1 Nédélec finite elements

In 1980 Nédélec [Nédélec \(1980\)](#) introduced $H(\text{curl}, \Omega)$ -conforming elements of the first type for cubes. For $\ell \in \mathbb{N}$ we define

$$\mathcal{Q}_{\text{Nédélec}}^\ell(K) := \mathcal{Q}^{\ell-1, \ell, \ell}(K) \times \mathcal{Q}^{\ell, \ell-1, \ell}(K) \times \mathcal{Q}^{\ell, \ell, \ell-1}(K) \quad \text{for all } K \in \mathcal{T}_H.$$

We skip the lengthy definitions of the degrees of freedom \mathcal{N}_K here and refer the reader to ([Nédélec, 1980](#), Definition 6) and ([Monk, 2003](#), Definition 6.4) for the details on the practical realization. Then, the Nédélec element defined on the reference cell is given as

$$\left\{ \widehat{K}, \mathcal{Q}_{\text{Nédélec}}^\ell(\widehat{K}), \mathcal{N}_{\widehat{K}} \right\}.$$

Using an affine transformation this defines a general Nédélec element $\{K, \mathcal{Q}_{\text{Nédélec}}^\ell(K), \mathcal{N}_K\}$. We denote the space of Nédélec's elements of the first type by

$$\mathbf{V}^\ell(\text{curl}, \mathcal{T}_H) := \{v_H \in H(\text{curl}, \Omega) : v_H|_K \in \mathcal{Q}_{\text{Nédélec}}^\ell(K) \text{ for all } K \in \mathcal{T}_H\}, \quad (5.7)$$

and further we define

$$\mathbf{V}_0^\ell(\text{curl}, \mathcal{T}_H) := H_0(\text{curl}, \Omega) \cap \mathbf{V}^\ell(\text{curl}, \mathcal{T}_H).$$

Of course the definition (5.7) of Nédélec elements is rather implicit, which follows from the conformity and unisolvence of the elements defined by the degrees of freedom. Important, however, is that we obtain the following interpolation error estimate that can be found in the original paper ([Nédélec, 1980](#), Theorem 6) or in ([Monk, 2003](#), Theorem 6.6).

Theorem 5.1.4. *Let $u \in H^{\ell+1}(\Omega; \mathbb{R}^3)$. There exists a global interpolation operator $\mathcal{I}_H : H^{\ell+1}(\Omega; \mathbb{R}^3) \rightarrow \mathbf{V}^\ell(\text{curl}, \mathcal{T}_H)$ for Nédélec elements of the first type, such that*

$$\|u - \mathcal{I}_H u\|_{H(\text{curl}, \Omega)} \leq CH^\ell |u|_{H^{\ell+1}(\Omega; \mathbb{R}^3)}. \quad (5.8)$$

As seen in (5.7) the space of Nédélec elements is a subspace of $H(\text{curl}, \Omega)$. We use this space for the discretization of the macroscopic Maxwell system. All quantities that are labeled with an H are macroscopic expressions, as the triangulation \mathcal{T}_H . For microscopic quantities we use h instead. The next section is dedicated to Lagrange elements.

5.1.2 Lagrange finite elements

As explained in Section 5.2.1 below we use Lagrange elements on the micro scale. Since the cell problems are posed on the unit cell Y we introduce another adjacent and shape regular triangulation with maximal cell diameter h . Hence, we denote it as \mathcal{T}_h .

The details may be found in ([Ern and Guermond, 2004](#), Section 1.2.4). Following ([Ern and Guermond, 2004](#), Definition 1.27) the degrees of freedom of a Lagrange element are always associated to point or nodal values.

Similar to the previous section we only give the characterization of the space.

$$S^k(\mathcal{T}_h) := \{v_h \in H^1(Y) : v_h|_K \in \mathcal{Q}^{k, k, k}(K) \text{ for all } K \in \mathcal{T}_h\}.$$

To include periodic or Dirichlet boundary conditions we also define

$$\begin{aligned} S_{\#}^k(\mathcal{T}_h) &:= \mathbf{H}_{\#}^1(Y) \cap S^k(\mathcal{T}_h), \\ S_0^k(\mathcal{T}_h) &:= \mathbf{H}_0^1(Y) \cap S^k(\mathcal{T}_h). \end{aligned}$$

Similar to Theorem 5.1.4 we get an interpolation error estimate as in (Ciarlet, 2002, Theorem 3.2.1) or (Monk, 2003, Theorem 6.11).

Theorem 5.1.5. *For the triangulation \mathcal{T}_h of Y and a function $u \in \mathbf{H}^{k+1}(Y; \mathbb{R})$ there exists an interpolation operator $\Pi_h : \mathbf{H}^{k+1}(Y; \mathbb{R}) \rightarrow S^k(\mathcal{T}_h)$ such that the following estimate holds*

$$\|u - \Pi_h u\|_{\mathbf{H}^1(Y; \mathbb{R})} \leq Ch^k |u|_{\mathbf{H}^{k+1}(Y; \mathbb{R})}.$$

Let us point out that the space $S^k(\mathcal{T}_h)$ is a subspace of $\mathbf{H}^1(Y; \mathbb{R})$. Thus, it is suitable for conforming finite element methods applied to the micro problems (4.39)-(4.43).

Next we comment on quadrature rules. This is an important point in the theory of finite element methods. Note that the definition (5.6) of the matrices involves integrals over the domain Ω . These are unlikely to be exactly calculable. Thus, we use quadrature rules to approximate these integrals.

5.1.3 Quadrature

On every element $K \in \mathcal{T}_H$ of the triangulation we choose a quadrature formula consisting of $Q_K \in \mathbb{N}$ quadrature points $x_K^q \in \bar{K}$ and quadrature weights $\gamma_K^q \in \mathbb{R}_{\geq 0}$, $q = 1, \dots, Q_K$. We assume that the quadrature is exact for polynomials in $\mathcal{Q}^{2\ell, 2\ell, 2\ell}(K)$, $\ell \in \mathbb{N}$. Hence, for $p \in \mathcal{Q}^{2\ell, 2\ell, 2\ell}(K)$ we have

$$\int_K p(x) \, dx = \sum_{q=1}^{Q_K} \gamma_K^q p(x_K^q). \quad (5.9)$$

Note that we use positive weights, e.g. Gaussian rules, to ensure that the resulting bilinear forms keep, for instance, their positivity.

In the next section we apply the finite element method with Nédélec elements to the macroscopic Maxwell system (4.32).

5.2 The HMM framework and the application to Maxwell's equations

In Section 4.4.8, we already derived a variational formulation of the effective system (4.32), which is

$$\begin{aligned} m^{\text{eff}}(\partial_t \mathbf{u}^{\text{eff}}(t), \Phi) + r^{\text{eff}}(\mathbf{u}^{\text{eff}}(t), \Phi) + \int_0^t g^{\text{eff}}(t-s; \mathbf{u}^{\text{eff}}(s), \Phi) \, ds + a(\mathbf{u}^{\text{eff}}(t), \Phi) \\ = m^{\text{eff}}(\mathbf{f}(t), \Phi) - (\mathbf{J}^{\text{eff}}(t) \mathbf{u}_0, \Phi) \quad \text{for all } \Phi \in \mathbf{V}^{\text{mac}}. \end{aligned}$$

Recall the definition of the effective parameters in (4.38a)-(4.38d). As explained in Section 5.1, we now choose a finite dimensional subspace of the space \mathbf{V}^{mac} , which in our case will be the spaces of Nédélec finite elements presented in Section 5.1.1. To be more precise, for the space $\mathbf{H}(\text{curl}, \Omega)$ we choose the

subspace $V^\ell(\text{curl}, \mathcal{T}_H)$ and for the space $\mathbf{H}_0(\text{curl}, \Omega)$ we choose $V_0^\ell(\text{curl}, \mathcal{T}_H)$ with suitable choices of the degree $\ell \in \mathbb{N}$ and the triangulation \mathcal{T}_H of the domain Ω . The discrete space is then defined as the product space

$$V_H := V_0^\ell(\text{curl}, \mathcal{T}_H) \times V^\ell(\text{curl}, \mathcal{T}_H)^{N_E} \times V^\ell(\text{curl}, \mathcal{T}_H) \times V^\ell(\text{curl}, \mathcal{T}_H)^{N_H}.$$

Thus, we eventually get the system

$$\mathfrak{M}\partial_t \mathbf{U}(t) + \mathfrak{R}\mathbf{U}(t) + \int_0^t \mathfrak{G}(t-s)\mathbf{U}(s) \, ds + \mathfrak{A}\mathbf{U}(t) = \mathfrak{g}(t) - \mathfrak{J}(t)\mathbf{U}(0), \quad (5.10)$$

with an analogue definition of the matrices as in (5.6) with δ replaced by eff . Additionally, we have the convolution kernel and the extra source, which are given as

$$\mathfrak{G}(t)_{i,j} := (\mathbf{G}^{\text{eff}}(t)\Phi_j, \Phi_i)_{L^2(\Omega; \mathbb{R}^n)}, \quad \mathfrak{J}(t)_{i,j} := (\mathbf{J}^{\text{eff}}(t)\Phi_j, \Phi_i)_{L^2(\Omega; \mathbb{R}^n)}.$$

The exact calculation of the integrals in the definition of the matrices is in general not possible. To overcome this problem one uses quadrature formulas. Thus, on V_H we use the inner product

$$(\Phi_H, \Psi_H)_H = \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \gamma_K^q \Phi_H(x_K^q) \cdot \Psi_H(x_K^q) \quad \text{for all } \Phi_H, \Psi_H \in V_H,$$

cf. (5.9). The discrete counterparts to the bilinear forms defined in (4.64) - (4.67) are $m_H^{\text{eff}}, r_H^{\text{eff}}, a_H : V_H \times V_H \rightarrow \mathbb{R}$ such that for every $\Phi_H, \Psi_H \in V_H$

$$m_H^{\text{eff}}(\Phi_H, \Psi_H) := (\mathbf{M}^{\text{eff}}\Phi_H, \Psi_H)_H, \quad (5.11)$$

$$r_H^{\text{eff}}(\Phi_H, \Psi_H) := (\mathbf{R}^{\text{eff}}\Phi_H, \Psi_H)_H, \quad (5.12)$$

$$a_H(\Phi_H, \Psi_H) := (\mathbf{A}\Phi_H, \Psi_H)_H. \quad (5.13)$$

Moreover, for $t \in [0, T]$ define $g_H^{\text{eff}} : [0, T] \times V_H \times V_H \rightarrow \mathbb{R}$ such that

$$g_H^{\text{eff}}(t; \Phi_H, \Psi_H) := (\mathbf{G}^{\text{eff}}(t)\Phi_H, \Psi_H)_H \quad \text{for all } \Phi_H, \Psi_H \in V_H. \quad (5.14)$$

Now the spatially discrete formulation of the effective Maxwell system reads

$$\begin{aligned} m_H^{\text{eff}}(\partial_t \mathbf{u}_H^{\text{eff}}(t), \Phi_H) + r_H^{\text{eff}}(\mathbf{u}_H^{\text{eff}}(t), \Phi_H) + \int_0^t g_H^{\text{eff}}(t-s; \mathbf{u}_H^{\text{eff}}(s), \Phi_H) \, ds + a_H(\mathbf{u}_H^{\text{eff}}(t), \Phi_H) \\ = m_H^{\text{eff}}(\mathbf{f}_H(t), \Phi_H) - (\mathbf{J}^{\text{eff}}(t)\mathbf{u}_{0,H}, \Phi_H)_H \quad \text{for all } \Phi_H \in V_H, \end{aligned} \quad (5.15)$$

where \mathbf{f}_H is defined such that for an approximation \mathbf{g}_H of \mathbf{g} we find

$$m_H^{\text{eff}}(\mathbf{f}_H(t), \Phi_H) = (\mathbf{g}_H(t), \Phi_H)_H \quad \text{for all } \Phi_H \in V_H.$$

Note that the approximation of the effective mass matrix by

$$\mathfrak{M}_{i,j} = m_H^{\text{eff}}(\Phi_{H,j}, \Phi_{H,i}) \approx m_H^{\text{eff}}(\Phi_{H,j}, \Phi_{H,i}), \quad (5.16)$$

only has an effect on the macroscopic level, i.e., here we only introduce a quadrature error. In principle this yields a spatially discrete system but the effective parameter itself is not known so far. The problem

is that we have no knowledge about the matrix-valued effective parameters except that they are defined as mean values of cell correctors. Hence, we do not know the cell correctors analytically and thus we are not able to evaluate the parameters. This situation is exactly where the Heterogeneous Multiscale Method (HMM) [Abdulle et al. \(2012\)](#) is applicable. On the macroscopic level we have a problem that can in principle be solved, but we are lacking the information of the effective parameters. This information can be resolved by the use of a microscopic solver. Let us mention that a straightforward approach is to solve the cell problems with a suitable finite element method itself and then use these approximations to get the effective parameters. This is feasible in the context of purely periodic materials and in that setting it is probably even the best approach. However, in this work we go one step further and allow for at least locally periodic materials. In this scenario we would have to solve infinitely many cell problems, which is of course impossible. To overcome this problem we use the Heterogeneous Multiscale Method (HMM) introduced in [E and Engquist \(2003\)](#) and exposed in [Abdulle et al. \(2012\)](#). We present the idea of the HMM in the next section. After that we briefly recapture the results obtained for the vacuum Maxwell system in [Section 5.2.2](#) and then discuss the difficulties in our setting in [Section 5.2.3](#). Finally, in [Section 5.3](#) we present the Finite Element Heterogeneous Multiscale Method (FE-HMM) for the effective Maxwell system of passive materials and analyze the semi-discrete error of the method.

5.2.1 Approximating the effective parameters on the microscale

Consider the mass matrix components and their approximation as shown in [\(5.16\)](#). The question is how we can compute these entries. With a quadrature formula such as [\(5.9\)](#) we get

$$\begin{aligned} \mathfrak{M}_{i,j} &= (\mathbf{M}^{\text{eff}} \Phi_j, \Phi_i)_{L^2(\Omega; \mathbb{R}^n)} = \int_{\Omega} \mathbf{M}^{\text{eff}}(x) \Phi_j(x) \cdot \Phi_i(x) \, dx \\ &\approx \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \gamma_K^q \mathbf{M}^{\text{eff}}(x_K^q) \Phi_j(x_K^q) \cdot \Phi_i(x_K^q). \end{aligned} \quad (5.17)$$

Here we immediately see that we only need to know the effective parameter for every macroscopic quadrature point. As we see later it is helpful to use transformed cell problems and parameters. Therefore, we introduce for each quadrature point x_K^q a so-called sampling domain of size delta by $Y^\delta(x_K^q) = x_K^q + \delta Y$. Then the transformed cell problem (cf. [\(4.20\)](#)) and parameter (cf. [\(4.23\)](#)) are given as

$$\int_{Y^\delta(x_K^q)} \mathbf{M}\left(x_K^q, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^M(x_K^q, y)) \cdot \nabla_y v(y) \, dy = 0 \quad \text{for all } v \in \mathbf{H}_{\#}^1(Y^\delta(x_K^q)), \quad (5.18)$$

and

$$(\mathbf{M}^{\text{eff}}(x_K^q))_{k,\ell} = \frac{1}{|Y^\delta(x_K^q)|} \int_{Y^\delta(x_K^q)} \mathbf{M}\left(x_K^q, \frac{y}{\delta}\right) (\mathbf{e}_k + \nabla_y w_k^M(x_K^q, y)) \cdot (\mathbf{e}_\ell + \nabla_y w_\ell^M(x_K^q, y)) \, dy. \quad (5.19)$$

The details of the transformation can be found in the [Appendix A.2](#). Let us point out that the correctors w_ℓ^M in [\(4.20\)](#) and [\(5.18\)](#) do not coincide. Nevertheless, in what follows we only use the corrector as solution of [\(5.18\)](#) and thus we drop the δ -dependence in the notation. For convenience, we write $\bar{x} = x_K^q$, $\bar{\gamma} = \gamma_K^q$ and abbreviate the mean as $\bar{f} \cdot dy = \frac{1}{|Y|} \int_Y f \cdot dy$. Moreover, we insert the transformed representation of

the parameter into (5.17). Together with the notations from Remark 4.4.4 and the transposed Jacobian $\mathbf{D}_y^T \bar{w}$ this yields

$$\begin{aligned} \mathfrak{M}_{i,j} &\approx \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \Phi_i(\bar{x})^T \mathbf{M}^{\text{eff}}(\bar{x}) \Phi_j(\bar{x}) \\ &= \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \Phi_i(\bar{x})^T \int_{Y^\delta(\bar{x})} (\mathbf{I} + \mathbf{D}_y^T w^M(\bar{x}, y))^T \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) (\mathbf{I} + \mathbf{D}_y^T w^M(\bar{x}, y)) \, dy \Phi_j(\bar{x}) \\ &= \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \int_{Y^\delta(\bar{x})} [(\mathbf{I} + \mathbf{D}_y^T w^M(\bar{x}, y)) \Phi_i(\bar{x})]^T \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) [(\mathbf{I} + \mathbf{D}_y^T w^M(\bar{x}, y)) \Phi_j(\bar{x})] \, dy. \end{aligned}$$

The idea now is to reformulate the cell problems and to solve directly for the components of the expression $w^M(\bar{x}, y) \Phi_i(\bar{x})$, which are the solution of

$$\int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^M(\bar{x}, y)) \cdot \nabla_y v(y) \, dy \Phi_i(\bar{x}) = 0 \quad \text{for all } v \in \mathbf{H}_\#^1(Y^\delta(x_K^q)).$$

In other words, if we set $\Psi_i(\bar{x}, y) = (\mathbf{I} + \mathbf{D}_y^T w^M(\bar{x}, y)) \Phi_i(\bar{x})$ we get

$$\mathfrak{M}_{i,j} \approx \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \int_{Y^\delta(\bar{x})} \Psi_i(\bar{x}, y)^T \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \Psi_j(\bar{x}, y) \, dy, \quad (5.20)$$

where Ψ_i solves

$$\int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \Psi_i(\bar{x}, y) \cdot \nabla_y v(y) \, dy = 0 \quad \text{for all } v \in \mathbf{H}_\#^1(Y^\delta(x_K^q)).$$

Setting

$$\bar{\Psi}_{i,\text{lin}}(\bar{x}, y) := \bar{y} \cdot \Phi_i(\bar{x}), \quad \bar{\Psi}_{i,\#}(\bar{x}, y) := w^M(\bar{x}, y) \cdot \Phi_i(\bar{x}), \quad \bar{\Psi}_i(\bar{x}, y) := \bar{\Psi}_{i,\text{lin}}(\bar{x}, y) + \bar{\Psi}_{i,\#}(\bar{x}, y), \quad (5.21)$$

we get

$$\nabla_y \bar{\Psi}_i(\bar{x}, y) = \Psi_i(\bar{x}, y).$$

With this the new formulation of the cell problem, which is advantageous for the implementation, is given as:

Find $\bar{\Psi}_{i,\#} \in \mathbf{H}_\#^1(Y^\delta(x_K^q))$ such that

$$\int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Psi}_{i,\#}(\bar{x}, y) \cdot \nabla_y v(y) \, dy = - \int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Psi}_{i,\text{lin}}(\bar{x}) \cdot \nabla_y v(y) \, dy,$$

for all $v \in \mathbf{H}_\#^1(Y^\delta(x_K^q))$. Still, the problem remains that the exact solution of the described problem is in general not available. Before we tackle this problem, let us give a central remark.

Remark 5.2.1 (Knowledge about periodicity). *The definition of the parameters involves the solution of cell problems and unit cells. So far we made the assumption that we know the period length δ of the*

heterogeneous parameters exactly. In this case we use the periodic boundary conditions imposed in the derivation of the cell problems. In a more general setting where we have no knowledge about the period length we can come up with a slightly different approach: We choose a parameter $\kappa > \delta$ and define a new and bigger sampling domain $Y^\kappa(x_K^q) = x_K^q + \kappa Y$ around every macroscopic quadrature point x_K^q . Since we have no knowledge about the periodicity, we simply impose Dirichlet boundary conditions in this setting. For a detailed study on the effects of boundary conditions see [Yue and E \(2007\)](#).

In the view of Remark 5.2.1 we introduce a triangulation \mathcal{T}_h of the sampling domain $Y^\kappa(x_K^q)$. As function space we choose Lagrange finite elements of degree $k \in \mathbb{N}$ with either periodic $V^h = S_{\#}^k(\mathcal{T}_h)$ or Dirichlet boundary conditions $V^h = S_0^k(\mathcal{T}_h)$. Now the discrete micro problem is to find $\Psi_{i,\#}^h(\bar{x}, \cdot) \in V^h$ such that

$$\int_{Y^\kappa(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Psi}_{i,\#}^h(\bar{x}, y) \cdot \nabla_y v^h(y) \, dy = - \int_{Y^\kappa(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Psi}_{i,\text{lin}}(\bar{x}) \cdot \nabla_y v^h(y) \, dy, \quad (5.22)$$

for all $v^h \in V^h$. We use the solution of problem (5.22) to set

$$\Psi_i^h(\bar{x}, y) := \nabla_y \bar{\Psi}_{i,\text{lin}}(\bar{x}, y) + \nabla_y \bar{\Psi}_{i,\#}^h(\bar{x}, y) \quad \text{for } i = 1, \dots, N_{V_H},$$

and define the approximation of the mass matrix by

$$\mathfrak{M}_{i,j} \approx m^{\text{HMM}}(\Phi_j, \Phi_i) := \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \int_{Y^\kappa(\bar{x})} \Psi_i^h(\bar{x}, y)^T \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \Psi_j^h(\bar{x}, y) \, dy. \quad (5.23)$$

Note that in (5.23) besides the approximation by quadrature we also introduce an error by the introduction of the finite element solutions of the micro problem. This error is analyzed in Section 5.3. In addition to the definition of the bilinear form $m^{\text{HMM}}(\cdot, \cdot)$ in (5.23) we give another possibility in the next section, which will be useful in the error analysis later. We point out that the formulation in (5.23) is the one we use for implementation.

HMM material parameter

For the error analysis it is useful that we can introduce an HMM parameter such that the bilinear form from (5.23) is equivalent to

$$m^{\text{HMM}}(\Phi_j, \Phi_i) = (\mathbf{M}^{\text{HMM}} \Phi_j, \Phi_i)_H = \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \mathbf{M}^{\text{HMM}}(\bar{x}) \Phi_j(\bar{x}) \cdot \Phi_i(\bar{x}), \quad \text{for all } i, j = 1, \dots, N_{V_H}.$$

Note that the test functions in the right-hand side changed. This parameter \mathbf{M}^{HMM} is given as

$$(\mathbf{M}^{\text{HMM}}(\bar{x}))_{k,\ell} := \int_{Y^\kappa(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \left(\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M},h}(\bar{x}, y) \right) \cdot \left(\mathbf{e}_k + \nabla_y w_k^{\text{M},h}(\bar{x}, y) \right) \, dy, \quad (5.24)$$

where $w_\ell^{\text{M},h}(\bar{x}, \cdot) \in V^h$ for $\ell = 1, \dots, n$ is solution of

$$\int_{Y^\kappa(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \left(\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M},h}(\bar{x}, y) \right) \cdot \nabla_y v^h(y) \, dy = 0 \quad \text{for all } v^h \in V^h. \quad (5.25)$$

Error between the effective and the HMM parameter

In the error analysis we have to bound the expression

$$\|\mathbf{M}^{\text{eff}}(x_K^q) - \mathbf{M}^{\text{HMM}}(x_K^q)\|_F, \quad (5.26)$$

for every quadrature point x_K^q where $\|\cdot\|_F$ denotes the Frobenius norm and \mathbf{M}^{eff} and \mathbf{M}^{HMM} are given in (5.19) and (5.24), respectively. The expression in (5.26) is called the HMM error. To derive an estimate, we split it in the micro error and the modeling error. Therefore, in view of Remark 5.2.1 we set the infinite dimensional space \mathbf{V}^{mic} either as $\mathbf{H}_{\#}^1(Y^\kappa(x_K^q))$ and $\frac{\kappa}{\delta} \in \mathbb{N}$ or as $\mathbf{H}_0^1(Y^\kappa(x_K^q))$ and $\kappa > \delta$. Now, we introduce another parameter, which is given by the exact solution on the sampling domain

$$(\mathbf{M}^{\text{eff},\kappa}(\bar{x}))_{k,\ell} = \frac{1}{|Y^\kappa(\bar{x})|} \int_{Y^\kappa(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(\bar{x}, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(\bar{x}, y)) \, dy, \quad (5.27)$$

where $w_\ell^{\text{M}}(\bar{x}, \cdot) \in \mathbf{V}^{\text{mic}}$ solves

$$\int_{Y^\kappa(x_K^q)} \mathbf{M}\left(x_K^q, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x_K^q, y)) \cdot \nabla_y v(y) \, dy = 0 \quad \text{for all } v \in \mathbf{V}^{\text{mic}}. \quad (5.28)$$

Note that the parameter (5.27) is given by the exact solution of the micro problems that we use for the HMM parameter and therefore its error to the HMM parameter is a standard finite element error. The transformation to the new sampling domain $Y^\kappa(x_K^q)$ changes the corrector as shown in the Appendix A.2. Nevertheless, we keep the notation as w_ℓ^{M} since we only use the corrector given in (5.28) in the remainder. Now, we split the HMM error as follows

$$\|\mathbf{M}^{\text{eff}}(x_K^q) - \mathbf{M}^{\text{HMM}}(x_K^q)\|_F \leq \|\mathbf{M}^{\text{eff}}(x_K^q) - \mathbf{M}^{\text{eff},\kappa}(x_K^q)\|_F + \|\mathbf{M}^{\text{eff},\kappa}(x_K^q) - \mathbf{M}^{\text{HMM}}(x_K^q)\|_F. \quad (5.29)$$

The first expression on the right-hand side of (5.29) is the modeling error, which covers the error introduced by the sampling domain and, in the case of unknown period length, the Dirichlet boundary conditions. The second term is the micro error, which measures how good the discretization of the cell problem is. For the standard parameter \mathbf{M}^{eff} and its counterparts both the modeling and the micro error have been analyzed. For the modeling error we have the following result from (Abdulle et al., 2012, Lemma 4.8, Theorem 4.9).

Lemma 5.2.2. (i) *If the local problems (5.28) are solved with periodic boundary values, i.e., $\mathbf{V}^{\text{mic}} = \mathbf{H}_{\#}^1(Y^\kappa(x_K^q))$, we have for $\frac{\kappa}{\delta} \in \mathbb{N}$*

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \|\mathbf{M}^{\text{eff}}(x_K^q) - \mathbf{M}^{\text{eff},\kappa}(x_K^q)\|_F = 0.$$

(ii) *If the local problems (5.28) are solved with homogeneous Dirichlet boundary conditions, i.e., $\mathbf{V}^{\text{mic}} = \mathbf{H}_0^1(Y^\kappa(x_K^q))$, we have for $\kappa > \delta$*

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \|\mathbf{M}^{\text{eff}}(x_K^q) - \mathbf{M}^{\text{eff},\kappa}(x_K^q)\|_F \leq C \left(\frac{\delta}{\kappa} + \kappa \right),$$

with a constant $C > 0$ independent of h and δ .

As one immediately observes the modeling error vanishes in the case where we know the period length, because in this setting it is perfectly fine to use periodic boundary conditions. On the other hand, we can still get good results if we have no knowledge about the period.

The next result from (Abdulle, 2012, Lemma 5.2) or (Abdulle, 2005, Lemma 3.3) bounds the micro error. We also present the proof here because it gives ideas how we get similar bounds for the additional parameters \mathbf{R}^{eff} and \mathbf{G}^{eff} later.

Lemma 5.2.3. *Assume that for every quadrature point x_K^q we have $w_\ell^{\text{M}}(x_K^q, \cdot) \in \text{H}^2(Y^\kappa(x_K^q))$ such that for all $\ell = 1, \dots, n$ it holds*

$$\left| w_\ell^{\text{M}}(x_K^q, \cdot) \right|_{\text{H}^2(Y^\kappa(x_K^q))} \leq C\delta^{-1} \sqrt{|Y^\kappa(x_K^q)|}.$$

Then, we have the following estimate

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \left\| \mathbf{M}^{\text{eff}, \kappa}(x_K^q) - \mathbf{M}^{\text{HMM}}(x_K^q) \right\|_F \leq C \left(\frac{h}{\delta} \right)^2.$$

Proof. We abbreviate $\bar{x} := x_K^q$ and use for $\ell = 1, \dots, N$ the notation $\nabla_y \varphi_\ell(x, y) := (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(x, y))$ and $\nabla_y \varphi_\ell^h(x, y) := (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}, h}(x, y))$. Moreover, we define the bilinear form s_m^κ such that for all $\phi, \psi \in \text{H}^1(Y^\kappa(\bar{x}))$

$$s_m^\kappa(\phi, \psi) := \int_{Y^\kappa(\bar{x})} \mathbf{M} \left(\bar{x}, \frac{y}{\delta} \right) \nabla_y \phi(y) \cdot \nabla_y \psi(y) \, dy. \quad (5.30)$$

Thus, we find with a productive zero

$$\begin{aligned} |\mathbf{M}^{\text{eff}, \kappa}(\bar{x})_{k, \ell} - \mathbf{M}^{\text{HMM}}(\bar{x})_{k, \ell}| &= \frac{1}{|Y^\kappa(\bar{x})|} \left| s_m^\kappa(\varphi_\ell(\bar{x}), \varphi_k(\bar{x})) - s_m^\kappa(\varphi_\ell^h(\bar{x}), \varphi_k^h(\bar{x})) \right| \\ &= \frac{1}{|Y^\kappa(\bar{x})|} \left| s_m^\kappa(\varphi_\ell(\bar{x}), \varphi_k(\bar{x})) - s_m^\kappa(\varphi_\ell(\bar{x}), \varphi_k^h(\bar{x})) + s_m^\kappa(\varphi_\ell(\bar{x}), \varphi_k^h(\bar{x})) - s_m^\kappa(\varphi_\ell^h(\bar{x}), \varphi_k^h(\bar{x})) \right| \\ &= \frac{1}{|Y^\kappa(\bar{x})|} \left| s_m^\kappa(\varphi_\ell(\bar{x}), \varphi_k(\bar{x}) - \varphi_k^h(\bar{x})) + s_m^\kappa(\varphi_\ell(\bar{x}) - \varphi_\ell^h(\bar{x}), \varphi_k^h(\bar{x})) \right|. \end{aligned}$$

The first expression vanishes since $\varphi_k(\bar{x}) - \varphi_k^h(\bar{x}) \in \text{V}^{\text{mic}}$. Therefore, we have a test function suitable for the cell problem for w_ℓ^{M} , i.e.,

$$s_m^\kappa(\varphi_\ell(\bar{x}), v) = 0 \quad \text{for all } v \in \text{V}^{\text{mic}}.$$

On the other hand due to the symmetry of the parameter \mathbf{M} we also find that

$$s_m^\kappa(\varphi_\ell(\bar{x}) - \varphi_\ell^h(\bar{x}), \varphi_k(\bar{x})) = 0,$$

since $\varphi_\ell(\bar{x}) - \varphi_\ell^h(\bar{x}) \in \text{V}^{\text{mic}}$. We thus subtract this zero from the equality and get

$$\begin{aligned} |\mathbf{M}^{\text{eff}, \kappa}(\bar{x})_{k, \ell} - \mathbf{M}^{\text{HMM}}(\bar{x})_{k, \ell}| &= \frac{1}{|Y^\kappa(\bar{x})|} \left| s_m^\kappa(\varphi_\ell(\bar{x}) - \varphi_\ell^h(\bar{x}), \varphi_k^h(\bar{x}) - \varphi_k(\bar{x})) \right| \\ &\leq \frac{C_{\mathbf{M}}}{|Y^\kappa(\bar{x})|} \left\| \nabla_y (\varphi_\ell(\bar{x}) - \varphi_\ell^h(\bar{x})) \right\|_{\text{L}^2(Y^\kappa(\bar{x}))} \left\| \nabla_y (\varphi_k^h(\bar{x}) - \varphi_k(\bar{x})) \right\|_{\text{L}^2(Y^\kappa(\bar{x}))} \\ &= \frac{C_{\mathbf{M}}}{|Y^\kappa(\bar{x})|} \left\| \nabla_y (w_\ell^{\text{M}}(\bar{x}) - w_\ell^{\text{M}, h}(\bar{x})) \right\|_{\text{L}^2(Y^\kappa(\bar{x}))} \left\| \nabla_y (w_k^{\text{M}, h}(\bar{x}) - w_k^{\text{M}}(\bar{x})) \right\|_{\text{L}^2(Y^\kappa(\bar{x}))}, \end{aligned}$$

where we used the boundedness of the parameter \mathbf{M} from (4.46). With the standard finite element error bound from (Ciarlet, 2002, Theorem 3.2.2), i.e.,

$$\left\| \nabla_y \left(w_\ell^{\mathbf{M}}(\bar{x}) - w_\ell^{\mathbf{M},h}(\bar{x}) \right) \right\|_{L^2(Y^\kappa(\bar{x}))} \leq Ch \left| w_\ell^{\mathbf{M}}(\bar{x}) \right|_{\mathbf{H}^2(Y^\kappa(\bar{x}))},$$

we now get

$$\left| \mathbf{M}^{\text{eff},\kappa}(\bar{x})_{k,\ell} - \mathbf{M}^{\text{HMM}}(\bar{x})_{k,\ell} \right| \leq \frac{Ch^2}{|Y^\kappa(\bar{x})|} \left| w_\ell^{\mathbf{M}}(\bar{x}) \right|_{\mathbf{H}^2(Y^\kappa(\bar{x}))} \left| w_k^{\mathbf{M}}(\bar{x}) \right|_{\mathbf{H}^2(Y^\kappa(\bar{x}))}.$$

The assumption on the cell corrector yields the stated bound. \square

If the regularity of the corrector is higher we get the following corollary.

Corollary 5.2.4. *Assume that for $k \in \mathbb{N}$ we have $w_\ell^{\mathbf{M}}(x_K^q, \cdot) \in \mathbf{H}^{k+1}(Y^\kappa(x_K^q))$ such that for every $\ell = 1, \dots, n$ and every quadrature point x_K^q it holds*

$$\left| w_\ell^{\mathbf{M}}(x_K^q, \cdot) \right|_{\mathbf{H}^{k+1}(Y^\kappa(x_K^q))} \leq C \delta^{-k} \sqrt{|Y^\kappa(x_K^q)|}.$$

Then, we have the following estimate

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \left\| \mathbf{M}^{\text{eff},\kappa}(x_K^q) - \mathbf{M}^{\text{HMM}}(x_K^q) \right\|_F \leq C \left(\frac{h}{\delta} \right)^{2k}.$$

Proof. In the proof of Lemma 5.2.3 use the higher regularity of the corrector to get the bound

$$\left\| \nabla_y \left(w_\ell^{\mathbf{M}}(\bar{x}) - w_\ell^{\mathbf{M},h}(\bar{x}) \right) \right\|_{L^2(Y)} \leq Ch^k \left| w_\ell^{\mathbf{M}}(\bar{x}) \right|_{\mathbf{H}^{k+1}(Y)},$$

and then apply the assumption on the \mathbf{H}^{k+1} -norm of the corrector. \square

As already mentioned, this result is essential in the analysis of the Heterogeneous Multiscale Method since it bounds the so-called micro error of the scheme. A main contribution of this work is to establish similar results for the other parameters arising in the general Maxwell setting. Before we study the general system, we present some results obtained for the Maxwell system in vacuum.

5.2.2 Results for the classical Maxwell system

In this section we briefly present results for the Maxwell system without any damping. For details see Hochbruck et al. (2019). Here neither a polarization nor a magnetization is present and therefore the dimension of the problem is $n = 6$. Thus, the space \mathbf{V}^{mac} is given as

$$\mathbf{V}^{\text{mac}} = \mathbf{H}_0(\text{curl}, \Omega) \times \mathbf{H}(\text{curl}, \Omega).$$

The effective Maxwell system then reads:

Find $\mathbf{u}^{\text{eff}} : [0, T] \rightarrow \mathbf{V}^{\text{mac}}$ such that for all $\Phi \in \mathbf{V}^{\text{mac}}$ it holds

$$\begin{aligned} m^{\text{eff}}(\partial_t \mathbf{u}^{\text{eff}}(t), \Phi) + a(\mathbf{u}^{\text{eff}}(t), \Phi) &= m^{\text{eff}}(f^{\text{eff}}(t), \Phi), \\ \mathbf{u}^{\text{eff}}(0) &= \mathbf{u}_0. \end{aligned} \tag{5.31}$$

Here, the bilinear forms are given as in (4.64) and (4.66). The right-hand side is defined as

$$m^{\text{eff}}(f^{\text{eff}}(t), \Phi) = \left(\left(\begin{array}{c} -\mathbf{J}(t) \\ \mathbf{0}_3 \end{array} \right), \Phi \right)_{L^2(\Omega; \mathbb{R}^6)},$$

where \mathbf{J} is the electric current density from (3.1a). Compare with (4.68) for the general case. The formulation of the FE-HMM in this setting has been derived in Section 5.2.1. Here the discrete function space reads

$$V_H = V_0^\ell(\text{curl}, \mathcal{T}_H) \times V^\ell(\text{curl}, \mathcal{T}_H).$$

The variational form is:

Find $\mathbf{u}_H^{\text{HMM}}: [0, T] \rightarrow V_H$ such that for all $\Phi^H \in V_H$ it holds

$$\begin{aligned} m^{\text{HMM}}(\partial_t \mathbf{u}_H^{\text{HMM}}(t), \Phi^H) + a_H(\mathbf{u}_H^{\text{HMM}}(t), \Phi^H) &= m^{\text{HMM}}(f^{\text{HMM}}(t), \Phi^H), \\ \mathbf{u}_H^{\text{HMM}}(0) &= \mathbf{u}_{0,H}. \end{aligned} \quad (5.32)$$

The HMM bilinear form is given as in (5.23). Moreover, for an approximation $\mathbf{J}_H: [0, T] \rightarrow V^\ell(\text{curl}, \mathcal{T}_H)$ of \mathbf{J} we defined f^{HMM} such that

$$m^{\text{HMM}}(f^{\text{HMM}}(t), \Phi_H) = \left(\left(\begin{array}{c} -\mathbf{J}_H(t) \\ \mathbf{0}_3 \end{array} \right), \Phi_H \right).$$

Together with the Lemmas 5.2.2 and 5.2.3 or Corollary 5.2.4 we get the semi-discrete error bound as in (Hochbruck et al., 2019, Theorem 4.5) for Nédélec elements of order $\ell \in \mathbb{N}$ on the macro level and Lagrange elements of order $k \in \mathbb{N}$ on the micro level.

Theorem 5.2.5. *For $\ell \geq 1$ let $\mathbf{u}^{\text{eff}} \in C^1(0, T; H^{\ell+1}(\Omega))$ be the solution of (5.31). Moreover, let $\mathbf{u}^{\text{HMM}} \in L^\infty(0, T; V_H)$ be the solution of (5.32) at time $t \in (0, T)$. If*

$$|w^{\text{M}}|_{k+1, Y^\kappa(x_K^q)} \leq C\delta^{-k} \sqrt{|Y^\kappa(x_K^q)|} \quad \text{for all } x_K^q,$$

and

$$\mathbf{M}^{\text{eff}}|_K \in W^{\ell+1, \infty}(K)^{6 \times 6}, \quad \|\mathbf{M}^{\text{eff}}\|_{\ell+1, \infty, K} \leq \tilde{C}, \quad \text{for all } K \in \mathcal{T}_H,$$

with a constant \tilde{C} independent of δ , h and H , we get the error estimate

$$\begin{aligned} \|\mathbf{u}^{\text{HMM}}(t) - \mathbf{u}^{\text{eff}}(t)\|_{0, \Omega} &\leq C(1+t) \left(\|\mathbf{u}_{0,H} - \mathcal{I}_H \mathbf{u}_0\|_{0, \Omega} + \|\mathbf{J}_H - \mathbf{J}\|_{L^\infty(0, t; L^2(\Omega; \mathbb{R}^3))} \right. \\ &\quad \left. + H^\ell \left(\|\mathbf{u}^{\text{eff}}\|_{L^\infty(0, t; H^{\ell+1}(\Omega))} + \|\partial_t \mathbf{u}^{\text{eff}}\|_{L^\infty(0, t; H^{\ell+1}(\Omega))} \right) \right. \\ &\quad \left. + \left(\left(\frac{h}{\delta} \right)^{2k} + e_{\text{mod}} \right) \|\partial_t \mathbf{u}^{\text{eff}}\|_{L^\infty(0, t; H^{\ell+1}(\Omega))} \right), \end{aligned}$$

with $e_{\text{mod}} = 0$ under the assumption of Lemma 5.2.2 (i) or $e_{\text{mod}} < \frac{\delta}{\kappa} + \kappa$ under the assumptions of Lemma 5.2.2 (ii).

The main goal of the rest of this chapter is to show a similar bound for the more general Maxwell system and its discretization. The key difference between the two is the convolution that occurs due to the damping.

5.2.3 Difficulties in the approximation of time-dependent parameters

Let us briefly comment on the difficulties that occur in the error analysis of the general Maxwell case. We already saw in Chapter 4 that the effective Maxwell system changes in comparison to the vacuum Maxwell model. The main difference is the presence of the additional convolution and the extra source: The convolution introduces the difficulty that the effective convolution kernel is time dependent and this dependence leads to issues in both the analysis and the numerics. Moreover, the time dependence is given by a time-dependent corrector, that has to be analyzed separately. Not only the time-dependent parameter introduces a problem but already the effective damping parameter has a different structure than the effective parameter in the vacuum case. Therefore, we have to take care in the analysis of this parameter which is not time dependent at all. We thus have to derive results for the parameters \mathbf{R} and $\mathbf{G}(t)$ that are similar to Lemma 5.2.3 for \mathbf{M} . Our aim is to use similar techniques as in the proof of Theorem 5.2.5. But, since the convolution is not covered in the unified error analysis in Hipp et al. (2019), which is used for the effective vacuum Maxwell system, we have to generalize the proof here. This short overview shows how the introduction of damping in the Maxwell system, which is always present in matter, complicates the effective system. Nevertheless, we already gave the plan of how to circumvent these difficulties.

5.3 A FE-HMM algorithm for the dispersive Maxwell system

After the discussion of some general ideas and results for the classic Maxwell system in Section 5.2 we now turn to the general Maxwell system including damping. The starting point is the effective system we derived in Section 4.4 and which is given in (4.68). Recall that we seek $\mathbf{u}^{\text{eff}}: [0, T] \rightarrow \mathbf{V}^{\text{mac}}$ such that

$$\begin{aligned} m^{\text{eff}}(\partial_t \mathbf{u}^{\text{eff}}(t), \Phi) + r^{\text{eff}}(\mathbf{u}^{\text{eff}}(t), \Phi) + \int_0^t g^{\text{eff}}(t-s; \mathbf{u}^{\text{eff}}(s), \Phi) \, ds + a(\mathbf{u}^{\text{eff}}(t), \Phi) \\ = m^{\text{eff}}(\mathbf{f}(t), \Phi) - (\mathbf{J}^{\text{eff}}(t) \mathbf{u}_0, \Phi) \quad \text{for all } \Phi \in \mathbf{V}^{\text{mac}}. \end{aligned} \quad (5.33)$$

In (5.15) we already derived the discrete version:

Find $\mathbf{u}_H^{\text{eff}}: [0, T] \rightarrow \mathbf{V}_H$ such that

$$\begin{aligned} m_H^{\text{eff}}(\partial_t \mathbf{u}_H^{\text{eff}}(t), \Phi_H) + r_H^{\text{eff}}(\mathbf{u}_H^{\text{eff}}(t), \Phi_H) + \int_0^t g_H^{\text{eff}}(t-s; \mathbf{u}_H^{\text{eff}}(s), \Phi_H) \, ds + a_H(\mathbf{u}_H^{\text{eff}}(t), \Phi_H) \\ = m_H^{\text{eff}}(\mathbf{f}_H(t), \Phi_H) - (\mathbf{J}^{\text{eff}}(t) \mathbf{u}_{0,H}, \Phi_H)_H \quad \text{for all } \Phi_H \in \mathbf{V}_H. \end{aligned}$$

In Section 5.2.1 we further derived the bilinear form $m^{\text{HMM}}(\cdot, \cdot)$ defined in (5.23) that is actually computable, whereas $m_H^{\text{eff}}(\cdot, \cdot)$ still involves the exact solution of cell problems. In this section we follow the same ideas to derive the HMM formulation for the bilinear forms $r_H^{\text{eff}}(\cdot, \cdot)$ and $g_H^{\text{eff}}(t; \cdot, \cdot)$.

Derivation of r^{HMM}

For the damping parameter the derivation of the HMM counterpart is very similar to that of m^{HMM} . This is due to the fact that the same cell correctors are involved in its computation. As in Section 5.2.1

we start with the approximation of the matrix as

$$\mathfrak{R}_{i,j} = r^{\text{eff}}(\Phi_j, \Phi_i) \approx r_H^{\text{eff}}(\Phi_j, \Phi_i) \quad \text{for all } i, j = 1, \dots, N_{V_H},$$

where the bilinear form on the right-hand side is defined in (5.12). Let $k, \ell \in \{1, \dots, n\}$. The components of the effective parameter \mathbf{R}^{eff} are given in (4.25) as

$$(\mathbf{R}^{\text{eff}}(x))_{k,\ell} = \int_Y \mathbf{R}(x, y) (\mathbf{e}_\ell + \nabla_y w_\ell^M(x, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^M(x, y)) \, dy.$$

We transform this problem to the sampling domain $Y^\delta(\bar{x})$, where $\bar{x} = x_K^q$ is a macroscopic quadrature point

$$(\mathbf{R}^{\text{eff}}(\bar{x}))_{k,\ell} = \int_{Y^\delta(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^M(\bar{x}, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^M(\bar{x}, y)) \, dy. \quad (5.34)$$

Again note that the correctors have changed by transformation, but we keep the naming. In Section 5.2.1, we saw that it is beneficial to search for

$$\Psi_i(\bar{x}, y) = \nabla_y \bar{\Psi}_{i,\text{lin}}(\bar{x}, y) + \nabla_y \bar{\Psi}_{i,\#}(\bar{x}, y),$$

which are defined in (5.21), where $\bar{\Psi}_{i,\#}(\bar{x}, \cdot) \in H_{\#}^1(Y^\delta(x_K^q))$ solves

$$\int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Psi}_{i,\#}(\bar{x}, y) \cdot \nabla_y v(y) \, dy = - \int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Psi}_{i,\text{lin}}(\bar{x}) \cdot \nabla_y v(y) \, dy,$$

for all $v \in H_{\#}^1(Y^\delta(x_K^q))$. Now, the approximation of the matrix component can be computed similar to (5.20) by

$$\mathfrak{R}_{i,j} \approx \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \int_{Y^\delta(\bar{x})} \Psi_i(\bar{x}, y)^T \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \Psi_j(\bar{x}, y) \, dy.$$

Here, we have the same problem as for the matrix \mathbf{M} that the computation still involves exact solutions of PDEs, but we already showed how to deal with this. Following Remark 5.2.1, we introduce the solution $\bar{\Psi}_{i,\#}^h$ of (5.22), which is the approximation of $\bar{\Psi}_{i,\#}$ by a finite element method on a possibly different sampling domain. Finally, the HMM bilinear form corresponding to r^{eff} is

$$r^{\text{HMM}}(\Phi_j, \Phi_i) := \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \int_{Y^\delta(\bar{x})} \Psi_i^h(\bar{x}, y)^T \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \Psi_j^h(\bar{x}, y) \, dy. \quad (5.35)$$

Derivation of g^{HMM}

We now turn to the delay integral and its convolution kernel g^{eff} . Recall the definition of the discrete bilinear form in (5.14). The matrix components for $t \in [0, T]$ are

$$\mathfrak{G}(t)_{i,j} = g^{\text{eff}}(t; \Phi_j, \Phi_i) \approx g_H^{\text{eff}}(t; \Phi_j, \Phi_i) \quad \text{for all } i, j = 1, \dots, N_{V_H}.$$

Again choose $k, \ell \in \{1, \dots, n\}$. The components of the effective and transformed parameter \mathbf{G}^{eff} are given as (compare (4.26))

$$(\mathbf{G}^{\text{eff}}(t, \bar{x}))_{k,\ell} = \int_{Y^\delta(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell(t, \bar{x}, y) \cdot (\mathbf{e}_k + \nabla_y w_k^M(\bar{x}, y)) \, dy,$$

where we keep the notation of the correctors, although they changed by transformation. Following the lines of Section 5.2.1 we write

$$\begin{aligned} \mathfrak{G}(t)_{i,j} &\approx \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \Phi_i(\bar{x})^T \mathbf{G}^{\text{eff}}(t, \bar{x}) \Phi_j(\bar{x}) \\ &= \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \int_{Y^\delta(\bar{x})} [(\mathbf{I} + \mathbf{D}_y^T w^M(\bar{x}, y)) \Phi_i(\bar{x})]^T \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) [\mathbf{D}_y^T \bar{w}(t, \bar{x}, y) \Phi_j(\bar{x})] dy. \end{aligned}$$

We already explained how to deal with the first expression in the integral. This is the standard cell corrector. The new expression we have to evaluate is $\bar{w}(t, \bar{x}, y) \Phi_j(\bar{x})$. The idea is the same as before. We multiply the transformed cell problem for \bar{w}_ℓ (compare with (4.33), (4.34)) and its initial value problem by the macroscopic basis function. This yields

$$\int_{Y^\delta(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \partial_t \nabla_y \bar{w}_\ell(t, \bar{x}, y) + \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell(t, \bar{x}, y) \right] \cdot \nabla_y v(y) dy \Phi_j(\bar{x}) = 0,$$

and

$$\int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell(0, \bar{x}, y) \cdot \nabla_y v(y) dy \Phi_j(\bar{x}) = - \int_{Y^\delta(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^M(\bar{x}, y)) \cdot \nabla_y v(y) dy \Phi_j(\bar{x}),$$

for all $v \in \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$. Setting $\Theta_j(t, \bar{x}, y) = \mathbf{D}_y^T \bar{w}(t, \bar{x}, y) \Phi_j(\bar{x})$ and $\Psi_i(\bar{x}, y) = (\mathbf{I} + \mathbf{D}_y^T w^M(\bar{x}, y)) \Phi_i(\bar{x})$ as in (5.20) we get

$$\mathfrak{G}(t)_{i,j} \approx \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \int_{Y^\delta(\bar{x})} \Psi_i(\bar{x}, y)^T \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \Theta_j(t, \bar{x}, y) dy.$$

Here, Θ_j solves

$$\int_{Y^\delta(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \partial_t \Theta_j(t, \bar{x}, y) + \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \Theta_j(t, \bar{x}, y) \right] \cdot \nabla_y v(y) dy = 0,$$

and the $\Theta_j(0)$ is solution of

$$\int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \Theta_j(0, \bar{x}, y) \cdot \nabla_y v(y) dy = - \int_{Y^\delta(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \Psi_j(\bar{x}, y) \cdot \nabla_y v(y) dy.$$

Now we set

$$\bar{\Theta}_{j,\#}(t, \bar{x}, y) := \bar{w}(t, \bar{x}, y) \cdot \Phi_j(\bar{x}),$$

and get $\nabla_y \bar{\Theta}_{j,\#}(t, \bar{x}, y) = \Theta_j(t, \bar{x}, y)$. We see that $\bar{\Theta}_{j,\#}(t, \bar{x}, \cdot) \in \mathbf{H}_{\#}^1(Y^\delta(x_K^q))$ solves

$$\int_{Y^\delta(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \partial_t \nabla_y \bar{\Theta}_{j,\#}(t, \bar{x}, y) + \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Theta}_{j,\#}(t, \bar{x}, y) \right] \cdot \nabla_y v(y) dy = 0,$$

with

$$\int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Theta}_{j,\#}(0, \bar{x}, y) \cdot \nabla_y v(y) dy = - \int_{Y^\delta(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Psi}_j(\bar{x}, y) \cdot \nabla_y v(y) dy,$$

for all $v \in H_{\#}^1(Y^\delta(x_K^q))$. The right-hand side $\bar{\Psi}_j$ is defined in (5.21).

The final step is to use a finite element scheme to get an approximation $\bar{\Theta}_{j,\#}^h(t, \bar{x}, \cdot) \in V^h$ of $\bar{\Theta}_{j,\#}(t, \bar{x}, \cdot) \in V^{\text{mic}}$, which solves

$$\int_{Y^\kappa(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \partial_t \nabla_y \bar{\Theta}_{j,\#}^h(t, \bar{x}, y) + \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Theta}_{j,\#}^h(t, \bar{x}, y) \right] \cdot \nabla_y v^h(y) \, dy = 0,$$

with

$$\int_{Y^\kappa(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Theta}_{j,\#}^h(0, \bar{x}, y) + \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) (\nabla_y \bar{\Psi}_{j,\text{lin}}(\bar{x}, y) + \nabla_y \bar{\Psi}_{j,\#}^h(\bar{x}, y)) \right] \cdot \nabla_y v^h(y) \, dy = 0,$$

for all $v^h \in V^h$. We set

$$\Theta_j^h(t, \bar{x}, y) := \nabla_y \bar{\Theta}_{j,\#}^h(t, \bar{x}, y),$$

and the HMM bilinear form corresponding to g^{eff} is

$$g^{\text{HMM}}(t, \Phi_j, \Phi_i) := \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_\kappa} \bar{\gamma} \int_{Y^\kappa(\bar{x})} \Psi_i^h(\bar{x}, y)^T \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \Theta_j^h(t, \bar{x}, y) \, dy. \quad (5.36)$$

Derivation of J^0

The final expression of the system (5.10) that includes an effective parameter is

$$\mathfrak{J}(t)_{i,j} := (\mathbf{J}^{\text{eff}}(t) \Phi_j, \Phi_i)_{L^2(\Omega; \mathbb{R}^n)} \approx (\mathbf{J}^{\text{eff}}(t) \Phi_j, \Phi_i)_H \quad \text{for all } i, j = 1, \dots, N_{V_H}.$$

Here the components of the effective and transformed parameter \mathbf{J}^{eff} are given as (compare (4.27))

$$(\mathbf{J}^{\text{eff}}(t, \bar{x}))_{\kappa, \ell} = \int_{Y^\delta(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y w_\ell^0(t, \bar{x}, y) \cdot (\mathbf{e}_\kappa + \nabla_y w_\kappa^M(\bar{x}, y)) \, dy.$$

We proceed as for the convolution kernel with the difference to take the cell problems (4.35) and (4.36). This results in

$$(\mathbf{J}^{\text{HMM}}(t) \Phi_j, \Phi_i)_H := \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_\kappa} \bar{\gamma} \int_{Y^\kappa(\bar{x})} \Psi_i^h(\bar{x}, y)^T \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \Theta_j^{0,h}(t, \bar{x}, y) \, dy,$$

where as above

$$\Theta_j^{0,h}(t, \bar{x}, y) := \nabla_y \bar{\Theta}_{j,\#}^{0,h}(t, \bar{x}, y),$$

and $\bar{\Theta}_{j,\#}^{0,h}(t, \bar{x}, \cdot) \in V^h$ solves

$$\int_{Y^\kappa(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \partial_t \nabla_y \bar{\Theta}_{j,\#}^{0,h}(t, \bar{x}, y) + \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Theta}_{j,\#}^{0,h}(t, \bar{x}, y) \right] \cdot \nabla_y v^h(y) \, dy = 0,$$

with

$$\int_{Y^\kappa(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Theta}_{j,\#}^{0,h}(0, \bar{x}, y) + \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{\Psi}_{j,\text{lin}}(\bar{x}, y) \right] \cdot \nabla_y v^h(y) \, dy,$$

for all $v^h \in V^h$.

The FE-HMM Maxwell system

Combining the results from this section we get the Finite Element Heterogeneous Multiscale Method for the general Maxwell system as

$$m^{\text{HMM}}(\partial_t \mathbf{u}^{\text{HMM}}(t), \Phi_H) + r^{\text{HMM}}(\mathbf{u}^{\text{HMM}}(t), \Phi_H) + \int_0^t g^{\text{HMM}}(t-s; \mathbf{u}^{\text{HMM}}(s), \Phi_H) ds + a_H(\mathbf{u}^{\text{HMM}}(t), \Phi_H) = m^{\text{HMM}}(\mathbf{f}^{\text{HMM}}(t), \Phi_H) - (\mathbf{J}^{\text{HMM}}(t) \mathbf{u}_{0,H}, \Phi_H)_H \quad \text{for all } \Phi_H \in V_H, \quad (5.37)$$

where \mathbf{f}^{HMM} is defined such that

$$m^{\text{HMM}}(\mathbf{f}^{\text{HMM}}(t), \Phi_H) = (\mathbf{g}_H(t), \Phi_H)_H \quad \text{for all } \Phi_H \in V_H.$$

The bilinear forms are given in (5.23), (5.35) and (5.36) as well as in (5.13). The rest of this chapter is dedicated to the error analysis of the error

$$\|\mathbf{u}^{\text{eff}}(t) - \mathbf{u}^{\text{HMM}}(t)\|_{L^2(\Omega; \mathbb{R}^n)},$$

between the solution \mathbf{u}^{eff} of (5.33) and the solution \mathbf{u}^{HMM} of (5.37). The outline of this analysis is as follows. As pointed out for the classic Maxwell system we need to bound the error in the parameters. This is rather easy in the case of the parameter \mathbf{R}^{eff} with which we start in the next section. But it is a challenge for the time-dependent parameter \mathbf{G}^{eff} . Especially we need an error analysis for the corresponding cell problems, which are so-called Sobolev equations. We deal with these equations and the parameter \mathbf{G}^{eff} in Section 5.3.2. After that in Section 5.3.3, we show the wellposedness of the system (5.37). Eventually we fit the results together in Section 5.3.4 to get the requested semi-discrete error estimate for the FE-HMM.

5.3.1 Error analysis of the damping parameter

We start the analysis of the micro error with the damping parameter. Recall the definition of the effective transformed parameter from (5.34)

$$(\mathbf{R}^{\text{eff}}(\bar{x}))_{k,\ell} = \int_{Y^\delta(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(\bar{x}, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(\bar{x}, y)) dy.$$

For the error analysis it is again useful to reformulate the bilinear form (5.35) such that

$$r^{\text{HMM}}(\Phi_j, \Phi_i) := \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \Phi_i(\bar{x})^T \mathbf{R}^{\text{HMM}}(\bar{x}) \Phi_j(\bar{x}),$$

with the HMM parameter

$$(\mathbf{R}^{\text{HMM}}(\bar{x}))_{k,\ell} = \int_{Y^\kappa(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M},h}(\bar{x}, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M},h}(\bar{x}, y)) dy. \quad (5.38)$$

Here the corrector $w_\ell^{\text{M},h}$, $\ell = 1, \dots, n$ solves the cell problem on the sampling domain given in (5.25). Following the idea of the error analysis for the standard parameter and the Remark 5.2.1, we introduce the parameter

$$(\mathbf{R}^{\text{eff},\kappa}(\bar{x}))_{k,\ell} = \int_{Y^\kappa(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) (\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(\bar{x}, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(\bar{x}, y)) dy, \quad (5.39)$$

where $w_\ell^M(\bar{x}, \cdot) \in V^{\text{mic}}$ solves (5.28). Next we bound the modeling and the micro error separately. We start with the modeling error, which behaves exactly as the one for the parameter \mathbf{M}^{eff} in Lemma 5.2.2 since the same cell correctors are involved.

Lemma 5.3.1. (i) *If the local problems (5.28) are solved with periodic boundary values, i.e., we use the space $V^{\text{mic}} = H_{\#}^1(Y^\kappa(x_K^q))$, we have for $\frac{\kappa}{\delta} \in \mathbb{N}$*

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \|\mathbf{R}^{\text{eff}}(x_K^q) - \mathbf{R}^{\text{eff}, \kappa}(x_K^q)\|_F = 0.$$

(ii) *If the local problems (5.28) are solved with homogeneous Dirichlet boundary conditions, i.e., we use the space $V^{\text{mic}} = H_0^1(Y^\kappa(x_K^q))$, we have for $\kappa > \delta$*

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \|\mathbf{R}^{\text{eff}, \kappa}(x_K^q) - \mathbf{R}^{\text{eff}}(x_K^q)\|_F \leq C \left(\frac{\delta}{\kappa} + \kappa \right),$$

with a constant $C > 0$ independent of h and δ .

For the micro error things change in contrast to the standard parameter. This is obvious if one looks at the proof of Lemma 5.2.3. The proof involves the cell problems which are closely related to the parameter \mathbf{M} itself. In our current setting the cell problems are still valid but the parameter has changed from \mathbf{M} to \mathbf{R} . The next lemma shows how the micro error can be bounded.

Lemma 5.3.2. *For every quadrature point x_K^q we assume $w_\ell^M(x_K^q, \cdot), \bar{w}_\ell(0, x_K^q, \cdot) \in H^2(Y^\kappa(x_K^q))$ such that for every $\ell = 1, \dots, n$ it holds*

$$|w_\ell^M(x_K^q, \cdot)|_{H^2(Y^\kappa(x_K^q))}, |\bar{w}_\ell(0, x_K^q, \cdot)|_{H^2(Y^\kappa(x_K^q))} \leq C \delta^{-1} \sqrt{|Y^\kappa(x_K^q)|}. \quad (5.40)$$

Then, there exists a constant $C > 0$ independent of h and δ such that we have a bound on the Frobenius norm

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \|\mathbf{R}^{\text{eff}, \kappa}(x_K^q) - \mathbf{R}^{\text{HMM}}(x_K^q)\|_F \leq C \left(\frac{h}{\delta} \right)^2.$$

Proof. For $k = 1, \dots, N$ we introduce the short notation $\nabla_y \varphi_k(x, y) := (\mathbf{e}_k + \nabla_y w_k^M(x, y))$ as well as $\nabla_y \varphi_k^h(x, y) := (\mathbf{e}_k + \nabla_y w_k^{M, h}(x, y))$. Similar to the proof of Lemma 5.2.3 we introduce a bilinear form s_r^κ by

$$s_r^\kappa(\phi, \psi) := \int_{Y^\kappa(x_K^q)} \mathbf{R} \left(x_K^q, \frac{y}{\delta} \right) \nabla_y \phi(y) \cdot \nabla_y \psi(y) \, dy \quad \text{for all } \phi, \psi \in H^1(Y^\kappa(x_K^q)). \quad (5.41)$$

We investigate the difference, suppressing the x_K^q variable and find

$$|(\mathbf{R}^{\text{eff}, \kappa})_{k, \ell} - (\mathbf{R}^{\text{HMM}})_{k, \ell}| = \frac{1}{|Y^\kappa|} |s_r^\kappa(\varphi_\ell, \varphi_k) - s_r^\kappa(\varphi_\ell^h, \varphi_k^h)|.$$

We add and subtract $s_r^\kappa(\varphi_\ell^h, \varphi_k)$, which leads to

$$|(\mathbf{R}^{\text{eff}, \kappa})_{k, \ell} - (\mathbf{R}^{\text{HMM}})_{k, \ell}| = \frac{1}{|Y^\kappa|} |s_r^\kappa(\varphi_\ell - \varphi_\ell^h, \varphi_k) + s_r^\kappa(\varphi_\ell^h, \varphi_k - \varphi_k^h)|.$$

Recall the definition in (5.30). Now we use the transposed cell problem for $\bar{w}_k(0)$ (cf. 4.34) tested with $\varphi_\ell - \varphi_\ell^h$, i.e.,

$$s_m^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k(0)) + s_r^\kappa(\varphi_\ell - \varphi_\ell^h, \varphi_k) = 0,$$

in the first expression and add the cell problem for $\bar{w}_\ell(0)$ tested with $\varphi_k^h - \varphi_k$, i.e.,

$$s_m^\kappa(\bar{w}_\ell(0), \varphi_k^h - \varphi_k) + s_r^\kappa(\varphi_\ell, \varphi_k^h - \varphi_k) = 0.$$

This yields

$$\begin{aligned} |(\mathbf{R}^{\text{eff},\kappa})_{k,\ell} - (\mathbf{R}^{\text{HMM}})_{k,\ell}| &= \frac{1}{|Y^\kappa|} \left| -s_m^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k(0)) + s_r^\kappa(\varphi_\ell^h, \varphi_k - \varphi_k^h) \right. \\ &\quad \left. + s_m^\kappa(\bar{w}_\ell(0), \varphi_k^h - \varphi_k) + s_r^\kappa(\varphi_\ell, \varphi_k^h - \varphi_k) \right| \\ &= \frac{1}{|Y^\kappa|} \left| s_r^\kappa(\varphi_\ell^h - \varphi_\ell, \varphi_k - \varphi_k^h) + s_m^\kappa(\bar{w}_\ell(0), \varphi_k^h - \varphi_k) \right. \\ &\quad \left. - s_m^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k(0)) \right|. \end{aligned}$$

Finally, we use the standard cell problems (5.28) and (5.25) for $w_\ell^M, w_\ell^{M,h}$ and the transposed ones for $w_k^M, w_k^{M,h}$ tested with $\bar{w}_k^h(0)$ and $\bar{w}_\ell^h(0)$ respectively and add the resulting zeros, i.e.

$$s_m^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k^h(0)) = 0, \quad s_m^\kappa(\bar{w}_\ell^h(0), \varphi_k^h - \varphi_k) = 0,$$

which leads to

$$\begin{aligned} |(\mathbf{R}^{\text{eff},\kappa})_{k,\ell} - (\mathbf{R}^{\text{HMM}})_{k,\ell}| &= \frac{1}{|Y^\kappa|} \left| s_r^\kappa(\varphi_\ell^h - \varphi_\ell, \varphi_k - \varphi_k^h) + s_m^\kappa(\bar{w}_\ell(0) - \bar{w}_\ell^h(0), \varphi_k^h - \varphi_k) \right. \\ &\quad \left. + s_m^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k^h(0) - \bar{w}_k(0)) \right| \\ &= \frac{1}{|Y^\kappa|} \left| s_r^\kappa(w_\ell^{M,h} - w_\ell^M, w_k^M - w_k^{M,h}) + s_m^\kappa(\bar{w}_\ell(0) - \bar{w}_\ell^h(0), w_k^{M,h} - w_k^M) \right. \\ &\quad \left. + s_m^\kappa(w_\ell^M - w_\ell^{M,h}, \bar{w}_k^h(0) - \bar{w}_k(0)) \right|. \end{aligned}$$

An application of the Cauchy-Schwarz inequality and the boundedness of the parameters yields

$$\begin{aligned} |(\mathbf{R}^{\text{eff},\kappa})_{k,\ell} - (\mathbf{R}^{\text{HMM}})_{k,\ell}| &\leq C \frac{1}{|Y^\kappa|} \left\| \nabla_y (w_\ell^{M,h} - w_\ell^M) \right\|_{L^2(Y^\kappa; \mathbb{R}^N)} \left\| \nabla_y (w_k^M - w_k^{M,h}) \right\|_{L^2(Y^\kappa; \mathbb{R}^N)} \\ &\quad + C \frac{1}{|Y^\kappa|} \left\| \nabla_y (\bar{w}_\ell(0) - \bar{w}_\ell^h(0)) \right\|_{L^2(Y^\kappa; \mathbb{R}^N)} \left\| \nabla_y (w_k^{M,h} - w_k^M) \right\|_{L^2(Y^\kappa; \mathbb{R}^N)} \\ &\quad + C \frac{1}{|Y^\kappa|} \left\| \nabla_y (w_\ell^M - w_\ell^{M,h}) \right\|_{L^2(Y^\kappa; \mathbb{R}^N)} \left\| \nabla_y (\bar{w}_k^h(0) - \bar{w}_k(0)) \right\|_{L^2(Y^\kappa; \mathbb{R}^N)}. \end{aligned}$$

Now we apply the standard finite element error estimate (Ciarlet, 2002, Theorem 3.2.2) and obtain

$$\begin{aligned} |(\mathbf{R}^{\text{eff},\kappa})_{k,\ell} - (\mathbf{R}^{\text{HMM}})_{k,\ell}| &\leq C \frac{1}{|Y^\kappa|} h^2 |w_\ell^M|_{H^2(Y^\kappa; \mathbb{R}^N)} |w_k^M|_{H^2(Y^\kappa; \mathbb{R}^N)} \\ &\quad + C \frac{1}{|Y^\kappa|} h^2 |\bar{w}_\ell(0)|_{H^2(Y^\kappa; \mathbb{R}^N)} |w_k^M|_{H^2(Y^\kappa; \mathbb{R}^N)} \\ &\quad + C \frac{1}{|Y^\kappa|} h^2 |w_\ell^M|_{H^2(Y^\kappa; \mathbb{R}^N)} |\bar{w}_k(0)|_{H^2(Y^\kappa; \mathbb{R}^N)}. \end{aligned}$$

With the assumption on the regularity of the cell correctors we end up with the asserted estimate

$$|(\mathbf{R}^{\text{eff},\kappa})_{k,\ell} - (\mathbf{R}^{\text{HMM}})_{k,\ell}| \leq C \left(\frac{h}{\delta} \right)^2.$$

□

As for the parameter \mathbf{M}^{eff} we again get higher order estimates if the correctors have higher regularity.

Corollary 5.3.3. *Assume for $k \in \mathbb{N}$ that $w_\ell^{\text{M}}(x_K^q, \cdot), \bar{w}_\ell(0, x_K^q, \cdot) \in \mathbf{H}^{k+1}(Y^\kappa(x_K^q))$ such that for every $\ell = 1, \dots, n$ and every quadrature point x_K^q it holds*

$$|w_\ell^{\text{M}}(x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\kappa(x_K^q))}, |\bar{w}_\ell(0, x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\kappa(x_K^q))} \leq C\delta^{-k} \sqrt{|Y^\kappa(x_K^q)|}. \quad (5.42)$$

Then there exists a constant $C > 0$ independent of h and δ such that we have a bound on the Frobenius norm

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \|\mathbf{R}^{\text{eff}, \kappa}(x_K^q) - \mathbf{R}^{\text{HMM}}(x_K^q)\|_F \leq C \left(\frac{h}{\delta}\right)^{2k}.$$

Proof. Follow the lines of Lemma 5.3.2 and use higher regularity of the corrector. \square

After stating a short remark about the regularity assumption we continue our analysis with the time-dependent parameter.

Remark 5.3.4. *We point out that the regularity assumption (5.40) is for example satisfied if the parameters \mathbf{M} and \mathbf{R} satisfy*

$$\mathbf{M}^\delta|_K, \mathbf{R}^\delta|_K \in W^{1, \infty}(K; \mathbb{R}^{n \times n}), \quad |\mathbf{M}_{k, \ell}^\delta|_{W^{1, \infty}(K)}, |\mathbf{R}_{k, \ell}^\delta|_{W^{1, \infty}(K)} \leq C\delta^{-1},$$

for all $K \in \mathcal{T}_H$ and $k, \ell = 1, \dots, n$. See (Abdulle, 2012, Remark 5.1) for the details. For the higher regularity as in (5.42) we need again more regularity of the parameters.

We thus got a bound for the modeling and the micro error concerning the effective parameter \mathbf{R}^{eff} . In the next section we deal with the parameter related to the convolution. Especially we have to take a closer look at the related cell problem.

5.3.2 Error estimates for the Sobolev equation and the time-dependent parameter

In the proofs of Lemma 5.2.3 and 5.3.2 we used the result (Ciarlet, 2002, Theorem 3.2.2) to bound the \mathbf{H}^1 -error between cell correctors and their respective finite element approximations. The result was developed for elliptic problems and was thus directly applicable. In the forthcoming error analysis for the convolution kernel similar expressions occur, but this time these are not solutions of elliptic problems but of so-called Sobolev equations, which have a parabolic character. This is due to the structure of the cell problem for the corrector \bar{w} , which is in fact a Sobolev equation. Thus, before we study the micro error related to the convolution kernel we examine the error of a finite element method for the Sobolev equation.

Error estimate for Sobolev equation

We already analyzed the wellposedness of the homogeneous Sobolev equation in Section 4.4.5, where we also showed a stability result for the solution in Lemma 4.4.16. For the error analysis we consider the inhomogeneous equation. The results presented in this section are related to Bekkouche et al. (2019);

Thomé (1984) and follow the lines of Hipp et al. (2019). Recall the definition of the bilinear forms $s_m(\cdot, \cdot)$ and $s_r(\cdot, \cdot)$ in (4.45) and (4.53). Let $f \in C(0, T; \mathbf{V}^{\text{mic}})$ and seek $\bar{w} : [0, T] \rightarrow \mathbf{V}^{\text{mic}}$ such that

$$s_m(\partial_t \bar{w}(t), v) + s_r(\bar{w}(t), v) = s_m(f(t), v), \quad \text{for all } v \in \mathbf{V}^{\text{mic}}. \quad (5.43)$$

With the same technique as for the homogeneous equation (4.54) we find the solution by the variation of constants formula

$$\bar{w}(t) = e^{-\mathcal{S}t} \bar{w}(0) + \int_0^t e^{-\mathcal{S}(t-s)} f(s) \, ds,$$

with \mathcal{S} given in (4.55). Since the parameter \mathbf{R} is assumed to be positive semi-definite the contraction property (4.61) yields the stability bound

$$\|\bar{w}(t)\|_{s_m} \leq \|e^{-\mathcal{S}t} \bar{w}(0)\|_{s_m} + \int_0^t \|e^{-\mathcal{S}(t-s)} f(s)\|_{s_m} \, ds \leq \|\bar{w}(0)\|_{s_m} + \int_0^t \|f(s)\|_{s_m} \, ds.$$

As in (Thomé, 1984, Chapter 1) we introduce a finite dimensional subspace $\mathbf{V}^h \subseteq \mathbf{V}^{\text{mic}}$ with the typical finite element approximation property that for $r \geq 2$

$$\inf_{v^h \in \mathbf{V}^h} \|\nabla_y (u - v^h)\|_{L^2(Y; \mathbb{R}^N)} \leq Ch^{k-1} \|u\|_{\mathbf{H}^k(Y; \mathbb{R}^N)} \quad 1 \leq k \leq r \text{ for all } u \in \mathbf{H}^k(Y; \mathbb{R}^N) \cap \mathbf{V}^{\text{mic}}. \quad (5.44)$$

We point out that the subspace \mathbf{V}^h resulting from Lagrange elements introduced in Section 5.1.2 satisfies the property (5.44). Next we define the semi-discrete counterpart to (5.43) on \mathbf{V}^h with inhomogeneity: For $f \in C(0, T; \mathbf{V}^h)$ find $\bar{w}^h : [0, T] \rightarrow \mathbf{V}^h$ such that

$$s_m^h(\partial_t \bar{w}^h(t), v^h) + s_r^h(\bar{w}^h(t), v^h) = s_m^h(f^h(t), v^h) \quad \text{for all } v^h \in \mathbf{V}^h. \quad (5.45)$$

Similar to the continuous case we introduce the operator $\mathcal{S}^h : \mathbf{V}^h \rightarrow \mathbf{V}^h$ such that

$$s_r^h(\phi^h, \psi^h) = s_m^h(\mathcal{S}^h \phi^h, \psi^h) \quad \text{for all } \phi, \psi \in \mathbf{V}^h. \quad (5.46)$$

This yields that the problem is equivalent to

$$s_m^h(\partial_t \bar{w}^h(t), v^h) + s_m^h(\mathcal{S}^h \bar{w}^h(t), v^h) = s_m^h(f^h(t), v^h) \quad \text{for all } v^h \in \mathbf{V}^h.$$

The operator \mathcal{S}^h inherits all the properties of the operator \mathcal{S} . Thus, the semi-discrete problem is also wellposed and its solution satisfies the stability bound in the discrete norm

$$\|\bar{w}^h(t)\|_{s_m^h} \leq \|\bar{w}^h(0)\|_{s_m^h} + \int_0^t \|f^h(s)\|_{s_m^h} \, ds. \quad (5.47)$$

Since we have

$$s_m(\phi^h, \psi^h) = s_m^h(\phi^h, \psi^h), \quad s_r(\phi^h, \psi^h) = s_r^h(\phi^h, \psi^h) \quad \text{for all } \phi^h, \psi^h \in \mathbf{V}^h, \quad (5.48)$$

the discretization is conforming. The last ingredient we need for the error analysis is the s_m -orthogonal projection on \mathbf{V}^h denoted by Π^h , i.e.,

$$s_m(\Pi^h \phi, \psi^h) = s_m(\phi, \psi^h) \quad \text{for all } \phi \in \mathbf{V}^{\text{mic}}, \psi^h \in \mathbf{V}^h. \quad (5.49)$$

Lemma 5.3.5. *The s_m -orthogonal projection (5.49) satisfies for $1 \leq k \leq r$ and $w \in \mathbf{H}^k(Y; \mathbb{R}^N) \cap \mathbf{H}_{\#}^1(Y; \mathbb{R}^N)$ the estimate*

$$\|\Pi^h w - w\|_{s_m} \leq Ch^{k-1} \|w\|_{\mathbf{H}^k(Y; \mathbb{R}^N)}.$$

Proof. Due to Galerkin orthogonality, i.e.,

$$s_m(\Pi^h \phi - \phi, \psi^h) = 0 \quad \text{for all } \phi \in \mathbf{V}^{\text{mic}}, \psi^h \in \mathbf{V}^h,$$

and for $\tilde{v} \in \mathbf{V}^h$ with $\|\nabla_y(w - \tilde{v})\|_{\mathbf{L}^2(Y; \mathbb{R}^N)} = \inf_{v^h \in \mathbf{V}^h} \|\nabla_y(w - v^h)\|_{\mathbf{L}^2(Y; \mathbb{R}^N)}$ we find

$$\begin{aligned} \|\Pi^h w - w\|_{s_m}^2 &= s_m(\Pi^h w - w, \Pi^h w - w) = s_m(\Pi^h w - w, \Pi^h w) - s_m(\Pi^h w - w, w) \\ &= s_m(\Pi^h w - w, \tilde{v}^h) - s_m(\Pi^h w - w, w) = s_m(\Pi^h w - w, \tilde{v}^h - w) \\ &\leq \|\Pi^h w - w\|_{s_m} \|\tilde{v}^h - w\|_{s_m} \leq \|\Pi^h w - w\|_{s_m} \sqrt{C_{\mathbf{M}}} \|\nabla_y(\tilde{v}^h - w)\|_{\mathbf{L}^2(Y; \mathbb{R}^N)}. \end{aligned}$$

From (5.44) we get the result. \square

We can now give the semi-discrete error estimate for the Sobolev equation. This result estimates the error in the \mathbf{H}^1 -norm, which is the right norm in the setting of Sobolev equations and the effective parameters in homogenization.

Theorem 5.3.6. *Let \bar{w} be the solution of (4.54) and \bar{w}^h be the solution of*

$$s_m^h(\partial_t \bar{w}^h(t), v^h) + s_r^h(\bar{w}^h(t), v^h) = 0 \quad \text{for all } v^h \in \mathbf{V}^h. \quad (5.50)$$

Assume that \bar{w} satisfies $\bar{w} \in C^1(0, T; \mathbf{H}^k(Y))$. Then there exists a constant $C > 0$ independent of h and t such that

$$\|\bar{w}(t) - \bar{w}^h(t)\|_{s_m} \leq \|e^h(0)\|_{s_m^h} + Ch^{k-1} \left[\|\bar{w}(t)\|_{\mathbf{H}^k(Y)} + \int_0^t \|\bar{w}(s)\|_{\mathbf{H}^k(Y)} \, ds \right],$$

where

$$e^h(0) = \Pi^h \bar{w}(0) - \bar{w}^h(0).$$

Proof. We follow ideas of (Hipp et al., 2019, Theorem 2.8) and split the error in a projection error and a discretization error

$$\|\bar{w}(t) - \bar{w}^h(t)\|_{s_m} \leq \|\bar{w}(t) - \Pi^h \bar{w}(t)\|_{s_m} + \|\Pi^h \bar{w}(t) - \bar{w}^h(t)\|_{s_m}.$$

Consider the discretization error $e^h(t) = \Pi^h \bar{w}(t) - \bar{w}^h(t)$. For $v^h \in \mathbf{V}^h$ we find with (5.48), (4.56), (5.50), and (5.46)

$$\begin{aligned} s_m^h(\partial_t e^h(t), v^h) &= s_m(\partial_t e^h(t), v^h) = s_m(\partial_t \Pi^h \bar{w}(t), v^h) - s_m(\partial_t \bar{w}^h(t), v^h) \\ &= s_m(\partial_t \bar{w}(t), v^h) - s_m(\partial_t \bar{w}^h(t), v^h) \\ &= -s_m(\mathcal{S} \bar{w}(t), v^h) + s_m(\mathcal{S}^h \bar{w}^h(t), v^h) \\ &= -s_m(\Pi^h \mathcal{S} \bar{w}(t), v^h) + s_m(\mathcal{S}^h \bar{w}^h(t), v^h). \end{aligned}$$

We get

$$s_m(\partial_t e^h(t), v^h) + s_m(\mathcal{S}^h e^h(t), v^h) = s_m((\mathcal{S}^h \Pi^h - \Pi^h \mathcal{S}) \bar{w}(t), v^h) \quad \text{for all } v^h \in V^h.$$

Thus, the error itself solves an inhomogeneous semi-discrete equation of the structure (5.45). The application of the discrete stability bound (5.47) yields

$$\|e^h(t)\|_{s_m^h} \leq \|e^h(0)\|_{s_m^h} + \int_0^t \|(\mathcal{S}^h \Pi^h - \Pi^h \mathcal{S}) \bar{w}(s)\|_{s_m^h} ds.$$

It remains to estimate the integrand $\|(\mathcal{S}^h \Pi^h - \Pi^h \mathcal{S}) \bar{w}(s)\|_{s_m^h}$ for $s \in [0, t]$. Note that we have a discrete function $\xi^h(s) := (\mathcal{S}^h \Pi^h - \Pi^h \mathcal{S}) \bar{w}(s) \in V^h$ and thus using (5.46), (5.48), (5.49) and (4.55) we obtain

$$\begin{aligned} \|\xi^h(s)\|_{s_m^h}^2 &= s_m^h(\xi^h(s), \xi^h(s)) = s_m^h(\mathcal{S}^h \Pi^h \bar{w}(s), \xi^h(s)) - s_m^h(\Pi^h \mathcal{S} \bar{w}(s), \xi^h(s)) \\ &= s_r^h(\Pi^h \bar{w}(s), \xi^h(s)) - s_m(\Pi^h \mathcal{S} \bar{w}(s), \xi^h(s)) = s_r(\Pi^h \bar{w}(s), \xi^h(s)) - s_m(\mathcal{S} \bar{w}(s), \xi^h(s)) \\ &= s_r(\Pi^h \bar{w}(s), \xi^h(s)) - s_r(\bar{w}(s), \xi^h(s)) = s_r(\Pi^h \bar{w}(s) - \bar{w}(s), \xi^h(s)) \\ &\leq C_R \|\Pi^h \bar{w}(s) - \bar{w}(s)\|_{V^{\text{mic}}} \|\xi^h(s)\|_{V^{\text{mic}}} \leq \frac{C_R}{\alpha} \|\Pi^h \bar{w}(s) - \bar{w}(s)\|_{s_m} \|\xi^h(s)\|_{s_m}. \end{aligned}$$

Combining these results and using Lemma 5.3.5 yields

$$\begin{aligned} \|\bar{w}(t) - \bar{w}^h(t)\|_{s_m} &\leq \|\bar{w}(t) - \Pi^h \bar{w}(t)\|_{s_m} + \|e^h(0)\|_{s_m^h} + \frac{C_R}{\alpha} \int_0^t \|(\Pi^h \bar{w}(s) - \bar{w}^h(s))\|_{s_m} ds \\ &\leq \|e^h(0)\|_{s_m^h} + Ch^{k-1} \left[\|\bar{w}(t)\|_{\mathbb{H}^k(Y; \mathbb{R}^N)} + \int_0^t \|\bar{w}(s)\|_{\mathbb{H}^k(Y; \mathbb{R}^N)} ds \right], \end{aligned}$$

which implies the claimed estimate. \square

With this crucial bound we can now proceed with the analysis of the micro and the modeling error of the convolution kernel. The procedure is similar to that of Section 5.3.1. We first introduce an HMM parameter and then an intermediate quantity. This splitting is exactly the differentiation in micro and modeling error.

Error bound for the micro error of the convolution kernel

First recall that we defined the effective convolution kernel for all $t \in [0, T]$ in (4.26). The transformed counterpart on the sampling domain is

$$(\mathbf{G}^{\text{eff}}(t, \bar{x}))_{k, \ell} = \int_{Y^\delta(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell(t, \bar{x}, y) \cdot (\mathbf{e}_k + \nabla_y w_k^{\text{M}}(\bar{x}, y)) dy,$$

where we again point out that the correctors changed, but we stick to the notation. The reformulation of the HMM bilinear form (5.36) is

$$g^{\text{HMM}}(t; \Phi_j, \Phi_i) := \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \bar{\gamma} \Phi_i(\bar{x})^T \mathbf{G}^{\text{HMM}}(t, \bar{x}) \Phi_j(\bar{x}),$$

where the HMM parameter is defined as

$$\left(\mathbf{G}^{\text{HMM}}(t, \bar{x})\right)_{k,\ell} = \int_{Y^\kappa(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell^h(t, \bar{x}, y) \cdot \left(\mathbf{e}_k + \nabla_y w_k^{\text{M},h}(\bar{x}, y)\right) dy. \quad (5.51)$$

The cell corrector $w_k^{\text{M},h}$ solves the cell problem (5.25) and the corrector \bar{w}_ℓ^h is the finite element approximation of \bar{w}_ℓ , i.e., it solves:

Find $\bar{w}_\ell^h(\cdot, \bar{x}, \cdot) : [0, T] \rightarrow V^h$ such that

$$\int_{Y^\kappa(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \partial_t \nabla_y \bar{w}_\ell^h(t, \bar{x}, y) + \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell^h(t, \bar{x}, y) \right] \cdot \nabla_y v^h(y) dy = 0, \quad (5.52)$$

and

$$\int_{Y^\kappa(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell^h(0, \bar{x}, y) \cdot \nabla_y v^h(y) dy = - \int_{Y^\kappa(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \left(\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M},h}(\bar{x}, y)\right) \cdot \nabla_y v^h(y) dy, \quad (5.53)$$

for all $v^h \in V^h$. With the HMM parameter at hand we also need the intermediate parameter that is related to Remark 5.2.1. It is defined with the continuous solution of the transformed cell problem

$$\left(\mathbf{G}^{\text{eff},\kappa}(t, \bar{x})\right)_{k,\ell} = \int_{Y^\kappa(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell(t, \bar{x}, y) \cdot \left(\mathbf{e}_k + \nabla_y w_k^{\text{M}}(\bar{x}, y)\right) dy, \quad (5.54)$$

where w_k^{M} solves (5.28) and \bar{w}_ℓ is the solution of

$$\int_{Y^\kappa(\bar{x})} \left[\mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \partial_t \nabla_y \bar{w}_\ell(t, \bar{x}, y) + \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell(t, \bar{x}, y) \right] \cdot \nabla_y v(y) dy = 0, \quad (5.55)$$

with initial value $\bar{w}_\ell(0)$ solution of

$$\int_{Y^\kappa(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y \bar{w}_\ell(0, \bar{x}, y) \cdot \nabla_y v(y) dy = - \int_{Y^\kappa(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \left(\mathbf{e}_\ell + \nabla_y w_\ell^{\text{M}}(\bar{x}, y)\right) \cdot \nabla_y v(y) dy, \quad (5.56)$$

for all $v \in V^{\text{mic}}$. Note that we again stick to the notation \bar{w}_ℓ although the corrector changed due to transformation.

To our knowledge the analysis of the modeling error for the time dependent parameters is an open question. Since the expectation is, that this error vanishes in the periodic setting as for the stationary parameters we concentrate on the micro error. Unfortunately in this time-dependent case we are not able to get the expected convergence rate from the Lemmas 5.2.3 and 5.3.2. In particular, in contrast to the other parameters we only get first-order convergence in the micro error result. After showing this estimate we present some ideas of a possible starting point for a better convergence proof.

Lemma 5.3.7. *Assume that for every quadrature point x_K^q we have $w_\ell^{\text{M}}(x_K^q, \cdot) \in \mathbf{H}^2(Y^\kappa(x_K^q))$ and $\bar{w}_\ell(t, x_K^q, \cdot) \in \mathbf{H}^2(Y^\kappa(x_K^q))$ for all $t \in [0, T]$ such that for every $\ell = 1, \dots, n$ it holds*

$$\left|w_\ell^{\text{M}}(x_K^q, \cdot)\right|_{\mathbf{H}^2(Y^\kappa(x_K^q))} \leq C\delta^{-1} \sqrt{|Y^\kappa(x_K^q)|}, \quad \left|\bar{w}_\ell(0, x_K^q, \cdot)\right|_{\mathbf{H}^2(Y^\kappa(x_K^q))} \leq C\delta^{-1} \sqrt{|Y^\kappa(x_K^q)|},$$

for a constant $C > 0$. Then, there exists a constant $C > 0$ independent of t , h and δ such that we have a bound on the Frobenius norm

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \left\| \mathbf{G}^{\text{eff},\kappa}(t, x_K^q) - \mathbf{G}^{\text{HMM}}(t, x_K^q) \right\|_F \leq C(1+t) \left(\frac{h}{\delta}\right).$$

Proof. The proof of first-order convergence is considerably shorter than the second-order proofs in Lemmas 5.2.3 and 5.3.2. With the same notation as in Lemma 5.3.2, i.e., $\nabla_y \varphi_k(x, y) := (\mathbf{e}_k + \nabla_y w_k^M(x, y))$ and $\nabla_y \varphi_k^h(x, y) := (\mathbf{e}_k + \nabla_y w_k^{M,h}(x, y))$ for $k = 1, \dots, N$, we investigate the difference, suppressing the x_K^q variable. We use the bilinear forms s_m^κ and s_r^κ from (5.30) and (5.41). The error can be expressed by (5.54) and (5.51) as

$$|(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| = \frac{1}{|Y^\kappa|} |s_r^\kappa(\bar{w}_\ell(t), \varphi_k) - s_r^\kappa(\bar{w}_\ell^h(t), \varphi_k^h)|.$$

We add and subtract $s_r^\kappa(\bar{w}_\ell(t), \varphi_k^h)$ and obtain

$$|(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| = \frac{1}{|Y^\kappa|} |s_r^\kappa(\bar{w}_\ell(t), \varphi_k - \varphi_k^h) + s_r^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \varphi_k^h)|. \quad (5.57)$$

Next, we use the boundedness of the parameter \mathbf{R} as well as the boundedness of the solutions $\bar{w}_\ell(t)$ and φ_k^h from Lemmas 4.4.15 and 4.4.16. This yields with a constant $C > 0$

$$\begin{aligned} |(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| &\leq \frac{C}{|Y^\kappa|} \|\bar{w}_\ell(t)\|_{s_m} \|w_k^M - w_k^{M,h}\|_{s_m} + \frac{C}{|Y^\kappa|} \|\bar{w}_\ell(t) - \bar{w}_\ell^h(t)\|_{s_m} \|\varphi_k^h\|_{s_m} \\ &\leq \frac{C}{\sqrt{|Y^\kappa|}} \|w_k^M - w_k^{M,h}\|_{s_m} + \frac{C}{\sqrt{|Y^\kappa|}} \|\bar{w}_\ell(t) - \bar{w}_\ell^h(t)\|_{s_m}. \end{aligned}$$

The final step is to use the finite element error estimate from (Ciarlet, 2002, Theorem 3.2.2) in the first expression and the error result from Theorem 5.3.6 in the second one. This gives

$$\begin{aligned} &|(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| \\ &\leq \frac{C}{\sqrt{|Y^\kappa|}} h |w_k^M|_{\text{H}^2(Y^\kappa)} + \frac{C}{\sqrt{|Y^\kappa|}} \left(\|\bar{w}_\ell(0) - \bar{w}_\ell^h(0)\|_{s_m} + Ch \left(|\bar{w}_\ell(t)|_{\text{H}^2(Y^\kappa)} + \int_0^t |\bar{w}_\ell(s)|_{\text{H}^2(Y^\kappa)} \, ds \right) \right). \end{aligned}$$

Note that the difference of the initial values is again bounded with the result from (Ciarlet, 2002, Theorem 3.2.2). With the assumption on the correctors and Theorem 4.4.17 we get the final result

$$|(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| \leq C(1+t) \frac{h}{\delta}.$$

□

We again point out that this result could possibly be improved since the order of convergence seems to be non-optimal. Thus, in the remainder of this section we comment on the difficulties. At least we are able to get the requested second-order convergence for the initial value of the micro error. The starting point of the discussion is the equation (5.57). Instead of the direct use of the boundedness of \mathbf{R} as in Lemma 5.3.7 we test the transposed cell problem for $\bar{w}_k(0)$ (cf. (5.56)) with $\bar{w}_\ell(t) - \bar{w}_\ell^h(t) \in \mathbf{V}^{\text{mic}}$, i.e.,

$$s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k(0)) + s_r^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \varphi_k) = 0,$$

and subtract it from (5.57) to get

$$\begin{aligned} |(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| &= \frac{1}{|Y^\kappa|} \left| s_r^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \varphi_k^h - \varphi_k) + s_r^\kappa(\bar{w}_\ell(t), \varphi_k - \varphi_k^h) \right. \\ &\quad \left. - s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k(0)) \right|. \end{aligned}$$

In the second expression we use the problem for $\bar{w}_\ell(t)$ (5.55) tested with $\varphi_k - \varphi_k^h \in \mathbf{V}^{\text{mic}}$, i.e.,

$$s_m^\kappa(\partial_t \bar{w}_\ell(t), \varphi_k - \varphi_k^h) + s_r^\kappa(\bar{w}_\ell(t), \varphi_k - \varphi_k^h) = 0,$$

which yields

$$\begin{aligned} |(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| &= \frac{1}{|Y^\kappa|} \left| s_r^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \varphi_k^h - \varphi_k) - s_m^\kappa(\partial_t \bar{w}_\ell(t), \varphi_k - \varphi_k^h) \right. \\ &\quad \left. - s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k(0)) \right|. \end{aligned} \quad (5.58)$$

We further test the transposed cell problems for w_k^{M} and $w_k^{\text{M},h}$ (cf. (5.28) and (5.25)) with $\partial_t \bar{w}_\ell^h(t) \in V^h$, i.e.,

$$s_m^\kappa(\partial_t \bar{w}_\ell^h(t), \varphi_k) = 0, \quad s_m^\kappa(\partial_t \bar{w}_\ell^h(t), \varphi_k^h) = 0,$$

and subtract the two equations above. Adding the result to (5.58) leads to

$$\begin{aligned} |(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| &= \frac{1}{|Y^\kappa|} \left| s_r^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \varphi_k^h - \varphi_k) + s_m^\kappa(\partial_t \bar{w}_\ell^h(t) - \partial_t \bar{w}_\ell(t), \varphi_k - \varphi_k^h) \right. \\ &\quad \left. - s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k(0)) \right|. \end{aligned} \quad (5.59)$$

The remaining expression that is not in the right form with differences in both arguments is

$$-s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k(0)).$$

We proceed by using the representations of the solutions we found in Section 4.4.5 and the first part of Section 5.3.2. Observe the following equation including the operators \mathcal{S} and \mathcal{S}^h from (4.55) and (5.46) respectively

$$\begin{aligned} s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k^h(0)) &= s_m^\kappa(\exp(-\mathcal{S}t) \bar{w}_\ell(0) - \exp(-\mathcal{S}^h t) \bar{w}_\ell^h(0), \bar{w}_k^h(0)) \\ &= s_m^\kappa(\exp(-\mathcal{S}t) (\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k^h(0)) + s_m^\kappa((\exp(-\mathcal{S}t) - \exp(-\mathcal{S}^h t)) \bar{w}_\ell^h(0), \bar{w}_k^h(0)). \end{aligned}$$

Moreover, we find again with the solution representation

$$\begin{aligned} s_m^\kappa(\exp(-\mathcal{S}t) (\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k(0)) &= s_m^\kappa((\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \exp(-\mathcal{S}t) \bar{w}_k(0)) \\ &= s_m^\kappa((\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k(t)). \end{aligned}$$

Additionally, we can test the cell problems for $\bar{w}_\ell(0)$ and $\bar{w}_\ell^h(0)$ (cf. (5.56) and (5.53)) with $\bar{w}_k^h(t)$, which results in

$$s_m^\kappa((\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k^h(t)) + s_r^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k^h(t)) = 0.$$

Now again $\varphi_\ell - \varphi_\ell^h \in \mathbf{V}^{\text{mic}}$ is an admissible test function for $\bar{w}_k(t)$. Thus, we get from (5.55)

$$s_m^\kappa(\varphi_\ell - \varphi_\ell^h, \partial_t \bar{w}_k(t)) + s_r^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k(t)) = 0.$$

Moreover, $\partial_t \bar{w}_\ell^h(t) \in V^h$ is suitable for both standard cell problems (cf. (5.28) and (5.25)), i.e.,

$$s_m^\kappa(\varphi_\ell, \partial_t \bar{w}_k^h(t)) = 0, \quad s_m^\kappa(\varphi_\ell^h, \partial_t \bar{w}_k^h(t)) = 0.$$

Using these results in (5.59) finally yields

$$\begin{aligned}
& |(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| \\
&= \frac{1}{|Y^\kappa|} \left| s_r^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \varphi_k^h - \varphi_k) + s_m^\kappa(\partial_t \bar{w}_\ell^h(t) - \partial_t \bar{w}_\ell(t), \varphi_k - \varphi_k^h) \right. \\
&\quad - s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k(0)) + s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k^h(0)) \\
&\quad - s_m^\kappa(\exp(-\mathcal{S}t)(\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k^h(0)) - s_m^\kappa((\exp(-\mathcal{S}t) - \exp(-\mathcal{S}^h t)) \bar{w}_\ell^h(0), \bar{w}_k^h(0)) \\
&\quad + s_m^\kappa(\exp(-\mathcal{S}t)(\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k(0)) - s_m^\kappa((\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k(t)) \\
&\quad + s_m^\kappa((\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k^h(t)) + s_r^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k^h(t)) \\
&\quad \left. - s_r^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k(t)) - s_m^\kappa(\varphi_\ell - \varphi_\ell^h, \partial_t \bar{w}_k(t)) + s_m^\kappa(\varphi_\ell, \partial_t \bar{w}_k^h(t)) - s_m^\kappa(\varphi_\ell^h, \partial_t \bar{w}_k^h(t)) \right|.
\end{aligned}$$

Collecting the terms results in

$$\begin{aligned}
& |(\mathbf{G}^{\text{eff},\kappa}(t))_{k,\ell} - (\mathbf{G}^{\text{HMM}}(t))_{k,\ell}| \\
&= \frac{1}{|Y^\kappa|} \left| s_r^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \varphi_k^h - \varphi_k) + s_m^\kappa(\partial_t \bar{w}_\ell^h(t) - \partial_t \bar{w}_\ell(t), \varphi_k - \varphi_k^h) \right. \\
&\quad + s_m^\kappa(\bar{w}_\ell(t) - \bar{w}_\ell^h(t), \bar{w}_k^h(0) - \bar{w}_k(0)) + s_m^\kappa(\exp(-\mathcal{S}t)(\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k(0) - \bar{w}_k^h(0)) \\
&\quad + s_m^\kappa((\bar{w}_\ell(0) - \bar{w}_\ell^h(0)), \bar{w}_k^h(t) - \bar{w}_k(t)) + s_r^\kappa(\varphi_\ell - \varphi_\ell^h, \bar{w}_k^h(t) - \bar{w}_k(t)) \\
&\quad \left. + s_m^\kappa(\varphi_\ell - \varphi_\ell^h, \partial_t \bar{w}_k^h(t) - \partial_t \bar{w}_k(t)) - s_m^\kappa((\exp(-\mathcal{S}t) - \exp(-\mathcal{S}^h t)) \bar{w}_\ell^h(0), \bar{w}_k^h(0)) \right|.
\end{aligned} \tag{5.60}$$

However, we have a new expression which is not in the form with differences, i.e.,

$$-s_m^\kappa((\exp(-\mathcal{S}t) - \exp(-\mathcal{S}^h t)) \bar{w}_\ell^h(0), \bar{w}_k^h(0)). \tag{5.61}$$

Let us point out some things about the expression in (5.60). First note that if one considers the initial time $t = 0$ the remaining expression in (5.61) vanishes. Thus, with the same techniques as in Lemma 5.3.7 we get the desired second-order of convergence. Nevertheless, for $t > 0$ we still only get first-order of convergence. The numerical experiment in Chapter 7 suggest that, at least in the test cases considered, the expected second-order convergence rate can be observed. It is, however, an open question whether this is always true.

Again a direct consequence of the proof of Lemma 5.3.7 is that under higher regularity assumptions we get better convergence.

Corollary 5.3.8. *Assume for $k \in \mathbb{N}$ that $w_\ell^M(x_K^q, \cdot) \in \mathbf{H}^{k+1}(Y^\kappa(x_K^q))$ and $\bar{w}_\ell(t, x_K^q, \cdot) \in \mathbf{H}^{k+1}(Y^\kappa(x_K^q))$ for all $t \in [0, T]$ such that for every $\ell = 1, \dots, n$ it holds*

$$|w_\ell^M(x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\kappa(x_K^q))} \leq C\delta^{-k} \sqrt{|Y^\kappa(x_K^q)|}, \quad |\bar{w}_\ell(0, x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\kappa(x_K^q))} \leq C\delta^{-k} \sqrt{|Y^\kappa(x_K^q)|},$$

for a constant $C > 0$. Then, there exists a constant $C > 0$ independent of t, h and δ such that we have a bound on the Frobenius norm

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_K\}} \|\mathbf{G}^{\text{eff},\kappa}(t, x_K^q) - \mathbf{G}^{\text{HMM}}(t, x_K^q)\|_F \leq C(1+t) \left(\frac{h}{\delta}\right)^k.$$

Proof. Follow the lines of Lemma 5.3.7 and use the higher order results, especially from Theorem 5.3.6. \square

Micro error of extra source

As already observed before the results concerning the convolution can be directly transferred to the extra source \mathbf{J}^{eff} . Thus, we define the HMM extra source

$$(\mathbf{J}^{\text{HMM}}(t, \bar{x}))_{k,\ell} = \int_{Y^\kappa(\bar{x})} \mathbf{R}\left(\bar{x}, \frac{y}{\delta}\right) \nabla_y w_\ell^{0,h}(t, \bar{x}, y) \cdot \left(\mathbf{e}_k + \nabla_y w_k^{\text{M},h}(\bar{x}, y)\right) dy, \quad (5.62)$$

and obtain the following result.

Lemma 5.3.9. *Assume for $k \in \mathbb{N}$ that $w_\ell^{\text{M}}(x_K^q, \cdot) \in \mathbf{H}^{k+1}(Y^\kappa(x_K^q))$ and $\bar{w}_\ell(t, x_K^q, \cdot), w_\ell^0(t, x_K^q, \cdot) \in \mathbf{H}^{k+1}(Y^\kappa(x_K^q))$ for all $t \in [0, T]$ such that for every $\ell = 1, \dots, n$ there is a constant $C > 0$ with*

$$|w_\ell^{\text{M}}(x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\kappa(x_K^q))} \leq C\delta^{-k} \sqrt{|Y^\kappa(x_K^q)|}, \quad |w_\ell^0(0, x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\kappa(x_K^q))} \leq C\delta^{-k} \sqrt{|Y^\kappa(x_K^q)|}.$$

Then there exists a constant $C > 0$ independent of t, h and δ such that we have a bound on the Frobenius norm

$$\sup_{K \in \mathcal{T}_H, q \in \{1, \dots, Q_\kappa\}} \|\mathbf{J}^{\text{eff}, \kappa}(t, x_K^q) - \mathbf{J}^{\text{HMM}}(t, x_K^q)\|_F \leq C(1+t) \left(\frac{h}{\delta}\right)^k.$$

Here we again do not get the desired second-order rate. Be aware that the initial value $w_\ell^0(0)$ is the solution of (4.36). This is different compared to (4.34). Nevertheless, with the same techniques as in the previous section we can derive a second-order result for the error in the initial value. Although we were not able to derive the expected bounds on the micro error we now turn to the semi-discrete error analysis of the macroscopic Maxwell HMM system. Nevertheless, let us mention that a refinement in the proofs of the micro error in Lemmas 5.3.7 and 5.3.9 directly lead to better bounds in the space-discrete error result below.

5.3.3 Wellposedness of the semi-discrete system

Let us first comment on the relation between the effective and the HMM system. First note that due to the property $\mathbf{V}_H \subseteq \mathbf{V}^{\text{mac}}$ the discrete Maxwell operator inherits its properties from the continuous one. Moreover, the HMM parameters (5.24), (5.38), (5.51) and (5.62) satisfy the same bounds as the effective parameters as in Lemma 4.4.19.

Lemma 5.3.10. *The parameters $\mathbf{M}^{\text{HMM}}, \mathbf{R}^{\text{HMM}}$ are bounded, i.e., $\mathbf{M}^{\text{HMM}}, \mathbf{R}^{\text{HMM}} \in \mathbf{L}^\infty(\Omega; \mathbb{R}^{n \times n})$. In addition the parameter \mathbf{M}^{HMM} is coercive with the same constant α as \mathbf{M} .*

Moreover, the time-dependent parameters $\mathbf{G}^{\text{HMM}}(t)$ and $\mathbf{J}^{\text{HMM}}(t)$ are bounded for all $t \in [0, T]$, i.e., $\mathbf{G}^{\text{HMM}}(t), \mathbf{J}^{\text{HMM}}(t) \in \mathbf{L}^\infty(\Omega; \mathbb{R}^{n \times n})$. As a direct consequence the bilinear forms $m^{\text{HMM}}, r^{\text{HMM}}$ and $g^{\text{HMM}}(t)$ are bounded.

Proof. The proof is as in Lemma 4.4.19 just with the use of the discrete cell problems. \square

We now show that the HMM system (5.37) is well posed.

Theorem 5.3.11. *Assume that $\mathbf{u}^\delta, \mathbf{g}^\delta$ and the parameters \mathbf{M} and \mathbf{R} satisfy the assumptions from Theorem 4.4.6. Then the HMM system (5.37) has a unique solution $\mathbf{u}^{\text{HMM}} \in \mathbf{W}^{1,1}(0, T; \mathbf{V}_H)$.*

Proof. Note that the space V_H is finite dimensional and has a basis $\{\Phi_1, \dots, \Phi_{N_{V_H}}\}$. Therefore, we can represent the solution in this basis

$$\mathbf{u}^{\text{HMM}}(t, x) = \sum_{i=1}^{N_{V_H}} \mathbf{U}_i(t) \Phi_i(x).$$

Thus, we can rewrite the system (5.37) using the techniques from finite element theory in Section 5.1. With the matrices

$$\begin{aligned} \mathfrak{M}_{i,j}^H &= m^{\text{HMM}}(\Phi_j, \Phi_i), & \mathfrak{R}_{i,j}^H &= r^{\text{HMM}}(\Phi_j, \Phi_i), & \mathfrak{G}_{i,j}^H(t) &= g^{\text{HMM}}(t; \Phi_j, \Phi_i), \\ \mathfrak{A}_{i,j} &= a_H(\Phi_j, \Phi_i), & \mathfrak{g}_i^H(t) &= m^{\text{HMM}}(\mathbf{f}^{\text{HMM}}(t), \Phi_i), & \mathfrak{J}_{i,j}^H(t) &= (\mathbf{J}^{\text{HMM}}(t) \Phi_j, \Phi_i), \end{aligned}$$

we define the integro-differential matrix system as

$$\mathfrak{M}^H \partial_t \mathbf{U}(t) + \mathfrak{R}^H \mathbf{U}(t) + \int_0^t \mathfrak{G}^H(t-s) \mathbf{U}(s) ds + \mathfrak{A} \mathbf{U}(t) = \mathfrak{g}^H(t) - \mathfrak{J}^H(t) \mathbf{U}(0).$$

Reverting the reformulation from Section 4.4.4, i.e., from (4.37) to (4.15a) yields the equivalent system

$$\mathfrak{M}^H \partial_t \mathbf{U}(t) + \tilde{\mathfrak{R}}^H \mathbf{U}(t) + \partial_t \int_0^t \tilde{\mathfrak{G}}^H(t-s) \mathbf{U}(s) ds + \mathfrak{A} \mathbf{U}(t) = \mathfrak{g}^H(t) - \mathfrak{J}^H(t) \mathbf{U}(0).$$

We integrate this equation over $[0, t]$ and find

$$\mathfrak{M}^H \mathbf{U}(t) + \int_0^t (\tilde{\mathfrak{R}}^H + \tilde{\mathfrak{G}}^H(t-s) + \mathfrak{A}) \mathbf{U}(s) ds = \mathfrak{M}^H \mathbf{U}(0) + \int_0^t \mathfrak{g}^H(s) - \mathfrak{J}^H(s) \mathbf{U}(0) ds,$$

where we observe that we can apply Lemma 4.4.20. \square

In addition to the existence and uniqueness of a solution to the HMM system we also need a stability bound for the solution. Thanks to the space discretization, which satisfies $V_H \subseteq V^{\text{mac}}$, and due to the assumptions on the quadrature, this bound follows with the same techniques as in the continuous setting in (4.74) just with the discrete norms, i.e.,

$$\begin{aligned} \|\mathbf{u}^{\text{HMM}}(t)\|_{V_H} &\leq \exp\left(\int_0^t \|\mathbf{G}^{\text{HMM}}\|_{L^1(0,s;L^\infty(\Omega; \mathbb{R}^{n \times n}))} ds\right) \left[\frac{1}{\alpha} t \|\mathfrak{g}\|_{L^\infty(0,t;V_H)} \right. \\ &\quad \left. + \left(1 + \frac{1}{\alpha} \|\mathbf{J}^{\text{HMM}}\|_{L^1(0,t;L^\infty(\Omega; \mathbb{R}^{n \times n}))}\right) \|\mathbf{u}_0\|_{V_H} \right]. \end{aligned} \quad (5.63)$$

After this wellposedness study of the macroscopic HMM system we will now turn to the semi-discrete error analysis.

5.3.4 Semi-discrete a priori error analysis

We analyze the error between the solutions of the effective system (5.33) and the HMM system (5.37). We recall the effective equation

$$\begin{aligned} m^{\text{eff}}(\partial_t \mathbf{u}^{\text{eff}}(t), \Phi) + r^{\text{eff}}(\mathbf{u}^{\text{eff}}(t), \Phi) + \int_0^t g^{\text{eff}}(t-s; \mathbf{u}^{\text{eff}}(s), \Phi) ds + a(\mathbf{u}^{\text{eff}}(t), \Phi) \\ = m^{\text{eff}}(\mathbf{f}(t), \Phi) - (\mathbf{J}^{\text{eff}}(t) \mathbf{u}_0, \Phi) \quad \text{for all } \Phi \in V^{\text{mac}}, \end{aligned} \quad (5.64)$$

and the HMM system

$$\begin{aligned}
& m^{\text{HMM}}(\partial_t \mathbf{u}^{\text{HMM}}(t), \Phi_H) + r^{\text{HMM}}(\mathbf{u}^{\text{HMM}}(t), \Phi_H) + \int_0^t g^{\text{HMM}}(t-s; \mathbf{u}^{\text{HMM}}(s), \Phi_H) \, ds \\
& + a_H(\mathbf{u}^{\text{HMM}}(t), \Phi_H) = m^{\text{HMM}}(\mathbf{f}^{\text{HMM}}(t), \Phi_H) - (\mathbf{J}^{\text{HMM}}(t) \mathbf{u}_{0,H}, \Phi_H)_H \\
& \text{for all } \Phi_H \in V_H.
\end{aligned} \tag{5.65}$$

Recall that we have $N = 2 + N_E + N_H$, $n = 3N$ and assume $\ell \geq 1$. For the error analysis we use the following Hilbert spaces equipped with inner products or norms

$$\begin{aligned}
X &:= L^2(\Omega; \mathbb{R}^n), & (\phi, \psi)_X &= m^{\text{eff}}(\phi, \psi), \\
V^{\text{mac}} &:= H_0(\text{curl}, \Omega)^{1+N_E} \times H(\text{curl}, \Omega)^{1+N_H}, & (\phi, \psi)_{V^{\text{mac}}} &= (\phi, \psi)_{H(\text{curl}, \Omega)^N}, \\
V_H &:= V_0^\ell(\text{curl}, \mathcal{T}_H)^{1+N_E} \times V^\ell(\text{curl}, \mathcal{T}_H)^{1+N_H}, & (\phi^H, \psi^H)_{V_H} &= m^{\text{HMM}}(\phi^H, \psi^H), \\
Z &:= H^{\ell+1}(\Omega; \mathbb{R}^n), & \|\phi\|_Z &= \|\phi\|_{H^{\ell+1}(\Omega; \mathbb{R}^n)}.
\end{aligned}$$

Due to the properties of \mathbf{M} in (3.26) and the resulting properties for the bilinear forms m^{eff} in Lemma 4.4.19 and m^{HMM} in Lemma 5.3.10 the induced norms are equivalent to the standard $L^2(\Omega; \mathbb{R}^n)$ -norm, i.e., for all $\Phi \in X$ and $\Phi_H \in V_H$ we get

$$\sqrt{\alpha} \|\Phi\|_{L^2(\Omega; \mathbb{R}^n)} \leq \|\Phi\|_X \leq \sqrt{C_{\mathbf{M}}} \|\Phi\|_{L^2(\Omega; \mathbb{R}^n)}, \tag{5.66a}$$

$$\sqrt{\alpha} \|\Phi_H\|_{L^2(\Omega; \mathbb{R}^n)} \leq \|\Phi_H\|_{V_H} \leq \sqrt{C_{\mathbf{M}}} \|\Phi_H\|_{L^2(\Omega; \mathbb{R}^n)}. \tag{5.66b}$$

An immediate consequence is

$$\frac{\sqrt{\alpha}}{\sqrt{C_{\mathbf{M}}}} \|\Phi_H\|_X \leq \|\Phi_H\|_{V_H} \leq \frac{\sqrt{C_{\mathbf{M}}}}{\sqrt{\alpha}} \|\Phi_H\|_X. \tag{5.67}$$

In Section 5.1.1 we introduced the standard interpolation operator for Nédélec elements \mathcal{I}_H and we extend this to higher dimensions by component-wise application.

Since \mathbf{M}^{HMM} is positive definite we may introduce $\mathcal{P}_H : X \rightarrow V_H$ such that

$$m^{\text{HMM}}(\mathcal{P}_H \Phi, \Phi_H) = m^{\text{eff}}(\Phi, \Phi_H) \quad \text{for all } \Phi \in X, \Phi_H \in V_H. \tag{5.68}$$

We are now in the position to give the first error estimate, which is the starting point for the semi-discrete error analysis. The procedure follows Hipp et al. (2019) and Hochbruck et al. (2019), where we first derive a general error estimate. This estimate is successively refined afterwards using estimates on conformity errors. Observe that although $V_H \subseteq V^{\text{mac}}$, the HMM method fits in the setting of a non-conforming finite element method since the bilinear forms do not coincide. The micro and modeling errors will enter these conformity errors. We now state the most general error estimate similar to (Hipp et al., 2019, Theorem 2.8).

Theorem 5.3.12. *Let \mathbf{u}^{HMM} and \mathbf{u}^{eff} be the solutions of (5.65) and (5.64), respectively, and assume*

$\mathbf{u}^{\text{eff}} \in C^1([0, T]; Z)$. Then the error of the semi-discrete HMM-solution is bounded by

$$\begin{aligned}
& \|\mathbf{u}^{\text{HMM}}(t) - \mathbf{u}^{\text{eff}}(t)\|_X \\
& \leq \exp\left(\int_0^t \|\mathbf{G}^{\text{HMM}}\|_{L^1(0,s;L^\infty(\Omega;\mathbb{R}^{n \times n}))} \, ds\right) \left[\left(1 + \frac{1}{\alpha} \|\mathbf{J}^{\text{HMM}}\|_{L^1(0,t;L^\infty(\Omega;\mathbb{R}^{n \times n}))}\right) \|\mathbf{u}_{0,H} - \mathcal{I}_H \mathbf{u}_0\|_{V_H} \right. \\
& \quad + \frac{t}{\alpha} \|\mathbf{f}^{\text{HMM}} - \mathcal{P}_H \mathbf{f}\|_{L^\infty(0,t;V_H)} + \frac{t}{\alpha} \|(\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{R}^{\text{eff}} - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{R}^{\text{HMM}} \mathcal{I}_H) \mathbf{u}^{\text{eff}}\|_{L^\infty(0,t;V_H)} \\
& \quad + \frac{t}{\alpha} \|(\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{A} - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{A}_H \mathcal{I}_H) \mathbf{u}^{\text{eff}}\|_{L^\infty(0,t;V_H)} \\
& \quad + \frac{t}{\alpha} \sup_{s \in [0,t]} \left\| \int_0^s (\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{G}^{\text{eff}}(s-r) - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{G}^{\text{HMM}}(s-r) \mathcal{I}_H) \mathbf{u}^{\text{eff}}(r) \, dr \right\|_{V_H} \\
& \quad + \frac{t}{\alpha} \sup_{s \in [0,t]} \|\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{J}^{\text{eff}}(s) \mathbf{u}_0 - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{J}^{\text{HMM}}(s) \mathcal{I}_H \mathbf{u}_0\|_{V_H} + \frac{t}{\alpha} \|(\mathcal{P}_H - \mathcal{I}_H) \partial_t \mathbf{u}^{\text{eff}}\|_{L^\infty(0,t;V_H)} \Big] \\
& \quad + \|(\mathcal{I}_H - \text{id}) \mathbf{u}^{\text{eff}}(t)\|_X .
\end{aligned} \tag{5.69}$$

Proof. In the proof we abbreviate $\mathbf{u} = \mathbf{u}^{\text{eff}}$ and $\mathbf{u}^H = \mathbf{u}^{\text{HMM}}$ and do the same for the other expressions with superscript either $^{\text{eff}}$ or $^{\text{HMM}}$. We introduce the discrete error $\mathbf{e}_H(t) := \mathbf{u}^H(t) - \mathcal{I}_H \mathbf{u}(t) \in V_H$ and denote by id the identity operator. Observe that due to (5.66), we have

$$\|\mathbf{u}^H(t) - \mathbf{u}(t)\|_X \leq \|\mathbf{e}_H(t)\|_X + \|(\mathcal{I}_H - \text{id}) \mathbf{u}(t)\|_X \leq \frac{\sqrt{C_M}}{\sqrt{\alpha}} \|\mathbf{e}_H(t)\|_{V_H} + \|(\mathcal{I}_H - \text{id}) \mathbf{u}(t)\|_X . \tag{5.70}$$

To examine the discrete error we consider its time derivative: For any $\Phi_H \in V_H$ we get

$$\begin{aligned}
m^H(\partial_t \mathbf{e}_H(t), \Phi_H) &= m^H(\partial_t \mathbf{u}^H(t) - \mathcal{I}_H \partial_t \mathbf{u}(t), \Phi_H) \\
&= m^H(\partial_t \mathbf{u}^H(t) - \mathcal{P}_H \partial_t \mathbf{u}(t), \Phi_H) + m^H((\mathcal{P}_H - \mathcal{I}_H) \partial_t \mathbf{u}(t), \Phi_H) ,
\end{aligned} \tag{5.71}$$

and we rewrite the first part on the right-hand side of (5.71) using the effective and HMM systems (5.64) and (5.65) as well as (5.68)

$$\begin{aligned}
m^H(\partial_t \mathbf{u}^H(t) - \mathcal{P}_H \partial_t \mathbf{u}(t), \Phi_H) &= m^H(\partial_t \mathbf{u}^H(t), \Phi_H) - m(\partial_t \mathbf{u}(t), \Phi_H) \\
&= m^H(\mathbf{f}^H(t), \Phi_H) - r^H(\mathbf{u}^H(t), \Phi_H) - \int_0^t g^H(t-s; \mathbf{u}^H(s), \Phi_H) \, ds - a_H(\mathbf{u}^H(t), \Phi_H) \\
& \quad - (\mathbf{J}^H(t) \mathbf{u}_{0,H}, \Phi_H)_H - m(\mathbf{f}(t), \Phi_H) + r(\mathbf{u}(t), \Phi_H) \\
& \quad + \int_0^t g(t-s; \mathbf{u}(s), \Phi_H) \, ds + a(\mathbf{u}(t), \Phi_H) + (\mathbf{J}(t) \mathbf{u}_0, \Phi_H) \\
&= -r^H(\mathbf{e}_H(t), \Phi_H) - \int_0^t g^H(t-s; \mathbf{e}_H(s), \Phi_H) \, ds - a_H(\mathbf{e}_H(t), \Phi_H) - (\mathbf{J}^H(t) \mathbf{e}_H(0), \Phi_H)_H \\
& \quad + m^H(\mathbf{f}^H(t), \Phi_H) - m(\mathbf{f}(t), \Phi_H) + r(\mathbf{u}(t), \Phi_H) - r^H(\mathcal{I}_H \mathbf{u}(t), \Phi_H) \\
& \quad + a(\mathbf{u}(t), \Phi_H) - a_H(\mathcal{I}_H \mathbf{u}(t), \Phi_H) + \int_0^t g(t-s; \mathbf{u}(s), \Phi_H) \, ds - \int_0^t g^H(t-s; \mathcal{I}_H \mathbf{u}(s), \Phi_H) \, ds \\
& \quad + m(\mathbf{M}^{-1} \mathbf{J}(t) \mathbf{u}_0, \Phi_H) - m^H((\mathbf{M}^H)^{-1} \mathbf{J}^H(t) \mathcal{I}_H \mathbf{u}_0, \Phi_H) .
\end{aligned}$$

In the last two expressions we used the facts

$$(\mathbf{J}(t)\mathbf{u}_0, \Phi_H) = m(\mathbf{M}^{-1}\mathbf{J}(t)\mathbf{u}_0, \Phi_H), \quad (\mathbf{J}^H(t)\mathcal{I}_H\mathbf{u}_0, \Phi_H)_H = m^H\left((\mathbf{M}^H)^{-1}\mathbf{J}^H(t)\mathcal{I}_H\mathbf{u}_0, \Phi_H\right).$$

Recall the definitions in (4.64)-(4.67) of the continuous bilinear forms and the definitions in (5.23), (5.35), (5.13) and (5.36) of their discrete counterparts. Together with (5.71) and the property (5.68) we get a differential equation for the error \mathbf{e}_H

$$\begin{aligned} & m^H(\partial_t \mathbf{e}_H(t), \Phi_H) + r^H(\mathbf{e}_H(t), \Phi_H) + \int_0^t g^H(t-s; \mathbf{e}_H(s), \Phi_H) \, ds + a_H(\mathbf{e}_H(t), \Phi_H) \\ &= m^H(\mathbf{f}^H(t), \Phi_H) - (\mathbf{J}^H(t)\mathbf{e}_H(0), \Phi_H)_H - m(\mathbf{f}(t), \Phi_H) + r(\mathbf{u}(t), \Phi_H) - r^H(\mathcal{I}_H\mathbf{u}(t), \Phi_H) \\ & \quad + a(\mathbf{u}(t), \Phi_H) - a_H(\mathcal{I}_H\mathbf{u}(t), \Phi_H) + \int_0^t g(t-s; \mathbf{u}(s), \Phi_H) \, ds - \int_0^t g^H(t-s; \mathcal{I}_H\mathbf{u}(s), \Phi_H) \, ds \\ & \quad + m(\mathbf{M}^{-1}\mathbf{J}(t)\mathbf{u}_0, \Phi_H) - m^H((\mathbf{M}^H)^{-1}\mathbf{J}^H(t)\mathcal{I}_H\mathbf{u}_0, \Phi_H) + m^H((\mathcal{P}_H - \mathcal{I}_H)\partial_t\mathbf{u}(t), \Phi_H) \\ &= -(\mathbf{J}^H(t)\mathbf{e}_H(0), \Phi_H)_H + m^H(\tilde{\mathbf{f}}^H(t), \Phi_H), \end{aligned}$$

for the right-hand side given as

$$\begin{aligned} m^H(\tilde{\mathbf{f}}^H(t), \Phi_H) &= m^H(\mathbf{f}^H(t) - \mathcal{P}_H\mathbf{f}(t), \Phi_H) + m^H((\mathcal{P}_H\mathbf{M}^{-1}\mathbf{R} - (\mathbf{M}^H)^{-1}\mathbf{R}^H\mathcal{I}_H)\mathbf{u}(t), \Phi_H) \\ & \quad + m^H((\mathcal{P}_H\mathbf{M}^{-1}\mathbf{A} - (\mathbf{M}^H)^{-1}\mathbf{A}_H\mathcal{I}_H)\mathbf{u}(t), \Phi_H) \\ & \quad + m^H\left(\int_0^t (\mathcal{P}_H\mathbf{M}^{-1}\mathbf{G}(t-s) - (\mathbf{M}^H)^{-1}\mathbf{G}^H(t-s)\mathcal{I}_H)\mathbf{u}(s) \, ds, \Phi_H\right) \\ & \quad + m^H((\mathcal{P}_H\mathbf{M}^{-1}\mathbf{J}(t) - (\mathbf{M}^H)^{-1}\mathbf{J}^H(t)\mathcal{I}_H)\mathbf{u}_0, \Phi_H) + m^H((\mathcal{P}_H - \mathcal{I}_H)\partial_t\mathbf{u}(t), \Phi_H). \end{aligned}$$

At this point we use the stability estimate (5.63) for this discrete system. Note that we need $\tilde{\mathbf{f}} \in L^\infty(0, t; \mathbf{V}_H)$, which is the case due to the assumption on \mathbf{u} and the properties of the parameters. Thus, the error is bounded by

$$\begin{aligned} \|\mathbf{e}_H(t)\|_{\mathbf{V}_H} &\leq \exp\left(\int_0^t \|\mathbf{G}^H\|_{L^1(0,s;L^\infty(\Omega;\mathbb{R}^{n \times n}))} \, ds\right) \left[\left(1 + \frac{1}{\alpha} \|\mathbf{J}^H\|_{L^1(0,t;L^\infty(\Omega;\mathbb{R}^{n \times n}))}\right) \|\mathbf{e}_H(0)\|_{\mathbf{V}_H} \right. \\ & \quad + \frac{t}{\alpha} \|\mathbf{f}^H - \mathcal{P}_H\mathbf{f}\|_{L^\infty(0,t;\mathbf{V}_H)} + \frac{t}{\alpha} \|(\mathcal{P}_H\mathbf{M}^{-1}\mathbf{R} - (\mathbf{M}^H)^{-1}\mathbf{R}^H\mathcal{I}_H)\mathbf{u}\|_{L^\infty(0,t;\mathbf{V}_H)} \\ & \quad + \frac{t}{\alpha} \|(\mathcal{P}_H\mathbf{M}^{-1}\mathbf{A} - (\mathbf{M}^H)^{-1}\mathbf{A}_H\mathcal{I}_H)\mathbf{u}\|_{L^\infty(0,t;\mathbf{V}_H)} \\ & \quad + \frac{t}{\alpha} \sup_{s \in [0,t]} \left\| \int_0^s (\mathcal{P}_H\mathbf{M}^{-1}\mathbf{G}(s-r) - (\mathbf{M}^H)^{-1}\mathbf{G}^H(s-r)\mathcal{I}_H)\mathbf{u}(r) \, dr \right\|_{\mathbf{V}_H} \\ & \quad \left. + \frac{t}{\alpha} \sup_{s \in [0,t]} \|\mathcal{P}_H\mathbf{M}^{-1}\mathbf{J}(s)\mathbf{u}_0 - (\mathbf{M}^H)^{-1}\mathbf{J}^H(s)\mathcal{I}_H\mathbf{u}_0\|_{\mathbf{V}_H} + \frac{t}{\alpha} \|(\mathcal{P}_H - \mathcal{I}_H)\partial_t\mathbf{u}\|_{L^\infty(0,t;\mathbf{V}_H)} \right]. \end{aligned}$$

Along with (5.70) we showed the result. \square

The bound in (5.69) consists of many expressions. We proceed by bounding each of those expressions separately. The first two steps are given in Hipp et al. (2019). We start with $\|(\mathcal{P}_H - \mathcal{I}_H)\partial_t\mathbf{u}^{\text{eff}}\|_{L^\infty(0,t;\mathbf{V}_H)}$. From (5.67) we can bound the expression with the following Lemma.

Lemma 5.3.13. For all $\Phi \in Z$ we have with a constant $C > 0$

$$\|(\mathcal{P}_H - \mathcal{I}_H)\Phi\|_{V_H} \leq C \|(\text{id} - \mathcal{I}_H)\Phi\|_X + \max_{\|\Psi_H\|_{V_H}=1} |\Delta m(\mathcal{I}_H\Phi, \Psi_H)|,$$

where for Φ_H, Ψ_H we define

$$\Delta m(\Phi_H, \Psi_H) := m^{\text{eff}}(\Phi_H, \Psi_H) - m^{\text{HMM}}(\Phi_H, \Psi_H). \quad (5.72)$$

Proof. For all $\Phi_H \in V_H$ we have

$$\|\Phi_H\|_{V_H} = \max_{\|\Psi_H\|_{V_H}=1} m^{\text{HMM}}(\Phi_H, \Psi_H).$$

We thus get for $\Phi \in Z$ with the property of \mathcal{P}_H in (5.68)

$$\begin{aligned} \|(\mathcal{P}_H - \mathcal{I}_H)\Phi\|_{V_H} &= \max_{\|\Psi_H\|_{V_H}=1} m^{\text{HMM}}((\mathcal{P}_H - \mathcal{I}_H)\Phi, \Psi_H) \\ &= \max_{\|\Psi_H\|_{V_H}=1} [m^{\text{HMM}}(\mathcal{P}_H\Phi, \Psi_H) - m^{\text{HMM}}(\mathcal{I}_H\Phi, \Psi_H)] \\ &= \max_{\|\Psi_H\|_{V_H}=1} [m^{\text{eff}}(\Phi, \Psi_H) - m^{\text{eff}}(\mathcal{I}_H\Phi, \Psi_H) + m^{\text{eff}}(\mathcal{I}_H\Phi, \Psi_H) - m^{\text{HMM}}(\mathcal{I}_H\Phi, \Psi_H)] \\ &= \max_{\|\Psi_H\|_{V_H}=1} [m^{\text{eff}}((\text{id} - \mathcal{I}_H)\Phi, \Psi_H) + \Delta m(\mathcal{I}_H\Phi, \Psi_H)] \\ &\leq \max_{\|\Psi_H\|_{V_H}=1} \|\Psi_H\|_X \|(\text{id} - \mathcal{I}_H)\Phi\|_X + \max_{\|\Psi_H\|_{V_H}=1} |\Delta m(\mathcal{I}_H\Phi, \Psi_H)| \\ &\leq \frac{\sqrt{C_M}}{\sqrt{\alpha}} \|(\text{id} - \mathcal{I}_H)\Phi\|_X + \max_{\|\Psi_H\|_{V_H}=1} |\Delta m(\mathcal{I}_H\Phi, \Psi_H)|. \end{aligned}$$

□

We observe that this gives an estimate in terms of an interpolation error, which we can control, and a conformity error of the bilinear forms m^{eff} and m^{HMM} . We find a similar structure in the remaining expressions. As the next step we treat the error $\|(\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1}\mathbf{A} - (\mathbf{M}^{\text{HMM}})^{-1}\mathbf{A}_H\mathcal{I}_H)\mathbf{u}^{\text{eff}}\|_{L^\infty(0,t;V_H)}$, which is the error between the continuous and discrete Maxwell operator. This result is also the prototype for the following ones.

Lemma 5.3.14. Let $\Phi \in Z$. Then there exists a constant $C > 0$ such that

$$\|(\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1}\mathbf{A} - (\mathbf{M}^{\text{HMM}})^{-1}\mathbf{A}_H\mathcal{I}_H)\Phi\|_{V_H} \leq C \|(\text{id} - \mathcal{I}_H)\Phi\|_{V_{\text{mac}}} + \max_{\|\Psi_H\|_{V_H}=1} |\Delta a(\mathcal{I}_H\Phi, \Psi_H)|,$$

where for Φ_H, Ψ_H we define

$$\Delta a(\Phi_H, \Psi_H) := a(\Phi_H, \Psi_H) - a_H(\Phi_H, \Psi_H). \quad (5.73)$$

Proof. For $\Psi_H \in V_H$ we get with (5.68), (4.64), (5.23), (4.66), and (5.13) together with the continuity of $a(\cdot, \cdot)$ the following estimate

$$\begin{aligned} &m^{\text{HMM}}((\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1}\mathbf{A} - (\mathbf{M}^{\text{HMM}})^{-1}\mathbf{A}_H\mathcal{I}_H)\Phi, \Psi_H) \\ &= m^{\text{HMM}}(\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1}\mathbf{A}\Phi, \Psi_H) - m^{\text{HMM}}((\mathbf{M}^{\text{HMM}})^{-1}\mathbf{A}_H\mathcal{I}_H\Phi, \Psi_H) \\ &= a(\Phi, \Psi_H) - a_H(\mathcal{I}_H\Phi, \Psi_H) = a((\text{id} - \mathcal{I}_H)\Phi, \Psi_H) + a(\mathcal{I}_H\Phi, \Psi_H) - a_H(\mathcal{I}_H\Phi, \Psi_H) \\ &\leq C \|(\text{id} - \mathcal{I}_H)\Phi\|_{V_{\text{mac}}} \|\Psi_H\|_X + \Delta a(\mathcal{I}_H\Phi, \Psi_H) \\ &\leq C \|(\text{id} - \mathcal{I}_H)\Phi\|_{V_{\text{mac}}} \frac{\sqrt{C_M}}{\sqrt{\alpha}} \|\Psi_H\|_{V_H} + \Delta a(\mathcal{I}_H\Phi, \Psi_H). \end{aligned}$$

This yields

$$\begin{aligned} & \left\| (\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{A} - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{A}_H \mathcal{I}_H) \Phi \right\|_{V_H} \\ &= \max_{\|\Psi_H\|_{V_H}=1} m^{\text{HMM}} \left((\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{A} - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{A}_H \mathcal{I}_H) \Phi, \Psi_H \right) \\ &\leq \frac{\sqrt{C_M}}{\sqrt{\alpha}} C \left\| (\text{id} - \mathcal{I}_H) \Phi \right\|_{V_{\text{mac}}} + \max_{\|\Psi_H\|_{V_H}=1} |\Delta a(\mathcal{I}_H \Phi, \Psi_H)|. \end{aligned}$$

□

For the expression $\left\| (\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{R}^{\text{eff}} - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{R}^{\text{HMM}} \mathcal{I}_H) \mathbf{u}^{\text{eff}} \right\|_{L^\infty(0,t;V_H)}$ we follow the same approach as in Lemma 5.3.14 since the structure is basically the same. We use that the parameter \mathbf{R}^{eff} is bounded, i.e., $\mathbf{R}^{\text{eff}} \in L^\infty(\Omega; \mathbb{R}^{n \times n})$. This leads to the following lemma.

Lemma 5.3.15. *Let $\Phi \in Z$. Then there is a constant $C > 0$ such that*

$$\begin{aligned} & \left\| (\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{R}^{\text{eff}} - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{R}^{\text{HMM}} \mathcal{I}_H) \Phi \right\|_{V_H} \\ &\leq C \left\| (\text{id} - \mathcal{I}_H) \Phi \right\|_X + \max_{\|\Psi_H\|_{V_H}=1} |\Delta r(\mathcal{I}_H \Phi, \Psi_H)|_{V_H}. \end{aligned}$$

Again for $\Phi_H, \Psi_H \in V_H$ use the notation

$$\Delta r(\Phi_H, \Psi_H) := r^{\text{eff}}(\Phi_H, \Psi_H) - r^{\text{HMM}}(\Phi_H, \Psi_H). \quad (5.74)$$

Proof. We follow the proof of Lemma 5.3.14, which yields

$$\left\| (\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{R}^{\text{eff}} - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{R}^{\text{HMM}} \mathcal{I}_H) \Phi \right\|_{V_H} = \max_{\|\Psi_H\|_{V_H}=1} \left[r^{\text{eff}}((\text{id} - \mathcal{I}_H) \Phi, \Psi_H) + \Delta r(\mathcal{I}_H \Phi, \Psi_H) \right].$$

The bound (4.72) for $\Phi \in Z$ and $\Psi_H \in V_H$

$$|r^{\text{eff}}(\Phi, \Psi_H)| \leq C \|\Phi\|_X \|\Psi_H\|_{V_H},$$

yields the result. □

Now we turn our focus to the time-dependent parameters and the corresponding forms. Let us start with a result concerning the convolution. In (5.69) the corresponding expression is

$$\sup_{s \in [0,t]} \left\| \int_0^s (\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{G}^{\text{eff}}(s-r) - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{G}^{\text{HMM}}(s-r) \mathcal{I}_H) \mathbf{u}(r) \, dr \right\|_{V_H}.$$

Although this expression is time dependent we recognize again the same structure as before. Since the convolution kernel is bounded, we again get the desired result.

Lemma 5.3.16. *For all $s \in [0, T]$ and $\Phi \in C^1([0, t]; Z)$ we have for a constant $C > 0$ the bound*

$$\begin{aligned} & \left\| \int_0^s (\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{G}^{\text{eff}}(s-r) - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{G}^{\text{HMM}}(s-r) \mathcal{I}_H) \Phi(r) \, dr \right\|_{V_H} \\ &\leq C \left\| (\text{id} - \mathcal{I}_H) \Phi \right\|_{L^\infty(0,s;X)} + \max_{\|\Psi_H\|_{V_H}=1} \left| \int_0^s \Delta g(s-r; \mathcal{I}_H \Phi(r), \Psi_H) \, dr \right|, \end{aligned}$$

where for $\Phi_H, \Psi_H \in V_H$ we define

$$\Delta g(t; \Phi_H, \Psi_H) := g^{\text{eff}}(t; \Phi_H, \Psi_H) - g^{\text{HMM}}(t; \Phi_H, \Psi_H). \quad (5.75)$$

Proof. With the same ideas as in Lemma 5.3.14 we get from (5.68), (4.67) and (5.36)

$$\begin{aligned}
& \left\| \int_0^s [\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{G}^{\text{eff}}(s-r) - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{G}^{\text{HMM}}(s-r) \mathcal{I}_H] \Phi(r) \, dr \right\|_{\mathbf{V}_H} \\
&= \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} m^{\text{HMM}} \left(\int_0^s [\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{G}^{\text{eff}}(s-r) - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{G}^{\text{HMM}}(s-r) \mathcal{I}_H] \Phi(r) \, dr, \Psi_H \right) \\
&= \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} \int_0^s [g^{\text{eff}}(s-r; \Phi(r), \Psi_H) - g^{\text{HMM}}(s-r; \mathcal{I}_H \Phi(r), \Psi_H)] \, dr \\
&= \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} \int_0^s g^{\text{eff}}(s-r; (\text{id} - \mathcal{I}_H) \Phi(r), \Psi_H) \, dr + \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} \int_0^s \Delta g(s-r; \mathcal{I}_H \Phi(r), \Psi_H) \, dr.
\end{aligned}$$

In (4.73) we saw that the bilinear form $g^{\text{eff}}(t; \Phi, \Psi_H)$ is bounded independent of t , i.e., there exists a constant $C > 0$ with

$$g^{\text{eff}}(t; \Phi, \Psi_H) \leq C \|\Phi\|_{L^2(\Omega; \mathbb{R}^n)} \|\Psi_H\|_{L^2(\Omega; \mathbb{R}^n)}.$$

This bound yields the final estimate. \square

The next lemma is an estimate for the error that stems from the extra source term. We again make use of the structure we already know where we take advantage of the boundedness again.

Lemma 5.3.17. *For all $s \in [0, T]$ and $\Phi \in \mathbf{Z}$, we have for a constant $C > 0$*

$$\begin{aligned}
& \left\| \mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{J}^{\text{eff}}(s) \mathbf{u}_0 - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{J}^{\text{HMM}}(s) \mathcal{I}_H \mathbf{u}_0 \right\|_{\mathbf{V}_H} \\
& \leq C \|(\text{id} - \mathcal{I}_H) \Phi\|_X + \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} |\Delta j(s; \mathcal{I}_H \Phi, \Psi_H)|,
\end{aligned}$$

where for $\Phi_H, \Psi_H \in \mathbf{V}_H$ we define

$$\Delta j(s; \Phi_H, \Psi_H) := (\mathbf{J}^{\text{eff}}(s) \Phi_H, \Psi_H) - (\mathbf{J}^{\text{HMM}}(s) \Phi_H, \Psi_H)_H. \quad (5.76)$$

Proof. With the boundedness of $\mathbf{J}^{\text{eff}}(s)$ independent of s from Lemma 4.4.19 we get

$$\begin{aligned}
& \left\| [\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{J}^{\text{eff}}(s) - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{J}^{\text{HMM}}(s) \mathcal{I}_H] \Phi \right\|_{\mathbf{V}_H} \\
&= \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} m^{\text{HMM}} ([\mathcal{P}_H(\mathbf{M}^{\text{eff}})^{-1} \mathbf{J}^{\text{eff}}(s) - (\mathbf{M}^{\text{HMM}})^{-1} \mathbf{J}^{\text{HMM}}(s) \mathcal{I}_H] \Phi, \Psi_H) \\
&= \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} [(\mathbf{J}^{\text{eff}}(s) \Phi, \Psi_H) - (\mathbf{J}^{\text{HMM}}(s) \mathcal{I}_H \Phi, \Psi_H)_H] \\
&= \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} (\mathbf{J}^{\text{eff}}(s) (\text{id} - \mathcal{I}_H) \Phi, \Psi_H) + \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} \Delta j(s; \mathcal{I}_H \Phi, \Psi_H) \\
&\leq \frac{1}{\sqrt{\alpha}} \|\mathbf{J}^{\text{eff}}(s)\|_{L^\infty(\Omega)} \|(\text{id} - \mathcal{I}_H) \Phi\|_X + \max_{\|\Psi_H\|_{\mathbf{V}_H}=1} |\Delta j(s; \mathcal{I}_H \Phi, \Psi_H)|.
\end{aligned}$$

\square

We insert the inequalities from Lemma 5.3.13 to Lemma 5.3.17 into the result from Theorem 5.3.12. Thus, (5.69) becomes

$$\begin{aligned}
& \|\mathbf{u}^{\text{HMM}}(t) - \mathbf{u}^{\text{eff}}(t)\|_X \\
& \leq \exp\left(\int_0^t \|\mathbf{G}^{\text{HMM}}\|_{L^1(0,s;L^\infty(\Omega;\mathbb{R}^{n \times n}))} ds\right) \left[\left(1 + \frac{1}{\alpha} \|\mathbf{J}^{\text{HMM}}\|_{L^1(0,t;L^\infty(\Omega;\mathbb{R}^{n \times n}))}\right) \|\mathbf{u}_{0,H} - \mathcal{I}_H \mathbf{u}_0\|_{V_H} \right. \\
& \quad + \frac{t}{\alpha} \|\mathbf{f}^{\text{HMM}} - \mathcal{P}_H \mathbf{f}\|_{L^\infty(0,t;V_H)} + \frac{t}{\alpha} \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} |\Delta r(\mathcal{I}_H \mathbf{u}^{\text{eff}}(s), \Psi_H)| \\
& \quad + \frac{t}{\alpha} C \|(\text{id} - \mathcal{I}_H) \mathbf{u}^{\text{eff}}\|_{L^\infty(0,t;V^{\text{mac}})} + \frac{t}{\alpha} \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} |\Delta a(\mathcal{I}_H \mathbf{u}^{\text{eff}}(s), \Psi_H)| \\
& \quad + \frac{t}{\alpha} \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} \left| \int_0^s \Delta g(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| \\
& \quad + \frac{t}{\alpha} \sup_{s \in [0,t]} C \|(\text{id} - \mathcal{I}_H) \mathbf{u}_0\|_X + \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} |\Delta \mathbf{J}^{\text{eff}}(s; \mathcal{I}_H \mathbf{u}_0, \Psi_H)| \\
& \quad + \frac{t}{\alpha} C \|(\text{id} - \mathcal{I}_H) \partial_t \mathbf{u}^{\text{eff}}\|_{L^\infty(0,t;X)} + \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} |\Delta m(\mathcal{I}_H \partial_t \mathbf{u}^{\text{eff}}(s), \Psi_H)| \Big] \\
& \quad + \|(\mathcal{I}_H - \text{id}) \mathbf{u}^{\text{eff}}(t)\|_X .
\end{aligned} \tag{5.77}$$

The estimate is now given in terms of interpolation and conformity errors plus errors in the data. Hence, we have to analyze the conformity errors, where $\Delta m(\mathcal{I}_H \partial_t \mathbf{u}^{\text{eff}}, \Psi_H)$ and $\Delta a(\mathcal{I}_H \mathbf{u}^{\text{eff}}, \Psi_H)$ have already been analyzed in (Hochbruck et al., 2019, Lemma 4.1). Note that in Hochbruck et al. (2019) the classical Maxwell system is analyzed without polarization. The approach is nevertheless analogous in this context for these two error contributions. We recall the result here.

Lemma 5.3.18. *Assume that the corrector corresponding to the parameter \mathbf{M} satisfies*

$$|w^{\mathbf{M}}(x_K^q, \cdot)|_{\mathbb{H}^{k+1}(Y^\eta(x_K^q))} \leq C \delta^{-k} \sqrt{|Y^\delta(x_K^q)|} \quad \text{for all } x_K^q ,$$

for a constant $C > 0$ and furthermore

$$\mathbf{M}^{\text{eff}}|_K \in W^{\ell+1,\infty}(K; \mathbb{R}^{n \times n}), \quad \|\mathbf{M}^{\text{eff}}\|_{W^{\ell+1,\infty}(K)} \leq C ,$$

for all $K \in \mathcal{T}_H$ with a different constant $C > 0$ independent of δ and H . Then, for all $\Phi \in Z$ and $\Psi_H \in V_H$ we get

$$|\Delta m(\Phi, \Psi_H)| \leq C \left(H^\ell + \left(\frac{h}{\delta}\right)^{2k} + e_{\text{mod}} \right) \|\Phi\|_{\mathbb{H}^{\ell+1}(\Omega; \mathbb{R}^n)} \|\Psi_H\|_{L^2(\Omega; \mathbb{R}^n)} , \tag{5.78}$$

with e_{mod} as in Theorem 5.2.5. Moreover, for all $\Phi_H, \Psi_H \in V_H$ we find

$$|\Delta a(\Phi_H, \Psi_H)| = 0 . \tag{5.79}$$

Proof. Let $\Phi \in Z$ and $\Psi_H \in V_H$. Recall the definition of Δm from (5.72) and note that we can extend m_H^{eff} and m^{HMM} to $Z \times Z$ due to the continuous embedding of Z into $C(\Omega)$. Now split Δm into the HMM and the quadrature error

$$|\Delta m(\Phi, \Psi_H)| \leq |\Delta m_{\text{quad}}(\Phi, \Psi_H)| + |\Delta m_{\text{HMM}}(\Phi, \Psi_H)| ,$$

where for $\Phi \in Z$ and $\Psi \in V_H$ we set

$$\begin{aligned}\Delta m_{\text{Quad}}(\Phi, \Psi_H) &= m^{\text{eff}}(\Phi, \Psi_H) - m_H^{\text{eff}}(\Phi, \Psi_H), \\ \Delta m_{\text{HMM}}(\Phi, \Psi_H) &= m_H^{\text{eff}}(\Phi, \Psi_H) - m^{\text{HMM}}(\Phi, \Psi_H).\end{aligned}$$

We use the definition of the effective and HMM parameters, which gives

$$\begin{aligned}|\Delta m_{\text{HMM}}(\Phi, \Psi_H)| &= \left| \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \gamma_K^q [\mathbf{M}^{\text{eff}}(x_K^q) - \mathbf{M}^{\text{HMM}}(x_K^q)] \Phi(x_K^q) \cdot \Psi_H(x_K^q) \right| \\ &\leq \sup_{K,q} \|\mathbf{M}^{\text{eff}}(x_K^q) - \mathbf{M}^{\text{HMM}}(x_K^q)\|_F \sum_K \left| \sum_q \gamma_K^q \Phi(x_K^q) \cdot \Psi_H(x_K^q) \right|.\end{aligned}$$

We again use a productive zero and (Ciarlet, 2002, Theorem 4.1.5) to derive

$$\begin{aligned}\sum_K \left| \sum_q \gamma_K^q \Phi(x_K^q) \cdot \Psi_H(x_K^q) \right| &\leq \sum_K \left[\left| \sum_q \gamma_K^q \Phi(x_K^q) \cdot \Psi_H(x_K^q) - \int_K \Phi(x) \cdot \Psi_H(x) dx \right| + \left| \int_K \Phi(x) \cdot \Psi_H(x) dx \right| \right] \\ &\leq \sum_K \left[CH^{\ell+1} \|\Phi\|_{H^{\ell+1}(K)} \|\Psi_H\|_{H^1(K)} + \|\Phi\|_{L^2(K)} \|\Psi_H\|_{L^2(K)} \right].\end{aligned}$$

With the inverse inequality (Ciarlet, 2002, Theorem 3.2.6), and $\|\Phi\|_{L^2(K)} \leq \|\Phi\|_{H^{\ell+1}(K)}$ we get

$$\begin{aligned}\sum_K \left| \sum_q \gamma_K^q \Phi(x_K^q) \cdot \Psi_H(x_K^q) \right| &\leq \sum_K \left[CH^\ell \|\Phi\|_{H^{\ell+1}(K)} \|\Psi_H\|_{L^2(K)} + \|\Phi\|_{H^{\ell+1}(K)} \|\Psi_H\|_{L^2(K)} \right] \\ &\leq (CH^\ell + 1) \|\Phi\|_{H^{\ell+1}(\Omega)} \|\Psi_H\|_{L^2(\Omega)},\end{aligned}$$

where we applied the Cauchy-Schwarz inequality to the discrete product over $K \in \mathcal{T}_H$ in the last step. The final bound for $\Delta m_{\text{HMM}}(\Phi, \Psi_H)$ now follows from (5.29) together with Lemma 5.2.2 and Corollary 5.2.4 as

$$|\Delta m_{\text{HMM}}(\Phi, \Psi_H)| \leq C \left(\left(\frac{h}{\delta} \right)^{2k} + e_{\text{mod}} \right) (CH^\ell + 1) \|\Phi\|_{H^{\ell+1}(\Omega)} \|\Psi_H\|_{L^2(\Omega)}.$$

For the remainder $\Delta m_{\text{Quad}}(\Phi, \Psi_H)$ we use the same techniques as above since the structure is the same, i.e., we have

$$\begin{aligned}|\Delta m_{\text{Quad}}(\Phi, \Psi_H)| &= \left| \int_{\Omega} \mathbf{M}^{\text{eff}}(x) \Phi(x) \cdot \Psi_H(x) dx - \sum_K \sum_q \gamma_K^q \mathbf{M}^{\text{eff}}(x_K^q) \Phi(x_K^q) \cdot \Psi_H(x_K^q) \right| \\ &\leq CH^{\ell+1} \sum_K \|\Phi\|_{H^{\ell+1}(K)} \|\Psi_H\|_{H^1(K)} \leq CH^\ell \sum_K \|\Phi\|_{H^{\ell+1}(K)} \|\Psi_H\|_{L^2(K)} \\ &\leq CH^\ell \|\Phi\|_{H^{\ell+1}(\Omega)} \|\Psi_H\|_{L^2(\Omega)}.\end{aligned}$$

Thus, we have shown (5.78). Now let $\Phi_H, \Psi_H \in V_H$. It holds with (4.66) and (5.13)

$$\begin{aligned}|\Delta a(\Phi_H, \Psi_H)| &= \left| \int_{\Omega} \mathbf{A} \Phi_H(x) \cdot \Psi_H(x) dx - \sum_{K \in \mathcal{T}_H} \sum_{q=1}^{Q_K} \gamma_K^q \mathbf{A} \Phi_H(x_K^q) \cdot \Psi_H(x_K^q) \right| \\ &\leq \sum_K \left| \int_K \mathbf{A} \Phi_H(x) \cdot \Psi_H(x) dx - \sum_q \gamma_K^q \mathbf{A} \Phi_H(x_K^q) \cdot \Psi_H(x_K^q) \right|.\end{aligned}$$

The assumption on the quadrature in (5.9) to be exact for polynomials in $\mathcal{Q}^{2\ell, 2\ell, 2\ell}$ yields (5.79). \square

As in Lemma 5.3.18 we also need bounds on $\Delta r(\cdot, \cdot)$, $\Delta g(t; \cdot, \cdot)$ and $\Delta j(t; \cdot, \cdot)$. Despite slightly different assumptions on the correctors related to these parameters, the bounds on these bilinear forms follow the same idea as those in Lemma 5.3.18. For the conformity error in r we get the following result.

Lemma 5.3.19. *Assume that the correctors w^M and $\bar{w}(0)$ satisfy*

$$|w_\ell^M(x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\eta(x_K^q))}, |\bar{w}_\ell(0, x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\eta(x_K^q))} \leq C\delta^{-k} \sqrt{|Y^\delta(x_K^q)|} \quad \text{for all } x_K^q,$$

and furthermore

$$\mathbf{R}^{\text{eff}}|_K \in \mathbf{W}^{\ell+1, \infty}(K; \mathbb{R}^{n \times n}), \quad \|\mathbf{R}^{\text{eff}}\|_{\mathbf{W}^{\ell+1, \infty}(K)} \leq C,$$

for all $K \in \mathcal{T}_H$ with a constant $C > 0$ independent of δ and H . Then, for all $\Phi \in \mathbf{Z}$ and $\Psi_H \in \mathbf{V}_H$ we get

$$|\Delta r(\Phi, \Psi_H)| \leq C \left(H^\ell + \left(\frac{h}{\delta} \right)^{2k} + e_{\text{mod}} \right) \|\Phi\|_{\mathbf{H}^{\ell+1}(\Omega; \mathbb{R}^n)} \|\Psi_H\|_{\mathbf{L}^2(\Omega; \mathbb{R}^n)}, \quad (5.80)$$

with e_{mod} as in Theorem 5.2.5.

For the time-dependent bilinear forms based on the weaker micro error estimate from Lemma 5.3.7 and 5.3.9 we get a similar result, which only gives first-order convergence. Apart from that the proofs are identical. Although we have not analyzed the modeling error for the time dependent parameters we keep track of it in the following results. We first give the result for the convolution kernel.

Lemma 5.3.20. *Assume that the correctors w^M and $\bar{w}(0)$ satisfy*

$$|w_\ell^M(x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\eta(x_K^q))}, |\bar{w}_\ell(0, x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\eta(x_K^q))} \leq C\delta^{-k} \sqrt{|Y^\delta(x_K^q)|} \quad \text{for all } x_K^q, t \in [0, T],$$

and furthermore

$$\mathbf{M}^{\text{eff}}|_K, \mathbf{R}^{\text{eff}}|_K \in \mathbf{W}^{\ell+1, \infty}(K; \mathbb{R}^{n \times n}), \quad \|\mathbf{M}^{\text{eff}}\|_{\mathbf{W}^{\ell+1, \infty}(K)}, \|\mathbf{R}^{\text{eff}}\|_{\mathbf{W}^{\ell+1, \infty}(K)} \leq C,$$

for all $K \in \mathcal{T}_H$ with a constant $C > 0$ independent of δ and H . Then, for all $t \in [0, T]$, $\Phi \in \mathbf{Z}$ and $\Psi_H \in \mathbf{V}_H$ we find

$$|\Delta g(t; \Phi, \Psi_H)| \leq \left(CH^\ell + C(1+t) \left(\left(\frac{h}{\delta} \right)^k + e_{\text{mod}} \right) \right) \|\Phi\|_{\mathbf{H}^{\ell+1}(\Omega; \mathbb{R}^n)} \|\Psi_H\|_{\mathbf{L}^2(\Omega; \mathbb{R}^n)}. \quad (5.81)$$

Proof. The proof follows the proof of Lemma 5.3.18 and uses the fact that we can bound the effective parameter independent of t as shown in Lemma 4.4.19 and the micro error estimate in Corollary 5.3.8. \square

As already mentioned the procedure for the extra source is the same as for the convolution kernel. We thus get the following result.

Lemma 5.3.21. *Assume that the correctors w^M and $w^0(0)$ satisfy*

$$|w_\ell^M(x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\eta(x_K^q))}, |w_\ell^0(0, x_K^q, \cdot)|_{\mathbf{H}^{k+1}(Y^\eta(x_K^q))} \leq C\delta^{-k} \sqrt{|Y^\delta(x_K^q)|} \quad \text{for all } x_K^q, t \in [0, T],$$

and furthermore

$$\mathbf{M}^{\text{eff}}|_K, \mathbf{R}^{\text{eff}}|_K \in W^{\ell+1,\infty}(K; \mathbb{R}^{n \times n}), \quad \|\mathbf{M}^{\text{eff}}\|_{W^{\ell+1,\infty}(K)}, \|\mathbf{R}^{\text{eff}}\|_{W^{\ell+1,\infty}(K)} \leq C,$$

for all $K \in \mathcal{T}_H$ with a constant $C > 0$ independent of δ and H . Then, for all $t \in [0, T]$, $\Phi \in \mathbf{Z}$ and $\Psi_H \in \mathbf{V}_H$ it holds

$$|\Delta j(t; \Phi, \Psi_H)| \leq \left(CH^\ell + C(1+t) \left(\left(\frac{h}{\delta} \right)^k + e_{\text{mod}} \right) \right) \|\Phi\|_{H^{\ell+1}(\Omega; \mathbb{R}^n)} \|\Psi_H\|_{L^2(\Omega; \mathbb{R}^n)}. \quad (5.82)$$

Proof. Again we apply the same proof as in Lemma 5.3.18 but this time with the use of Lemma 5.3.9 and the boundedness of $\mathbf{J}^{\text{eff}}(t)$. \square

These preliminary lemmas give us bounds for the conformity errors. We use those estimates together with the bounds for the interpolation errors, which follow from the property of the interpolation \mathcal{I}_H from Theorem 5.1.4, i.e.,

$$\begin{aligned} \|(\text{id} - \mathcal{I}_H)\mathbf{u}\|_{L^\infty(0,t; \mathbf{V}^{\text{mac}})} &\leq CH^\ell \|\mathbf{u}\|_{L^\infty(0,t; \mathbf{Z})}, \\ \|(\text{id} - \mathcal{I}_H)\partial_t \mathbf{u}\|_{L^\infty(0,t; \mathbf{X})} &\leq CH^\ell \|\partial_t \mathbf{u}\|_{L^\infty(0,t; \mathbf{Z})}. \end{aligned}$$

As it turns out we still need a result that states that the interpolation operator from Theorem 5.1.4 keeps its property also in the norm of the discrete space \mathbf{V}_H . Note that for $\phi \in \mathbf{Z}$ this is not a direct consequence of the norm equivalence (5.66) since $\phi - \mathcal{I}_H \phi$ is in general not an element of \mathbf{V}_H . The result is found in (Hochbruck et al., 2019, Lemma 4.4). As mentioned in the proof of Lemma 5.3.18 we may extend the definition of $\|\cdot\|_{\mathbf{V}_H}$ to \mathbf{Z} .

Lemma 5.3.22. *For $\phi \in \mathbf{Z}$ the following estimate holds*

$$\|\phi - \mathcal{I}_H \phi\|_{\mathbf{V}_H} \leq CH^\ell \|\phi\|_{H^{\ell+1}(\Omega; \mathbb{R}^3)}.$$

Now we have all results at hand that we need to give the final semi-discrete error bound.

Theorem 5.3.23. *Let \mathbf{u}^{HMM} and \mathbf{u}^{eff} be the solutions of (5.65) and (5.64) respectively and assume that $\mathbf{u}^{\text{eff}} \in C^1([0, T]; \mathbf{Z})$ and $\mathbf{u}^{\text{HMM}} \in C^0(0, T; \mathbf{V}_H)$. Then the error of the semi-discrete HMM-solution is bounded by*

$$\begin{aligned} &\|\mathbf{u}^{\text{HMM}}(t) - \mathbf{u}^{\text{eff}}(t)\|_{\mathbf{X}} \\ &\leq C \exp \left(\int_0^t \|\mathbf{G}^{\text{HMM}}\|_{L^1(0,s; L^\infty(\Omega; \mathbb{R}^{n \times n}))} \, ds \right) (1+t) \left[\|\mathbf{u}_{0,H} - \mathcal{I}_H \mathbf{u}_0\|_{\mathbf{V}_H} + \|\mathbf{f}^{\text{HMM}} - \mathcal{P}_H \mathbf{f}\|_{L^\infty(0,t; \mathbf{V}_H)} \right. \\ &\quad \left. + \left(H^\ell + \left(\frac{h}{\delta} \right)^{2k} + e_{\text{mod}} \right) \left[\|\mathbf{u}^{\text{eff}}\|_{L^\infty(0,t; \mathbf{Z})} + \|\partial_t \mathbf{u}^{\text{eff}}\|_{L^\infty(0,t; \mathbf{Z})} \right] \right. \\ &\quad \left. + (1+t) \left(H^\ell + t \left(\left(\frac{h}{\delta} \right)^k + e_{\text{mod}} \right) \right) \|\mathbf{u}^{\text{eff}}\|_{L^\infty(0,t; \mathbf{Z})} \right]. \end{aligned} \quad (5.83)$$

Proof. The starting point is the refined estimate (5.77). We use the interpolation error estimates in (5.77), which yields

$$\begin{aligned}
& \|\mathbf{u}^{\text{HMM}}(t) - \mathbf{u}^{\text{eff}}(t)\|_X \\
& \leq C \exp\left(\int_0^t \|\mathbf{G}^{\text{HMM}}\|_{L^1(0,s;L^\infty(\Omega;\mathbb{R}^{n \times n}))} ds\right) (1+t) \left[\|\mathbf{f}^{\text{HMM}} - \mathcal{P}_H \mathbf{f}\|_{L^\infty(0,t;V_H)} \right. \\
& \quad + \left(1 + \|\mathbf{J}^{\text{HMM}}\|_{L^\infty(0,t;L^\infty(\Omega;\mathbb{R}^{n \times n}))}\right) \|\mathbf{u}_{0,H} - \mathcal{I}_H \mathbf{u}_0\|_{V_H} + \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} |\Delta r(\mathcal{I}_H \mathbf{u}^{\text{eff}}(s), \Psi_H)| \\
& \quad + H^\ell |\mathbf{u}^{\text{eff}}|_{L^\infty(0,t;Z)} + \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} |\Delta a(\mathcal{I}_H \mathbf{u}^{\text{eff}}(s), \Psi_H)| \\
& \quad + \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} \left| \int_0^s \Delta g(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| \\
& \quad + H^\ell |\mathbf{u}_0|_Z + \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} |\Delta \mathbf{J}^{\text{eff}}(s; \mathcal{I}_H \mathbf{u}_0, \Psi_H)| \\
& \quad \left. + H^\ell |\partial_t \mathbf{u}^{\text{eff}}|_{L^\infty(0,t;Z)} + \sup_{s \in [0,t]} \max_{\|\Psi_H\|_{V_H}=1} |\Delta m(\mathcal{I}_H \partial_t \mathbf{u}^{\text{eff}}(s), \Psi_H)| \right]. \tag{5.84}
\end{aligned}$$

Next we use the Lemmas 5.3.18, 5.3.19, 5.3.20 and 5.3.21. However, this is not possible directly since the arguments $\mathcal{I}_H \mathbf{u}^{\text{eff}}$ and $\mathcal{I}_H \partial_t \mathbf{u}^{\text{eff}}$ are not in the space Z . The procedure to overcome this is similar for all the remaining expressions Δm , Δr , Δa , Δg and Δj . Therefore, we demonstrate it for the most complicated one. Recall the definition of Δg in (5.75). We introduce a productive zero and use the triangle inequality, to derive

$$\begin{aligned}
& \left| \int_0^s \Delta g(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| \leq \left| \int_0^s g^{\text{eff}}(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) dr - \int_0^s g^{\text{eff}}(s-r; \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| \\
& \quad + \left| \int_0^s g^{\text{eff}}(s-r; \mathbf{u}^{\text{eff}}(r), \Psi_H) dr - \int_0^s g^{\text{HMM}}(s-r; \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| \\
& \quad + \left| \int_0^s g^{\text{HMM}}(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) dr - \int_0^s g^{\text{HMM}}(s-r; \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| \\
& = \left| \int_0^s g^{\text{eff}}(s-r; (\mathcal{I}_H - \text{id}) \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| + \left| \int_0^s \Delta g(s-r; \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| \\
& \quad + \left| \int_0^s g^{\text{HMM}}(s-r; (\mathcal{I}_H - \text{id}) \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right|.
\end{aligned}$$

Now we use the boundedness of the bilinear forms g^{eff} and g^{HMM} given in (4.73) and Lemma 5.3.10, which yields

$$\begin{aligned}
& \left| \int_0^s \Delta g(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) dr \right| \leq \int_0^s C \|(\mathcal{I}_H - \text{id}) \mathbf{u}^{\text{eff}}(r)\|_X \|\Psi_H\|_X dr \\
& \quad + \int_0^s |\Delta g(s-r; \mathbf{u}^{\text{eff}}(r), \Psi_H)| dr + \int_0^s C \|(\mathcal{I}_H - \text{id}) \mathbf{u}^{\text{eff}}(r)\|_{V_H} \|\Psi_H\|_{V_H} dr.
\end{aligned}$$

With this, Lemma 5.3.22 and the properties of the interpolation in (5.8) we get

$$\begin{aligned} & \max_{\|\Psi_H\|_{V_H}=1} \left| \int_0^s \Delta g(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) \, dr \right| \leq \max_{\|\Psi_H\|_{V_H}=1} \int_0^s |\Delta g(s-r; \mathbf{u}^{\text{eff}}(r), \Psi_H)| \, dr \\ & \quad + \int_0^s C \sqrt{\frac{C_M}{\alpha}} \|(\mathcal{I}_H - \text{id}) \mathbf{u}^{\text{eff}}(r)\|_X \, dr + \int_0^s C \|(\mathcal{I}_H - \text{id}) \mathbf{u}^{\text{eff}}(r)\|_{V_H} \, dr \\ & \leq \max_{\|\Psi_H\|_{V_H}=1} \int_0^s |\Delta g(s-r; \mathbf{u}^{\text{eff}}(r), \Psi_H)| \, dr + \int_0^s C H^\ell |\mathbf{u}^{\text{eff}}(r)|_{H^{\ell+1}(\Omega)} \, dr. \end{aligned}$$

Here the result from Lemma 5.3.20 is applicable for the first expression. We find

$$\begin{aligned} & \max_{\|\Psi_H\|_{V_H}=1} \left| \int_0^s \Delta g(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) \, dr \right| \\ & \leq \int_0^s \left(C H^\ell + C(1+s-r) \left(\left(\frac{h}{\delta} \right)^k + e_{\text{mod}} \right) \right) \|\mathbf{u}^{\text{eff}}(r)\|_{H^{\ell+1}(\Omega; \mathbb{R}^n)} \, dr + \int_0^s C H^\ell |\mathbf{u}^{\text{eff}}(r)|_{H^{\ell+1}(\Omega)} \, dr, \end{aligned}$$

and finally

$$\begin{aligned} & \sup_{s \in [0, t]} \max_{\|\Psi_H\|_{V_H}=1} \left| \int_0^s \Delta g(s-r; \mathcal{I}_H \mathbf{u}^{\text{eff}}(r), \Psi_H) \, dr \right| \\ & \leq \int_0^t \left(C H^\ell + C(1+t-r) \left(\left(\frac{h}{\delta} \right)^k + e_{\text{mod}} \right) \right) \, dr \|\mathbf{u}^{\text{eff}}\|_{L^\infty(0, t; Z)}. \end{aligned}$$

With this bound and similar bounds for the other conformity errors we eventually get the result. \square

Let us comment on the semi-discrete error estimate (5.83). First as mentioned in Section 4.4.8 we can get better bounds for the stability estimate, which improves the growth rate. If we have a bound without exponential growth, this also transfers to the error estimate. Moreover, a refined bound on the growth of the H^2 -norm of the solution of the Sobolev equation would also improve the bound.

At the end of this section let us summarize what we achieved so far. Starting from the general heterogeneous Maxwell system (4.11) we used the method of homogenization in Chapter 4 to derive an effective system. After that in this chapter we derived a space discretization for this system. Moreover, we analyzed the error due to the finite element methods on the macroscopic and microscopic level. Still, the system (5.37) is continuous in the time variable. Thus, in the next chapter we propose a time discretization.

Time discretization and approximation of the convolution

In the previous chapters we derived the space discrete effective Maxwell system using the Finite Element Heterogeneous Multiscale Method. The space discrete system resulting from this method is given in (5.37). In the proof of Theorem 5.3.11 and in Section 5.1 we derived the integro-differential matrix equation

$$\mathfrak{M}^H \partial_t \mathbf{U}(t) + \mathfrak{R}^H \mathbf{U}(t) + \int_0^t \mathfrak{G}^H(t-s) \mathbf{U}(s) ds + \mathfrak{A} \mathbf{U}(t) = \mathfrak{g}^H(t) - \mathfrak{J}^H(t) \mathbf{U}(0), \quad (6.1)$$

which is an equivalent formulation of (5.37). The aim of this section is to derive suitable time integration schemes that numerically solve the equation (6.1). Let us comment on results related to the FE-HMM for Maxwell's equations as well as the time integration of integro-differential Maxwell systems.

As presented in Section 5.2.2, the Heterogeneous Multiscale Method has been analyzed in Hochbruck et al. (2019) where the authors also considered the time integration of the resulting Maxwell system (5.32). Although this system does not inherit any memory effects, the fully discrete error has been analyzed including the error of the HMM. On the other hand integro-differential Maxwell systems with macroscopic polarization or magnetization have been analyzed. The Debye model from Section 3.1.3 is an example of this class. In Li (2007) the authors studied the time discretization of these models and analyzed the resulting method. To our knowledge there are no results concerning the combination of the HMM for dispersive media and the time discretizations for the resulting integro-differential system. In this thesis we do not close this gap but the combination of the references Hochbruck et al. (2019) and Li (2007) should yield an error estimate for the time discretization of the FE-HMM.

The structure of this chapter is as follows. We start by briefly introducing standard time integration schemes for first-order ordinary differential equations in the next section. The main part of this chapter is dedicated to the approximation of the convolution integral. There are different possibilities one may choose for this approximation. We decided to use the approach of recursive convolution first proposed in Luebbers et al. (1990), which seems to suit the structure of the convolution kernels very well. This

method is explained in Section 6.2 where we also discuss why we use this scheme. Finally, in Section 6.3 we present our fully discrete scheme and show some error estimates for special cases of convolution kernels.

Throughout this chapter, in all quantities, we drop the notation indicating a dependence on the space discretization, i.e., the macroscopic solution \mathbf{u}^{HMM} is denoted as \mathbf{u} and micro solutions $w_\ell^{\text{M},h}$ are denoted as w_ℓ^{M} .

6.1 Time integration of first-order ordinary differential equations

The theory in this section is standard and may be found in Hairer and Wanner (1996); Hairer et al. (1993, 2006). Before we start to consider the convolution integral in the next section, we give a brief overview of time integration of ordinary differential equations of the form

$$\mathfrak{M}\partial_t\mathbf{u}(t) + \mathfrak{A}\mathbf{u}(t) = \mathfrak{f}(t). \quad (6.2)$$

For that purpose we introduce a uniform subdivision of $[0, T]$ in N_T sub-intervals. Thus, the time step is $\Delta t = \frac{T}{N_T}$ and for $n = 0, \dots, N_T$ (not to be confused with the dimension of the Maxwell system) we use the notation

$$t_n = n\Delta t, \quad \mathbf{u}^n = \mathbf{u}(t_n).$$

We thus have

$$t_0 = 0, \quad t_{N_T} = T.$$

There are many possible ways to discretize the equation (6.2) for example Runge–Kutta methods, exponential integrators, Krylov subspace methods and many more. The time integrators considered in this work all fit in the class of Runge–Kutta schemes.

In general, we are interested in successively computing solutions starting from the initial value $\mathbf{u}^0 = \mathbf{u}(0)$. The schemes we use in this work are the explicit and implicit Euler method as well as the explicit Heun’s method and the implicit Crank–Nicolson scheme. The last two are second order methods whereas the first two have convergence order one. Let us recall how Runge–Kutta schemes are derived. We start by the reformulation of (6.2) as

$$\partial_t\mathbf{u}(t) = \mathfrak{M}^{-1}(\mathfrak{f}(t) - \mathfrak{A}\mathbf{u}(t)).$$

We find that

$$\mathbf{u}(t_n + \Delta t) = \mathbf{u}(t_n) + \int_0^{\Delta t} \partial_t\mathbf{u}(t_n + s) \, ds = \mathbf{u}(t_n) + \int_0^{\Delta t} \mathfrak{M}^{-1}(\mathfrak{f}(t_n + s) - \mathfrak{A}\mathbf{u}(t_n + s)) \, ds, \quad (6.3)$$

and apply a quadrature rule to the integral. Depending on the choice of the quadrature we get various schemes, where we only recapture four classical representatives in the rest of this thesis.

6.1.1 Explicit schemes

We start by repeating the explicit time integrators. These schemes have restricted stability regions, which yields step size restrictions. For hyperbolic systems, as the macroscopic Maxwell system, these are known as CFL conditions of the form

$$\Delta t \leq CH.$$

This step size restriction makes the explicit methods unattractive for stiff ODEs, such as the discretization of the Maxwell system. For parabolic problems the step size restriction is given as

$$\Delta t \leq Ch^2,$$

but we point out that the microscopic problems are not stiff and thus the application of explicit scheme to the time dependent cell problems is a cheap and fast alternative.

Explicit Euler

The easiest method is the explicit Euler scheme, which originates from (6.3) by the use of the left rectangular rule. The new solution \mathbf{u}^{n+1} is then computed by

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t \mathfrak{M}^{-1} (\mathfrak{f}^n - \mathfrak{A} \mathbf{u}^n).$$

The explicit Euler scheme is of first order.

Heun's method

The other explicit method we consider is Heun's method. In (6.3) first use the trapezoidal rule. This yields

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{2} \mathfrak{M}^{-1} (\mathfrak{f}^{n+1} - \mathfrak{A} \mathbf{u}^{n+1} + \mathfrak{f}^n - \mathfrak{A} \mathbf{u}^n),$$

and we use an explicit Euler step to approximate \mathbf{u}^{n+1} . Thus, the complete 2 step scheme is given as

$$\tilde{\mathbf{u}}^{n+1} = \mathbf{u}^n + \Delta t \mathfrak{M}^{-1} (\mathfrak{f}^n - \mathfrak{A} \mathbf{u}^n), \tag{6.4a}$$

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{2} \mathfrak{M}^{-1} (\mathfrak{f}^n + \mathfrak{f}^{n+1} - \mathfrak{A} (\mathbf{u}^n + \tilde{\mathbf{u}}^{n+1})). \tag{6.4b}$$

Heun's method is of second order with a slightly larger stability region but still we get a step size restriction condition.

These two explicit schemes have been analyzed for the time integration of the Sobolev equation in [Bekkouche et al. \(2019\)](#). As counterparts to the explicit schemes, which lack in stability, we recapture two implicit schemes in the next section. These are used for the macroscopic time integration.

6.1.2 Implicit schemes

The explicit schemes from the previous section may have severe step size restrictions. An alternative approach that yields stable methods are implicit time integration schemes. The drawback of these methods is that the inversion of the stiffness matrix \mathfrak{A} is necessary, which causes higher computational effort.

Implicit Euler

Let us first recall the implicit Euler method, i.e.,

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t \mathfrak{M}^{-1} (\mathfrak{f}^{n+1} - \mathfrak{A} \mathbf{u}^{n+1}),$$

which is equivalent to

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t (\mathfrak{M} + \Delta t \mathfrak{A})^{-1} (\mathfrak{f}^{n+1} - \mathfrak{A} \mathbf{u}^n).$$

This method is unconditionally stable and as the explicit Euler scheme of first order. It originates from (6.3) by the use of a right rectangular rule.

Crank-Nicolson

We also consider the Crank–Nicolson scheme, which results from the trapezoidal rule as

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{2} \mathfrak{M}^{-1} (\mathfrak{f}^n - \mathfrak{A} \mathbf{u}^n + \mathfrak{f}^{n+1} - \mathfrak{A} \mathbf{u}^{n+1}).$$

In contrast to the explicit Heun scheme we keep the \mathbf{u}^{n+1} on the right-hand side. This can be written as

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{2} \left(\mathfrak{M} + \frac{\Delta t}{2} \mathfrak{A} \right)^{-1} (\mathfrak{f}^n + \mathfrak{f}^{n+1} - 2\mathfrak{A} \mathbf{u}^n).$$

The Crank–Nicolson schemes is a second order method which is unconditionally stable. The macroscopic Maxwell system results in a stiff ODE. Thus, we prefer the use of an implicit scheme for its time discretization. This choice is supported by [Hochbruck et al. \(2015b\)](#) where the authors show that implicit schemes can even outperform explicit methods for Maxwell’s equations.

After this short repetition of time integration, we turn our attention to the problems caused by the convolution.

6.1.3 Problems for integro-differential equations

Let us point out that in general Runge–Kutta schemes are derived for ODEs of the form

$$\partial_t \mathbf{u}(t) = \mathfrak{f}(t, \mathbf{u}(t)).$$

Here the right-hand side \mathfrak{f} depends on the time t and the solution $\mathbf{u}(t)$ at that time. Thus, the system (6.1) does not fit in this setting since the whole history of the solution is present in the convolution integral. This is the reason why we have to derive a suitable time integration scheme for the integro-differential system. The starting point for this derivation is again (6.3) but now including the convolution, i.e.,

$$\mathbf{u}(t_n + \Delta t) = \mathbf{u}(t_n) + \int_0^{\Delta t} \left(\mathfrak{f}(t_n + s, \mathbf{u}(t_n + s)) + \int_0^{t_n + s} \mathfrak{G}(t_n + s - r) \mathbf{u}(r) dr \right) ds. \quad (6.5)$$

The question is how to approximate the convolution integral. Therefore, the next section is concerned with the approximation of the convolution.

6.2 Approximating the convolution

The most challenging part in the equation (6.1) and thus in (6.5) is the convolution. The direct approach for its approximation is to use a quadrature for the convolution integral. We briefly show this method in the next section. Moreover, we highlight why it is not feasible to use this brute-force quadrature. The approximation we choose relies on observations on the structure of the convolution kernel, which we exploit. Therefore, in Section 6.2.2 we introduce the method of recursive convolution where we assume that the convolution kernel is a decaying exponential. This method is not only reducing the amount of solutions we have to store but also the convolution kernel itself is only required to be known for less time steps. In Section 6.2.3, we comment on the use of memory variables, which again uses an exponential structure of the convolution kernel. In the context of convolution approximation the method of convolution quadrature proposed in Lubich (1988) is widely used. Nevertheless, to our knowledge it has not been used in the context of dispersive Maxwell systems in time-domain.

6.2.1 Numerical quadrature

The easiest way to approximate a convolution integral is to use a brute-force quadrature. To be consistent with the time discretization of the ODE we choose the same quadrature as for the respective Runge–Kutta scheme. This is useful since the solution is then evaluated at the same steps as for the time integration scheme. Assume that we use either the implicit Euler scheme or the Crank–Nicolson method. A generalization to other Runge–Kutta schemes is straight forward as it relies on the respective quadrature formula of the Runge–Kutta scheme.

We approximate the convolution integral at a time step t_n by splitting it in n sub-intervals and using the trapezoidal rule

$$\int_0^{t_n} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds \approx \sum_{k=0}^{n-1} \frac{\Delta t}{2} (\mathfrak{G}(t_n - t_{k+1}) \mathbf{u}(t_{k+1}) + \mathfrak{G}(t_n - t_k) \mathbf{u}(t_k)) .$$

The drawback of this approach is that we have to store all previous solutions to evaluate the convolution. Moreover, the convolution kernel has to be evaluated for all $t_n - t_k$, $k = 0, \dots, n$. Together this yields a tremendous amount of storage that is used for the evaluation of the convolution. Nevertheless, the analysis of this approximation is probably straight forward since both, the Runge–Kutta schemes and quadrature formulas, are understood very well. Over that, this method is applicable for every convolution kernel.

Still, this approach is not feasible for implementation and therefore a different approach is considered in the next section.

6.2.2 Recursive convolution

Instead of the brute-force quadrature we choose to use recursive quadrature. Let us first comment on the idea why we chose to use this approach. There are basically two reasons. On the one hand one observes in classic models such as the Drude or the Debye model that the convolution kernel is a decaying exponential. On the other hand we have the definition (4.26) of the effective parameter and its time evolution is given by the correctors $\bar{w}_\ell(t)$. Thus, these solutions of the cell problems (4.33) determine the time evolution of the convolution kernel. We saw in Section 4.4.5 that the convolution kernel is always bounded by

its values at time $t = 0$. Moreover, in Section 4.4.7 we showed that there are constellations where the convolution kernel is actually exponentially decaying. Thus, we think that the approximation by a linear combination of exponential functions may yield a good approximation of the convolution kernel. In Section 7.2.1 we show the decay property of some examples of micro structures and there we observe again that the approximation via exponentials is very accurate. But here we also hit the point that we did not yet consider, the quality of the approximation via exponential functions. The problem is that although we showed that the general solution of the Sobolev equation is bounded we only showed the decaying property for the special case where the damping parameter is strictly positive. This restricts us in the error estimate since we do not have a convergence result for the approximation with exponential functions. For the case where the convolution kernel is in fact an exponential, Li (2007) showed that the recursive convolution leads to suitable convergence results. We now present the recursive convolution which is found in Banks et al. (2006), Kelley and Luebbers (1996), and Luebbers et al. (1990).

The main assumption is that the convolution kernel is a matrix exponential, i.e., for full rank matrices $\mathfrak{G}_0, \mathfrak{G}_1 \in \mathbb{R}^{N_{\text{VH}} \times N_{\text{VH}}}$ we have

$$\mathfrak{G}(t) = \mathfrak{G}_0 \exp(-\mathfrak{G}_1 t) . \quad (6.6)$$

The key observation is that for $s \in [0, t_n]$ we get

$$\begin{aligned} \mathfrak{G}(t_n - s) &= \mathfrak{G}_0 \exp(-\mathfrak{G}_1(t_n - s)) = \mathfrak{G}_0 \exp(-\mathfrak{G}_1(t_n - t_{n-1} + t_{n-1} - s)) \\ &= \mathfrak{G}_0 \exp(-\mathfrak{G}_1(t_n - t_{n-1})) \exp(-\mathfrak{G}_1(t_{n-1} - s)) = \mathfrak{G}_0 \exp(-\mathfrak{G}_1 \Delta t) \exp(-\mathfrak{G}_1(t_{n-1} - s)) \\ &= \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{G}(0) \exp(-\mathfrak{G}_1(t_{n-1} - s)) = \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{G}(t_{n-1} - s) . \end{aligned} \quad (6.7)$$

We denote the convolution at time t_n with \mathfrak{J}^n and using equality (6.7) we find

$$\begin{aligned} \mathfrak{J}^n &= \int_0^{t_n} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds = \int_0^{t_{n-1}} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds + \int_{t_{n-1}}^{t_n} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds \\ &= \int_0^{t_{n-1}} \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{G}(t_{n-1} - s) \mathbf{u}(s) \, ds + \int_{t_{n-1}}^{t_n} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds \\ &= \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{J}^{n-1} + \int_{t_{n-1}}^{t_n} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds . \end{aligned} \quad (6.8)$$

In equation (6.8) we see the recursive structure. Thus, for the convolution at time step t_n we only need the convolution at the previous time step t_{n-1} as well as the solution and convolution kernel that are needed for the evaluation of the remainder integral

$$\int_{t_{n-1}}^{t_n} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds .$$

The approximation of this integral can be done by classic quadrature rules as in Section 6.2.1. At this point there is no problem with storage since this integral is only evaluated in the time interval $[t_{n-1}, t_n]$. If we choose for example the trapezoidal rule as in Siushansian and LoVetri (1997) we get the approximation

as

$$\int_{t_{n-1}}^{t_n} \mathfrak{G}(t_n - s) \mathbf{u}(s) \, ds \approx \frac{\Delta t}{2} (\mathfrak{G}(0) \mathbf{u}(t_n) + \mathfrak{G}(\Delta t) \mathbf{u}(t_{n-1})) .$$

Here we see that the solution is only required for the current and previous time step whereas the convolution kernel is still only necessary for $t = 0$ and $t = \Delta t$. This approximation yields the following second order recursive definition of the convolution

$$\mathfrak{J}^0 = \mathbf{0} , \tag{6.9}$$

$$\mathfrak{J}^n = \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{J}^{n-1} + \frac{\Delta t}{2} (\mathfrak{G}(0) \mathbf{u}^n + \mathfrak{G}(\Delta t) \mathbf{u}^{n-1}) . \tag{6.10}$$

Of course the choice of the quadrature for the remainder integral is not restricted to the trapezoidal rule. But one should take care that the quadrature has the same order as the macroscopic time integration scheme. In the case of a first order method the natural choice is a rectangular quadrature rule whereas in the case of a second order scheme such as the Heun or the Crank–Nicolson method the right choice is a quadrature rule that also is of second order, i.e., trapezoidal or midpoint rule. Thus for higher order time integration schemes also higher order quadrature is mandatory to get the right order of convergence for the time integration. With the recursive definition of the convolution in (6.8) and the approximation in equation (6.10) we have a promising tool to implement an efficient time integration scheme for the macroscopic Maxwell system.

Let us point out that the recursive convolution approach for first or second order time integration schemes only needs knowledge of the convolution kernel for the two time points 0 and Δt . Since the time dependence of the effective convolution kernel is given by the correctors \bar{w}_ℓ this means that we only have to solve the microscopic problems for the time interval $[0, \Delta t]$ and not for the whole macroscopic time interval $[0, T]$.

We now comment on the combination of the time integration schemes from Section 6.1 and the recursive convolution. We consider the general system

$$\partial_t \mathbf{u}(t) = \mathfrak{M}^{-1} \left(\mathfrak{f}(t) - \mathfrak{A} \mathbf{u}(t) - \int_0^t \mathfrak{G}(t-r) \mathbf{u}(r) \, dr \right) .$$

From the derivation of Runge–Kutta schemes we get

$$\mathbf{u}(t_n + \Delta t) = \mathbf{u}(t_n) + \int_0^{\Delta t} \mathfrak{M}^{-1} \left(\mathfrak{f}(t_n + s) - \mathfrak{A} \mathbf{u}(t_n + s) - \int_0^{t_n+s} \mathfrak{G}(t_n + s - r) \mathbf{u}(r) \, dr \right) \, ds .$$

Let us demonstrate how we derive the recursive Crank–Nicolson scheme. We choose the trapezoidal rule for the outer integral, which yields

$$\begin{aligned} \mathbf{u}(t_n + \Delta t) = \mathbf{u}(t_n) + \frac{\Delta t}{2} \mathfrak{M}^{-1} & \left(\mathfrak{f}(t_n) - \mathfrak{A} \mathbf{u}(t_n) - \int_0^{t_n} \mathfrak{G}(t_n - r) \mathbf{u}(r) \, dr \right. \\ & \left. + \mathfrak{f}(t_n + \Delta t) - \mathfrak{A} \mathbf{u}(t_n + \Delta t) - \int_0^{t_n+\Delta t} \mathfrak{G}(t_n + \Delta t - r) \mathbf{u}(r) \, dr \right) . \end{aligned}$$

Now using the recursive definition of the convolution from (6.10) again with a trapezoidal rule yields the

discrete scheme

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{2} \mathfrak{M}^{-1} (\mathfrak{f}^n + \mathfrak{f}^{n+1} - \mathfrak{A}(\mathbf{u}^n + \mathbf{u}^{n+1}) - \mathfrak{J}^n - \mathfrak{J}^{n+1}), \quad (6.11a)$$

$$\mathfrak{J}^{n+1} = \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{J}^n + \frac{\Delta t}{2} (\mathfrak{G}(0) \mathbf{u}^{n+1} + \mathfrak{G}(\Delta t) \mathbf{u}^n). \quad (6.11b)$$

For the implicit Euler method we proceed in the same way as for the Crank–Nicolson scheme, with the only difference that we use the right rectangular rule. This yields

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t \mathfrak{M}^{-1} (\mathfrak{f}^{n+1} - \mathfrak{A} \mathbf{u}^{n+1} - \mathfrak{J}^{n+1}), \quad (6.12a)$$

$$\mathfrak{J}^{n+1} = \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{J}^n + \Delta t \mathfrak{G}(0) \mathbf{u}^{n+1}. \quad (6.12b)$$

In general we suggest using the same quadrature rule for the approximation of the convolution as for the macroscopic time integration.

In the next section we briefly comment on the use of memory variables instead of recursive convolution. After that in Section 6.3, we present the fully discrete scheme using the recursive convolution presented in this section.

6.2.3 Memory variables

This technique may also be known as the auxiliary differential equation (ADE) method. The assumption (6.6) that the convolution kernel is an exponential could be used in another way. In

$$\mathfrak{M} \partial_t \mathbf{u}(t) + \mathfrak{A} \mathbf{u}(t) + \int_0^t \mathfrak{G}(t-s) \mathbf{u}(s) ds = \mathfrak{f}(t),$$

we replace the convolution integral by a memory variable

$$\mathfrak{J}(t) := \int_0^t \mathfrak{G}(t-s) \mathbf{u}(s) ds.$$

This yields

$$\mathfrak{M} \partial_t \mathbf{u}(t) + \mathfrak{A} \mathbf{u}(t) + \mathfrak{J}(t) = \mathfrak{f}(t). \quad (6.13)$$

The naming is no coincidence. Evaluating the memory variable at a time t_n should yield the same as the convolution \mathfrak{J}^n .

Now consider the time derivative of the convolution using the exponential structure

$$\begin{aligned} \partial_t \mathfrak{J}(t) &= \partial_t \int_0^t \mathfrak{G}(t-s) \mathbf{u}(s) ds = \int_0^t \partial_t \mathfrak{G}(t-s) \mathbf{u}(s) ds + \mathfrak{G}(0) \mathbf{u}(t) \\ &= - \int_0^t \mathfrak{G}_0 \mathfrak{G}_1 \exp(-\mathfrak{G}_1(t-s)) \mathbf{u}(s) ds + \mathfrak{G}(0) \mathbf{u}(t) \\ &= -\mathfrak{G}_0 \mathfrak{G}_1 \mathfrak{G}_0^{-1} \int_0^t \mathfrak{G}_0 \exp(-\mathfrak{G}_1(t-s)) \mathbf{u}(s) ds + \mathfrak{G}_0 \mathbf{u}(t) \\ &= -\mathfrak{G}_0 \mathfrak{G}_1 \mathfrak{G}_0^{-1} \int_0^t \mathfrak{G}(t-s) \mathbf{u}(s) ds + \mathfrak{G}_0 \mathbf{u}(t). \end{aligned}$$

This is a differential equation for the convolution

$$\partial_t \mathcal{J}(t) + \mathfrak{G}_0 \mathfrak{G}_1 \mathfrak{G}_0^{-1} \mathcal{J}(t) = \mathfrak{G}_0 \mathbf{u}(t). \quad (6.14)$$

Combining the equations (6.13) and (6.14) yields

$$\begin{pmatrix} \mathfrak{M} & \\ & \mathfrak{G}_0^{-1} \end{pmatrix} \partial_t \begin{pmatrix} \mathbf{u}(t) \\ \mathcal{J}(t) \end{pmatrix} + \begin{pmatrix} \mathfrak{A} & \mathbf{I}_{N_{\text{vH}}} \\ -\mathbf{I}_{N_{\text{vH}}} & \mathfrak{G}_1 \mathfrak{G}_0^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{u}(t) \\ \mathcal{J}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{f}(t) \\ \mathbf{0}_{N_{\text{vH}}} \end{pmatrix}. \quad (6.15)$$

For this system we can directly use a standard time integration scheme from Section 6.1. The implicit Euler scheme applied to (6.15) reads

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t \mathfrak{M}^{-1} (\mathbf{f}^{n+1} - \mathfrak{A} \mathbf{u}^{n+1} - \mathcal{J}^{n+1}), \quad (6.16a)$$

$$\mathcal{J}^{n+1} = \mathcal{J}^n + \Delta t \mathfrak{G}_0 (\mathbf{u}^{n+1} - \mathfrak{G}_1 \mathfrak{G}_0^{-1} \mathcal{J}^{n+1}). \quad (6.16b)$$

In comparison to (6.12), the evolution of the memory variable or the convolution, respectively, is different. To be more precise in the recursive convolution approach we need the value $\mathfrak{G}(\Delta t)$ of the kernel. In contrast, for the memory variable method we need to know the growth rate \mathfrak{G}_1 .

For this simple method one may use the Taylor expansion for the exponential (6.6) to derive

$$\mathfrak{G}(\Delta t)^{-1} = \mathfrak{G}_0^{-1} + \Delta t \mathfrak{G}_1 \mathfrak{G}_0^{-1} + \mathcal{O}(\Delta t^2).$$

Inserting this representation in (6.12b) and neglecting higher order terms one can show the relation between the two update formulas (6.12b) and (6.16b).

The next section is dedicated to the fully discrete scheme for the effective Maxwell system. We use the recursive convolution presented in Section 6.2.2 combined with a time integration scheme from Section 6.1.

6.3 The fully discrete scheme and its error analysis

In Chapter 5 the space discretization of the system (4.32) has been derived. In the first parts of this chapter we showed classic time integration schemes as well as some techniques for the approximation of the convolution. In this section we combine all these methods, which eventually yields the fully discrete scheme. The starting point is again the space discrete system (6.1). In the next section we show our approach for the time integration and in Section 6.3.2 we present the fully discrete error analysis.

6.3.1 Recursive FE-HMM scheme for approximation of the effective solution

Consider the semi-discrete system (6.1) where we already applied our FE-HMM from Chapter 5.1 to the macroscopic Maxwell system in space

$$\mathfrak{M} \partial_t \mathbf{U}(t) + \mathfrak{R} \mathbf{U}(t) + \int_0^t \mathfrak{G}(t-s) \mathbf{U}(s) \, ds + \mathfrak{A} \mathbf{U}(t) = \mathbf{g}(t) - \mathcal{J}(t) \mathbf{U}(0). \quad (6.17)$$

For convenience we introduce a right-hand side \mathbf{f} as

$$\mathbf{f}(t) = \mathbf{g}(t) - \mathcal{J}(t) \mathbf{U}(0).$$

We comment on the evaluation of \mathfrak{f} later. Let us now focus on the convolution. As pointed out in Section 6.2 we assume that the convolution kernel is an exponential. The following procedure is similar to the one in Chapter 5 in the sense that we first discretize on the macroscopic level and then on the microscopic level.

Macroscopic time stepping

On the macro level we have a stiff ODE due to the Maxwell operator. Thus, we use either the recursive implicit Euler or the recursive Crank–Nicolson scheme for time discretization. Both schemes have been defined in Section 6.2.2.

Let us first present the implicit Euler scheme for the macroscopic Maxwell system

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t \mathfrak{M}^{-1} (\mathfrak{f}^{n+1} - \mathfrak{A} \mathbf{u}^{n+1} - \mathfrak{R} \mathbf{u}^{n+1} - \mathfrak{J}^{n+1}), \quad (6.18a)$$

$$\mathfrak{J}^{n+1} = \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{J}^n + \Delta t \mathfrak{G}(0) \mathbf{u}^{n+1}. \quad (6.18b)$$

The Crank–Nicolson method is derived as in (6.11)

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{2} \mathfrak{M}^{-1} (\mathfrak{f}^{n+1} + \mathfrak{f}^n - \mathfrak{A} \mathbf{u}^{n+1} - \mathfrak{A} \mathbf{u}^n - \mathfrak{R} \mathbf{u}^{n+1} - \mathfrak{R} \mathbf{u}^n - \mathfrak{J}^{n+1} - \mathfrak{J}^n), \quad (6.19a)$$

$$\mathfrak{J}^{n+1} = \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{J}^n + \frac{\Delta t}{2} (\mathfrak{G}(0) \mathbf{u}^{n+1} + \mathfrak{G}(\Delta t) \mathbf{u}^n). \quad (6.19b)$$

Let us highlight the properties of the algorithms (6.18) and (6.19). The matrices \mathfrak{M} , \mathfrak{R} and $\mathfrak{G}(0)$ are computed using the HMM. Thus, their computation relies on the solution of cell problems on sampling domains. Those are solved a priori such that all quantities except $\mathfrak{G}(\Delta t)$ are initially known for the time integration. This only missing expression may also be computed a priori. We use a different time stepping scheme on the micro level to evolve the cell correctors to Δt . With these solutions we then compute the convolution kernels $\mathfrak{G}(\Delta t)$.

Microscopic time stepping

For the macroscopic implicit Euler and Crank–Nicolson method we need the convolution kernel at Δt . As pointed out, the time dependence of the convolution kernel stems from the solution of the Sobolev equation. The space discretization in Section 5.3.2 is the starting point for our time stepping scheme. The space discrete formulation is given in (5.50). We applied a conforming finite element method with Lagrange elements introduced in Section 5.1.2. Thus, the system in (5.50) is equivalent to the system (6.2) with right-hand side $\mathfrak{f} \equiv 0$. As discussed in Bekkouche et al. (2019) this is a non-stiff problem and thus it is reasonable to apply explicit schemes to solve it. Therefore, we use the second order explicit Heun scheme (6.4) for the time discretization of the micro problem. Introduce a subdivision of the interval $[0, \Delta t]$ in N_{mic} sub-intervals with step size $\Delta t_{\text{mic}} = \frac{\Delta t}{N_{\text{mic}}}$. Due to the vanishing right-hand side we find for $n = 0, \dots, N_{\text{mic}} - 1$

$$\tilde{w}_\ell^{n+1} = \bar{w}_\ell^n - \Delta t_{\text{mic}} (\mathfrak{M}^{\text{mic}})^{-1} \mathfrak{R}^{\text{mic}} \bar{w}_\ell^n, \quad (6.20a)$$

$$\bar{w}_\ell^{n+1} = \bar{w}_\ell^n - \frac{\Delta t_{\text{mic}}}{2} (\mathfrak{M}^{\text{mic}})^{-1} \mathfrak{R}^{\text{mic}} \left(\bar{w}_\ell^n + \tilde{w}_\ell^{n+1} \right). \quad (6.20b)$$

Here the initial value \bar{w}_ℓ^0 is an approximation resulting from a stationary PDE (5.53), which we again solve with Lagrange finite elements. It is also used to compute $\mathfrak{G}(0)$. Let us stress that the final time

for the scheme (6.20) is Δt , the macroscopic step size. With the solution $\bar{w}_\ell^{N_{\text{mic}}}$ we approximate the convolution kernel $\mathfrak{G}(\Delta t)$, which is used in the macroscopic schemes (6.18) and (6.19), respectively.

The right-hand side and extra source

Let us comment on the evaluation of the extra source. For both macroscopic scheme we need to evaluate the right-hand side, i.e.,

$$\mathfrak{f}(t_{n+1}) = \mathfrak{g}(t_{n+1}) - \mathfrak{J}(t_{n+1})\mathbf{u}(0).$$

For the recursive convolution approach we use the assumption (6.6). This transfers to the cell correctors \bar{w}_ℓ since those define the time dependence of the convolution kernel. As pointed out, the cell problems for \bar{w}_ℓ and w_ℓ^0 only differ in the initial value. Thus, we assume that the structure of the cell corrector w_ℓ^0 is again given by a decaying exponential. Therefore, it is reasonable to assume for the extra source that there exists full rank matrices $\mathfrak{J}_0, \mathfrak{J}_1 \in \mathbb{R}^{N_{\text{vH}} \times N_{\text{vH}}}$ such that

$$\mathfrak{J}(t)\mathbf{u}_0 = \mathfrak{J}_0 \exp(-\mathfrak{J}_1 t)\mathbf{u}_0.$$

Consequently, a recursive definition of the extra source is given as

$$\mathfrak{J}^{n+1}\mathbf{u}_0 = \mathfrak{J}(\Delta t)\mathfrak{J}(0)^{-1}\mathfrak{J}^n\mathbf{u}_0. \quad (6.21)$$

As for the convolution kernel we only need to compute $\mathfrak{J}(0)$ and $\mathfrak{J}(\Delta t)$. This is done before the macroscopic time evolution.

With the evaluation of the right-hand side we have all ingredient at hand for the fully discrete scheme, which is summarized in the next section.

The fully discrete scheme

In this section we eventually present our fully discrete scheme. In order to solve the macroscopic Maxwell system modeling wave propagation in locally periodic structures we start on the micro level.

For each sampling domain we first solve the stationary cell problems (5.25), (5.53). The latter solution is used as initial value for the evolution problem (5.52). We evolve using the Heun scheme (6.20) until $t = \Delta t$, using N_{mic} time steps of size Δt_{mic} . After that, all necessary solutions are available to compute the macro finite element matrices.

Thus, we get back to the macro level where we use the recursive convolution in combination with either the implicit Euler (6.18) or the Crank–Nicolson scheme (6.19) to compute the time evolution of the Maxwell system. A sketch of the space discrete part of the scheme is shown in Figure 6.1.

In the next section we comment on error estimates concerning the time discretization and the fully discrete scheme.

6.3.2 Fully discrete error analysis

In this section we comment on the fully discrete error analysis of the recursive FE-HMM scheme. In contrast to the previous chapter we do not provide a rigorous error analysis here. In this chapter we aim for an applicable scheme rather than its analysis. Moreover, we point out that for general cases the

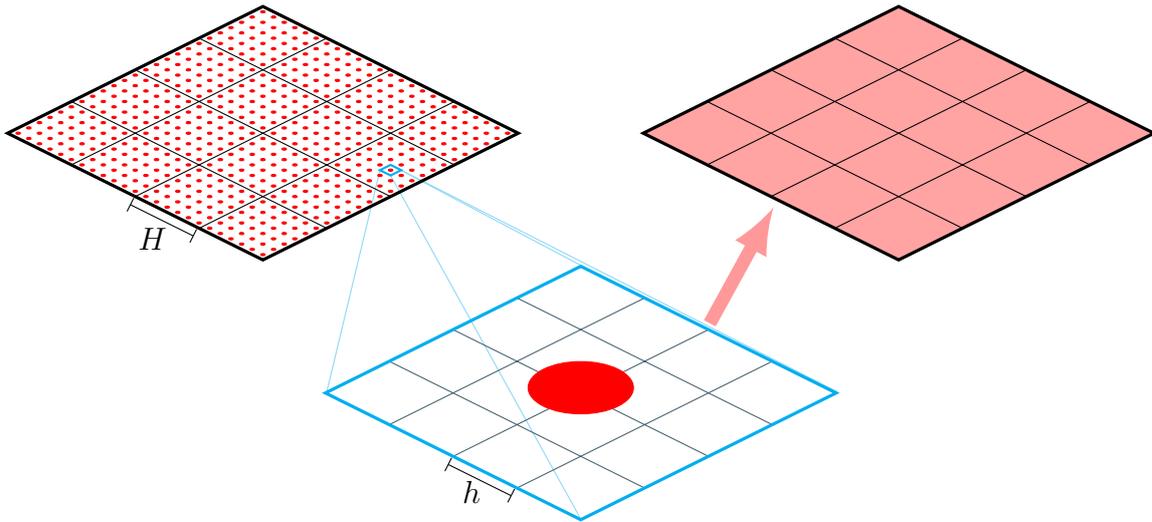


Figure 6.1: Sketch of the HMM: To solve a macroscopic problem with a (locally) periodic structure we solve a micro problem of size δ . Here the small scale is resolvable. Thus, the information of the micro problem is used to derive an effective parameter that has no microscopic oscillations. This effective problem may be solved using the original mesh size H .

recursive convolution approach does not apply since the assumption (6.6) is not satisfied. Nevertheless, let us present results for various sub problems of the recursive FE-HMM. We emphasize that this section is a good starting point for the fully discrete error analysis of the presented scheme.

As in Chapter 5 first consider the error that occurs due to the approximation of the effective parameters. All error estimates from that chapter are continuous in time. In Section 5.3.2, we analyzed the error of the space discretization of the Sobolev equation. Moreover, we showed an estimate on the micro error in Lemma 5.3.7. The first step in the direction of a fully discrete error estimate is to show an estimate for the error between the exact effective convolution kernel at time t_n denoted as $\mathfrak{G}(t_n)$ and its fully discrete approximation $\mathfrak{G}^{H,n}$. In Lemma 5.3.7, we considered the error of the exact kernel to the space discrete approximation $\mathfrak{G}^H(t_n)$. Hence, we have to estimate the error

$$\mathfrak{G}^{H,n} - \mathfrak{G}^H(t_n).$$

Here we may use the result concerning the fully discrete scheme for Sobolev equations given in (Bekkouche et al., 2019, Corollary 3.15). The authors show that the error between the solution $\bar{w}_\ell(t_n)$ and its approximation $\bar{w}_\ell^{h,n}$ by k th-order Lagrange elements and Heun's method (6.20) is bounded by

$$\left\| \bar{w}_\ell(t_n) - \bar{w}_\ell^{h,n} \right\|_{\mathbf{H}^1(Y; \mathbb{R}^N)} \leq C (h^k + \Delta t_{\text{mic}}^2).$$

This result holds true if the solution has the regularity $\bar{w}_\ell \in C^3(0, T; \mathbf{H}^{k+1}(Y; \mathbb{R}^N))$.

The next step in the analysis of the recursive FE-HMM method is the quantification of the error due to the exponential structure. To be more precise, the assumption (6.6) introduces a modeling error. To our knowledge there is no general result about the approximation of solutions to Sobolev equations by exponential functions. Still, we point out that the numerical experiments in Chapter 7 show that this

approximation is accurate at least for various examples. In a nutshell there is an error of the form

$$\min_{\mathfrak{G}_0, \mathfrak{G}_1} \|\mathfrak{G}^{H,n} - \mathfrak{G}_0 \exp(-\mathfrak{G}_1 t_n)\|_F,$$

which has to be studied for each given microscopic structure separately.

If the convolution kernel is actually given as an exponential we may use results obtained for classic polarization models. Here we reach the macro level and thus the final step in the error analysis is to estimate the error of the recursive convolution approach. The main contribution to the error analysis of integro-differential Maxwell system is found in [Li \(2007\)](#) and [Li and Chen \(2008\)](#), where in both references the recursive convolution technique is used for time integration. The authors analyze the implicit Euler method and the Crank–Nicolson scheme combined with recursive convolution for first and second order Maxwell systems. For the analysis of time integration schemes with the use of memory variables in dispersive media as in Section 6.2.3 we refer to [Li \(2011\)](#); [Li and Zhang \(2010\)](#). In the first reference the authors analyze the Crank–Nicolson method in time, whereas in the second reference the explicit leap-frog scheme is analyzed. Let us mention that one observes the expected convergence, i.e., second order in time for both schemes. In [Lanteri and Scheid \(2013\)](#) a discontinuous Galerkin approach in space is combined with a leap-frog scheme in time to solve a Maxwell–Debye system with the memory variable method. Finally, we highlight [Hochbruck et al. \(2019\)](#) where the fully discrete error of the FE-HMM for the vacuum Maxwell system (5.31) has been analyzed for algebraically stable Runge–Kutta schemes. This is also a starting point for the analysis of the memory variable method since here we directly apply classic Runge–Kutta schemes.

CHAPTER 7

Implementation and numerical experiments

In this chapter we concentrate on the implementation of the recursive FE-HMM algorithm. We present the approach of an efficient implementation, especially for the computation of the homogenized parameters and the approximation of the convolution. The implementation of the program is done in the C++ library `deal.II` (Arndt et al., 2020). Moreover, we present some numerical experiments to show the functionality of the code in particular for cases that are not covered by the theory.

We start to describe the algorithmic approach in the next section where we comment on the approximation of the parameters in Section 7.1.1 followed by a discussion of the convolution in Section 7.1.2. This presentation of the results is followed by numerical examples related to the microscopic level in Section 7.2. In Section 7.3, we consider the macroscopic scheme and test the fully discrete scheme.

7.1 Algorithmic approach

Let us start with a short overview of the structure of `deal.II` programs. The solution of a PDE is basically split in three parts. First the `setup_system` function is called. Here the degrees of freedom are distributed and the memory is allocated. The second function is the `assemble_system` function, where actually the finite element matrices are assembled. The last step is the `solve` function, which is used to either solve a linear system of equations in the time independent case or introduces some kind of time integration scheme for the evolution of the solution.

This division also suits the presentation from the previous chapters. More precisely, the Heterogeneous Multiscale Method from Chapter 5 is found in the assembling process. The recursive convolution introduced in Chapter 6, however, is implemented in the `solve` function.

The code presented here is an extension of the code provided in Hochbruck et al. (2019). We extended this program to the use of the `FESystem` class. This enables an easy extension to systems with more unknowns than the electric and magnetic field, which is important for the polarization and magnetization we consider. Moreover, it is straight forward to use other than the curl conforming finite elements for

the latter mentioned fields.

The code is provided [<https://www.doi.org/10.5445/IR/1000129217>].

7.1.1 Efficient parallel computation of the HMM parameters

The computation of the HMM parameters significantly depends on the structure of the heterogeneous ones. More precisely, if the parameter is purely periodic it is completely independent of the macroscopic scale. Thus, all cell problems can be solved on the unit cell Y or one chosen sampling domain $Y^\kappa(\bar{x})$. In contrast, for a locally periodic parameter this is not true since it depends on the macroscopic scale. For this case we use the sampling domains $Y^\kappa(\bar{x})$, which incorporate the macroscopic dependence. An important observation is that all these transformed cell problems are independent of each other. Thus, here is a potential for parallel computations. More exactly, a distribution of these independent problems to different processes yields a parallel assembly of the finite element matrices.

The starting point for this parallel approach is the distribution of the cells, and thus the triangulation, to different processes. Here we make use of the interface of `deal.II` to the `p4est` library (Burstedde et al., 2011), which provides parallel distribution of meshes. Moreover, we use the `Trilinos` library (Trilinos Project Team, 2020 (accessed February 1, 2021)) for parallel linear algebra. Details on the distribution using `p4est` are found in Bangerth et al. (2012). The parallel solution of the cell problems is the point where the HMM becomes really efficient. The distribution is part of the `setup_system` function. We remark that this setup is also the point where we save allocated memory due to the use of the recursive convolution. More precisely, if we used the brute force quadrature from Section 6.2.1 we would have to store all solutions. With the recursive approach this reduces to the solution from the last step in time integration. Thus, we only store one solution, which is evolved through time. The same holds true for the convolution kernel. Again with brute force quadrature we would have to store and compute all kernels. Due to the recursive Crank–Nicolson scheme we only need two convolution kernel finite element matrices.

After the system has been set up and the triangulation has been distributed, the next step is to call the function `assemble_system`. Let us briefly comment how these matrices are assembled in the homogeneous case, i.e., assume that \mathbf{M}^{eff} is known explicitly. This is classical for finite element methods. First we run a loop over all cells of our triangulation. Within this loop we start to iterate over all quadrature points of the present cell. Since we know the parameters, we evaluate them at the quadrature points. The innermost loop runs over the degrees of freedom per cell and here we actually compute the element matrix corresponding to the cell. Then leaving the loops over the degrees of freedom and the quadrature points, the only thing that remains is to add the contribution of the cell to the global finite element matrix. See Algorithm 1 for the pseudo code.

As explained in Chapter 5, the execution of line 3 in Algorithm 1 is in general not possible, since there is no explicit representation of the effective parameters. Therefore we introduced the HMM in Section 5.2.1. The application of a finite element method to the micro problem (5.22) yields a linear system for every quadrature point \bar{x} and every macroscopic basis function

$$\mathfrak{M}^{\text{mic}}(\bar{x})\psi_i(\bar{x}) = \mathfrak{g}_i^{\text{mic}}(\bar{x}). \quad (7.1)$$

The solution is used to compute an approximation of the finite element matrix as in (5.23). Thus, for every macroscopic quadrature point \bar{x} we solve the discrete cell problem on the sampling domain $Y^\kappa(\bar{x})$.

Algorithm 1 Assembly of finite element matrix

```

1: for  $K \in \mathcal{T}_H$  do
2:   for  $q \in \{1, \dots, Q_K\}$  do
3:     evaluate  $\mathbf{M}^{\text{eff}}(x_K^q)$ 
4:     for  $i, j \in K$  do
5:       compute  $\mathfrak{M}_{i,j}^K$ 
6:     end for
7:   end for
8:    $\mathfrak{M} \leftarrow \mathfrak{M} + \mathfrak{M}^K$ 
9: end for

```

In (7.1) we see that the finite element matrix on the microscopic level is independent of the macroscopic basis function. Therefore, we first assemble the microscopic finite element matrix for a given quadrature point and then loop over the macroscopic basis functions that are not zero on the considered cell. Here we assemble the right-hand side, which depends on these basis functions. Finally, we solve the cell problems. With these solutions we compute the element matrix, which is eventually added to the macroscopic finite element matrix. The pseudo code for the above computation is shown in Algorithm 2.

Algorithm 2 Assembly of HMM finite element matrix

```

1: for  $K \in \mathcal{T}_H$  do
2:   for  $q \in \{1, \dots, Q_K\}$  do
3:     assemble  $\mathfrak{M}^{\text{mic}}(x_K^q)$ 
4:     for  $i \in K$  do
5:       assemble  $\mathfrak{g}_i^{\text{mic}}(x_K^q)$ 
6:       solve (7.1)
7:       compute  $\mathfrak{M}^K(x_K^q)$ 
8:     end for
9:      $\mathfrak{M}^K \leftarrow \mathfrak{M}^K + \mathfrak{M}^K(x_K^q)$ 
10:  end for
11:   $\mathfrak{M} \leftarrow \mathfrak{M} + \mathfrak{M}^K$ 
12: end for

```

In Algorithm 2, it is clearly evident that the different cell problems are independent of each other. Thus, a distribution of the cells to different processes is suitable.

We point out that a generalization of the above algorithms to the other HMM parameters, especially the time dependent ones, is straight forward. In addition to the matrix $\mathfrak{M}^{\text{mic}}$ we have to assemble $\mathfrak{R}^{\text{mic}}$. Moreover, the initial problem for \bar{w}_ℓ^h uses a different right-hand side. Of course, in the time dependent case the solution of the cell problems iterates over all microscopic time steps, i.e., we solve (6.20).

7.1.2 Time integration and convolution approximation

The time integration procedure has been introduced in the previous chapter. Here, we comment on the implementation of the schemes. Note, that both, the implicit Euler scheme (6.18) and the Crank–Nicolson

method (6.19) are only implicit in the update of the new solution but not in the convolution. Thus, we rewrite the both schemes by inserting (6.19b) and (6.18b) in (6.19a) and (6.18a), respectively. For the implicit Euler scheme this yields

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \Delta t (\mathfrak{M} + \Delta t (\mathfrak{A} + \mathfrak{R} + \Delta t \mathfrak{G}(0)))^{-1} (\mathfrak{f}^{n+1} - (\mathfrak{A} + \mathfrak{R} + \Delta t \mathfrak{G}(0)) \mathbf{u}^n - \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1} \mathfrak{J}^n). \quad (7.2)$$

The update for the Crank–Nicolson scheme is

$$\begin{aligned} \mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{2} \left(\mathfrak{M} + \frac{\Delta t}{2} \left(\mathfrak{A} + \mathfrak{R} + \frac{\Delta t}{2} \mathfrak{G}(0) \right) \right)^{-1} & \left(\mathfrak{f}^{n+1} + \mathfrak{f}^n - 2 \left(\mathfrak{A} + \mathfrak{R} + \frac{\Delta t}{4} (\mathfrak{G}(\Delta t) + \mathfrak{G}(0)) \right) \mathbf{u}^n \right. \\ & \left. - (\mathbf{I} + \mathfrak{G}(\Delta t) \mathfrak{G}(0)^{-1}) \mathfrak{J}^n \right). \end{aligned} \quad (7.3)$$

The updates in (7.2) and (7.3) basically consist of two steps. First, we evaluate the expression in the right most parenthesis. Thereafter, we evaluate the multiplication by the inverse, which is actually done by solving the equivalent linear system. Eventually, we update the solution and perform the recursive quadrature of the convolution as in (6.19b) and (6.18b).

In addition to the recursive FE-HMM algorithm for the macroscopic Maxwell problem we also provide a code for solving the cell problems separately. This code is used in Section 7.2.1 where we consider four different microscopic structures and their corresponding effective parameters. This program is based on the step 4 tutorial from the `deal.II` page.

Remark 7.1.1. *We remark that all linear system that occur in our algorithm are solved using either the CG or GMRES method implemented in the library.*

Exponential least squares fit

In the next section we frequently use exponential least squares fitting. This is to check how accurate the assumption in (6.6) is. The method we use for this fitting is the `lsqcurvefit` function in Matlab. The data is taken from the components of the parameters that were approximated by the solution of the cell problems. The function also provides a residual error that is stated several times.

7.2 Numerical examples on the microscopic level

In this section we are finally in the position to show numerical experiments of the Recursive FE-HMM method. Before we actually consider the macroscopic Maxwell system, we show results on the microscopic level. Therefore, in the next section we introduce four different microscopic structures, show the corresponding effective parameters and demonstrate the convergence of the numerical approximation. For all these parameters we examine how close the corresponding convolution kernel is to an exponential function.

7.2.1 Microstructures and cell correctors in the conductivity setting

We study four different material parameters in the setting of (4.2), which can be equivalently written as (4.11) for $N_E = N_H = 0$. Thus, for a given permittivity and conductivity we aim to solve the system (4.37) with effective parameters as in (4.38). In all examples below the permeability $\mu = \mu_0$ is assumed to be the vacuum permeability. Therefore, we only have to compute the entries of the effective parameters that belong to the electric field.

All the parameter definitions that follow are on the unit cell $Y = (-\frac{1}{2}, \frac{1}{2})^3$. This cell is exactly what is used to determine the effective parameters in the process of homogenization and represents the microscopic scale. The macroscopic definition of the heterogeneous parameters is straight forward due to periodicity.

Layered material

The first parameter we study is a layered material that varies only in y_1 direction. Consider the permittivity and conductivity given as

$$\varepsilon(y) = \begin{cases} \varepsilon_1, & y_1 < 0 \\ \varepsilon_2, & \text{else} \end{cases}, \quad \sigma(y) = \begin{cases} \sigma_1, & y_1 < 0 \\ \sigma_2, & \text{else} \end{cases} \quad \text{for } y \in Y. \quad (7.4)$$

This material is isotropic in both composites. As shown in the Appendix A.3, in this setting it is possible to compute the effective parameters analytically. They are given as

$$\varepsilon^{\text{eff}} = \begin{pmatrix} \frac{2\varepsilon_1\varepsilon_2}{\varepsilon_1+\varepsilon_2} & 0 & 0 \\ 0 & \frac{\varepsilon_1+\varepsilon_2}{2} & 0 \\ 0 & 0 & \frac{\varepsilon_1+\varepsilon_2}{2} \end{pmatrix}, \quad \sigma^{\text{eff}} = \begin{pmatrix} \frac{2(\varepsilon_1^2\sigma_2+\varepsilon_2^2\sigma_1)}{(\varepsilon_1+\varepsilon_2)^2} & 0 & 0 \\ 0 & \frac{\sigma_1+\sigma_2}{2} & 0 \\ 0 & 0 & \frac{\sigma_1+\sigma_2}{2} \end{pmatrix},$$

$$\mathbf{G}^{\text{eff}}(t) = \begin{pmatrix} -\frac{2(\varepsilon_1\sigma_2-\varepsilon_2\sigma_1)^2}{(\varepsilon_1+\varepsilon_2)^3} e^{-\frac{\sigma_1+\sigma_2}{\varepsilon_1+\varepsilon_2}t} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{J}^0(t) = \begin{pmatrix} -\frac{(\varepsilon_1\sigma_2-\varepsilon_2\sigma_1)(\varepsilon_1-\varepsilon_2)}{(\varepsilon_1+\varepsilon_2)^2} e^{-\frac{\sigma_1+\sigma_2}{\varepsilon_1+\varepsilon_2}t} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It is reasonable that the effective material parameters are anisotropic. Observe how the direction of the layer is visible in the effective parameters. Especially, the convolution kernel and the extra source, which are only present due to homogenization vanish for the y_2 and y_3 component. Note that in this setting the convolution kernel and the additional source are in fact exponential functions.

For the numerical experiments we choose $\varepsilon_1 = 1 = \sigma_1$, $\varepsilon_2 = 2$ and $\sigma_2 = 4$ as well as the final time $T = 10$. In Figure 7.1 we show the only non-zero component of the stationary and time dependent corrector at initial time and at $t = 5$. Note, how the jump in the parameters is covered in the correctors. They

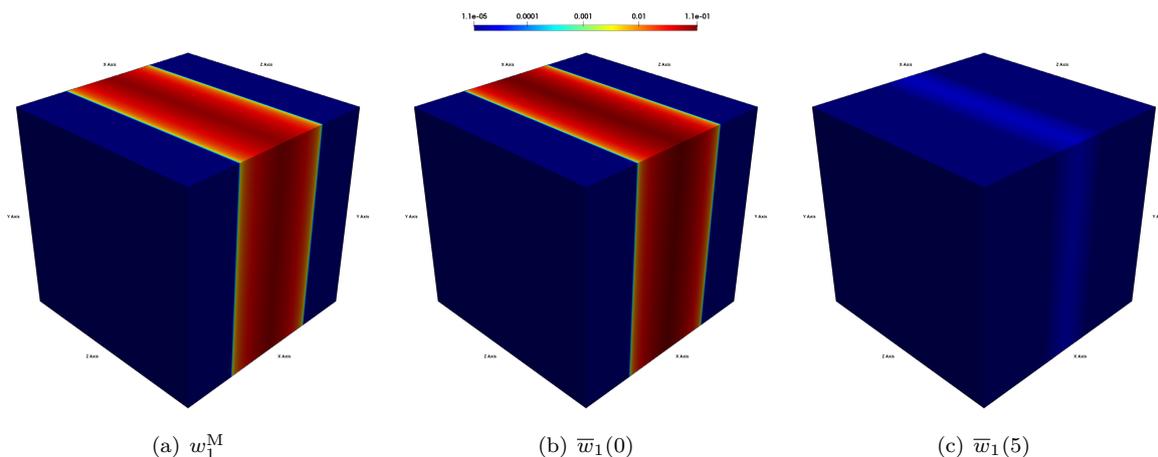


Figure 7.1: Correctors for layered material. 7.1(a) shows the stationary corrector related to the permittivity ε . In 7.1(b) we show the initial value of the time dependent corrector, which depends on both ε and σ . 7.1(c) shows the time dependent corrector after 512 time steps at $T = 5$. Note, that we use a logarithmic scale, which highlights the exponential decay of the corrector \bar{w}_1 over time.

evolve linearly in the same direction. Additionally, the stationary corrector w^M and the time dependent corrector \bar{w} have a very similar shape. Finally, the decay of the corrector \bar{w} is clearly evident as well. Next we present numerical results concerning the layered parameter. In this setting we use the exact effective parameters and compute the errors with respect to those. More precisely, we show the errors

$$\|\varepsilon^{\text{eff}} - \varepsilon^{\text{HMM}}\|_F, \quad \|\sigma^{\text{eff}} - \sigma^{\text{HMM}}\|_F, \quad \max_{t \in [0, T]} \|\mathbf{G}^{\text{eff}}(t) - \mathbf{G}^{\text{HMM}}(t)\|_F,$$

for various values of h with fixed $\Delta t = \frac{10}{1024}$ computed with the second order Heun method. Note that this is the micro error and that the modeling error in this case vanishes due to periodicity.

Let us mention that due to the structure of the parameter, which has its discontinuity in the middle of the domain, we get very good approximations if we hit this jump with our triangulation, i.e., for values of $h = \frac{1}{2^k}$ for $k \in \mathbb{N}$. This explains our choice of discretization parameters. The spatial errors for first order Lagrange elements are shown in Figure 7.2 and Table 7.1. If we consider this parameter as a microscopic structure the theory from Chapter 5 is not valid, since the parameter has not enough regularity. However, for the elliptic and parabolic problems on the unit cell we expect an order reduction due to the discontinuity, which suggests first order of convergence for this parameter choice.

As next step we consider the error in time for the explicit Euler and the Heun method. We explained in Chapter 6 that those schemes have been analyzed in Bekkouche et al. (2019) for the time integration of Sobolev equations. Thus, we expect first order of convergence for the explicit Euler and second order for the Heun scheme. In Figure 7.3 and Table 7.2 we show the corresponding convergence plot and rates. We observe that both schemes match their proposed convergence rates. In this example even very coarse time steps yield good results.

Following the convergence study we briefly comment on the exponential structure. In this example we have the exact representation of the only non zero component of the convolution kernel as

$$\mathbf{G}_{11}^{\text{eff}}(t) = -\frac{8}{27}e^{-\frac{5}{3}t}.$$

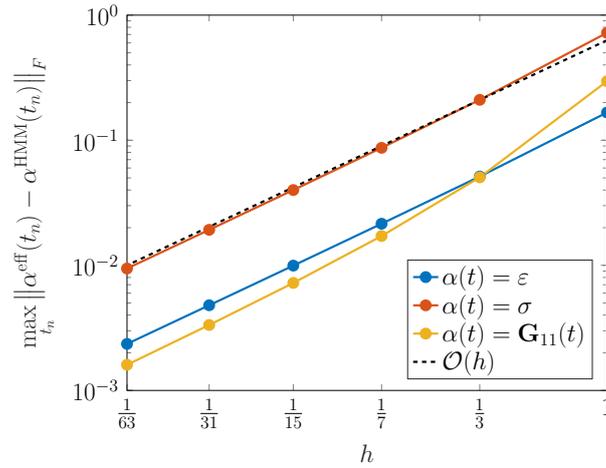


Figure 7.2: Convergence in space for second order Heun method with time step $\Delta t = \frac{10}{1024}$ and first order Lagrange elements.

h / dofs	$\ \varepsilon^{\text{eff}} - \varepsilon^{\text{HMM}}\ _F$		$\ \sigma^{\text{eff}} - \sigma^{\text{HMM}}\ _F$		$\max_{t \in [0, T]} \ \mathbf{G}^{\text{eff}}(t) - \mathbf{G}^{\text{HMM}}(t)\ _F$	
	rate		rate		rate	
1/8	0.16667	-	0.72222	-	0.2963	-
$\frac{1}{3}/64$	0.05128	1.07	0.21039	1.12	0.05051	1.61
$\frac{1}{7}/512$	0.02151	1.03	0.08695	1.04	0.01715	1.27
$\frac{1}{15}/4096$	0.00995	1.01	0.04000	1.02	0.00723	1.13
$\frac{1}{31}/32768$	0.00480	1.01	0.01923	1.01	0.00334	1.07
$\frac{1}{63}/262144$	0.00236	1.00	0.00943	1.00	0.00160	1.03

Table 7.1: Convergence rates for layered parameter for fixed time step $\Delta t = \frac{10}{1024}$ with second order Heun method and first order Lagrange elements.

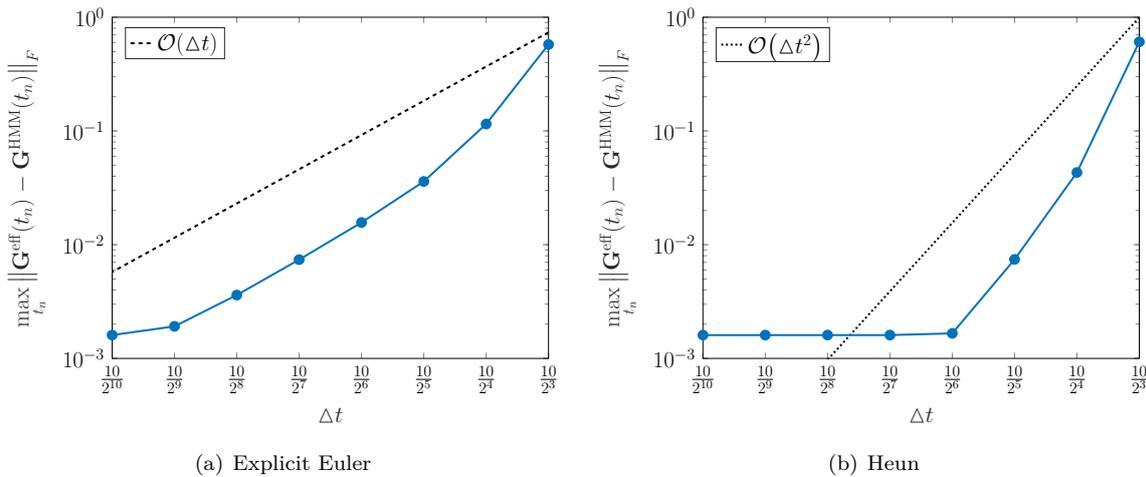


Figure 7.3: Convergence in time for the layered material parameters for the first order explicit Euler scheme and second order explicit Heun's method for $h = \frac{1}{63}$.

Δt	$\max_{t \in [0, T]} \ \mathbf{G}^{\text{eff}}(t) - \mathbf{G}^{\text{HMM}}(t)\ _F$			
	Explicit Euler		Heun	
	rate		rate	
$\frac{10}{8}$	0.57555	-	0.60788	-
$\frac{10}{16}$	0.11510	2.32	0.04316	3.82
$\frac{10}{32}$	0.03597	1.68	0.00742	2.54
$\frac{10}{64}$	0.01568	1.20	0.00166	2.16
$\frac{10}{128}$	0.00739	1.09	0.00160	0.05
$\frac{10}{256}$	0.00361	1.03	0.00160	0
$\frac{10}{512}$	0.00191	0.91	0.00160	0
$\frac{10}{1024}$	0.00160	0.26	0.00160	0

Table 7.2: Convergence of the explicit Euler scheme and the Heun method for fixed $h = \frac{1}{63}$ and 262144 degrees of freedom

We apply an exponential least squares fitting to the data that originates from the finest resolution in space and time, i.e., $h = \frac{1}{63}$ and $\Delta t = \frac{10}{1024}$. The result using the default settings from Matlabs `lsqcurvefit` function is the exponential function $-0.294613e^{-1.65536t}$. The squared error in the ℓ_2 -norm of the fitting to our data is

$$\sum_{k=0}^{1024} (-0.294613e^{-1.65536k\Delta t} - \mathbf{G}_{11}^{\text{HMM}}(k\Delta t))^2 \approx 2.1141 \times 10^{-7}.$$

The maximal error between the exact kernel and the exponential fit is

$$\max_{t \in [0, T]} \left| -\frac{8}{27}e^{-\frac{5}{3}t} + 0.294613e^{-1.65536t} \right| = \left| -\frac{8}{27} + 0.294613 \right| \approx 0.00168.$$

Here we find precisely the error that we observe for the finest resolutions in the Tables 7.1 and 7.2. This indicates that the error of the exponential fit is dominated by the error of our finite element method, at least for the resolution presented here.

The layered parameter from this section shows the expected rates of convergence in time. In space, however, we lose one order of convergence due to the discontinuity of the parameter. Nevertheless, here we have an exact solution and even more the convolution kernel in this setting is explicitly given as an exponential function. We exploit this fact in Section 7.3.2.

Continuous parameter

In this section we consider a microscopic structure that is given by the smooth parameter choice

$$\begin{aligned} \varepsilon(y) &= \frac{1}{2} (a + \sin(2\pi y_1)) (a + \sin(2\pi y_2)) (a + \sin(2\pi y_3)), \\ \sigma(y) &= \frac{1}{2} (b + \sin(2\pi y_1)) (b + \sin(2\pi y_2)) (b + \sin(2\pi y_3)), \end{aligned} \tag{7.5}$$

and thus ε and σ are both elements of $C_{\#}^{\infty}(Y; \mathbb{R})$. Thus, if we choose this as microscopic structure, the theoretical results in the Lemmas 5.2.3, 5.3.2 and 5.3.7 hold true. In our simulations we use the specific

choice

$$a = \sqrt{2}, \quad b = 2.$$

As shown in Jikov et al. (1994) the corresponding effective permittivity is the identity, i.e., $\varepsilon^{\text{eff}} = \mathbf{I}_{3 \times 3}$. In Figure 7.4, we show a contour plot of this parameter and a plot of a slice for $y_2 = \frac{1}{4}$.

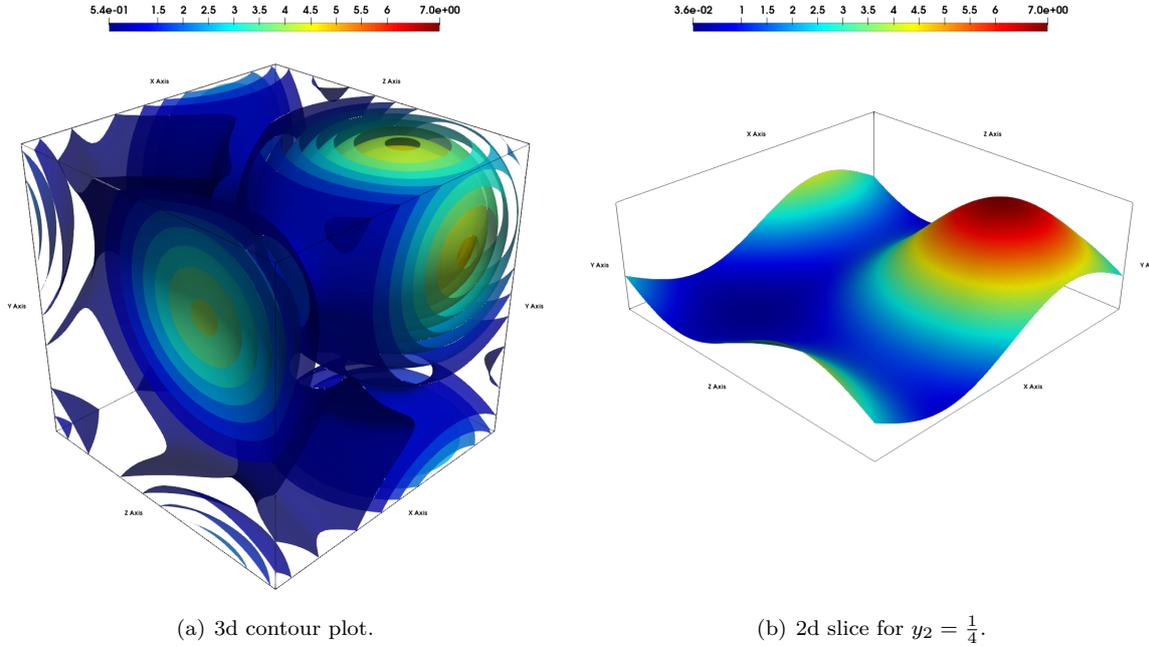


Figure 7.4: Smooth periodic parameter. The maximal value is $\frac{(\sqrt{2}+1)^3}{2}$ for $y = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.

For the computation of the parameter we used $h = \frac{1}{64}$ and the Heun method with $\Delta t = \frac{10}{1024}$ and end time $T = 10$. The first component of the correctors for this parameter are shown in Figure 7.5. Observe in Figure 7.5(a) that the first component of the stationary corrector only varies in the y_1 -direction but is not linear as for the layered parameter. The time dependent corrector has variations in all space directions. Moreover, the change of the corrector over time is evident by comparing Figure 7.5(b) with Figure 7.5(c). In addition to the change in the shape of the corrector we clearly see its decay over time.

We point out that this parameter choice leads to an isotropic effective parameter set, and thus we only show the first component of the correctors since the others have the same shape with respect to the space direction.

Before we study the convergence rates we comment on the error computation without exact solution. We compute the error between different levels of refinement either in space or time. For errors err_i and err_{i+1} on neighboring levels we compute the experimental order of convergence (EOC) as

$$\text{EOC} = \log_2 \left(\frac{\text{err}_i}{\text{err}_{i+1}} \right).$$

In Figure 7.6(a) and Table 7.3 we show the convergence rates for the spatial discretization with fixed time step $\Delta t = \frac{10}{1024}$ and Heun's method. For all parameters we observe second order of convergence for first order elements as expected.

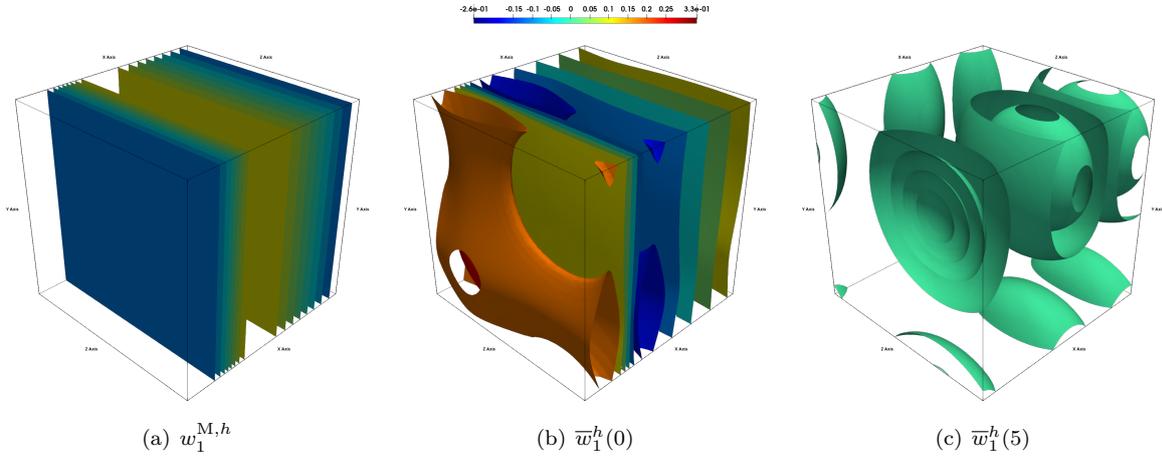


Figure 7.5: Contour plot of the first components of the correctors related to the continuous parameter. 7.5(a) shows the stationary corrector related to ε . 7.5(b) and 7.5(c) are the correctors at the initial value $t = 0$ and after 512 time steps ($\Delta t = \frac{10}{1024}$) with $h = \frac{1}{64}$. The shape of the time dependent corrector changes. Moreover, the decay over time is evident.

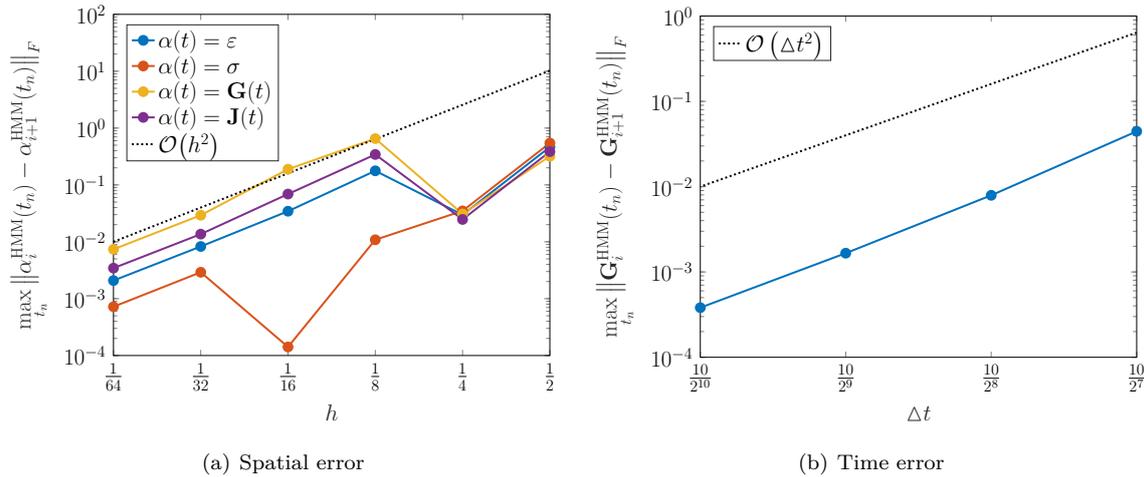


Figure 7.6: Convergence for the continuous parameter. The spatial errors are computed with fixed time step $\Delta t = \frac{10}{1024}$, whereas the time error uses fixed mesh size $h = \frac{1}{64}$.

h/i	$\ \alpha_i^{\text{HMM}} - \alpha_{i+1}^{\text{HMM}}\ _F$				$\max_{t_n} \ \alpha_i^{\text{HMM}}(t_n) - \alpha_{i+1}^{\text{HMM}}(t_n)\ _F$			
	$\alpha = \varepsilon$		$\alpha = \sigma$		$\alpha = \mathbf{G}$		$\alpha = \mathbf{J}$	
	EOC		EOC		EOC		EOC	
$\frac{1}{2}/1$	0.46502	-	0.54481	-	0.31914	-	0.38524	-
$\frac{1}{4}/2$	0.02983	3.96	0.03495	3.96	0.0309	3.37	0.02471	3.96
$\frac{1}{8}/3$	0.17706	-2.57	0.01086	1.69	0.65253	-4.40	0.34326	-3.8
$\frac{1}{16}/4$	0.03448	2.36	0.00014	6.26	0.18901	1.79	0.06909	2.31
$\frac{1}{32}/5$	0.00827	2.06	0.00291	-4.36	0.0293	2.69	0.01363	2.34
$\frac{1}{64}/6$	0.00208	1.99	0.00072	2.01	0.00737	1.99	0.00344	1.99

Table 7.3: Convergence rates for the continuous parameter between different levels of spatial refinement for fixed time step $\Delta t = \frac{10}{1024}$ with the second order Heun method.

$\Delta t/i$	$\max_{t \in [0, T]} \ \mathbf{G}_i^{\text{HMM}}(t) - \mathbf{G}_{i+1}^{\text{HMM}}(t)\ _F$	
	Heun	
	EOC	
$\frac{10}{128}/1$	0.04450	-
$\frac{10}{256}/2$	0.00792	2.49
$\frac{10}{512}/3$	0.00166	2.25
$\frac{10}{1024}/4$	0.00038	2.12

Table 7.4: Convergence of Heun's method for fixed $h = \frac{1}{64}$ and different levels of refinement in time.

The convergence in time is shown in Figure 7.6(b) and Table 7.4. In the convergent region we see the expected second order of convergence. The computation fails for time steps greater or equal $\Delta t = \frac{10}{32}$, which is probably due to the step size restriction for the explicit scheme.

In this example it is still possible to get exact results for the stationary effective parameters. To our knowledge this is not true for the time dependent convolution kernel and extra source. Thus, we take a look at the time evolution of these parameters as results of our computation. In Figure 7.7, we show the convolution kernel and the extra source evolving in time. Moreover, we present two possible exponential functions that are fitted to the data from our computation. The errors between the fitting with one exponential function and the data are

$$\sum_{k=0}^{1024} (\mathbf{G}_{11}^{\text{HMM}}(k\Delta t) - 0.697425e^{-3.679396k\Delta t})^2 \approx 0.0018,$$

$$\sum_{k=0}^{1024} (\mathbf{J}_{11}^{\text{HMM}}(k\Delta t) - 0.477970e^{-3.313265k\Delta t})^2 \approx 6.4539 \times 10^{-4}.$$

An improvement is the approximation with a sum of two exponential functions. The overall error for this

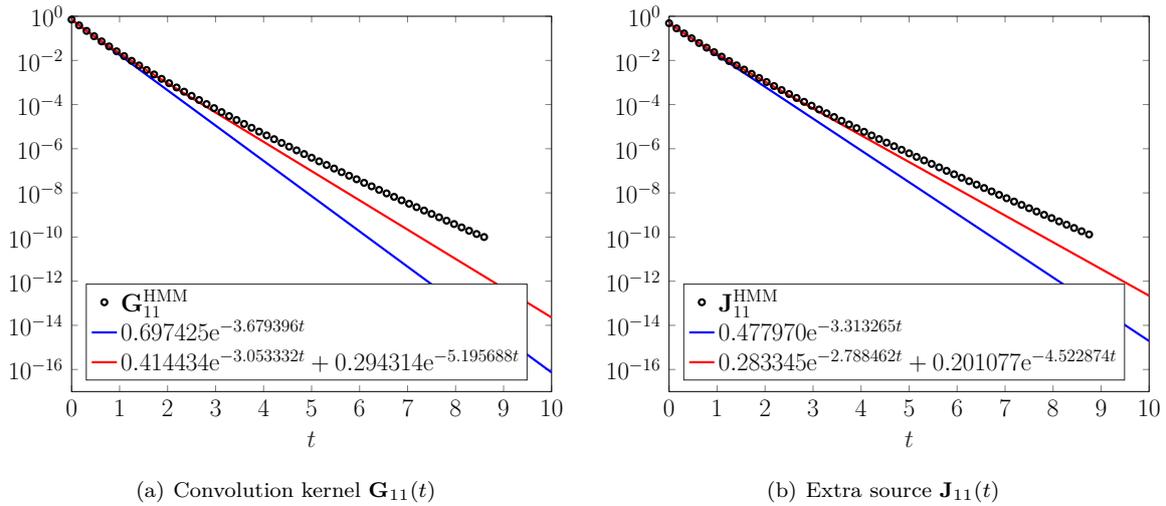


Figure 7.7: Semi logarithmic plot with exponentially fitted functions of absolute value of convolution kernel (a) and extra source (b) for the continuous parameter.

fitting is

$$\sum_{k=0}^{1024} (\mathbf{G}_{11}^{\text{HMM}}(k\Delta t) - 0.414434e^{-3.053332k\Delta t} - 0.294314e^{-5.195688k\Delta t})^2 \approx 1.9817 \times 10^{-6},$$

$$\sum_{k=0}^{1024} (\mathbf{J}_{11}^{\text{HMM}}(k\Delta t) - 0.283345e^{-2.788462k\Delta t} - 0.201077e^{-4.522874k\Delta t})^2 \approx 5.3744 \times 10^{-7}.$$

Thus, even in this complicated example we find exponential functions that are close to the data and therefore to the effective parameters. Hence, the recursive convolution scheme is applicable. Moreover, the use of a sum of exponential functions for the fitting yields a generalization, which is probably more accurate.

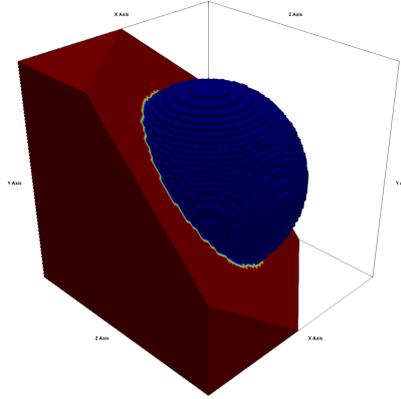
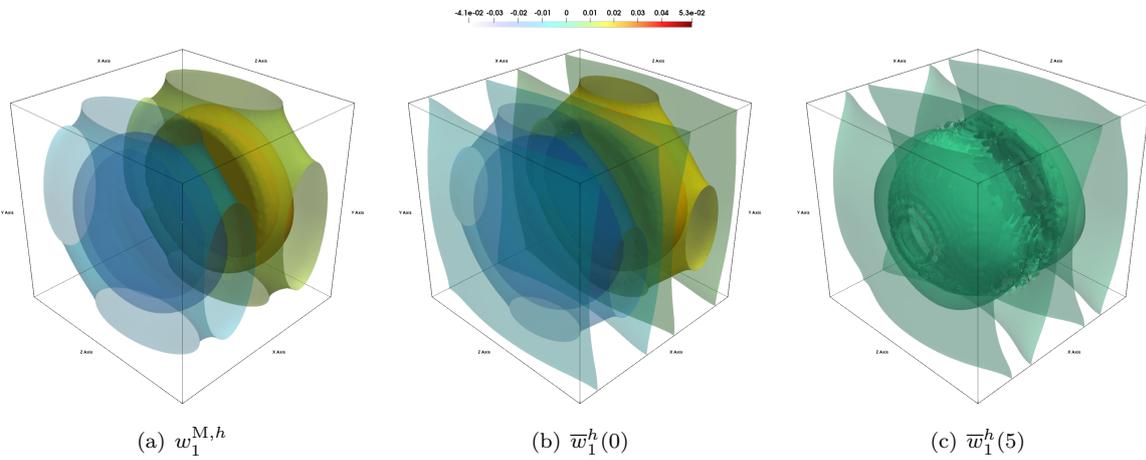
If we use this parameter as microscopic structure in our scheme the regularity enables us to apply the theory from Chapter 5. We show results for this setting in Section 7.3.3, but we emphasize that this regularity is not to be expected in practice. Especially, for composite materials discontinuous parameters are natural to occur.

Circular inclusion

The next example we consider is a circular inclusion. This could correspond to a gas that is enclosed within some other material. Here we are more general than in the first two examples in the sense that we allow for local periodicity. Thus, for a macroscopic $x \in \Omega$ we define a radius $0 \leq r(x) \leq 0.5$. Now for $y \in Y$ consider

$$\varepsilon(x, y) = \begin{cases} \varepsilon_1, & \|y\|_2 < r(x) \\ \varepsilon_2, & \text{else} \end{cases}, \quad \sigma(x, y) = \begin{cases} \sigma_1, & \|y\|_2 < r(x) \\ \sigma_2, & \text{else} \end{cases}. \quad (7.6)$$

This parameter is locally periodic since the radius depends on the macroscopic variable. We highlight the study of this parameter in Banks et al. (2006), which also covers the Debye setting and different radii. In Figure 7.8, we show the inclusion for the radius $r(x) = 0.4$. For now, we fix this radius.

Figure 7.8: The unit cell with circular inclusion with radius $r = 0.4$.Figure 7.9: Contour plots of correctors corresponding to the circular inclusion with radius $r = 0.4$.

Moreover, we choose the values $\varepsilon_1 = 1 = \sigma_1$, $\varepsilon_2 = 2$ and $\sigma_2 = 4$. Due to the symmetry of the circular inclusion, the effective parameters are isotropic. This again corresponds to corrector components that are similar with respect to their respective space direction. Hence, we only show the first components of the cell correctors in Figure 7.9. For the computation we used $h = \frac{1}{64}$, final time $T = 10$ with 1024 steps and the second order Heun scheme. We observe that the circular inclusion is represented in the correctors themselves and as in the layered case the initial value of \bar{w}_1^h is similar to the stationary corrector $w_1^{M,h}$. The decay of the time dependent corrector is also evident. In Figure 7.10 and Table 7.5 we show the error between neighboring levels of spatial refinement for fixed time step $\Delta t = \frac{10}{1024}$. Although we do not see a uniform rate, the overall decay is evident.

In Table 7.6, we show the experimental order of convergence in time for fixed $h = \frac{1}{64}$. The proposed second order of convergence is reflected in the data. For the finest resolution in space and time, the effective permittivity and conductivity are given as

$$\varepsilon^{\text{eff}} \approx 1.69493 \mathbf{I}_3, \quad \sigma^{\text{eff}} \approx 3.0421 \mathbf{I}_3.$$

For the time dependent parameters we again use an exponential fitting. The result concerning the

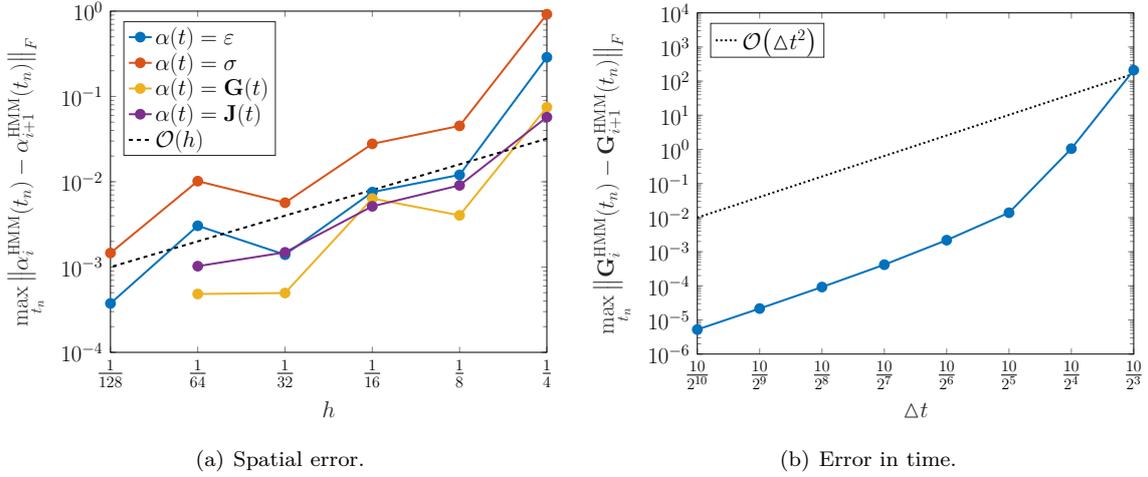


Figure 7.10: Convergence plots for circular inclusion with second order Heun method. Figure 7.10(a) shows the spatial convergence for fixed time step $\Delta t = \frac{10}{1024}$. In Figure 7.10(b), we fixed $h = \frac{1}{64}$ and used different time steps.

	$\ \alpha_i^{\text{HMM}} - \alpha_{i+1}^{\text{HMM}}\ _F$		$\max_{t_n} \ \alpha_i^{\text{HMM}}(t_n) - \alpha_{i+1}^{\text{HMM}}(t_n)\ _F$					
	$\alpha = \varepsilon$		$\alpha = \sigma$		$\alpha = \mathbf{G}$		$\alpha = \mathbf{J}$	
h/i	EOC		EOC		EOC		EOC	
$\frac{1}{4}/1$	0.28785	-	0.92075	-	0.07481	-	0.05721	-
$\frac{1}{8}/2$	0.01204	4.58	0.04518	4.35	0.00404	4.21	0.00907	2.66
$\frac{1}{16}/3$	0.00756	0.67	0.02784	0.70	0.00636	-0.66	0.00515	0.82
$\frac{1}{32}/4$	0.00140	2.44	0.00568	2.29	0.00050	3.68	0.00149	1.79
$\frac{1}{64}/5$	0.00305	-1.13	0.01017	-0.84	0.00048	0.04	0.00102	0.54
$\frac{1}{128}/6$	0.00037	3.02	0.00146	2.80	-	-	-	-

Table 7.5: Convergence between different levels of spatial refinement for circular inclusion. The time integration scheme is Heun's method with $\Delta t = \frac{10}{1024}$.

	$\max_{t_n} \ \mathbf{G}_i^{\text{HMM}}(t_n) - \mathbf{G}_{i+1}^{\text{HMM}}(t_n)\ _F$	
	Heun	
$\Delta t/i$	EOC	
$\frac{10}{8}/2$	209.115	-
$\frac{10}{16}/3$	1.04921	7.64
$\frac{10}{32}/4$	0.01391	6.24
$\frac{10}{64}/5$	0.00218	2.67
$\frac{10}{128}/6$	0.00042	2.39
$\frac{10}{256}/7$	0.00009	2.18
$\frac{10}{512}/8$	0.00002	2.09
$\frac{10}{1024}/9$	0.00001	2.05

Table 7.6: Convergence for circular inclusion of Heun's method for fixed $h = \frac{1}{64}$.

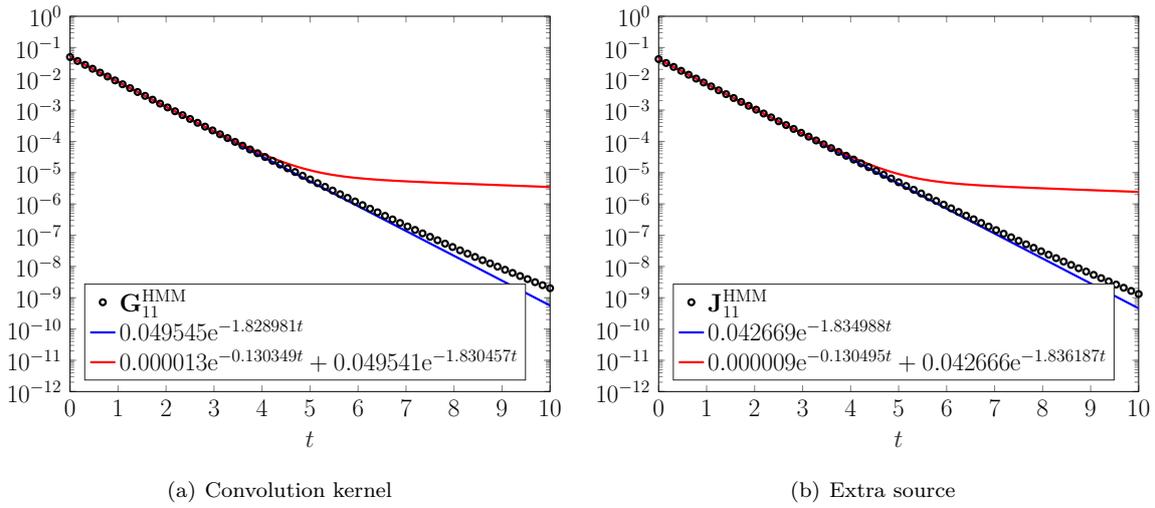


Figure 7.11: Semi logarithmic plot of convolution kernel (a) and extra source (b) including fitted exponential functions for circular inclusion.

convolution kernel is shown in Figure 7.11. Here we see that one exponential function fits the data very well. To be exact, the error of the approximation

$$\mathbf{G}_{11}^{\text{HMM}}(t) \approx 0.049545e^{-1.828981t},$$

is given as

$$\sum_{k=0}^{1024} (\mathbf{G}_{11}^{\text{HMM}}(k\Delta t) - 0.049545e^{-1.828981k\Delta t})^2 \approx 3.2837 \times 10^{-8}.$$

For the extra source we find the approximation

$$\mathbf{J}_{11}^{\text{HMM}}(t) \approx 0.042669e^{-1.834988t},$$

which has an error that is of the same order as the one for the convolution kernel. Therefore, we are in a situation where we expect that the use of the recursive convolution method is beneficial since the kernel is close to an exponential.

Split ring resonator

Let us now consider a micro structure that occurs in the field of meta materials. These materials are artificial composites that may have (electromagnetic) properties that are not observed in nature. In Figure 7.12, we show a unit cell from which a meta material may be constructed by periodicity. This cell consists of eight split ring resonators. Each resonator consist of two concentric rings, which have a cut at opposite sides. The exact definition of the parameter in the unit cell is found in the `microproblem.cc` program. The part in Figure 7.12 that is red has permittivity $\varepsilon = 2$ and conductivity $\sigma = 4$. The permittivity in the rest of the cell is equal to ε_0 , i.e., $\varepsilon = 1$ whereas the conductivity is split further. The ring resonators are fixed on crossing plates, which have conductivity $\sigma = 1$. All the rest of the cell has zero conductivity.

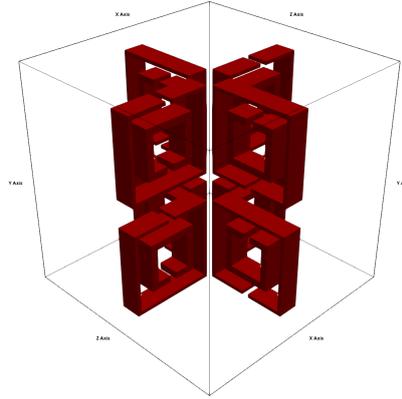


Figure 7.12: Split ring resonator unit cell.

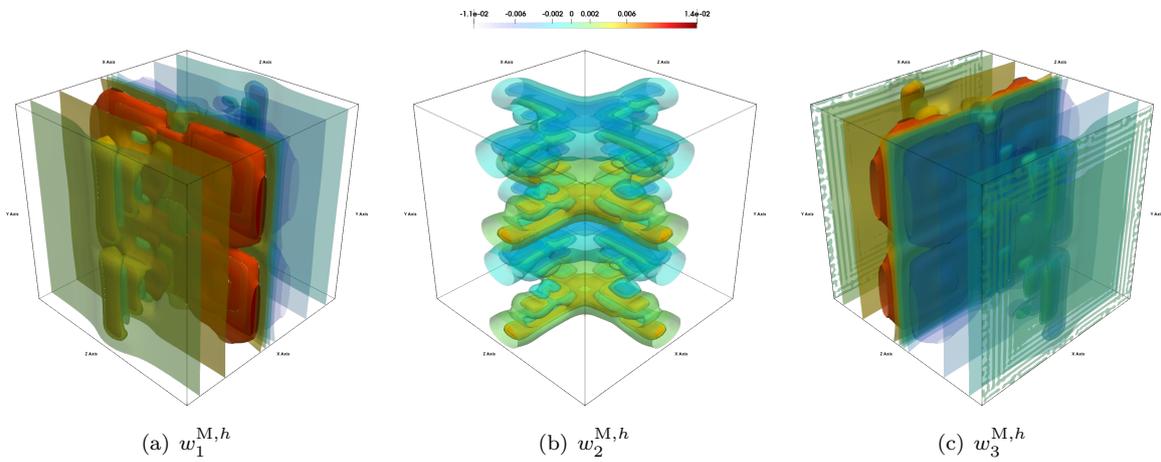


Figure 7.13: Stationary corrector components of the split ring resonator unit cell.

In Figure 7.13, we show all components of the stationary corrector. The first and third component are similar with respect to their space direction. The second component, however, looks different. This observation is also valid for the initial values of the time dependent corrector \bar{w}^h as seen in Figure 7.14. Again observe that the stationary corrector and the initial values of the time dependent corrector have a similar shape. Figure 7.15 shows these components at $t = 5$. Here the change of the scale is important to recognize the decay in all components.

Our computation results in the following effective anisotropic stationary parameters

$$\varepsilon^{\text{eff}} = \begin{pmatrix} 1.04896 & 0 & 0 \\ 0 & 1.0527 & 0 \\ 0 & 0 & 1.04896 \end{pmatrix}, \quad \sigma^{\text{eff}} = \begin{pmatrix} 2.69208 \times 10^{-1} & 0 & 0 \\ 0 & 3.22116 \times 10^{-1} & 0 \\ 0 & 0 & 2.69208 \times 10^{-1} \end{pmatrix}.$$

In Table 7.7, we show the relative error between different levels of refinement. Here it is important to note that the thickness of the resonators in our example is 0.1. This implies that we need at least 10 refinements in each space dimension to be able to resolve the split ring resonator in our simulations.

The two components $\mathbf{G}_{11}^{\text{HMM}}(t)$ and $\mathbf{G}_{22}^{\text{HMM}}(t)$ of the convolution kernel and the extra source including

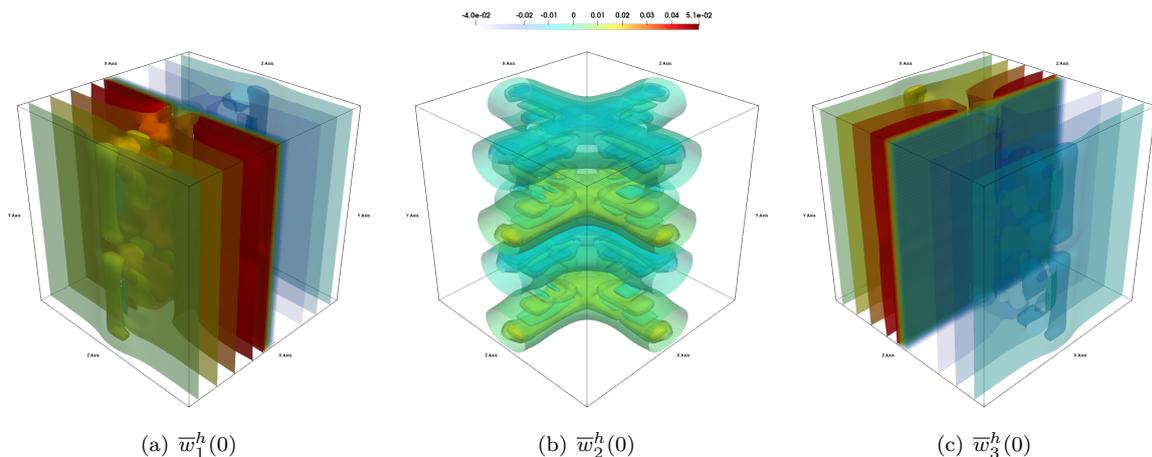


Figure 7.14: Initial value of the components of the time dependent corrector corresponding to the split ring resonator.

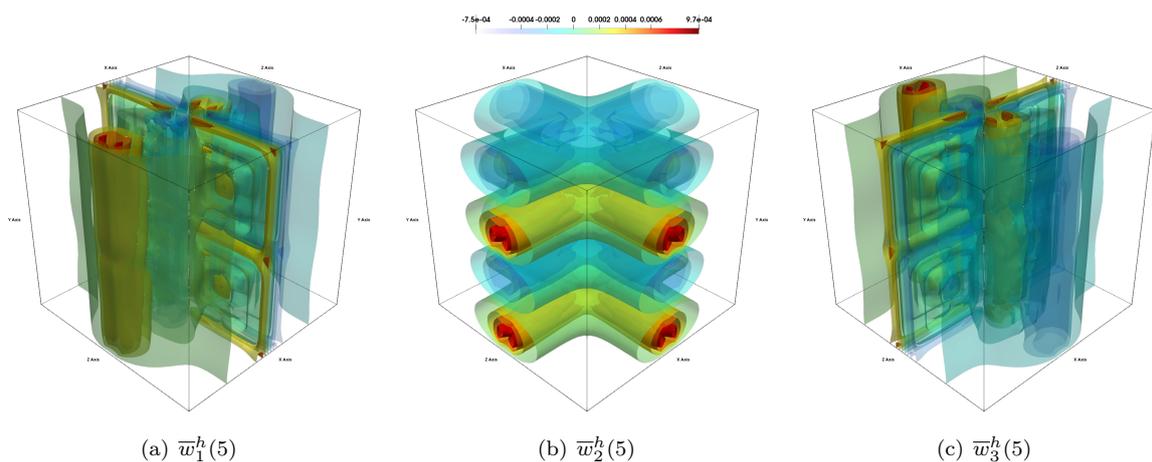


Figure 7.15: Value of the components of the time dependent corrector corresponding to the split ring resonator after 512 time steps, i.e., at $t = 5$.

h/i	$\ \alpha_i^{\text{HMM}} - \alpha_{i+1}^{\text{HMM}}\ _F$		$\max_{t_n} \ \alpha_i^{\text{HMM}}(t_n) - \alpha_{i+1}^{\text{HMM}}(t_n)\ _F$					
	$\alpha = \varepsilon$		$\alpha = \sigma$		$\alpha = \mathbf{G}$		$\alpha = \mathbf{J}$	
	EOC		EOC		EOC		EOC	
$\frac{1}{8}/1$	0.07310	-	0.59439	-	0.12084	-	0.02407	-
$\frac{1}{16}/2$	0.01793	2.03	0.04039	3.88	0.10001	0.27	0.02048	0.23
$\frac{1}{32}/3$	0.00758	1.24	0.07951	-0.98	0.03394	1.56	0.00674	1.60
$\frac{1}{64}/4$	0.01130	-0.58	0.04286	0.89	0.01723	0.98	0.00693	-0.04
$\frac{1}{128}/5$	0.00599	0.92	0.03928	0.13	-	-	-	-

Table 7.7: Convergence table for split ring resonator. We used Heun's method with stepsize $\Delta t = \frac{1}{168}$

	$\max_{t_n} \ \mathbf{G}_i^{\text{HMM}}(t_n) - \mathbf{G}_{i+1}^{\text{HMM}}(t_n)\ _F$	
	Heun	
$\Delta t/i$	EOC	
$\frac{5}{4}/2$	72.72913	-
$\frac{5}{8}/3$	0.16402	8.79
$\frac{5}{16}/4$	0.01654	3.31
$\frac{5}{32}/5$	0.00281	2.56
$\frac{5}{64}/6$	0.00057	2.31
$\frac{5}{128}/7$	0.00013	2.15
$\frac{5}{256}/8$	0.00003	2.07
$\frac{5}{512}/9$	0.00001	2.04

Table 7.8: Convergence for split ring resonator of Heun's method for fixed $h = \frac{1}{64}$.

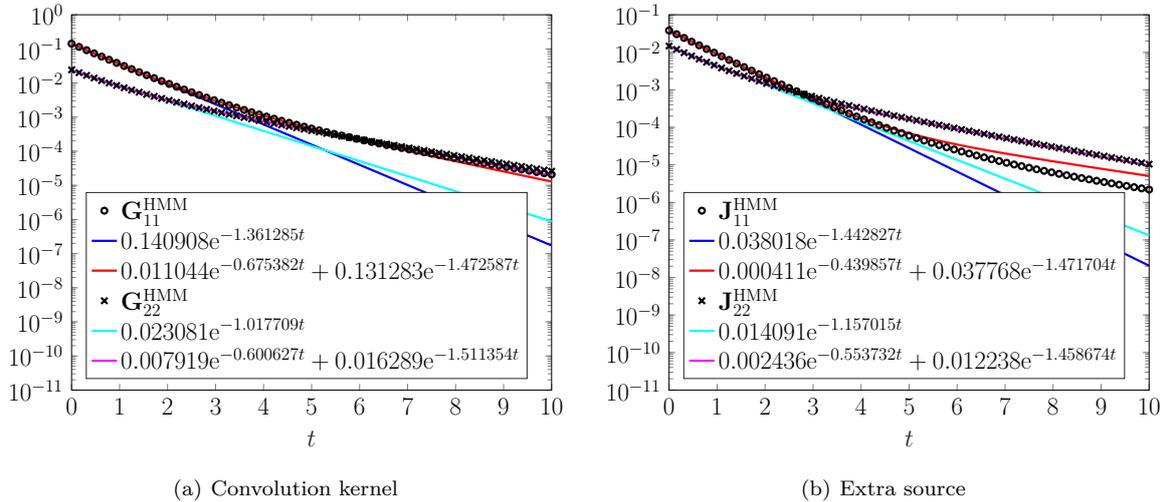


Figure 7.16: Exponential fit for absolute value of convolution kernel (a) $\mathbf{G}_{11}^{\text{HMM}}(t)$, $\mathbf{G}_{22}^{\text{HMM}}(t)$ and extra source (b) $\mathbf{J}_{11}^{\text{HMM}}(t)$, $\mathbf{J}_{22}^{\text{HMM}}(t)$ of SRR.

fitted exponential functions are shown in Figure 7.16. Here we observe that still for this complex parameter the approximation using at least two exponential functions is very accurate. The error is of order 10×10^{-7} , which is a lot smaller than the spatial error for this resolution. Thus, in a macroscopic Maxwell system it is feasible to use the recursive convolution or a memory variable approach.

We point out that the parameter values in this setting are not physically relevant. Nevertheless, this example shows that our numerical method covers this kind of structures and an experiment with physically relevant parameters should be possible.

After this study of different micro structures for the Maxwell system with conductivity we proceed to consider an example for the Maxwell–Debye setting from Section 3.1.3.

7.2.2 Microstructure in the Debye setting

Let us consider a more complex setting than the one from the previous examples. Here we consider the Maxwell–Debye system discussed in Section 3.1.3 and Section 4.6. Thus, the heterogeneous parameters are the instantaneous permittivity $\varepsilon_\infty^\delta$ and the relaxation time τ^δ . Moreover, we assume that the static frequency ε_s is constant, which yields $\Delta\varepsilon^\delta = \varepsilon_s - \varepsilon_\infty^\delta$. The parameters are chosen to vary only in y_1 direction as

$$\varepsilon(y) = \begin{cases} 1, & y_1 < 0.1 \\ 2, & \text{else} \end{cases}, \quad \tau(y) = \begin{cases} 1, & y_1 < 0.1 \\ 3, & \text{else} \end{cases}, \quad \text{for } y \in Y, \quad \varepsilon_s = 4, \quad \mu = 1.$$

As in the layered setting for conductivity we computed the exact effective parameters that correspond to the effective Maxwell system (4.77). We find the effective mass parameters

$$\varepsilon^{\text{eff}} = \begin{pmatrix} \frac{5}{4} & & \\ & \frac{7}{5} & \\ & & \frac{7}{5} \end{pmatrix}, \quad \mathbf{M}_P^{\text{eff}} = \begin{pmatrix} \frac{5}{13} & & \\ & \frac{2}{5} & \\ & & \frac{2}{5} \end{pmatrix},$$

and the effective damping

$$\mathbf{R}_{EE}^{\text{eff}} = \begin{pmatrix} \frac{35}{12} & & \\ & \frac{31}{15} & \\ & & \frac{31}{15} \end{pmatrix}, \quad \mathbf{R}_{EP}^{\text{eff}} = \mathbf{R}_{PE}^{\text{eff}} = \begin{pmatrix} -\frac{145}{156} & & \\ & -\frac{11}{15} & \\ & & -\frac{11}{15} \end{pmatrix}, \quad \mathbf{R}_{PP}^{\text{eff}} = \begin{pmatrix} \frac{155}{507} & & \\ & \frac{4}{15} & \\ & & \frac{4}{15} \end{pmatrix}.$$

The most interesting part are the effective convolution kernels, which may be computed using the formulation from Section 4.6 and the techniques from the Appendix A.3. All these kernels are given as linear combination of two exponential functions,

$$\begin{aligned} \mathbf{G}_{EE}^{\text{eff}}(t) &= \mathbf{G}_{EE}^1 e^{-\frac{20-3\sqrt{43}}{26}t} + \mathbf{G}_{EE}^2 e^{-\frac{20+3\sqrt{43}}{26}t}, \\ \mathbf{G}_{EP}^{\text{eff}}(t) &= \mathbf{G}_{PE}^{\text{eff}}(t) = \mathbf{G}_{EP}^1 e^{-\frac{20-3\sqrt{43}}{26}t} + \mathbf{G}_{EP}^2 e^{-\frac{20+3\sqrt{43}}{26}t}, \\ \mathbf{G}_{PP}^{\text{eff}}(t) &= \mathbf{G}_{PP}^1 e^{-\frac{20-3\sqrt{43}}{26}t} + \mathbf{G}_{PP}^2 e^{-\frac{20+3\sqrt{43}}{26}t}, \end{aligned}$$

where $\mathbf{G}_{EE}^1, \mathbf{G}_{EE}^2, \mathbf{G}_{EP}^1, \mathbf{G}_{EP}^2, \mathbf{G}_{PP}^1, \mathbf{G}_{PP}^2 \in \mathbb{R}^{3 \times 3}$ are constant matrices. Thus, although the damping parameter in the Maxwell–Debye setting is only positive semi-definite we find an exponentially decaying convolution kernel, although with two exponential functions. This example with exact solution shows that a generalization of our scheme to more than one exponential function is mandatory.

7.3 Numerical examples for the HMM

Let us now turn to the macroscopic Maxwell and the recursive FE-HMM. We start with a reformulation of Maxwell systems in dimensionless form. Then, in Section 7.3.2 we show a numerical experiment with exact solution in the purely periodic setting based on the layer parameter from the previous section. The HMM error is examined in Section 7.3.3.

7.3.1 Dimensionless Maxwell system

In our numerical experiments we consider various Maxwell systems where we always face the problem of possibly very small or big parameter values. For example, the permittivity of free space, which enters the

Maxwell system is of order 1×10^{-12} . Since these small values could lead to precision errors we derive a rescaled system, which is also dimensionless. In our numerical examples we consider either the Maxwell system with conductivity (3.10) with $\mu_r = 1$ or the Maxwell–Debye system introduced in Section 3.1.3. Thus, consider a general Maxwell system with polarization

$$\begin{aligned}\varepsilon_0 \varepsilon \partial_t \mathbf{E}(t, x) + \partial_t \mathbf{P}(t, x) + \sigma \mathbf{E}(t, x) - \operatorname{curl} \mathbf{H}(t, x) &= -\mathbf{J}(t, x), \\ \mu_0 \partial_t \mathbf{H}(t, x) + \operatorname{curl} \mathbf{E}(t, x) &= \mathbf{0}.\end{aligned}$$

Recall the definition of the vacuum speed of light and the vacuum impedance

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \approx 3 \times 10^8 \text{ m s}^{-1}, \quad Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} \approx 120\pi \text{ V A}^{-1}.$$

For a given unit electric field strength E_0 and a reference length scale L we introduce the following unit-free variables

$$\begin{aligned}\bar{t} &= \frac{ct}{L}, & \bar{x} &= \frac{x}{L}, & \bar{\sigma} &= \sigma LZ_0, \\ \bar{\mathbf{E}}(\bar{t}, \bar{x}) &= \frac{1}{E_0} \mathbf{E}(t, x), & \bar{\mathbf{H}}(\bar{t}, \bar{x}) &= \frac{Z_0}{E_0} \mathbf{H}(t, x), & \bar{\mathbf{P}}(\bar{t}, \bar{x}) &= \frac{cZ_0}{E_0} \mathbf{P}(t, x), & \bar{\mathbf{J}}(\bar{t}, \bar{x}) &= \frac{LZ_0}{E_0} \mathbf{J}(t, x).\end{aligned}$$

From these variables we get the dimensionless Maxwell system as

$$\begin{aligned}\varepsilon \partial_{\bar{t}} \bar{\mathbf{E}}(\bar{t}, \bar{x}) + \partial_{\bar{t}} \bar{\mathbf{P}}(\bar{t}, \bar{x}) + \bar{\sigma} \bar{\mathbf{E}}(\bar{t}, \bar{x}) - \operatorname{curl}_{\bar{x}} \bar{\mathbf{H}}(\bar{t}, \bar{x}) &= -\bar{\mathbf{J}}(\bar{t}, \bar{x}), \\ \partial_{\bar{t}} \bar{\mathbf{H}}(\bar{t}, \bar{x}) + \operatorname{curl}_{\bar{x}} \bar{\mathbf{E}}(\bar{t}, \bar{x}) &= \mathbf{0}.\end{aligned}$$

Let us also consider the case of a Debye polarization, i.e.,

$$\partial_t \mathbf{P}(t, x) + \frac{1}{\tau} \mathbf{P}(t, x) - \varepsilon_0 \frac{\Delta \varepsilon}{\tau} \mathbf{E}(t, x) = \mathbf{0}.$$

With the scaled variables from above we find the dimensionless counterpart as

$$\partial_{\bar{t}} \bar{\mathbf{P}}(\bar{t}, \bar{x}) + \frac{1}{\bar{\tau}} \bar{\mathbf{P}}(\bar{t}, \bar{x}) - \frac{\Delta \varepsilon}{\bar{\tau}} \bar{\mathbf{E}}(\bar{t}, \bar{x}) = \mathbf{0},$$

where the scaled relaxation time is defined as

$$\bar{\tau} = \frac{\tau c}{L}.$$

Observe that the quantities ε and $\Delta \varepsilon$ are already unit-free and thus no rescaling is needed for those. At the end of this short section on scaling we present the both dimensionless Maxwell system that are used in the remainder. Here, we write all rescaled components without the bar. On the one hand we consider the dimensionless Maxwell system with conductivity

$$\varepsilon \partial_t \mathbf{E}(t, x) + \sigma \mathbf{E}(t, x) - \operatorname{curl} \mathbf{H}(t, x) = -\mathbf{J}(t, x), \quad (7.7a)$$

$$\partial_t \mathbf{H}(t, x) + \operatorname{curl} \mathbf{E}(t, x) = \mathbf{0}, \quad (7.7b)$$

and on the other hand the unit-free Maxwell–Debye system

$$\varepsilon \partial_t \mathbf{E}(t, x) + \frac{\Delta \varepsilon}{\tau} \mathbf{E}(t, x) - \frac{1}{\tau} \mathbf{P}(t, x) + \sigma \mathbf{E}(t, x) - \operatorname{curl} \mathbf{H}(t, x) = -\mathbf{J}(t, x), \quad (7.8a)$$

$$\partial_t \mathbf{P}(t, x) + \frac{1}{\tau} \mathbf{P}(t, x) - \frac{\Delta \varepsilon}{\tau} \mathbf{E}(t, x) = \mathbf{0}, \quad (7.8b)$$

$$\partial_t \mathbf{H}(t, x) + \operatorname{curl} \mathbf{E}(t, x) = \mathbf{0}. \quad (7.8c)$$

These two dimensionless Maxwell systems will be used in the remaining part of this chapter.

7.3.2 A test case with exact solution

The first macroscopic example we consider is that of a layered material, which only varies in one space dimension. Representative for this case is the first parameter presented in Section 7.2.1. As shown, this situation allows to compute the effective parameters analytically. Moreover, in this setting we already saw that the convolution kernel is indeed an exponential and thus the approach via recursive convolution is perfectly justified. A drawback of this example is that the parameter is discontinuous and thus the micro error estimate does not hold true. Nevertheless, we still see the convergence rates on the macroscopic level.

Let us give the configuration of the present example. We consider the domain $\Omega = (0, 1)^3$ and use perfectly conducting boundary conditions. On the unit cell $Y = (-\frac{1}{2}, \frac{1}{2})^3$ we define the parameter as before

$$\varepsilon(y) = \begin{cases} 1, & y_1 < 0 \\ 2, & \text{else} \end{cases}, \quad \sigma(y) = \begin{cases} 1, & y_1 < 0 \\ 4, & \text{else} \end{cases}, \quad \mu(y) = 1.$$

We extend these parameters periodically to the whole \mathbb{R}^3 which yields 1-periodic functions. The highly oscillatory parameters are given as usual by

$$\varepsilon^\delta(x) = \varepsilon\left(\frac{x}{\delta}\right), \quad \sigma^\delta(x) = \sigma\left(\frac{x}{\delta}\right) \quad \text{for } x \in \mathbb{R}^3.$$

With these definitions the heterogeneous Maxwell system with conductivity reads

$$\begin{aligned} \varepsilon^\delta(x) \partial_t \mathbf{E}^\delta(t, x) + \sigma^\delta(x) \mathbf{E}^\delta(t, x) - \operatorname{curl} \mathbf{H}^\delta(t, x) &= -\mathbf{J}(t, x), \\ \partial_t \mathbf{H}^\delta(t, x) + \operatorname{curl} \mathbf{E}^\delta(t, x) &= \mathbf{0}, \\ \operatorname{div} (\varepsilon^\delta(x) \mathbf{E}^\delta(t, x)) &= 0, \\ \operatorname{div} \mathbf{H}^\delta(t, x) &= 0. \end{aligned}$$

As shown in Chapter 4 the effective system becomes

$$\varepsilon^{\text{eff}} \partial_t \mathbf{E}^{\text{eff}}(t) + \sigma^{\text{eff}} \mathbf{E}^{\text{eff}}(t) + \int_0^t \mathbf{G}^{\text{eff}}(t-s) \mathbf{E}^{\text{eff}}(s) ds - \operatorname{curl} \mathbf{H}^{\text{eff}}(t) = -\mathbf{J}(t), \quad (7.9a)$$

$$\partial_t \mathbf{H}^{\text{eff}}(t) + \operatorname{curl} \mathbf{E}^{\text{eff}}(t) = \mathbf{0}, \quad (7.9b)$$

where the parameters are given as

$$\varepsilon^{\text{eff}} = \begin{pmatrix} \frac{4}{3} & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \frac{3}{2} \end{pmatrix}, \quad \sigma^{\text{eff}} = \begin{pmatrix} \frac{16}{9} & 0 & 0 \\ 0 & \frac{5}{2} & 0 \\ 0 & 0 & \frac{5}{2} \end{pmatrix}, \quad \mathbf{G}^{\text{eff}}(t) = \begin{pmatrix} \frac{-8}{27} e^{-\frac{5t}{3}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (7.10)$$

The right-hand side in (7.9) is defined as

$$\mathbf{J}(t, x) = \mathbf{f}(t, x) + \mathbf{J}^{\text{eff}}(t) \mathbf{E}_0^{\text{eff}}(x),$$

with given current density \mathbf{f} and extra source \mathbf{J}^{eff} as

$$\mathbf{J}^{\text{eff}}(t) = \begin{pmatrix} \frac{2}{9} e^{-\frac{5t}{3}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In this case we found that the effective convolution kernel \mathbf{G}^{eff} is indeed an exponential function. Moreover, the only nonzero component is $\mathbf{G}_{11}^{\text{eff}}$. Thus, by setting

$$\mathbf{P}^{\text{eff}}(t) = \int_0^t \mathbf{G}^{\text{eff}}(t-s) \mathbf{E}^{\text{eff}}(s) \, ds,$$

and differentiating we find an ODE for the first component of the polarization. The other components vanish, i.e.,

$$\partial_t \mathbf{P}_1^{\text{eff}}(t) = -\frac{5}{3} \mathbf{P}_1^{\text{eff}}(t) - \frac{8}{27} \mathbf{E}_1^{\text{eff}}(t), \quad \mathbf{P}_1^{\text{eff}}(0) = 0, \quad \mathbf{P}_2^{\text{eff}}(t) \equiv \mathbf{P}_3^{\text{eff}}(t) \equiv 0.$$

This definition of the polarization is similar to the one in (3.13) originating from the Debye model. Substituting this in the homogeneous Maxwell system (7.9) yields

$$\varepsilon^{\text{eff}} \partial_t \mathbf{E}^{\text{eff}}(t) + \sigma^{\text{eff}} \mathbf{E}^{\text{eff}}(t) + \mathbf{P}^{\text{eff}}(t) - \text{curl} \mathbf{H}^{\text{eff}}(t) = -\mathbf{J}(t), \quad (7.11a)$$

$$\partial_t \mathbf{P}_1^{\text{eff}}(t) + \frac{5}{3} \mathbf{P}_1^{\text{eff}}(t) + \frac{8}{27} \mathbf{E}_1^{\text{eff}}(t) = \mathbf{0}, \quad (7.11b)$$

$$\partial_t \mathbf{H}^{\text{eff}}(t) + \text{curl} \mathbf{E}^{\text{eff}}(t) = \mathbf{0}. \quad (7.11c)$$

The two equivalent systems (7.9) and (7.11) are prototypes for the recursive FE-HMM from Section 6.3 and the memory variable approach in Section 6.2.3. Therefore, we may compare both methods. Since we know the convolution kernel exactly, we are also able to construct an exact solution, which is given as

$$\mathbf{E}^{\text{eff}}(t, x) = \begin{pmatrix} \mathbf{E}_1(t) \cos(2\pi x) \sin(2\pi y) \sin(2\pi z) \\ \mathbf{E}_2(t) \sin(2\pi x) \cos(2\pi y) \sin(2\pi z) \\ \mathbf{E}_3(t) \sin(2\pi x) \sin(2\pi y) \cos(2\pi z) \end{pmatrix}, \quad \mathbf{H}^{\text{eff}}(t, x) = \begin{pmatrix} \mathbf{H}_1(t) \sin(2\pi x) \cos(2\pi y) \cos(2\pi z) \\ \mathbf{H}_2(t) \cos(2\pi x) \sin(2\pi y) \cos(2\pi z) \\ \mathbf{H}_3(t) \cos(2\pi x) \cos(2\pi y) \sin(2\pi z) \end{pmatrix}.$$

This is the cavity solution constructed such that the perfectly conducting boundary conditions are satisfied. The exact definition of this solution is given in the Appendix A.4. As final time, we choose $T = 5$.

Figure 7.17 shows the maximal L^2 -error on the macro scale for both, the recursive convolution approach and the memory variable method. The proposed first convergence order is observed for both schemes. Note that we used the effective parameters for the computation and thus, no micro error is present. In fact this experiment is dedicated more to the recursive convolution approach than to the HMM part of our scheme.

The time discretization is probably also fine enough as seen in Figure 7.18. Here we show the difference in the time discretization, again for the two effective systems (7.9) and (7.11). We observe that both methods perform similarly and show second order of convergence. For the classical Crank–Nicolson scheme applied to system (7.11) this is expected. For the recursive Crank–Nicolson scheme applied to the integro-differential system (7.9) this is not obvious. As mentioned in Chapter 6 there are results for the recursive time integration but to our knowledge the Crank–Nicolson scheme applied to a first order integro-differential Maxwell system has not been analyzed.

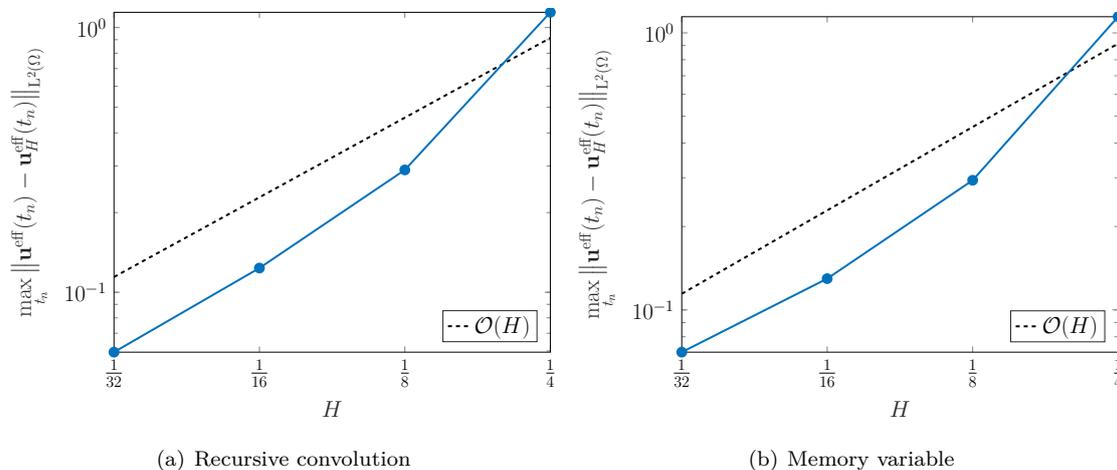


Figure 7.17: Maximal L²-error between effective solution and macro approximation for fixed time step $\Delta t = \frac{5}{512}$ and different macro discretizations H .

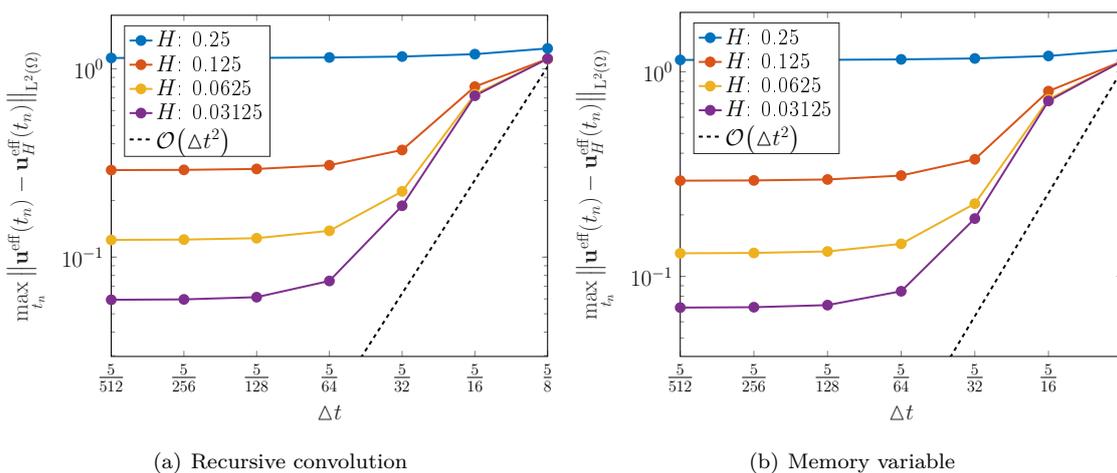


Figure 7.18: Maximal L²-error for various fixed macro discretizations and varying time step size.

7.3.3 Micro error analysis and locally periodic setting

In this section we study the micro error that is analyzed in Lemmas 5.2.3, 5.3.2, 5.3.7 and 5.3.9. The microscopic structures considered here all result from the parameters introduced in Section 7.2.1. For the computations we use the computational domain $\Omega = (0, 1)^3$ and $\delta = 2^{-6}$. The initial value is chosen as

$$\mathbf{E}^{\text{HMM}}(0, x) = \frac{1}{2} \begin{pmatrix} 2 \cos(2\pi x) \sin(2\pi y) \sin(2\pi z) \\ -3 \sin(2\pi x) \cos(2\pi y) \sin(2\pi z) \\ 1 \sin(2\pi x) \sin(2\pi y) \cos(2\pi z) \end{pmatrix}, \quad \mathbf{H}^{\text{HMM}}(0, x) = \mathbf{0}_3.$$

The heterogeneous system in whose solution we are interested in is

$$\begin{aligned} \varepsilon^\delta(x) \partial_t \mathbf{E}^\delta(t, x) + \sigma^\delta(x) \mathbf{E}^\delta(t, x) - \text{curl } \mathbf{H}^\delta(t, x) &= \mathbf{0}, \\ \partial_t \mathbf{H}^\delta(t, x) + \text{curl } \mathbf{E}^\delta(t, x) &= \mathbf{0}, \end{aligned}$$

where we point out that the current density is zero. Thus, the energy of this system decays as shown in Section 3.2.3. As an approximation we solve for the HMM solution that is given as

$$\begin{aligned} \varepsilon^{\text{HMM}} \partial_t \mathbf{E}^{\text{HMM}}(t) + \sigma^{\text{HMM}} \mathbf{E}^{\text{HMM}}(t) + \int_0^t \mathbf{G}^{\text{HMM}}(t-s) \mathbf{E}^{\text{HMM}}(s) \, ds - \text{curl } \mathbf{H}^{\text{HMM}}(t) &= -\mathbf{J}^{\text{HMM}}(t) \mathbf{E}_0, \\ \partial_t \mathbf{H}^{\text{HMM}}(t) + \text{curl } \mathbf{E}^{\text{HMM}}(t) &= \mathbf{0}. \end{aligned}$$

This system involves an extra right-hand side due to the homogenization. Nevertheless, in the following examples we observe that the energy of the HMM system is again decaying. Furthermore, we present results on the actual HMM error.

Continuous micro structure

Let us consider a microscopic structure that consists of the continuous parameter from Section 7.2.1. Thus, as in the previous section we define the heterogeneous parameters ε^δ and σ^δ as

$$\varepsilon^\delta(x) = \varepsilon \left(\frac{x}{\delta} \right), \quad \sigma^\delta(x) = \sigma \left(\frac{x}{\delta} \right) \quad \text{for } x \in \mathbb{R}^3,$$

where ε and σ are the periodic extensions of the parameters defined in (7.5) on the unit cell Y . In this setting the parameters are regular enough such that the results from Chapter 5 hold true. The final computation time is $T = 1$ with $\Delta t = 2^{-8}$. In Figure 7.19(a), we show the evolution of the energy over time and observe the decay of the aforementioned.

Thus, in this example the HMM system seems to be dissipative. In Corollary 4.4.22, we did not see this behavior and therefore the question arises how to improve the stability estimate. In Figure 7.19(b) and Table 7.9 we show the HMM error, which behaves not as proposed, but better.

In fact, we observe second order of convergence in $\frac{1}{\delta}$, which is not covered in our semi-discrete error estimate in Theorem 5.3.23. We already discussed this gap in Section 5.3.2 where we analyzed the micro error of the convolution kernel. In this numerical result we see why it is reasonable to expect higher order convergence also for this time dependent parameter.

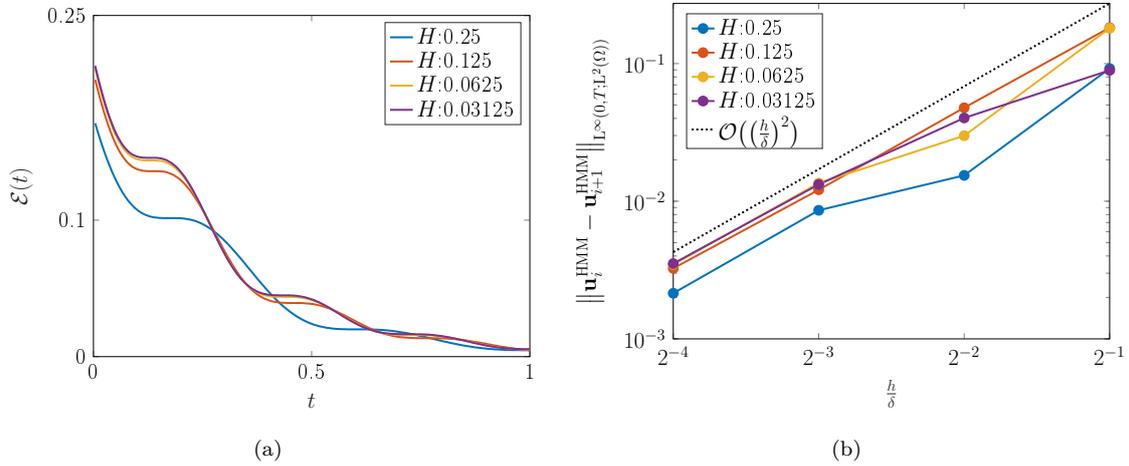


Figure 7.19: (a) Energy of the solution to the continuous microstructure. (b) The HMM error for the continuous microstructure. We show different macroscopic discretizations and the convergence in $\frac{h}{\delta}$ with fixed $\Delta t = 2^{-8}$.

	$\ \mathbf{u}_i^{\text{HMM}} - \mathbf{u}_{i+1}^{\text{HMM}}\ _{L^\infty(0,T;L^2(\Omega))}$							
	$H = 0.25$		$H = 0.125$		$H = 0.0625$		$H = 0.03125$	
$\frac{h}{\delta}/i$	EOC		EOC		EOC		EOC	
$\frac{1}{2}/1$	0.09205	-	0.18232	-	0.18058	-	0.08934	-
$\frac{1}{4}/2$	0.01540	2.58	0.04778	1.93	0.02993	2.59	0.04027	1.15
$\frac{1}{8}/3$	0.00859	0.84	0.01211	1.98	0.01354	1.14	0.01324	1.60
$\frac{1}{16}/4$	0.00215	2.00	0.00325	1.90	0.00347	1.96	0.00353	1.91

Table 7.9: Estimated order of convergence for a microscopic structure consisting of the continuous parameter. The time step is $\Delta t = 2^{-8}$. We observe second order of convergence.

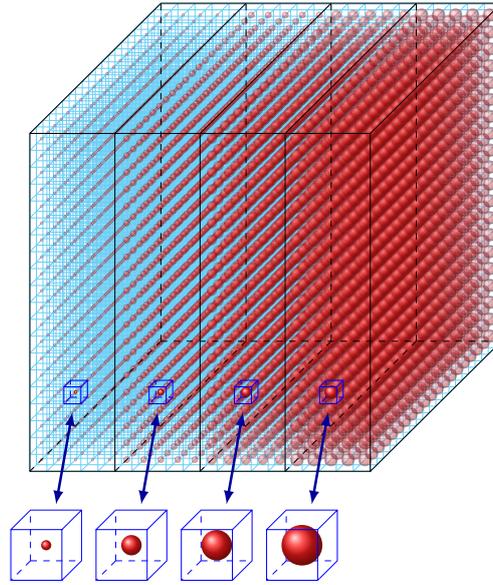


Figure 7.20: Configuration with locally periodic inclusions.

Circular inclusion

The second example we consider is a locally periodic micro structure given by the circular inclusion (7.6) for varying radii. We divide the domain Ω in four slices in x_1 -direction. In every of these subsets we use a different radius for the inclusions on the microscopic scale. To be exact, we use

$$r(x) = \begin{cases} 0.1, & 0 \leq x_1 < \frac{1}{4} \\ 0.2, & \frac{1}{4} \leq x_1 < \frac{1}{2} \\ 0.3, & \frac{1}{2} \leq x_1 < \frac{3}{4} \\ 0.4, & \frac{3}{4} \leq x_1 \leq 1 \end{cases},$$

combined with the definition of the circular inclusion in (7.6), which we extend by periodicity to \mathbb{R}^3 . A sketch of the configuration is shown in Figure 7.20. The heterogeneous parameters are thus for every $x \in \Omega$ given as

$$\varepsilon^\delta(x) = \varepsilon\left(x, \frac{x}{\delta}\right), \quad \sigma^\delta(x) = \sigma\left(x, \frac{x}{\delta}\right).$$

Note that the effective parameters and their HMM approximations now depend on the macroscopic space variable as well, i.e., for $x \in \Omega$ and $t \in [0, T]$ we find the dependencies

$$\varepsilon^{\text{HMM}}(x), \quad \sigma^{\text{HMM}}(x), \quad \mathbf{G}^{\text{HMM}}(t, x), \quad \mathbf{J}^{\text{HMM}}(t, x).$$

In the setting of a locally periodic parameter we may use the parallel assembly of the macroscopic matrices as explained in Section 7.1.1. In this example, however, we pre-computed the parameters for the four different slices. The final time of computation is again chosen to be $T = 1$ and we used step size $\Delta t = 2^{-9}$. In this setting the micro error results from Chapter 5 do not hold true since the parameter is discontinuous.

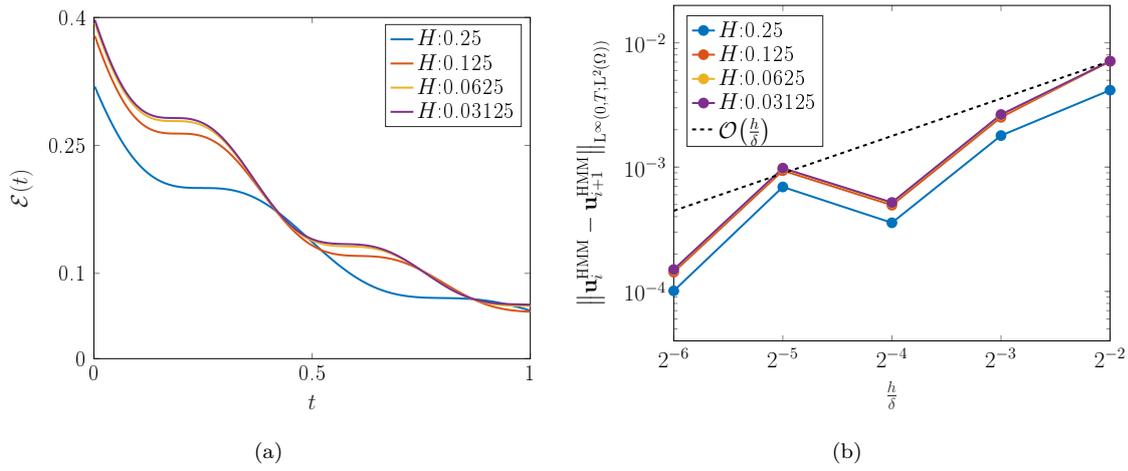


Figure 7.21: (a) Energy of the solution to the locally periodic circular inclusion. (b) HMM error for fixed time step $\Delta t = 2^{-9}$ and various macroscopic discretization sizes H .

	$\ \mathbf{u}_i^{\text{HMM}} - \mathbf{u}_{i+1}^{\text{HMM}}\ _{L^\infty(0,T;L^2(\Omega))}$							
	$H = 0.25$		$H = 0.125$		$H = 0.0625$		$H = 0.03125$	
$\frac{h}{\delta}/i$	EOC		EOC		EOC		EOC	
$\frac{1}{4}/1$	0.00416	-	0.00713	-	0.00709	-	0.00712	-
$\frac{1}{8}/2$	0.00179	1.21	0.00252	1.5	0.00262	1.43	0.00265	1.42
$\frac{1}{16}/3$	0.00036	2.33	0.00049	2.35	0.00051	2.35	0.00052	2.35
$\frac{1}{32}/4$	0.00069	-0.96	0.00094	-0.92	0.00097	-0.92	0.00098	-0.92
$\frac{1}{64}/5$	0.00010	2.78	0.00014	2.71	0.00015	2.71	0.00015	2.71

Table 7.10: Estimated order of convergence for a micro structure consisting of locally periodic circular inclusions. The time step is $\Delta t = 2^{-9}$.

In Figure 7.21(a), we show the energy of the solution over time for various macroscopic space discretizations and fixed $h = 2^{-7}$. Again we observe the decay that is not covered in our results. The HMM error is shown in Figure 7.21(b) and Table 7.10. These results indicate first order of convergence in $\frac{h}{\delta}$. Even with an order reduction due to the discontinuity our results do not cover this setting. This is again due to the non-optimal first order estimate for the time dependent parameters.

CHAPTER 8

Conclusion and Outlook

In the final chapter of this thesis we recapture our results and give an outlook for future research. We presented and analyzed a numerical method to model electromagnetic wave propagation in heterogeneous, (locally) periodic media. The focus within our analysis and the numerical scheme are dispersive models, that we incorporate already in the heterogeneous structure. The main contribution of this thesis is the semi-discrete error analysis of the Heterogeneous Multiscale Method (HMM) applied to a homogenized time-dependent Maxwell system. For that purpose we derived central properties of the effective parameters, showed the wellposedness of the effective and the HMM system and proved micro error estimates. Moreover, we proposed a fully discrete scheme to solve three-dimensional Maxwell's equations. This scheme incorporates effective dispersive material properties. This is, to our knowledge, the first method to cover these effects. A central assumption for this method is the exponential structure of the convolution kernel. Although this structure is not guaranteed, we showed the special case of a positive definite damping parameter that ensures this exponential decay. Moreover, we performed several experiments with different heterogeneities and presented that the resulting effective kernels may be accurately approximated at least by a linear combination of exponential functions. Finally, we tested the fully discrete scheme in several numerical examples. Thus, this thesis demonstrates that the abstract homogenization results may indeed be efficiently implemented in a numerical scheme to simulate the effective material behavior of heterogeneous materials with dispersive effects.

Concerning possible extensions of our results we start with the stability estimates for the integro-differential Maxwell system in Theorem 4.4.21 and Corollary 4.4.22. Those estimates seem to be non-optimal. At least for the examples we considered, the observed growth was much slower than predicted by our results. From the physical point of view, a related question is whether the mixture of stable materials may yield an unstable effective behavior. The next non-optimal result are the micro error estimate in Lemmas 5.3.7 and 5.3.9 for the time dependent parameters. We already discussed that the first order convergence is not satisfactory, which is also supported by our numerical experiments that suggest second order convergence. A better understanding of the cell problems, i.e., the Sobolev equation might be

helpful to derive a better estimate. Most importantly, the fully discrete error analysis has to be carried out. The major challenge probably is the estimation of the error due to the exponential fitting, which is related to the field of model error estimation. Apart from that, the analysis of either the recursive convolution or the memory variable method combined with the HMM should be possible using the techniques discussed in Section 6.3.2. Finally, the proposed scheme has to be tested on more powerful computers for considerably larger problems. For realistic simulations physically relevant boundary conditions have to be implemented, for e.g. the perfectly matched layers (PML) (Berenger, 1994).

APPENDIX A

Appendix

A.1 The Debye model

We consider dielectric materials, that are insulators which can be polarized by an electric field. More precisely, an applied electric field can effect the material in two ways, first it could induce electrical dipoles that do not exist without the field and try to align them in field direction. Second it aligns already existing dipoles. The combination of both effects is called the polarization of the material. In experiments it has been observed that the underlying law between the polarization and the applied electric field can in many cases be approximated by a linear material law, i.e.

$$\mathbf{P} = \varepsilon_0 \chi \mathbf{E}, \quad (\text{A.1})$$

with the permittivity constant of vacuum ε_0 and a material parameter χ , called the dielectric susceptibility, depending on the frequency ω . In addition to the electric field strength \mathbf{E} we have a relation between the polarization and the electric displacement \mathbf{D} that is given by the constitutive relation

$$\mathbf{D} = \varepsilon_0 \varepsilon_r \mathbf{E}, \quad (\text{A.2})$$

where ε_r is called the relative dielectric constant of the material and $\varepsilon_0 \varepsilon_r$ is called the permittivity. The well-known material law we need is

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} = \varepsilon_0 (1 + \chi) \mathbf{E}. \quad (\text{A.3})$$

From (A.2) and (A.3) we find

$$\varepsilon_r = 1 + \chi,$$

and note here that these parameters depend on the frequency ω of the electric field and thus we call $\varepsilon_r(\omega)$ the dielectric function. In order to determine the susceptibility χ we have to consider models for polarization.

A.1.1 Mechanisms of polarization

There are four kinds of polarization mechanisms:

1. Electronic polarization,
2. Orientation polarization,
3. Ionic polarization,
4. Interface polarization.

The first two concepts will be discussed in the following.

Electronic polarization

The electronic polarization describes the effect of an applied electric field to a single atom. It is the displacement of the center of charges of electrons with respect to the nucleus and this polarization effect occurs in every atom and molecule. The dipole moment \mathbf{p}_{ep} induced in an atom where we assume for simplicity that the charges are distributed uniformly in a sphere of radius R around the nucleus is

$$\mathbf{p}_{ep} = 4\pi\varepsilon_0 R^3 \mathbf{E}.$$

The polarizability α_{ep} is defined as the induced dipole moments per unit electric field intensity and we thus have

$$\alpha_{ep} = 4\pi\varepsilon_0 R^3. \quad (\text{A.4})$$

A central equality we need is the connection between the polarization \mathbf{P} , the dipole moment per unit volume, and the polarizability α_{ep} . This is obtained by observing that the dipole moment of an atom is given by $\alpha_{ep}\mathbf{E}$ and denoting with N the number of atoms per unit volume. We find the connection as

$$\mathbf{P} = N\alpha_{ep}\mathbf{E}, \quad (\text{A.5})$$

and inserting (A.1) and (A.4) gives a first definition of the electric susceptibility and with it of the relative permittivity

$$\chi = 4\pi NR^3, \quad \varepsilon_r = 1 + 4\pi NR^3.$$

The problem with this approach is, that we assume that neighboring molecules do not affect each other but this is wrong, for example for high pressures. One has to consider the internal field \mathbf{E}_i , also known as Lorentz field, and the corresponding dipole moment $\mathbf{p}_i = \alpha_{ep}\mathbf{E}_i$. This is shown in (Raju, 2003, Section 2.3). The result are the equations

$$\mathbf{P} = N\alpha_{ep}\mathbf{E}_i, \quad \mathbf{E}_i = \frac{2 + \varepsilon_r}{3} \mathbf{E}.$$

The combination with the definition of \mathbf{P} from (A.1) yields

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{N\alpha_{ep}}{3\varepsilon_0}.$$

Introducing the molar volume V , given by $\frac{M}{\rho}$, where M is the molecular weight, ρ the density and the Avogadro number $N_A = NV$ we end up with the Clausius-Mossotti equation for molar polarizability P_m given as

$$P_m = \frac{\varepsilon_r - 1}{\varepsilon_r + 2} \frac{M}{\rho} = \frac{N_A \alpha_{ep}}{3\varepsilon_0}.$$

This equation describes the effect of electronic polarization very well as long as we consider non-polar materials without permanent dipole moments. To handle permanent dipoles we have to consider orientation polarization, which is done in the next section.

Orientation polarization

We consider materials which can be affected by orientation polarization and thus already have dipoles, which get oriented by the electric field. We look at a material with existing dipoles which can rotate freely, the main example for that is liquid water since every H₂O molecule is a dipole with random orientation with respect to the other molecules. The polarization of such a material is zero in the absence of an electric field. The question is how an electric field changes the polarization. Let us start with the potential energy U of a dipole in an electric field

$$U = -\mathbf{p} \cdot \mathbf{E},$$

where again \mathbf{p} is the dipole moment. Denoting with θ the angle between the dipole \mathbf{p} and the field \mathbf{E} we get

$$U = -|\mathbf{p}||\mathbf{E}| \cos(\theta).$$

How does the distribution of the dipoles due to the electric field look like? For this purpose consider a solid angle v that is formed between θ and $\theta + \delta$. Furthermore we introduce the temperature T , the Boltzmann constant k and a constant N_0 which depends on the total number of dipoles. Now the number of dipoles confined to the solid angle v is given by the Boltzmann distribution as

$$N(\theta) = N_0 \exp\left(\frac{|\mathbf{p}||\mathbf{E}| \cos(\theta)}{kT}\right) v. \quad (\text{A.6})$$

The surface area between the angles θ and $\theta + \delta$ on a sphere of radius r is

$$S = 2\pi r^2 \sin(\theta) \delta,$$

and with this we define the solid angle by

$$v = \frac{S}{r^2} = 2\pi \sin(\theta) \delta.$$

Substituting this in (A.6) we find

$$N(\theta) = N_0 \exp\left(\frac{|\mathbf{p}||\mathbf{E}| \cos(\theta)}{kT}\right) 2\pi \sin(\theta) \delta.$$

A dipole of permanent moment \mathbf{p} making an angle θ with the direction of the electric field contributes a moment $p_E = |\mathbf{p}| \cos(\theta)$ and thus the contribution of all dipoles in v is

$$p_E(\theta) = N(\theta) p_E = N(\theta) |\mathbf{p}| \cos(\theta).$$

We now get the average dipole moment \bar{p}_E per dipole in the direction of the electric field by the ratio of the dipole moment due to all molecules divided by the number of dipoles

$$\bar{p}_E = \frac{\int_0^\pi p_E(\theta) d\theta}{\int_0^\pi N(\theta) d\theta} = \frac{\int_0^\pi N(\theta) |\mathbf{p}| \cos(\theta) d\theta}{\int_0^\pi N(\theta) d\theta} = \frac{\int_0^\pi \exp\left(\frac{|\mathbf{p}||\mathbf{E}| \cos(\theta)}{kT}\right) 2\pi \sin(\theta) \delta |\mathbf{p}| \cos(\theta) d\theta}{\int_0^\pi \exp\left(\frac{|\mathbf{p}||\mathbf{E}| \cos(\theta)}{kT}\right) 2\pi \sin(\theta) \delta d\theta}.$$

We substitute $\frac{|\mathbf{p}||\mathbf{E}|}{kT} = x$ as well as $y = \cos(\theta)$ and obtain

$$\frac{\bar{p}_E}{|\mathbf{p}|} = \frac{\int_{-1}^1 y \exp(yx) dy}{\int_{-1}^1 \exp(yx) dy} = \left(\coth(x) - \frac{1}{x} \right) = L(x),$$

with the Langevin function $L(x)$. The values of $x = \frac{|\mathbf{p}||\mathbf{E}|}{kT}$ are usually small, i.e. $x < 1$ for real materials. For these small values we use the linear part of the Taylor expansion around zero to approximate the Langevin function

$$L(x) \approx \frac{1}{3}x,$$

and this yields

$$\bar{p}_E = |\mathbf{p}|L(x) \approx \frac{|\mathbf{p}|}{3}x = \frac{|\mathbf{p}|}{3} \frac{|\mathbf{p}||\mathbf{E}|}{kT} = \frac{|\mathbf{p}|^2|\mathbf{E}|}{3kT}.$$

We now introduce the polarizability α_0 due to orientation polarization by the equality $\alpha_0 = \frac{\bar{p}_E}{|\mathbf{E}|}$ and find

$$\alpha_0 = \frac{|\mathbf{E}|^2}{3kT},$$

and with this size we can define the orientation polarization of the dielectric as

$$\mathbf{P} = N\alpha_0\mathbf{E} = \frac{N|\mathbf{p}|^2}{3kT}\mathbf{E}. \quad (\text{A.7})$$

A.1.2 The model

The combination of the two previous sections, more precisely the equations (A.5) and (A.7), leads to the Debye equation

$$\mathbf{P} = N \left(\alpha_{ep} + \frac{|\mathbf{P}|^2}{3kT} \right) \mathbf{E}.$$

The molar polarizability now follows as in Section A.1.1

$$P_m = \frac{\varepsilon_r - 1}{\varepsilon_r + 2} \frac{M}{\varrho} = \frac{N_A}{3\varepsilon_0} \left(\alpha_{ep} + \frac{|\mathbf{P}|^2}{3kT} \right).$$

The frequency depended permittivity

The dependence of the molar polarization on the frequency is given in Debye (1929) equation (64) as

$$\frac{\hat{\varepsilon}_r(\omega) - 1}{\hat{\varepsilon}_r(\omega) + 2} \cdot \frac{M}{\varrho} = P_m(\omega) = \frac{N_A}{3\varepsilon_0} \left[\alpha_{ep} + \frac{|\mathbf{P}|^2}{3kT} \cdot \frac{1}{1 + i\omega\tau} \right], \quad (\text{A.8})$$

and this defines the frequency dependent relative permittivity by

$$\hat{\varepsilon}_r(\omega) = \frac{1 + 2\frac{\varrho}{M}P_m(\omega)}{1 - \frac{\varrho}{M}P_m(\omega)}. \quad (\text{A.9})$$

Now we introduce two dielectric constants, namely ε_∞ and ε_s that are the dielectric constants at very high ($\omega \rightarrow \infty$) and at zero frequency ($\omega = 0$), respectively. These constants are determined by

$$\begin{aligned} \frac{\varepsilon_\infty - 1}{\varepsilon_\infty + 2} \cdot \frac{M}{\varrho} &= \frac{N_A}{3\varepsilon_0} \alpha_{ep} \\ \frac{\varepsilon_s - 1}{\varepsilon_s + 2} \cdot \frac{M}{\varrho} &= \frac{N_A}{3\varepsilon_0} \left[\alpha_{ep} + \frac{|\mathbf{P}|^2}{3kT} \right]. \end{aligned}$$

Inserting these equalities in the representation of the polarization (A.8) yields

$$P_m(\omega) = \frac{M}{\varrho} \left[\frac{\varepsilon_\infty - 1}{\varepsilon_\infty + 2} + \frac{1}{1 + i\omega\tau} \left(\frac{\varepsilon_s - 1}{\varepsilon_s + 2} - \frac{\varepsilon_\infty - 1}{\varepsilon_\infty + 2} \right) \right].$$

The next step is to combine the equation (A.9) with the just gained representation of the polarization. We finally get

$$\widehat{\varepsilon}_r(\omega) = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + i\omega\tau \frac{\varepsilon_s + 2}{\varepsilon_\infty + 2}}. \quad (\text{A.10})$$

This is almost the well known equation of the Debye relative permittivity in the frequency domain.

Deriving the time domain formulation

The relative permittivity depending on the frequency in the Debye case can be expressed as

$$\widehat{\varepsilon}_r(\omega) = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + i\omega\tau}. \quad (\text{A.11})$$

We start with the frequency formulation of the constitutive law for the electric density

$$\widehat{\mathbf{D}}(\omega) = \varepsilon_0 \widehat{\varepsilon}_r(\omega) \widehat{\mathbf{E}}(\omega) = \varepsilon_0 \left(\varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + i\omega\tau} \right) \widehat{\mathbf{E}}(\omega),$$

and apply the inverse Fourier–Laplace transform to this equation, which is defined as

$$f(t) = \mathcal{F}^{-1}(\widehat{f}(\omega))(t) = \frac{1}{\sqrt{2\pi}} \int \widehat{f}(\omega) \exp(i\omega t) \, d\omega.$$

This results in

$$\begin{aligned} \mathbf{D}(t) &= \left(\mathcal{F}^{-1} \widehat{\mathbf{D}}(\omega) \right) (t) = \varepsilon_0 \mathcal{F}^{-1} \left(\left(\varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{1 + i\omega\tau} \right) \widehat{\mathbf{E}}(\omega) \right) (t) \\ &= \varepsilon_0 \mathcal{F}^{-1} \left(\varepsilon_\infty \widehat{\mathbf{E}}(\omega) \right) (t) + \varepsilon_0 \frac{\varepsilon_s - \varepsilon_\infty}{\tau} \mathcal{F}^{-1} \left(\frac{\tau}{1 + i\omega\tau} \widehat{\mathbf{E}}(\omega) \right) (t) \\ &= \varepsilon_0 \varepsilon_\infty \mathbf{E}(t) + \varepsilon_0 \frac{\varepsilon_s - \varepsilon_\infty}{\tau} \mathcal{F}^{-1} \left(\frac{\tau}{1 + i\omega\tau} \widehat{\mathbf{E}}(\omega) \right) (t). \end{aligned}$$

We use the convolution theorem which states

$$\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g),$$

or equivalently

$$\mathcal{F}^{-1}(\widehat{f}) * \mathcal{F}^{-1}(\widehat{g}) = \mathcal{F}^{-1}(\widehat{f\widehat{g}}),$$

and this gives

$$\begin{aligned} \mathcal{F}^{-1} \left(\frac{\tau}{1 + i\omega\tau} \widehat{\mathbf{E}}(\omega) \right) (t) &= \mathcal{F}^{-1} \left(\frac{\tau}{1 + i\omega\tau} \right) (t) * \mathbf{E}(t) = \exp\left(-\frac{t}{\tau}\right) \Theta(t) * \mathbf{E}(t) \\ &= \int_0^t \exp\left(-\frac{t-s}{\tau}\right) \mathbf{E}(s) \, ds. \end{aligned}$$

We finally get

$$\mathbf{D}(t) = \varepsilon_0 \varepsilon_\infty \mathbf{E}(t) + \varepsilon_0 \frac{\varepsilon_s - \varepsilon_\infty}{\tau} \int_0^t \exp\left(-\frac{t-s}{\tau}\right) \mathbf{E}(s) \, ds, \quad (\text{A.12})$$

and add this constitutive relation to the Maxwell's equations in a domain Ω .

$$\begin{aligned} \partial_t \mathbf{D} &= \operatorname{curl} \mathbf{H}, & \text{in } (0, T) \times \Omega, \\ \mu \partial_t \mathbf{H} &= -\operatorname{curl} \mathbf{E}, & \text{in } (0, T) \times \Omega, \\ \operatorname{div} \mathbf{D} &= 0, & \operatorname{div} \mu \mathbf{H} = 0, & \text{in } (0, T) \times \Omega, \\ \mathbf{E}(0, x) &= \mathbf{E}_0(x), & \mathbf{H}(0, x) &= \mathbf{H}_0(x), & \text{in } \Omega. \end{aligned}$$

Here the parameter μ is the permeability of the material we consider. An alternative form of this polarization model is to write a differential equation for the polarization \mathbf{P} such that the constitutive relation

$$\mathbf{D}(t) = \varepsilon_0 \varepsilon_\infty \mathbf{E}(t) + \mathbf{P}(t), \quad (\text{A.14})$$

corresponds to (A.12). From (A.12) and (A.14) we see that the polarization satisfies

$$\mathbf{P}(t) = \varepsilon_0 \frac{\varepsilon_s - \varepsilon_\infty}{\tau} \int_0^t \exp\left(-\frac{t-s}{\tau}\right) \mathbf{E}(s) \, ds.$$

We use the variation-of-constants formula to see that \mathbf{P} solves

$$\tau \partial_t \mathbf{P}(t) + \mathbf{P}(t) = \varepsilon_0 (\varepsilon_s - \varepsilon_\infty) \mathbf{E}(t),$$

with the starting value $\mathbf{P}(0) = 0$.

A.2 Transformation of cell problems and parameters

Consider the cell problem (4.20) posed on the unit cell $Y = (-\frac{1}{2}, \frac{1}{2})^3$: Find $w_\ell^M(\bar{x}, \cdot) \in \mathbf{H}_\#^1(Y)$ such that

$$\int_Y \mathbf{M}(\bar{x}, y) (\mathbf{e}_k + \nabla_y w_\ell^M(\bar{x}, y)) \cdot \nabla_y v(y) \, dy = 0 \quad \text{for all } v \in \mathbf{H}_\#^1(Y), \quad (\text{A.15})$$

and the corresponding effective parameter, cf. (4.23),

$$(\mathbf{M}^{\text{eff}}(\bar{x}))_{k,\ell} = \int_Y \mathbf{M}(\bar{x}, y) (\mathbf{e}_\ell + \nabla_y w_\ell^M(\bar{x}, y)) \cdot (\mathbf{e}_k + \nabla_y w_k^M(\bar{x}, y)) \, dy.$$

Lemma A.2.1. *Let $\bar{x} \in \Omega$ and consider a sampling domain $Y^\delta(\bar{x})$. The effective parameter is given as*

$$(\mathbf{M}^{\text{eff}}(\bar{x}))_{k,\ell} = \int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{x}{\delta}\right) (\mathbf{e}_\ell + \nabla_x w_\ell^{\mathbf{M},\delta}(\bar{x}, x)) \cdot (\mathbf{e}_k + \nabla_x w_k^{\mathbf{M},\delta}(\bar{x}, x)) \, dx,$$

where $w_\ell^{\mathbf{M},\delta}(\bar{x}, \cdot) \in \mathbf{H}_\#^1(Y^\delta(\bar{x}))$ solves

$$\int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{x}{\delta}\right) (\mathbf{e}_\ell + \nabla_x w_\ell^{\mathbf{M},\delta}(\bar{x}, x)) \cdot \nabla_x v(x) \, dx = 0 \quad \text{for all } v \in \mathbf{H}_\#^1(Y^\delta(\bar{x})).$$

Proof. Extend w_ℓ^M by periodicity, such that $w_\ell^M(\bar{x}, y)$ is defined for all $y \in \mathbb{R}^d$ and set

$$w_\ell^{\mathbf{M},\delta}(\bar{x}, x) := \delta w_\ell^M\left(\bar{x}, \frac{x}{\delta}\right).$$

Thus, $w_\ell^{M,\delta}(\bar{x}, \cdot) \in H_{\#}^1(Y^\delta(\bar{x}))$ and $\nabla w_\ell^{M,\delta}(\bar{x}, x) = \nabla_y w_\ell^M(\bar{x}, \frac{x}{\delta})$. By transformation $y = \frac{x}{\delta}$ we find for all $v \in H_{\#}^1(Y^\delta(\bar{x}))$ the equality

$$\begin{aligned} & \int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{x}{\delta}\right) \left(\mathbf{e}_k + \nabla w_\ell^{M,\delta}(\bar{x}, x)\right) \cdot \nabla v(x) \, dx \\ &= \int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{x}{\delta}\right) \left(\mathbf{e}_k + \nabla_y w_\ell^M\left(\bar{x}, \frac{x}{\delta}\right)\right) \cdot \nabla v(x) \, dx \\ &= \delta^3 \int_{Y^1(\bar{x})} \mathbf{M}(\bar{x}, y) \left(\mathbf{e}_k + \nabla_y w_\ell^M(\bar{x}, y)\right) \cdot \nabla_y v(\delta y) \, dy. \end{aligned}$$

All functions are Y -periodic and therefore it does not matter over which period we integrate. Now (A.15) yields

$$\begin{aligned} & \int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{x}{\delta}\right) \left(\mathbf{e}_k + \nabla w_\ell^{M,\delta}(\bar{x}, x)\right) \cdot \nabla v(x) \, dx \\ &= \delta^3 \int_Y \mathbf{M}(\bar{x}, y) \left(\mathbf{e}_k + \nabla_y w_\ell^M(\bar{x}, y)\right) \cdot \nabla_y v(y) \, dy = 0. \end{aligned}$$

To show the alternative formulation of the effective parameter we proceed in the same way and use that $|Y^\delta(\bar{x})| = \delta^3$

$$\begin{aligned} & \int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{x}{\delta}\right) \left(\mathbf{e}_\ell + \nabla w_\ell^{M,\delta}(\bar{x}, x)\right) \cdot \left(\mathbf{e}_k + \nabla w_k^{M,\delta}(\bar{x}, x)\right) \, dx \\ &= \frac{1}{|Y^\delta(\bar{x})|} \int_{Y^\delta(\bar{x})} \mathbf{M}\left(\bar{x}, \frac{x}{\delta}\right) \left(\mathbf{e}_\ell + \nabla_y w_\ell^M\left(\bar{x}, \frac{x}{\delta}\right)\right) \cdot \left(\mathbf{e}_k + \nabla_y w_k^M\left(\bar{x}, \frac{x}{\delta}\right)\right) \, dx \\ &= \frac{\delta^3}{|Y^\delta(\bar{x})|} \int_{Y^1(\bar{x})} \mathbf{M}(\bar{x}, y) \left(\mathbf{e}_\ell + \nabla_y w_\ell^M(\bar{x}, y)\right) \cdot \left(\mathbf{e}_k + \nabla_y w_k^M(\bar{x}, y)\right) \, dy \\ &= \int_Y \mathbf{M}(\bar{x}, y) \left(\mathbf{e}_\ell + \nabla_y w_\ell^M(\bar{x}, y)\right) \cdot \left(\mathbf{e}_k + \nabla_y w_k^M(\bar{x}, y)\right) \, dy \\ &= (\mathbf{M}^{\text{eff}}(\bar{x}))_{k,\ell}. \end{aligned}$$

□

A.3 Effective parameters of isotropically layered material

We only give the most important steps here and skip most of the computations. To derive the exact formulation of the effective parameters we note that due to the layered structure of the parameter the cell problems reduce to one-dimensional problems on $Y = (-\frac{1}{2}, \frac{1}{2})$. Thus, for $w^\varepsilon \in H_{\#}^1(Y; \mathbb{R})$ consider the micro problem

$$\int_Y \varepsilon(y) (1 + \partial_y w^\varepsilon(y)) \partial_y v(y) \, dy = 0 \quad \text{for all } v \in H_{\#}^1(Y; \mathbb{R}). \quad (\text{A.16})$$

We decompose the unit cell according to the parameter in $Y_1 = (-\frac{1}{2}, 0)$ and $Y_2 = (0, \frac{1}{2})$ and do the same for the solution $w^\varepsilon(y) = \begin{cases} w_1^\varepsilon(y), & y \in Y_1 \\ w_2^\varepsilon(y), & y \in Y_2 \end{cases}$. Thus, we decompose the equation (A.16) accordingly and after an integration by parts we find the representations

$$\partial_y w_1^\varepsilon(y) = \frac{C_1}{\varepsilon_1} - 1, \quad \partial_y w_2^\varepsilon(y) = \frac{C_2}{\varepsilon_2} - 1.$$

The periodicity of the test function v and the continuity of w yield $C_1 = C_2$. Using the periodicity of the cell corrector w we find

$$0 = \int_Y \partial_y w^\varepsilon(y) dy = \int_{Y_1} \partial_y w_1^\varepsilon(y) dy + \int_{Y_2} \partial_y w_2^\varepsilon(y) dy = C_1 \frac{\varepsilon_1 + \varepsilon_2}{2\varepsilon_1\varepsilon_2} - 1,$$

from which we deduce $C_1 = \frac{2\varepsilon_1\varepsilon_2}{\varepsilon_1 + \varepsilon_2}$. The corresponding component of the effective parameter now evaluates as

$$\varepsilon_{11}^{\text{eff}} = \int_Y \varepsilon(y) (1 + \partial_y w^\varepsilon(y)) (1 + \partial_y w^\varepsilon(y)) dy = C_1 = \frac{2\varepsilon_1\varepsilon_2}{\varepsilon_1 + \varepsilon_2}.$$

The effective conductivity in this case is given as

$$\sigma_{11}^{\text{eff}} = \int_Y \sigma(y) (1 + \partial_y w^\varepsilon(y)) (1 + \partial_y w^\varepsilon(y)) dy = \frac{2(\varepsilon_1^2\sigma_2 + \varepsilon_2^2\sigma_1)}{(\varepsilon_1 + \varepsilon_2)^2}.$$

For the initial value $\bar{w}(0, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R})$ of the time dependent corrector the cell problem reads

$$\int_Y (\varepsilon(y) \partial_y \bar{w}(t, y) + \sigma(y) (1 + \partial_y w^\varepsilon(y))) \partial_y v(y) dy = 0 \quad \text{for all } v \in \mathbf{H}_{\#}^1(Y; \mathbb{R}).$$

This is solved using the same techniques as above. The only missing part is the time dependent problem for $\bar{w}(t, \cdot) \in \mathbf{H}_{\#}^1(Y; \mathbb{R})$, which is given as

$$\int_Y (\varepsilon(y) \partial_t \partial_y \bar{w}(t, y) + \sigma(y) \partial_y \bar{w}(t, y)) \partial_y v(y) dy = 0.$$

We again use the decomposition of the domain and the procedure from above. Additionally, we use the variation of constants formula to get

$$\partial_y \bar{w}_i(t, y) = e^{-\frac{\sigma_i}{\varepsilon_i} t} \partial_y \bar{w}_i(0, y) + \int_0^t e^{-\frac{\sigma_i}{\varepsilon_i} (t-s)} C(s) ds, \quad \text{for } i = 1, 2,$$

and a time dependent parameter $C(s)$. Now we again use the periodicity of the cell corrector

$$0 = \int_Y \partial_y \bar{w}(t, y) dy,$$

which results in an expression determining $C(s)$ in an integral equation. With an application of the Laplace transform this equation is solvable for the Laplace transform of $C(s)$. Finally, the inverse Laplace transform yields a representation of the time dependent parameter $C(s)$ and thus of $\partial_y \bar{w}_i(t, y)$ for $i = 1, 2$. This representation may then be used in the definition of the effective convolution kernel to determine

$$\int_Y \sigma(y) \partial_y \bar{w}(t, y) (1 + \partial_y w^\varepsilon(y)) dy = -\frac{2(\varepsilon_1\sigma_2 - \varepsilon_2\sigma_1)^2}{(\varepsilon_1 + \varepsilon_2)^3} e^{-\frac{\sigma_1 + \sigma_2}{\varepsilon_1 + \varepsilon_2} t}.$$

The approach for the extra source follows the same line as for the convolution kernel.

A.4 Exact solution for macroscopic Maxwell with layer material

The time dependent components are computed such that the equations with zero right-hand side are satisfied, i.e., (7.9b), (7.11b), (7.11c). Finally, the right-hand side \mathbf{f} is adjusted such that the equations (7.9a) and (7.11a) hold true. The time dependent components for the exact solution from Section 7.3.2 are given as

$$\begin{aligned} \mathbf{E}_1(t) &:= \frac{-3e^{-t}(48\pi^2 \cos(2\pi t) - 48\pi^2 \sin(2\pi t) + 8\pi \sin(\pi t) \cos(\pi t) + 8\pi \cos(\pi t)^2 - 4\pi e^{-\frac{t}{2}})}{128\pi^2 + 8} \\ &\quad + \frac{-3e^{-t}(-2 \sin(\pi t) \cos(\pi t) + 2 \cos(\pi t)^2 - 4\pi + 3 \cos(2\pi t) - 3 \sin(2\pi t) - e^{-\frac{t}{2}} - 1)}{128\pi^2 + 8}, \\ \mathbf{E}_2(t) &:= \cos(2\pi t)e^{-t}, \\ \mathbf{E}_3(t) &:= -\sin(2\pi t)e^{-t}, \\ \mathbf{H}_1(t) &:= \frac{4\pi(-(\pi + \frac{1}{2})e^{-t} \cos(2\pi t) + e^{-t}(\pi - \frac{1}{2}) \sin(2\pi t) + \pi + \frac{1}{2})}{4\pi^2 + 1}, \\ \mathbf{H}_2(t) &:= \frac{144(\frac{17}{9}\pi(\pi^3 - \frac{21}{68}\pi^2 + \frac{7}{136}\pi - \frac{3}{136})e^{-t} \cos(2\pi t) + e^{-t}\pi(\pi^3 + \frac{37}{36}\pi^2 + \frac{\pi}{24} + \frac{5}{72}) \sin(2\pi t))}{128\pi^4 + 40\pi^2 + 2} \\ &\quad + \frac{144(\frac{(\pi+\frac{1}{4})(\pi^2+\frac{1}{4})\pi}{9}e^{-\frac{3}{2}t} - \frac{10}{9}(\pi^2 + \frac{1}{16})(\pi^2 - \frac{\pi}{2} - \frac{1}{5}))}{128\pi^4 + 40\pi^2 + 2}, \\ \mathbf{H}_3(t) &:= \frac{-272(\frac{9}{17}\pi(\pi^3 - \frac{37}{36}\pi^2 + \frac{\pi}{24} - \frac{5}{72})e^{-t} \cos(2\pi t) + e^{-t}\pi(\pi^3 + \frac{21}{68}\pi^2 + \frac{7}{136}\pi + \frac{3}{136}) \sin(2\pi t))}{128\pi^4 + 40\pi^2 + 2} \\ &\quad + \frac{-272(\frac{(\pi+\frac{1}{4})(\pi^2+\frac{1}{4})\pi}{17}e^{-\frac{3}{2}t} - \frac{2}{17}(\pi^2 + \frac{1}{16})(\pi^2 - \frac{9}{2}\pi - 1))}{128\pi^4 + 40\pi^2 + 2}. \end{aligned}$$

List of symbols

- \mathcal{P}_H operator such that $m^{\text{HMM}}(\mathcal{P}_H\Phi, \Phi_H) = m^{\text{eff}}(\Phi, \Phi_H)$ for all $\Phi \in X, \Phi_H \in V_H$
- α coercivity constant of the parameter \mathbf{M}
- \mathbf{A} unweighted Maxwell operator
- $\mathcal{A}, \mathcal{B}, \mathcal{R}, \mathcal{S}$ operators
- \mathbf{B} magnetic displacement
- $\partial\Omega$ boundary of the domain Ω
- $C_{\mathbf{M}}$ $L^\infty(\Omega; \mathbb{R}^{n \times n})$ -norm of parameter \mathbf{M}
- $C_{\mathbf{R}}$ $L^\infty(\Omega; \mathbb{R}^{n \times n})$ -norm of parameter \mathbf{R}
- χ_e electric susceptibility of material
- χ_e^{in} instantaneous electric susceptibility of material
- χ_m magnetic susceptibility of material
- χ_m^{in} instantaneous magnetic susceptibility of material
- C generic constant independent of h, H, δ, κ and t if not stated differently
- \mathbf{D} electric displacement
- δ characteristic size of micro structure. As superscript indicates dependence on micro structure
- \mathbf{D}_y^T transposed jacobian matrix
- $^{\text{eff}}$ indicating that a quantity is effective and does not depend on microscopic scale
- $^{\text{eff}, \kappa}$ indicates that a quantity is related to the exact solution of a problem posed on the sampling domain $Y^\kappa(x_K^q)$
- \mathbf{e}_k k -th canonical basis vector in \mathbb{R}^d
- \mathbf{E} electric field strength
- \mathbf{E}_0 initial electric field strength
- \mathcal{E} energy of Maxwell system
- ε_∞ permittivity at maximum frequency
- ε_{in} instantaneous permittivity of material
- ε_r relative permittivity of material
- ε_s permittivity at zero frequency
- $\Delta\varepsilon$ difference $\varepsilon_s - \varepsilon_\infty$

- ε_0 permittivity of vacuum
 $|\cdot|$ Euclidean norm of a vector or 3-dimensional Lebesgue measure
 $\mathcal{F}(u), \widehat{u}$ Fourier–Laplace transform of a function u
 $\|\cdot\|_F$ Frobenius norm
 \mathbf{g} Right-hand side of Maxwell system
 \mathbf{H} magnetic field strength
 \mathbf{H}_0 initial magnetic field strength
HMM indicates that a quantity is HMM approximation of an effective one
 $\mathbf{I}_{d \times d}$ $d \times d$ identity matrix
id identity operator
 \mathcal{I}_H Nédélec interpolation operator for triangulation \mathcal{T}_H
 Π_h Lagrange interpolation operator for triangulation \mathcal{T}_h
 \mathbf{J} current density
 \mathbf{J}_0 current density that is not covered in Ohm’s law
 $\overline{\mathbf{M}}$ magnetization field
 \mathbb{M} magnetization collection
 f mean value integral
 $\mathfrak{M}, \mathfrak{R}, \mathfrak{G}, \mathfrak{A}, \mathfrak{J}, \mathfrak{g}, \mathfrak{f}$ macroscopic finite element matrices and vectors
 $\mathfrak{M}^H, \mathfrak{R}^H, \mathfrak{G}^H, \mathfrak{A}^H, \mathfrak{J}^H, \mathfrak{g}^H, \mathfrak{f}^H$ macroscopic finite element matrices and vectors resulting from the HMM bilinear forms
 \mathbf{M} mass parameter of Maxwell system, element of $\mathbb{R}^{n \times n}$
 μ_{in} instantaneous permeability of material
 μ_r relative permeability of material
 μ_0 permeability of vacuum
 n dimension of Maxwell system equal to $3(2 + N_E + N_H)$
 N_H number of magnetizations
 N_E number of polarizations
 \mathbf{n} unit outward normal to $\partial\Omega$
 Ω bounded simply connected domain in \mathbb{R}^3
 ω Frequency
 \mathbf{P} polarization field
 \mathbb{P} polarization collection
 \mathbf{P}_{in} instantaneous polarization field
 \mathbf{P}_{tot} total polarization field
 $\mathcal{Q}^{\ell, m, n}(K)$ space of polynomials of degree ℓ, m, n in the respective component on K
 x_K^q, γ_K^q q -th quadrature point and weight on K
 \mathbf{R} damping parameter of Maxwell system, element of $\mathbb{R}^{n \times n}$
 ϱ charge density
 σ conductivity
 τ relaxation time
 T final time
 Δt time step size

\mathcal{T}_H triangulation of the computational domain Ω

\mathcal{T}_h triangulation of the unitcell Y or the sampling domains $Y^\kappa \left(x \frac{q}{K} \right)$

$T(t), S(t)$ semigroups

\mathcal{T}^δ periodic unfolding operator

\mathbf{u} solution of Maxwell system

\mathbf{u}_0 initial solution of Maxwell system

w^M, \bar{w}, w^0 cell correctors

$w^{M,h}, \bar{w}^h, w^{0,h}$ discrete cell correctors

$Y^\kappa(\bar{x}) = \bar{x} + \kappa Y$ sampling domain

Y unit cell and periodicity domain

$\mathbf{0}_d$ d -dimensional vector consisting of zeros

$\mathbf{0}_{d \times n}$ $d \times n$ matrix of zeros

List of function spaces

- $a(\cdot, \cdot)$ bilinear form associated to Maxwell operator
 $m^{\text{HMM}}(\cdot, \cdot), r^{\text{HMM}}(\cdot, \cdot), g^{\text{HMM}}(\cdot; \cdot, \cdot)$ the HMM macro bilinear form associated to $\mathbf{M}^{\text{HMM}}, \mathbf{R}^{\text{HMM}}$ and \mathbf{G}^{HMM}
 $m^{\text{eff}}(\cdot, \cdot), r^{\text{eff}}(\cdot, \cdot), g^{\text{eff}}(\cdot; \cdot, \cdot)$ the effective macro bilinear forms associated to $\mathbf{M}^{\text{eff}}, \mathbf{R}^{\text{eff}}$ and $\mathbf{G}^{\text{eff}}(\cdot)$
 $m_H^{\text{eff}}(\cdot, \cdot), r_H^{\text{eff}}(\cdot, \cdot), g_H^{\text{eff}}(\cdot; \cdot, \cdot)$ the discrete effective macro bilinear form associated to $\mathbf{M}^{\text{eff}}, \mathbf{R}^{\text{eff}}$ and $\mathbf{G}^{\text{eff}}(\cdot)$
 $s_m^h(\cdot, \cdot), s_r^h(\cdot, \cdot)$ micro bilinear forms associated to \mathbf{M} or \mathbf{R}
 $s_m(\cdot, \cdot), s_r(\cdot, \cdot)$ micro bilinear forms associated to \mathbf{M} or \mathbf{R}
 $s_m^\kappa(\cdot, \cdot), s_r^\kappa(\cdot, \cdot)$ micro bilinear forms associated to \mathbf{M} or \mathbf{R} on the sampling domain $Y^\kappa(x_K^q)$
 $C^k(\Omega)$ space of k times differentiable functions
 $C_0^\infty(\Omega)$ space of smooth functions with compact support
 $C_\#^\infty(\Omega)$ space of smooth periodic functions
 $\mathcal{D}(\mathcal{A})$ domain of operator \mathcal{A}
 $\mathbf{H}(\text{curl}, \Omega)$ space of square integrable functions possessing weak curl
 $\mathbf{H}_0(\text{curl}, \Omega)$ closure of compactly supported functions with respect to the $\|\cdot\|_{\mathbf{H}(\text{curl}, \Omega)}$ -norm
 $\mathbf{H}^k(\Omega)$ the Hilbert spaces $W^{k,2}(\Omega)$
 $\mathbf{H}_\#^1(Y)$ space of periodic Sobolev functions with zero mean
 $\overline{\mathbf{H}}_\#^1(Y)$ space of periodic Sobolev functions
 (\cdot, \cdot) inner product on $L^2(\Omega)$
 $(\cdot, \cdot)_H$ discrete L^2 inner product
 $L^\infty(\Omega)$ essentially bounded measurable functions on Ω
 $L^p(\Omega)$ standard Lebesgue space of real-valued $L^p(\Omega)$ functions
 $\|\cdot\|_{X \rightarrow X}$ operator norm of linear operator
 $S_k^\#(\mathcal{T}_h), S_k^0(\mathcal{T}_h)$ Lagrange elements of order k on \mathcal{T}_h with either periodic or Dirichlet boundary conditions
 \mathbf{V}_H discrete macroscopic solution space, $\mathbf{V}_0^\ell(\text{curl}, \mathcal{T}_H) \times \mathbf{V}_0^\ell(\text{curl}, \mathcal{T}_H)^{NE} \times \mathbf{V}^\ell(\text{curl}, \mathcal{T}_H) \times \mathbf{V}^\ell(\text{curl}, \mathcal{T}_H)^{NH}$
 \mathbf{V}^h discrete microscopic solution space, either $S_k^\#(\mathcal{T}_h)$ or $S_k^0(\mathcal{T}_h)$
 $(\cdot, \cdot)_{\mathbf{V}_H}$ inner product induced by $m^{\text{HMM}}(\cdot, \cdot)$
 \mathbf{V}^{mac} macroscopic solution space, $\mathbf{H}_0(\text{curl}, \Omega) \times \mathbf{H}_0(\text{curl}, \Omega)^{NE} \times \mathbf{H}(\text{curl}, \Omega) \times \mathbf{H}(\text{curl}, \Omega)^{NH}$
 \mathbf{V}^{mic} microscopic solution space, either $\mathbf{H}_\#^1(Y^\kappa(x_K^q))$ or $\mathbf{H}_0^1(Y^\kappa(x_K^q))$

$(\cdot, \cdot)_{\text{Vmic}} (\nabla_y \cdot, \nabla_y \cdot)_{L^2(Y; \mathbb{R}^n)}$

$W^{k,p}(\Omega)$ standard Sobolev space of k -times weakly differentiable functions in $L^p(\Omega)$

X state space of Maxwell system given as $L^2(\Omega; \mathbb{R}^n)$

$(\cdot, \cdot)_X$ inner product induced by $m^{\text{eff}}(\cdot, \cdot)$

Z high order Sobolev space $H^\ell(\Omega)$

Bibliography

- A. Abdulle. On a priori error analysis of fully discrete heterogeneous multiscale FEM. *Multiscale Model. Simul.*, 4(2): 447–459, 2005. doi: 10.1137/040607137. URL <https://doi.org/10.1137/040607137>.
- A. Abdulle. The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs. In *Multiple scales problems in biomathematics, mechanics, physics and numerics*, volume 31 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, pages 133–181. Gakkōtoshō, Tokyo, 2009.
- A. Abdulle. Discontinuous Galerkin finite element heterogeneous multiscale method for elliptic problems with multiple scales. *Math. Comp.*, 81(278):687–713, 2012. doi: 10.1090/S0025-5718-2011-02527-5. URL <https://doi.org/10.1090/S0025-5718-2011-02527-5>.
- A. Abdulle and T. Pouchon. A priori error analysis of the finite element heterogeneous multiscale method for the wave equation over long time. *SIAM J. Numer. Anal.*, 54(3):1507–1534, 2016. doi: 10.1137/15M1025633. URL <https://doi.org/10.1137/15M1025633>.
- A. Abdulle, W. E. B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numer.*, 21:1–87, 2012. doi: 10.1017/S0962492912000025. URL <https://doi.org/10.1017/S0962492912000025>.
- A. Abdulle, M. J. Grote, and C. Stohrer. Finite element heterogeneous multiscale method for the wave equation: long-time effects. *Multiscale Model. Simul.*, 12(3):1230–1257, 2014. doi: 10.1137/13094195X. URL <https://doi.org/10.1137/13094195X>.
- R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.
- G. Allaire. Homogenization and two-scale convergence. *SIAM J. Math. Anal.*, 23(6):1482–1518, 1992. doi: 10.1137/0523084. URL <https://doi.org/10.1137/0523084>.
- D. Arndt, W. Bangerth, B. Blais, T. C. Clevenger, M. Fehling, A. V. Grayver, T. Heister, L. Heltai, M. Kronbichler, M. Maier, and et al. The deal.II library, Version 9.2. *J. Numer. Math.*, 28(3):131–146, 2020. doi: 10.1515/jnma-2020-0043. URL <https://doi.org/10.1515/jnma-2020-0043>.
- W. Bangerth, C. Burstedde, T. Heister, and M. Kronbichler. Algorithms and data structures for massively parallel generic adaptive finite element codes. *ACM Trans. Math. Softw.*, 38(2), 2012. doi: 10.1145/2049673.2049678. URL <https://doi.org/10.1145/2049673.2049678>.
- H. T. Banks, M. W. Buksas, and T. Lin. *Electromagnetic material interrogation using conductive interfaces and acoustic wavefronts*, volume 21 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. doi: 10.1137/1.9780898719871. URL <https://doi.org/10.1137/1.9780898719871>.
- H. T. Banks, V. A. Bokil, D. Cioranescu, N. L. Gibson, G. Griso, and B. Miara. Homogenization of periodically varying coefficients in electromagnetic materials. *J. Sci. Comput.*, 28(2-3):191–221, 2006. doi: 10.1007/s10915-006-9091-y. URL <https://doi.org/10.1007/s10915-006-9091-y>.
- F. Bekkouché, W. Chikouché, and S. Nicaise. Fully discrete approximation of general nonlinear Sobolev equations. *Afr. Mat.*, 30(1-2):53–90, 2019. doi: 10.1007/s13370-018-0626-9. URL <https://doi.org/10.1007/s13370-018-0626-9>.

- A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*. AMS Chelsea Publishing, Providence, RI, 2011. doi: 10.1090/chel/374. URL <https://doi.org/10.1090/chel/374>.
- J.-P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, 114(2):185–200, 1994. doi: 10.1006/jcph.1994.1159. URL <https://doi.org/10.1006/jcph.1994.1159>.
- V. A. Bokil, H. T. Banks, D. Cioranescu, and G. Griso. A multiscale method for computing effective parameters of composite electromagnetic materials with memory effects. *Quart. Appl. Math.*, 76(4):713–738, 2018. doi: 10.1090/qam/1503. URL <https://doi.org/10.1090/qam/1503>.
- A. Bossavit, G. Griso, and B. Miara. Modelling of periodic electromagnetic structures bianisotropic materials with memory effects. *J. Math. Pures Appl. (9)*, 84(7):819–850, 2005. doi: 10.1016/j.matpur.2004.09.015. URL <https://doi.org/10.1016/j.matpur.2004.09.015>.
- S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008. doi: 10.1007/978-0-387-75934-0. URL <https://doi.org/10.1007/978-0-387-75934-0>.
- C. Burstedde, L. C. Wilcox, and O. Ghattas. p4est: scalable algorithms for parallel adaptive mesh refinement on forests of octrees. *SIAM J. Sci. Comput.*, 33(3):1103–1133, 2011. doi: 10.1137/100791634. URL <https://doi.org/10.1137/100791634>.
- M. Cassier, P. Joly, and M. Kachanovska. Mathematical models for dispersive electromagnetic waves: an overview. *Comput. Math. Appl.*, 74(11):2792–2830, 2017. doi: 10.1016/j.camwa.2017.07.025. URL <https://doi.org/10.1016/j.camwa.2017.07.025>.
- P. Ciarlet, Jr., S. Fliss, and C. Stohrer. On the approximation of electromagnetic fields by edge finite elements. Part 2: A heterogeneous multiscale method for Maxwell’s equations. *Comput. Math. Appl.*, 73(9):1900–1919, 2017. doi: 10.1016/j.camwa.2017.02.043. URL <https://doi.org/10.1016/j.camwa.2017.02.043>.
- P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. doi: 10.1137/1.9780898719208. URL <https://doi.org/10.1137/1.9780898719208>.
- D. Cioranescu and P. Donato. *An introduction to homogenization*, volume 17 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1999.
- D. Cioranescu, A. Damlamian, and G. Griso. Periodic unfolding and homogenization. *C. R. Math. Acad. Sci. Paris*, 335(1):99–104, 2002. doi: 10.1016/S1631-073X(02)02429-9. URL [https://doi.org/10.1016/S1631-073X\(02\)02429-9](https://doi.org/10.1016/S1631-073X(02)02429-9).
- D. Cioranescu, A. Damlamian, and G. Griso. The periodic unfolding method in homogenization. *SIAM J. Math. Anal.*, 40(4):1585–1620, 2008. doi: 10.1137/080713148. URL <https://doi.org/10.1137/080713148>.
- P. J. W. Debye. *Polare Molekeln*. Hirzel, Leipzig, 1929.
- T. Dohnal, A. Lamacz, and B. Schweizer. Bloch-wave homogenization on large time scales and dispersive effective wave equations. *Multiscale Model. Simul.*, 12(2):488–513, 2014. doi: 10.1137/130935033. URL <https://doi.org/10.1137/130935033>.
- T. Dohnal, A. Lamacz, and B. Schweizer. Dispersive homogenized models and coefficient formulas for waves in general periodic media. *Asymptot. Anal.*, 93(1-2):21–49, 2015. doi: 10.3233/ASY-141280. URL <https://doi.org/10.3233/ASY-141280>.
- W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003. doi: 10.4310/CMS.2003.v1.n1.a8. URL <https://doi.org/10.4310/CMS.2003.v1.n1.a8>.
- W. E and B. Engquist. The heterogeneous multi-scale method for homogenization problems. In *Multiscale methods in science and engineering*, volume 44 of *Lect. Notes Comput. Sci. Eng.*, pages 89–110. Springer, Berlin, 2005. doi: 10.1007/3-540-26444-2_4. URL https://doi.org/10.1007/3-540-26444-2_4.
- Y. Efendiev and T. Y. Hou. *Multiscale finite element methods: Theory and applications*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2009. doi: 10.1007/978-0-387-09496-0. URL <https://doi.org/10.1007/978-0-387-09496-0>.
- K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. doi: 10.1007/b97696. URL <https://doi.org/10.1007/b97696>.
- A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004. doi: 10.1007/978-1-4757-4355-5. URL <https://doi.org/10.1007/978-1-4757-4355-5>.

- D. Gallistl and D. Peterseim. Computation of quasi-local effective diffusion tensors and connections to the mathematical theory of homogenization. *Multiscale Model. Simul.*, 15(4):1530–1552, 2017. doi: 10.1137/16M1088533. URL <https://doi.org/10.1137/16M1088533>.
- D. Gallistl, P. Henning, and B. Verfürth. Numerical homogenization of $\mathbf{H}(\text{curl})$ -problems. *SIAM J. Numer. Anal.*, 56(3):1570–1596, 2018. doi: 10.1137/17M1133932. URL <https://doi.org/10.1137/17M1133932>.
- P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 69 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011. doi: 10.1137/1.9781611972030. URL <https://doi.org/10.1137/1.9781611972030>.
- E. Hairer and G. Wanner. *Solving ordinary differential equations II: Stiff and differential-algebraic problems*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1996. doi: 10.1007/978-3-642-05221-7. URL <https://doi.org/10.1007/978-3-642-05221-7>.
- E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations I: Nonstiff problems*, volume 8 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1993. doi: 10.1007/978-3-540-78862-1. URL <https://doi.org/10.1007/978-3-540-78862-1>.
- E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration: Structure-preserving algorithms for ordinary differential equations*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. doi: 10.1007/3-540-30666-8. URL <https://doi.org/10.1007/3-540-30666-8>.
- C. Harder, D. Paredes, and F. Valentin. A family of multiscale hybrid-mixed finite element methods for the Darcy equation with rough coefficients. *J. Comput. Phys.*, 245:107–130, 2013. doi: 10.1016/j.jcp.2013.03.019. URL <https://doi.org/10.1016/j.jcp.2013.03.019>.
- P. Henning, A. Målqvist, and D. Peterseim. A localized orthogonal decomposition method for semi-linear elliptic problems. *ESAIM Math. Model. Numer. Anal.*, 48(5):1331–1349, 2014. doi: 10.1051/m2an/2013141. URL <https://doi.org/10.1051/m2an/2013141>.
- P. Henning, M. Ohlberger, and B. Verfürth. A new heterogeneous multiscale method for time-harmonic Maxwell’s equations. *SIAM J. Numer. Anal.*, 54(6):3493–3522, 2016. doi: 10.1137/15M1039225. URL <https://doi.org/10.1137/15M1039225>.
- D. Hipp, M. Hochbruck, and C. Stohrer. Unified error analysis for nonconforming space discretizations of wave-type equations. *IMA J. Numer. Anal.*, 39(3):1206–1245, 2019. doi: 10.1093/imanum/dry036. URL <https://doi.org/10.1093/imanum/dry036>.
- M. Hochbruck, T. Jahnke, and R. Schnaubelt. Convergence of an ADI splitting for Maxwell’s equations. *Numer. Math.*, 129(3):535–561, 2015a. doi: 10.1007/s00211-014-0642-0. URL <https://doi.org/10.1007/s00211-014-0642-0>.
- M. Hochbruck, T. Pažur, A. Schulz, E. Thawinan, and C. Wieners. Efficient time integration for discontinuous Galerkin approximations of linear wave equations. *ZAMM Z. Angew. Math. Mech.*, 95(3):237–259, 2015b. doi: 10.1002/zamm.201300306. URL <https://doi.org/10.1002/zamm.201300306>.
- M. Hochbruck, B. Maier, and C. Stohrer. Heterogeneous multiscale method for Maxwell’s equations. *Multiscale Model. Simul.*, 17(4):1147–1171, 2019. doi: 10.1137/18M1234072. URL <https://doi.org/10.1137/18M1234072>.
- T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189, 1997. doi: 10.1006/jcph.1997.5682. URL <https://doi.org/10.1006/jcph.1997.5682>.
- T. Hytönen, J. van Neerven, M. Veraar, and L. Weis. *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, volume 63 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer, Cham, 2016. doi: 10.1007/978-3-319-48520-1. URL <https://doi.org/10.1007/978-3-319-48520-1>.
- J. D. Jackson. *Classical electrodynamics*. John Wiley & Sons, Inc., New York, third edition, 1999.
- V. V. Jikov, S. M. Kozlov, and O. A. Oleĭnik. *Homogenization of differential operators and integral functionals*. Springer-Verlag, Berlin, 1994. doi: 10.1007/978-3-642-84659-5. URL <https://doi.org/10.1007/978-3-642-84659-5>.
- D. F. Kelley and R. J. Luebbers. Piecewise linear recursive convolution for dispersive media using FDTD. *IEEE Transactions on Antennas and Propagation*, 44(6):792–797, 1996. doi: 10.1109/8.509882. URL <https://doi.org/10.1109/8.509882>.
- L. D. Landau and E. M. Lifshitz. *Electrodynamics of continuous media*. Course of Theoretical Physics, Vol. 8. Translated from the Russian by J. B. Sykes and J. S. Bell. Pergamon Press, Oxford-London-New York-Paris; Addison-Wesley Publishing Co., Inc., Reading, Mass., 1960.
- S. Lanteri and C. Scheid. Convergence of a discontinuous Galerkin scheme for the mixed time-domain Maxwell’s equations in dispersive media. *IMA J. Numer. Anal.*, 33(2):432–459, 2013. doi: 10.1093/imanum/drs008. URL <https://doi.org/10.1093/imanum/drs008>.

- S. Lanteri, D. Paredes, C. Scheid, and F. Valentin. The multiscale hybrid-mixed method for the Maxwell equations in heterogeneous media. *Multiscale Model. Simul.*, 16(4):1648–1683, 2018. doi: 10.1137/16M110037X. URL <https://doi.org/10.1137/16M110037X>.
- J. Li. Error analysis of fully discrete mixed finite element schemes for 3-D Maxwell’s equations in dispersive media. *Comput. Methods Appl. Mech. Engrg.*, 196(33-34):3081–3094, 2007. doi: 10.1016/j.cma.2006.12.009. URL <https://doi.org/10.1016/j.cma.2006.12.009>.
- J. Li. Unified analysis of leap-frog methods for solving time-domain Maxwell’s equations in dispersive media. *J. Sci. Comput.*, 47(1):1–26, 2011. URL <https://doi.org/10.1007/s10915-010-9417-7>.
- J. Li and Y. Chen. Finite element study of time-dependent Maxwell’s equations in dispersive media. *Numer. Methods Partial Differential Equations*, 24(5):1203–1221, 2008. doi: 10.1002/num.20314. URL <https://doi.org/10.1002/num.20314>.
- J. Li and Y. Huang. *Time-domain finite element methods for Maxwell’s equations in metamaterials*, volume 43 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013. doi: 10.1007/978-3-642-33789-5. URL <https://doi.org/10.1007/978-3-642-33789-5>.
- J. Li and Z. Zhang. Unified analysis of time domain mixed finite element methods for Maxwell’s equations in dispersive media. *J. Comput. Math.*, 28(5):693–710, 2010. doi: 10.4208/jcm.1001-m3072. URL <https://doi.org/10.4208/jcm.1001-m3072>.
- C. Lubich. Convolution quadrature and discretized operational calculus. I. *Numer. Math.*, 52(2):129–145, 1988. doi: 10.1007/BF01398686. URL <https://doi.org/10.1007/BF01398686>.
- R. Luebbers, F. P. Hunsberger, K. S. Kunz, R. B. Standler, and M. Schneider. A frequency-dependent finite-difference time-domain formulation for dispersive materials. *IEEE Transactions on Electromagnetic Compatibility*, 32(3):222–227, 1990.
- A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603, 2014. doi: 10.1090/S0025-5718-2014-02868-8. URL <https://doi.org/10.1090/S0025-5718-2014-02868-8>.
- P. Monk. *Finite element methods for Maxwell’s equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2003. doi: 10.1093/acprof:oso/9780198508885.001.0001. URL <https://doi.org/10.1093/acprof:oso/9780198508885.001.0001>.
- J.-C. Nédélec. Mixed finite elements in \mathbb{R}^3 . *Numer. Math.*, 35(3):315–341, 1980. doi: 10.1007/BF01396415. URL <https://doi.org/10.1007/BF01396415>.
- G. Nguetseng. A general convergence result for a functional related to the theory of homogenization. *SIAM J. Math. Anal.*, 20(3):608–623, 1989. doi: 10.1137/0520043. URL <https://doi.org/10.1137/0520043>.
- A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983. doi: 10.1007/978-1-4612-5561-1. URL <https://doi.org/10.1007/978-1-4612-5561-1>.
- J. B. Pendry. Negative refraction makes a perfect lens. *Phys. Rev. Lett.*, 85:3966–3969, 2000. doi: 10.1103/PhysRevLett.85.3966. URL <https://doi.org/10.1103/PhysRevLett.85.3966>.
- J. B. Pendry, A. J. Holden, W. J. Stewart, and I. Youngs. Extremely low frequency plasmons in metallic mesostructures. *Phys. Rev. Lett.*, 76:4773–4776, 1996. doi: 10.1103/PhysRevLett.76.4773. URL <https://doi.org/10.1103/PhysRevLett.76.4773>.
- D. Peterseim. Variational multiscale stabilization and the exponential decay of fine-scale correctors. In *Building bridges: connections and challenges in modern approaches to numerical partial differential equations*, volume 114 of *Lect. Notes Comput. Sci. Eng.*, pages 341–367. Springer, [Cham], 2016.
- G. G. Raju. *Dielectrics in electric fields*. Power engineering; 19. Dekker, New York, 2003.
- E. Sánchez-Palencia. *Nonhomogeneous media and vibration theory*, volume 127 of *Lecture Notes in Physics*. Springer-Verlag, Berlin-New York, 1980. doi: 10.1007/3-540-10000-8. URL <https://doi.org/10.1007/3-540-10000-8>.
- R. A. Shelby, D. R. Smith, and S. Schultz. Experimental verification of a negative index of refraction. *Science*, 292(5514):77–79, 2001. doi: 10.1126/science.1058847. URL <https://doi.org/10.1126/science.1058847>.
- R. Siushansian and J. LoVetri. Efficient evaluation of convolution integrals arising in FDTD formulations of electromagnetic dispersive media. *Journal of Electromagnetic Waves and Applications*, 11(1):101–117, 1997. doi: 10.1163/156939397X00675. URL <https://doi.org/10.1163/156939397X00675>.
- D. R. Smith and N. Kroll. Negative refractive index in left-handed materials. *Phys. Rev. Lett.*, 85:2933–2936, 2000. doi: 10.1103/PhysRevLett.85.2933. URL <https://doi.org/10.1103/PhysRevLett.85.2933>.

- D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz. Composite medium with simultaneously negative permeability and permittivity. *Phys. Rev. Lett.*, 84:4184–4187, 2000. doi: 10.1103/PhysRevLett.84.4184. URL <https://doi.org/10.1103/PhysRevLett.84.4184>.
- L. Solymar and E. Shamonina. *Waves in metamaterials*. Oxford University Press, Oxford, 2009.
- A. Sturm. *Locally Implicit Time Integration for Linear Maxwell's Equations*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2017. doi: 10.5445/IR/1000069341. URL <https://doi.org/10.5445/IR/1000069341>.
- V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 1054 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1984. doi: 10.1007/BFb0071790. URL <https://doi.org/10.1007/BFb0071790>.
- T. Trilinos Project Team. *The Trilinos Project Website*, 2020 (accessed February 1, 2021). URL <https://trilinos.github.io>.
- V. G. Veselago. The electrodynamics of substances with simultaneously negative values of ϵ and μ . *Soviet Physics Uspekhi*, 10(4):509–514, 1968. doi: 10.1070/pu1968v010n04abeh003699. URL <https://doi.org/10.1070/pu1968v010n04abeh003699>.
- N. Wellander. Homogenization of the Maxwell equations. Case I. Linear theory. *Appl. Math.*, 46(1):29–51, 2001. doi: 10.1023/A:1013727504393. URL <https://doi.org/10.1023/A:1013727504393>.
- N. Wellander. Homogenization of the Maxwell equations. Case II. Nonlinear conductivity. *Appl. Math.*, 47(3):255–283, 2002. doi: 10.1023/A:1021797505024. URL <https://doi.org/10.1023/A:1021797505024>.
- N. Wellander and G. Kristensson. Homogenization of the Maxwell equations at fixed frequency. *SIAM J. Appl. Math.*, 64(1):170–195, 2003. doi: 10.1137/S0036139902403366. URL <https://doi.org/10.1137/S0036139902403366>.
- X. Yue and W. E. The local microscale problem in the multiscale modeling of strongly heterogeneous media: effects of boundary conditions and cell size. *J. Comput. Phys.*, 222(2):556–572, 2007. doi: 10.1016/j.jcp.2006.07.034. URL <https://doi.org/10.1016/j.jcp.2006.07.034>.
- Y. Zhang, L. Cao, W. Allegretto, and Y. Lin. Multiscale numerical algorithm for 3-D Maxwell's equations with memory effects in composite materials. *Int. J. Numer. Anal. Model. Ser. B*, 1(1):41–57, 2010.