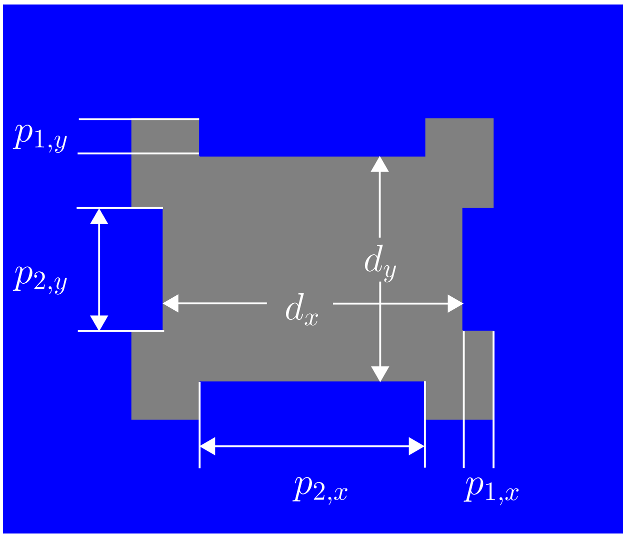
Reduced Basis Method for Electromagnetic Scattering Problems



Dissertation zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
am Fachbereich Mathematik und Informatik
an der Freien Universität Berlin

vorgelegt von
Jan Pomplun
im Oktober 2009

Betreuer und Erstgutachter:
PD Dr. Frank Schmidt

Zweitgutachter:
Professor Alfio Quarteroni, Dr. h.c.

Datum der Disputation: 5. Mai 2010

Acknowledgement

First of all I thank Frank Schmidt for introducing me to the field of computational nano-optics and giving me the opportunity to carry out this work in his group. I am truly greatful for his support, open and friendly attitude, and trust over the past years.

Working here at the Zuse Institute Berlin was wonderful. My thanks, therefore, also go to Professor Deuflhard for the possibility to carry out research in such a stimulating environment.

I am grateful to Lin Zschiedrich for his support. He was a valuable source of advice for theoretical, numerical, and implementational questions and also criticism, which always helps to improve the own work.

I thank my colleagues Sven Burger, Benjamin Kettner, Daniel Kojda, Daniel Lockau, Therese Pollok, and Achim Schädle for providing such a great atmosphere in our group.

I thank Daniel Lockau also for proofreading the manuscript.

I thank Professor Anthony Patera, Professor Yvon Maday, and Gianluigi Rozza for their great work during the CEA-EDF-INRIA summer school on the reduced basis method. Their lessons helped a lot, getting started in this field of research.

Regarding the work on inverse scatterometry, I thank Frank Scholze and Christian Laubis from the Physikalisch-Technische Bundesanstalt for the fruitful collaboration.

Für meine Familie

Contents

1 Introduction **9**

1.1 Motivation . . . 9

1.2 Earlier work . . . 10

1.3 Thesis contribution . . . 11

1.4 Thesis outline . . . 12

2 Preliminaries **13**

2.1 Mathematical background . . . 13

2.1.1 Functional analysis . . . 13

2.1.2 Exterior Calculus . . . 22

2.2 Maxwell's equations . . . 31

2.2.1 Maxwell's equations in differential form . . . 31

2.2.2 Maxwell's equations in integral form . . . 32

2.2.3 Time-harmonic Maxwell's equations . . . 33

2.2.4 Maxwell's equations with exterior calculus . . . 34

3 Electromagnetic scattering problem **39**

3.1 Strong formulation of the scattering problem . . . 39

3.2 Weak formulation of the scattering problem . . . 41

3.3 Transparent boundary conditions . . . 43

3.3.1 Perfectly matched layers . . . 44

3.3.2 Pole condition . . . 47

3.3.3 Dirichlet to Neumann operator . . . 47

3.4 Finite element discretization . . . 48

4 A posteriori error estimation **51**

4.1 Problem setup . . . 51

4.2 Residuum . . . 52

4.3 Inf-sup constant . . . 53

4.4 Primal error bounds . . . 55

4.5 Dual problem . . . 56

4.6 Effectivities . . . 58

5 Reduced basis method 61
5.1 Problem setup . . . 61
5.2 Reduced basis approximation . . . 63
5.2.1 Reduced basis spaces . . . 64
5.2.2 Online-offline decomposition . . . 65
5.2.3 Operation count . . . 67
5.3 A posteriori error estimation . . . 67
5.3.1 Error bounds . . . 67
5.3.2 Online-offline decomposition . . . 69
5.3.3 Operation count . . . 71
5.4 Dual problem . . . 71
5.5 Effectivities . . . 73
5.6 Greedy sampling strategy . . . 74
5.7 Affine geometry precondition . . . 78
5.8 Affine decomposition of Maxwell's equations . . . 80
5.9 Residuum estimator . . . 83
5.9.1 Classical sub-domain residuum method . . . 84
5.9.2 Reduced basis sub-domain residuum method . . . 86
5.9.3 Online-offline decomposition . . . 87
5.9.4 Operation count . . . 88
5.10 Multiple sources . . . 90
5.10.1 Problem setup . . . 90
5.10.2 Efficient reduced basis treatment . . . 91

6 Application examples 95
6.1 Incoming fields and outputs of interest . . . 96
6.2 2D phase shift mask . . . 98
6.3 3D periodic grating . . . 105
6.4 Inverse scatterometry . . . 109
6.4.1 Mathematical formulation . . . 111
6.4.2 EUV mask reconstruction . . . 113
6.5 Optical proximity correction . . . 126
6.5.1 Model problem . . . 128
6.5.2 Optimization problem . . . 131
6.5.3 Reduced basis approximation . . . 134

7 Conclusion and outlook 139

8 Zusammenfassung 141

1 Introduction

1.1 Motivation

A main objective of numerical analysis and modeling is the simulation of complex technological problems, arising in engineering and natural sciences. Numerical simulations help to understand, design and optimize, or control and characterize systems or components.

Usually the behaviour of a system is described by physical quantities like temperature, stress, or electromagnetic fields. These fields are solutions to partial differential equations (PDEs), which are stated on the domain of interest with appropriate boundary conditions. Since in general the analytical solution to a PDE is unavailable, a discretization procedure [70] such as finite element [51, 37], finite difference [105], discontinuous Galerkin [28], or finite volume method [43, 101] has to be applied. The discretized system is then solved numerically. For real world problems the numerical solution is usually expensive, regarding computational resources and time. Computational times can be of the order of seconds, up to hours and days, and even many problems can not be solved at all with reasonable effort due to their complexity.

In engineering applications like optimization or parameter estimation the discretized models have to be solved multiply for different configurations of the system under consideration, for example, regarding geometrical or material parameters. Hence, a large number of solutions for different parameters are required in reasonable time (*many-query* context), or a single solution has to be computed very fast (*real-time* context). Even for moderate problems these requirements can often not be met with above discretization methods.

In applications usually the output of interest is not the solution of the PDE itself, but some derived quantities. Hence, a method for fast and reliable evaluation of *input-output relationships* is desirable. The input are, for example, geometrical or material parameters of the system under consideration. The output is given implicitly as a functional of the field variable, which is the solution to the input parameter dependent PDE.

The *reduced basis method* offers a way to construct approximations to such input-output relationships, which can be evaluated very fast. The key is an *online-offline decomposition*. In a so-called *offline phase* the reduced model is built self-adaptively. In an actual application (*online phase*) only the reduced model is solved. Rigorous error estimation techniques allow to control and quantify the accuracy of the approximative reduced model, such that reduced basis solutions are reliable.

1.2 Earlier work

In the 1970s and 1980s the basic idea of using a small number of global shape functions [21] as a reduced basis was first applied to finite element problems in the field of linear and nonlinear structural analysis [57]. The method was further analyzed in [4, 19, 68, 75], including the field of error estimation. First extensions of the reduced basis method began into the field of fluid dynamics [24, 34, 61]. The drawback of these earlier works was that the operation count for a reduced basis computation still depended on the number of finite element degrees of freedom. In [69, 48, 99, 100] efficient online-offline decomposition strategies for affinely parametrized problems were developed together with a posteriori error estimation concepts. For non-affine parameter dependencies the empirical interpolation method [7] was introduced.

An overview of state-of-the-art reduced basis techniques for affinely parametrized linear coercive elliptic problems is given in [82, 59]. Results on parabolic and hyperbolic problems can be found in [76, 77, 22] and [42] respectively, and [55] covers non-affine parametrizations.

Application areas range from mechanics [82, 54, 50], fluid dynamics [72, 83, 81, 38], medicine [78, 71], heat and mass transfer [58, 22, 96] to acoustics [92], or quantum mechanics [15, 49]. Only little work has been done in the field of electromagnetics. In [16] a 2D electromagnetic cavity problem with parametrized material properties is considered as an example. In [107] the reduced basis method is used for fast computation of geometrically parametrized 2D electromagnetic scattering problems, however, only structured rectangular grids are used, and the field of error estimation is not covered.

The reduced basis method has further been applied to real-time and many-query applications like inverse problems [54], parameter estimation [45, 46], shape optimization [78, 79, 80], and optimal control [73].

Application of the reduced basis method to finite volume schemes can be found in [26, 27].

1.3 Thesis contribution

The online and offline costs of state-of-the-art reduced basis methods mainly depend on two quantities: first the dimension of the reduced basis space N and second the number of terms Q in the affine expansion of the system bilinear form. Strategies for the case of high values for Q are not addressed in current research. In [82] the problem of poor online and offline performance is mentioned together with the necessity of efficient "reduced basis triangulations" for geometrically parametrized problems. However, for complicated geometries such efficient triangulations are not always possible, or even efficient reduced basis triangulations lead to affine expansions with high Q. Especially error estimation then becomes extremely expensive and often infeasible. For some of our application examples, well-established techniques would lead to memory requirements of the order of Terabytes, making the reduced basis method non-applicable. Therefore, we develop a novel method for estimation of the residuum in the reduced basis context, which is a key ingredient for a posteriori error estimation concepts. The novel estimator is orders of magnitude faster than current techniques, leads to substantial memory savings, and enables application of the reduced basis method to a wider class of problems. Since the developed error estimator only contains an estimate to the exact norm of the residuum, the error bounds are not rigorous anymore. In numerical experiments, however, we will demonstrate, that the residuum estimate provides a very accurate approximation to the true residuum.

Especially problems in 3D lead to affine expansions with high number of terms and are, to our knowledge, not addressed in actual research. The application of the reduced basis method to geometrically parametrized finite element problems in three spatial dimensions, therefore, also lies in the focus of our work. The problem class of Maxwell scattering problems on unbounded domains, which is considered in the thesis, also has not been addressed so far.

Furthermore, we develop a new technique for efficient reduced basis computation of outputs of interest for systems subject to a large number of different sources. Usually, the online costs computing outputs of interest scales linearly with the number of sources. With our technique the costs become basically independent on the number of sources. In nano-optical applications multiple sources are a common situation, e.g., arising when a system is illuminated by a complex light source, which is modeled by a large number of different incoming fields.

1.4 Thesis outline

The thesis is structured as follows. In Chapter 2 we set the mathematical and physical background for investigation of the reduced basis method and its application to electromagnetic scattering problems. Besides recapitulating important results from functional analysis and exterior calculus, we review Maxwell's equations.

In Chapter 3 the mathematical formulation of electromagnetic scattering problems on unbounded domains is given. As an important aspect we give different concepts of transparent boundary conditions.

Since the reduced basis method gives approximative solutions to a given problem, error estimation is an important topic. This is subject of section 4. We derive a posteriori error bounds for non-coercive elliptic variational problems, which will be used estimating errors of reduced basis solutions.

Chapter 5 contains our main work on the reduced basis method. We develop our contributions and give state-of-the art results for efficient online-offline decomposition, error estimation, and construction of reduced basis spaces.

In Chapter 6 we finally apply the developed reduced basis techniques to a number of challenging nano-optical problems. These include a real world inverse scatterometry problem and shape optimization of nano-optical systems in the field of computational lithography.

2 Preliminaries

In the following we summarize theoretical background, necessary for mathematical formulation of our problem setup and analysis of the reduced basis method. We start with a review of important definitions and results from functional analysis. Then we cover important concepts of exterior calculus, which are used for formulation of the electromagnetic scattering problem and derivation of its affine parametrization.

Secondly we review Maxwell's equations, which serve as the fundamental equations, modeling electromagnetic scattering problems.

2.1 Mathematical background

2.1.1 Functional analysis

Tools from functional analysis will be used extensively in the analysis of the reduced basis method. Therefore, we review important results and definitions in the following, also introducing our notation. We will mostly follow [74, 102, 18], where omitted proofs can be found.

Hilbert space

Electromagnetic fields, which are computed numerically, are elements of certain Hilbert spaces. We start with the following definition:

Definition 1. Let X be a real (complex) vector space. A function $||\cdot||_X$ from X to $\mathbb{R}$ is called a **norm** if it satisfies the following conditions for all $x, y \in X$ and $\alpha \in \mathbb{R}(\mathbb{C})$:

(i) $||x||_X \geq 0 \quad \forall x \in X$,

(ii) $||x||_X = 0 \Leftrightarrow x = 0$,

(iii) $||\alpha x||_X = |\alpha| \, ||x||_X$,

(iv) $||x + y||_X \leq ||x||_X + ||y||_X$.

The pair $(X, ||\cdot||_X)$ is called **normed linear space**.

A second ingredient to a Hilbert space is a scalar product. We restrict us to complex vector spaces in the following:

Definition 2. Let X be a complex vector space. A function $(\cdot,\cdot)_X$ on $X \times X$ is called **scalar product** or **inner product** if it satisfies the following conditions for all $x, y, z \in X$ and $\alpha \in \mathbb{C}$:

(i) $(x,x)_X \geq 0$ and $(x,x)_X = 0 \Leftrightarrow x = 0$,

(ii) $(x, y+z)_X = (x,y)_X + (x,z)_X$,

(iii) $(x, \alpha y)_X = \alpha\,(x,y)_X$,

(iv) $(x,y)_X = \overline{(y,x)_X}$,

where the bar denotes complex conjugation. The pair $(X, (\cdot,\cdot)_X)$ is called **inner product space**.

Two elements x, y of an inner product space are said to be **orthogonal** if $(x,y)_X = 0$. For each inner product space one can define the norm $||x||_X = \sqrt{(x,x)_X}$. Having a norm at hand, completeness of a normed space is defined as follows:

Definition 3. A normed space in which all Cauchy sequences converge is called **complete**.

Now we have everything together for the definition of a Hilbert space:

Definition 4. A complete inner product space is called **Hilbert space**.

The dual space of a Hilbert space is defined as follows:

Definition 5. Let X be a Hilbert space. The space of all linear mappings form X to $\mathbb{C}$ is called **dual space** of X and is denoted by X'. The elements of X' are called **continuous linear functionals**. For $R, S \in X'$, $x \in X$ and $\alpha \in \mathbb{C}$ we define

$$(\alpha R)(x) = \alpha\ R(x),$$
$$(R+S)(x) = R(x) + S(x),$$

which makes X' a linear space. Furthermore, we can define a norm $||\cdot||_{X'}$ on X', called **dual norm**:

$$||S||_{X'} = \sup_{x \neq 0} \frac{||S(x)||_X}{||x||_X},$$

which makes X' a normed linear space. Instead of $S(x)$, we will often write Sx.

Now an important theorem can be stated, which establishes a connection between the elements of a Hilbert space and its dual space:

Theorem 1 (Riesz representation theorem)**.** Let X be a Hilbert space with dual X'. For each $S \in X'$, there is a unique $y_S \in X$ such that $S(x) = (x, y_S)_X$ for all $x \in X$. Furthermore, we have $||y_S||_X = ||S||_{X'}$. We call y_S the **Riesz representation** of S.

Proof. Here we only proof the equality of the norm of S with the norm of its Riesz representation. We start with the definition of the dual norm and use the Riesz representation of S:

$$\begin{aligned} ||S||_{X'} &= \sup_{x \neq 0} \frac{||S(x)||_X}{||x||_X} \\ &= \sup_{x \neq 0} \frac{|(x, y_S)_X|}{||x||_X}. \end{aligned}$$

Using the Cauchy-Schwarz inequality we get:

$$\begin{aligned} ||S||_{X'} &\leq \sup_{x \neq 0} \frac{||x||_X \, ||y_S||_X}{||x||_X} \\ &= ||y_S||_X \, . \end{aligned}$$

Since $y_S \in X$, we also have:

$$\begin{aligned} ||S||_{X'} &= \sup_{x \neq 0} \frac{|(x, y_S)_X|}{||x||_X} \\ &\geq \frac{|(y_S, y_S)_X|}{||y_S||_X} \\ &= ||y_S||_X \, , \end{aligned}$$

which concludes the proof. □

Bilinear forms

For application of the finite element method, PDEs are stated in weak form. This formulation involves bilinear forms:

Definition 6. Let X be a Hilbert space. A mapping

$$\begin{aligned} a : X \times X &\to \mathbb{R}\ (\mathbb{C}) \\ (x, y) &\mapsto a(x, y), \end{aligned}$$

is called a **bilinear (sesquilinear) form** if for each $\alpha, \beta \in \mathbb{R}\ (\mathbb{C})$ and $x, y, z \in X$ the following relations hold:

(i) $a(\alpha x + \beta y, z) = \alpha\, a(x, z) + \beta\, a(y, z)$,

(ii) $a(x, \alpha y + \beta z) = \overline{\alpha}\, a(x, y) + \overline{\beta}\, a(y, z)$.

For notational convenience, and since Maxwell's equations are stated on complex Hilbert spaces, we will give the following results only for sesquilinear forms. The Riesz representation theorem has an important corollary.

Corollary 1. Let $a(\cdot, \cdot)$ be a sesquilinear form which satisfies:

$$|a(x, y)| \leq C\, ||x||_X\, ||y||_X\,, \tag{2.1}$$

for all $x, y \in X$. Then there exists a unique bounded linear transformation T from X to X such that:

$$a(x, y) = (y, Tx)_X\,.$$

The norm of T is the smallest constant C satisfying (2.1).

Sesquilinear forms can have the following important properties [10]:

Definition 7. A sesquilinear form a on a normed linear space X is:

(i) **bounded** or (**continuous**) if there is a constant $\mathbb{R} \ni \gamma < \infty$ such that:

$$|a(x, y)| \leq \gamma\, ||x||_X\, ||y||_X\,, \quad \forall x, y \in X. \tag{2.2}$$

The smallest γ such that (2.2) holds is called the **continuity constant** of a.

(ii) **coercive** if there is a constant $\mathbb{R} \ni \alpha > 0$ such that:

$$|a(x,x)| \geq \alpha \, ||x||_X^2 \, , \quad \forall x \in X. \tag{2.3}$$

The largest α such that (2.3) holds is called the **coercivity constant** of a.

(iii) **hermitian** or **symmetric** if:

$$a(x,y) = \overline{a(y,x)} \, , \quad \forall x, y \in X.$$

For non-coercive sesquilinear forms the coercivity property can be generalized:

Definition 8. Let a be a sesquilinear form on a normed linear space X. The **inf-sup constant** of a is defined as:

$$\beta = \inf_{x \in X} \sup_{y \in X} \frac{|a(x,y)|}{||x||_X \, ||y||_X}. \tag{2.4}$$

A sesquilinear form is said to satisfy the **Babuška-Brezzi condition** if $\beta > 0$.

Solvability of variational problems

We look at the following variational problem:

Problem 1. Let a be a sesquilinear form on X and $f \in X'$.
Find $x \in X$ such that:

$$a(x,y) = f(y) \, , \quad \forall y \in X. \tag{2.5}$$

The following lemma states existence and uniqueness of a solution x for coercive sesquilinear forms [10]:

Lemma 1 (Lax-Milgram)**.** Suppose a is a bounded coercive sesquilinear form. Then for each $f \in X'$ there exists a unique solution $x \in X$ to (2.5) with

$$||x||_X \leq \frac{\gamma}{\alpha} \, ||f||_{X'} \, ,$$

where γ and α are the continuity and coercivity constant of a.

For the non-coercive case the lemma can be generalized [51]:

Lemma 2. Suppose a is a bounded sesquilinear form, which satisfies the Babuška-Brezzi condition. Then for each $f \in X'$ there exists a unique solution $x \in X$ to (2.5) with

$$||x||_X \leq \frac{\gamma}{\beta} \, ||f||_{X'} \, ,$$

where γ and β are the continuity and inf-sup constant of a.

L^p spaces

We now introduce important classes of Hilbert spaces. These are specific function spaces, which are used when stating PDEs weakly.

The integrals in the following are defined in the Lebesgue sense. Furthermore, in the following we assume that all functions are measurable. We say that some property holds almost everywhere (a.e.), if it holds on a set with Lebesgue measure zero.

From now on let $\Omega \subset \mathbb{R}^n$. We start with the definition of following spaces:

$$\begin{aligned} \mathcal{L}^p(\Omega) &:= \left\{ f : \Omega \to \mathbb{C} : \int_\Omega |f|^p < \infty \right\}, \; 1 \leq p < \infty, \\ ||f||^*_{\mathcal{L}^p(\Omega)} &:= \left(\int_\Omega |f|^p \right)^{1/p} . \end{aligned}$$

These are spaces of p-integrable functions. For the case $p = \infty$ we have to define the essential supremum:

Definition 9. Let f be a real-valued, measurable function. The **essential supremum** of f is defined by:

$$\operatorname{ess\,sup} f = \inf \left\{ c \in \mathbb{R} : |f| < c \text{ (a.e.)} \right\} .$$

Then we can define:

$$\begin{aligned} \mathcal{L}^\infty(\Omega) &:= \left\{ f : \Omega \to \mathbb{C} : \operatorname{ess\,sup} f < \infty \right\}, \\ ||f||^*_{\mathcal{L}^\infty(\Omega)} &:= \operatorname{ess\,sup} f. \end{aligned}$$

The problem with above spaces is that $||\cdot||^*_{\mathcal{L}^p(\Omega)}$ only defines a semi-norm on $\mathcal{L}^p(\Omega)$, since all functions which are zero a.e. have zero norm. Let us denote the kernel of the semi-norm by:

$$N_p := \{ f : f = 0 \text{ (a.e.)} \}.$$

This kernel is used to define an equivalence relationship $\sim_{N_p}$:

$$f \sim_{N_p} g \Leftrightarrow f - g = 0 \text{ (a.e.)}.$$

The space $L^p(\Omega)$ is then defined as the following quotient space:

$$L^p(\Omega) := \mathcal{L}^p(\Omega)/ \sim_{N_p} .$$

Although the members of L^p are equivalence classes $[f]$ of functions, a single representation f can be used, whenever the result does not depend on the specific representation. In the following we will mostly use the space $L^2(\Omega)$, i.e., the set of quadratically integrable functions. On $L^2(\Omega)$ we can define the following scalar product:

$$(u, v)_{L^2(\Omega)} := \int_\Omega \overline{u}\, v, \tag{2.6}$$

and the associated norm:

$$||u||_{L^2(\Omega)} := \left(\int_\Omega \overline{u}\, u\right)^{1/2},$$

which makes it a Hilbert space.

Sobolev spaces

For the definition of Sobolev spaces we need the concept of weak derivatives, which are defined, utilizing test functions.

Definition 10. Let $C_c^\infty(\Omega)$ denote the space of infinitely differentiable functions $\phi : \Omega \to \mathbb{R}$, with compact support in Ω. The members of $C_c^\infty(\Omega)$ will be called **test functions**.

Now let $L^p_{\text{loc}}(\Omega)$ be the set of locally integrable functions:

$$L^p_{\text{loc}}(\Omega) := \{f : f \in L^p(U) \text{ for all compact sub-domains } U \subset \Omega\}.$$

Then we define:

Definition 11. Suppose $u \in L^1_{\text{loc}}(\Omega)$, and α is a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$. Suppose there exists a function $v \in L^1_{\text{loc}}(\Omega)$ such that:

$$\int_\Omega u \partial_\alpha \phi = (-1)^{|\alpha|} \int_\Omega v\phi, \quad \forall \phi \in C_c^\infty(\Omega),$$

with $\partial_\alpha = \partial_{\alpha_1} \cdots \partial_{\alpha_n}$ and $|\alpha| = \sum_{i=1}^n \alpha_i$. Then we call v the α**th-weak partial derivative** of u and formally write $\partial_\alpha u = v$. If no such v exists, we say that u does not possess a weak αth-weak partial derivative.

If it exists, the weak derivative of a function is unique [18]. Now Sobolev spaces can be defined:

Definition 12. The **Sobolev space** is defined by

$$W^{k,p}(\Omega) = \left\{ u \in L^1_{\text{loc}}(\Omega) : \ \forall \alpha \text{ with } 0 \le |\alpha| \le k, \, \partial_\alpha u \text{ exists, and } \partial_\alpha u \in L^p(\Omega) \right\}.$$

For the important case $p = 2$, one writes $H^k(\Omega) = W^{k,2}(\Omega)$. On the Sobolev space $H^k(\Omega)$ a scalar product:

$$(u, v)_{H^k(\Omega)} = \sum_{|\alpha| \le k} \int_\Omega \partial_\alpha \overline{u} \partial_\alpha v,$$

and associated norm:

$$||u||_{H^k(\Omega)} = \left(\sum_{|\alpha| \le k} \int_\Omega \partial_\alpha \overline{u} \partial_\alpha u \right)^{1/2},$$

can be defined, which makes it a Hilbert space.

$H(\mathbf{curl})$ spaces

We already introduced the concept of weak partial derivatives, acting on scalar functions. The solution to Maxwell's equations, however, are vectorial functions in $\mathbb{R}^3$. In the following we will introduce the appropriate Sobolev space [51]. The definition of the L^2 scalar product (2.6) can trivially be extended to vectorial functions. With $\mathbf{u} = (u_1, u_2, u_3) \in (L^2(\Omega))^3$ and $\mathbf{v} = (v_1, v_2, v_3) \in (L^2(\Omega))^3$ we define the $(L^2(\Omega))^3$ inner product and norm according to:

$$(\mathbf{u}, \mathbf{v})_{(L^2(\Omega))^3} = \int_\Omega \sum_{j=1}^{3} \overline{u_i} \, v_i,$$

$$||\mathbf{u}||_{(L^2(\Omega))^3} = \sqrt{(\mathbf{u}, \mathbf{u})_{(L^2(\Omega))^3}}.$$

In Maxwell's equations the **curl**-operator appears as differential operator and has to be generalized for functions in $(L^2(\Omega))^3$, i.e., in the weak sense. Since the **curl**-operator only involves partial derivatives, this offers no principle difficulties:

$$\mathbf{curl}\ \mathbf{u} = (\partial_2 u_3 - \partial_3 u_2, \partial_3 u_1 - \partial_1 u_3, \partial_1 u_2 - \partial_2 u_1,).$$

All derivatives are understood in the weak sense.

The variational formulation of Maxwell's equations is stated on the following Sobolev space:

Definition 13. The space of three dimensional vectorial functions $\mathbf{u}$ in $\left(L^2(\Omega)\right)^3$ with **curl** in $\left(L^2(\Omega)\right)^3$ is defined as:

$$\mathrm{H}\left(\mathbf{curl},\Omega\right) = \left\{\mathbf{u} \in \left(L^2(\Omega)\right)^3 : \mathbf{curl}\,\mathbf{u} \in \left(L^2(\Omega)\right)^3\right\}. \tag{2.7}$$

The norm in the space $\mathrm{H}\left(\mathbf{curl},\Omega\right)$ is defined by:

$$||\mathbf{u}||_{\mathrm{H}(\mathbf{curl},\Omega)} = \left(||\mathbf{u}||^2_{(L^2(\Omega))^3} + ||\mathbf{curl}\,\mathbf{u}||^2_{(L^2(\Omega))^3}\right)^{1/2}.$$

Boundary conditions

For the solutions of PDEs an important class of boundary conditions are Dirichlet boundary conditions, i.e., the value of a function is fixed on the boundary of the domain of interest Ω. For simplicity we will consider zero or essential Dirichlet boundary conditions in the following. Since the boundary $\Gamma = \partial\Omega$ has measure zero, a straightforward definition like "$f|_\Gamma = 0$" makes no sense for functions in $L^2(\Omega)$ or $\mathrm{H}\left(\mathbf{curl},\Omega\right)$. Instead we have the following definition:

Definition 14. The space of functions in $\mathrm{H}\left(\mathbf{curl},\Omega\right)$ satisfying essential Dirichlet boundary conditions is defined by:

$$\mathrm{H}_0\left(\mathbf{curl},\Omega\right) = \text{closure of } \left(C_0^\infty(\Omega)\right)^3 \text{ in the } \mathrm{H}\left(\mathbf{curl},\Omega\right)\text{-norm}, \tag{2.8}$$

where $C_0^\infty(\Omega)$ is the set of infinitely differentiable functions vanishing on $\partial\Omega$.

For motivation of above definition we first give the following theorem, which also serves as an alternative definition of the space $\mathrm{H}\left(\mathbf{curl},\Omega\right)$:

Theorem 2. Suppose Ω is a bounded Lipschitz domain in $\mathbb{R}^3$. Then the closure of $\left(C^\infty(\overline{\Omega})\right)^3$ in the $\mathrm{H}\left(\mathbf{curl},\Omega\right)$-norm is $\mathrm{H}\left(\mathbf{curl},\Omega\right)$.

The proof can be found in [51]. The theorem shows that each function in $\mathrm{H}\left(\mathbf{curl},\Omega\right)$ is a limit of functions in $\left(C^\infty(\overline{\Omega})\right)^3$ in the $\mathrm{H}\left(\mathbf{curl},\Omega\right)$-norm. The definition of essential Dirichlet boundary conditions for $\mathrm{H}_0\left(\mathbf{curl},\Omega\right)$ (2.8) is realized by taking the closure in the same norm but of, the function space $\left(C_0^\infty(\Omega)\right)^3$, whose elements satisfy the boundary condition.

2.1.2 Exterior Calculus

In this Section we summerize results about differential forms, which are used for formulation of Maxwell's equations in coordinate-independent form. In our presentation we mainly follow [36, 20].

Our goal is to describe the following integrals:

$$\int_L Adx + Bdy + Cdz,$$
$$\iint_S Ddydz + Edzdx + Fdxdy,$$
$$\iiint_V Gdxdydz,$$

where L, S, V is a line, surface, and volume in $\mathbb{R}^3$. Differential geometry and the so-called exterior calculus allow an elegant description of these expressions. L, S, V are manifolds and the integrands are exterior differential forms, which will be defined in the following Sections.

Manifolds

For the present work it is sufficient to consider manifolds as subsets of $\mathbb{R}^n$. In this Section we will use the terms diffeomorph, differentiable, etc. in the C^∞-sense.

Definition 15. A subset $M \subset \mathbb{R}^n$ is called a **k-dimensional sub-manifold** of $\mathbb{R}^n$, if for each $p \in M$ there is an open neighbourhood W of p in $\mathbb{R}^n$ and a diffeomorphism $H : W \overset{\cong}{\to} W'$ onto an open subset $W' \subset \mathbb{R}^n$ such that:

$$H(M \cap W) = (\mathbb{R}^k \times 0^{n-k}) \cap W'.$$

The mapping H is called **exterior chart**.

This means, for each point p of a k-dimensional sub-manifold M, we can find an open neighbourhood W in $\mathbb{R}^n$ such that $M \cap W$ is diffeomorph to $\mathbb{R}^k$. Examples are smooth curves in $\mathbb{R}^3$, which are 1-dimensional sub-manifolds, or the boundary of a torus, which is a two dimensional sub-manifolds of $\mathbb{R}^3$.

Considering sub-manifolds of $\mathbb{R}^n$, the mapping H is called **exterior** chart, since its domain includes points which are not in M. Taking the intersection of W and M we get the (interior) chart:

Definition 16. Let M be a k-dimensional sub-manifold of $\mathbb{R}^n$ and $H : W \to W'$ an exterior chart. The **chart** h is defined by the restriction of H as follows:

$$h : U \overset{\cong}{\to} U', \tag{2.9}$$

with **chart domain** $U = M \cap W$ and $U' \times 0^{n-k} = (\mathbb{R}^k \times 0^{n-k}) \cap W'$. The inverse mapping is called a local **parametrization** of M:

$$\varphi := h^{-1} : U' \to U. \tag{2.10}$$

Next we have to construct tangent spaces of sub-manifolds. We start with the following definition.

Definition 17. Let M be a k-dimensional sub-manifold of $\mathbb{R}^n$ and $p \in M$. Suppose $\alpha : (-\epsilon, \epsilon) \to M$ is a differentiable curve with $\alpha(0) = p$, then $\dot{\alpha}(0) \in \mathbb{R}^n$ is called **tangential vector** of M at point p.

Tangential vectors can be defined in several alternative ways [36, 20], however, the above definition is probably most illustrative.

Definition 18. Let M be a k-dimensional sub-manifold of $\mathbb{R}^n$ and $p \in M$. The set of all tangential vectors at p is called **tangential space** of M at point p and is denoted by T_pM.

At each $p \in M$ the tangential space T_pM is a k-dimensional vector space. Furthermore, we can collect all tangential spaces of a manifold and get:

Definition 19. The **tangential bundle** of a manifold M is defined by.

$$TM = \bigcup_p (p, T_pM) .$$

Alternating r-Forms

Next we define objects which operate on r-dimensional vector spaces.

Definition 20. Let V be a k-dimensional vector space. We call

$$\omega : \underbrace{V \times \cdots \times V}_{r\text{-times}} \to \mathbb{R}$$

an **alternating r-form** on V if it is multi-linear and alternating, i.e., for $v_i, v_j \in V$:

$$\omega(\cdots v_i \cdots v_j \cdots) = -\omega(\cdots v_j \cdots v_i \cdots).$$

In order to motivate this definition, we look at a special alternating 3-form on the vector space $\mathbb{R}^3$. Let $v_1, v_2, v_3 \in \mathbb{R}^3$. Then the determinant $\det([v_1, v_2, v_3])$ of the matrix, whose i-th column is v_i, is an alternating 3-form, which measures the volume spanned by the parallel-epiped (v_1, v_2, v_3). Loosely spoken we can say that alternating r-forms measure r-dimensional volumes. Different alternating r-forms, acting on the same vector space, can be added and multiplied by real numbers, which gives rise to the following definition.

Definition 21. The vector space of alternating r-forms on V is denoted by $\operatorname{Alt}^r V$. Its dimension is $\dim \operatorname{Alt}^r V = \binom{k}{r}$, where $k = \dim V$. We define $\operatorname{Alt}^0 V = \mathbb{R}$

We can define the product of an alternating r-form ω and s-form η. A straightforward definition like $\omega(v_1, \ldots, v_r)\eta(v_{r+1}, \ldots, v_{r+s})$, however, would give a non-alternating form. Therefore, the product has to be anti-symmetrized, which is done defining the wedge product.

Definition 22. Let $\omega \in \operatorname{Alt}^r V$ and $\eta \in \operatorname{Alt}^s V$. Then the wedge product $\wedge : \operatorname{Alt}^r V \times \operatorname{Alt}^s V \to \operatorname{Alt}^{r+s} V$ is defined by:

$$(\omega \wedge \eta)(v_1, ..., v_{r+s}) = \frac{1}{r!s!} \sum_{\tau \in \mathcal{S}_{r+s}} \operatorname{sgn} \tau \; \omega\left(v_{\tau(1)}, ..., v_{\tau(r)}\right) \eta\left(v_{\tau(r+1)}, ..., v_{\tau(r+s)}\right),$$

where $\mathcal{S}_{r+s}$ is the set of all permutations of $(1, \ldots, r+s)$. The wedge product has the following important properties. It is:

(i) bilinear,

(ii) associative $(\eta \wedge \rho) \wedge \phi = \eta \wedge (\rho \wedge \phi)$,

(iii) anti-commutative $\eta \wedge \rho = (-1)^{rs} \rho \wedge \eta$,

(iv) $\eta \wedge 1 = \eta$ for $1 \in \operatorname{Alt}^0 V = \mathbb{R}$.

We notice that alternating 1-forms are linear mappings from V to the real numbers, i.e., $\operatorname{Alt}^0 V \cong V'$. The basis vectors of $\operatorname{Alt}^0 V$ are often denoted by dx^i. Using the wedge product, the basis vectors of Alt^r can be given as:

$$dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_r}\,, \text{ with } \mu_1 < \mu_2 < \cdots < \mu_r. \tag{2.11}$$

If the basis is given in non ascending order, it can be brought to above form, using property (iii) of Definition 22 of the wedge product.

Differential r-Forms

In order to use alternating r-forms on the tangent spaces of a manifold M, we have to assign each $p \in M$ its own alternating r-form.

Definition 23. Let M be a k-dimensional sub-manifold of $\mathbb{R}^n$. We define a **differential r-form** (or shortly r-form) as a mapping:

$$\omega : M \to \operatorname{Alt}^r TM = \bigcup_p \operatorname{Alt}^r T_pM$$

$$p \mapsto \omega_p,$$

with $\omega_p \in \operatorname{Alt}^r T_pM$:

$$\omega_p : \underbrace{T_pM \times \cdots \times T_pM}_{r\text{-times}} \to \mathbb{R}.$$

The vector space of r-forms on a manifold M is denoted by $\Omega^r M$. We define $\Omega^0 M = C^\infty(M, \mathbb{R})$ as the set of smooth functions on M.

Using the representation of the basis of $\operatorname{Alt}^r T_pM$ (2.11), each differential form can be given as follows:

Corollary 2. Let M be a k-dimensional sub-manifold of $\mathbb{R}^n$ and U a chart domain of M. Then $\omega \in \Omega^r U$ can be given by its component functions:

$$\omega = \sum_{\mu_1 < \cdots < \mu_r} \omega_{\mu_1 \ldots \mu_r} du^{\mu_1} \wedge \cdots \wedge du^{\mu_r}. \qquad (2.12)$$

with $\omega_{\mu_1 \ldots \mu_r} = \omega_{\mu_1 \ldots \mu_r}(p)$.

The wedge product of alternating r-forms is transferred pointwise to differential r-forms:

Definition 24. Let M be a k-dimensional sub-manifold of $\mathbb{R}^n$. The wedge product $\wedge : \Omega^r M \times \Omega^s M \to \Omega^{r+s} M$ is defined by:

$$(\omega \wedge \eta)_p = \omega_p \wedge \eta_p \in \operatorname{Alt}^{r+s} T_pM\ ,\ \forall p \in M.$$

We will see later that the wedge product of differential forms on sub-manifolds in $\mathbb{R}^3$ can be interpreted as well-known scalar or vector products of vector fields. As an important concept next we introduce the exterior derivative, which translates the differential operators **grad**, **curl**, and **div** to the language of differential forms. The exterior derivative d is a mapping from $\Omega^r M \to \Omega^{r+1} M$. It is defined by the following definition and theorem:

Definition 25. Let M be a k-dimensional sub-manifold of $\mathbb{R}^n$. There exists a unique sequence

$$\Omega^0 M \xrightarrow{d} \Omega^1 M \xrightarrow{d} \Omega^2 M \xrightarrow{d} \dots$$

of linear mappings which fulfills:

(i) for $f \in \Omega^0 M$, $df \in \Omega^1 M$ is the usual differential of f,

(ii) $d \circ d = 0$,

(iii) product rule: for $\omega \in \Omega^r M$: $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^r \omega \wedge d\eta$.

$d\omega$ is called the **exterior derivative** of ω.

The component representation of the 1-form df on a chart domain U is analogously to (2.12) given by:

$$df = \sum_{\mu=1}^{k} \partial_\mu f du^\mu. \tag{2.13}$$

It can be shown that the coordinate representation of the $r+1$-form $d\omega$ is given by:

$$d\omega = \sum_{\mu_1 < \dots < \mu_r} d(\omega_{\mu_1 \dots \mu_r}) \wedge du^{\mu_1} \wedge \dots \wedge du^{\mu_r}, \tag{2.14}$$

where $d(\omega_{\mu_1 \dots \mu_r})$ is the differential of the component function $\omega_{\mu_1 \dots \mu_r}(p) \in \Omega^0 M$ defined in (2.13). The coordinate representation (2.14) of the exterior derivative will be used to translate d to the differential operators **grad**, **curl**, and **div** of classical vector analysis.

Stokes theorem

Now we are ready to integrate differential forms. For definition of the integral of a differential form, the coordinate representation (2.12) is utilized. For simplicity we look at differential forms whose support is restricted to a single chart domain.

Definition 26. Let M be a k-dimensional oriented [36] sub-manifold of $\mathbb{R}^n$, $\omega \in \Omega^k M$, and U a chart domain of M. Let $\operatorname{supp} \omega \subset U$ and $\varphi : U' \to U$ a parametrization of U. If the k-dimensional Riemann integral

$$\int_{U'} \omega_{\varphi(u')} d^k u' =: \int_M \omega$$

exists, then it does not depend on the choice of the chart and is called **integral of ω over M**.

We notice that k-forms are integrated over k-dimensional manifolds, hence 1-forms are integrated along lines, 2-forms over surfaces, and 3-forms over volumes. Using the exterior derivative, the classical Stokes and Gauß integral theorems can be given in a uniform way:

Theorem 3 (Stokes theorem)**.** Let M be a k-dimensional oriented submanifold with boundary ∂M [36] and $\omega \in \Omega^{k-1}M$ with compact support. Then the following identity holds:

$$\int_M d\omega = \int_{\partial M} \omega. \tag{2.15}$$

Translation to classical vector analysis

Differential forms and the exterior derivative can be interpreted as objects from classical vector analysis. In this Section we give the translation isomorphisms, starting with differential forms.

Differential forms In Definition 21 of alternating r-forms we had the result that $\dim \operatorname{Alt}^r V = \binom{k}{r}$, where k is the dimension of the vector space V. Since later we consider Maxwell's equations in three dimensions we compute:

$$\begin{aligned}
\dim \operatorname{Alt}^0 \mathbb{R}^3 &= 1,\\
\dim \operatorname{Alt}^1 \mathbb{R}^3 &= 3,\\
\dim \operatorname{Alt}^2 \mathbb{R}^3 &= 3,\\
\dim \operatorname{Alt}^3 \mathbb{R}^3 &= 1.
\end{aligned}$$

This motivates the following isomorphism. Let $M \subset \mathbb{R}^3$, and define $\mathcal{F}(M) := C^\infty(M, \mathbb{R}^3)$ as the set of smooth scalar functions and $\mathcal{V}(M) := (C^\infty(M, \mathbb{R}^3))^3$ as the set of smooth vectorial functions in $\mathbb{R}^3$. In the definition of differential forms 23 we already established the connection $\Omega^0 M = \mathcal{F}(M)$. Furthermore, we have:

$$\begin{aligned}
\Omega^1 M &\cong \mathcal{V}(M),\\
\Omega^2 M &\cong \mathcal{V}(M),\\
\Omega^3 M &\cong \mathcal{F}(M).
\end{aligned}$$

How do the elements of $\Omega^r(M)$ act on elements of their domain?

First let $a \in \Omega^1 M$ be a 1-form given in component representation:

$$a = a_1 dx^1 + a_2 dx^2 + a_3 dx^3,$$

where $\{dx^1, dx^2, dx^3\}$ is the dual basis to the standard basis in $\mathbb{R}^3$. Then if $\vec{v} \in \mathbb{R}^3$ we have:

$$\begin{aligned} a(\vec{v}) &= a_1 dx^1(\vec{v}) + a_2 dx^2(\vec{v}) + a_3 dx^3(\vec{v}) \\ &= a_1 v_1 + a_2 v_2 + a_3 v_3, \end{aligned}$$

where v_1, v_2, v_3 are the components of $\vec{v}$ in the standard basis of $\mathbb{R}^3$. Hence, the isomorphism $\Omega^1 M \cong \mathcal{V}(M)$ of a 1-form a is given by:

$$\begin{aligned} &a \xmapsto{\cong} (\vec{a}, \cdot)_{\mathbb{R}^3}\,, \text{ with} \\ &\vec{a} = (a_1, a_2, a_3)\,. \end{aligned} \tag{2.16}$$

Now we consider a differential 2-form $b \in \Omega^2 M$:

$$b = b_1 dx^2 \wedge dx^3 + b_2 dx^3 \wedge dx^1 + b_3 dx^1 \wedge dx^2,$$

and calculate how it acts on two input vectors $\vec{v}, \vec{w}$:

$$\begin{aligned} b(\vec{v}, \vec{w}) &= b_1 dx^2 \wedge dx^3(\vec{v}, \vec{w}) + b_2 dx^3 \wedge dx^1(\vec{v}, \vec{w}) + b_3 dx^1 \wedge dx^2(\vec{v}, \vec{w}) \\ &= b_1 (v_2 w_3 - v_3 w_2) + b_2 (v_3 w_1 - v_1 w_3) + b_3 (v_1 w_2 - v_2 w_1)\,. \end{aligned}$$

This is the Laplace expansion of the determinant $\det\left(\left[\vec{b}, \vec{v}, \vec{w}\right]\right)$, with $\vec{b} = (b_1, b_2, b_3)$. The isomorphism $\Omega^2 M \cong \mathcal{V}(M)$ of a 2-form b is, therefore, given by:

$$\begin{aligned} &b \xmapsto{\cong} \det\left(\vec{b}, \cdot\,, \cdot\right), \text{ with} \\ &\vec{b} = (b_1, b_2, b_3)\,. \end{aligned} \tag{2.17}$$

Finally we look at a 3-form $c \in \Omega^3 M$:

$$c = \rho\, dx^1 \wedge dx^2 \wedge dx^3,$$

and calculate how it acts on three input vectors $\vec{u}, \vec{v}, \vec{w}$:

$$\begin{aligned} c(\vec{u}, \vec{v}, \vec{w}) &= \rho\, dx^1 \wedge dx^2 \wedge dx^3(\vec{u}, \vec{v}, \vec{w}) \\ &= \rho \det(\vec{u}, \vec{v}, \vec{w}). \end{aligned}$$

The corresponding isomorphism $\Omega^3 M \cong \mathcal{F}(M)$ therefore is:

$$c \overset{\cong}{\longmapsto} \rho \det(\cdot\,,\cdot\,,\cdot)\,. \tag{2.18}$$

In summary we can interpret 0- and 3-forms as scalar fields and 1- and 2-forms as vector fields. The corresponding isomorphisms are thereby very illustrative. The answer of a 1-form vector field to a tangential vector of a curve, is the projected component of the vector field in this tangential direction. The answer of a 2-form vector field to two vectors, spanning an area, is the flux of the vector field through this area. And the answer of a 3-form scalar field to three vectors, defining a volume, is a scalar density times this volume. This illustrates the connection to the integration of differential forms.

Exterior derivative Next we analyze, how the exterior derivative acting on differential forms can be interpreted as classical differential operators **grad**, **curl**, and **div**, acting on their isomorphic scalar and vector fields. We start with a differential 0-form f. By definition the exterior derivative d acting on f is simply the differential of f:

$$df = \frac{\partial f}{\partial x_1} dx^1 + \frac{\partial f}{\partial x_2} dx^2 + \frac{\partial f}{\partial x_3} dx^3.$$

Hence, the components of the vector field corresponding to the 1-form df are $\vec{v} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}\right)$, and we have:

$$df \overset{\cong}{\longmapsto} \mathbf{grad}\, f. \tag{2.19}$$

Now let a be a 1-form. We compute:

$$\begin{aligned} da &= d\left(\sum_{i=1}^{3} a_i dx^i\right) \\ &= \sum_{i=1}^{3} (da_i) \wedge dx^i \\ &= \sum_{i=1}^{3} \left(\sum_{j=1}^{3} \frac{\partial a_i}{\partial x_j} dx^j\right) \wedge dx^i \\ &= \left(\frac{\partial a_2}{\partial x_3} - \frac{\partial a_3}{\partial x_2}\right) dx^2 \wedge dx^3 \\ &\quad + \left(\frac{\partial a_3}{\partial x_1} - \frac{\partial a_1}{\partial x_3}\right) dx^3 \wedge dx^1 + \left(\frac{\partial a_1}{\partial x_2} - \frac{\partial a_2}{\partial x_1}\right) dx^1 \wedge dx^2. \end{aligned}$$

This gives the isomorphism for the exterior derivative acting on a 1-form a:

$$da \overset{\cong}{\longmapsto} \textbf{curl}\, \vec{a}. \tag{2.20}$$

Finally let b be a 2-form:

$$\begin{aligned}
db &= d\left(b_1 dx^2 \wedge dx^3 + b_2 dx^3 \wedge dx^1 + b_3 dx^1 \wedge dx^2\right)\\
db &= \sum_{i=1}^{3} \frac{\partial b_1}{\partial x_i} dx^i \wedge dx^2 \wedge dx^3 + \sum_{i=1}^{3} \frac{\partial b_2}{\partial x_i} dx^i \wedge dx^3 \wedge dx^1 + \sum_{i=1}^{3} \frac{\partial b_3}{\partial x_i} dx^i \wedge dx^1 \wedge dx^2\\
&= \frac{\partial b_1}{\partial x_1} dx^1 \wedge dx^2 \wedge dx^3 + \frac{\partial b_2}{\partial x_2} dx^2 \wedge dx^3 \wedge dx^1 + \frac{\partial b_3}{\partial x_3} dx^3 \wedge dx^1 \wedge dx^2\\
&= \left(\frac{\partial b_1}{\partial x_1} + \frac{\partial b_2}{\partial x_2} + \frac{\partial b_3}{\partial x_3}\right) dx^1 \wedge dx^2 \wedge dx^3,
\end{aligned}$$

where we used $dx^i \wedge dx^i = 0$. This gives the isomorphism for db:

$$db \overset{\cong}{\longmapsto} \textbf{div}\, \vec{b}. \tag{2.21}$$

Wedge product The wedge product of differential forms also has analogons to classical vector products. Again, let f be a 0-form, a and a_2 1-forms, b a 2-form, and c a 3-form. Then the wedge product has the following isomorphisms:

$$\begin{aligned}
f \wedge a &\overset{\cong}{\longmapsto} f\vec{a},\\
f \wedge b &\overset{\cong}{\longmapsto} f\vec{b},\\
f \wedge c &\overset{\cong}{\longmapsto} f\vec{\rho},\\
a \wedge a_2 &\overset{\cong}{\longmapsto} \vec{a} \times \vec{a}_2,\\
a \wedge b &\overset{\cong}{\longmapsto} \vec{a} \cdot \vec{b}.
\end{aligned} \tag{2.22}$$

2.2 Maxwell's equations

In this Section we review Maxwell's equations, which describe the dynamics of electromagnetic fields. In our presentation we follow [35, 56].

2.2.1 Maxwell's equations in differential form

Maxwell's equations can be given in different formulations. We start with their differential form:

$$
\begin{aligned}
&\text{Coulombs law:} && \mathbf{div}\ \mathbf{D} = \rho, && (2.23a)\\
&\text{Ampères law:} && \mathbf{curl}\ \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}, && (2.23b)\\
&\text{Faraday's law of induction:} && \mathbf{curl}\ \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, && (2.23c)\\
&\text{absence of magnetic monopoles:} && \mathbf{div}\ \mathbf{B} = 0. && (2.23d)
\end{aligned}
$$

All of above fields depend on the spatial variable $\mathbf{x}$ and time t. $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic field strength, $\mathbf{D}$ is the electric displacement, and $\mathbf{B}$ the magnetic flux density. ρ is the macroscopic charge density and $\mathbf{J}$ is the macroscopic electric current density. The time derivative of the electric displacement in Ampères law was added by Maxwell and named displacement current. Since Maxwell was the first to state the above equations in the correct form, they were named after him.

Taking the time derivative of (2.23a) and adding the divergence of (2.23b), gives the continuity equation for the electric charge:

$$
\frac{\partial \rho}{\partial t} + \mathbf{div}\ \mathbf{J} = 0,
$$

where we used $\mathbf{div}\ \mathbf{curl}\ = 0$.

Maxwell's equations are under-determined. We have 8 equations for 12 field components of $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$, and $\mathbf{H}$. Furthermore, $\mathbf{E}$ and $\mathbf{B}$ are not independent. Due to (2.23d) we can introduce the vector potential $\mathbf{A}$ such that:

$$
\mathbf{B} = \mathbf{curl}\ \mathbf{A}.
$$

Using this relation in (2.23c) we get

$$
\mathbf{curl}\ \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0.
$$

Since the curl of above expression vanishes, it can be represented as the gradient of a scalar potential ϕ:

$$-\mathbf{grad}\ \phi = \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}.$$

Hence, we have 4 independent quantities, fixing $\mathbf{E}$ and $\mathbf{B}$. This means, we need six equations in addition to Maxwell's equations, in order to get unique solutions to the system. These are so-called constitutive relations, which describe the dependence of the electric displacement and magnetic field on the electric field and magnetic flux density:

$$\mathbf{D} = \mathbf{D}(\mathbf{E}, \mathbf{B}), \tag{2.24a}$$

$$\mathbf{H} = \mathbf{H}(\mathbf{E}, \mathbf{B}). \tag{2.24b}$$

Above relations are used to describe the behaviour of matter in the presence of electric and magnetic fields.

2.2.2 Maxwell's equations in integral form

Now we give the integral representation of Maxwell's equations. Let V be a volume with boundary ∂V and A a surface with boundary ∂A in $\mathbb{R}^3$:

Coulombs law:
$$\oiint_{\partial V} \mathbf{D} \cdot d\mathbf{f} = \iiint_V \rho dV, \tag{2.25a}$$

Ampères law:
$$\iint_A \left(\frac{\partial}{\partial t}\mathbf{D} + \mathbf{J} \right) \cdot d\mathbf{f} = \oint_{\partial A} \mathbf{H} \cdot d\mathbf{r}, \tag{2.25b}$$

Faraday's law of induction:
$$-\frac{\partial}{\partial t} \iint_A \mathbf{B} \cdot d\mathbf{f} = \oint_{\partial A} \mathbf{E} \cdot d\mathbf{r}, \tag{2.25c}$$

absence of magnetic monopoles:
$$\oiint_{\partial V} \mathbf{B} \cdot d\mathbf{f} = 0. \tag{2.25d}$$

We notice that the electric and magnetic field $\mathbf{E}$ and $\mathbf{H}$ are integrated along one dimensional lines. They can be interpreted as differential 1-forms, which were introduced in Section 2.1.2. The electric displacement $\mathbf{D}$, magnetic flux density $\mathbf{B}$, and macroscopic electric current density $\mathbf{J}$ are integrated over surfaces and are differential 2-forms. Finally the charge density ρ is integrated over a volume and can be interpreted as a differential 3-form.

The integral representation of Maxwell's equations allows to analyze how the electromagnetic field behaves at a material interface S [56]:

$$\mathbf{n} \cdot [\mathbf{D}] = \rho_S, \tag{2.26a}$$
$$\mathbf{n} \times [\mathbf{H}] = 0, \tag{2.26b}$$
$$\mathbf{n} \times [\mathbf{E}] = 0, \tag{2.26c}$$
$$\mathbf{n} \cdot [\mathbf{B}] = 0, \tag{2.26d}$$

where $[\cdot]$ denotes the jump of the field across S, $\mathbf{n}$ is the normal vector of S, and ρ_S the surface charge density. Hence, the tangential component of $\mathbf{E}$ and $\mathbf{H}$, and the normal component of $\mathbf{D}$ (in the absence of a surface charge) and $\mathbf{B}$ are continuous across surfaces.

2.2.3 Time-harmonic Maxwell's equations

For stationary electromagnetic problems, which are in the focus of this work, the time dependence in Maxwell's equations can be separated making a time-harmonic ansatz. Choosing:

$$\mathbf{E}(\mathbf{x}, t) = \hat{\mathbf{E}}(\mathbf{x}, \omega) e^{-i\omega t}, \tag{2.27}$$

with frequency ω, the time derivative becomes $\frac{\partial}{\partial t} \rightarrow -i\omega$. Making the same ansatz for $\mathbf{D}$, $\mathbf{H}$, $\mathbf{B}$, $\mathbf{J}$, and ρ, Maxwell's equations read:

$$\textbf{div}\ \hat{\mathbf{D}} = \hat{\rho}, \tag{2.28a}$$
$$\textbf{curl}\ \hat{\mathbf{H}} + i\omega\hat{\mathbf{D}} = \hat{\mathbf{J}}, \tag{2.28b}$$
$$\textbf{curl}\ \hat{\mathbf{E}} - i\omega\hat{\mathbf{B}} = 0, \tag{2.28c}$$
$$\textbf{div}\ \hat{\mathbf{B}} = 0. \tag{2.28d}$$

Physical fields are always real-valued, whereas the solutions to time-harmonic Maxwell's equations are now complex-valued, containing phase information. Taking the real part of a solution $\Re\left(\hat{\mathbf{E}}\right)$, one recovers the real-valued field at a specific time t.

The continuity equation in the time-harmonic case reads:

$$\textbf{div}\ \hat{\mathbf{J}} - i\omega\hat{\rho} = 0 \tag{2.29}$$

In the following we will only consider time-harmonic fields and drop the hats of the complex-valued, so-called phasors.

Furthermore, we only consider linear material relationships (2.24a), (2.24b). The constitutive relationships can then be stated as:

$$\mathbf{D} = \epsilon\mathbf{E}, \tag{2.30a}$$
$$\mathbf{B} = \mu\mathbf{H}, \tag{2.30b}$$

where ϵ and μ is the permittivity and permeability tensor respectively. Finally, we assume that the macroscopic electric current density can be separated into a part which depends linearly on the electric field via a conductivity (charge carriers in material) and an impressed part originating from external sources:

$$\mathbf{J} = \sigma \mathbf{E} + \mathbf{J}_{\mathrm{i}}, \tag{2.31}$$

where σ is the conductivity tensor. We want to reduce the full coupled system of Maxwell's equations to a single equation for the electric field $\mathbf{E}$. Therefore, we use the linear constitutive relationships (2.30a) and (2.30b) and Maxwell's equations (2.28b) and (2.28c):

$$\begin{aligned}
\mathbf{curl}\ \mathbf{E} - i\omega \mathbf{B} &= 0 && \Leftrightarrow \\
\mu^{-1}\mathbf{curl}\ \mathbf{E} - i\omega \mathbf{H} &= 0 && \Leftrightarrow \\
\mathbf{curl}\ \mu^{-1}\mathbf{curl}\ \mathbf{E} - i\omega\, \mathbf{curl}\ \mathbf{H} &= 0 && \Leftrightarrow \\
\mathbf{curl}\ \mu^{-1}\mathbf{curl}\ \mathbf{E} - i\omega \left(\mathbf{J} - i\omega \mathbf{D}\right) &= 0 && \Leftrightarrow \\
\mathbf{curl}\ \mu^{-1}\mathbf{curl}\ \mathbf{E} - i\omega \left(\sigma - i\omega\epsilon\right) \mathbf{E} &= i\omega \mathbf{J}_{\mathrm{i}} && \Leftrightarrow \\
\mathbf{curl}\ \mu^{-1}\mathbf{curl}\ \mathbf{E} - \omega^2 \varepsilon \mathbf{E} &= i\omega \mathbf{J}_{\mathrm{i}},
\end{aligned}$$

with the complex permittivity tensor $\varepsilon = \epsilon + \frac{i}{\omega}\sigma$. Because of its importance, we state the final equation again:

$$\mathbf{curl}\ \mu^{-1}\mathbf{curl}\ \mathbf{E} - \omega^2 \varepsilon \mathbf{E} = i\omega \mathbf{J}_{\mathrm{i}}. \tag{2.32}$$

This will be the formulation of Maxwell's equations used throughout the thesis.

2.2.4 Maxwell's equations with exterior calculus

Finally we state Maxwell's equations with differential forms. The integral formulation (2.25a)-(2.25d) showed that $\mathbf{E}$ and $\mathbf{H}$ can be interpreted as 1-forms, $\mathbf{D}$, $\mathbf{J}$, and $\mathbf{B}$ as 2-forms, and ρ as a 3-form. Since $\mathbf{D}$ and $\mathbf{B}$ are 2-forms, $d\mathbf{D}$ and $d\mathbf{B}$ correspond to $\mathbf{div}\ \mathbf{D}$ and $\mathbf{div}\ \mathbf{B}$. The exterior derivative acting on the 1-forms $\mathbf{E}$ and $\mathbf{H}$ corresponds to the **curl** operator, c.f., (2.23b) and (2.23c). Using the exterior derivative, Maxwell's equations can then be stated as follows:

$$\begin{aligned}
d\mathbf{D} &= \rho, \\
d\mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} &= \mathbf{J}, \\
d\mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0, \\
d\mathbf{B} &= 0.
\end{aligned} \tag{2.33}$$

The constitutive relationships (2.30a) and (2.30b) connect 1- and 2-forms, and involve the Hodge operator [20]:

$$\begin{aligned} \mathbf{D} &= *_\varepsilon \mathbf{E}, \\ \mathbf{B} &= *_\mu \mathbf{H}. \end{aligned} \tag{2.34}$$

Using these relationships, the curl-curl equation for the electric field (2.32) can be given as:

$$d *_\mu^{-1} d\mathbf{E} - \omega^2 *_\varepsilon \mathbf{E} = i\omega \mathbf{J}_\mathrm{i}. \tag{2.35}$$

Coordinate transformation

For application of the reduced basis method to geometrically parametrized PDEs, we have to know how Maxwell's equations behave under a coordinate transformation. The curl-curl equation for the electric field (2.35) is stated covariantly. It holds in each coordinate system. However, since we have to use a coordinate representation, when we solve Maxwell's equations, we have to investigate, how the components of differential forms transform. Let us consider a coordinate transformation G form $\overline{\Omega}$ to Ω:

$$\begin{aligned} G : \overline{\Omega} &\to \Omega \\ x_i &= G_i(\overline{x}_j). \end{aligned}$$

The Jacobian of G is defined as:

$$\begin{aligned} J_{ij} &= \frac{\partial G_i}{\partial \overline{x}_j}, \\ |J| &= \det\left(J\right). \end{aligned}$$

The coordinate differentials then transform according to:

$$dx^i = \frac{\partial G_i}{\partial \overline{x}_j} d\overline{x}^j = J_{ij} d\overline{x}^j.$$

In this Section we make use of the summation convention, which states that a summation has to be carried out over indices, which appear twice in a single expression. We first consider a 1-form a and express it in the original and transformed coordinate system:

$$\begin{aligned} a &= a_i dx^i \\ &= \overline{a}_i d\overline{x}^j \\ &= \overline{a}_j J_{ji}^{-1} dx^i. \end{aligned}$$

It follows that the components of a 1-form transform according to

$$\vec{a} = J^{-T}\,\vec{\overline{a}}. \tag{2.36}$$

Now we consider a 2-form b:

$$\begin{aligned} b &= b_1\,dx^2 \wedge dx^3 + b_2\,dx^3 \wedge dx^1 + b_3\,dx^1 \wedge dx^2 \qquad (2.37)\\ &= \overline{b}_1\,d\overline{x}^2 \wedge d\overline{x}^3 + \overline{b}_2\,d\overline{x}^3 \wedge d\overline{x}^1 + \overline{b}_3\,d\overline{x}^1 \wedge d\overline{x}^2\\ &= \overline{b}_1\,J^{-1}_{2m}dx^m \wedge J^{-1}_{3n}dx^n + \overline{b}_2\,J^{-1}_{3m}dx^m \wedge J^{-1}_{1n}dx^n + \overline{b}_3\,J^{-1}_{1m}dx^m \wedge J^{-1}_{2n}dx^n \qquad (2.38) \end{aligned}$$

Let us look at the factor of $\overline{b}_1$ as an example:

$$\begin{aligned} J^{-1}_{2m}dx^m \wedge J^{-1}_{3n}dx^n &= \left(J^{-1}_{22}J^{-1}_{33} - J^{-1}_{23}J^{-1}_{32}\right)dx^2 \wedge dx^3\\ &\quad + \left(J^{-1}_{23}J^{-1}_{31} - J^{-1}_{21}J^{-1}_{33}\right)dx^3 \wedge dx^1\\ &\quad + \left(J^{-1}_{21}J^{-1}_{32} - J^{-1}_{22}J^{-1}_{31}\right)dx^1 \wedge dx^2\\ &= \left(J^{-1}\right)^{\mathrm{adj}}_{11}dx^2 \wedge dx^3 + \left(J^{-1}\right)^{\mathrm{adj}}_{12}dx^3 \wedge dx^1 + \left(J^{-1}\right)^{\mathrm{adj}}_{13}dx^1 \wedge dx^2, \end{aligned}$$

where A^{adj} is the adjunct of matrix A. With:

$$\begin{aligned} A^{-1} &= \frac{1}{|A|}\left(A^{\mathrm{adj}}\right)^T \quad \text{and}\\ |A| &= \frac{1}{|A^{-1}|}, \end{aligned}$$

we can rewrite the last equation and get:

$$J^{-1}_{2m}dx^m \wedge J^{-1}_{3n}dx^n = \frac{1}{|J|}\left(J^T_{11}dx^2 \wedge dx^3 + J^T_{12}dx^3 \wedge dx^1 + J^T_{13}dx^1 \wedge dx^2\right).$$

The factors of $\overline{b}_2$ and $\overline{b}_3$ in (2.38) give corresponding expressions, such that (2.38) can be written as:

$$\begin{aligned} b =& \frac{1}{|J|}\left(\overline{b}_1J^T_{11} + \overline{b}_2J^T_{21} + \overline{b}_3J^T_{31}\right)dx^2 \wedge dx^3\\ &+ \frac{1}{|J|}\left(\overline{b}_1J^T_{12} + \overline{b}_2J^T_{22} + \overline{b}_3J^T_{32}\right)dx^3 \wedge dx^1\\ &+ \frac{1}{|J|}\left(\overline{b}_1J^T_{13} + \overline{b}_2J^T_{23} + \overline{b}_3J^T_{33}\right)dx^1 \wedge dx^2. \end{aligned}$$

A comparison with (2.37) gives the transformation rule for the components of a 2-form:

$$\vec{b} = \frac{1}{|J|}J\,\vec{\overline{b}}. \tag{2.39}$$

Finally we consider a 3-form:

$$\begin{aligned} c &= \rho\, dx^1 \wedge dx^2 \wedge dx^3 \\ &= \overline{\rho}\, d\overline{x}^1 \wedge d\overline{x}^2 \wedge d\overline{x}^3 \\ &= \overline{\rho}\, J^{-1}_{1l} J^{-1}_{2m} J^{-1}_{3n} dx^l \wedge dx^m \wedge dx^n \\ &= \overline{\rho}\, |J^{-1}| dx^1 \wedge dx^2 \wedge dx^3 \\ &= \frac{1}{|J|} \overline{\rho}\, dx^1 \wedge dx^2 \wedge dx^3. \end{aligned}$$

This gives the transformation rule for a 3-form:

$$\rho = \frac{1}{|J|}\overline{\rho}. \tag{2.40}$$

Now we analyze how the components of the permittivity and permeability tensors in Maxwell's Equations transform:

$$d *_\mu^{-1} d\mathbf{E} - \omega^2 *_\varepsilon \mathbf{E} = i\omega \mathbf{J}_i.$$

We start with the permittivity tensor, and use transformation rule (2.39) of a 2-form. Remembering that $*_\varepsilon \mathbf{E}$ is a 2-form gives:

$$*_\varepsilon \mathbf{E} = \frac{1}{|J|} J\, \overline{*_\varepsilon \mathbf{E}}.$$

Now we use:

$$\overline{*_\varepsilon \mathbf{E}} = \overline{*_\varepsilon}\, \overline{\mathbf{E}},$$

and transformation rule (2.36) for 1-forms, to transform $\overline{\mathbf{E}}$:

$$\begin{aligned} *_\varepsilon \mathbf{E} &= \frac{1}{|J|} J\, \overline{*_\varepsilon \mathbf{E}} \\ &= \frac{1}{|J|} J\, \overline{*_\varepsilon}\, \overline{\mathbf{E}} \\ &= \frac{1}{|J|} J\, \overline{*_\varepsilon} J^T \mathbf{E}. \end{aligned}$$

This gives the following transformation rule for the components of the permittivity tensor:

$$\overline{*_\varepsilon} = |J| J^{-1} *_\varepsilon J^{-T}. \tag{2.41}$$

For the permeability tensor we first use transformation rule (2.36) and respect that $*_\mu^{-1} d\mathbf{E}$ is a 1-form:

$$*_\mu^{-1} d\mathbf{E} = J^{-T}\, \overline{*_\mu^{-1} d\mathbf{E}}.$$

Again we make use of:

$$\overline{*_\mu^{-1} d\mathbf{E}} = \overline{*_\mu^{-1}}\ \overline{d\mathbf{E}},$$

and apply transformation rule (2.39) to the 2-form $d\mathbf{E}$:

$$\begin{aligned} *_\mu^{-1} d\mathbf{E} &= J^{-T}\, \overline{*_\mu^{-1}}\ \overline{d\mathbf{E}} \\ &= J^{-T}\ \overline{*_\mu^{-1}}\, |J| J^{-1} d\mathbf{E}. \end{aligned}$$

This gives the transformation rule for the components of the inverse permeability tensor:

$$\overline{*_\mu^{-1}} = \frac{1}{|J|} J^T *_\mu^{-1} J. \tag{2.42}$$

3 Electromagnetic scattering problem

Having set up the mathematical background and Maxwell's equations, we now derive the formulation of an electromagnetic scattering problem, which is suited for discretization with finite elements. Since scattering problems are per se stated on unbounded domains, we thereby also comment on transparent boundary conditions. In particular we focus on the perfectly matched layer (PML) method [8], which we use in numerical simulations.

3.1 Strong formulation of the scattering problem

The setup for a scattering problem is depicted in Fig. 3.1. The unbounded domain $\mathbb{R}^3$ is divided into an interior domain Ω_{int} with boundary Γ and an exterior domain $\mathbb{R}^3 \setminus \Omega_{\text{int}}$. We assume that Γ is a polygon with outward pointing normal vector $\mathbf{n}$. The incoming electric source field $\mathbf{E}_{\text{in}}$ is a solution to Maxwell's equations in the exterior. It enters the domain Ω_{int} and is scattered by an obstacle S. The scattered field $\mathbf{E}_{\text{sc}}$ leaves the interior domain and is, therefore, strictly outward radiating. The field within the interior domain Ω_{int} is denoted by $\mathbf{E}$.

The strong formulation for the scattering problem is given by:

Problem 2. Find $\mathbf{E}$ such that:

(i) The electric field $\mathbf{E}$ fulfills Maxwell's equations in the interior domain:

$$d *_{\mu}^{-1} d\mathbf{E} - \omega^2 *_{\varepsilon} \mathbf{E} = 0 \text{ in } \Omega_{\text{int}}. \tag{3.1}$$

(ii) The scattered field $\mathbf{E}_{\text{sc}}$ fulfills Maxwell's equations in the exterior domain:

$$d *_{\mu}^{-1} d\mathbf{E}_{\text{sc}} - \omega^2 *_{\varepsilon} \mathbf{E}_{\text{sc}} = 0 \text{ in } \mathbb{R}^3 \setminus \Omega_{\text{int}}.$$

(iii) Boundary condition at Γ: the tangential component of the electric field is continuous:

$$\mathbf{n} \times (\mathbf{E}_{\text{in}} + \mathbf{E}_{\text{sc}} - \mathbf{E})|_{\Gamma} = 0,$$

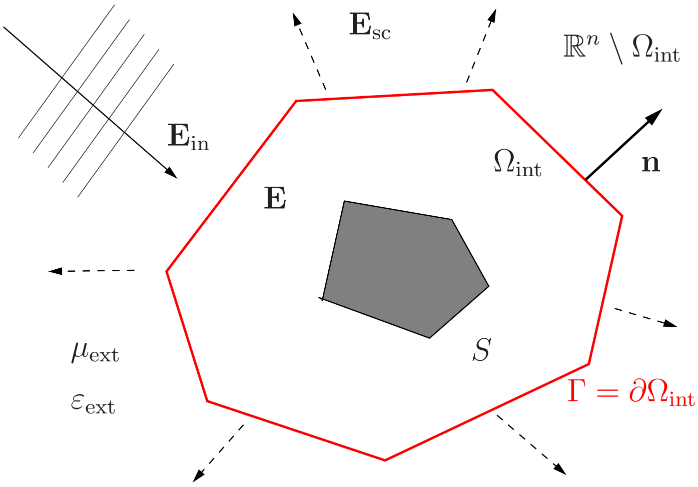

$\mathbf{E}_{\mathrm{sc}}$
$\mathbb{R}^n \setminus \Omega_{\mathrm{int}}$
$\mathbf{E}_{\mathrm{in}}$
Ω_{int}
$\mathbf{n}$
$\mathbf{E}$
S
μ_{ext}
$\varepsilon_{\mathrm{ext}}$
$\Gamma = \partial\Omega_{\mathrm{int}}$

Figure 3.1: Setup of a scattering problem. The interior domain Ω_{int} contains the scatterer S and is embedded into an infinite exterior $\mathbb{R}^3$ with relative permittivity $\varepsilon_{\mathrm{ext}}$ and permeability μ_{ext}. The incoming electric field $\mathbf{E}_{\mathrm{in}}$ is entering the interior domain via the boundary Γ and is the source for the electric field $\mathbf{E}$ inside Ω_{int}. The scattered field $\mathbf{E}_{\mathrm{sc}}$ is originated within Ω_{int}. It is, therefore, strictly outward radiating.

where the incoming field $\mathbf{E}_{\mathrm{in}}$ has to fulfill Maxwell's equations in a neighborhood of the boundary Γ in the exterior.

(iv) Radiating boundary condition: $\mathbf{E}_{\mathrm{sc}}$ is strictly outward radiating.
E.g., Silver-Müller radiation condition:

$$\lim_{r\to\infty} r \left(\mathbf{curl}\, \mathbf{E}_{\mathrm{sc}}(\mathbf{r}) \times \mathbf{r}_0 - i \frac{\omega\sqrt{\varepsilon_{\mathrm{ext}}\mu_{\mathrm{ext}}}}{c} \mathbf{curl}\, \mathbf{E}_{\mathrm{sc}}(\mathbf{r}) \right) = 0$$

$$\text{uniformly continuous in each direction } \mathbf{r}_0,$$

where $\mathbf{r}$ is the coordinate vector in $\mathbb{R}^3$, r its norm, $\mathbf{r}_0 = \frac{\mathbf{r}}{r}$, $\varepsilon_{\mathrm{ext}}$ and μ_{ext} the relative permittivity and permeability in the exterior, and c the speed of light in vacuum.

Above setup is a coupled interior-exterior problem for the electric field with no impressed sources, i.e.:

$$\mathbf{J}_{\mathrm{i}} = 0.$$

The data of the incoming field $\mathbf{E}_{\mathrm{in}}$ enters via the continuity condition stated on the boundary Γ.

3.2 Weak formulation of the scattering problem

For discretization with finite elements we have to derive a weak formulation of the strong scattering Problem 2.

In the following we will denote the electric field by u. We start with the curl-curl form of Maxwell's equations (2.35), test it with the complex conjugate $\bar{v}$ of a function v, and integrate over $\mathbb{R}^3$:

$$0 = \int\limits_{\mathbb{R}^3} \bar{v} d *_\mu^{-1} du - \omega^2 \bar{v} *_\varepsilon u.$$

The first term is integrated by parts, where we assume that u vanishes at infinity. This gives:

$$0 = \int\limits_{\mathbb{R}^3} d\bar{v} \wedge *_\mu^{-1} du - \omega^2 \bar{v} \wedge *_\varepsilon u. \tag{3.2}$$

According to Fig. 3.1 the infinite domain is divided into an interior and exterior: $\mathbb{R}^3 = \Omega_{\text{int}} \cup (\mathbb{R}^3 \setminus \Omega_{\text{int}})$. In the exterior we decompose the electric field by $u = u_{\text{in}} + u_{\text{sc}}$, which gives:

$$\begin{aligned} 0 = &\int\limits_{\Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du - \omega^2 \bar{v} \wedge *_\varepsilon u \\ &+ \int\limits_{\mathbb{R}^3 \setminus \Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du_{\text{sc}} - \omega^2 \bar{v} \wedge *_\varepsilon u_{\text{sc}} + \int\limits_{\mathbb{R}^3 \setminus \Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du_{\text{in}} - \omega^2 \bar{v} \wedge *_\varepsilon u_{\text{in}} \ , \end{aligned} \tag{3.3}$$

where the continuity condition:

$$\mathbf{n} \times (u_{\text{in}} + u_{\text{sc}} - u)|_\Gamma = 0, \tag{3.4}$$

has to hold. Now we use the product rule:

$$\begin{aligned} d(v \wedge u) &= (dv) \wedge u - v \wedge du & \Leftrightarrow \\ (dv) \wedge u &= d(v \wedge u) + v \wedge du, & \end{aligned}$$

on the last integral in (3.3):

$$\int\limits_{\mathbb{R}^3\backslash\Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du_{\text{in}} - \omega^2 \bar{v} \wedge *_\varepsilon u_{\text{in}} =$$

$$\int\limits_{\mathbb{R}^3\backslash\Omega_{\text{int}}} d(\bar{v} \wedge *_\mu^{-1} du_{\text{in}}) + \bar{v} \wedge d *_\mu^{-1} du_{\text{in}} - \omega^2 \bar{v} \wedge *_\varepsilon u_{\text{in}} =$$

$$\int\limits_{\mathbb{R}^3\backslash\Omega_{\text{int}}} d(\bar{v} \wedge *_\mu^{-1} du_{\text{in}}) + \bar{v} \wedge \underbrace{\left(d *_\mu^{-1} du_{\text{in}} - \omega^2 *_\varepsilon u_{\text{in}}\right)}_{=0} = \int\limits_{\mathbb{R}^3\backslash\Omega_{\text{int}}\Gamma} \bar{v} \wedge *_\mu^{-1} du_{\text{in}},$$

where we used Stoke's theorem (2.15) and the fact that u_{in} fulfills Maxwell's equations in the exterior. Equation (3.3) then reads:

$$\begin{aligned} 0 = &\int\limits_{\Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du - \omega^2 \bar{v} \wedge *_\varepsilon u \\ &+ \int\limits_{\mathbb{R}^3\backslash\Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du_{\text{sc}} - \omega^2 \bar{v} \wedge *_\varepsilon u_{\text{sc}} + \int\limits_{\mathbb{R}^3\backslash\Omega_{\text{int}}\Gamma} \bar{v} \wedge *_\mu^{-1} du_{\text{in}}. \end{aligned} \tag{3.5}$$

Now we want to incorporate the continuity condition (3.4) and therefore define:

$$\tilde{u}_{\text{sc}} = u_{\text{sc}} + \mathscr{P}[u_{\text{in}}], \tag{3.6}$$

where $\mathscr{P}[u_{\text{in}}]$ is an operator which interpolates the tangential component of u_{in} on Γ, and $\mathscr{P}[u_{\text{in}}]$ has only support in a small region $\Omega_\delta \subset \mathbb{R}^3 \setminus \Omega$ in the exterior [84]. We use this ansatz in (3.5) and get:

$$\int\limits_{\Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du - \omega^2 \bar{v} \wedge *_\varepsilon u + \int\limits_{\mathbb{R}^3\backslash\Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} d\tilde{u}_{\text{sc}} - \omega^2 \bar{v} \wedge *_\varepsilon \tilde{u}_{\text{sc}} = f[u_{\text{in}}](v),$$

with:

$$f[u_{\text{in}}](v) = \int\limits_{\Omega_\delta} d\bar{v} \wedge *_\mu^{-1} d\mathscr{P}[u_{\text{in}}] - \omega^2 \bar{v} \wedge *_\varepsilon \mathscr{P}[u_{\text{in}}] - \int\limits_{\Gamma} \bar{v} \wedge *_\mu^{-1} du_{\text{in}}. \tag{3.7}$$

Since u and $\tilde{u}_{\text{sc}}$ are continuous across Γ, we can glue them together and define:

$$\tilde{u} = \left\{ \begin{array}{cl} u & x \in \Omega_{\text{int}} \\ \tilde{u}_{\text{sc}} & x \in \mathbb{R}^3 \setminus \Omega_{\text{int}} \end{array} \right. . \tag{3.8}$$

This gives:

$$\int\limits_{\mathbb{R}^3} d\bar{v} \wedge *_{\mu}^{-1} d\tilde{u} - \omega^2 \bar{v} \wedge *_{\varepsilon} \tilde{u} = f[u_{\text{in}}](v).$$

We introduce the following abbreviations:

$$a^{\text{id}}(u, v) := \int\limits_{\mathbb{R}^3} d\bar{v} \wedge *_{\mu}^{-1} d\tilde{u} - \omega^2 \bar{v} \wedge *_{\varepsilon} \tilde{u}, \tag{3.9}$$

$$f(v) = f[u_{\text{in}}](v), \tag{3.10}$$

where “id” refers to infinite domain. Now we can state the weak formulation of the electromagnetic scattering problem.

Problem 3. Find $u \in \mathrm{H}\left(\mathbf{curl}\,, \mathbb{R}^3\right)$ such that:

$$a^{\text{id}}(u, v) = f(v)\,, \quad \forall v \in \mathrm{H}\left(\mathbf{curl}\,, \mathbb{R}^3\right). \tag{3.11}$$

Using the isomorphisms of exterior calculus introduced in Section 2.1.2, we can give a^{id} and f in the notation of classical vector analysis:

$$a^{\text{id}}(\vec{u}, \vec{v}) := \int\limits_{\mathbb{R}^3} \left(\mathbf{curl}\, \bar{\vec{v}} \cdot \mu^{-1} \cdot \mathbf{curl}\, \tilde{\vec{u}} - \omega^2\, \bar{\vec{v}} \cdot \varepsilon \cdot \tilde{\vec{u}} \right) dV, \tag{3.12}$$

$$\begin{aligned} f[\vec{u}_{\text{in}}](\vec{v}) = & \int\limits_{\Omega_\delta} \left(\mathbf{curl}\, \bar{\vec{v}} \cdot \mu^{-1} \mathbf{curl}\, \mathscr{P}[\vec{u}_{\text{in}}] - \omega^2\, \bar{\vec{v}} \cdot \varepsilon \cdot \mathscr{P}[\vec{u}_{\text{in}}] \right) dV \\ & - \int\limits_{\Gamma} \left(\bar{\vec{v}} \times \left[\mu^{-1} \cdot \mathbf{curl}\, \vec{u}_{\text{in}} \right] \right) \cdot d\vec{A}. \end{aligned} \tag{3.13}$$

3.3 Transparent boundary conditions

The weak formulation of the scattering Problem (3.11) is stated on an unbounded domain, making a straightforward discretization and numerical computation difficult. Truncating the domain and stating Dirichlet, Neumann, or Robin boundary conditions, results in artificial reflections at the introduced artificial boundaries. Therefore, transparent (or radiating) boundary conditions have to be introduced. These boundary conditions have to assure that the radiation condition holds, i.e., the scattered field is strictly outward radiating.

3.3.1 Perfectly matched layers

In the present work we use the perfectly matched layer (PML) method as a transparent boundary condition [8, 108]. For its application a special coordinate system is introduced in the exterior, which includes a generalized distance variable ξ. The basic idea of the PML method is a complex coordinate stretching in radial direction ξ in the exterior $\mathbb{R}^3 \setminus \Omega_{\text{int}}$:

$$\tilde{\xi} = (1 + i\sigma)\xi \,, \quad 0 < \sigma \in \mathbb{R}.$$

The transformed scattered field is then the complex continuation of the original field:

$$\tilde{u}_{\text{sc}}^{\text{PML}}(\cdot) = \tilde{u}_{\text{sc}}([1 + i\sigma] \cdot). \tag{3.14}$$

The complex continuation $\tilde{u}_{\text{sc}}^{\text{PML}}$ is used in the exterior, c.f. ansatz (3.8):

$$\tilde{u}^{\text{PML}} = \left\{ \begin{array}{cl} u & x \in \Omega_{\text{int}} \\ \tilde{u}_{\text{sc}}^{\text{PML}} & x \in \mathbb{R}^3 \setminus \Omega_{\text{int}} \end{array} \right. . \tag{3.15}$$

Since $\tilde{u}_{\text{sc}}$ is outward radiating, its complex continuation is exponentially decreasing [8]. After sufficient decay of the scattered field, the exterior domain can then be truncated according to:

$$\mathbb{R}^3 \setminus \Omega_{\text{int}} \to \Omega_{\text{PML}}. \tag{3.16}$$

The integral over the infinite exterior in the definition of the bilinear form a^{id} (3.9) is truncated accordingly:

$$\int\limits_{\Omega} d\bar{v} \wedge *_{\mu}^{-1} d\tilde{u}^{\text{PML}} - \omega^2 \bar{v} \wedge *_{\varepsilon} \tilde{u}^{\text{PML}} = f[u_{\text{in}}](v), \tag{3.17}$$

where we set $\Omega = \Omega_{\text{int}} \cup \Omega_{\text{PML}}$. On the artificial boundary of Ω, which is introduced by the PML truncation (3.16), zero Dirichlet or Neumann boundary conditions are stated for u^{PML}. It is worth noting that according to Section 2.2.4, the complex coordinate stretching (i.e. coordinate transformation) simply results in a transformation of the permittivity and permeability tensors. Hence, the finite element discretization of the exterior domain Ω_{PML} offers no additional difficulties.

For brevity we introduce:

$$a(u, v) = \int\limits_{\Omega} d\bar{v} \wedge *_{\mu}^{-1} d\tilde{u}^{\text{PML}} - \omega^2 \bar{v} \wedge *_{\varepsilon} \tilde{u}^{\text{PML}}, \tag{3.18}$$

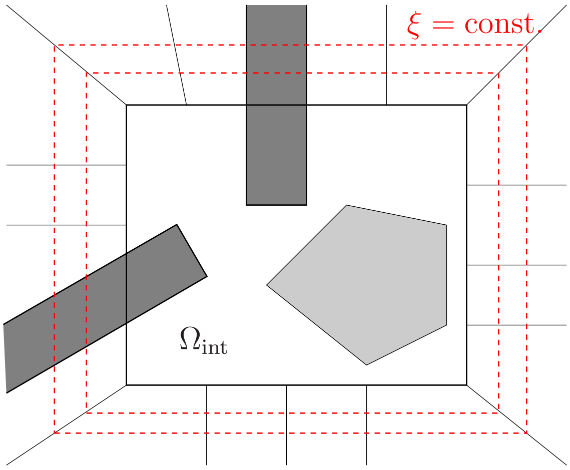


Figure 3.2: Interior domain Ω_{int} with structured exterior and corresponding prismatoidal coordinate system. Two coordinate lines of constant ξ, referring to generalized distance variable are shown.

and suppress the dependence of the right hand side functional on u_{in}:

$$f(v) = f[u_{\text{in}}](v).$$

Now we state the electromagnetic scattering problem with perfectly matched layers as transparent boundary conditions:

Problem 4. Find $u \in \mathrm{H}\,(\mathbf{curl}, \Omega)$ such that

$$a(u, v) = f(v)\,, \quad \forall v \in \mathrm{H}\,(\mathbf{curl}, \Omega)\,. \tag{3.19}$$

This is the continuous problem formulation, which we will discretize with finite elements.

Radial coordinate system

In the following we give details of our realization of the PML and the radial coordinate system. For simplicity we consider the 2D case. For application to typical nano-optical systems, the exterior coordinate system should allow discretization of multiply structured exterior domains, which could contain semi-infinite layers or waveguide structures.

We use a prismatoidal coordinate system with semi-infinite rays in the exterior, as depicted in Fig. 3.2. The coordinate system for a specific ray is shown in Fig. 3.3. Since Maxwell's equations are rotational and translational

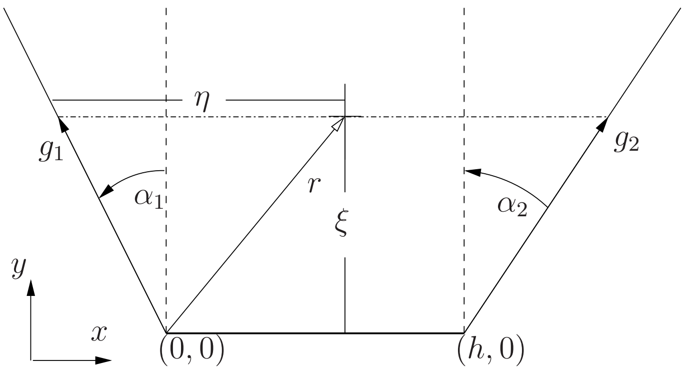


Figure 3.3: Prismatoidal coordinate system for exterior rays, c.f. Fig. 3.2.

invariant, we located the baseline of the exterior prism between $(0,0)$ and $(h,0)$. The coordinate transformation from the prismatoidal to Cartesian coordinates is given by:

$$\begin{aligned} \begin{pmatrix} x(\eta,\xi) \\ y(\eta,\xi) \end{pmatrix} &= g_1(\xi) + \eta\left(\begin{pmatrix} h \\ 0 \end{pmatrix} + g_2(\xi) - g_1(\xi)\right), \\ g_1(\xi) &= \frac{\xi}{\zeta}\begin{pmatrix} -\tan\alpha_1 \\ 1 \end{pmatrix}, \\ g_2(\xi) &= \frac{\xi}{\zeta}\begin{pmatrix} \tan\alpha_2 \\ 1 \end{pmatrix}. \end{aligned} \tag{3.20}$$

Hence, ξ describes the distance of a point to the baseline of the prism under normal projection, and η is a generalized angular variable. The scaling factor ζ is used to construct a global distance variable, i.e., connecting neighbouring prisms: if a segment i and neighbour segment $i+1$ are glued together, the relation

$$\zeta_{i+1} = \zeta_i \frac{\cos\alpha_{1,i}}{\cos\alpha_{2,i+1}}, \tag{3.21}$$

has to hold.

Figure 3.2 shows $\xi = const$ coordinate lines for demonstration. As explained before, using the PML method, ξ is replaced by $(1+\sigma)\xi$, and Maxwell's equations are formulated in the new coordinate system. The Jacobian of (3.20), which is needed for transformation of the permeability and permittivity tensors, is given by:

$$J = \begin{pmatrix} \frac{h\zeta + a(1+\sigma)\xi}{h\zeta} & (1+\sigma)\frac{-a_1 h + a\eta}{h\zeta} & 0 \\ 0 & \frac{1+\sigma}{\zeta} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3.22}$$

where $a_1 = \tan\alpha_1$ and $a = \tan\alpha_1 + \tan\alpha_2$.

3.3.2 Pole condition

Another elegant principle, which can be used to state transparent boundary conditions, is the pole condition [85]. In order to characterize outgoing waves, the following definition is utilized:

Definition 27 (Pole condition in 1D)**.** A function $u : \mathbb{R}_+ \to \mathbb{C}$ satisfies the pole condition if its Laplace transform $\hat{u}$ has a holomorphic extension to the lower half complex plane.

Outgoing waves are defined as functions satisfying the Pole condition. The generalization to higher spatial dimensions is done by introduction of a coordinate system in the exterior, which includes a generalized distance variable, e.g., the prismatoidal coordinate system given in the previous section. Then the Laplace transform of a function with respect to this distance variable is utilized in the definition of the pole condition. A discretization scheme for the pole condition leads to discretization of the Laplace transform $\hat{u}$, where its analyticity in the lower half complex plane is respected in the ansatz space. The pole condition has spectral convergence properties in the number of degrees of freedom for this representation [31, 32, 52, 86].

3.3.3 Dirichlet to Neumann operator

Transparent boundary conditions can also be formulated, using a so-called Dirichlet to Neumann (DtN) operator. For its introduction we test the curl-curl equations (3.1) with a function v, but only integrate over the interior domain Ω_{int}:

$$\int\limits_{\Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du - \omega^2 \bar{v} \wedge *_\varepsilon u = \int\limits_{\partial\Omega_{\text{int}}} \bar{v} \wedge *_\mu^{-1} du.$$

We already integrated the curl-curl term by parts. The 1-form $*_\mu^{-1} du$ is continuous at interfaces, such that we can express u by the scattered field u_{sc} and the incoming field u_{in}, which have support in the exterior of Ω_{int}:

$$\int\limits_{\Omega_{\text{int}}} d\bar{v} \wedge *_\mu^{-1} du - \omega^2 \bar{v} \wedge *_\varepsilon u - \int\limits_{\partial\Omega_{\text{int}}} \bar{v} \wedge *_\mu^{-1} du_{\text{sc}} = \int\limits_{\partial\Omega_{\text{int}}} \bar{v} \wedge *_\mu^{-1} du_{\text{in}}.$$

In contrast to the incoming field du_{in}, Neumann data for the scattered field du_{sc} is not known. Therefore, a DtN operator is introduced, which maps Dirichlet data on the boundary $\partial\Omega_{\text{int}}$ onto Neumann data:

$$\text{DtN}(u) = *_\mu^{-1} du|_{\partial\Omega_{\text{int}}}.$$

Substituting $u_{\text{sc}} = u - u_{\text{in}}$ gives a closed formulation for the unknown interior field u:

$$\int\limits_{\Omega_{\text{int}}} d\bar{v} \wedge *_{\mu}^{-1} du - \omega^2 \bar{v} \wedge *_{\varepsilon}\, u - \int\limits_{\partial\Omega_{\text{int}}} \bar{v} \wedge \text{DtN}(u) =$$
$$\int\limits_{\partial\Omega_{\text{int}}} \bar{v} \wedge \text{DtN}(u_{\text{in}}) + \int\limits_{\partial\Omega_{\text{int}}} \bar{v} \wedge *_{\mu}^{-1} du_{\text{in}}.$$

In order to construct the DtN operator explicitly, the boundary value problem in the exterior for arbitrary Dirichlet data on $\partial\Omega_{\text{int}}$ has to be solved. This is only possible for simple cases, e.g., homogeneous exterior domains. Discretization of the DtN operator results in a non-local dense operator on the degrees of freedom of u on $\partial\Omega_{\text{int}}$, whereas for the PML and pole condition approach the system matrix of the exterior degrees of freedom remains sparse.

3.4 Finite element discretization

The weak formulation of a PDE offers an elegant discretization scheme. The key idea is the restriction of the continuous variational formulation (3.19) to a finite dimensional space V_h. Here we choose:

$$V_h \subset \text{H}\left(\mathbf{curl}, \Omega\right), \quad \dim V_h = \mathcal{N} < \infty, \tag{3.23}$$

i.e., the finite element space is a subspace of the original space $\text{H}\left(\mathbf{curl}, \Omega\right)$. In this case V_h is referred to as conforming finite element space [51, 10]. The discrete scattering problem corresponding to Problem 4 then reads:

Problem 5. Find $u \in V_h$ such that

$$a(u, v) = f(v), \quad \forall v \in V_h. \tag{3.24}$$

This will be our so-called **truth approximation**.

In order to obtain a numerical scheme for solution of the truth approximation, we construct a basis $B = \{\varphi_1, \ldots, \varphi_{\mathcal{N}}\}$ of V_h. Then we can expand the solution u according to:

$$u = \sum_{i=1}^{\mathcal{N}} \alpha_i \varphi_i.$$

Since each element in V_h can be expanded in this basis, it is sufficient to require:

$$a(u,v) = f(v)\,, \quad \forall v \in B,$$

instead of (3.24). This gives:

$$a\left(\sum_{i=1}^{\mathcal{N}} \alpha_i \varphi_i, \varphi_j\right) = \sum_{i=1}^{\mathcal{N}} a(\varphi_i, \varphi_j)\alpha_i = f(\varphi_j), \quad j = 1, \ldots, \mathcal{N}\,,$$

which gives a linear system of equations for the unknown coefficients α_i.

An important question is how to construct the space V_h. In order to avoid spurious modes, this has to be done with care for Maxwell's equations, as will be explained in the following [51]. We introduce the sequence:

$$H^1(\Omega)/\mathbb{R} \xrightarrow{\mathbf{grad}} H(\mathbf{curl}\,, \Omega) \xrightarrow{\mathbf{curl}} H(\mathbf{div}\,, \Omega) \xrightarrow{\mathbf{div}} L^2(\Omega), \tag{3.25}$$

which is called de Rham complex [20, 36]. Here elements of the quotient space $H^1(\Omega)/\mathbb{R}$ are in the same equivalence class if they only differ by a constant function. On simply connected subsets $\Omega \subset \mathbb{R}^3$ this sequence is exact, i.e.:

- the **grad** -operator has a trivial kernel on $H^1(\Omega)/\mathbb{R}$,
- the range of **grad** on $H^1(\Omega)/\mathbb{R}$ is a subset of $H(\mathbf{curl}\,, \Omega)$, and it is exactly the kernel of the **curl** -operator,
- the range of the **curl** -operator on $H(\mathbf{curl}\,, \Omega)$ is a subset of $H(\mathbf{div}\,, \Omega)$, and it is exactly the kernel of the **div** -operator,
- the range of **div** on $H(\mathbf{div}\,, \Omega)$ is $L^2(\Omega)$.

For the corresponding conforming finite element spaces:

$$\begin{aligned} U_h &\subset H^1(\Omega), \\ V_h &\subset H(\mathbf{curl}\,, \Omega), \\ W_h &\subset H(\mathbf{div}\,, \Omega), \\ Z_h &\subset L^2(\Omega), \end{aligned} \tag{3.26}$$

we also want that an exact sequence holds:

$$U_h/\mathbb{R} \xrightarrow{\mathbf{grad}} V_h \xrightarrow{\mathbf{curl}} W_h \xrightarrow{\mathbf{div}} Z_h. \tag{3.27}$$

If this is fulfilled, each function $v \in V_h$, which lies in the kernel of the **curl**-operator, is the gradient of a scalar function, i.e.:

$$\mathbf{curl}\, v = 0 \implies \exists\, \phi \in U_h : v = \mathbf{grad}\, \phi.$$

Then there are no unphysical spurious modes in V_h, which lie in the kernel of the **curl**-operator, but can not be expressed as the gradient of a scalar potential. In our work we use Nedelec elements for discretization [53, 51], which fulfill the discrete exact sequence.

For a detailed overview of the finite element method for electromagnetic field computations and appropriate construction of finite element spaces we refer to [29, 106, 88].

4 A posteriori error estimation

The reduced basis method computes approximate solutions to the truth approximation of a PDE. For efficient construction of a reduced basis approximation, and in order to quantify the reliability of a reduced basis solution, estimation of errors is very important. In this chapter we, therefore, derive a posteriori error bounds for approximate solutions to a PDE. Our setting is the following. We consider an exact problem with solution u and an approximated problem with solution $\tilde{u}$. Our goal is to derive bounds for the difference between u and $\tilde{u}$, and outputs of interest, which are computed from these solutions respectively.

The following results carry over directly to the reduced basis setting. Later u will be identified with the truth approximation and $\tilde{u}$ with the reduced basis solution.

4.1 Problem setup

The exact problem we consider is given as follows:

Problem 6. Compute the output of interest

$$s = l^{o}\left(u\right), \tag{4.1}$$

where u is the solution to the following problem:
Find $u \in X$ such that:

$$a(u,v) = f(v)\,, \quad \forall v \in X, \tag{4.2}$$

which is the standard Galerkin scheme from a finite element discretization. Trial and test space are equal. Equation (4.1) defines the so-called output of interest, which is given by a functional, acting on the solution of (4.2). Now suppose, we do not solve (4.2), but compute an approximate solution on the reduced subspace:

$$\tilde{X} \subset X. \tag{4.3}$$

The variational problem on the reduced space reads:

Problem 7. Compute the output of interest

$$\tilde{s} = l^{\mathrm{o}}\left(\tilde{u}\right), \tag{4.4}$$

where $\tilde{u}$ is the solution to the following problem:
Find $\tilde{u} \in \tilde{X}$ such that:

$$a(\tilde{u}, \tilde{v}) = f(\tilde{v}), \quad \forall \tilde{v} \in \tilde{X}. \tag{4.5}$$

The goal of a posteriori error estimation is to quantify the error of the solution:

$$||u - \tilde{u}||_X , \tag{4.6}$$

and of the output of interest:

$$|s - \tilde{s}| = |l^{\mathrm{o}}\left(u\right) - l^{\mathrm{o}}\left(\tilde{u}\right)| ,$$

without computing u itself.

The weak form of Maxwell's equations leads to non-coercive sesquilinear forms a. We will, therefore, focus on this case in the following [47, 16]. A posteriori error analysis for reduced basis approximations in the elliptic case can be found in [59, 82].

4.2 Residuum

The residuum $r^{\mathrm{pr}}(\cdot;\tilde{u})$ of an approximate solution $\tilde{u}$ plays a key role in a posteriori error estimation. It is defined by:

$$\begin{aligned} r^{\mathrm{pr}}(\cdot;\tilde{u}) &: X \to \mathbb{C} \\ & v \mapsto r^{\mathrm{pr}}(v;\tilde{u}) := f(v) - a(\tilde{u}, v), \end{aligned} \tag{4.7}$$

hence, $r^{\mathrm{pr}}(\cdot;\tilde{u}) \in X'$. The superscript "pr" thereby denotes the primal residuum. Later we will introduce a dual problem with corresponding dual residuum r^{du}.

The error e of the approximate solution:

$$e = u - \tilde{u}, \tag{4.8}$$

is connected to the residuum via the following error residuum relationship:

$$a(e, v) = r^{\mathrm{pr}}(v;\tilde{u}), \quad \forall v \in X. \tag{4.9}$$

Hence, the residuum can be interpreted as a source, from which the error could be computed directly. However, this has costs equal to solution of the original exact problem (4.2).

The Galerkin scheme 'orthogonalizes' the components of the residuum against the test space:

$$a(e, \tilde{v}) = r^{\mathrm{pr}}(\tilde{v}; \tilde{u}) = 0\,, \quad \forall \tilde{v} \in \tilde{X}, \tag{4.10}$$

which is referred to as Galerkin orthogonality.

If the bilinear form a is coercive, it induces a scalar product. Then the solution of the Galerkin method has the following property:

$$\tilde{u} = \arg\min_{\tilde{v} \in \tilde{X}} ||\tilde{v} - u||_a\,,$$

i.e., the Galerkin solution is the element of the trial space which has the minimum error, measured in the so-called energy norm $||\cdot||_a$, induced by a [10]. For non-coercive bilinear forms, Galerkin orthogonality (4.10) still holds, however, there is no underlying minimization principle for the error.

For the construction of error bounds, the dual norm of the residuum has to be determined. Therefore, first we define the Riesz representation $\hat{e}^{\mathrm{pr}}$ of the residuum by:

$$(v, \hat{e}^{\mathrm{pr}})_X = r^{\mathrm{pr}}(v; \tilde{u}), \quad \forall v \in X. \tag{4.11}$$

According to the Riesz representation theorem, the dual norm of the residuum is equal to the norm of $\hat{e}^{\mathrm{pr}}$:

$$||r^{\mathrm{pr}}(\cdot; \tilde{u})||_{X'} = ||\hat{e}^{\mathrm{pr}}||_X\,. \tag{4.12}$$

4.3 Inf-sup constant

A second ingredient to a posteriori error bounds is the inf-sup constant. For a given sesquilinear form the definition of the inf-sup constant (2.4), however, offers no obvious instruction, how to compute it. Therefore, in the following we reformulate the definition, such that the square of the inf-sup constant β^2 can be interpreted as the solution to an eigenvalue problem.

Let a be a bounded non-coercive sesquilinear form. We recall the definition of the inf-sup constant:

$$\beta = \inf_{v \in X} \sup_{w \in X} \frac{|a(v, w)|}{||v||_X\, ||w||_X}.$$

Since $a(v,\cdot)$ is a linear form on X, the definition can also be formulated as:

$$\beta = \inf_{v\in X} \frac{||a(v,\cdot)||_{X'}}{||v||_X}. \tag{4.13}$$

According to Corollary 1, we can find a representation operator T of the functional $a(v,\cdot) \in X'$ such that:

$$a(v,\cdot) = (\cdot, Tv)_X \,, \quad \forall v \in X, \tag{4.14}$$

with

$$\begin{aligned} T : X &\to X \\ v &\mapsto Tv. \end{aligned} \tag{4.15}$$

The inf-sup constant can then be given as follows:

$$\begin{aligned} \beta &= \inf_{v\in X} \frac{||a(v,\cdot)||_{X'}}{||v||_X} \\ &= \inf_{v\in X} \frac{||(\cdot, Tv)_X||_{X'}}{||v||_X} \\ &= \inf_{v\in X} \frac{||Tv||_X}{||v||_X} && (4.16) \\ &= \inf_{v\in X} \frac{(Tv, Tv)_X}{||v||_X\,||Tv||_X} \\ &= \inf_{v\in X} \frac{a(v, Tv)}{||v||_X\,||Tv||_X}, && (4.17) \end{aligned}$$

where we used $||(\cdot, Tv)_X||_{X'} = ||Tv||_X$. Squaring Eq. (4.16) gives:

$$\beta^2 = \inf_{v\in X} \frac{(Tv, Tv)_X}{(v, v)_X}. \tag{4.18}$$

Now we interpret the right hand side of (4.18) as the Rayleigh quotient of the following symmetric, positive-definite generalized eigenvalue problem:
Find pairs (χ, λ) with $\chi \in X$ and $\lambda \in \mathbb{R}$ such that:

$$(T\chi, Tv)_X = \lambda\,(v, \chi)_X\,, \quad \forall v \in X. \tag{4.19}$$

According to (4.18) the square of the inf-sup constant β^2 corresponds to the minimum eigenvalue of (4.19):

$$\beta^2 = \lambda_{\min}. \tag{4.20}$$

4.4 Primal error bounds

Now we can give an estimate for the error of the approximate solution $\tilde{u}$ in the X-norm:

Theorem 4. The error e (4.8) of the approximate solution $\tilde{u}$ is bounded by:

$$||e||_X \leq \frac{1}{\beta} \, ||r^{\mathrm{pr}}(\cdot;\tilde{u})||_{X'} \, . \tag{4.21}$$

Proof. From (4.17) we find with $v = u - \tilde{u}$:

$$\begin{aligned} \beta \, ||u - \tilde{u}||_X \, ||T\,(u - \tilde{u})||_X &\leq a(u - \tilde{u}, T\,(u - \tilde{u})) \\ &= r^{\mathrm{pr}}(T\,(u - \tilde{u})\,;\tilde{u}) \\ &\leq ||r^{\mathrm{pr}}(\cdot;\tilde{u})||_{X'} \, ||T\,(u - \tilde{u})||_X \, , \end{aligned}$$

where we used the error residuum relationship (4.9). Consequently we have:

$$||u - \tilde{u}||_X \leq \frac{1}{\beta} \, ||r^{\mathrm{pr}}(\cdot;\tilde{u})||_{X'} \, . \tag{4.22}$$

□

According to bound (4.21) we define the estimator for the error of the approximate solution in X-norm by:

$$\Delta^X = \frac{1}{\beta} \, ||r^{\mathrm{pr}}(\cdot;\tilde{u})||_{X'} \, . \tag{4.23}$$

Now we can quantify the error of the output of interest (4.4).

Lemma 3. The error of the output of interest (4.4) is bounded by:

$$|s - \tilde{s}| \leq \frac{1}{\beta} \, ||l^{\mathrm{o}}\,(\cdot)||_{X'} \, ||r^{\mathrm{pr}}(\cdot;\tilde{u})||_{X'} \, . \tag{4.24}$$

Proof. Since the output of interest is given by a bounded linear function, we find:

$$\begin{aligned} |s - \tilde{s}| &= |l^{\mathrm{o}}\,(u) - l^{\mathrm{o}}\,(\tilde{u})\,| \\ &= |l^{\mathrm{o}}\,(u - \tilde{u})\,| \\ &\leq ||l^{\mathrm{o}}_{.}\,(\cdot)||_{X'} \, ||e||_X \\ &\leq \frac{1}{\beta} \, ||l^{\mathrm{o}}\,(\cdot)||_{X'} \, ||r^{\mathrm{pr}}(\cdot;\tilde{u})||_{X'} \, , \end{aligned}$$

where we used estimate (4.21) for $||e||_X$. □

The error estimator for the output of interest is according to (4.24) given by:

$$\Delta^{\mathrm{o}} = \frac{1}{\beta} \left|\left|l^{\mathrm{o}}\left(\cdot\right)\right|\right|_{X'} \left|\left|r^{\mathrm{pr}}(\cdot;\tilde{u})\right|\right|_{X'}. \tag{4.25}$$

Hence, the error bound for the output of interest converges with the same rate as the error bound for the X-norm of the solution.

4.5 Dual problem

It is possible to increase the convergence rate of the error bound for the output of interest by introducing a dual problem [62]:

Problem 8. Find $w \in X^{\mathrm{du}}$ such that:

$$a(v,w) = -l^{\mathrm{o}}\left(v\right), \quad \forall v \in X^{\mathrm{du}}. \tag{4.26}$$

As for the primal problem, we define an approximate dual problem:

Problem 9. Find $\tilde{w} \in \tilde{X}^{\mathrm{du}}$ such that:

$$a(\tilde{v},\tilde{w}) = -l^{\mathrm{o}}\left(\tilde{v}\right), \quad \forall \tilde{v} \in \tilde{X}^{\mathrm{du}}, \tag{4.27}$$

with $\tilde{X}^{\mathrm{du}} \subset X^{\mathrm{du}}$.

If the bilinear (sesquilinear) form a is (complex-) symmetric:

$$a(v,w) = a(w,v), \quad \forall v, w \in X, \tag{4.28}$$

and the system right hand side is equal to the output functional:

$$f(\cdot) = l^{\mathrm{o}}\left(\cdot\right), \tag{4.29}$$

the primal and dual problem are equivalent. This is called the compliant case, and the following lemma holds.

Lemma 4. For a compliant input-output relationship the error in the output of interest is bounded by:

$$|s - \tilde{s}| \leq \frac{1}{\beta} \left|\left|r^{\mathrm{pr}}(\cdot;\tilde{u})\right|\right|^2_{X'}. \tag{4.30}$$

Proof. We use linearity of the output of interest, the compliance properties (4.28), (4.29), Galerkin orthogonality (4.10), and the error residuum relationship (4.9):

$$\begin{aligned}|s-\tilde{s}| =&|f(u-\tilde{u})|\\ =&|a(u,u-\tilde{u})|\\ =&|a(u-\tilde{u},u)|\\ =&|a(u-\tilde{u},u)-a(u-\tilde{u},\tilde{u})|\\ =&|a(u-\tilde{u},u-\tilde{u})|\\ =&|r^{\mathrm{pr}}(u-\tilde{u};\tilde{u})|\\ \leq& \,||r^{\mathrm{pr}}(\cdot;\tilde{u})||_{X'}\,||e||_X\,.\end{aligned}$$

Applying bound (4.21) for $||e||_X$ concludes the proof. □

Hence, in the compliant case the convergence rate of the error bound of the output of interest is doubled. The compliant case is easy to treat in the reduced basis context, since no reduced basis solution for the dual problem has to be computed.

However, for the Maxwell scattering problem neither a is complex symmetric, nor are the primal right hand side and outputs of interest equal. In the following we, therefore, derive error bounds for the non-compliant case. According to the primal error (4.8), we define the dual error:

$$e^{\mathrm{du}} := w-\tilde{w}, \tag{4.31}$$

and dual residuum of a solution to the dual Problem (4.27):

$$r^{\mathrm{du}}(\cdot;\tilde{w}) = -l^{\mathrm{o}}\,(\cdot) - a(\cdot,\tilde{w}). \tag{4.32}$$

The dual error residuum relationship is given by:

$$a(v,e^{\mathrm{du}}) = r^{\mathrm{du}}(v;\tilde{w})\,, \quad \forall v \in X^{\mathrm{du}}. \tag{4.33}$$

Galerkin orthogonality for the dual solution reads:

$$r^{\mathrm{du}}(\tilde{v};\tilde{w}) = 0\,, \quad \forall \tilde{v} \in \tilde{X}^{\mathrm{du}}. \tag{4.34}$$

The approximate solution of the dual problem can be used to correct the output of interest of the primal problem [62]. We define:

$$\tilde{s}^{\mathrm{pd}} = \tilde{s} - r^{\mathrm{pr}}(\tilde{w};\tilde{u}), \tag{4.35}$$

hence, the output is corrected by the primal residuum, evaluated at the dual solution.

The following lemma motivates this definition:

Lemma 5. The error of the dual corrected output of interest is bounded by:

$$|s - \tilde{s}^{\mathrm{pd}}| \leq \frac{1}{\beta} \left|\left|r^{\mathrm{pr}}(\cdot; \tilde{u})\right|\right|_{X'} \left|\left|r^{\mathrm{du}}(\cdot; \tilde{w})\right|\right|_{X'}. \tag{4.36}$$

Proof. We find:

$$\begin{aligned} |s - \tilde{s}^{\mathrm{pd}}| &= |s - \tilde{s} + r^{\mathrm{pr}}(\tilde{w}; \tilde{u})| \\ &= |l^{\mathrm{o}}(u) - l^{\mathrm{o}}(\tilde{u}) + f(\tilde{w}) - a(\tilde{u}, \tilde{w})| \\ &= | - a(u, w) + a(\tilde{u}, w) + a(u, \tilde{w}) - a(\tilde{u}, \tilde{w})| \\ &= | - a(u - \tilde{u}, w) + a(u - \tilde{u}, \tilde{w})| \\ &= |a(u - \tilde{u}, \tilde{w} - w)|. \end{aligned}$$

Furthermore, we have:

$$\begin{aligned} |a(u - \tilde{u}, \tilde{w} - w)| &= |r^{\mathrm{du}}(u - \tilde{u}; \tilde{w})| \\ &\leq \left|\left|r^{\mathrm{du}}(\cdot; \tilde{w})\right|\right|_{X'} ||u - \tilde{u}||_X \\ &= \left|\left|r^{\mathrm{du}}(\cdot; \tilde{w})\right|\right|_{X'} ||e||_X, \end{aligned}$$

which together with bound (4.21) for $||e||_X$ concludes the proof. □

Corresponding to (4.36) we define the error estimator for the dual corrected output of interest:

$$\Delta^o_{\mathrm{pd}} = \frac{1}{\beta} \left|\left|r^{\mathrm{pr}}(\cdot; \tilde{u})\right|\right|_{X'} \left|\left|r^{\mathrm{du}}(\cdot; \tilde{w})\right|\right|_{X'}. \tag{4.37}$$

Comparing this primal-dual error bound with the primal error bound (4.25), we notice that the constant norm of the output of interest in (4.25) is replaced by the norm of the dual residuum. Hence, an accurate approximation to the dual problem, can increase the convergence rate for the output of interest.

4.6 Effectivities

In order to quantify the performance of error estimators, we define the effectivity of an estimator by:

$$\text{effectivity} = \frac{\text{estimated error}}{\text{true error}}.$$

For the error estimators introduced in this chapter we define:

$$\eta^X = \frac{\Delta^X}{||u - \tilde{u}||_X}, \tag{4.38}$$

$$\eta^o = \frac{\Delta^o}{|s - \tilde{s}|}, \tag{4.39}$$

$$\eta^o_{\mathrm{pd}} = \frac{\Delta^o_{\mathrm{pd}}}{|s - \tilde{s}^{\mathrm{pd}}|}, \tag{4.40}$$

which quantify the performance of the error estimator for the approximate solution in X-norm and the output of interest using the primal and primal-dual approximations. According to their construction, the estimators are rigorous, i.e.:

$$\eta^X,\, \eta^o,\, \eta^o_{\mathrm{pd}} \geq 1.$$

This means the error is never underestimated. For sharpness we desire that the estimators are as close to unity as possible.

5 Reduced basis method

In this chapter we develop the reduced basis method for fast and reliable solution of parametrized elliptic non-coercive PDEs. The reduced basis method allows to split up the solution process of a parametrized problem into an expensive offline and a cheap online phase [82]. In the offline phase a reduced model to the truth approximation is constructed self-adaptively. In the online phase the reduced system can then be solved orders of magnitude faster than the original problem.

In the following we start with the formulation of the problem setup and develop established state-of-the-art reduced basis techniques, regarding construction of the reduced basis, online-offline decomposition, and a posteriori error estimation [82].

Then we develop the affine decomposition for geometrically parametrized 3D Maxwell scattering problems, which is the key for efficient online-offline decomposition in the reduced basis context.

Examining online and offline computational costs, we will observe that there is need for alternative error estimation techniques, when dealing with affine decompositions consisting of a large number of terms. We introduce a novel residuum estimator, inspired by the sub-domain residuum method of a posteriori error estimation of finite element solutions [1], which will lead to substantial savings, regarding computational time and memory requirements.

Finally, we develop a novel reduced basis technique for dealing with systems subject to a large number of different sources, which is a typical situation in many nano-optical applications.

5.1 Problem setup

In our setting the connection between input and output of a parametrized system is stated via a PDE. The input enters as parameters to the PDE, and usually the result is a physical field, like a temperature distribution or electric field strength. Often the output of interest is not the field solution itself, but some quantity which is derived from the field. Here, we consider outputs, which are given by linear operators acting on the field solution.

In our work we consider geometrical parameters as inputs to the PDE.

The input parameters are then p real numbers, e.g., the width, height, or length of objects, sidewall angles, radii etc. In applications the range of these parameters is usually bounded. We assume that the input parameters ν are restricted to a bounded parameter box $D \subset \mathbb{R}^p$.

Variable material parameters like permittivities of objects in an optical system, would offer no additional difficulty. In fact, assembling of a parametrized system with variable material parameters is simpler than in the case of geometrical parameters.

Mathematically the continuous input-output relationship is stated as follows:

Problem 10. For given tuple of input parameters $\nu \in D \subset \mathbb{R}^p$, compute the output of interest:

$$s(\nu) = l^{\mathrm{o}}\left(u(\nu)\right), \tag{5.1}$$

where $u(\nu) \in X$ is the solution to the following problem:
Find $u(\nu) \in X$ such that:

$$a\left(u(\nu), v; \nu\right) = f\left(v\right), \quad \forall v \in X. \tag{5.2}$$

The PDE is stated in weak form, and the input parameters enter as parameters to the bilinear form $a\left(\cdot, \cdot; \nu\right)$. The output of interest is given by a linear functional $l^{\mathrm{o}}\left(\cdot\right) \in \left(X\right)'$, evaluated at the parameter dependent solution $u(\nu)$.

Usually the continuous mathematical model (5.2) can not be solved exactly. Hence, a numerical method has to be applied. In the reduced basis context the finite element method is the method of choice. The discretized input-output relationship is then given by:

Problem 11. For given tuple of input parameters $\nu \in D \subset \mathbb{R}^p$, compute the output of interest:

$$s^{\mathcal{N}}(\nu) = l^{\mathrm{o}}\left(u^{\mathcal{N}}(\nu)\right), \tag{5.3}$$

where $u^{\mathcal{N}}(\nu) \in X^{\mathcal{N}}$ is the solution to the following problem:
Find $u^{\mathcal{N}}(\nu) \in X^{\mathcal{N}}$ such that:

$$a\left(u^{\mathcal{N}}(\nu), v; \nu\right) = f\left(v\right), \quad \forall v \in X^{\mathcal{N}}. \tag{5.4}$$

The dimension of the finite element space is $\dim X^{\mathcal{N}} = \mathcal{N}$, and $l^{\mathrm{o}}\left(\cdot\right) \in \left(X^{\mathcal{N}}\right)'$ is the linear output functional.

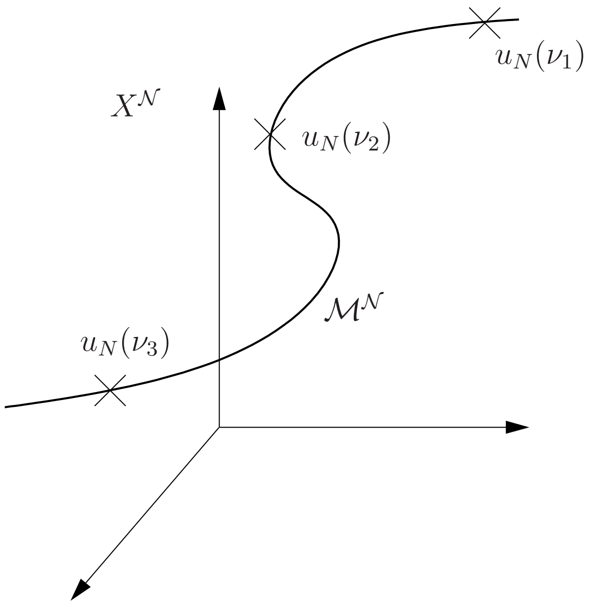


Figure 5.1: Sub-manifold of possible solutions $\mathcal{M}^{\mathcal{N}}$ in truth approximation space $X^{\mathcal{N}}$ for a single input parameter. Three snapshot solutions $u_N(\nu_i)$ are depicted.

Problem 11 is the so-called truth approximation. The variational formulation (5.4) on a finite element space $X^{\mathcal{N}}$ leads to a large sparse matrix equation of dimension $\mathcal{N}$, which has to be solved numerically. In nano-optical applications $\mathcal{N}$ is typically between 10^5 and several millions.

In applications a single evaluation of the truth approximation can take several minutes or up to hours, which often rules out real-time and many-query applications.

5.2 Reduced basis approximation

The main purpose of the reduced basis (RB) method is the construction and rapid evaluation of approximated input-output relationships for a given truth approximation.

The reduced basis method can be motivated as follows [82]. First, we define the manifold of all possible solutions to the truth approximation:

$$\mathcal{M}^{\mathcal{N}} = \left\{ u^{\mathcal{N}}(\nu) \text{ is a solution to (5.4)} \mid \nu \in D \right\}. \tag{5.5}$$

This manifold is a subset of the finite element space $\mathcal{M}^{\mathcal{N}} \subset X^{\mathcal{N}}$, as depicted in Fig. 5.1. Now suppose the manifold $\mathcal{M}^{\mathcal{N}}$ can be approximated "with good quality" by a low dimensional space X_N:

$$\mathcal{M}^{\mathcal{N}} \approx X_N \subset X^{\mathcal{N}}.$$

Then it is reasonable to assume that solutions to following reduced problem are good approximations to the truth approximation:

Problem 12. For given tuple of input parameters $\nu \in D \subset \mathbb{R}^p$, compute the output of interest:

$$s_N(\nu) = l^o \left(u_N(\nu) \right), \tag{5.6}$$

where $u_N(\nu) \in X_N$ is the solution to the following problem:
Find $u_N(\nu) \in X_N$ such that:

$$a \left(u_N(\nu), v; \nu \right) = f \left(v \right), \quad \forall v \in X_N. \tag{5.7}$$

The space X_N will be referred to as reduced basis space. Its dimension is given by $\dim X_N = N$.

If $N \ll \mathcal{N}$, the resulting reduced system (5.7) can be solved much faster than the truth approximations.

5.2.1 Reduced basis spaces

An important question is how to construct a space X_N with good approximation properties to $\mathcal{M}^{\mathcal{N}}$. In the reduced basis context the space X_N is built from so-called snapshots. These are solutions to the truth approximation (5.4) for a set of fixed parameter values $\nu^i \in D$, as depicted in Fig. 5.1. The corresponding reduced basis space is referred to as Lagrange space. In order to approximate $\mathcal{M}^{\mathcal{N}}$, it is also possible to include first or higher order derivatives of snapshot solutions with respect to the parameters into the reduced basis. This gives so-called Taylor and Hermite reduced basis spaces.

For construction of reduced basis spaces we first define a sequence of hierarchical subsets of the parameter domain [59]. Let $\nu_i \in D\,,\ i = 1, \dots, N_{\max}$. Then we define the following nested sets:

$$S_i = \{\nu_1, \dots, \nu_i\}\ , \quad i = 1, \dots, N_{\max}, \tag{5.8}$$

which have the property:

$$S_1 \subset S_2 \subset \cdots \subset S_N \subset \cdots \subset S_{N_{\max}}. \tag{5.9}$$

The Lagrange reduced basis space $W_N^{\mathcal{N}}$ of dimension N is then defined by:

$$W_N^{\mathcal{N}} = \operatorname{span}\left\{u^{\mathcal{N}}(\nu) \text{ is a solution to (5.4)} \mid \nu \in S_N\right\}. \tag{5.10}$$

For sufficiently large N this space should provide a good approximation to $\mathcal{M}^{\mathcal{N}}$. In numerical examples we will see that already for small reduced basis dimensions N, usually of the the order $O(10)$, the reduced system provides very accurate approximations to the truth approximation.

5.2.2 Online-offline decomposition

The definition of the reduced basis space (5.10) already shows a main part of the offline computational costs, when constructing a reduced basis approximation. Since a reduced basis space consists of snapshot solutions, the truth approximation has to be solved N times for its construction. Therefore, if one is only interested in a few evaluations of the input-output relationship for known parameter values, the reduced basis method offers no advantage. If, however, one is interested in real-time solutions, or has to perform many evaluations, it becomes favourable to pay the price of a time-consuming offline phase for the benefit of rapid online evaluations.

Let us have a look at the costs of solving the reduced variational formulation (5.7). Suppose we have given a basis

$$B_N^{\mathcal{N}} = \left\{\zeta_q^{\mathcal{N}} \mid q = 1, \ldots, N\right\} \tag{5.11}$$

of the reduced basis space $X_N = \operatorname{span} B_N^{\mathcal{N}}$. Then we can expand each reduced basis solution in this basis:

$$u_N(\nu) = \sum_{q=1}^{N} \alpha_q(\nu) \zeta_q^{\mathcal{N}}. \tag{5.12}$$

Galerkin projection of the reduced system onto the reduced basis $v \in B_N^{\mathcal{N}}$ gives:

$$\sum_{q=1}^{N} \alpha_q(\nu) a\left(\zeta_q^{\mathcal{N}}, \zeta_n^{\mathcal{N}}; \nu\right) = f\left(\zeta_n^{\mathcal{N}}\right), \quad n = 1, \ldots, N, \tag{5.13}$$

which is a N dimensional system of equations for the unknown parameter dependent coefficients $\alpha_q(\nu)$. Hence, costs for solution of this equation only depends on the dimension of the reduced basis space, which is usually very small.

The right hand side is a parameter independent vector of dimension N:

$$f^N = \left(f\left(\zeta_n^{\mathcal{N}}\right)\right)_{n=1,\ldots,N}, \tag{5.14}$$

which can be assembled offline. However, assembling of the parameter dependent matrix in (5.13):

$$A_N(\nu) = \left(a\left(\zeta_q^{\mathcal{N}}, \zeta_n^{\mathcal{N}}; \nu\right)\right)_{q,n=1,\ldots,N} \tag{5.15}$$

depends on the number of degrees of freedoms of the truth approximation, which we want to avoid for efficient online-offline decomposition. The key is the affine decomposition of the parameter dependent system bilinear form $a\left(\cdot, \cdot; \nu\right)$ as follows:

$$a\left(v, u; \nu\right) = \sum_{m=1}^{Q} \Theta_m(\nu) a_m(v, u), \tag{5.16}$$

where we have parameter dependent functions $\Theta_m(\nu)$ and parameter independent bilinear forms $a_m(\cdot, \cdot)$. If we can construct such a decomposition of the system bilinear form, we can assemble the following parameter independent matrices in the offline phase:

$$A_N^m = \left(a_m(\zeta_q^{\mathcal{N}}, \zeta_n^{\mathcal{N}})\right)_{q,n=1,\ldots,N}, \quad m = 1, \ldots, Q. \tag{5.17}$$

The parameter dependent system matrix (5.15) is then assembled in the online step according to:

$$A_N(\nu) = \sum_{m=1}^{Q} \Theta_m(\nu) A_N^m.$$

The assembling costs now only depend on the RB dimension N and the number of terms Q in the affine decomposition. Since the right hand side $\left(f(\zeta_q^N)\right)_{q=1,\ldots,N}$ of the reduced system (5.13) is also computed offline, the assembling costs of (5.13) become independent on $\mathcal{N}$.

Also for computation of output of interest (5.6) an online-offline decomposition is possible. Since the output functional is linear, the output can be computed by:

$$s_N(\nu) = \sum_{q=1}^{N} \alpha_q(\nu) l^{\mathrm{o}}\left(\zeta_q^{\mathcal{N}}\right). \tag{5.18}$$

The $\mathcal{N}$ dependent quantities $l^o\left(\zeta_q^{\mathcal{N}}\right)$ are computed offline. Hence, evaluation of the reduced input-output relationship becomes independent on the degrees of freedom of the truth approximation. Therefore, the accuracy of the truth approximation can be chosen conservatively by discretization with high $\mathcal{N}$, without increasing the online costs. Of course offline computational costs depend on $\mathcal{N}$.

5.2.3 Operation count

The online operation count is $O(N^2Q)$ for the assembling phase, $O(N^3)$ for solution of the linear system, and $O(N)$ for evaluation of output of interest.

Offline N FEM solutions of dimension $\mathcal{N}$ have to be computed. If the reduced basis space provides a good approximation to the truth approximation space $\mathcal{M}^{\mathcal{N}}$, it is obvious that each new snapshot, which is added to the reduced basis, has to be orthogonalized against the previous ones for stability reasons. This leads to offline costs of $O(N^2\mathcal{N})$ computing the scalar products, e.g., in the Gram-Schmidt orthogonalization procedure.

Furthermore, the Q constant matrices A_N^m (5.17) in the affine decomposition have to be computed. This is done by projection of the affine decomposition of the finite element matrix onto the reduced basis, which has offline costs of $O(QN^2\mathcal{N})$.

5.3 A posteriori error estimation

The result of a reduced basis computation is an approximative solution to the truth approximation. In order to quantify the reliability of a reduced basis solution, error estimation is very important. Furthermore, error estimators can be used for greedy construction of reduced basis spaces [82].

For efficient online-offline decomposition in the reduced basis context, we have the requirement that online evaluation costs of the error estimator should be independent on the number $\mathcal{N}$ of FEM degrees of freedom.

5.3.1 Error bounds

The estimation of errors of a reduced basis solution and output of interest directly resembles the setup in Section 4, where rigorous a posteriori error bounds were derived: the reduced basis space is a subspace of the finite element space: $X_N \subset X^{\mathcal{N}}$, c.f. Eq. (4.3). In the following we translate our results to the reduced basis setup.

First we define the error of the reduced basis solution:

$$e(\nu) = u^{\mathcal{N}}(\nu) - u_N(\nu) \tag{5.19}$$

and the error of the output of interest

$$e^{\mathrm{o}}(\nu) =|s^{\mathcal{N}}(\nu) - s_N(\nu)|. \tag{5.20}$$

For brevity we drop the ν dependence of u_N and $u^{\mathcal{N}}$ in the following. The primary residuum of the reduced basis solution is given by:

$$r^{\mathrm{pr}}\left(\cdot; u_N; \nu\right) = f(\cdot) - a\left(u_N, \cdot; \nu\right), \tag{5.21}$$

with $r^{\mathrm{pr}}\left(\cdot; u_N; \nu\right) \in \left(X^{\mathcal{N}}\right)'$. The parameter dependent inf-sup constant of $a\left(\cdot, \cdot; \nu\right)$ is defined by:

$$\beta(\nu) = \inf_{v \in X^{\mathcal{N}}} \sup_{w \in X^{\mathcal{N}}} \frac{\left|a\left(v, w; \nu\right)\right|}{||v||_{X^{\mathcal{N}}}\, ||w||_{X^{\mathcal{N}}}}. \tag{5.22}$$

According to Section 4.3, the inf-sup constant can be computed from a generalized eigenvalue problem. Therefore, first we define the parameter dependent representation operator T_ν of the functional $a\left(v, \cdot; \nu\right) \in \left(X^{\mathcal{N}}\right)'$:

$$a\left(v, \cdot; \nu\right) = \left(\cdot, T_\nu v\right)_{X^{\mathcal{N}}}, \quad \forall v \in X^{\mathcal{N}}, \tag{5.23}$$

with:

$$\begin{aligned} T_\nu : X^{\mathcal{N}} &\to X^{\mathcal{N}} \\ v &\mapsto T_\nu v. \end{aligned} \tag{5.24}$$

Now we state the following generalized eigenvalue problem:
Find pairs $(\chi(\nu), \lambda(\nu))$ with $\chi(\nu) \in X^{\mathcal{N}}$ and $\lambda(\nu) \in \mathbb{R}$ such that:

$$\left(T_\nu \chi(\nu), T_\nu v\right)_{X^{\mathcal{N}}} = \lambda(\nu)\left(v, \chi(\nu)\right)_{X^{\mathcal{N}}}, \quad \forall v \in X^{\mathcal{N}}. \tag{5.25}$$

We then have $\beta(\nu)^2 = \lambda_{\min}(\nu)$, where $\lambda_{\min}(\nu)$ is the minimum eigenvalue of (5.25).

According to the primal error estimate (4.23) derived in Section 4.4, we now define the estimator for the error of the reduced basis solution in $X^{\mathcal{N}}$-norm by:

$$\Delta^{X^{\mathcal{N}}}(\nu) = \frac{1}{\beta(\nu)} ||r^{\mathrm{pr}}\left(\cdot; u_N; \nu\right)||_{(X^{\mathcal{N}})'}, \tag{5.26}$$

and for the output of interest according to estimate (4.25):

$$\Delta^{\mathrm{o}}(\nu) = \frac{1}{\beta(\nu)} ||l^{\mathrm{o}}\left(\cdot\right)||_{(X^{\mathcal{N}})'}\, ||r^{\mathrm{pr}}\left(\cdot; u_N; \nu\right)||_{(X^{\mathcal{N}})'}. \tag{5.27}$$

According to Theorem 4 and Lemma 3 both estimators are rigorous:

$$||e(\nu)||_{X^{\mathcal{N}}} \leq \Delta^{X^{\mathcal{N}}}(\nu), \tag{5.28}$$

$$|s^{\mathcal{N}}(\nu) - s_N(\nu)| \leq \Delta^{\mathrm{o}}(\nu). \tag{5.29}$$

5.3.2 Online-offline decomposition

In the following we derive the online-offline decomposition for the reduced basis error estimators (5.26) and (5.27).

Dual norm of residuum

A key ingredient to the error estimators is the dual norm of the residuum of a reduced basis solution:

$$||r^{\mathrm{pr}}\left(\cdot;u_N;\nu\right)||_{(X^{\mathcal{N}})'}\,. \tag{5.30}$$

For numerical evaluation of this expression we utilize the Riesz representation of the residuum, which is defined by:

$$\left(v,\hat{e}^{\mathrm{pr}}\left(\nu\right)\right)_{X^{\mathcal{N}}}=r^{\mathrm{pr}}\left(v;u_N;\nu\right),\quad \forall v\in X^{\mathcal{N}}. \tag{5.31}$$

Above definition is an elliptic problem, from which $\hat{e}^{\mathrm{pr}}\left(\nu\right)$ can be computed. According to the Riesz representation theorem, the dual norm can then be evaluated by:

$$||r^{\mathrm{pr}}\left(\cdot;u_N;\nu\right)||_{(X^{\mathcal{N}})'}=||\hat{e}^{\mathrm{pr}}\left(\nu\right)||_{X^{\mathcal{N}}}\,. \tag{5.32}$$

However, solving elliptic problem (5.31) is as expensive as solving the truth approximation itself. Also computation of the $X^{\mathcal{N}}$-norm of $\hat{e}^{\mathrm{pr}}\left(\nu\right)$ depends on the number of FEM degrees of freedom, which we want to avoid. For an efficient online-offline decomposition again the affine decomposition of the bilinear form $a\left(\cdot,\cdot;\nu\right)$ (5.16) is utilized. According to the definition of the primal residuum (5.21), we find:

$$\begin{aligned} r^{\mathrm{pr}}\left(\cdot;u_N;\nu\right) &= f\left(\cdot\right)-\sum_{m=1}^{Q}\Theta_m(\nu)a_m(u_N,\cdot)\\ &= f\left(\cdot\right)-\sum_{q=1}^{N}\sum_{m=1}^{Q}\Theta_m(\nu)\alpha_q(\nu)a_m(\zeta_q^{\mathcal{N}},\cdot), \end{aligned} \tag{5.33}$$

where we inserted the expansion of reduced basis solution u_N into a basis of X_N in the second line. This formulation gives rise to the definition of following elliptic problems:

$$f\left(v\right)=\left(v,b\right)_{X^{\mathcal{N}}}\qquad ,\forall v\in X^{\mathcal{N}}, \tag{5.34a}$$

$$a_q(\zeta_m^{\mathcal{N}},v)=\left(v,L_q^m\right)_{X^{\mathcal{N}}}\qquad ,\forall v\in X^{\mathcal{N}}, \tag{5.34b}$$

which are independent on the parameters ν and can be solved offline. Using expansion (5.33) and (5.31), the Riesz representation of the residuum can then be expressed in terms of solutions of above problems:

$$\hat{e}^{\mathrm{pr}}(\nu) = b - \sum_{q=1}^{N}\sum_{m=1}^{Q} \Theta_m(\nu)\alpha_q(\nu) L_q^m. \tag{5.35}$$

The norm of $\hat{e}^{\mathrm{pr}}(\nu)$ is given by:

$$\begin{aligned} ||\hat{e}^{\mathrm{pr}}(\nu)||_{X^{\mathcal{N}}}^2 = (b,b)_{X^{\mathcal{N}}} - 2\Re\left\{\sum_{q=1}^{N}\sum_{m=1}^{Q}\Theta_m(\nu)\alpha_q(\nu)\left(b, L_q^m\right)_{X^{\mathcal{N}}}\right\} \\ + \sum_{q,q'=1}^{N}\sum_{m,m'=1}^{Q}\Theta_m(\nu)\Theta_{m'}(\nu)\alpha_q(\nu)\alpha_{q'}(\nu)\left(L_q^m, L_{q'}^{m'}\right)_{X^{\mathcal{N}}}. \end{aligned} \tag{5.36}$$

The scalar products $(b,b)_{X^{\mathcal{N}}}$, $\left(b, L_q^m\right)_{X^{\mathcal{N}}}$, and $\left(L_q^m, L_{q'}^{m'}\right)_{X^{\mathcal{N}}}$ can be computed offline, such that evaluation of the dual norm of the residuum becomes independent on $\mathcal{N}$.

Inf-sup constant

The second ingredient for computation of error bounds (5.26) and (5.27) is the parameter dependent inf-sup constant $\beta(\nu)$. According to (5.25), it can be obtained from a generalized eigenvalue problem of size $\mathcal{N}$. Hence, an online-offline decomposition for the computation of $\beta(\nu)$ is necessary for an efficient evaluation of error bounds in the reduced basis context. State-of-the-art methods determine a lower bound to the inf-sup constant $0 < \beta(\nu)^{\mathrm{LB}} < \beta(\nu)$, such that the estimators remain rigorous. The most recent development is the successive constraint method (SCM), which constructs a lower bound by solving a linear optimization problem [33, 16]. Especially when resonances occur in the system, the inf-sup constant β varies over several orders of magnitude and tends to zero. This leads to dramatically overestimated errors and, hence, ineffective error bounds.

The problem setup we are interested in are electromagnetic scattering problems on unbounded domains. For our examples the inf-sup constant hardly depends on the geometrical parameters. It is basically constant, as will be demonstrated numerically. With nearly constant inf-sup constant, it is not a great loss of effectivity, using a parameter independent lower bound β_0 for the error bounds:

$$\beta_0 = \min_{\nu\in D}\beta(\nu). \tag{5.37}$$

Realizing this approach we take the minimum over a finite subset $\tilde{D} \subset D$. In order to minimize computational costs, it is favourable to take the minimum over all snapshot parameters. Taking a snapshot we already have the necessary LU decomposition of the system matrix $A(\nu)$ at hand, which is used in the Lanczos procedure computing $\beta(\nu)$.

Of course with this approach, the computed error bounds are not rigorous anymore, and in case there are resonances in the system, above approach is not feasible. For scattering problems with transparent boundary conditions there are, however, only resonances with complex eigenvalues ω in the system [41]. The scattering problem (3.19) with real-valued ω is always invertible, and the magnitude of the inf-sup constant should depend on the distance of the given real incoming frequency ω to the closest complex resonance in the system. At least for our scattering examples with no "cavity-like" character these complex resonances and therewith the inf-sup constant does not depend significantly on the geometry, even for larger parameter variations.

5.3.3 Operation count

The online operation count for computation of the dual norm of the residuum is $O(N^2Q^2)$, according to decomposition (5.36).

Offline $N \cdot Q$ finite element problems of dimension $\mathcal{N}$ have to be solved, to obtain the Riesz representations (5.34a) and (5.34b). However, only a single expensive LU decomposition of the matrix defining the norm on $X^{\mathcal{N}}$ has to be performed and $N \cdot Q$ forward backward substitutions, which gives costs of $O(NQ\mathcal{N})$. The costs for computation of scalar products (5.36) are $O(N^2Q^2\mathcal{N})$.

5.4 Dual problem

In Section 4.5 we showed that usage of an approximation to the dual (or adjoint) Problem (4.26) in the non-compliant case can increase the convergence rate of the output of interest.

The truth approximation of the dual problem for the input-output relationship Problem 11 is defined by:

Problem 13. Find $w^{\mathcal{N}} \in X^{\mathcal{N}}_{\mathrm{du}}$ such that:

$$a\left(v, w^{\mathcal{N}}; \nu\right) = -l^{\mathrm{o}}\left(v\right), \quad \forall v \in X^{\mathcal{N}}_{\mathrm{du}}. \tag{5.38}$$

The solution $w^{\mathcal{N}}$ to the dual problem can be used to correct the output of interest according to (4.35):

$$s^{\mathcal{N}}_{\text{pd}}(\nu) = s^{\mathcal{N}}(\nu) - r^{\text{pr}}\left(w^{\mathcal{N}}; u^{\mathcal{N}}; \nu\right). \tag{5.39}$$

If the dual problem would be solved on the same space as the primal problem, i.e. $X^{\mathcal{N}}_{\text{du}} = X^{\mathcal{N}}$, the correction would be zero, since $r^{\text{pr}}\left(w^{\mathcal{N}}; u^{\mathcal{N}}; \nu\right) = 0$ due to Galerkin orthogonality.

In the reduced basis context the dual problem reads:

Problem 14. Find $w_N \in X^{\text{du}}_N$ such that:

$$a\left(v, w_N; \nu\right) = -l^{\text{o}}\left(v\right), \quad \forall v \in X^{\text{du}}_N. \tag{5.40}$$

Also here a separate dual reduced basis space has to be constructed, in order to correct the output of interest of the primal reduced basis problem:

$$s^{\text{pd}}_N\left(\nu\right) = s_N(\nu) - r^{\text{pr}}\left(w_N; u_N; \nu\right). \tag{5.41}$$

We define the error of the dual corrected output of interest by:

$$e^{\text{o}}_{\text{pd}}(\nu) = |s^{\mathcal{N}}(\nu) - s^{\text{pd}}_N\left(\nu\right)|. \tag{5.42}$$

Especially for many outputs of interest the increased convergent rate might not pay of, due to the necessity of solving a dual problem for each output of interest separately.

Online-offline decomposition

The online-offline decomposition strategy for the dual problem is given as follows. Since the system right hand side $f\left(\cdot\right)$ and the output of interest $l^{\text{o}}\left(\cdot\right)$ differ, one is forced to construct a separate reduced basis space for the primal and dual reduced basis problem. Let us denote the dual basis by:

$$B^{\mathcal{N}}_{N,\text{du}} = \left\{\zeta^{\mathcal{N}}_{q,\text{du}} |\, q = 1, \ldots, N\right\}. \tag{5.43}$$

Then the solution of the dual reduced basis approximation can be given by:

$$w_N(\nu) = \sum_{q'=1}^{N} \alpha_{q',\text{du}}(\nu)\zeta^{\mathcal{N}}_{q',\text{du}}. \tag{5.44}$$

The online-offline decomposition of the dual reduced basis system is similar to the primal system, given in Section 5.2.2. Again the affine decomposition of the system bilinear form is the key ingredient. For efficient computation of the dual correction (5.41) we also utilize the affine decomposition:

$$\begin{aligned} r^{\mathrm{pr}}\left(w_N;u_N;\nu\right) =& f\left(w_N\right) - a\left(u_N,w_N;\nu\right) \\ =& \sum_{q'=1}^{N} \alpha_{q',\mathrm{du}}(\nu) f\left(\zeta^{\mathcal{N}}_{q',\mathrm{du}}\right) \\ & - \sum_{q,q'=1}^{N} \sum_{m=1}^{Q} \Theta_m(\nu)\alpha_q(\nu)\alpha_{q',\mathrm{du}}(\nu) a_m(\zeta^{\mathcal{N}}_q, \zeta^{\mathcal{N}}_{q',\mathrm{du}}). \end{aligned} \tag{5.45}$$

The parameter independent quantities $f\left(\zeta^{\mathcal{N}}_{q',\mathrm{du}}\right)$ and $a_m(\zeta^{\mathcal{N}}_q, \zeta^{\mathcal{N}}_{q',\mathrm{du}})$ can be computed offline. Hence, evaluation of the dual correction becomes independent on the number of FEM degrees of freedom.

Error bounds

According to Section 4, the error estimator for the dual corrected output of interest (4.37) is given by:

$$\Delta^{\mathrm{o}}_{\mathrm{pd}}(\nu) = \frac{1}{\beta(\nu)} \left|\left|r^{\mathrm{pr}}\left(\cdot;u_N;\nu\right)\right|\right|_{(X^{\mathcal{N}})'} \left|\left|r^{\mathrm{du}}\left(\cdot;w_N;\nu\right)\right|\right|_{(X^{\mathcal{N}})'}. \tag{5.46}$$

The online-offline decomposition for computation of the norm of the dual residuum follows exactly that of the primal residuum given in Section 5.3.2.

Operation count

The online and offline operation count for solution of the dual problem is equal to the primal reduced basis problem. In addition the dual correction (5.45) has to be determined. Computation of the $\mathcal{N}$ dependent quantities in (5.45) has offline costs of $O(QN^2\mathcal{N})$. Online evaluation of the dual correction then has costs of $O(QN^2)$.

For several outputs of interest, in general, all dual online and offline costs have to be multiplied by the number of output functionals, since a separate dual correction has to be computed for each output. This can make the usage of dual corrections infeasible.

5.5 Effectivities

In order to quantify the performance of the error estimators, we introduce effectivities according to Section 4.6. They measure the ratio between estimated and true error:

$$\eta^H(\nu) = \frac{\Delta^H(\nu)}{||u^{\mathcal{N}} - u_N||_{X^{\mathcal{N}}}}, \tag{5.47a}$$

$$\eta^{\mathrm{o}}(\nu) = \frac{\Delta^{\mathrm{o}}(\nu)}{|s^{\mathcal{N}}(\nu) - s_N(\nu)|}, \tag{5.47b}$$

$$\eta^{\mathrm{o}}_{\mathrm{pd}}(\nu) = \frac{\Delta^{\mathrm{o}}_{\mathrm{pd}}(\nu)}{|s^{\mathcal{N}}(\nu) - s^{\mathrm{pd}}_N(\nu)|}, \tag{5.47c}$$

for the reduced basis solution in $X^{\mathcal{N}}$-norm, and the output of interest using the primal and primal-dual reduced basis approximations. Since the estimators are rigorous, we have $\eta^H(\nu)$, $\eta^{\mathrm{o}}(\nu)$, $\eta^{\mathrm{o}}_{\mathrm{pd}}(\nu) \geq 1$.

5.6 Greedy sampling strategy

In Section 5.2.1 we introduced Lagrange reduced basis spaces. These spaces consist of snapshot solutions of the truth approximation for N parameter values in the parameter domain $S_N \subset D$. An important question is, how to choose the snapshot parameter set S_N. Since the online costs of reduced basis computations depend on N, our goal is to construct reduced basis spaces of smallest possible dimension with good approximation properties. Hence, we have to choose S_N, such that a maximum of information about the manifold of possible solutions $\mathcal{M}^{\mathcal{N}}$ (5.5) is included in the reduced basis space X_N.

For selection of snapshot parameters, we use a greedy algorithm [82], which is explained in the following. First a so-called training set $\Xi_{\mathrm{train}} \in D$ of candidate snapshot has to be defined, from which the reduced basis snapshot parameters are chosen iteratively. The first snapshot parameter is chosen randomly from the training sample. Now suppose, a reduced basis approximation of dimension i is already built. For selection of the next snapshot parameter, a reduced basis error estimator $\Delta_i(\nu)$ as introduced in Section 5.3 is utilized. The error estimate using the current reduced basis approximation is computed for each parameter in the training set. Then the snapshot for the parameter corresponding to the largest estimated error is included into the current basis.

The motivation for this procedure is that a candidate snapshot with large error is not well approximated by the current basis, and its inclusion adds a “maximum of new information” into the reduced basis.

In an actual implementation a criterion which defines, if the reduced basis approximation is sufficiently accurate, could be introduced by:

$$\Delta_i(\nu) < \epsilon\,, \quad \forall \nu \in \Xi_{\mathrm{train}},$$

Algorithm 1 Greedy construction of reduced basis space X_N

```
choose training sample $\Xi_{\text{train}} \in D$, maximum reduced basis dimension $N_{\max} \geq 1$, and error threshold $\epsilon$
choose $\nu^1 \in \Xi_{\text{train}}$ randomly
compute snapshot $u_N(\nu^1)$
$X_1 = u_N(\nu^1)$
orthonormalize $X_1$
compute $\beta(\nu^1)$
$\hat{\beta}^{\text{LB}} = \beta(\nu^1)$
construct $\Delta_1$
$\nu^2 = \arg\max_{\nu \in \Xi_{\text{train}}} \Delta_1(\nu)$
$i = 2$
while $(\Delta_{i-1}(\nu^i) \geq \epsilon)$ && $(i \leq N_{\max})$ do
    compute snapshot $u_N(\nu^i)$
    $X_i = X_{i-1} \cup u_N(\nu^i)$
    orthonormalize $X_i$
    compute $\beta(\nu^i)$
    if $\beta(\nu^i) < \hat{\beta}^{\text{LB}}$ then
        $\hat{\beta}^{\text{LB}} = \beta(\nu^i)$
    end if
    construct $\Delta_i$
    $\nu^{i+1} = \arg\max_{\nu \in \Xi_{\text{train}}} \Delta_i(\nu)$
    $i = i + 1$
end while
$X_N = \text{span}\{X_N\}$
```

where the threshold ϵ has to be specified by the user. This guarantees that the reduced basis approximation is sufficiently accurate over the training set. Furthermore, the greedy sampling can be stopped, if the reduced basis space reaches a maximum dimension $N_{\max}$. The greedy construction algorithm, which is used in our numerical examples is stated in Algorithm 1. It also includes computation of an estimate to the global lower bound of the inf-sup constant, as explained in Section 5.3.2.

The step "construct Δ_i" includes offline steps for projection of the affine decomposition onto the reduced basis and for construction of error estimator given in Sections 5.2.2 and 5.3.2 respectively. It is reasonable to choose the estimator for the output of interest (5.27) in Algorithm 1.

If we include the dual problem, Algorithm 1 is modified, and a dual reduced basis space X_N^{du} is constructed in addition to the primal reduced basis space.

Algorithm 2 Greedy construction of primal and dual reduced basis spaces X_N and X_N^{du}

```
choose training sample $\Xi_{\mathrm{train}} \in D$, maximum reduced basis dimension $N_{\max} \geq 1$, and error threshold $\epsilon$
choose $\nu^1 \in \Xi_{\mathrm{train}}$ randomly
compute primal snapshot $u_N(\nu^1)$
$X_1 = u_N(\nu^1)$
orthonormalize $X_1$
compute $\beta(\nu^1)$
$\hat{\beta}^{\mathrm{LB}} = \beta(\nu^1)$
construct $\Delta_1$
$\nu_{\mathrm{du}}^1 = \arg \max\limits_{\nu \in \Xi_{\mathrm{train}}} \Delta_1(\nu)$
compute dual snapshot $w_N(\nu_{\mathrm{du}}^1)$
$X_1^{\mathrm{du}} = w_N(\nu_{\mathrm{du}}^1)$
orthonormalize $X_1^{\mathrm{du}}$
construct $\Delta_2$
$\nu^2 = \arg \max\limits_{\nu \in \Xi_{\mathrm{train}}} \Delta_2(\nu)$
$i = 2$
while $(\Delta_{2i-2}(\nu^i) \geq \epsilon)$ && $(i \leq N_{\max})$ do
    compute primal snapshot $u_N(\nu^i)$
    $X_i = X_{i-1} \cup u_N(\nu^i)$
    orthonormalize $X_i$
    compute $\beta(\nu^i)$
    if $\beta(\nu^i) < \hat{\beta}^{\mathrm{LB}}$ then
        $\hat{\beta}^{\mathrm{LB}} = \beta(\nu^i)$
    end if
    construct $\Delta_{2i-1}$
    $\nu_{\mathrm{du}}^i = \arg \max\limits_{\nu \in \Xi_{\mathrm{train}}} \Delta_1(\nu)$
    compute dual snapshot $w_N(\nu_{\mathrm{du}}^i)$
    $X_i^{\mathrm{du}} = X_{i-1}^{\mathrm{du}} \cup w_N(\nu_{\mathrm{du}}^i)$
    orthonormalize $X_i^{\mathrm{du}}$
    construct $\Delta_{2i}$
    $\nu^{i+1} = \arg \max\limits_{\nu \in \Xi_{\mathrm{train}}} \Delta_i(\nu)$
    $i = i + 1$
end while
$X_N = \mathrm{span}\{X_N\}$
$X_N^{\mathrm{du}} = \mathrm{span}\left\{X_N^{\mathrm{du}}\right\}$
```

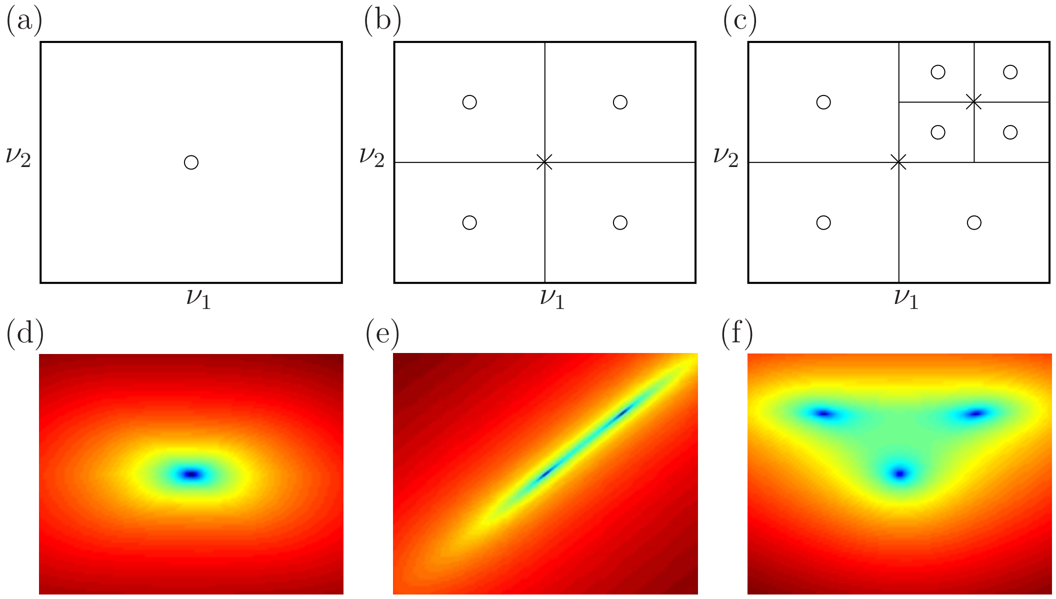


Figure 5.2: (a), (b), (c) Iterative enlargement of training sample. Crosses: snapshots - circles: candidate snapshots. (a) Initial training sample (b), (c) training sample after first and second snapshot. (d), (e), (f) Logarithm of error estimate over parameter domain for one, two, and three dimensional reduced basis approximation – corresponds to numerical example of Section 6.3 with variable parameters width $\nu_1 = W$ and length $\nu_2 = H$. The figures have different scaling (blue: small error bound, red: large error bound).

This is given in Algorithm 2, and of course the primal-dual error estimator for output of interest (5.27) is chosen for selection of the snapshot parameters.

For training samples Ξ_{train} with many elements, computation of the maximum $\arg \max_{\nu \in \Xi_{\text{train}}} \Delta_i(\nu)$ can become very expensive, especially for higher dimensional parameter spaces. To circumvent this, we use an adaptive refinement strategy for the training sample [66]. We start with a small Ξ_{train} and refine it, after a snapshot has been made, as follows. We assume the parameter domain D is a p-dimensional cube in $\mathbb{R}^p$. At the beginning the training sample only includes the center of this cube, which is consequently chosen as snapshot parameter. The chosen parameter is removed from Ξ_{train}, and the cube is subdivided into 2^p new cubes, see Fig. 5.2(a), (b), and (c) for illustration in a 2D parameter space. The 2^p new center points of the cubes are included into the training sample, and the next parameter is chosen. The corresponding center point of the cube is removed from the training sample again, and the cube divided into 2^p new cubes as before. Therewith the training sample grows moderately by $2^p - 1$, when a snapshot is made.

A motivation for this adaptive approach is given in Figures 5.2(d), (e), and (f). The reduced basis error bound over a 2D parameter domain is shown for a one-, two-, and three-dimensional reduced basis approximation. For the type of wave equation we are considering, a snapshot at a specific point in the parameter space mainly leads to a local improvement of the reduced basis approximation. Therefore, it seems reasonable to include candidate snapshots with increasing density into the training sample. In general, the adaptive approach could lead to failures, if a candidate snapshot is very well approximated by a different parameter, which is already included in the basis. This candidate might then not be included in the basis and "blocks" the inclusion of new close candidates into the training sample. A way out could be an occasional dense sweep over the parameter domain during adaptive construction of the basis, in order to find such problematic cases. However, we did not observe above problems in our applications.

Other approaches to adaptive greedy strategies can be found in [25, 26].

5.7 Affine geometry precondition

In the preceeding sections we saw that affine decomposition of the bilinear form of the truth approximation (5.4) is essential for efficient online-offline decomposition. The solution of the reduced basis approximation (5.7), as well as error estimation (5.36) becomes independent on the number of FEM degrees of freedom $\mathcal{N}$. In the following we show for which types of parametrized geometries an affine decomposition of the system bilinear form can be constructed. These geometries are said to hold the affine geometry precondition [82].

We start by introducing a mapping G, which maps a reference configuration of the geometry with corresponding parameters ν_{ref} onto a new configuration:

$$\begin{aligned} G : \Omega \times D \rightarrow& \Omega, \\ (\mathbf{x}, \nu) \mapsto& G(\mathbf{x}; \nu), \end{aligned} \tag{5.48}$$

where $D \subset \mathbb{R}^p$ is the parameter domain, and $\Omega \subset \mathbb{R}^n$. By construction G becomes the identity on Ω for ν_{ref}:

$$G(\cdot; \nu_{\text{ref}}) = \mathbb{1}_\Omega(\cdot).$$

Next we introduce the Jacobian J of G:

$$\begin{aligned} J : \mathbb{R}^n \times D \rightarrow& \mathbb{R}^n, \\ (x, \nu) \mapsto& J(x; \nu). \end{aligned} \tag{5.49}$$

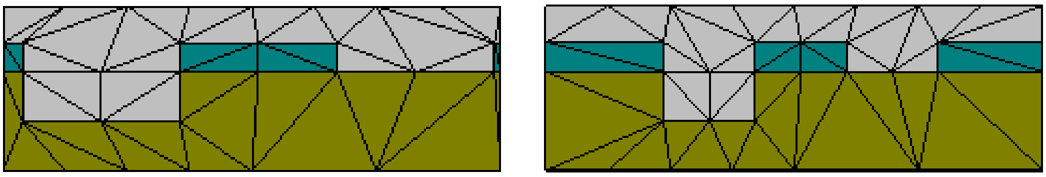


Figure 5.3: Topologically equivalent grids which can be mapped by piecewise linear transformation onto each other.

The affine geometry precondition is defined as follows:

Definition 28. Let $\Omega \subset \mathbb{R}^n$ and G be a mapping as defined in (5.48) with Jacobian J. The pair (Ω, G) holds the **affine geometry precondition**, if there exists a partition of open subsets $P = \{\Omega_1, \dots, \Omega_K\}$ of Ω with the following properties:

(i) $\Omega = \bigcup_{i=1}^{K} \overline{\Omega_i}$,

(ii) $\Omega_i \bigcap \Omega_j = \emptyset$, for $i \neq j$,

(iii) $J(\mathbf{x}; \nu)$ is constant in $\mathbf{x}$ on each Ω_i for fixed $\nu \in D$.

If (iii) holds for the Jacobian J, it can be written as:

$$J(\cdot; \nu) = \sum_{i=1}^{K} J_i(\nu)\, \chi_{\Omega_i}(\cdot), \tag{5.50}$$

where $\chi_{\Omega_i}(\cdot)$ is the characteristic function of Ω_i define by:

$$\begin{aligned} \chi_{\Omega_i} : \Omega &\to \Omega \\ (\mathbf{x}) &\mapsto \chi_{\Omega_i}(\mathbf{x}) = \left\{ \begin{array}{ll} 1: & \mathbf{x} \in \Omega_i \\ 0: & \text{else} \end{array} \right. . \end{aligned} \tag{5.51}$$

We will refer to the sets Ω_i as meta cells of the reduced basis partition of the geometry. For example, the mapping of a triangulation of a geometry onto a topologically equivalent triangulation, as depicted in Fig. 5.3, fulfills the affine geometry precondition. Since the mapping of each triangle onto a deformed triangle has a constant Jacobian, the composed mapping of all triangles has a piecewise constant Jacobian. Having a geometrically parametrized problem which holds the affine geometry precondition, it is possible to construct an affine decomposition of the system bilinear form:

$$a(v,u;\nu)=\sum_{m=1}^{Q}\Theta_m(\nu)a_m(v,u).$$

An explicit derivation of the affine decomposition for Maxwell's equations will be given in the following section.

5.8 Affine decomposition of Maxwell's equations

For construction of the affine decomposition of geometrically parametrized Maxwell scattering problems, we focus on the three dimensional case: $\Omega\subset\mathbb{R}^3$.

Let us recall, how Maxwell's equations transform under a coordinate transformation G. The sesquilinear form of the scattering problem reads according to (3.18):

$$a(u,v)=\int_{\Omega} d\bar{v}\wedge *_{\mu}^{-1}d\tilde{u}^{\mathrm{PML}}-\omega^2\bar{v}\wedge *_{\varepsilon}\tilde{u}^{\mathrm{PML}}.$$

The transformation rules (2.41) and (2.42) describe how $*_\mu$ and $*_\varepsilon$ transform under G. Since G is parameter dependent, the sesquilinear form also becomes parameter dependent:

$$\begin{aligned}a\left(u,v;\nu\right)=&\int_{\Omega} d\bar{v}\wedge\left(\frac{1}{|J(\mathbf{x};\nu)|}J(\mathbf{x};\nu)^T *_{\mu}^{-1}(\mathbf{x})J(\mathbf{x};\nu)\right)du\\&-\int_{\Omega}\omega^2\bar{v}\wedge\left(|J(\mathbf{x};\nu)|J(\mathbf{x};\nu)^{-1} *_{\varepsilon}(\mathbf{x})J(\mathbf{x};\nu)^{-T}\right)u.\end{aligned}\tag{5.52}$$

Now suppose that the parametrized geometry (Ω,G) holds the affine geometry precondition, as defined in Section 5.7. Furthermore, we require that the permeability and permittivity tensors $*_\mu, *_\varepsilon$ respect the reduced basis partition of Ω into $\bigcup_{i=1}^{K}\overline{\Omega_i}$, such that they take constant values $*_{\mu,i}$ and $*_{\varepsilon,i}$ on each Ω_i. Then we can define the parameter dependent stiffness tensor:

$$\begin{aligned}S(\mathbf{x};\nu)=&\frac{1}{|J(\mathbf{x};\nu)|}J(\mathbf{x};\nu)^T *_{\mu}^{-1} J(\mathbf{x};\nu)\\=&\sum_{i=1}^{K}\frac{1}{|J_i(\nu)|}J_i(\nu)^T *_{\mu,i}^{-1} J_i(\nu)\,\chi_{\Omega_i}(\mathbf{x})\\=&\sum_{i=1}^{K}S^i(\nu)\,\chi_{\Omega_i}(\mathbf{x}),\end{aligned}\tag{5.53}$$

with:

$$S^i(\nu) = \frac{1}{|J_i(\nu)|} J_i(\nu)^T *_{\mu,i}^{-1} J_i(\nu), \tag{5.54}$$

and the parameter dependent mass tensor:

$$\begin{aligned} M(\mathbf{x};\nu) =& |J(\mathbf{x};\nu)|\ J(\mathbf{x};\nu)^{-1} *_\varepsilon (\mathbf{x}) J(\mathbf{x};\nu)^{-T} \\ =& \sum_{i=1}^{K} |J_i(\nu)|\ J_i(\nu)^{-1} *_{\varepsilon,i} J_i(\nu)^{-T}\, \chi_{\Omega_i}(\mathbf{x}) \\ =& \sum_{i=1}^{K} M^i(\nu)\, \chi_{\Omega_i}(\mathbf{x}), \end{aligned} \tag{5.55}$$

with:

$$M^i(\nu) = |J_i(\nu)|\ J_i(\nu)^{-1} *_{\varepsilon,i} J_i(\nu)^{-T}. \tag{5.56}$$

First we insert (5.53) into the stiffness integral of (5.52):

$$\begin{aligned} &\int_\Omega d\bar{v} \wedge \left(\frac{1}{|J(\mathbf{x};\nu)|} J(\mathbf{x};\nu)^T *_\mu^{-1} (\mathbf{x}) J(\mathbf{x};\nu) \right) du = \\ &\int_\Omega d\bar{v} \wedge \sum_{i=1}^{K} S^i(\nu)\, \chi_{\Omega_i}(\mathbf{x}) du = \\ &\sum_{i=1}^{K} \int_{\Omega_i} d\bar{v} \wedge S^i(\nu)\, du. \end{aligned}$$

Let us denote by S^i_{jk} the jk-component of the symmetric 3 by 3 tensor S^i. Furthermore, for clarity we use the classical notion $du \cong \nabla \times \vec{u}$. Expanding S^i into components, we then get (note the symmetry in the second line):

$$\begin{aligned} \sum_{i=1}^{K} S^i_{11}(\nu) \int\limits_{\Omega_i} (\nabla \times \vec{\bar{v}}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} (\nabla \times \vec{u}) + \\ S^i_{12}(\nu) \int\limits_{\Omega_i} (\nabla \times \vec{\bar{v}}) \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} (\nabla \times \vec{u}) + \cdots + \\ S^i_{33}(\nu) \int\limits_{\Omega_i} (\nabla \times \vec{\bar{v}}) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} (\nabla \times \vec{u}), \end{aligned}$$

and accordingly for the mass integral of (5.52):

$$\sum_{i=1}^{K} M_{11}^i(\nu) \int_{\Omega_i} \bar{\vec{v}} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \vec{u} +$$

$$M_{12}^i(\nu) \int_{\Omega_i} \bar{\vec{v}} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \vec{u} + \cdots +$$

$$M_{33}^i(\nu) \int_{\Omega_i} \bar{\vec{v}} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \vec{u}.$$

Hence, we arrive at an affine decomposition of the form:

$$a(v, u; \nu) = \sum_{i=1}^{K} \sum_{j=1}^{L} \Theta_i^j(\nu) a_i^j(v, u), \tag{5.57}$$

where the parameter dependent functions $S_{kl}^i(\nu)$ and $M_{mn}^i(\nu)$ were rearranged and relabeled by $\Theta_i^j(\nu)$. We could rearrange the sums over i and j into a single sum (which would give an affine decomposition of form (5.16)), however, for implementation of the following sub-domain residuum method we will need this form of the affine decomposition. The important property is that the integral of the sesquilinear form a_i^j is only carried out over the meta cell Ω_i, for $j = 1, \ldots, L$.

For Maxwell's equations in 3D we have in general $L = 12$ terms on each Ω_i, 2×6 terms for the components of the symmetric 3 by 3 tensors S^i and M^i. For other elliptic operators a similar decomposition can be derived, where the constant L depends on the type of equation that is transformed. For an arbitrarily polarized electric field in 2D we have $L = 8$, for TM polarization (electric field in 2D plane) and TE polarization (Helmholtz equation) we have $L = 4$.

Construction of piecewise affine mappings

Finally we comment on the construction of the mapping G, such that the parametrized geometry holds the affine geometry precondition. Let us look at the transformation of a reference tetrahedron with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ onto a tetrahedron with vertices x^0, x^1, x^2, x^3, with $x^i \in \mathbb{R}^3$. The mapping G of this transformation is given by:

$$G(x, y, z) = \begin{pmatrix} x_1^1 - x_1^0 & x_1^2 - x_1^0 & x_1^3 - x_1^0 \\ x_2^1 - x_2^0 & x_2^2 - x_2^0 & x_2^3 - x_2^0 \\ x_3^1 - x_3^0 & x_3^2 - x_3^0 & x_3^3 - x_3^0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

where x_j^i is the j-th Cartesian component of x^i. Hence, the Jacobian Matrix is constant:

$$J = \begin{pmatrix} x_1^1 - x_1^0 & x_1^2 - x_1^0 & x_1^3 - x_1^0 \\ x_2^1 - x_2^0 & x_2^2 - x_2^0 & x_2^3 - x_2^0 \\ x_3^1 - x_3^0 & x_3^2 - x_3^0 & x_3^3 - x_3^0 \end{pmatrix}.$$

Since the transformation of a tetrahedron onto another deformed tetrahedron can be constructed via an intermediate transformation to a reference tetrahedron, the joined Jacobian is also constant. If we now subdivide Ω into tetrahedrons and define G such that it produces a topologically equivalent tetrahedral mesh, its Jacobian is piecewise constant as desired. The same holds for transformation of topologically equivalent triangular meshes in 2D, and even curvy triangles can be used for reduced basis triangulations [82].

In our implementation the reduced basis triangulation comes from a coarse Delaunay triangulation of the computational domain, which respects constraint edges of the geometry. Often it is possible to summarize several tetrahedrons (triangles in 2D) into a single meta cell, which transform with same constant Jacobian. This reduces the number of terms in the affine decomposition of the system bilinear form.

5.9 Residuum estimator

The residuum based error estimator given in Section 5.3 uses the dual norm of the residuum of a reduced basis approximation to construct an error bound. The dual norm is evaluated via the Riesz representation of the residuum. The online-costs regarding computational time and memory requirements are $O(Q^2N^2)$, where N is the reduced basis dimension and Q the number of terms in the affine expansion of the system bilinear form.

Especially when constructing the affine decomposition for a geometrically parametrized problem in 3D, Q can become very large $\propto 10^3$. With a typical reduced basis dimension of $N \propto 10^2$ the operation count and memory requirements are of the order of 10^{10}. This corresponds to Terabytes of data and makes error estimation in the presented form basically impossible. Furthermore, the online costs for solution of the reduced basis system is only $O(QN^2)$, a factor Q cheaper than error estimation. In the following, therefore, our goal is to construct a much cheaper error estimator, with costs also of the order $O(QN^2)$ [66].

The estimator is inspired by the sub-domain residuum method, which is used for a posteriori error estimation of finite element solutions [3, 1]. Before considering the reduced basis setup, we give the idea of the method in the following.

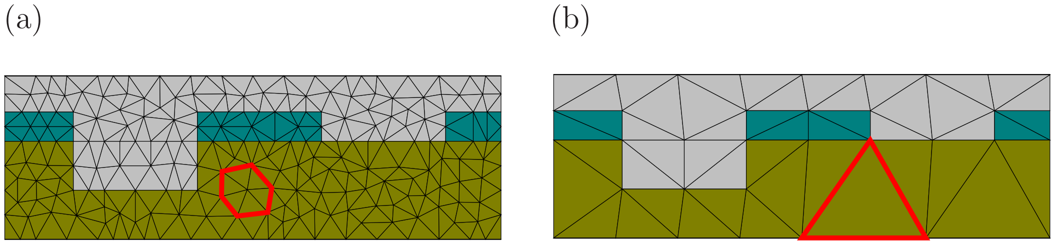


Figure 5.4: (a) Finite element and (b) reduced basis triangulation with exemplary sub-domains for sub-domain residuum method.

5.9.1 Classical sub-domain residuum method

The basic idea of the sub-domain residuum method is to solve the error residuum relationship approximatively, by restricting it to several independent problems on sub-domains of the computational domain Ω. Let us consider the following variational problem:
Find $u \in X(\Omega)$ such that:

$$a(u, v) = f(v), \quad \forall v \in X(\Omega), \tag{5.58}$$

where $a(\cdot, \cdot)$ is a coercive bilinear form and $f \in (X(\Omega))'$. Let $X^{\mathcal{N}}(\Omega) \subset X(\Omega)$ and $u^{\mathcal{N}}$ a solution to (5.58) on $X^{\mathcal{N}}(\Omega)$ given by:
Find $u^{\mathcal{N}} \in X^{\mathcal{N}}(\Omega)$ such that:

$$a(u^{\mathcal{N}}, v) = f(v), \quad \forall v \in X^{\mathcal{N}}(\Omega). \tag{5.59}$$

Now we want to quantify the error $e = u - u^{\mathcal{N}}$. In principle the error residuum relationship (4.9) can be used to compute e:

$$a(e, v) = r(v; u) = f(v) - a(u^{\mathcal{N}}, v), \quad \forall v \in X(\Omega),$$

which is a boundary value problem of same complexity as the original Problem (5.58). This problem will now be simplified as follows.

Let $\mathcal{P}$ be a partition of Ω into patches, e.g., triangles from finite element discretization. The set $\{\vartheta_n\}_{n \in W}$ denotes the first-order Lagrange functions, defined by the element vertices x_m of the partition. The Lagrange basis functions are characterized by:

$$\vartheta_n(x_m) = \delta_{mn}, \tag{5.60}$$

and have the important property:

$$\sum_{n \in W} \vartheta_n(x) = 1, \quad x \in \Omega. \tag{5.61}$$

Furthermore, we define the support of the basis functions:

$$\tilde{\Omega}_n = \operatorname{supp} \vartheta_n \,, \quad n \in W. \tag{5.62}$$

Figure 5.4(a) shows a single sub-domain $\tilde{\Omega}_n$ for a finite element triangulation. Property (5.61) is now used in the error residuum relationship (5.58):

$$\begin{aligned} a(e, v) = a(e, \sum_{n\in W} \vartheta_n v) &= f(\sum_{n\in W} \vartheta_n v) - a(u^{\mathcal{N}}, \sum_{n\in W} \vartheta_n v) \\ = \sum_{n\in W} a(e, \vartheta_n v) &= \sum_{n\in W} \left[f(\vartheta_n v) - a(u^{\mathcal{N}}, \vartheta_n v) \right] , \end{aligned}$$

which can be written as follows:

$$\sum_{n\in W} \left[a(e, \vartheta_n v) - f(\vartheta_n v) + a(u^{\mathcal{N}}, \vartheta_n v) \right] = 0. \tag{5.63}$$

The function $\vartheta_n v$ only has support on $\tilde{\Omega}_n$ and vanishes on $\partial\Omega_n$, hence, $(\vartheta_n v) \in X_0(\tilde{\Omega}_n)$, which is a subset of elements from $X(\tilde{\Omega}_n)$ with zero Dirichlet values on the boundary $\partial\tilde{\Omega}_n$. This gives rise to the definition of following restricted bilinear forms and functionals:

$$a_n(v, w) = a(v, w) \,, \qquad \forall v, w \in X_0(\tilde{\Omega}_n), \tag{5.64}$$

$$f_n(v) = f(v) \,, \qquad \forall v \in X_0(\tilde{\Omega}_n), \tag{5.65}$$

hence, $a_n(v, w) : X_0(\tilde{\Omega}_n) \times X_0(\tilde{\Omega}_n) \mapsto \mathbb{R}(\mathbb{C})$ and $f_n(\cdot) \in \Big(X_0(\tilde{\Omega}_n) \Big)'$.

The sub-domain residuum method consists of solving the following problems, which are motivated by the partition introduced in (5.63):
For $n \in W$ find $e_n \in X_0(\tilde{\Omega}_n)$ such that:

$$a_n(e_n, v) - f_n(v) + a_n(u^{\mathcal{N}}, v) = 0 \,, \quad \forall v \in X_0(\tilde{\Omega}_n). \tag{5.66}$$

These equations can be understood as localized error residuum relationships. The error estimator Δ_n associated with the sub-domain $\tilde{\Omega}_n$ is defined by:

$$\Delta_n = ||e_n||_{a_n} \,, \tag{5.67}$$

where $||\cdot||_{a_n}$ is the energy norm induced by the coercive bilinear form $a_n(\cdot, \cdot)$:

$$||v||_{a_n} = \sqrt{a_n(v, v)} \,, \quad \forall v \in X_0(\tilde{\Omega}_n). \tag{5.68}$$

The global estimate for the error in energy norm is obtained by summing up the local contributions:

$$\Delta = \sqrt{\sum_{n\in W} \Delta_n^2}. \tag{5.69}$$

In [1] the equivalence of this estimator to the true error $||e||_a$ is proven.

The simplification of the sub-domain residuum method is the division of the error residuum relationship into several smaller problems (5.66) on sub-domains.

5.9.2 Reduced basis sub-domain residuum method

We can adapt the idea of the sub-domain residuum method for construction of a cheap residuum estimator in the reduced basis context. The meta cells $\Omega_i\,,\ i = 1, \ldots, K$ introduced in the derivation of the affine decomposition in Section 5.7 appear as natural candidates for the sub-domains, as depicted in Fig. 5.4(b).The key analogon to the classical sub-domain residuum method is the following representation of the identity:

$$\sum_{i=1}^{K} \chi_{\Omega_i}(x) = 1, \quad x \in \Omega, \tag{5.70}$$

which is used instead of (5.61). Since we want to derive an estimate for the dual norm of the residuum, we start with the characterization of the Riesz representation of the residuum (5.31) and use the affine decomposition (5.57) of the sesquilinear form:

$$(v, \hat{e}^{\mathrm{pr}}(\nu))_{X^{\mathcal{N}}} = f(v) - \sum_{i=1}^{K}\sum_{j=1}^{L} \Theta_i^j(\nu) a_i^j(u_N, v)\,, \quad \forall v \in X^{\mathcal{N}}. \tag{5.71}$$

Now we insert the representation of unity (5.70):

$$\begin{aligned}
\left(\sum_{m=1}^{K} \chi_{\Omega_m} v, \hat{e}^{\mathrm{pr}}(\nu)\right)_{X^{\mathcal{N}}} &= f\left(\sum_{m=1}^{K} \chi_{\Omega_m} v\right) - \sum_{i=1}^{K}\sum_{j=1}^{L} \Theta_i^j(\nu) a_i^j\left(u_N, \sum_{m=1}^{K} \chi_{\Omega_m} v\right)\\
\sum_{m=1}^{K} (\chi_{\Omega_m} v, \hat{e}^{\mathrm{pr}}(\nu))_{X^{\mathcal{N}}} &= \sum_{m=1}^{K}\left(f(\chi_{\Omega_m} v) - \sum_{i=1}^{K}\sum_{j=1}^{L} \Theta_i^j(\nu) a_i^j(u_N, \chi_{\Omega_m} v)\right)\\
\sum_{m=1}^{K} (\chi_{\Omega_m} v, \hat{e}^{\mathrm{pr}}(\nu))_{X^{\mathcal{N}}} &= \sum_{m=1}^{K}\left(f(\chi_{\Omega_m} v) - \sum_{j=1}^{L} \Theta_m^j(\nu) a_m^j(u_N, v)\right).
\end{aligned} \tag{5.72}$$

In the last step we used:

$$a_i^j(u_N, \chi_{\Omega_m} v) = \delta_{im} a_i^j(u_N, v),$$

which holds, since $a_i^j(u_N, v)$ only depends on data of u_N and v on Ω_i, and $\chi_{\Omega_m}(\mathbf{x}) = 0$ for $\mathbf{x} \notin \Omega_m$, c.f. discussion after (5.57). Furthermore, $f(\chi_{\Omega_m} \cdot)$ only depends on data on Ω_m.

Inspired by the sub-domain residuum method (5.66), we solve the following localized problems:
For $m = 1, \ldots, K$ find $\hat{e}_m^{\mathrm{pr}}(\nu) \in X_0^{\mathcal{N}}(\Omega_m)$ such that:

$$(v, \hat{e}_m^{\mathrm{pr}}(\nu))_{X^{\mathcal{N}}(\Omega_m)} = \left(f(v) - \sum_{j=1}^{L} \Theta_m^j(\nu) a_m^j(u_N, v) \right), \quad \forall v \in X_0^{\mathcal{N}}(\Omega_m). \tag{5.73}$$

In the classical sub-domain residuum method the choice of homogeneous Dirichlet boundary conditions is motivated by the property that each Lagrange basis function vanishes on the boundary of its support. The corresponding functions χ_{Ω_i} do not have this property. Here we chose homogeneous Dirichlet boundary conditions for computational convenience. Following the error estimate (5.69) of the sub-domain residuum method, the reduced basis residuum estimate is:

$$\epsilon^{\mathrm{pr}}(\nu) = \sqrt{\sum_{m=1}^{K} ||\hat{e}_m^{\mathrm{pr}}(\nu)||^2_{X^{\mathcal{N}}(\Omega_m)}}, \tag{5.74}$$

which is the main result of this section.

In order to quantify the performance of the residuum estimator in numerical examples, we define the following effectivity:

$$\tilde{\eta}^{\mathrm{res}}(\nu) = \frac{\epsilon^{\mathrm{pr}}(\nu)}{||\hat{e}^{\mathrm{pr}}(\nu)||_{X^{\mathcal{N}}}}, \tag{5.75}$$

which measures the ratio between estimated and true residuum.

5.9.3 Online-offline decomposition

For the introduced residuum estimator an online-offline decomposition can be constructed. The following localized problems have to be solved offline:

$$\begin{aligned} f(v) &= (v, b_m)_{X^{\mathcal{N}}(\Omega_m)}, \quad \forall v \in X_0^{\mathcal{N}}(\Omega_m), \\ a_m^j(\zeta_q^{\mathcal{N}}, v) &= \left(v, L_{jm}^q\right)_{X^{\mathcal{N}}(\Omega_m)}, \quad \forall v \in X_0^{\mathcal{N}}(\Omega_m). \end{aligned} \tag{5.76}$$

The localized contributions to the estimate $\epsilon^{\mathrm{pr}}(\nu)$ in (5.74) are then obtained by:

$$\begin{aligned}||\hat{e}^{\mathrm{pr}}_m(\nu)||^2_{X^{\mathcal{N}}(\Omega_m)} = & ||b_m||^2_{X^{\mathcal{N}}(\Omega_m)} \\ & - \sum_{j'=1}^{L}\sum_{q'=1}^{N} \Theta^{j'}_m(\nu)\alpha_{q'}(\nu)\left(b_m, L^{q'}_{j'm}\right)_{X^{\mathcal{N}}(\Omega_m)} \\ & - \sum_{j=1}^{L}\sum_{q=1}^{N} \Theta^{j}_m(\nu)\alpha_{q}(\nu)\left(L^{q}_{jm}, b_m\right)_{X^{\mathcal{N}}(\Omega_m)} \\ & + \sum_{j,j'=1}^{L}\sum_{q,q'=1}^{N} \Theta^{j}_m(\nu)\alpha_{q}(\nu)\Theta^{j'}_m(\nu)\alpha_{q'}(\nu)\left(L^{q}_{jm}, L^{q'}_{j'm}\right)_{X^{\mathcal{N}}(\Omega_m)} . \end{aligned} \tag{5.77}$$

The terms $||b_m||^2_{X^{\mathcal{N}}(\Omega_m)}$, $\left(b_m, L^{q'}_{j'm}\right)_{X^{\mathcal{N}}(\Omega_m)}$, $\left(L^{q}_{jm}, L^{q'}_{j'm}\right)_{X^{\mathcal{N}}(\Omega_m)}$ can also be computed offline, such that evaluation of estimate (5.74) becomes independent on the number of finite element degrees of freedom $\mathcal{N}$.

5.9.4 Operation count

In the offline phase the elliptic problems (5.76) for determination of the Riesz representation have to be solved. Without the sub-domain residuum method this corresponds to NQ forward backward substitution (FBS) with total costs $O(\mathcal{N}NQ)$, a single LU decomposition of a matrix of size $\mathcal{N}$, and computation of scalar products with costs of $O(N^2Q^2\mathcal{N})$, c.f. Section 5.3.3. With the sub-domain residuum method we have to perform K LU decompositions with an average size of $\mathcal{N}/K$ and $NLK = NQ$ FBSs. However, each FBS is only performed on a meta cell, which leads to costs of $O(\mathcal{N}NQ/K) = O(\mathcal{N}NL)$. Computing the scalar products in (5.77) has costs of $O(K \cdot N^2L^2(\mathcal{N}/K)) = O(N^2L^2\mathcal{N})$.

In the online step the computational time and memory costs computing the exact dual norm of the residuum are $O(N^2Q^2)$. With the sub-domain residuum method the leading term according to (5.74) and (5.77) is $O(N^2L^2K) = O(N^2LQ)$.

Using the proposed sub-domain residuum method in the reduced basis context, memory and offline and online computational costs are saved by a factor of K, corresponding to the number of meta cells. In our application especially in 3D we have $K \propto 100$, which results in huge savings, making the reduced basis method applicable to a much larger class of applications. Furthermore, the factor L only depends on the PDE which is solved. L corresponds to the number of affine terms in the matrix expansion on a meta cell (5.57) and is therefore bounded. For 3D Maxwell's equations we have a maximum of $L = 12$ terms, corresponding to the entries in the symmetric 3×3-permittivity and permeability tensors. A constant L means that the sub-domain residuum estimator has online costs of $O(N^2Q)$, which is equal to the assembling costs of the reduced basis system as desired. For the class of problems considered in [82] we have $L = 6$, for 2D Helmholtz problems $L = 4$, etc.

5.10 Multiple sources

In applications it is often necessary to simulate the response of a system for multiple sources. In nano-optics an important example are complex, e.g., so-called dipole or quadrupole [44], illuminations. These can be modeled by a set of P plane waves with different amplitudes, incident angles, and polarizations, which have to be simulated independently.

5.10.1 Problem setup

The truth approximation to a setup with multiple sources reads:

Problem 15. For $i = 1, \ldots, P$ and given tuple of input parameters $\nu \in D \subset \mathbb{R}^p$, compute the outputs of interest:

$$s_i^{\mathcal{N}}(\nu) = l^{\mathrm{o}}\left(u_i^{\mathcal{N}}(\nu)\right), \tag{5.78}$$

where $u_i^{\mathcal{N}}(\nu) \in X^{\mathcal{N}}$ is the solution to the following problem:
Find $u_i^{\mathcal{N}}(\nu) \in X^{\mathcal{N}}$ such that:

$$a\left(u_i^{\mathcal{N}}(\nu), v; \nu\right) = f_i\left(v\right), \quad \forall v \in X^{\mathcal{N}}. \tag{5.79}$$

Here we assume that the system bilinear form does not depend on the source i. For time-harmonic Maxwell's equations this is the case if different sources have same frequency ω, since the incoming frequency ω enters into $a\left(\cdot, \cdot; \nu\right)$, c.f. (3.9). In a Bloch-periodic setting, furthermore, the different sources must have the same phase difference across the domain, since this phase difference also enters the system bilinear form.

If the system bilinear form does not depend on the source, the finite element computation of the truth approximation is basically independent on the number of right hand sides P. Only a single LU-decomposition of the system matrix has to be performed and P forward backward substitutions for different right hand sides. The forward backward substitutions are thereby much cheaper to compute than the LU-decomposition.

For formulation of the reduced basis problem, the finite element space $X^{\mathcal{N}}$ in truth approximation (5.79) is replaced by spaces of global functions $X_N^i \subset X^{\mathcal{N}}$:

Problem 16. For sources $i = 1, \ldots, P$ and given tuple of input parameters $\nu \in D \subset \mathbb{R}^p$, compute the outputs of interest:

$$s_N^i(\nu) = l^{\mathrm{o}}\left(u_N^i(\nu)\right), \tag{5.80}$$

where $u_N^i(\nu) \in X_N^i$ is the solution to the following problem:
Find $u_N^i(\nu) \in X_N^i$ such that:

$$a\left(u_N^i(\nu), v; \nu\right) = f_i(v), \quad \forall v \in X_N^i. \tag{5.81}$$

The space X_N^i is spanned by snapshot solutions to truth approximation (5.79) for different parameter values ν:

$$X_N^i = \operatorname{span}\left\{u_i^{\mathcal{N}}(\nu_1), \ldots, u_i^{\mathcal{N}}(\nu_N)\right\}. \tag{5.82}$$

In contrast to the truth approximation (5.79), the trial and test spaces in the reduced basis problem (5.81) differ for different i, and the reduced basis system matrices will, therefore, also depend on i. This is due to the fact that each reduced basis space X_N^i consists of snapshots, which are computed with different sources $f_i(\cdot)$. Hence, in general P reduced basis systems have to assembled and solved online for P sources. If P becomes large (e.g. $\propto 100$), this dramatically reduces the performance of the reduced basis approximation.

5.10.2 Efficient reduced basis treatment

In the following we will motivate and explain, how to obtain a reduced basis solution for above setting with online costs basically independent on P [67]. Let us have a look at the reduced basis solution to a fixed source j. It is given by:

$$u_N^j = \sum_{q=1}^{N} \alpha_q^j(\nu)\zeta_q^{\mathcal{N},j}, \tag{5.83}$$

where α_q^j are the reduced basis coefficients and $\zeta_q^{\mathcal{N},j}$ the elements of the reduced basis for source j. The primal residuum then reads:

$$r^{\mathrm{pr}}\left(v; u_N^j; \nu\right) = f_j(v) - \sum_{q=1}^{N} \alpha_q^j(\nu) a\left(\zeta_q^{\mathcal{N},j}, v; \nu\right), \quad \forall v \in X^{\mathcal{N}}. \tag{5.84}$$

The elements of the reduced basis are computed from snapshot solutions with parameters ν_q. Let us assume for simplicity that they are not orthogonalized against each other. Then the elements of the reduced basis are solutions to:
Find $\zeta_q^{\mathcal{N},j} \in X^{\mathcal{N}}$ such that:

$$a\left(\zeta_q^{\mathcal{N},j}, v; \nu_q\right) = f_j(v), \quad \forall v \in X^{\mathcal{N}}. \tag{5.85}$$

Now we introduce the Riesz representation of functional $f_j\left(\cdot\right)$:

$$f_j\left(v\right)=\left(v,\tilde{f}_j\right)_{X^{\mathcal{N}}}, \quad \forall v\in X^{\mathcal{N}}, \tag{5.86}$$

with $\tilde{f}_j\in X^{\mathcal{N}}$ and utilize the representation operator of the functional $a\left(\zeta_q^{\mathcal{N},j},\cdot;\nu\right)\in\left(X^{\mathcal{N}}\right)'$, introduced in (5.23):

$$a\left(\zeta_q^{\mathcal{N},j},\cdot;\nu\right)=\left(\cdot,T_\nu\zeta_q^{\mathcal{N},j}\right)_{X^{\mathcal{N}}}, \quad \forall\zeta_q^{\mathcal{N},j}\in X^{\mathcal{N}}. \tag{5.87}$$

The snapshot Problem (5.85) is then equivalent to:
Find $\zeta_q^{\mathcal{N},j}\in X^{\mathcal{N}}$ such that:

$$\left(v,T_{\nu_q}\zeta_q^{\mathcal{N},j}\right)_{X^{\mathcal{N}}}=\left(v,\tilde{f}_j\right)_{X^{\mathcal{N}}}, \quad \forall v\in X^{\mathcal{N}}, \tag{5.88}$$

from which we conclude:

$$\zeta_q^{\mathcal{N},j}=\left(T_{\nu_q}\right)^{-1}\tilde{f}_j. \tag{5.89}$$

The primal residuum (5.84) can now be rewritten as follows:

$$\begin{aligned} r^{\mathrm{pr}}\left(v;u_N^j;\nu\right)=&f_j\left(v\right)-\sum_{q=1}^{N}\alpha_q^j(\nu)a\left(\zeta_q^{\mathcal{N},j},v;\nu\right)\\ =&\left(v,\tilde{f}_j\right)_{X^{\mathcal{N}}}-\sum_{q=1}^{N}\alpha_q^j(\nu)a\left(\left(T_{\nu_q}\right)^{-1}\tilde{f}_j,v;\nu\right)\\ =&\left(v,\tilde{f}_j\right)_{X^{\mathcal{N}}}-\sum_{q=1}^{N}\alpha_q^j(\nu)\left(v,T_\nu\left(T_{\nu_q}\right)^{-1}\tilde{f}_j\right)_{X^{\mathcal{N}}}\\ =&\left(v,\left\{\mathbb{1}_{S^{\mathcal{N}}}-T_\nu\left[\sum_{q=1}^{N}\alpha_q^j(\nu)\left(T_{\nu_q}\right)^{-1}\right]\right\}\tilde{f}_j\right)_{X^{\mathcal{N}}}, \end{aligned} \tag{5.90}$$

where $\mathbb{1}_{S^{\mathcal{N}}}$ is the identity on the space of Riesz representations of all sources:

$$S^{\mathcal{N}}=\operatorname{span}\left\{\tilde{f}_1,\ldots,\tilde{f}_P\right\}. \tag{5.91}$$

Hence, the residuum will be small, and the truth approximation will be well approximated if the operator:

$$T_\nu\left[\sum_{q=1}^{N}\alpha_q^j(\nu)\left(T_{\nu_q}\right)^{-1}\right] \tag{5.92}$$

provides a good approximation to $\mathbb{1}_{S^{\mathcal{N}}}$ or in other words:

$$\sum_{q=1}^{N} \alpha_q^j(\nu) \left(T_{\nu_q}\right)^{-1} \approx T_\nu^{-1}, \quad \text{on } S^{\mathcal{N}}.$$

The key observation is that the operator (5.92) does not depend on the functional $\tilde{f}_j$ itself.

Instead of computing reduced basis coefficients $\alpha_q^j(\nu)$ for all sources j, we therefore compute the coefficients only for a single source i and use them for all other sources, i.e.:

$$\alpha_q^j(\nu) = \overline{\alpha}_q(\nu) := \alpha_q^i(\nu)\,, \quad j = 1, \ldots, P\,,\ q = 1, \ldots, N, \tag{5.93}$$

hence, we use a single representative $\tilde{f}_i$ from $S^{\mathcal{N}}$ to compute the reduced basis coefficients, and hope that the operator (5.92) provides a good approximation to unity on the whole space $S^{\mathcal{N}}$. We will show in numerical examples that this technique provides very good results.

As motivated above, in the offline phase the snapshots $\zeta_q^{\mathcal{N},j}$ for different sources j are computed at the same parameters ν_q. Then only a single LU-decomposition has to be computed for determination of snapshots solutions $\zeta_q^{\mathcal{N},j}$ for all sources $j = 1, \ldots, P$. The offline costs, therefore, do not increase significantly with increasing number of sources, since the LU decomposition is much more expensive then FBSs for all sources.

For numerical stability the reduced basis for source i is orthogonalized:

$$\hat{\zeta}_r^{\mathcal{N},i} = \sum_{q=1}^{N} \gamma_{rq} \zeta_q^{\mathcal{N},i}, \tag{5.94}$$

where $\hat{\zeta}_r^{\mathcal{N},i}$ then is the orthonormal basis. The outputs of interest for all sources j are computed by:

$$s_N^j(\nu) = \sum_{r=1}^{N} \overline{\alpha}_r(\nu) l^{\mathrm{o}} \left(\hat{\zeta}_r^{\mathcal{N},j}\right), \tag{5.95}$$

where the basis functions $\hat{\zeta}_r^{\mathcal{N},j}$ for sources $j \neq i$ are given by:

$$\hat{\zeta}_r^{\mathcal{N},j} = \sum_{q=1}^{N} \gamma_{rq} \zeta_q^{\mathcal{N},j}, \quad i = 1, \ldots, P, \tag{5.96}$$

i.e., the snapshots of all sources j are "orthogonalized" (transformed) with the same coefficients γ_{rq} as for source i.

Unfortunately the costs for error estimation of outputs of interest depends on the number of sources P and is, therefore, infeasible for a large number of sources. In numerical examples we will, however, see that the convergence rate for the outputs of all sources is basically equal. Hence, a small estimate for single source i gives confidence that reduced basis outputs for all sources provide accurate results.

6 Application examples

In this chapter we apply the developed reduced basis techniques to 2D and 3D electromagnetic scattering problems. The finite element solver used for implementation of the reduced basis method is JCMsuite, developed by the Zuse-Institute Berlin and JCMwave GmbH for the numerical solution of Maxwell's equations [13]. JCMsuite provides high-order edge elements, adaptive grid refinement, and transparent boundary conditions for multiply structured exterior domains. It offers the possibility for electromagnetic field computations in a wide range of nano-optical applications, including waveguide structures [11], nano resonators [40], DUV phase masks [12], and other nano-structured materials [17, 39, 30, 14]. The application field of our examples is computational lithography [44].

Lithography is a technique, which is nowadays used for fabrication of basically all integrated circuits [44, 103]. The basic principle is depicted in Fig. 6.1. In the lithographical process the design pattern of a circuit is imaged via a photomask onto a wafer, which is covered with photo sensitive resist. In illuminated areas the resist is chemically transformed and afterwards developed. This gives the desired patterns on the waver.

The following numerical examples will deal with application of the reduced basis method to simulation of light transmission through geometrically parametrized photomasks.

We start with a 2D and a 3D mask example, before we consider two more complex applications in the field of inverse scatterometry and mask pattern optimization. We will analyze the performance and convergence of the developed certified reduced basis techniques, where an evaluation of the proposed sub-domain residuum method and efficient treatment of multiple sources will be an important aspect.

The first real-time application will deal with a real world 2D inverse scatterometry problem. Our goal is fast reconstruction of a grating profile of an extreme ultraviolet (EUV) mask from experimental scatterometric data. This work was done in collaboration with the German national standards and metrology institute (Physikalisch-Technische Bundesanstalt, PTB) and the Advanced Mask Technology Center (AMTC) in Dresden, Germany.

Finally, we consider a challenging 3D example from optical proximity correction (OPC) [44], where the layout of a photomask is optimized, in order to

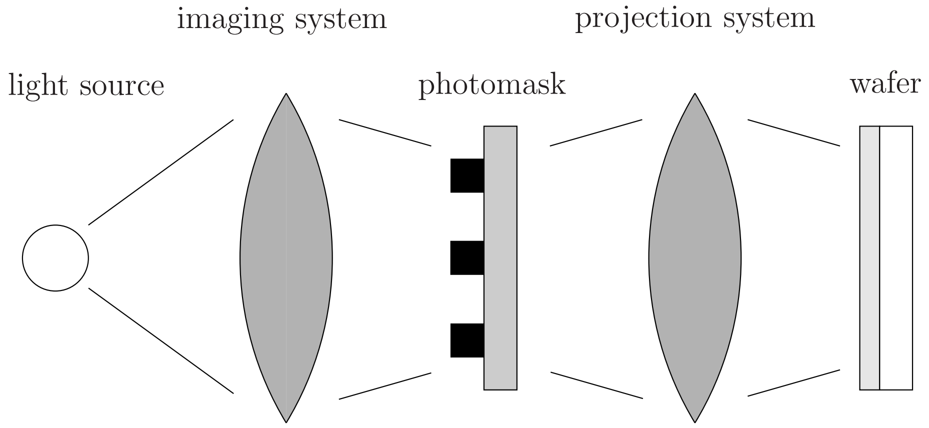


Figure 6.1: Principle setup of optical lithography: a photomask is illuminated by light, passing through an imaging system. The light distribution created by the mask absorber pattern is projected onto a wafer, which is covered with photosensitive resist.

obtain a desired resist pattern on the wafer. This is a many-query application. Thereby we model a complex light source with $P = 74$ sources and have to compute several thousands of outputs of interest. Computational times of the truth approximation are of the order of hours for this example.

All computational times given in this section correspond to single CPU times on a 2.6 GHz AMD Opteron processor.

Before we apply the reduced basis method, we comment on the type of sources and outputs of interest, considered in the numerical examples.

6.1 Incoming fields and outputs of interest

The incoming source fields of the scattering examples will be plane waves, given by:

$$u_{\text{in}}(\mathbf{x}) = \mathbf{E}_0 \exp\left(i\left(n_{\text{ext}}\mathbf{k}_0\right)\cdot\mathbf{x}\right). \tag{6.1}$$

The vector $\mathbf{E}_0$ describes the electric field amplitude, $\mathbf{k}_0$ the vacuum wave vector of the incoming field, and $n_{\text{ext}} = \sqrt{\epsilon_{\text{ext}}\mu_{\text{ext}}}$ the refractive index of the exterior material. Plane waves are often the appropriate model in nano-optics, since the scattering objects are usually much smaller than, e.g., the spot size of a laser. Hence, compared to the scatterer the source field is treated as if it would extend to infinity.

The solution of an electromagnetic scattering problem is the electric field u on Ω. This solution is referred to as near field. In applications the near

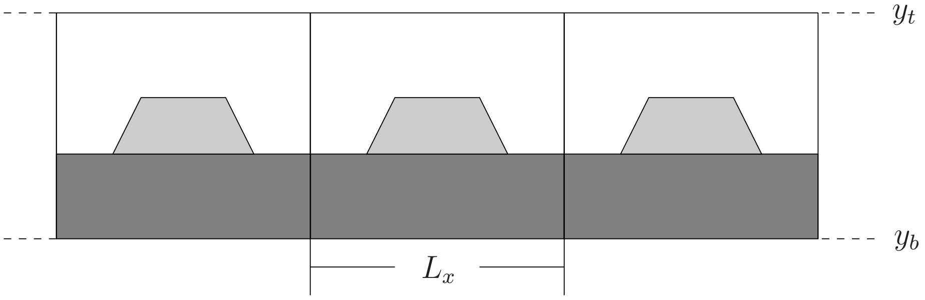


Figure 6.2: Periodic geometry of width L_x. Far field coefficients are computed on top ($y = y_t$) or bottom boundary ($y = y_b$).

field itself is often not of interest. More important are quantities which can be derived from the near field and are, e.g., easily experimentally accessible and basically determine the functionality of the system under consideration. The most important output of interest from a scattering experiment is the so-called far field, which describes the electromagnetic field far away from the scatterer. In the numerical examples we will consider periodic geometries. A periodic domain in 2D is periodified in one spatial dimension (here the x-direction) and in 3D in two spatial dimension (here the x- and y-direction), see Fig. 6.2. For these types of geometries the far field consists of discrete diffraction orders. The diffraction orders are complex Fourier coefficients of the near field on a periodic boundary. They are defined by:

$$\begin{aligned} \hat{\mathbf{E}}_l &= \frac{1}{L_x}\int_{x=0}^{L_x} dx\, u(x, y = y_b, y_t)\, e^{-ik_{x,l}x}, \\ \hat{\mathbf{E}}_{l,m} &= \frac{1}{L_x L_y}\int_{x=0}^{L_x} dx \int_{y=0}^{L_y} dy\, u(x, y, z = z_b, z_t)\, e^{-i\left(k_{x,l}x + k_{y,m}y\right)}, \end{aligned} \tag{6.2}$$

where $\hat{\mathbf{E}}_l$ is used for 2D domains and $\hat{\mathbf{E}}_{l,m}$ for 3D domains. L_x (and L_y in 3D) is the dimension of the periodic boundary, which is located at y_b or y_t, see Fig. 6.2. The k-vectors take values:

$$\begin{aligned} k_{x,l} &= k_{x,0} + \frac{2\pi}{L_x} l \quad \text{with } l \in \mathbb{Z}, \\ k_{y,m} &= k_{y,0} + \frac{2\pi}{L_y} m \quad \text{with } m \in \mathbb{Z}, \end{aligned} \tag{6.3}$$

where $k_{x,0}$ and $k_{y,0}$ are the x- and y-component of the k-vector ($n_{\text{ext}}\mathbf{k}_0$) of the incoming field (6.1). In the 2D (3D) case the y-(z-)component of the k-vector of diffraction orders is given by:

$$\begin{aligned} k_{y,l} &= \sqrt{n_{\text{ext}}^2 \left|\mathbf{k}_0\right|^2 - k_{x,l}^2} && (2D), \\ k_{z,lm} &= \sqrt{n_{\text{ext}}^2 \left|\mathbf{k}_0\right|^2 - k_{x,l}^2 - k_{y,m}^2} && (3D). \end{aligned}$$

In the far field only so-called propagating diffraction orders are visible. These correspond to k-vectors fulfilling the following relation:

$$\begin{aligned} n_{\text{ext}}^2 \left|\mathbf{k}_0\right|^2 - k_{x,l}^2 &> 0 && (2D), \\ n_{\text{ext}}^2 \left|\mathbf{k}_0\right|^2 - k_{x,l}^2 - k_{y,m}^2 &> 0 && (3D). \end{aligned} \tag{6.4}$$

The corresponding modes $\hat{\mathbf{E}}_l$ and $\hat{\mathbf{E}}_{l,m}$ propagate undamped towards infinity. In contrast, the amplitudes of evanescent modes decrease exponentially with increasing distance from the scatterer:

$$\begin{aligned} &e^{-y\sqrt{k_{x,l}^2 - n_{\text{ext}}^2 \left|\mathbf{k}_0\right|^2}} && (2D), \\ &e^{-z\sqrt{k_{x,l}^2 + k_{y,m}^2 - n_{\text{ext}}^2 \left|\mathbf{k}_0\right|^2}} && (3D). \end{aligned}$$

These evanescent modes are not visible in the far field.

Our outputs of interest are, therefore, defined by expanding the field data on the boundary of the computational domain into a Fourier series (6.2), where the k-vectors satisfy (6.4). Linearity of this output functional follows from linearity of integration.

6.2 2D phase shift mask

Figure 6.3 shows the 2D cross section of our first numerical example. A Chromium on glass (CoG) phase shift mask together with the intensity of the electric field is depicted. For this example we choose three geometrical parameters $\{d_1, d_2, d_3\}$ for construction of the reduced basis. The three dimensional parameter domain is given by $D = [340\,\text{nm}, 420\,\text{nm}]^3$, hence

$$\begin{aligned} d_1 &\in [340\,\text{nm}; 420\,\text{nm}], \\ d_2 &\in [340\,\text{nm}; 420\,\text{nm}], \\ d_3 &\in [340\,\text{nm}; 420\,\text{nm}]. \end{aligned} \tag{6.5}$$

The dimensions of the computational domain are $1520\,\text{nm} \times 500\,\text{nm}$. The wavelength of the incoming plane wave is $\lambda = 193\,\text{nm}$. The incoming wave vector is given by:

$$\mathbf{k}_0 = \left(0, 3.25 \cdot 10^7, 0\right),$$

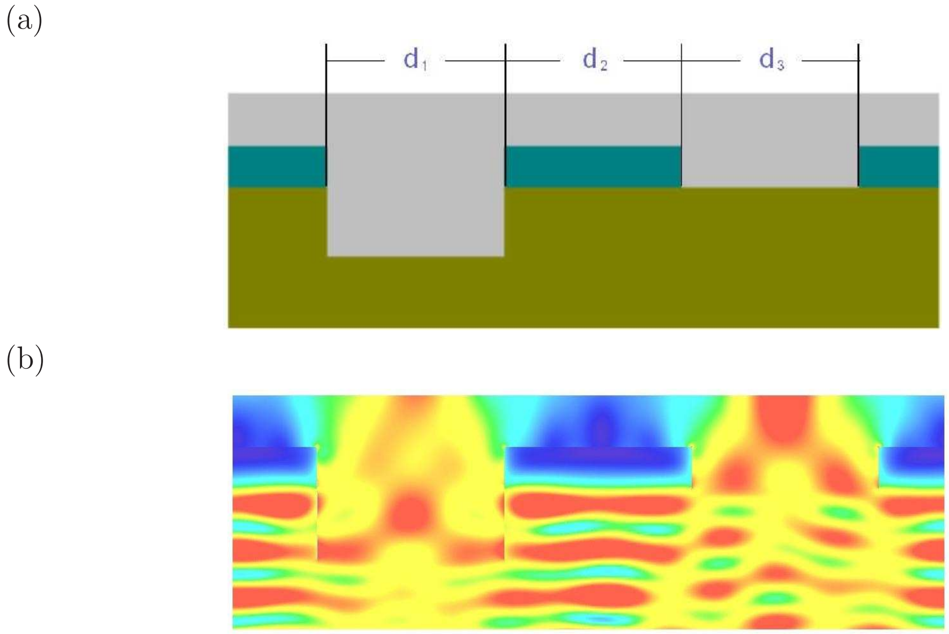


Figure 6.3: (a) Parametrized 2D phase shift mask for reduced basis computation. (b) Intensity of electric field obtained from FEM computation.

hence, we have normal incidence from below. The incoming field vector is given by:

$$\mathbf{E}_0 = (1, 0, 1) \,. \tag{6.6}$$

FEM discretization yields a system with $\mathcal{N} = 236832$ unknowns. The affine decomposition of the system matrix gives $Q = 151$ terms. Two separate reduced basis approximations were built, using a primal-only and primal-dual approximation. For analysis of the reduced basis approximation, the exact solution of the truth approximation was computed for 200 points, chosen randomly in the parameter domain.

First we look at the performance of the sub-domain residuum estimator, introduced in Section 5.9. For each of the 200 points in the random parameter ensemble Ξ the true and estimated residua were compared for increasing dimension of the reduced basis up to 125, i.e., we have 25000 data points for comparison. Figure 6.4(a) shows the estimated residuum $\epsilon^{\mathrm{pr}}(\nu)$ (5.74) in dependence on the true dual norm of the residuum $||\hat{e}^{\mathrm{pr}}(\nu)||_{X^{\mathcal{N}}}$. The corresponding effectivity of the residuum estimate is given in Fig. 6.4(b). The effectivity varies only very little between 2.0 and 5.0 over 6 orders of magnitude of the true residuum, which demonstrates a very good performance of

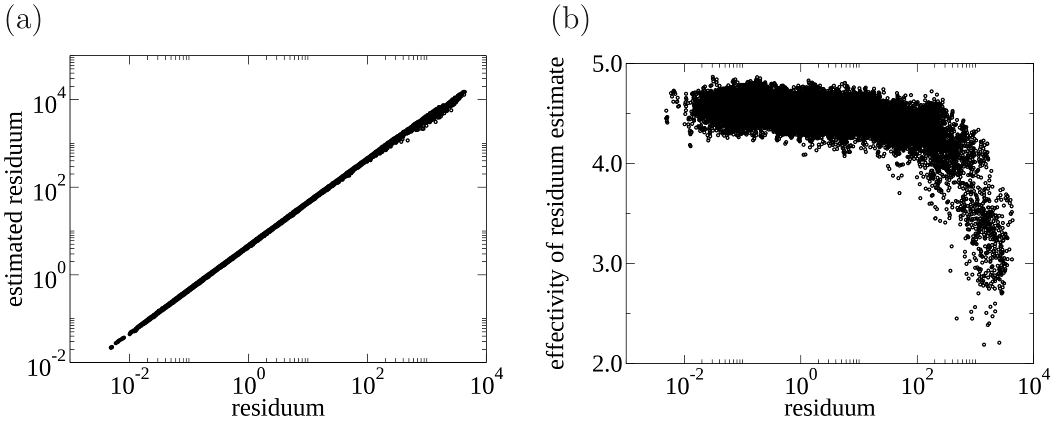


Figure 6.4: (a) Estimated dual norm of residuum (5.74) in dependence on exact dual norm of residuum. (b) Effectivity of residuum estimate (5.75) in dependence on true residuum. (2D mask, Section 6.2)

the reduced basis sub-domain residuum method.

The lower bound for the inf-sup constant (5.37), which was obtained during construction of the reduced basis was $\beta_0 = 3.0176 \cdot 10^{-3}$. The inf-sup constants over the random parameter ensemble varied from $3.0202 \cdot 10^{-3}$ to $3.0316 \cdot 10^{-3}$, which gives a maximum relative error of $\approx 5 \cdot 10^{-3}$ and justifies usage of a constant estimate for the lower bound. Next we look at the convergence of the reduced basis solution in the $\mathrm{H}(\mathbf{curl}, \Omega)$-norm. Figures 6.5(a), (b) show the error and the corresponding error bound. We observe exponential convergence with increasing dimension of the reduced basis approximation. The effectivity of the estimate is shown in Fig. 6.5(c). The effectivities stay between 200 and 900 and increase moderately with decreasing error of the reduced basis solution. In applications the convergence of the error in the output of interest is of importance. We compare a reduced basis approximation, using only the primal reduced basis and a primal-dual reduced basis with corrected output of interest. We restrict us to one vector component of the zeroth transmitted diffraction order, i.e., $l = 0$ in (6.3). The convergence of the primal and primal-dual errors (5.20), (5.42) and the corresponding error bounds (5.27) and (5.46) are shown in Figures 6.6 and 6.7. Again we observe exponential convergence. Incorporation of the dual correction dramatically increases the convergence rate. For a mean target accuracy of 10^{-2} for the error bound, the primal approximation has a dimension of 125, c.f. Fig. 6.6. With dual correction a primal and dual reduced basis approximation of only dimension 28 is needed. Finally, Fig. 6.8 shows the corresponding effectivities of the error estimates for the primal and primal-dual approximation. The effectivities of

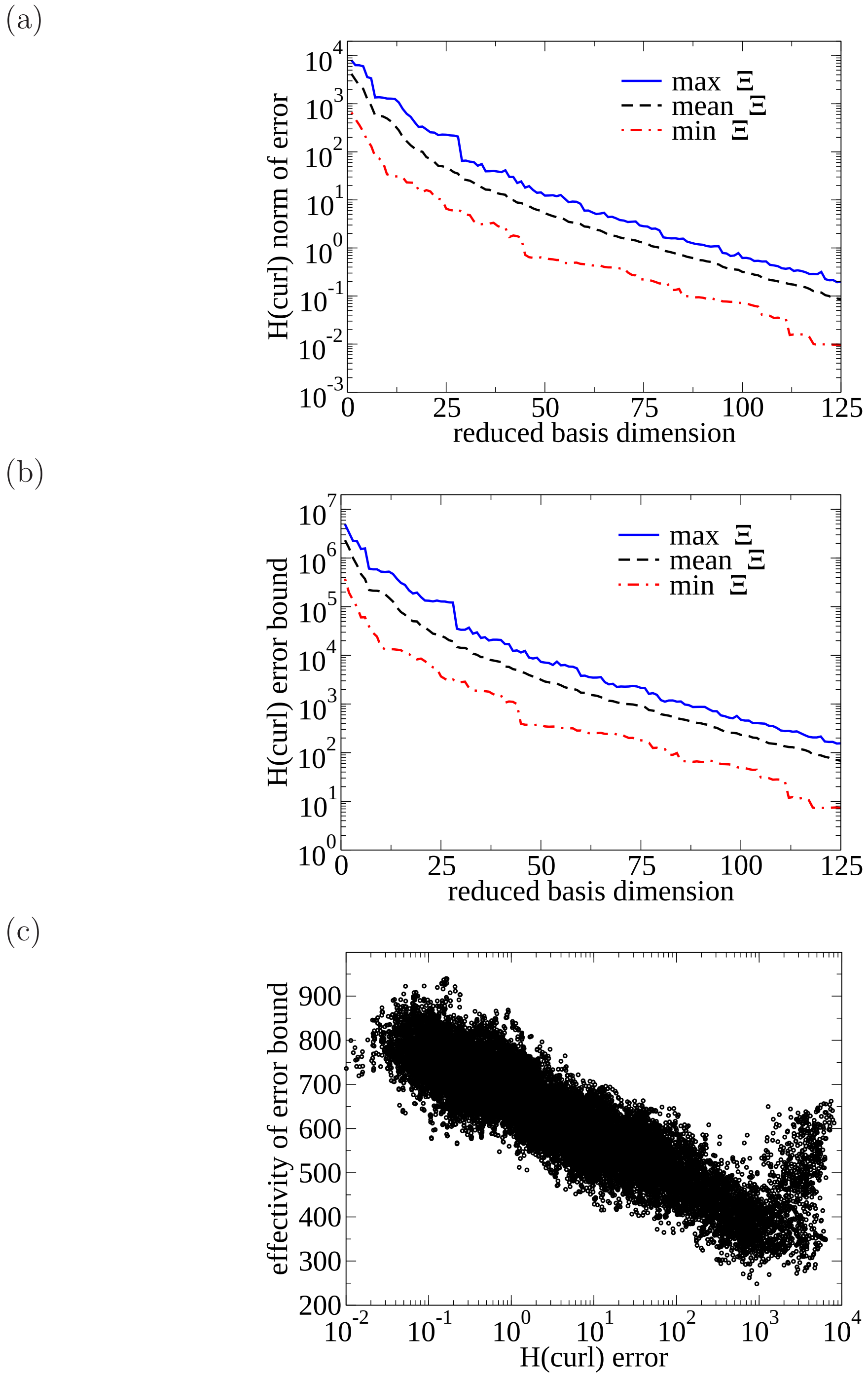


Figure 6.5: (a) Reduced basis solution error (4.8) in $\mathrm{H}\left(\mathbf{curl},\Omega\right)$-norm and (b) corresponding error bound (4.23) in dependence on reduced basis dimension. The mean, minimum and maximum values over the random parameter ensemble Ξ are shown. (c) Effectivity of $\mathrm{H}\left(\mathbf{curl},\Omega\right)$-estimator (5.47a) in dependence on reduced basis solution error. (2D mask, Section 6.2)

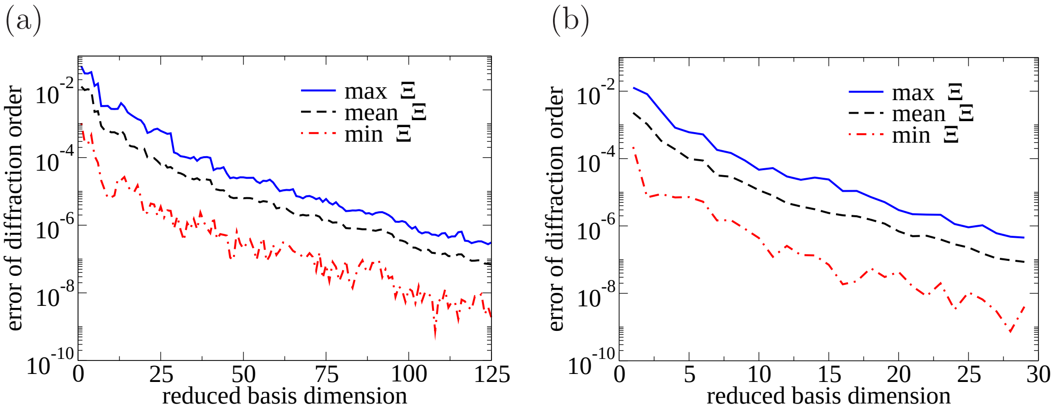


Figure 6.6: Convergence of output of interest in dependence on reduced basis dimension, (a) using only the primal and (b) primal-dual reduced basis approximation. The dimension of the dual reduced basis was set equal to the dimension of the primal reduced basis. The dual corrected solution converges much faster, compare the different scalings of the x-axes. (2D mask, Section 6.2)

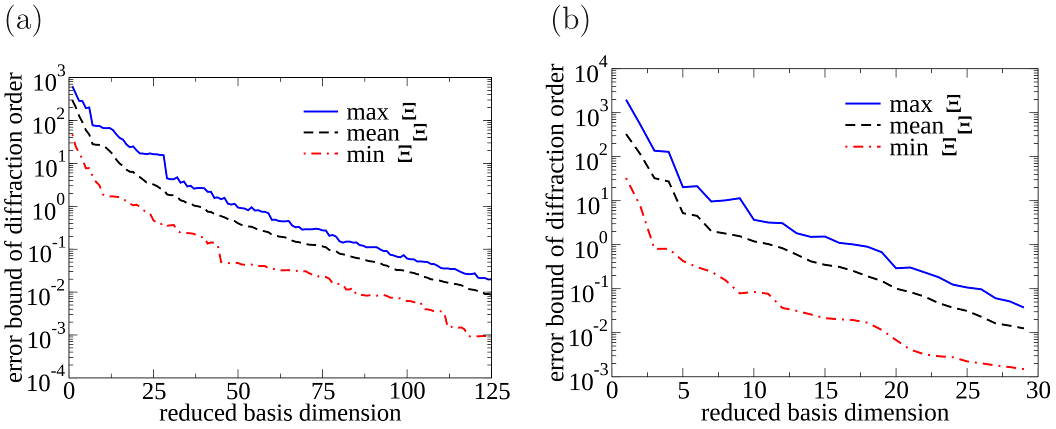


Figure 6.7: Convergence of error bound for output of interest in dependence on reduced basis dimension (a) using only the primal and (b) primal-dual reduced basis approximation. The dimension of the dual reduced basis was set equal to the dimension of the primal reduced basis. The dual corrected solution converges much faster, compare the different scalings of the x-axes. (2D mask, Section 6.2)

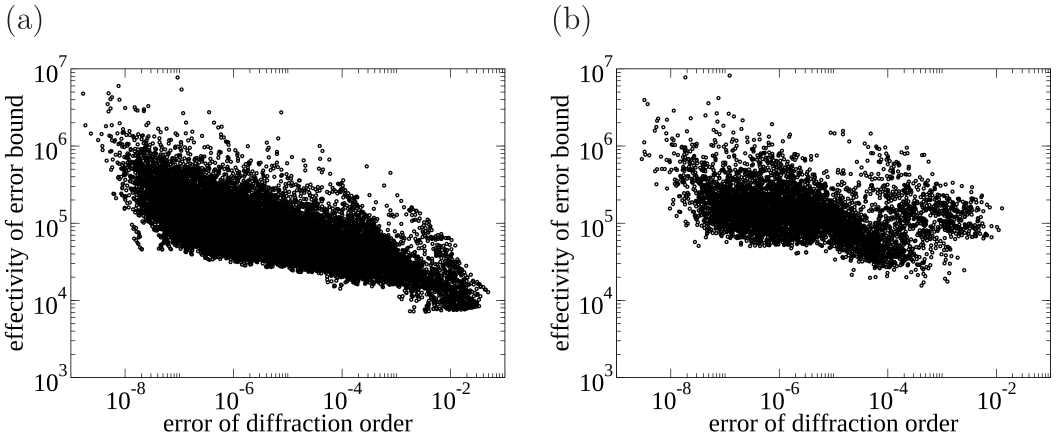


Figure 6.8: Effectivities of error estimates for output of interest in dependence on the error of the output of interest (a) using the primal-only reduced basis and (b) with dual correction. The dimension of the dual reduced basis was set equal to the dimension of the primal reduced basis. (2D mask, Section 6.2)

computational times	primal only ($N = 125$)	primal-dual ($N = 28$)
output of interest	0.08 s	0.05 s
error estimation	1.5 s	0.4 s

Table 6.1: Comparison of computational times for primal and primal-dual reduced basis computation for comparable error bound. Truth approximation computational time is 40 s.

both cases are comparable but quite high mostly between 10^4 and 10^6.

Solution of the full FEM problem takes about 40 s. Reduced basis computational times are given in Table 6.1. The reduced basis computational times are about 500 times smaller than for the truth approximation.

Multiple sources

In Section 5.10 we developed a technique for efficient treatment of problems with multiple sources in the reduced basis context. In order to quantify the performance of this method, we consider the 2D phase shift mask example with three different incoming plane waves as sources. These are defined by

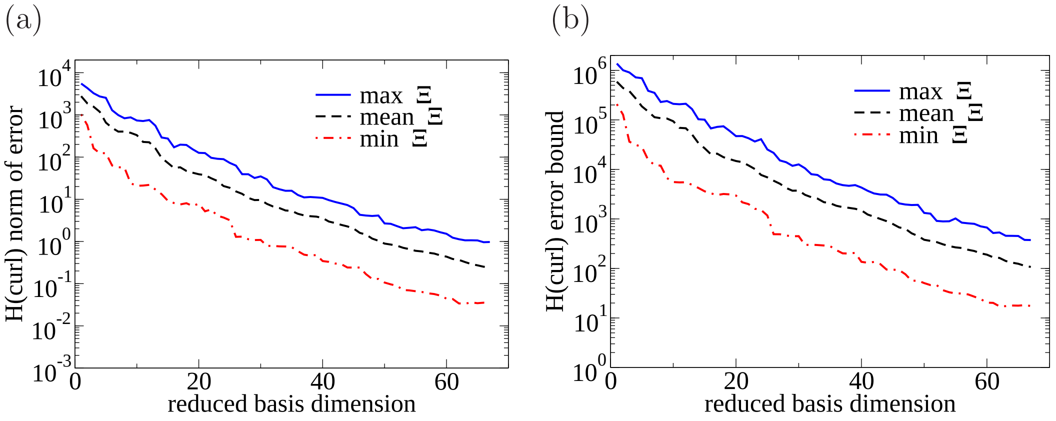


Figure 6.9: (a) Reduced basis solution error (4.8) in $\mathrm{H}\left(\mathbf{curl},\Omega\right)$-norm and (b) corresponding error bound (4.23) in dependence on reduced basis dimension. The mean, minimum and maximum values over the random parameter ensemble Ξ are shown. (2D mask with multiple sources, Section 6.2)

the following electric field and wave vectors:

$$\begin{aligned}
\mathbf{k}_0^1 &= \left(0, 3.25\cdot 10^7, 0\right)\,, & \mathbf{E}_0^1 &= \left(0,0,1\right),\\
\mathbf{k}_0^2 &= \left(0, 3.25\cdot 10^7, 0\right)\,, & \mathbf{E}_0^2 &= \left(1,0,0\right),\\
\mathbf{k}_0^3 &= \left(4.13\cdot 10^6, 3.23\cdot 10^7, 0\right)\,, & \mathbf{E}_0^3 &= \left(0,0,1\right),
\end{aligned}$$

i.e., we consider different polarization and incidence angle. As explained in Section 5.10, we build a reduced basis approximation only for source field 1. The corresponding reduced basis solution defines the coefficients $\overline{\alpha}(\nu)$, which are also used to compute the outputs of interest for source fields 2 and 3, according to Eq. (5.93). We compare the reduced basis approximation to the truth approximation for an ensemble Ξ of 100 random points in parameter space. For each source we computed two outputs of interest, corresponding to the 0th, and 1st diffraction order. First we look at the convergence of the reduced basis solution for source 1 in $\mathrm{H}\left(\mathbf{curl},\Omega\right)$-norm in Fig. 6.9. The error as well as the error bound converge exponentially with increasing reduced basis dimension. Hence, the reduced basis approximation for source 1 provides accurate results. Figure 6.10 shows the convergence of the mean error of outputs of interest for all sources over the random parameter ensemble. For sources 1 and 3 the output corresponds to the z-component of the diffraction order (TE polarization) and to the x-component for source 2 (TM polarization) - note that the y-component can be computed from the x-component using the

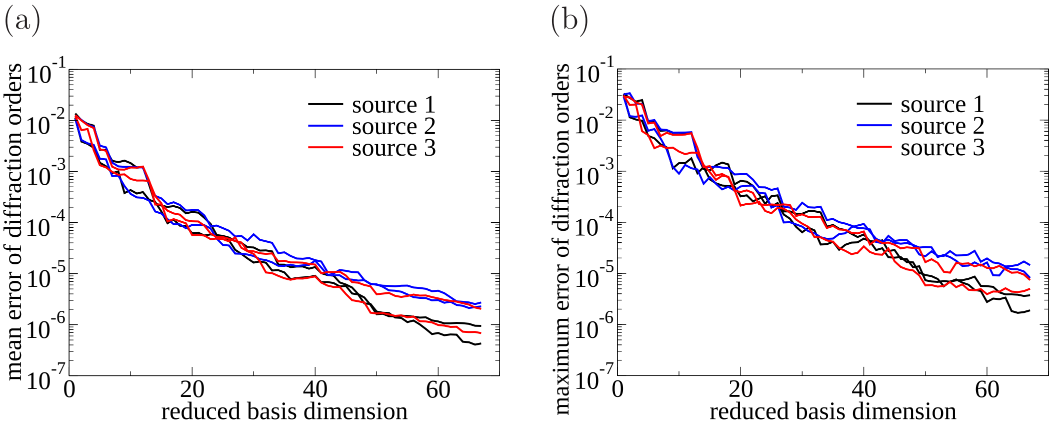


Figure 6.10: Convergence of reduced basis output of interest for multiple sources in dependence on reduced basis dimension. (a) Mean error and (b) maximum error over an ensemble Ξ of 100 random points in parameter space is shown. For each source, two independent outputs of interest were computed. (2D mask with multiple sources, Section 6.2)

divergence condition for the electric field $\mathbf{div}\ E = 0$. We observe that all outputs converge exponentially and basically with the same rate. The output of interest for source 1 converges slightly faster, which is not surprising, since the reduced basis system was constructed for this source. The numerical results show that our technique can be used for efficient construction of reduced basis approximations for systems with multiple sources. The online computational time is, thereby, independent on the number of sources.

6.3 3D periodic grating

Our next example is a 3D periodic grating structure, as shown in Fig. 6.11. Variable geometrical parameters width w, height h, and length l of the grating are chosen for construction of the reduced basis. The parameter domain is $D = [550\,\text{nm}; 650\,\text{nm}] \times [150\,\text{nm}; 250\,\text{nm}] \times [75\,\text{nm}; 125\,\text{nm}]$, i.e.:

$$\begin{aligned} w &\in [550\,\text{nm}; 650\,\text{nm}], \\ h &\in [150\,\text{nm}; 250\,\text{nm}], \\ l &\in [75\,\text{nm}; 125\,\text{nm}]. \end{aligned} \tag{6.7}$$

The dimensions of the computational domain are $800\,\text{nm} \times 400\,\text{nm} \times 200\,\text{nm}$, and it is periodified in x- and y-direction. The wavelength of the incoming

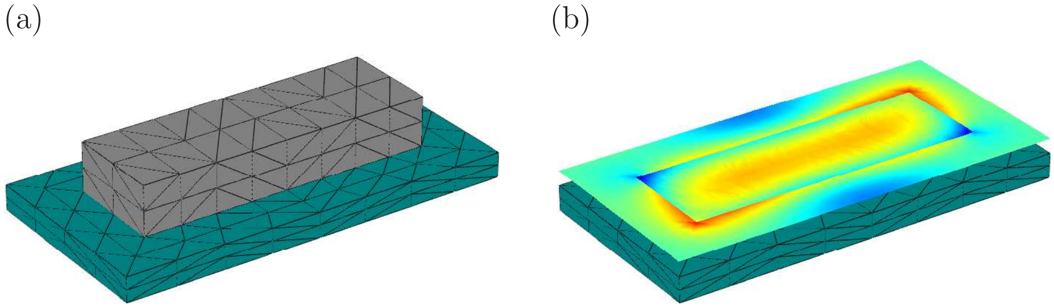


Figure 6.11: (a) 3D periodic grating for reduced basis computations. Parameters are height, width and length of the grating (depicted in grey). (b) Intensity of electric field obtained from FEM computation.

light is $\lambda = 532\,\mathrm{nm}$ with normal incidence from above. FEM discretization gives a system with $\mathcal{N} = 89820$ unknowns for the truth approximation. The affine decomposition of the system matrix gives $Q = 1148$ terms. Thereby meta cells of the geometry, which undergo the same affine mapping, were merged together, in order to reduce Q, e.g., all tetrahedrons within the bar. Looking at the rather simple parametrized geometry with a rectangular bar of variable length, width, and height, this number seems quite high. If one would allow variable grid points on the outer boundary of the computational domain, this simple example could be decomposed into a relatively small number of bricks with variable length, width, and height. The Jacobians of the transformation of each brick would, furthermore, only have entries on the diagonal and no "shear" terms – c.f. [82], where a 2D rectangle of variable length and width is considered. This would reduce the number of affine terms. However, for formulation of the scattering problem, variable grid points on the boundary lead to difficulties: for construction of the right hand side functional $f[u_{\mathrm{in}}](\cdot)$ (3.7), the incoming field has to be interpolated on the boundary, which leads to a non-affine parameter dependence of the system right hand side for variable points on the boundary. Therefore, we keep boundary points fixed, which leads to a number of sheared meta cells also in the substrate below the bar, c.f. Figure 6.11, and in the region filled with air above the bar, which is not depicted. Especially these sheared tetrahedrons in 3D seldom undergo the same affine mapping and contribute to Q.

A primal-only and primal-dual reduced basis approximation with respective dimensions of 99 and 29 was built. Again we compare the reduced basis approximation to the exact FEM solution over a random parameter ensemble Ξ of 200 points.

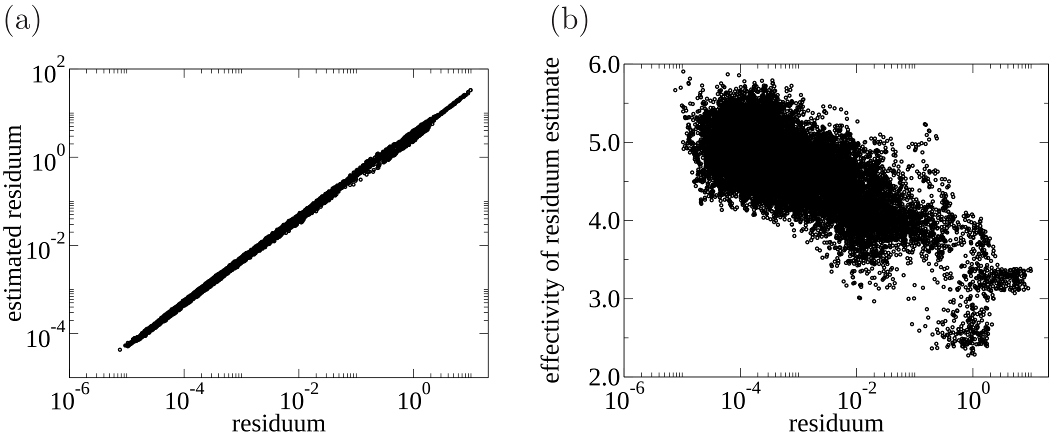


Figure 6.12: (a) Estimated dual norm of residuum (5.74) in dependence on exact dual norm of residuum. (b) Effectivity of residuum estimate (5.75) in dependence on true residuum. (3D mask, Section 6.3)

computational times	primal only ($N = 99$)	primal-dual ($N = 29$)
output of interest	0.3 s	0.4 s
error estimation	5 s	2.3 s

Table 6.2: Comparison of computational times for primal and primal-dual reduced basis computation for comparable error bound. Truth approximation computational time is 225 s.

We start quantifying the performance of the sub-domain residuum method. Figure 6.12(a) shows a scatter plot of the estimated residuum in dependence on the true residuum for 19800 points. The corresponding effectivity of the residuum estimate (5.75) is given in Fig. 6.12(b). Like for the 2D case, we have small and uniform effectivities between 2.0 and 6.0 for the sub-domain residuum method over 6 orders of magnitude for the true residuum, which again demonstrates the good performance of method.

The lower bound for the inf-sup constant determined during construction of the reduced basis was $\beta_0 = 8.4930 \cdot 10^{-3}$. Over the random parameter ensemble it varied between $8.5060 \cdot 10^{-3}$ and $8.8720 \cdot 10^{-3}$. This gives a maximum relative error of $\approx 4 \cdot 10^{-2}$. For the error and error bound of the reduced basis solution in $\mathrm{H}(\mathbf{curl}, \Omega)$-norm, we observe exponential convergence, Fig. 6.13(a), (b). The effectivity of the error bound shown in Fig. 6.13(c) stays between 200 and 700.

As an example for the output of interest we choose one component of the -1st reflected diffraction order. The output of interest convergences exponentially, and again the dual correction leads to much faster convergence,

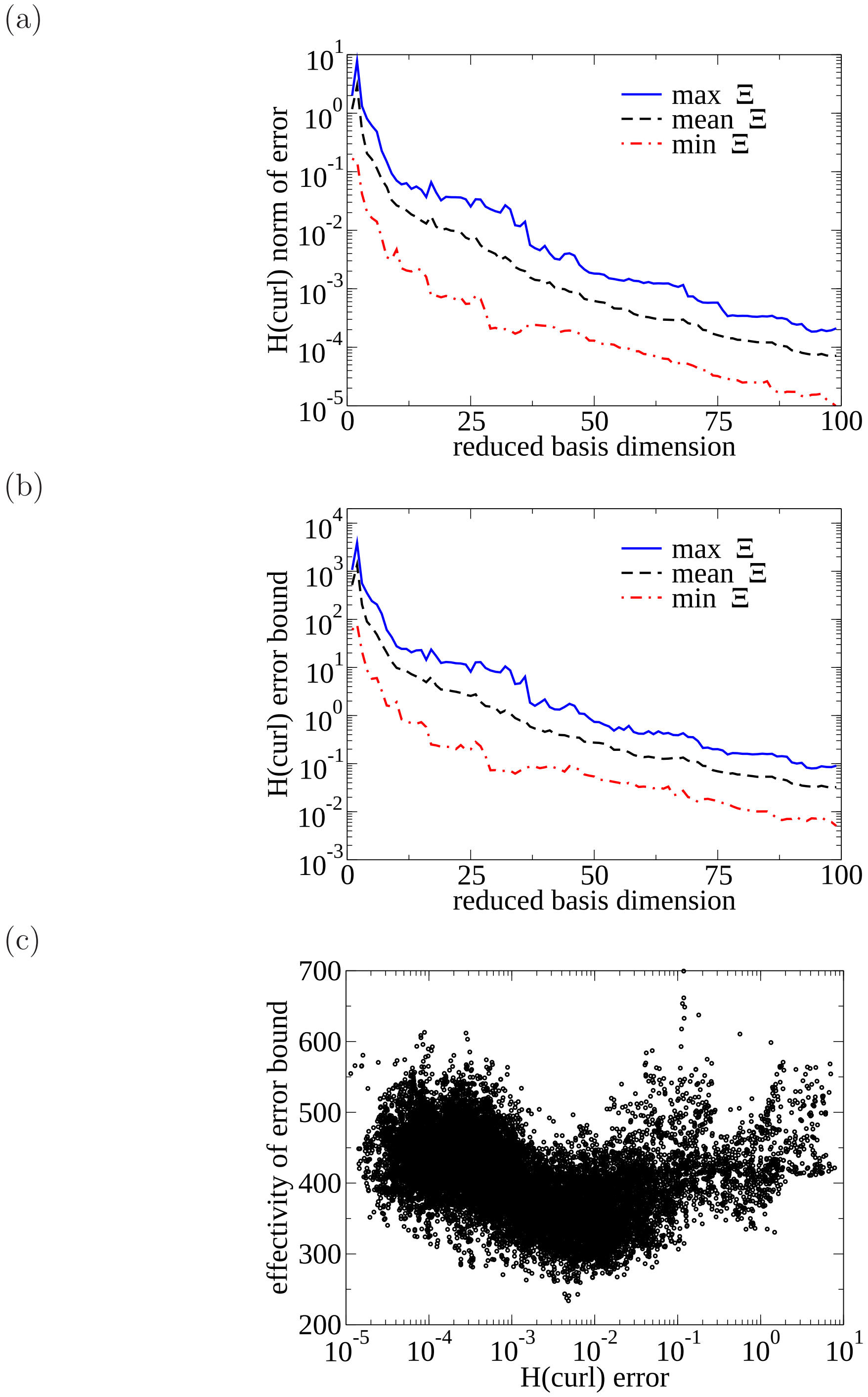


Figure 6.13: (a) Reduced basis solution error (4.8) in $\mathrm{H}(\mathbf{curl}, \Omega)$-norm and (b) corresponding error bound (4.23) in dependence on reduced basis dimension. The mean, minimum and maximum values over the random parameter ensemble Ξ are shown. (c) Effectivity of $\mathrm{H}(\mathbf{curl}, \Omega)$-estimator (5.47a) in dependence on reduced basis solution error. (3D mask, Section 6.3)

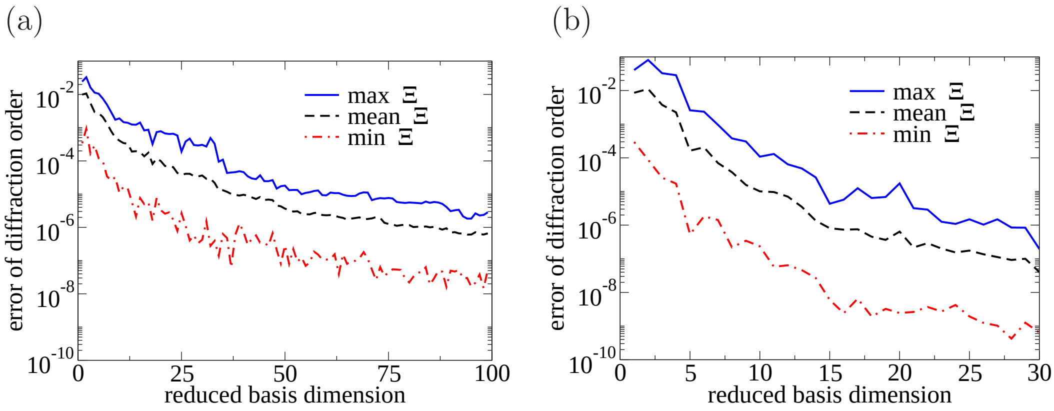


Figure 6.14: Convergence of reduced basis output of interest in dependence on reduced basis dimension (a) using only the primal reduced basis and (b) with dual correction. The dimension of the dual reduced basis was set equal to the dimension of the primal reduced basis. The dual corrected solution converges much faster, compare the different scalings of the x-axes. (3D mask, Section 6.3)

see Figures 6.14 and 6.15. The effectivities for the error of the output of interest are quite high, mostly between 10^4 and 10^6 like in the 2D case, as shown in Fig. 6.16.

The computational time for the truth approximation is about 225 s. The computational times of the reduced basis approximation are given in Table 6.2. Although the dimension of the dual reduced basis is smaller, we have a slightly larger computational time for the output, since two separate reduced basis systems have to be assembled and solved. In summary the reduced basis computational times are about 500 times smaller than for the truth approximation.

6.4 Inverse scatterometry

Inverse scatterometry is a metrology technique, which deduces properties of a system by analyzing fields or particles, which are scattered from the system. Here we focus on optical scatterometry, where the system under investigation is illuminated by a light source. The measured properties, which are of interest, are often geometrical or material parameters ν. In general the scattered light carries no direct information about the system under consideration, hence, an inverse problem has to be solved: given the measured response of a system, determine the parameters ν of the system, such that

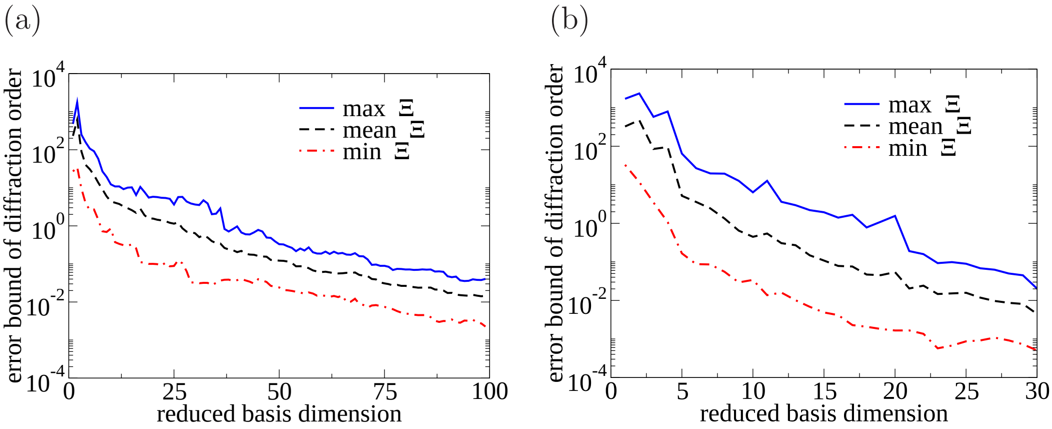


Figure 6.15: Convergence of reduced basis error bound for output of interest in dependence on reduced basis dimension (a) using only the primal reduced basis and (b) with dual correction. The dimension of the dual reduced basis was set equal to the dimension of the primal reduced basis. The dual corrected solution converges much faster, compare the different scalings of the x-axes. (3D mask, Section 6.3)

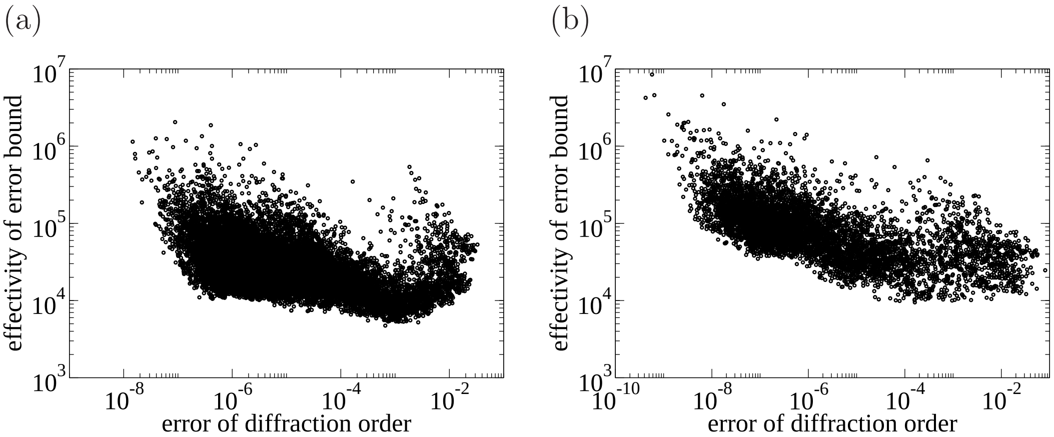


Figure 6.16: Effectivities of error estimates for output of interest in dependence on the error of the output of interest (a) using only the primal reduced basis and (b) with dual correction. The dimension of the dual reduced basis was set equal to the dimension of the primal reduced basis. (3D mask, Section 6.3)

the response is reproduced.

Solution of the inverse problem usually requires a large number of solutions of the forward problem. In this many-query context, application of the reduced basis method leads to significant savings in computational time.

6.4.1 Mathematical formulation

The inverse problem is usually solved by simulating the system under investigation and trying to find good agreement to measured data. Since experimental measurements always carry errors, and also the model of a real world system is never exact, the measured response will not be reproduced exactly. Hence, the inverse problem is transformed into a minimization problem: given the response of a system, determine the parameters ν which minimize the error between measured and simulated response in some sense.

In order to state the minimization problem, we first define the response of a system. Since it will correspond to a set of m experimentally measured values, we will define it as an element of $\mathbb{R}^m$:

$$\begin{aligned} s_{\text{exp}} \in \mathbb{R}^m : &\quad \text{measured (experimental) response,} \\ s_{\text{sim}} \in \mathbb{R}^m : &\quad \text{simulated response.} \end{aligned}$$

In order to quantify the error between measurement and simulation, a metric d on $\mathbb{R}^m$ has to be chosen:

$$d(s_{\text{exp}}, s_{\text{sim}}(\nu)).$$

The function $d(s_{\text{exp}}, \cdot)$ then defines a cost functional, which has to be minimized [2]. The inverse scattering problem can be stated as follows:

Problem 17. Given the experimental response $s_{\text{exp}} \in \mathbb{R}^m$ of a system, find $\nu_{\min} \in D$ such that:

$$\nu_{\min} = \min_{\nu \in D} d(s_{\text{exp}}, s_{\text{sim}}(\nu)), \tag{6.8}$$

where the simulated response

$$s_{\text{sim}}(\nu) = (s_1(\nu), \ldots, s_m(\nu)) \,, \tag{6.9}$$

is given as the outputs of interest

$$s_i(\nu) = l_i^{\text{o}} \left(u(\nu)\right) \,, \quad i = 1, \ldots, m \,, \tag{6.10}$$

to the solution of the following problem:
Find $u(\nu) \in X$ such that:

$$a \left(u(\nu), v; \nu\right) = f \left(v\right) , \quad \forall v \in X. \tag{6.11}$$

We notice the familiar input-output relationship, which usually has to be evaluated several times, solving the inverse problem. Instead of computing the truth approximation, a reduced basis approximation can be used.

The parameter domain D of the minimization Problem (6.8) has to be specified further. In the optical setting ν describes the permittivity distribution (we assume constant permeability) in the computational domain. All possible distributions define the so-called set of admissible geometries. In order to arrive at a well-posed minimization problem, we have to restrict the set of admissible geometries; see e.g. [6, 5] for theoretical results on the inverse reconstruction of grating profiles. In our example we will define the class of admissible geometries by a finite number of geometrical parameters. Therewith, we arrive at a well-posed finite dimensional optimization problem, and the parameter domain D fits into the reduced basis setting.

As optimization algorithm we use the Gauß-Newton method [2]. Therefore, we need derivative information, according to:

$$\partial_\nu d(s_{\text{exp}}, s_{\text{sim}}(\nu)),$$

which describes the change of the cost functional with respect to the input parameters. Hence, the derivatives of reduced basis outputs of interest with respect to the input parameters have to be computed:

$$\partial_\nu s_N(\nu) = \partial_\nu l^{\text{o}}\left(u_N\right) = l^{\text{o}}\left(\partial_\nu u_N\right).$$

Here $\partial_\nu u_N$ is the Frechet-derivative of the reduced basis solution. Using the expansion of u_N into the reduced basis (5.12) and linearity of the output functional $l^{\text{o}}(\cdot)$, above expression can be evaluated by:

$$\partial_\nu s_N(\nu) = \sum_{q=1}^{N} \left(\partial_\nu \alpha_q(\nu)\right) l^{\text{o}}\left(\zeta_q^{\mathcal{N}}\right), \tag{6.12}$$

where the quantities $l^{\text{o}}\left(\zeta_q^{\mathcal{N}}\right)$ are already available for computation of the output of interest. The derivative of the reduced basis coefficients with respect to ν are obtained as follows. The reduced basis system gives a linear system of equations for the reduced basis coefficient vector $\alpha(\nu) = (\alpha_1(\nu), \ldots, \alpha_N(\nu))$:

$$A_N(\nu)\alpha(\nu) = f,$$

where $A_N(\nu)$ refers to the parameter dependent reduced basis system matrix and f to the right hand side vector. Taking the derivative with respect to ν gives:

$$\left(\partial_\nu A_N(\nu)\right)\alpha(\nu) + A_N(\nu)\partial_\nu \alpha(\nu) = 0,$$

from which we conclude:

$$\partial_\nu \alpha(\nu) = -A_N(\nu)^{-1} \left(\partial_\nu A_N(\nu)\right) \alpha(\nu).$$

Once $\alpha(\nu)$ has been computed for the primal solution, the derivative is available at low costs, because the LU decomposition of $A_N(\nu)$ can be reused. The assembling of $\partial_\nu A_N(\nu)$ offers no additional difficulties, since only the derivatives of the known parameter dependent functions in the affine decomposition with respect to the parameters have to be computed:

$$\partial_\nu A_N(\nu) = \sum_{m=1}^{Q} \left(\partial_\nu \Theta_m(\nu)\right) A_N^m.$$

6.4.2 EUV mask reconstruction

In the following we apply the reduced basis technique to an inverse scatterometry application, namely the reconstruction of the pattern profile of so-called extreme ultraviolet (EUV) masks [60, 65]. We analyze experimental data, obtained from scatterometric measurements, carried out by the German national standards and metrology institute (Physikalisch-Technische Bundesanstalt, PTB) at the electron synchrotron BESSY II [98, 97]. The experimental data was taken from an EUV test mask, fabricated by the Advanced Mask Technology Center (AMTC, Dresden). In addition to the scatterometric measurements, direct atomic force (AFM) and scanning electron (SEM) microscopic measurements of the pattern profiles were performed by AMTC. The reconstructed results can, therefore, be compared and validated.

In the following we give a short introduction to EUV technology and describe the experimental setup. Then we define the truth approximation and construct the reduced basis approximation, which will be used for solution of the inverse scatterometric problem.

EUV lithography and metrology

Due to constant miniaturization of integrated circuits, the wavelength of light, which is used in the lithographic process, also has to decrease [44]. EUV lithography is a possible candidate for production of future generation computer technology. Figure 6.17 shows the principle setup. Due to the short wavelength of EUV light ($\approx 13\,\text{nm}$), novel reflective masks and optical systems have to be used in the production process, since there are no appropriate transparent materials in the EUV range.

The quality of pattern profiles of EUV masks becomes important, e.g., due to shadowing effects at oblique incidence illumination [94, 95]. Consequently,

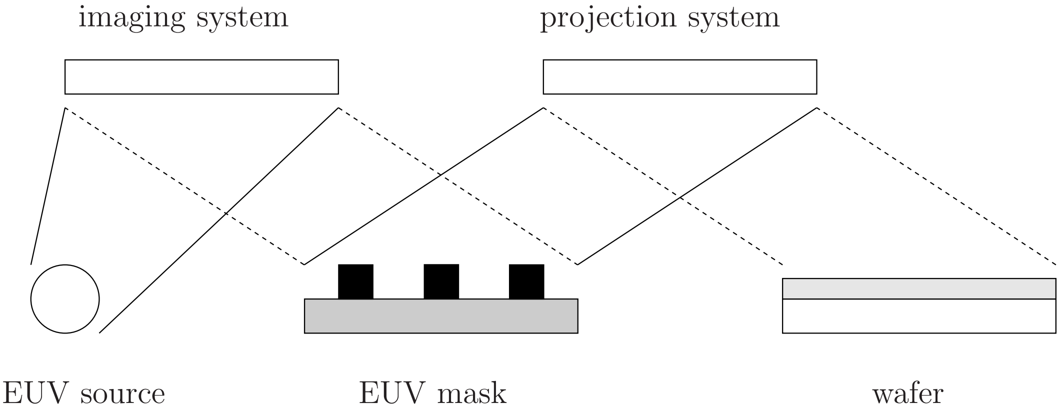


Figure 6.17: Setup of EUV lithography. An EUV mask is illuminated by EUV light, passing through a reflecting imaging system. The light distribution created by the mask absorber pattern is projected onto a wafer, which is covered with photosensitive resist.

there is a need for adequate destruction free pattern profile metrology techniques, allowing characterization of mask features down to a typical size of 100 nm and below [104]. The desired accuracy is, thereby, of the order of 1 nm, which is challenging even for direct atomic force (AFM) or scanning electron microscopic (SEM) measurements. These methods, e.g., suffer from surface charges and the need of the definition of edge operators, to reconstruct the mask topology. Furthermore, it is extremely time consuming and not practical, to scan large areas of a mask with direct microscopical measurements. Therefore, in the following we analyze capabilities of mask metrology by inverse EUV scatterometry. The reduced basis method will, thereby, reduce the reconstruction time significantly.

EUV mask

The principle assembly of an EUV mask is depicted in Fig. 6.18. In contrast to photomasks used nowadays in lithography, EUV masks are not transparent, but reflect the incident light. Thereby, an alternating Bragg multilayer stack, consisting of Molybdenum and Silicon, acts as mirror [89]. The Bragg mirror is covered with a resist pattern, corresponding to the desired image on the wafer.

Experimental data for inverse scatterometry was taken from an EUV test mask, fabricated by AMTC. The mask is subdivided into 11 (labeled A-K) times 11 (labeled 1-11) test fields.

Each of these test fields, furthermore, consists of a number of differently

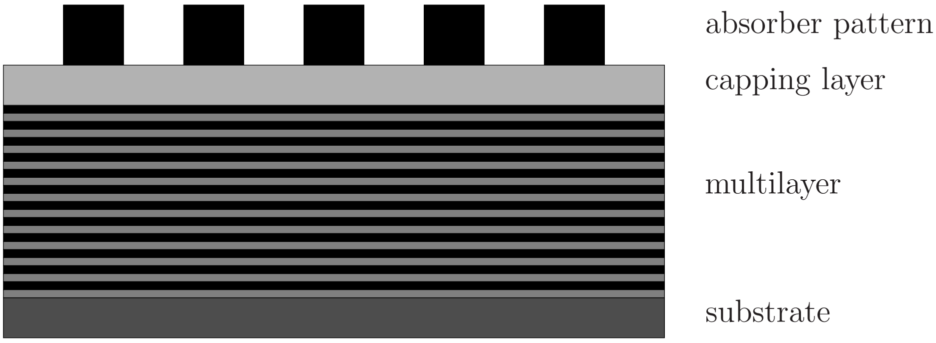


Figure 6.18: Principle assembly of an EUV mask. A multilayer acting as Bragg mirror for EUV radiation is deposited on a substrate. The capping layer is partially covered by the absorber pattern, corresponding to the desired reflected image.

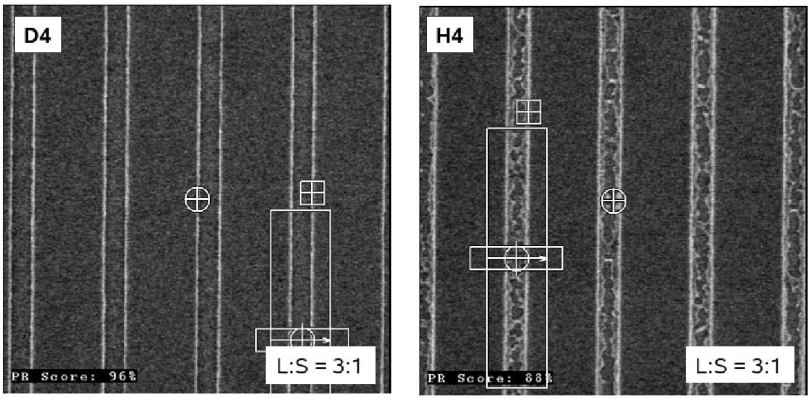


Figure 6.19: Top down SEM image of uniformity fields D4 and H4. It can be seen that the lines of field H4 are not completely etched.

patterned areas, e.g.:

- bright field: unpatterned, i.e., no absorber,
- dark field: completely covered with absorber material,
- CD (critical dimension) uniformity: covered with 1D periodic absorber lines,

where the term critical dimension refers to the width of the absorber lines. We will focus on CD uniformity fields, with design parameters given in Table 6.3. Figure 6.19 shows a top down microscopic image. We consider fields labeled by D4, H4, F6, D8, H8. The actual CD is the most important parameter, which we want to reconstruct.

mask parameters	nominal values
ARC + TaN-absorber thickness	67 nm
SiO_2-buffer thickness	10 nm
Si-capping layer thickness	11 nm
multilayer	Mo/Si
pitch	720 nm
CD	540 nm

Table 6.3: Design parameters of EUV test mask used for inverse scatterometry

EUV scatterometry

Single wavelength scatterometry, the analysis of light diffracted from a periodic structure, is a well suited tool for analysis of the geometry of EUV masks [91]. Since scatterometry only needs a light source and a simple detector with no imaging lens system, its setup is inexpensive and offers no additional technical challenges. Figure 6.20(a) shows a sketch of the experimental setup. Light of fixed wavelength and fixed incident angle is reflected from the mask, and the intensity of the reflected light is measured in dependence on the diffraction angle. Usage of EUV light for mask characterization is advantageous, because it fits the small feature sizes on EUV masks. Diffraction phenomena are minimized, and of course the appropriate wavelength of the resonant structure of the underlying multilayer is chosen. Light is not only reflected at the top interface of the mask, but all layers in the stack contribute to reflection. Therefore, one expects that EUV radiation provides much more information on relevant EUV mask features than conventional long wavelength methods [90]. The presented EUV measurements often provide up to 20 or more non-evanescent diffraction orders.

In experimental measurements the incidence angle was fixed at $\theta_{\text{in}} = 6°$, and the electric field was TE polarized. Figure 6.20(b) shows the intensity of a set of experimentally determined diffraction orders of the EUV mask for a CD uniformity field. These diffraction orders define the experimental response s_{exp}. We want to determine the absorber profile numerically, whose simulated diffraction orders, i.e., simulated response s_{sim}, fits the experimental best.

Since EUV light is reflected from the mask multilayer mirror, the intensities of the diffraction orders depend not only on the absorber profile but also on the underlying multilayer [64, 104]. This is demonstrated experimentally in Fig. 6.21(b). The wavelength dependence of the plain multilayer reflectivity is shown in comparison to the intensity of several diffraction orders. The

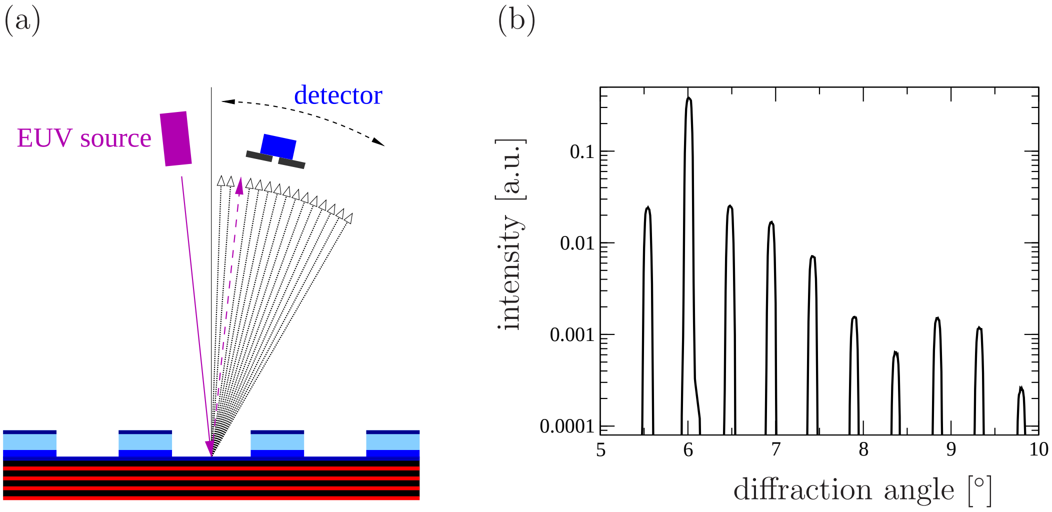


Figure 6.20: (a) Experimental setup for EUV scatterometry with fixed incidence angle θ_{in} and variable angle of detection θ_{out}. (b) Result of single wavelength scatterometry measurement of periodic mask pattern. Diffraction orders appear as peaks with finite width, the zeroth diffraction peak is centered around 6°.

curves resemble each other to a great extend. Hence, accurate modeling of the underlying multilayer is a crucial and also difficult part in inverse EUV scatterometry. Our model of the multilayer was built in advance to absorber profile reconstruction [63, 64]. Thereby, experimental bright field curves were fitted to 1D simulations of the plain multilayer. Figure 6.21(a) shows that measured and fitted multilayer curves agree quite well.

Another difficulty arises, because the multilayer is not homogeneous across the whole mask. Therefore, we also introduce a global multilayer scaling factor γ as a parameter into the reduced basis model, which scales the thicknesses of layers in the Bragg mirror. This parameter controls the relative position of the reflectance curve.

The experimental measurements were performed at three different wavelengths for each test field, corresponding to the left full width half maximum (FWHM) λ_1, center λ_2, and right FWHM wavelength λ_3 of the bright field curve as depicted in Fig. 6.21(a).

The experimental diffraction orders were measured with a relative accuracy of 1%. Noise in the EUV detector was of the order of $5 \cdot 10^{-5}$ relative to the incident light. Experimental diffraction orders with intensity below this limit were, therefore, not used for reconstruction, since their relative error is above 100%.

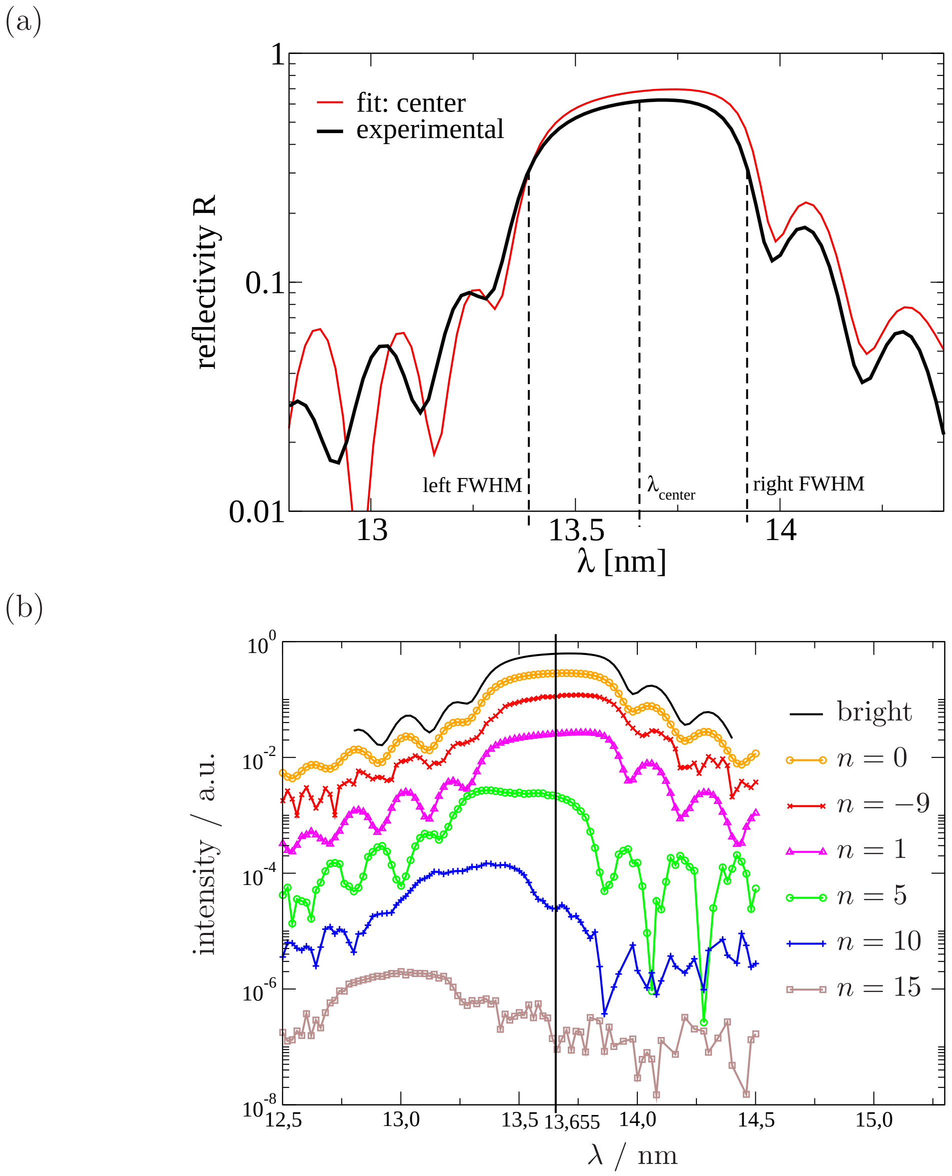


Figure 6.21: (a) Numerical fit of experimental bright field curve. Three wavelengths used for measurement of scatterometric data are shown. (b) Bright field measurement and diffraction orders in dependence on wavelength. Intensity of diffraction orders are scaled for better comparison. The wavelength $\lambda = 13.655\,\mathrm{nm}$ corresponds to the center wavelength of the bright field curve.

In [23] the uncertainty of the reconstruction results caused by measurement errors of the diffraction orders is analyzed and quantified. However, these uncertainties are smaller than those caused by modeling errors of the multilayer [64].

Truth approximation

Figure 6.22 shows the geometry of the EUV mask and a finite element solution for the electric field, which is the truth approximation. In the multilayer we observe a standing wave pattern, created by incoming and reflected light. In order to create the parametrized mask model, we choose 4 geometrical parameters:

- center CD w,
- height h of the absorber line,
- sidewall angle β of absorber stack,
- scaling factor γ for the multilayer.

The parameters are depicted in Fig. 6.23.

The outputs of interest are discrete diffraction orders. For the truth approximation we choose finite element degree $p = 6$. Discretization gives a system with $\mathcal{N} = 3032880$ unknowns. A single evaluation of the input-output relationship of the truth approximation takes about 420 s. For the inverse problem several of these evaluations have to be performed, which leads to reconstruction times in the order of hours, far too long for real-time application.

Reduced basis approximation

The affine decomposition of the mask model gives $Q = 262$ terms. For construction of the reduced basis approximation, we choose the parameter domain D as follows:

$$\begin{aligned} w &\in [530\text{nm}; 570\text{nm}], \\ h &\in [51.9\text{nm}; 57.9\text{nm}], \\ \beta &\in [0°; 10°], \\ \gamma &\in [0.996; 1.004]. \end{aligned} \tag{6.13}$$

Since the wavelength is not included as an input parameter to the model, three different reduced basis approximations are built, corresponding to the experimental wavelengths.

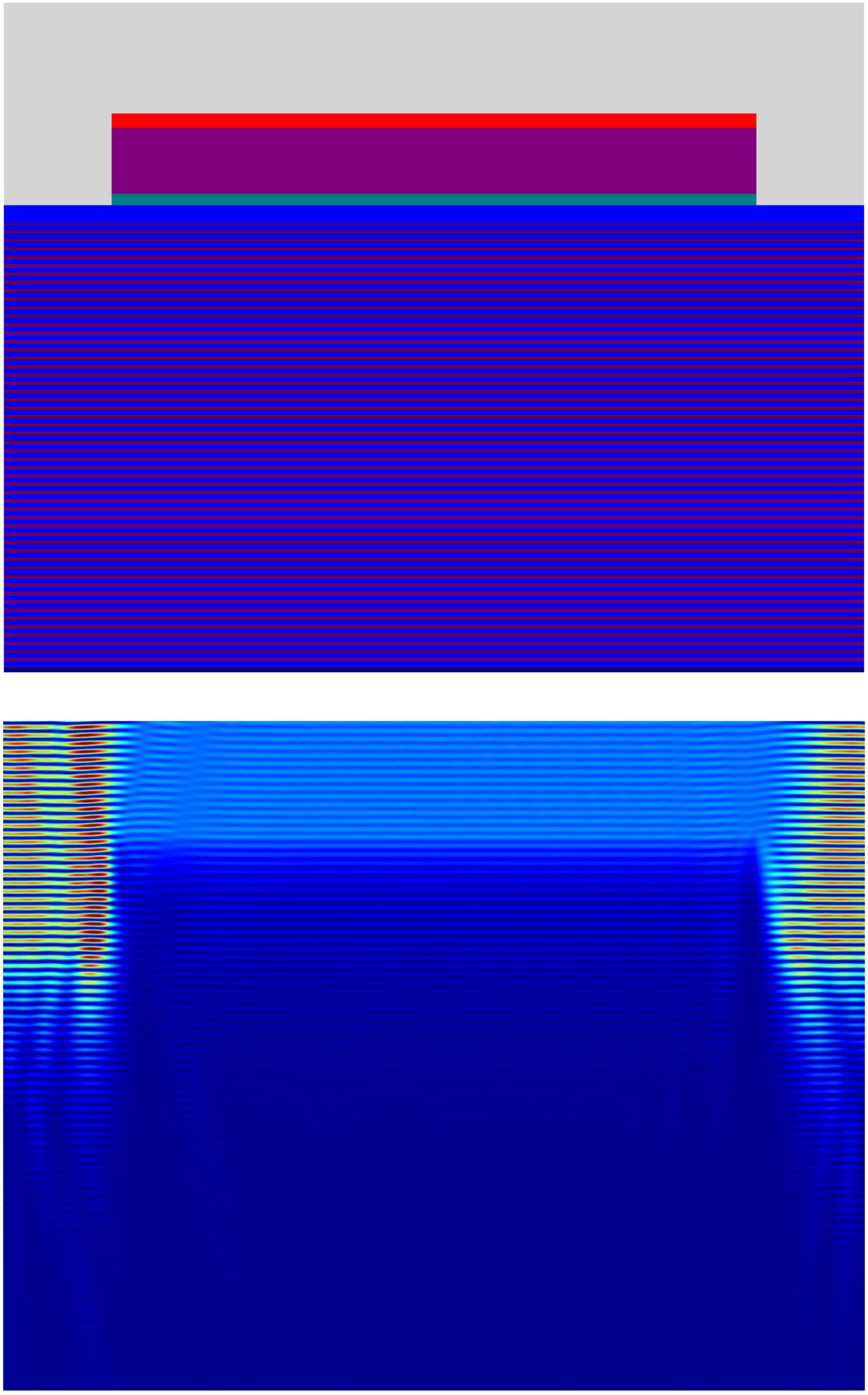


Figure 6.22: Geometry of EUV mask model and electric field intensity of finite element solution. Light is only reflected in regions not covered with absorber.

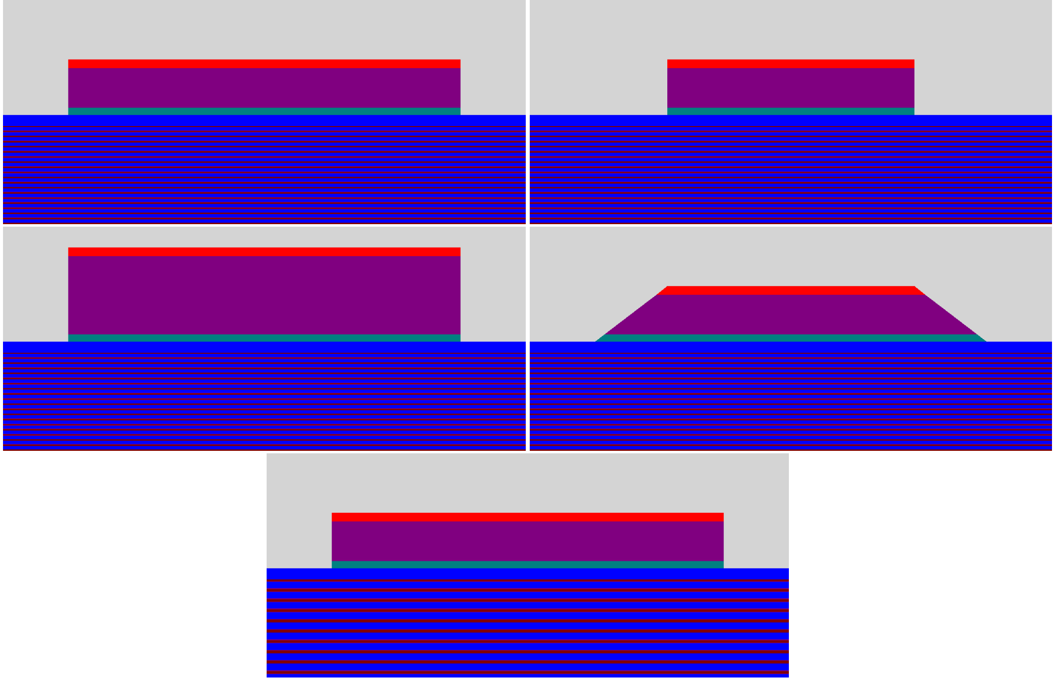


Figure 6.23: Illustration of geometrical parameters of EUV mask model – clockwise, starting at upper left: reference layout, center CD w, absorber stack sidewall angle β, multilayer scaling factor γ, absorber height h.

The final reduced basis dimension is chosen as $N = 80$. Due to the large number of outputs of interest, we construct a primal-only reduced basis approximation. Before solving the inverse problem, we present convergence results for λ_2 as an example. We compare the reduced basis approximation to the exact FEM solution for a random parameter ensemble Ξ of 100 points in the parameter domain.

First we look at the residuum estimator in Fig. 6.24. We observe small effectivities between 3.0 and 6.0 over several orders of magnitude of the true residuum, which again demonstrates the good performance of the sub-domain residuum method. Figure 6.25 shows the convergence of the reduced basis solution in $\mathrm{H}(\mathbf{curl}, \Omega)$-norm and the corresponding error bound. The convergence of a number of outputs of interest, namely the 0th, 5th, and 10th diffraction order, is given in Fig. 6.26 together with the corresponding bound for the 0th order. We observe exponential convergence with increasing dimension of the reduced basis. Again we have moderate effectivities for the error in $\mathrm{H}(\mathbf{curl}, \Omega)$-norm and high effectivities for the output of interest, given in Fig. 6.27.

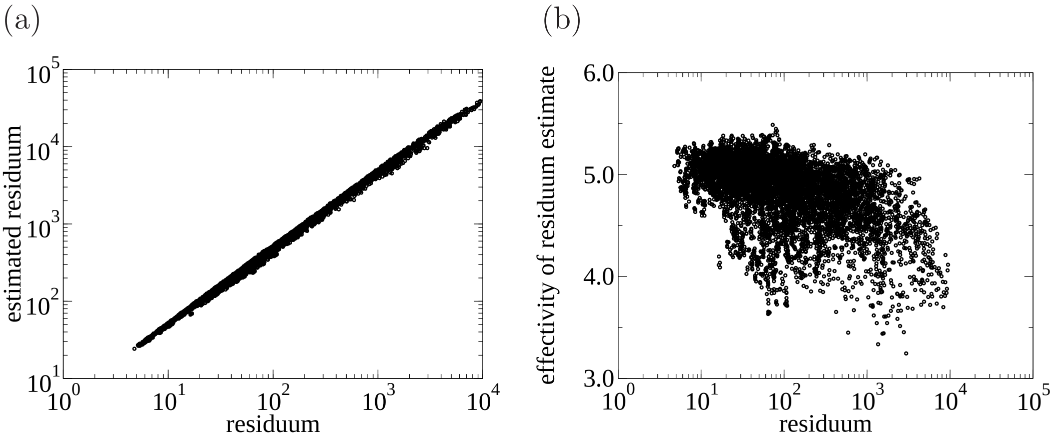


Figure 6.24: (a) Estimated dual norm of residuum (5.74) in dependence on exact dual norm of residuum. (b) Effectivity of residuum estimate (5.75) in dependence on true residuum. (EUV mask)

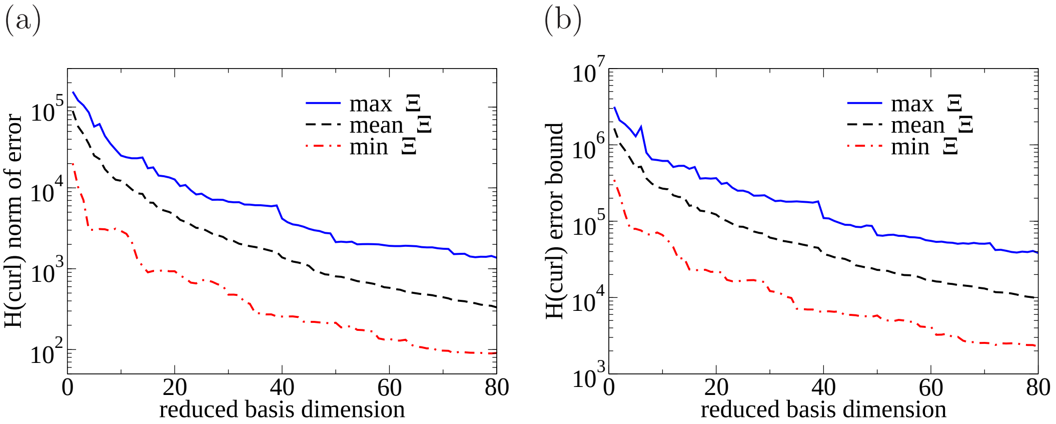


Figure 6.25: (a) Reduced basis solution error (4.8) in $\mathrm{H}(\mathbf{curl}, \Omega)$-norm and (b) corresponding error bound (4.23) in dependence on reduced basis dimension. The mean, minimum and maximum values over the random parameter ensemble Ξ are shown. (EUV mask)

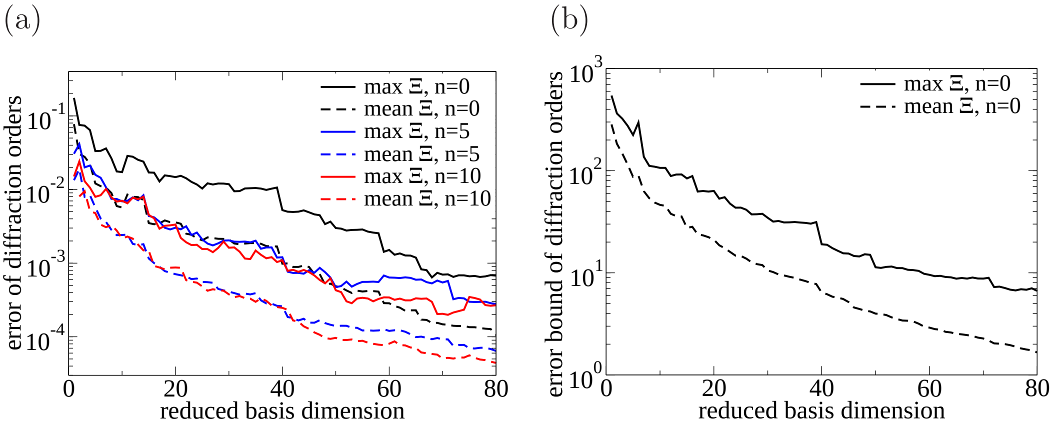


Figure 6.26: (a) Convergence of reduced basis output of interest in dependence on reduced basis dimension. The mean and maximum error over the assemble Ξ is shown for the 0th, 5th, and 10th diffraction orders. (b) Corresponding error bound for the 0th diffraction order. (EUV mask)

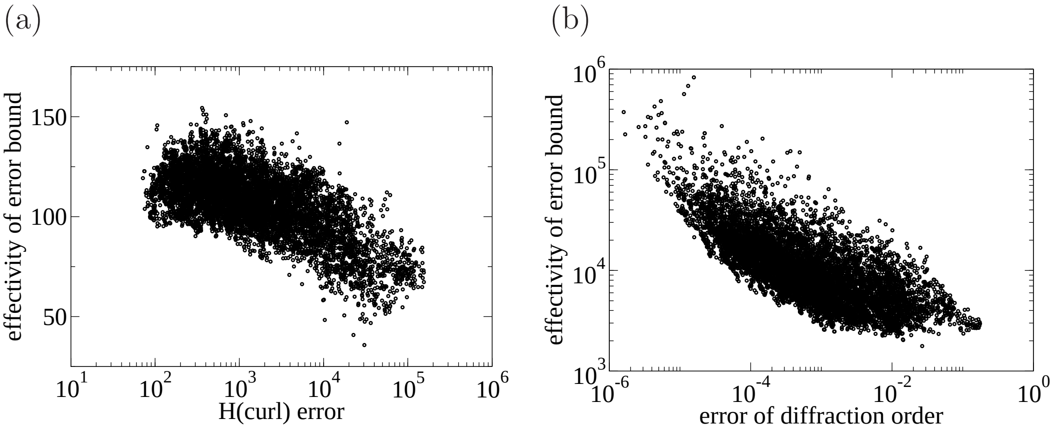


Figure 6.27: (a) Effectivity of H $(\mathbf{curl}, \Omega)$-estimator (5.47a) in dependence on reduced basis solution error. (b) Effectivity of error estimates for 0th diffraction order, c.f. Fig 6.26. (EUV mask)

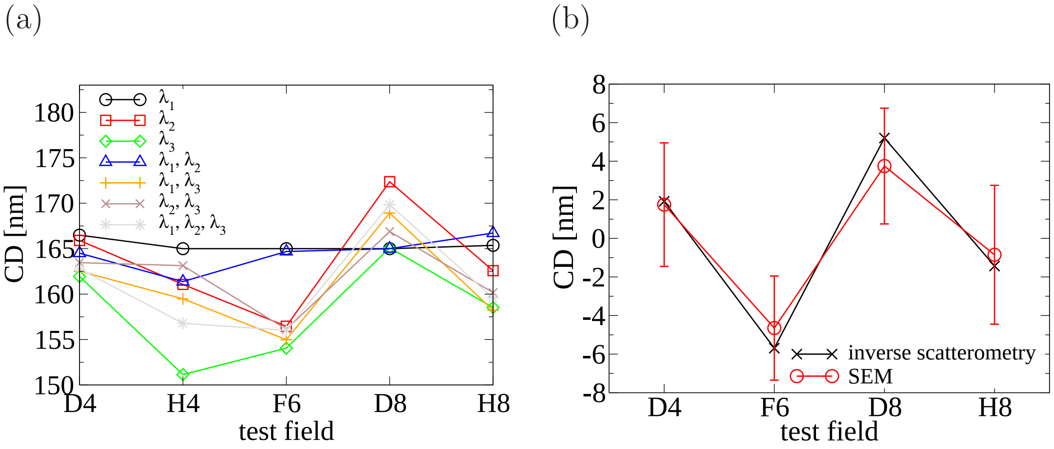


Figure 6.28: (a) Reconstructed CDs from inverse scatterometry using different subsets of experimental data. (b) Comparison of CDs obtained from direct scanning electron microscopy and inverse scatterometry after subtraction of mean values. The inverse scatterometry results correspond to best correlated results given in Table 6.4.

Also for this example, the inf-sup constant does not show a strong dependence on the input parameters. It varies between $2.918 \cdot 10^{-3}$ and $3.527 \cdot 10^{-3}$ over the random parameter ensemble. During construction of the reduced basis we obtained an estimate of $\beta_0 = 3.096 \cdot 10^{-3}$.

The online computational time for solution of the reduced basis problem is 0.15 s, which is about 3000 times faster than the truth approximation.

Reconstruction Results

The measured s_{exp} and simulated diffraction orders s_{sim} are used in the minimization problem (6.8). As metric, measuring their difference, we choose the sum of squared relative errors of m diffraction orders:

$$d(s_{\text{exp}}, s_{\text{sim}}(\nu)) = \sum_{n=1}^{m} \left(\frac{s_{\text{exp}}^n - s_{\text{sim}}^n(\nu)}{s_{\text{exp}}^n} \right)^2 . \tag{6.14}$$

Diffraction orders were measured at different wavelengths λ_1, λ_2, and λ_3 for each of the test fields. Figure 6.28(a) shows reconstruction results for the center CD of the line spacing (= pitch − absorber CD), using different subsets of the experimental data. For example for the curve "λ_1, λ_3", only experimental diffraction orders s_{exp}^n were used in (6.14), which were measured at wavelengths λ_1 and λ_3. Especially field $H4$ shows large deviations for different data subsets. The reason can be seen from a top down image of the absorber

data set	λ_1	λ_2	λ_3	λ_1, λ_2	λ_1, λ_3	λ_2, λ_3	λ_1, λ_2, λ_3
λ_1	1.0000	0.1223	0.2540	-0.2505	0.0754	0.2204	0.0142
λ_2	0.1223	1.0000	0.9909	-0.0665	0.9863	0.9944	0.9872
λ_3	0.2540	0.9909	1.0000	-0.1082	0.9729	**0.9991**	0.9651
λ_1, λ_2	-0.2505	-0.0665	-0.1082	1.0000	-0.2077	-0.1240	-0.1537
λ_1, λ_3	0.0754	0.9863	0.9729	-0.2077	1.0000	0.9818	0.9973
λ_2, λ_3	0.2204	0.9944	0.9991	-0.1240	0.9818	1.0000	0.9751
λ_1, λ_2, λ_3	0.0142	0.9872	0.9651	-0.1537	0.9973	0.9751	1.0000

Table 6.4: Cross correlation of reconstructed CDs of test fields using different experimental data sets. Maximum correlation can be found for reconstruction with wavelength λ_3, and wavelengths λ_2 and λ_3.

test field	CD [nm]	CD uniform. [nm]
D4	186.7	3.2
H4	168.0	7.3
F6	180.3	2.7
D8	188.7	3.0
H8	184.1	3.6

Table 6.5: SEM measurements of line profiles. CD uniformity is defined as 3σ over an ensemble of measurements. The CD is the mean over the ensemble.

lines of the field in Fig. 6.19. The absorber is not etched completely down to the capping layer, and the remaining material results in a large amount of diffusive scattering [90]. This effects the intensity of the measured diffraction orders and therewith the reconstruction results. Therefore, we will not consider field $H4$ in the following analysis.

Also the reconstructed CDs for all other fields given in Fig. 6.28(a) show differences for different subsets of experimental data. How can the “correct” CDs be extracted from these results and compared to the microscopically determined values?

For this task we first compute the cross correlation of the CDs of all fields, which were reconstructed with the different experimental subsets, i.e., the cross correlation between all curves given in Fig. 6.28(a). As shown in Table 6.4, many of the results are highly correlated (> 0.9). Some disagree completely, namely reconstructing with the single data set λ_1, and wavelengths λ_1 and λ_2. In order to obtain an estimate for the correct CD, we could for example average the CDs from two reconstructions, which are correlated best. Here,

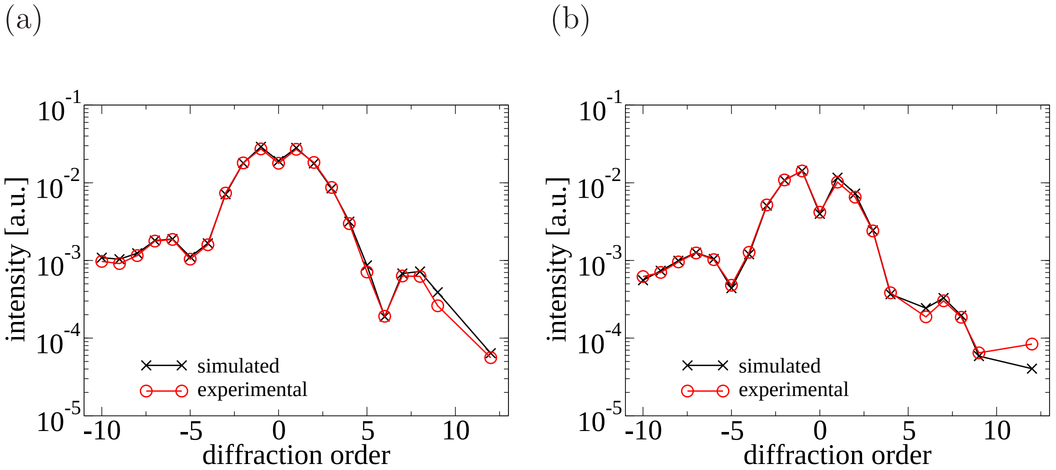


Figure 6.29: Experimental and best fitting simulated diffraction orders for reconstruction on field $D4$ for (a) wavelength λ_2 and (b) wavelength λ_3.

this means reconstruction with wavelength λ_3, and wavelengths λ_2 and λ_3. The result is shown in Fig. 6.28(b) in comparison to the SEM measurements. We subtracted the mean CD from both data sets, since also in the microscopic measurements determination of absolute CD values was difficult, and different methods like atomic force (AFM) or scanning electron microscopy (SEM) gave different absolute values [64]. We see very good agreement below 1 nm between direct microscopic measurements and inverse scatterometry. The uncertainty estimates for the microscopical results were obtained by several SEM measurements at different positions on the same test field.

Finally a comparison between experimental and best fitting diffraction intensities is given in Fig. 6.29 for field $D4$ as an example. We see remarkable agreement for almost all diffraction orders.

6.5 Optical proximity correction

In optical lithography photomasks are used to image desired patterns of integrated circuits onto wafers [44]. The ongoing miniaturization of these circuits drives the semiconductor industry and leads to CPUs with higher performance and memory elements with larger capacity. In order to produce structures with smaller feature size, absorber structures on the photomask also have to decrease in size.

Using optical lithography one is, however, confronted with the fundamental problem that optical systems act as low pass filters: consider for example a 2D

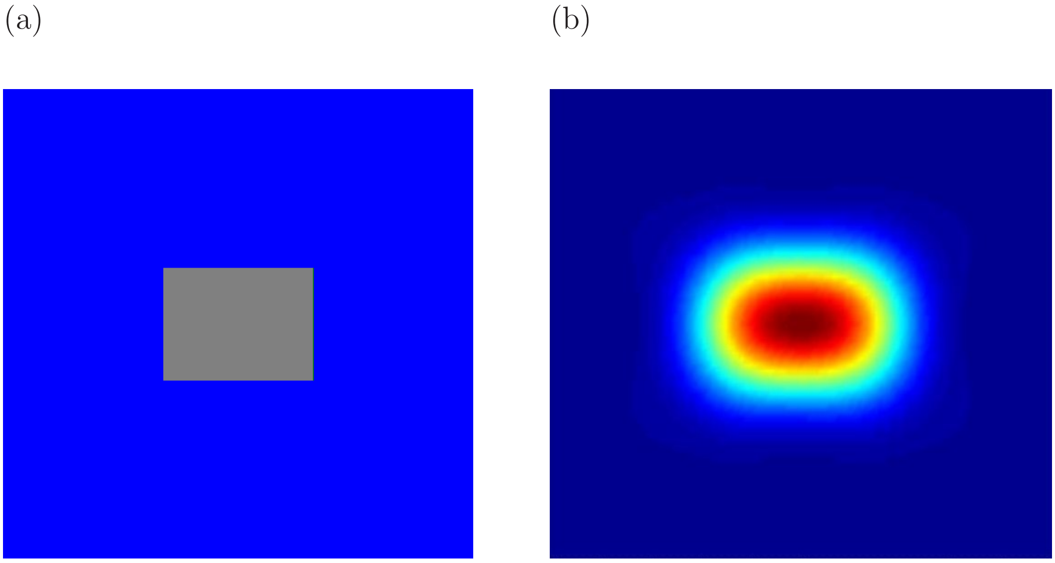


Figure 6.30: (a) Top view of mask layout for imaging of contact hole. Absorber material is shown in blue. (b) Far field image of contact hole shows corner rounding.

domain of width L_x, which is periodified in x-direction and an incoming plane wave under normal incidence with wavelength λ_0. As explained in Section 6.1, the spatial frequencies of propagating modes in the far field are given by:

$$k_{x,l} = l\frac{2\pi}{L_x}, \quad \text{with} \quad l = -\left\lfloor \tfrac{L_x}{\lambda_0} \right\rfloor, \ldots, 0, \ldots, +\left\lfloor \tfrac{L_x}{\lambda_0} \right\rfloor,$$

where we assumed vacuum in the exterior for simplicity, and $\lfloor \cdot \rfloor$ denotes rounding down. With decreasing structure size and therewith decreasing L_x, only low frequencies remain in the far field. The maximum spatial frequency in the far field is given by:

$$k_{x,l_{\max}} = \frac{2\pi}{L_x}\left\lfloor \frac{L_x}{\lambda_0} \right\rfloor \leq \frac{2\pi}{\lambda_0}.$$

Hence, the system acts as a low pass filter.

Figure 6.30 shows the consequence for the image of a rectangular hole. In the far field sharp corners are washed out, which is referred to as corner rounding. Optical proximity correction (OPC) is the adjustment of features on the photomask, in order to compensate this spatial filtering [44]. For the image of contact holes, OPC can be achieved with the introduction of serifs to the corners of the contact hole, as will be shown in the following. Our numerical example will be the optimization of such an OPC structure, in order to obtain a structure on the wafer which is closest to the desired rectangular shape.

6.5.1 Model problem

The geometry of the contact hole under consideration is depicted in Fig. 6.31. The size of the computational domain is $2.5\,\mu\text{m} \times 2.5\,\mu\text{m}$ with a height of 90 nm. The height of the Chromium absorber is 50 nm with a refractive index of:

$$n_{\text{Cr}} = 0.84 - 1.65i,$$

for a wavelength of $\lambda = 193\,\text{nm}$ [44]. The absorber (blue) is deposited on a silica substrate (gray). In x- and y-direction periodic boundary conditions are applied. The shape of serifs is described by 4 input parameters $p_{1,x}$, $p_{2,x}$, $p_{1,y}$, and $p_{2,y}$, defined in Fig. 6.31(b). The dimensions of the contact hole itself are fixed at:

$$\begin{aligned} d_x &= 800\,\text{nm}, \\ d_y &= 600\,\text{nm}. \end{aligned}$$

As incoming light we use so-called conventional illumination [44]. This is modeled by a set of incoming plane waves, whose incoming angles lie within a cone up to a certain maximum angle. In order to speed up computational time and make use of the efficient reduced basis setting with multiple sources, we use incoming plane waves with same wavelength and Bloch periodicity in x- and y-direction. The source can be visualized by a set of points in the k_x-k_y-plane, as shown in Fig. 6.32. For each of these incoming directions two orthogonal polarization states are simulated to mimic unpolarized light. This gives $P = 74$ sources in total.

Let us denote the far field coefficients of the system, obtained from illumination with the r-th source, by A_j^r. For coherent illumination the electric field intensity in the far field is given by:

$$I_{\text{coh}}(\mathbf{x}) = \left| \sum_{r=1}^{P} \sum_{j} A_j^r e^{i\mathbf{k}_j \cdot \mathbf{x}} \right|^2, \tag{6.15}$$

In this case a single FEM simulation is sufficient, where all P incoming plane waves are added to obtain the complex source.

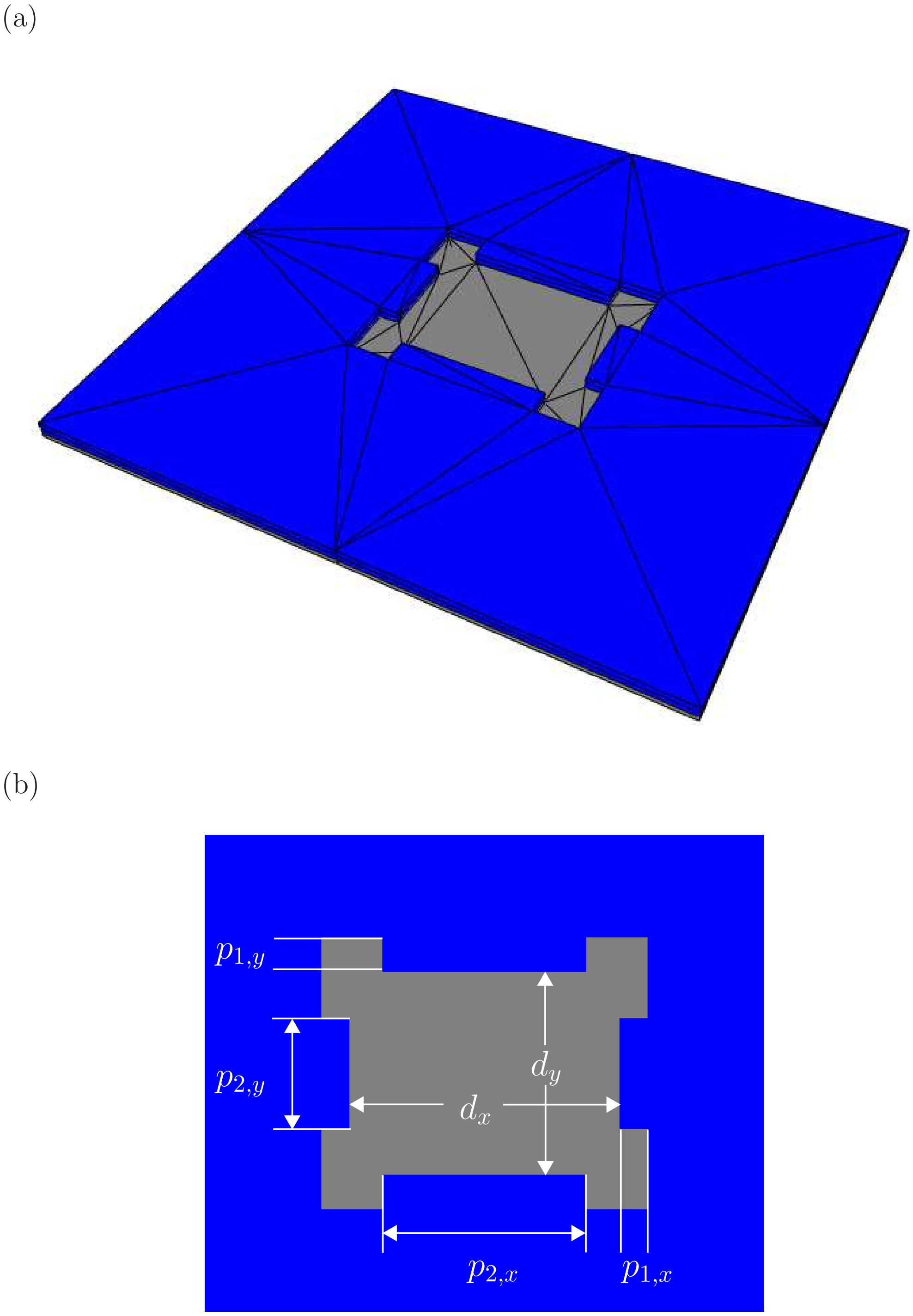


Figure 6.31: (a) 3D mask model used for finite element computation. (b) Definition of OPC parameters used for optimization; d_x and d_y are fixed.

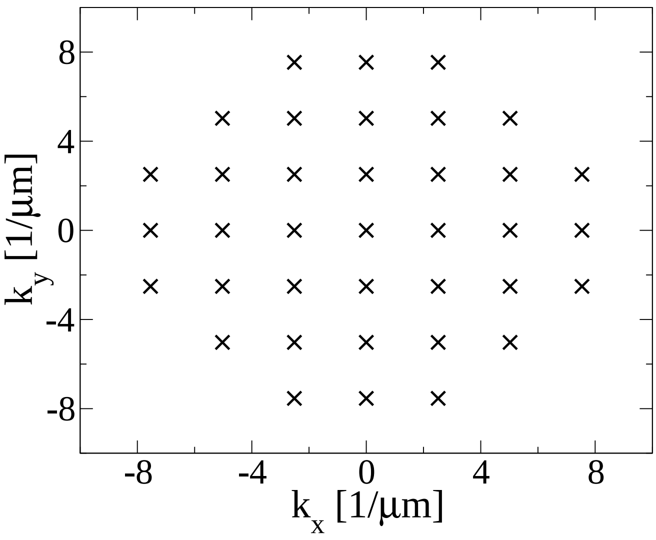


Figure 6.32: Modeling of conventional illumination: a set of plane waves with incoming directions up to a maximum incidence angle is used, i.e., maximum transversal k-vector. Each cross in k-space corresponds to two plane waves with orthogonal polarizations.

The far field for so-called partial coherent illumination, which we are considering, is given by [93, 103]:

$$I_{\text{incoh}}(\mathbf{x}) = \sum_{r=1}^{P} \left| \sum_{j} A_j^r e^{i\mathbf{k}_j \cdot \mathbf{x}} \right|^2 .$$

In order to compute this expression, the far field coefficients of a single source are added coherently, and then the resulting P intensities are added incoherently (taking the absolute value). In this case P near field simulations have to be performed separately, i.e., for each incoming plane wave. In application the partial coherent case is often important. For each of the sources we, therefore, have to compute the near field and corresponding far field coefficients separately. In this example there are 37 far field coefficients with 2 independent polarization states for each sources. This gives a total number of 5476 outputs of interest.

FEM discretization gives a system with $\mathcal{N} = 474720$ unknowns. The computational time for all sources is about $9300\,\mathrm{s} \approx 2.5h$. Hence, optimization of this structure can not be performed in reasonable time, using the truth approximation.

6.5.2 Optimization problem

In the following we describe a simple model, how the fabricated structure on the wafer, which we want to optimize, can be computed from the far field coefficients of the incoming fields.

Behind the photomask we place an optical system with a 4 : 1 reduction ratio, which is common for many wafer steppers [44]. The propagating modes computed from the near fields of all sources pass this system. For simplicity we assume that the optical system is aberration free and, therefore, only the direction of propagation of the modes in the far field is changed. Then we compute the aerial image [44] from the propagating modes after passing the optical system. The aerial image is the intensity distribution of the electric field in the plane where the wafer is located. However, it is computed assuming that only air occupies the space. For simplicity we assume that the wafer is located at $z = 0$, which gives following expression for the aerial image A:

$$A(x,y) = \sum_{r=1}^{P} \left| \sum_{l,m} A^r_{l,m} e^{i\left(k_{x,l}x + k_{y,m}y\right)} \right|^2 , \tag{6.16}$$

where $A_{l,m}$ and $k_{x,l}$ and $k_{y,m}$ are amplitudes and wave vector components of the propagating modes after passing the optical system.

The aerial image determines where the photo sensitive resist on the wafer is developed. Usually one defines a certain threshold σ, and all areas in the aerial image with intensity above this threshold are developed. The shape γ_σ of the structure on the wafer after development of the resist, is then given implicitly by:

$$\gamma_\sigma = \left\{ (x,y) \in \mathbb{R}^2 : A(x,y) = \sigma \right\} , \tag{6.17}$$

hence, γ_σ is the contour line of the aerial image at level σ.

Figure 6.33(a) shows the desired target rectangular structure on the wafer and a contour curve γ_σ, corresponding to the aerial image given in Fig. 6.30(b) with $\sigma = 0.4$. The target rectangle on the wafer has dimensions:

$$\begin{aligned} d_x^{\text{target}} &= 180\,\text{nm}, \\ d_y^{\text{target}} &= 120\,\text{nm}. \end{aligned}$$

We observe that the printed structure differs in size and due to corner rounding from the target structure. The level for the contour line can be adjusted such that the aerial image matches the target structure better. However, due to corner rounding, without OPC the shape is still ellipsoidal as demonstrated in Fig. 6.33(b).

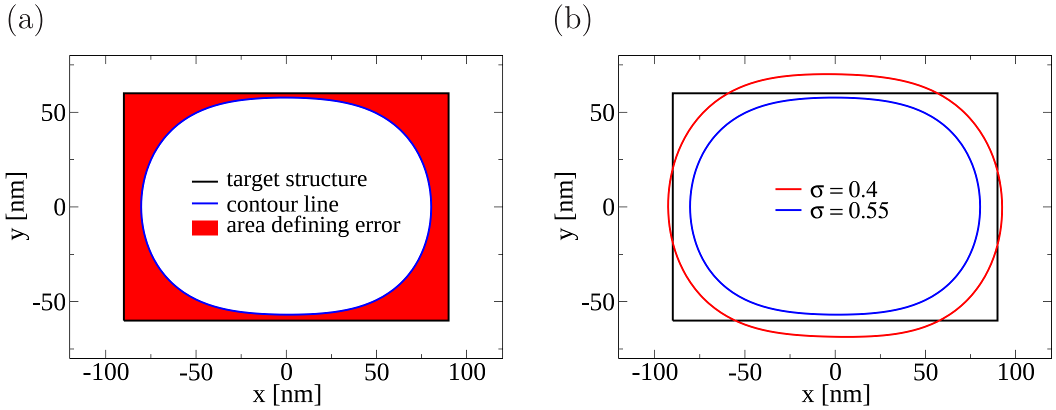


Figure 6.33: (a) Structure on wafer given as contour line of aerial image at $\sigma = 0.55$, i.e., 55% of maximum intensity. Furthermore, the target structure and area defining error functional are depicted. (b) Comparison of contour lines for different levels of σ.

For optimization of the mask layout, we have to define a cost functional. Therefore, let us denote by $\Gamma(\nu)$ the domain enclosed by the target shape γ_T and $\gamma_\sigma(\nu)$, as depicted in Fig. 6.33(a). The cost functional g is then given by the area of this domain:

$$g(\nu) = ||1||_{L^1(\Gamma(\nu))} \,. \tag{6.18}$$

We want to determine optimal parameters such that:

$$\nu_{\min} = \min_{\nu \in D} g(\nu). \tag{6.19}$$

In order to use a Gauß-Newton method for optimization, we need derivative information of the cost functional (6.18). Using for example a finite element representation of the aerial image (6.16), it offers no principle difficulties to compute the derivative of the cost functional:

$$\partial_\nu \, ||1||_{L^1(\Gamma(\nu))}$$

from the derivatives of the far field coefficients $\partial_\nu A_j^r$.

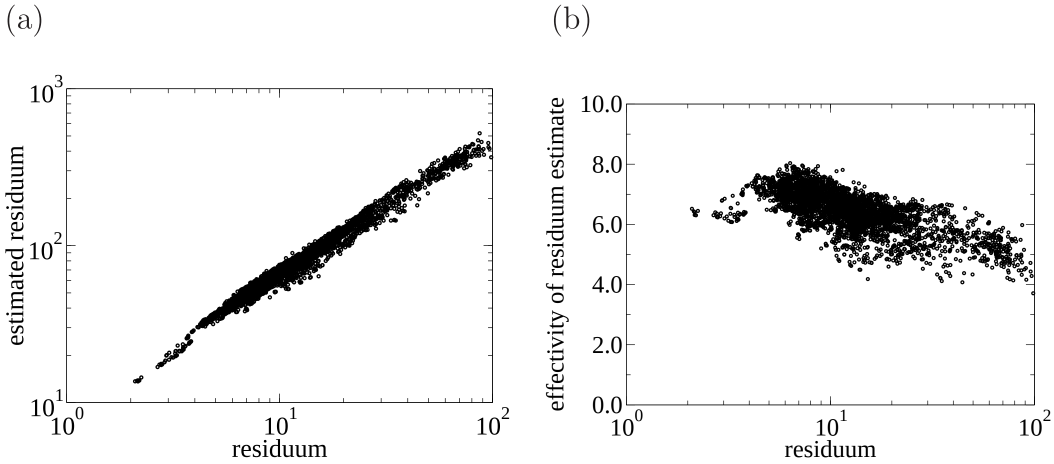


Figure 6.34: (a) Estimated dual norm of residuum (5.74) in dependence on exact dual norm of residuum. (b) Effectivity of residuum estimate (5.75) in dependence on true residuum. (OPC example)

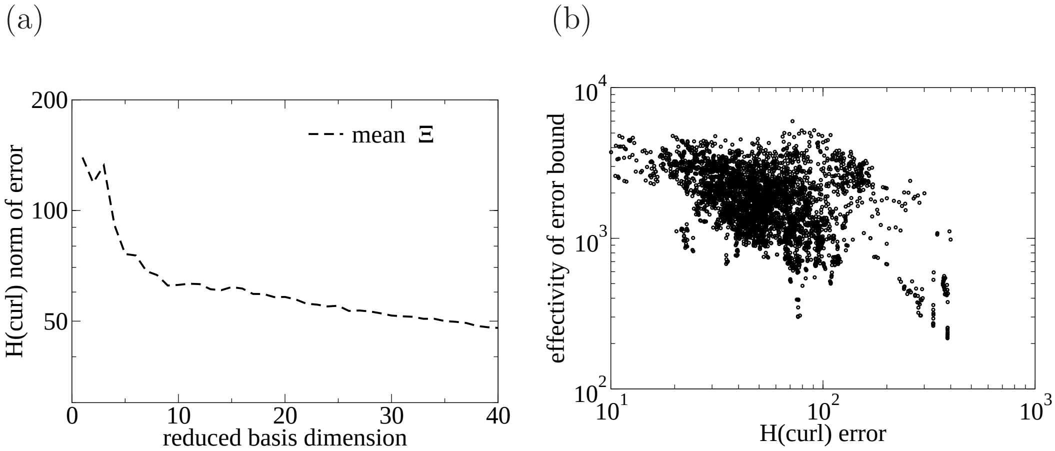


Figure 6.35: (a) Reduced basis solution error (4.8) in H (**curl**, Ω)-norm and (b) effectivity of H (**curl**, Ω)-estimator (5.47a) in dependence on reduced basis solution error. (OPC example)

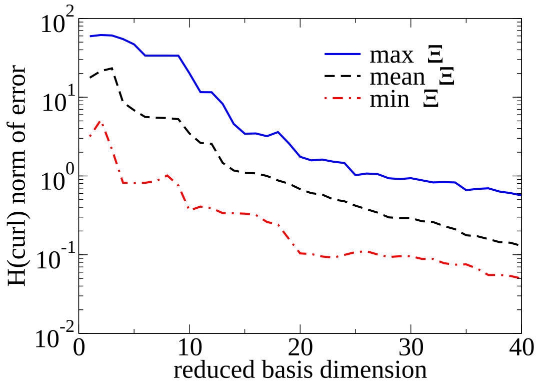


Figure 6.36: Reduced basis solution error (4.8) in $\mathrm{H}\left(\mathbf{curl},\Omega\right)$-norm for small parameter domain: $D = [217\,\mathrm{nm}; 233\,\mathrm{nm}] \times [145\,\mathrm{nm}; 155\,\mathrm{nm}] \times [117\,\mathrm{nm}; 133\,\mathrm{nm}] \times [145\,\mathrm{nm}; 155\,\mathrm{nm}]$, c.f. original parameter domain (6.20). (OPC example)

6.5.3 Reduced basis approximation

The affine decomposition of the truth approximation gives $Q = 1932$ terms. For the parameter domain D we choose:

$$\begin{aligned} p_{1x} &\in [145\,\mathrm{nm}; 305\,\mathrm{nm}], \\ p_{2x} &\in [100\,\mathrm{nm}; 200\,\mathrm{nm}], \\ p_{1y} &\in [45\,\mathrm{nm}; 205\,\mathrm{nm}], \\ p_{2y} &\in [100\,\mathrm{nm}; 200\,\mathrm{nm}]. \end{aligned} \tag{6.20}$$

Considering the high number of terms in the affine expansion, complexity of the truth approximation, regarding the number of incoming fields and computational times, and also the large parameter domain, this example can be seen as a very challenging problem for the reduced basis method. Again we use the technique for multiple sources developed in Section 5.10. Due to high computational times of $2.5h$ for a single snapshot and large memory requirements for the snapshots of all sources, we restrict the reduced basis dimension to $N = 40$.

We compare the reduced basis approximation to the truth approximation over a random parameter ensemble Ξ with 100 points in the parameter domain D. The good performance of the sub-domain residuum estimate is demonstrated in Fig. 6.34. We have low and homogeneous effectivities for the residuum estimate between 4.0 and 8.0. Figure 6.35 shows the convergence of the error in $\mathrm{H}\left(\mathbf{curl},\Omega\right)$-norm and the effectivities of the corresponding error estimate. After an initial drop, the error decreases very slowly for this example. However, this is not surprising since we consider a very large parameter

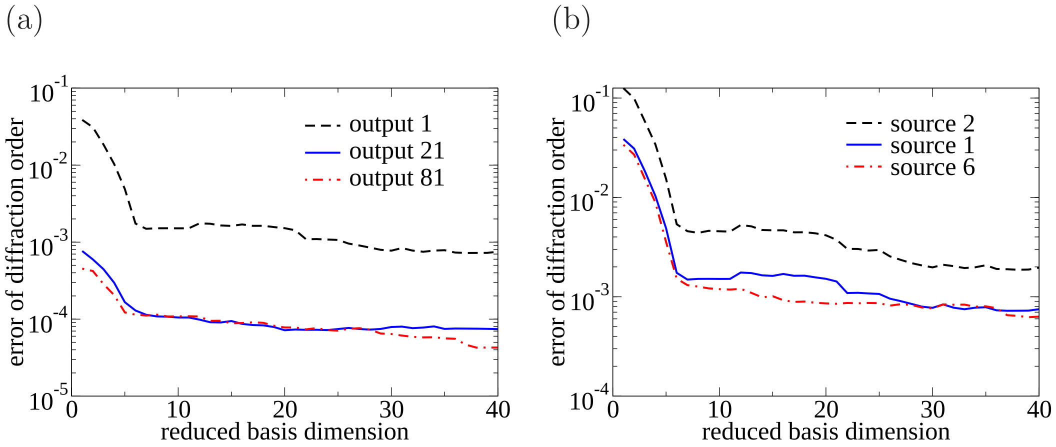


Figure 6.37: Convergence of reduced basis output of interest in dependence on reduced basis dimension (a) for source 1 and different outputs of interest, and (b) output of interest 1 and different sources. (OPC example)

space. For smaller parameter space the exponential convergence is again observable. This is shown in Fig 6.36, where a reduced basis was built for a parameter domain, which is 10 times smaller in each of the four parameter dimensions.

Convergence of the output of interest is shown in Fig. 6.37 for different diffraction orders and sources. We have the same situation as for the $\mathrm{H}(\mathbf{curl}, \Omega)$-norm, with a large initial drop of the error and a very slow convergence. However, the errors are at a relatively low level. We observe that the efficient treatment of multiple sources also works for this example. The outputs of interest for all sources converge with the same rate as for source 1, which was used for construction of the reduced basis system, see Fig. 6.37(b). The inf-sup constant varied between $\beta = 7.7 \cdot 10^{-5}$ and $8.8 \cdot 10^{-4}$ over the random parameter ensemble for this example, which is also due to the large parameter domain.

Since the convergence of the error of the reduced basis solution is very slow, the construction of a basis with rigorous small error bounds is not feasible. After performing the optimization in an application, it might be advisable, to compare the optimal reduced basis solution with the truth approximation for optimal parameters.

Despite the large number of sources and outputs of interest, the reduced basis computation only takes $1.1s$. This gives a speed up factor of about 8000 compared to the truth approximation and allows many-query application.

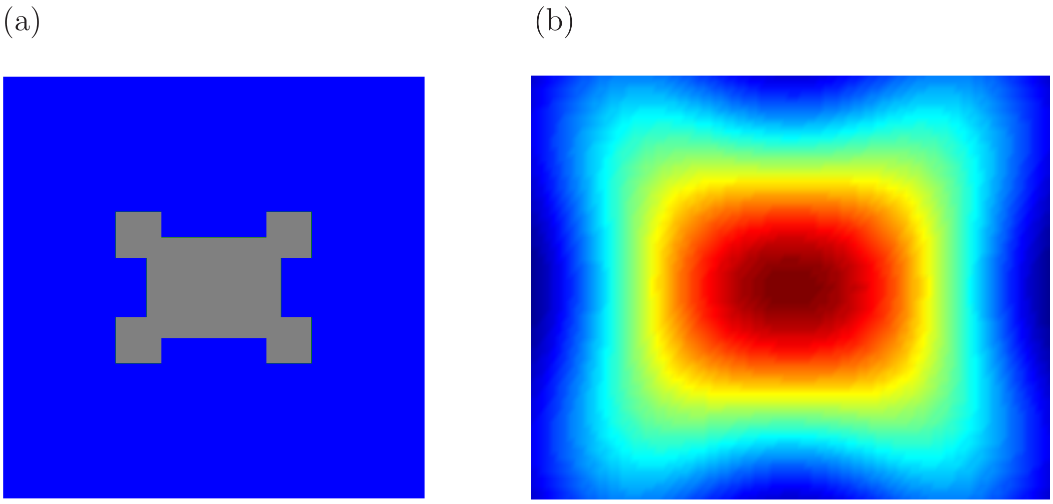


Figure 6.38: (a) Mask layout after OPC optimization and (b) corresponding aerial image.

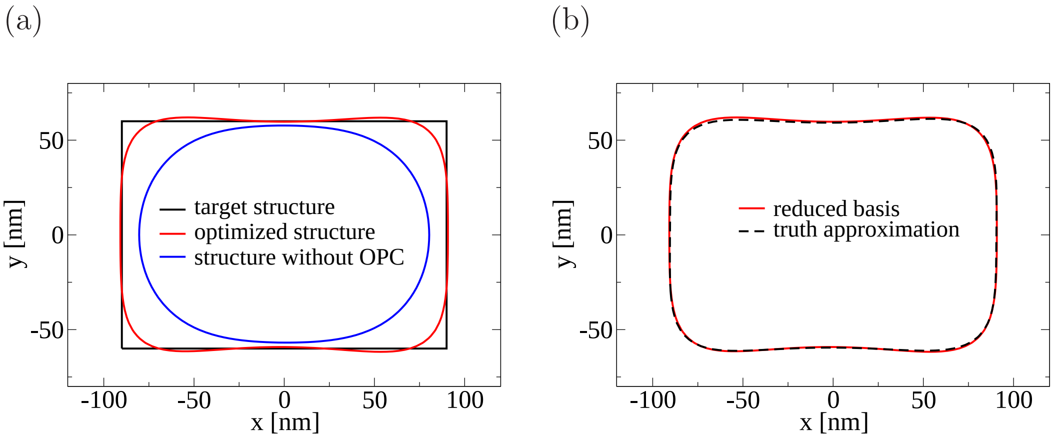


Figure 6.39: (a) Comparison of structure on wafer without and with optimized mask layout obtained from reduced basis computation. (b) Comparison of optimized reduced basis structure and corresponding truth approximation result ($\sigma = 0.55$).

Results

Figure 6.38 shows the optimized geometry of the photomask and the corresponding aerial image. The optimal values were found at:

$$\begin{aligned} p_{1x}^{\text{opt}} &= 295.0\,\text{nm}, \\ p_{2x}^{\text{opt}} &= 142.6\,\text{nm}, \\ p_{1y}^{\text{opt}} &= 152.1\,\text{nm}, \\ p_{2y}^{\text{opt}} &= 185.0\,\text{nm}, \end{aligned}$$

with a value for the cost functional (6.18) of $g(\nu^{\text{opt}}) = 1234.6\,\text{nm}^2$. The shape of the structure on the wafer without and with optimized serifs is depicted in Fig. 6.39(a). The optimized structure shows good agreement to the target structure. Of course corner rounding can not be avoided completely. Furthermore, a comparison of the optimal structure computed with the reduced model and obtained from the truth approximation is given in Fig. 6.39(b). We observe very good agreement. Largest deviations are of the order of 1 nm.

7 Conclusion and outlook

In the present work we developed efficient techniques for the reduced basis method, with a focus on application to real world nano-optical problems.

Especially in the field of a posteriori error estimation and multiple sources, established techniques were found to be infeasible and had to be further developed, in order to treat complex geometries in 2D and 3D and complex sources. Savings of computational costs of several orders of magnitude, could be demonstrated, compared to state-of-the-art methods.

In application examples our results showed that the reduced basis method is very well suited for complex engineering tasks like real-time inverse scatterometry, parameter estimation, and design optimization of optical systems.

Due to the encouraging results for quantitative scatterometry of EUV masks, the German national standards and metrology institute PTB uses the developed implementation as a prototype for evaluation of scatterometric measurements.

Future work will focus on the improvement of error estimation techniques. In general, this is a very difficult field for Maxwell's equations in the high-frequency regime and a topic of actual research [9, 87]. In our numerical examples the error estimates were highly correlated to the true errors and could, therefore, be used for efficient greedy construction of reduced basis spaces. However, the obtained error bounds were often largely overestimating errors.

Besides error estimation, other important topics for future development include the parametrization of exterior domains and parallelization of the developed methods for industrial applications, e.g., in the field of computational lithography.

8 Zusammenfassung

Eine Hauptaufgabe von numerischer Analysis und Modellierung ist die Simulation komplexer technologischer Probleme im ingenieur- und naturwissenschaftlichen Bereich. Simulationen helfen, Systeme oder Komponenten besser zu verstehen, zu designen, zu optimieren oder zu charakterisieren.

In vielen Anwendungsfeldern, wie numerischem Design, Parameterrekonstruktion oder bei inversen Problemen werden im Allgemeinen eine Vielzahl von Simulationen eines gegebenen Systems in Abhängigkeit von z.B. Geometrie- oder Materialparametern durchgeführt. Oft besteht dabei Echtzeitanforderung, so dass kurze Rechenzeiten des Vorwärtsproblems unverzichtbar sind. Vor allem für 3D-Probleme sind die Zeiten für die Berechnung einer einzigen Vorwärtslösung dafür jedoch oft zu lang.

Thema der vorliegenden Arbeit ist die Reduzierte Basis Methode, die zum Ziel hat, parametrisierte Probleme in obigen Anwendungsfeldern in Echtzeit zu lösen. Die Grundidee besteht darin, den Lösungsprozess in eine langsame Offline- und einen schnelle Online-Phase aufzuspalten. In der Offline-Phase wird das zu Grunde liegende Problem mehrmals rigoros gelöst, wobei längere Rechenzeiten in Kauf genommen werden. Diese Lösungen bilden die Basis eines reduzierten niedrigdimensionalen Systems, das man durch Projektion aus dem ursprünglichen Problem erhält. Im Online-Schritt wird lediglich das reduzierte Problem gelöst. Da die Reduzierte Basis Methode Näherungslösungen liefert, ist es für die Qualität und Verlässlichkeit der Rechnungen von großer Bedeutung, rigorose Fehlerschätzer zu konstruieren.

Anwendungsfeld dieser Arbeit ist das Gebiet “Computational Nano-Optics”, das sich mit der Lösung der Maxwellgleichungen in nanostrukturierten Systemen beschäftigt. Speziell werden Streuprobleme auf unbeschränkten, geometrisch parametrisierten 3D-Gebieten betrachtet. Vor allem auf dem Gebiet der a posteriori Fehlerschätzung sind bisherige “State-of-the-Art” Reduzierte Basis Methoden aufgrund extrem hohen Aufwands praktisch nicht durchführbar, um komplexe geometrisch parametrisierte Systeme in 2D und 3D zu behandeln. Daher wurde in der vorliegenden Arbeit ein neuer Fehlerschätzer entwickelt, der den Rechen- und Speicheraufwand um mehrere Größenordnungen reduziert. Dieser basiert auf Gebietszerlegungsmethoden, die auch für Fehlerschätzung von Finite Elemente Lösungen verwendet werden. Desweiteren wurde eine neue Technik für die Reduzierte Basis Methode entwickelt, die es erlaubt, die Reaktion von Systemen unter dem Einfluß einer Vielzahl von Quellen extrem effizient zu berechnen. Dies ist eine typische Situation in vielen nanooptischen Anwendungen, z.B. in der Lithographie.

Als numerische Beispiele wurde die Optimierung von Photomasken und die inverse Scatterometrie von EUV (extrem ultraviolett) Masken untersucht. Die Arbeiten zur inversen Scatterometrie wurden in Kollaboration mit der Physikalisch-Technischen Bundesanstalt (PTB) am Berliner Elektronensynchrotron BESSY II (experimentelle Messungen) und dem Advanced Mask Technology Center (Herstellung einer EUV Testmaske und Mikroskopie) durchgeführt. Aufgrund der vielversprechenden Ergebnisse wird eine Prototypimplementierung der in dieser Arbeit entwickelten Methoden für die Auswertung von Streuexperimenten an der PTB eingesetzt.

Bibliography

[1] M. Ainsworth and J. T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. John Wiley and Sons, 1. edition, 2000.

[2] W. Alt. *Nichtlineare Optimierung*. Vieweg Verlag, 1. edition, 2002.

[3] I. Babuška and W.C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM Journal on Numerical Analysis*, 15(4):736–754, 1978.

[4] E. Balmes. Parametric families of reduced finite element models: Theory and applications. *Mechanical Systems and Signal Processing*, 10(4):381–394, 1996.

[5] G. Bao, D. C. Dobson, and J. A. Cox. Mathematical studies in rigorous grating theory. *J. Opt. Soc. Am.*, 12(5):1029–1042, 1995.

[6] G. Bao and A. Friedman. Inverse problems for scattering by periodic structures. *Archive for Rational Mechanics and Analysis*, 132(1):49–72, 1994.

[7] M. Barrault, N.C. Nguyen, Y. Maday, and A.T. Patera. An "empirical interpolation" method: Application to efficient reduced-basis discretization of partial differential equations. *C. R. Acad. Sci. Paris, Ser. I*, 339:667–672, 2004.

[8] J. Bérenger. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, 114(2):185–200, 1994.

[9] D. Braess and J. Schöberl. Equilibrated Residual Error Estimator for Maxwell's Equations. *Mathematics of Computation*, 77(262):651–672, 2008.

[10] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 1. edition, 1994.

[11] S. Burger, R. Klose, A. Schädle, and F. Schmidt and L. Zschiedrich. FEM modelling of 3D photonic crystals and photonic crystal waveguides. In Y. Sidorin and C. A. Wächter, editors, *Integrated Optics: Devices, Materials, and Technologies IX*, volume 5728, pages 164–173. Proc. SPIE, 2005.

[12] S. Burger, R. Köhle, L. Zschiedrich, W. Gao, F. Schmidt, R. März, and C. Nölscher. Benchmark of FEM, Waveguide and FDTD Algorithms for Rigorous Mask Simulation. In J. T. Weed and P. M. Martin, editors, *Photomask Technology*, volume 5992, pages 378–389. Proc. SPIE, 2005.

[13] S. Burger, L. Zschiedrich, J. Pomplun, and F. Schmidt. JCMsuite: An Adaptive FEM Solver for Precise Simulations in Nano-Optics. In *Integrated Photonics and Nanophotonics Research and Applications, Computer Aided Design for Integrated and Nano Photonics (ITuE)*. Optical Society of America, 2008.

[14] S. Burger, L. Zschiedrich, J. Pomplun, F. Schmidt, B. Kettner, and D. Lockau. 3D Finite-Element Simulations of Enhanced Light Transmission Through Arrays of Holes in Metal Films. In *Numerical Methods in Optical Metrology*, volume 7390. Proc. SPIE, 2009.

[15] E. Cancés, C. Le Bris, Y. Maday, and G. Turinici. Towards reduced basis approaches in ab initio electronic structure computations. *J. Sci. Comput.*, 17:461–469, 2003.

[16] Y. Chen, J.S. Hesthaven, Y. Maday, and J. Rodriguez. A Monotonic Evaluation of Lower Bounds for inf-sup Stability Constants in the Frame of Reduced Basis Methods. *C. R. Acad. Sci. Paris, Ser. I*, 346, 2008.

[17] C. Enkrich, M. Wegener, S. Linden, S. Burger, L. Zschiedrich, F. Schmidt, C. Zhou, T. Koschny, and C. M. Soukoulis. Magnetic metamaterials at telecommunication and visible frequencies. *Phys. Rev. Lett.*, 95:203901, 2005.

[18] L.C. Evans. *Parital Differential Equations.* American Mathematical Society, 1. edition, 1998.

[19] J. P. Fink and W. C. Rheinboldt. On the error behavior of the reduced basis technique for nonlinear finite element approximations. *Z. Angew. Math. Mech.*, 63:21–28, 1983.

[20] H. Flanders. *Differential Forms with Applications to the Physical Sciences.* Dover, 1. edition, 1989.

[21] R. L. Fox and H. Miura. An Approximate Analysis Technique for Design Calculations. *AIAA*, 9(1):177–179, 1971.

[22] M. A. Grepl. *Reduced-Basis Approximation and A Posteriori Error Estimation for Parabolic Partial Differential Equations.* Phd thesis, Massachusetts Institute of Technology, Department of Mechanical Engineering, 2005.

[23] H. Gross, A. Rathsfeld, F. Scholze, R. Model, and M. Bär. Computational methods estimating uncertainties for profile reconstruction in scatterometry. In *Optical Micro- and Nanometrology in Microsystems Technology*, volume 6995. Proc. SPIE, 2008, in press.

[24] M. D. Gunzburger. *Finite Element Methods for Viscous Incompressible Flows: A Guide to Theory, Practice, and Algorithms.* Academic Press, Boston, 1989.

[25] B. Haasdonk and M. Ohlberger. Basis Construction for Reduced Basis Methods by Adaptive Parameter Grids. pages 116–119. International Conference on Adaptive Modeling and Simulation, 2007.

[26] B. Haasdonk and M. Ohlberger. Adaptive Basis Enrichment for the Reduced Basis Method Applied to Finite Volume Schemes. pages 471–478. 5th International Symposium on Finite Volumes for Complex Applications, 2008.

[27] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *Mathematical Modelling and Numerical Analysis*, 42(2):277–302, 2008.

[28] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications.* Springer, 2. edition, 2008.

[29] R. Hiptmair. Finite elements in computational electromagnetism. *Acta Numerica*, 11:237–339, 2002.

[30] J. Hoffmann, C. Hafner, P. Leidenberger, J. Hesselbarth, and S. Burger. Comparison of electromagnetic field solvers for the 3D analysis of plasmonic nano antennas. In *Numerical Methods in Optical Metrology*, volume 7390. Proc. SPIE, 2009.

[31] T. Hohage, F. Schmidt, and L. Zschiedrich. Solving Time-Harmonic Scattering Problems Based on the Pole Condition I:Theory. *SIAM J. Math. Anal.*, 35(1):183–210, 2003.

[32] T. Hohage, F. Schmidt, and L. Zschiedrich. Solving Time-Harmonic Scattering Problems Based on the Pole ConditionII: Convergence of the PML Method. *SIAM J. Math. Anal.*, 35(3):547–560, 2003.

[33] D. B. P. Huynh, G. Rozza, S. Sen, and A. T. Patera. A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. *C. R. Acad. Sci. Paris, Ser. I*, 345:473–478, 2007.

[34] K. Ito and S. S. Ravindran. A reduced-order method for simulation and control of fluid flows. *Journal of Computational Physics*, 143(2):403–425, 1998.

[35] J. D. Jackson. *Classical electrodynamics.* John Wiley and Sons, 3. edition, 1998.

[36] K. Jänich. *Mathematik 2 geschrieben für Physiker.* Springer, 2. edition, 2002.

[37] J. Jin. *The finite element method in electromagnetics.* John Wiley and Sons, 2. edition, 2002.

[38] P. S. Johansson, H. I. Andersson, and E. M. Rønquist. Reduced-basis modeling of turbulent plane channel flow. *Computers & Fluids*, 35:189–207, 2006.

[39] T. Kalkbrenner, U. Håkanson, A. Schädle, S. Burger, C. Henkel, and V. Sandoghdar. Optical microscopy using the spectral modifications of a nano-antenna. *Phys. Rev. Lett.*, 95:200801, 2005.

[40] M. Karl, B. Kettner, S. Burger, F. Schmidt, H. Kalt, and M. Hetterich. Dependencies of micro-pillar cavity quality factors calculated with finite element methods. *Optics Express*, 17(1144), 2009.

[41] S. Kim and J. E. Pasciak. The computation of resonances in open systems using a perfectly matched layer. *Mathematics of Computation*, 78(267):1375–1398, 2009.

[42] A. Tan Yong Kwang. *Reduced Basis Method for 2nd Order Wave Equation: Application to One-Dimensional Seismic Problem.* Master thesis, Massachusetts Institute of Technology, Department of Mechanical Engineering, 2006.

[43] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems.* Cambridge University Press, 1. edition, 2002.

[44] H. Levinson. *Principles of Lithography.* SPIE, 2. edition, 2004.

[45] G.R. Liu, J.H. Lee, A.T. Patera, Z. L. Yang, and K. Y. Lam. Inverse identification of thermal parameters using reduced-basis method. *Comput. Methods Appl. Mech. Engrg.*, 194:3090–3107, 2005.

[46] G.R. Liu, K. Zaw, and Y.Y. Wang. Rapid inverse parameter estimation using reduced-basis approximation with asymptotic error estimation. *Comput. Methods Appl. Mech. Engrg.*, 197:3898–3910, 2008.

[47] Y. Maday, A.T. Patera, and D.V. Rovas. A blackbox reduced-basis output bound method for noncoercive linear problems. In *Nonlinear Partial Differential Equations and their Applications*, volume 31, pages 533–569. Elsevier, 2002.

[48] Y. Maday, A.T. Patera, and G. Turinici. Global a priori convergence theory for reduced basis approximation of single parameter symmetric coercive elliptic partial differential equations. *C. R. Acad. Sci. Paris, Ser. I*, 335:289–294, 2002.

[49] Y. Maday and U. Razafison. A reduced basis method applied to the Restricted Hartree-Fock equations. *C. R. Acad. Sci. Paris, Ser. I*, 346:243–248, 2008.

[50] R. Milani, A. Quarteroni, and G. Rozza. Reduced basis method for linear elasticity problems with many parameters. *Comput. Methods Appl. Mech. Engrg.*, 197:4812–4829, 2008.

[51] P. Monk. *Finite Element Methods for Maxwell's Equations.* Oxford University Press, 2003.

[52] L. Nannen and A. Schädle. Transparent boundary conditions for Helmholtz-type problems using Hardy space infinite elements. *SIAM J. Sci. Comput.*, (submitted).

[53] J.C. Nedelec. Mixed finite elements in R^3. *Numer. Math.*, 35:315–341, 1980.

[54] N. C. Nguyen. *Reduced-Basis approximation and a posteriori error bounds for nonaffine nonlinear partial differential equations: Application to inverse analysis.* Phd thesis, National University of Singapore, 2005.

[55] N. C. Nguyen. A posteriori error estimation and basis adaptivity for reduced-basis approximation of nonaffine-parametrized linear elliptic differential equations. *J. Comput. Phys.*, 227:983–1006, 2007.

[56] W. Nolting. *Grundkurs Theoretische Physik 3.* Springer, 6. edition, 2002.

[57] A. K. Noor and J. M. Peters. Reduced Basis Techique for Nonlinear Analysis of Structures. *AIAA*, 18(4):455–462, 1980.

[58] A. T. Patera and E. M. Rønquist. Reduced basis approximation and a posteriori error estimation for a Boltzmann model. *Comput. Methods Appl. Mech. Engrg.*, 196:2925–2942, 2007.

[59] A.T. Patera and G. Rozza. *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations.* to appear in (tentative rubric) MIT Pappalardo Graduate Monographs in Mechanical Engineering, 1. edition, Copyright MIT 2006.

[60] J. Perlich, F.-M. Kamm, J. Rau, F. Scholze, and G.Ulm. Characterization of extreme ultraviolet masks by extreme ultraviolet scatterometry. *J. Vac. Sci. Technol. B*, 22:3059–3062, 2004.

[61] J. S. Peterson. The reduced basis method for incompressible viscous flow calculations. *SIAM J. Sci. Stat. Comput.*, 10(4):777–786, 1989.

[62] N. A. Pierce and M. B. Giles. Adjoint Recovery of Superconvergent Functionals from PDE Approximations. *SIAM Review*, 42(2):247–264, 2000.

[63] J. Pomplun, S. Burger, F. Schmidt, F. Scholze, C. Laubis, and U. Dersch. Finite Element Analysis of EUV Lithography. In H. Bosse, B. Bodermann, and R. M. Silver, editors, *Modeling Aspects in Optical Metrology*, volume 6617, page 18. Proc. SPIE, 2007.

[64] J. Pomplun, S. Burger, F. Schmidt, F. Scholze, C. Laubis, and U. Dersch. Metrology of EUV masks by EUV scatterometry and finite element analysis. In *Photomask and NGL Mask Technology XV*, volume 7028, page 24. Proc. SPIE, 2008.

[65] J. Pomplun, S. Burger, F. Schmidt, L. Zschiedrich, and F. Scholze. Rigorous FEM-simulation of EUV-masks: Influence of shape and material parameters. In P. M. Martin and R. J. Naber, editors, *Photomask Technology*, volume 6349, page 63493D. Proc. SPIE, 2006.

[66] J. Pomplun and F. Schmidt. Accelerated a posteriori error estimation for the reduced basis method with application to 3D electromagnetic scattering problems. *SIAM J. Sci. Comput. (accepted).*

[67] J. Pomplun and F. Schmidt. Reduced basis method for problems with multiple sources. *SIAM J. Sci. Comput. (submitted).*

[68] T. A. Porsching. Estimation of the error in the reduced basis method solution of nonlinear equations. *Mathematics of Computation*, 45(172):487–496, 1985.

[69] C. Prud'homme, D. Rovas, K. Veroy, Y. Maday, A.T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bounds methods. *Journal of Fluids Engineering*, 124(1):70–80, 2002.

[70] A. Quarteroni. *Numerical Models for Differential Problems.* Springer, 2. edition, 2009.

[71] A. Quarteroni and G. Rozza. Optimal Control and Shape Optimization of Aorto-Coronaric Bypass Anastomoses. *Mathematical Models and Methods in Applied Sciences*, 13(12):1801–1823, 2003.

[72] A. Quarteroni and G. Rozza. Numerical Solution of Parametrized Navier-Stokes Equations by Reduced Basis Methods. *Numerical Methods for PDEs*, 23(4):923–948, 2007.

[73] A. Quarteroni, G. Rozza, and A. Quaini. Reduced Basis Methods for Optimal Control of Advection-Diffusion Problems. In W. Fitzgibbon, R. Hoppe, J. Periaux, O. Pironneau, and Y. Vassilevski, editors, *Advances in Numerical Mathematics*, pages 193–216. Russian Academy of Sciences, Moscow and Department of Mathematics, University of Houston, 2007.

[74] M. Reed and B. Simon. *Methods of Modern Mathematical Physics I: Functional Analysis.* Academic Press, 2. edition, 1980.

[75] W. C. Rheinboldt. On the theory and error estimation of the reduced basis method for multi-parameter problems. *Nonlinear Analysis, Theory, Methods and Applications*, 21(11):849–858, 1993.

[76] D. V. Rovas. *Reduced-Basis Output Bound Methods for Parametrized Partial Differential Equations.* Phd thesis, Massachusetts Institute of Technology, Department of Mechanical Engineering, 2003.

[77] D. V. Rovas, L. Machiels, and Y. Maday. Reduced basis output bounds methods for parabolic problems. *AIAA*, 18(4):455–462, 1980.

[78] G. Rozza. Optimization, Control and Shape Design for an Arterial Bypass. *International Journal Numerical Methods for Fluids IJNMF*, 47(10-11):1411–1419, 2005.

[79] G. Rozza. Reduced Basis Methods for Elliptic Equations in subdomains with A-Posteriori Error Bounds and Adaptivity. *Applied Numerical Mathematics*, 55(4):404–423, 2005.

[80] G. Rozza. *Shape Design by Optimal Flow Control and Reduced Basis Techniques: Applications to Bypass Configurations in Haemodynamics.* Phd thesis, École Polytechnique Fédérale de Lausanne, 2005.

[81] G. Rozza. Reduced basis methods for Stokes equations in domains with non-affine parameter dependence. In *Computing and Visualization in Science*, volume 12, Issue 1, pages 23–35. Springer, 2009.

[82] G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics. *Archives of Computational Methods in Engineering*, 15(3):229–275, 2008.

[83] G. Rozza and K. Veroy. On the stability of reduced basis techniques for Stokes equations in parametrized domains. *Computer Methods in Applied Mechanics and Engineering*, 196(7):1244–1260, 2007.

[84] A. Schädle, L. Zschiedrich, S. Burger, R. Klose, and F. Schmidt. Domain decomposition method for Maxwell's equations: Scattering off periodic structures. *J. Comput. Phys.*, 226(1):477–493, 2007.

[85] F. Schmidt. *Solution of Interior-Exterior Helmholtz-Type Problems Based on the Pole Condition Concept: Theory and Algorithms.* Habilitation thesis, Free University Berlin, Fachbereich Mathematik und Informatik, 2002.

[86] F. Schmidt, T. Hohage, R. Klose, A. Schdle, and L. Zschiedrich. Pole condition: A numerical method for Helmholtz-type scattering problems with inhomogeneous exterior domain. *J. Comput. Appl. Math.*, 218(1):61–69, 2008.

[87] J. Schöberl. A posteriori error estimates for Maxwell Equations. *Mathematics of Computation*, 77(262):633–649, 2008.

[88] J. Schöberl and S. Zaglmayr. High order Nedelec elements with local complete sequence properties. *The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, 24(2):374–384, 2005.

[89] F. Scholze, C. Laubis, C. Buchholz, A. Fischer, A. Kampe, S. Plöger, F. Scholz, and G. Ulm. Polarization dependence of multilayer reflectance in the EUV spectral range. In M. J. Lercel, editor, *Emerging Lithographic Technologies XI*, volume 6517, pages 863–870. Proc. SPIE, 2006.

[90] F. Scholze, C. Laubis, U. Dersch, J. Pomplun, S. Burger, and F. Schmidt. The influence of line edge roughness and CD uniformity on EUV scatterometry for CD characterization of EUV masks. In H. Bosse, B. Bodermann, and R. M. Silver, editors, *Modeling Aspects in Optical Metrology*, volume 6617, page 1A. Proc. SPIE, 2007.

[91] F. Scholze, J. Tümmler, and G. Ulm. High-accuracy radiometry in the EUV range at the PTB soft X-ray radiometry beamline. volume 40, pages 224–228. Metrologia, 2003.

[92] S. Sen. *Reduced-Basis approximation and a posteriori error estimation for non-coercive elliptic problems: Application to acoustics.* Phd thesis, Massachusetts Institute of Technology, Department of Mechanical Engineering, 2007.

[93] W. Singer, M. Totzeck, and H. Gross. *Handbook of Optical Systems: Vol. 2. Physical Image Formation.* Wiley-VCH, 2. edition, 2005.

[94] M. Sugawara and I. Nishiyama. Impact of slanted absorber sidewall on printability in EUV lithography. volume 5992. Proc. SPIE, 2005.

[95] M. Sugawara, I. Nishiyama, and M. Takai. Influence of asymmetry of diffracted light on printability in EUV lithography. volume 5751, pages 721–732. Proc. SPIE, 2005.

[96] T. Tonn and K. Urban. A reduced-basis method for solving parameter-dependent convection-diffusion problems around rigid bodies. In P. Wesseling, E. Oñate, and J. Periaux, editors, *ECCOMAS CFD 2006 Proceedings.* TU Delft, 2006.

[97] J. Tümmler, G. Brandt, J. Eden, H. Scherr, F. Scholze, and G. Ulm. Characterization of the PTB EUV reflectometry facility for large EUVL optical components. volume 5037, pages 265–273. Proc. SPIE, 2003.

[98] G. Ulm, B. Beckhoff, R. Klein, M. Krumrey, H. Rabus, and R. Thornagel. The PTB radiometry laboratory at the BESSY II electron storage ring. volume 3444, pages 610–621. Proc. SPIE, 1998.

[99] K. Veroy. *Reduced Basis Methods Applied to Problems in Elasticity: Analysis and Applications*. Phd thesis, Massachusetts Institute of Technology, Department of Mechanical Engineering, 2003.

[100] K. Veroy, C. Prud'homme, and A. T. Patera. Reduced-basis approximation of the viscous Burgers equation: rigorous a posteriori error bounds. *C. R. Acad. Sci. Paris, Ser. I*, 337:619–624, 2003.

[101] T. Weiland. A discretization method for the solution of Maxwell's equations for six-component fields. *Archiv für Elektronik und Übertragungstechnik*, 31(3):116–120, 1977.

[102] D. Werner. *Funktionalanalysis*. Springer, 3. edition, 2000.

[103] A. K.-K. Wong. *Optical Imaging in Projection Microlithpgraphy*. SPIE, 1. edition, 2005.

[104] M. Wurm. *Über die dimensionelle Charakterisierung von Gitterstrukturen auf Fotomasken mit einem neuartigen DUV-Scatterometer*. Phd thesis, Friedrich-Schiller-Universität Jena Physikalisch-Astronomische Fakultät, 2008.

[105] K. Yee. Numerical solution of inital boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, 1966.

[106] S. Zaglmayr. *High Order Finite Elements for Electromagnetic Field Computation*. PhD thesis, Johannes Kepler Universität Linz, 2006.

[107] Z. Zhu and F. Schmidt. An efficient and robust mask model for lithography simulation. In H. J. Levinson and M. V. Dusa, editors, *Optical Microlithography XXI*, volume 6925, page 126. Proc. SPIE, 2008.

[108] L. Zschiedrich. *Transparent Boundary Conditions for Maxwell's Equations: Numerical Concepts beyond the PML Method*. PhD thesis, Freie Universität Berlin, Fachbereich Mathematik und Informatik, 2009.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, die vorliegende Arbeit selbständig und ausschließlich unter Verwendung der angegebenen Hilfsmittel und Quellen verfasst zu haben.

Berlin, den 23. Oktober 2009

Jan Pomplun