TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY
Julkaisu 811 • Publication 811

Pasi Raumonen

Mathematical Structures for Dimensional Reduction and Equivalence Classification of Electromagnetic Boundary Value Problems

Tampere 2009

Tampereen teknillinen yliopisto. Julkaisu 811
Tampere University of Technology. Publication 811

Pasi Raumonen

Mathematical Structures for Dimensional Reduction and Equivalence Classification of Electromagnetic Boundary Value Problems

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Rakennustalo Building, Auditorium RG202, at Tampere University of Technology, on the 26th of June 2009, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2009

ISBN 978-952-15-2164-5 (printed)
ISBN 978-952-15-2211-6 (PDF)
ISSN 1459-2045

Abstract

Boundary value problems (BVPs) are fundamental in electromagnetic engineering. The aim of this thesis is to introduce mathematical structures that can be exploited in a new way to formulate electromagnetic BVPs. The tools employed come from differential geometry and the theory of manifolds.

The structures offer a way to model electromagnetism in a coordinate-free manner, which is independent of the chosen metric. Differentiable manifolds and differential forms are used as models for space and electromagnetic fields, respectively. Together with the pullback, exterior derivative, and wedge product, they can be employed to introduce a formulation of electromagnetism that is invariant under diffeomorphisms.

Differential geometry enables us to formulate general electromagnetic BVPs, including static, initial value, and Cauchy problems, in a unified setting. Furthermore, under diffeomorphisms, equivalence of BVPs arises naturally and provides a unified theoretical setting for many traditional, seemingly different methods and approaches. Because of the diffeomorphism-invariance, in formulations of electromagnetic BVPs the metric of space is needed only to make the first connection between the model and the observations. The thesis introduces also $(3 + 1)$-decompositions of Maxwell's equations based on coordinate- and metric-free observer fields. A major results of this thesis is this unified aspect to BVPs and its applications to solution methods.

The structures used are also generic to all dimensions, which makes them natural tools to formulate electromagnetic BVPs of any dimension. In particular, another main result of this thesis is a symmetry-based theory of dimensional reduction of electromagnetic BVPs. It includes a dimensional reduction theorem that gives sufficient conditions for a BVP to be solved as a lower-dimensional BVP and also formulates the lower-dimensional BVP. Because the theory is completely independent of coordinates, metric, and dimension, differential geometric structures are virtually custom-made for it.

The thesis presents several applications and numerical examples, in which the structures offer new insight and benefits. These applications and examples include mesh generation problems, speeding up parametric models that include shape optimization and movement, open-boundary problems, invisibility cloaking, and dimensional reduction of helicoidal geometries.

Preface

This thesis and the study leading to it was carried out at the Department of Electronics, Tampere University of Technology. The thesis is part of a larger and longer term research project of the department aiming to find suitable mathematical structures for giving better understanding and more effective tools for numerical electromagnetic modeling.

I am grateful to prof. Lauri Kettunen for giving me the opportunity to do this work at best possible circumstances that one can reasonably hope for. I also want to thank him all the advice and the guidance he has given for my work.

I am also particularly grateful to Dr. Saku Suuriniemi who has been the actual day-to-day advisor of my work. His support and elaborate advice and our conversations have been central for this work.

The work of the last four and half years at the department has been a great experience and I have had good times with my colleagues at the department in work and in leisure. Special thanks are due to my colleagues Timo Tarhasaari, Janne Keränen, and Tuukka Nieminen. I also want to thank Lasse Söderlund and Maija-Liisa Paasonen for taking care of the administrative duties. Thanks are also due to Timo Lepistö for grammatical proofreading of this thesis.

This work was financially supported by Academy of Finland and Graduate School of Applied Electromagnetism.

Finally, I would like to thank my family and friends for all the support.

Contents

Abstract **i**

Preface **iii**

List of symbols **ix**

1 Introduction **1**

2 Mathematical structures and symmetry **5**

2.1 Mathematical structures . . . 5

2.2 Symmetry . . . 7

2.2.1 Mathematical definition of symmetry . . . 7

2.2.2 Instances of symmetry . . . 10

3 Basic differential geometric structures **13**

3.1 Topological space . . . 14

3.2 Metric space . . . 17

3.3 Vector space and its orientation . . . 18

3.4 Euclidean space . . . 20

3.5 Manifolds . . . 21

3.5.1 Topological manifold and charts . . . 22

3.5.2 Differentiable manifolds . . . 25

3.5.3 Tangent space . . . 28

3.5.4 Oriented manifolds . . . 30

3.5.5 Submanifolds . . . 31

3.5.6 Manifolds-with-boundary . . . 32

3.6 Lie groups . . . 33

3.7 Foliations of manifolds . . . 35

3.8 Analysis on manifolds . . . 36

3.8.1 Tangent bundle . . . 36

3.8.2 Cotangent bundle . . . 37

3.8.3 Differential forms 38
3.8.4 Pullback 39
3.8.5 Wedge product 40
3.8.6 Exterior derivative 42
3.8.7 Contraction and extension 43
3.8.8 Lie derivative 44
3.8.9 Integration on manifolds and Stokes's theorem 46
3.9 Metric structure for manifolds 48
3.9.1 Riemannian and semi-Riemannian manifolds 48
3.9.2 Representation of differential forms as proxy vector fields 51
3.9.3 Hodge-operator 53
3.10 Calculations in coordinates 54

4 Electromagnetic boundary value problems 57
4.1 Geometric decompositions of fields and the exterior derivative 59
4.1.1 Foliations of manifolds and observer structures 59
4.1.2 Geometric decomposition of fields 61
4.1.3 Geometric decomposition of the exterior derivative . . 65
4.2 Maxwell's equations 66
4.2.1 Maxwell's equations in spacetime 66
4.2.2 $(3 + 1)$-decomposition of Maxwell's equations 67
4.3 Decomposition of constitutive equations 70
4.3.1 Constitutive equations in spacetime 71
4.3.2 $(3 + 1)$-decomposition of constitutive equation 71
4.4 Electromagnetic BVPs with differential geometry 73
4.5 Equivalent formulations of a BVP 77
4.5.1 Equivalent differential equations 78
4.5.2 Equivalent boundary values 80
4.5.3 Equivalent constitutive equations 81
4.5.4 Equivalent BVP 82
4.6 Equivalent BVPs: Material parameters and chart 83
4.7 Metric and electromagnetic BVPs 87
4.7.1 Formulation of BVPs in practice 88
4.7.2 Constitutive equations and metric 89
4.7.3 Role of metric 91

5 Dimensional reduction of electromagnetic boundary value problems 93
5.1 Group action on a BVP domain 95
5.2 Group-invariant fields 98
5.3 Group-invariant constitutive equations 103

5.4 Unique solution of an invariant BVP is invariant 104
5.5 Orbit space 107
5.5.1 Manifold structure 108
5.5.2 Metric structure 111
5.6 Lower-dimensional BVPs for one-dimensional symmetry groups 112
5.6.1 Geometric decomposition of fields and the exterior derivative 114
5.6.2 Differential equations for reduced BVPs 116
5.6.3 Boundary values for reduced BVPs 119
5.6.4 Constitutive equations for reduced BVPs 123
5.6.5 Summary 127
5.7 Static and time-harmonic electromagnetics 129
5.7.1 Orbit space and geometric decompositions 130
5.7.2 Static and time-harmonic differential equations 131
5.7.3 Constitutive equations 132
5.8 Lower-dimensional BVPs for multi-dimensional symmetry groups 133
5.8.1 Group action and geometric decompositions 134
5.8.2 Reduced differential equations 136
5.8.3 Reduced boundary values 141
5.8.4 Reduced constitutive equations 143
5.9 Dimensional reduction theorem of electromagnetic BVPs . . . 147

6 Applications 149
6.1 Mesh generation problems 149
6.2 Open-boundary problems 152
6.3 Speeding up parametric models 155
6.4 Invisibility cloaking 156
6.5 Axisymmetric problems 159
6.6 Dimensional reduction 159

7 Examples 161
7.1 Parametric models: shape optimization 161
7.2 Dimensional reduction: helicoidal geometries 168

8 Conclusion 175

A Bibliography classification 179

Bibliography 181

Index **187**

List of symbols

$\mathcal{A}$ atlas, see Definition 3.24
BVP boundary value problem
B magnetic flux density
$B(p, r)$ open ball at point p with radius r
$\binom{n}{k}$ the binomial coefficient
$[c]$ tangential equivalence class of curve c, see Definition 3.29
C^r r-times continuously differentiable
d, d_M exterior derivative
d_τ horizontal exterior derivative
d_S spatial exterior derivative
D electric flux density
δ_{ij} Kronecker delta
$\mathcal{D}$ differentiable structure, see Definition 3.22
$dim(M)$ the dimension of space M
dis distance mapping, see Definition 3.9
∂M boundary of M
ϵ operator modeling permittivity
E electric field
E^n n-dimensional Euclidean space, see Definition 3.17
$(e_1, ..., e_n)$ the canonical basis of the Euclidean space $\mathbb{R}^n$
$\mathbb{F}$ field, usually the field of real or complex numbers
$\mathcal{F}$ the electromagnetic field two-form
$\flat$ metric induced isomorphism called flat
f_* pushforward of the mapping f, see Definition 3.31
f^* pullback of the mapping f, see Definition 3.50
$f^*\mathcal{A}$ pullback atlas, see Definition 3.27
$\mathcal{G}$ the excitation two-form
Gp orbit of p under group action of G, see definition 2.4
H magnetic field
$\mathcal{H}$ linear operator fixing the cohomology classes
H^n a half space of $(\mathbb{R}^n)$
i_X contraction with respect to X, see Definition 3.53
I_α extension with respect to X, see Definition 3.54
$\mathcal{J}$ the source three-form
J current density
κ cross-section, see Definition 5.9
$ker(\tau)$ kernel of a linear mapping τ
$\mathcal{L}_X$ Lie derivative with respect to X, see Definition 3.57
μ operator modeling permeability

ν inverse of μ
$\mathcal{O}$ orientation of a manifold, see Definition 3.33
$||v||$ norm of v
$\Omega(M)$ the set of all differential forms on M
$\Omega^k(M)$ the set of all differential k-forms on M
$\Omega_h(M)$ the set of all horizontal differential forms on M
$\Omega_h^k(M)$ the set of all horizontal differential k-forms on M
$\oplus$ direct sum
Φ a symmetric bilinear form
π canonical or natural projection
P_T a projection of differential forms induced by an observer
P_τ a projection of differential forms induced by an observer
$(\mathbb{R}, +)$ group of real numbers under addition
ρ charge density
S^1 Lie group of all complex numbers with absolute value one
sgn a mapping that maps even permutation to 1 and odd permutation to -1
$\sharp$ metric induced isomorphism called sharp
σ operator modeling conductivity
$\sim$ an equivalence relation
$span(T)$ the span of a vector T
$\star$ Hodge operator
$\mathcal{T}$ topology, see Definition 3.1
$T_p(M)$ tangent space of point p, see Definition 3.30
$T_p^*(M)$ cotangent space of point p, see Definition 3.45
$T(M)$ the set of all tangent vectors or the tangent bundle of M
$T^*(M)$ the set of all covectors or the cotangent bundle of M
$t\omega$ trace of ω, restriction of ω to boundary
υ Hodge-like operator, see Definition 4.6
$v_{||}$ horizontal component of v
$v_\perp$ vertical component of v
V^* dual space of a vector space V
vol a volume form, see Definition 3.67
$\wedge$ wedge product, see Definition 3.51

Chapter 1

Introduction

Electromagnetic engineering problems often involve solutions of mathematical problems called *boundary value problems* (BVP). Formulating an electromagnetic BVP is about specifying the domain and a set of Maxwell's equations that govern the electromagnetic fields inside the domain of the BVP. Moreover, for a unique solution to this system of equations, the constitutive equations, boundary values of fields, and possibly some constraints on cohomology classes must be specified. Various approaches and mathematical formalisms can be used to formulate electromagnetic BVPs. However, these different approaches and formalisms need not be equally good for formulating BVPs in different dimensions or for conceptual understanding the underlying physics and numerical solution methods.

The traditional approach to formulating electromagnetic BVPs is based on classical *vector analysis*. With classical vector analysis, the domain of the BVP is typically modeled with a single coordinate system, and the electric and magnetic fields are modeled with vector fields, Maxwell's equations are written with curls and divergences, and the constitutive equations are given using scalar or tensor fields. This conventional formalism has its merits, but it is based on strict initial assumptions. First, *one must choose a metric for the domain at the beginning of the modeling process*, after which most of the mathematical structures used in the formalism are defined with respect to this metric. Second, all structures of the formalism are built initially on three-dimensional domains, making the formalism *inherently three-dimensional*[1].

The above assumptions often make it challenging to apply the formalism. Because of the metric, expressions of the fields, Maxwell's equations,

[1]It is a particular property of a three-dimensional space that with a metric 2-vectors can be identified with 1-vectors. Consequently, it is possible to introduce vector analysis based solely on 1-vectors. This identification of 1- and 2-vectors is not possible in any other dimensions.

boundary values, and constitutive equations depend on a particular choice of the metric. Particularly, the differential operators grad, curl, and div are evaluated separately for each coordinate system. For instance, consider how the curl and divergence operators are represented in Cartesian and in cylindrical coordinate systems. The inherent three-dimensionality of vector analysis also makes it often a challenge to apply this formalism to other dimensions, because no natural counterparts exist for all the structures in other dimensions.

Vector analysis may give an illusion of an unnecessary dependency on metric and dimension. Particularly, the concept of *symmetry* has many applications in electromagnetic modeling and is often used to reduce the size of the BVP domain. If a BVP is symmetric, the domain consists of copies of some subdomain. However, symmetry is often not employed to its full potential but is instead understood in the restricted sense that the sizes of the subdomains that constitute the domain should be the same. Yet symmetry itself does not depend on such a metric. Furthermore, symmetry principles are often used without properly identifying them as such. For instance, the dimension of a BVP can often be reduced, if some component of the fields is fixed to zero. In summary, assumptions that make a formalism too rigid a construction can be limiting when one formulates and solves particular BVPs numerically with computers.

This thesis aims to bring up to date the traditional approach to formulate electromagnetic BVPs by introducing an modern alternative approach based on *differential geometry*. The needed mathematical structures are introduced and the benefits gained in formulating BVPs are demonstrated in several ways. The benefits follow from the flexibility that no assumptions are made about the metric and dimension of the domain. This allows a *clear separation of metric, orientation, and dimension, helping us to recognize those aspects of electromagnetism that do not depend on them.* Therefore, the differential geometric approach provides more accurate and precise geometrical and conceptual tools for modeling physics than the traditional approaches. For instance, the very idea of the electric field E is about electromotive forces along curves, and the electric flux D has to do with surfaces. Consequently, they should be modeled with objects that naturally correspond to curves and surfaces. Compare this to vector analysis, where $\mathbf{E}$ and $\mathbf{D}$ are both modeled with vector fields, not suggestive of any preferred geometric object.

In the differential geometric approach, space and spacetime are modeled with a mathematical structure called the differentiable manifold. With the manifold, the metric can be treated as a separate structure, and the multivariable and multivalue calculus can be defined without the metric and in a coordinate-free manner. Calculus employs so-called differential forms that model electromagnetic fields. The gradient, curl, and divergence are superseded by the exterior derivative, which is enough to impose Maxwell's equations on this formalism. Thus we can formulate most aspects of electromagnetic BVPs on a *differential topology* level. Particularly, this formulation is *invariant under diffeomorphisms*, which is an analog for general covariance under general differentiable change of coordinates.

Diffeomorphism-invariance allows us to define the *equivalence of BVPs under diffeomorphism*. The equivalence is a generalization of the traditional change of the coordinates procedure. Furthermore, the equivalence gives a *unified theoretical explanation for many traditional, seemingly different methods*. For instance, methods to solve open boundary problems and cloaking or "invisibility" can be explained in a unified manner based on formulations of equivalent BVPs. The equivalence suggest also new practical possibilities such as how to speed up parametric modeling.

The tools of differential geometry are not restricted to certain problems, such as static and time-harmonic problems, initial value problems, and Cauchy problems; rather they provide us with a unified setting of boundary value problems: all the problems consists of (partial) differential equations defined on a domain with a boundary such that the fields governed by the differential equations are pre-defined at the boundary. Consequently, equivalence under diffeomorphism can be established at once for all types of problems that accept the unified setting.

In the differential geometric setting of formulating BVPs, the metric is the tool that together with distance measurements provides a connection between model and observations: the manifold, its topology and charts, and the constitutive equations are first constructed using the metric. However, once the BVP is formulated, an equivalence of BVPs can be defined fully without the metric. Furthermore, the constitutive equations are relations between the fields and because the fields can be defined without metric, it follows that the relations do not depend on the choice of the metric. However, the representations of the relations with metric-dependent Hodge-operators do, of course, depend on the choice of the metric.

A general way is also introduced to decompose spacetime to space and time to derive $(3+1)$-decompositions of Maxwell's equations. This is done with the so-called observer structure, which can be characterized as a field of local observers. The observer structure can be defined in a coordinate- and metric-free manner.

All the structures introduced are *generic to all dimensions*, which makes them convenient for formulating BVPs in any dimension. In particular, this generality is exploited to derive a *theory of dimensional reduction of electromagnetic BVPs*. The theory provides sufficient conditions for solving a BVP as a lower-dimensional BVP and also formulates the lower-dimensional BVPs. The theory relies on *symmetry alone*, which implies that some components of solution fields are not assumed to vanish in some special coordinate system. In fact, all theories explaining dimensional reduction are based on symmetry, either implicitly or explicitly. Thus the *symmetry of BVPs is here examined independent of coordinates, metric, and dimension*. Consequently, one can recognize symmetries, such as helicoidal geometries, that are not obvious at the first glance.

Finally, let us briefly outline the content of this thesis. Chapter 2 discusses mathematical structures and the concept of symmetry, the key elements in this thesis. Chapter 3 defines the main differential geometric structures used in this thesis. After the preliminary chapters 2 and 3, discussion proceeds to the main engineering content of the thesis: Chapter 4 shows how to formulate general electromagnetic BVPs with the tools introduced in chapter 3 in a unified setting. Then the equivalence of BVPs under diffeomorphism is derived, and the role of the metric in electromagnetic BVPs is discussed. Furthermore, chapter 4 defines observer structures and derives general (3+1)-decompositions of Maxwell's equations. Chapter 5 presents the theory of dimensional reduction of electromagnetic BVPs, and chapters 6 and 7 present applications and numerical examples thereof.

Chapter 2

Mathematical structures and symmetry

Mathematical structures are the focus of this thesis; therefore, before introducing any structures, we should briefly explain what is meant by these structures. Another focus is the concept of symmetry and its uses. We give a precise meaning of the concept by defining it mathematically. Then we discuss several applications of symmetry in mathematics and physics.

2.1 Mathematical structures

Mathematical structures are defined using sets and adding more mathematical objects somehow incorporated in the sets. That is, a mathematical structure is a set with various mathematical objects such as relations and operations, which define what one can do with the elements of the set. A collection of associated mathematical objects is called the structure, and the set is called the underlying set. Thus mathematical structure is a universal term for constructions that unify particular mathematical set-constructions with concrete sets.

As an example of a mathematical structure, we give the definition of an important algebraic structure called the group:

Definition 2.1. A *group* $(G, *)$ is a set G together with a binary operation $* : G \times G \to G$, denoted by $*(a, b) = a * b$, that satisfies the following axioms:

(1) associativity: for all $a, b, c \in G$, the equation $(a * b) * c = a * (b * c)$ holds.

(2) identity element: there exists an element $e \in G$ such that for all $a \in G$, the equation $e * a = a * e = a$ holds.

(3) inverse element: for each $a \in G$, there exists an element b such that $a * b = b * a = e$ where e is an identity element.

An *Abelian group* is a group $(G, *)$ that satisfies an additional axiom:

(4) commutativity: for all $a, b \in G$, the equation $a * b = b * a$ holds.

Thus a group is a structure consisting of a set plus some binary operation on the set such that the binary operation satisfies certain axioms. That is, the mathematical object that is incorporated in the underlying set is a special kind of binary operation. The addition of real numbers is an example of a group, which is denoted by $(\mathbb{R}, +)$. Another example is the addition of vectors in a three-dimensional vector space. Thus there is a multitude of instances of groups, and the term *group structure* refers to what is common to all possible groups: group structure is an abstract construction that unifies or captures the essence of particular mathematical set-constructions such as $(\mathbb{R}, +)$ and the addition of vectors.

Even though $(\mathbb{R}, +)$ is a group, the particular group structure it has or the group it defines is not only about real numbers and their additions. That is, $(\mathbb{R}, +)$ defines a particular group using sets as a *language to define and communicate the properties of the group*. Similarly, the addition of vectors in the dimension n defines a totally different group. Hence set-constructions such as $(\mathbb{R}, +)$ are used to define a group, but the mere group itself is more elementary and abstract in the sense that $(\mathbb{R}, +)$ has *excess features that the group it defines does not have*, e.g., the order of real numbers.

To make the above point even clearer, let us look at the following example of two apparently different groups. Let V be a one-dimensional subspace of a three-dimensional vector space W. Then for each $v \in V$, there is a mapping $f_v : W \to W$, called translation by v such that $f_v(w) = w + v$ holds for all $w \in W$. The set of all such mappings f_v defined on W forms a group under the composition of mappings, because the composition of two translations is again a translation. Let us denote the set of all these translations by $T(W, V)$. Now we can put the real numbers and the translations of $T(W, V)$ into bijective correspondence such that the *group structures are preserved* under the correspondence: there is a mapping $g : \mathbb{R} \to T(W, V)$ such that $g(a + b) = g(a) \circ g(b)$ holds for all $a, b \in \mathbb{R}$. Clearly, the groups $(T(W, V), \circ)$ and $(\mathbb{R}, +)$ are in the group sense structurally fully identical, even though they have different underlying sets and binary operations. These two examples of groups are said to be *isomorphic*, and the mapping g is called an *group isomorphism*.

In general we use the term *isomorphism* for bijective *structure-preserving* mappings. The term *homomorphism* is used for structure-preserving mappings that are not bijective, i.e., for mappings that preserve the structures of the domain and codomain but that do not allow complete identification of the structures. If there are no general terms for the isomorphisms and homomorphisms of a certain structure, such as homeomorphism for the isomorphisms of topological space-structure, we add a suitable adjective before isomorphism and homomorphism, e.g., metrical isomorphism, to distinguish between different isomorphisms. Thus the term isomorphism is used in this thesis in the category theoretical sense [19].

Finally, let us comment briefly on the relations of physics and mathematics. First of all, any mathematical model used in physics is devoid of any physical significance without some interpretation, which connects the model to observations. In this sense, there are no "right" or "correct" mathematical structures to describe a physical phenomenon, because the structures themselves do not give meaning to physics. For example, even if most of people consider vector analysis formulation of electromagnetic theory more natural or intuitive, it does not mean that a formulation based on differential geometry is less correct a description of electromagnetism than the former. Of course, not all mathematical structures describe physics equally well or are even capable of allowing a useful interpretation. One goal of this thesis is to provide mathematical structures different from those of classical vector analysis for modeling electromagnetics and for showing that the structures are useful for numerical modeling and understanding physics.

2.2 Symmetry

The concept of symmetry is very common in physics. Evidently, it is a mathematical notion, and for this reason we first look at its mathematical definition. Only afterwards do we discuss its uses in physics.

2.2.1 Mathematical definition of symmetry

Intuitively, for many, symmetry has a strong visual meanings, and it is related to balance, harmony, and self-similarity. However, in mathematics, symmetry is a precisely defined concept, which is stripped of visual significance. Informally speaking, *symmetry is about something remaining the same under some transformations*. For example, let us look at Figure 2.1, which demonstrates "visual symmetry": the image is said to be symmetric because after 180-degree rotations and reflections with respect to diagonals, the figure appears exactly the same. Thus "remaining the same" means here that the image looks exactly the same, and the transformations are the above rotations and reflections.

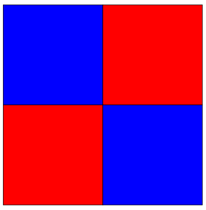

Figure 2.1: A figure that looks symmetric.

Formally speaking, symmetry is described with a set of objects, an equivalence relation on those objects, and a collection of bijective mappings from the set of objects back to itself. The elements of the set are the objects to be transformed (the points of the figure), the equivalence relation describes the notion of "remaining the same" (the points have the same color), and the mappings from the set back to itself are the transformations of the objects (the rotations and the reflections of the figure). Accordingly, we say that the set has a symmetry defined by a relation and a collection of mappings, if the points are equivalent to their images under the mappings. Thus Figure 2.1 has symmetry, because its points can be mapped to points with the same color.

Next, we give a proper definition of symmetry. First, we characterize the transformations: the transformations of a set under which the points may be equivalent to their images are called *symmetry transformations*. The reflexivity of an equivalence relation ($a \sim a$) implies that the identity mapping of the set is always a symmetry transformation. Furthermore, the symmetry property of equivalence relations (if $a \sim b$ then $b \sim a$) implies that the inverse of a symmetry transformation is again a symmetry transformation. Finally, the transitivity of equivalence relations (if $a \sim b$ and $b \sim c$ then $a \sim c$) implies that the composition of two symmetry transformations is also a symmetry transformation. Thus all *symmetry transformations are bijections, and they form a group under a composition of mappings.* This group is called the symmetry group of the set with respect to the equivalence relation.

Sometimes there may be multiple different equivalence relations and correspondingly multiple symmetries; for example, a physical model may contain both scalar fields and vector fields over some space. If we talk about symmetries of scalar and vector fields, then symmetry transformations map scalar and vectors fields to scalar and vector fields, respectively, over the same space. An equivalence relation for fields could be such that a field and its transformed field are pointwise the same. Now it is clear that equivalence relations on scalar and vector fields cannot be the same relations. Furthermore, the symmetry transformations for scalar and vector fields cannot be the same mappings either. Thus all equivalence relations may have their own symmetry groups. However, even if the elements of these symmetry groups are different and thus not directly comparable, the groups may be isomorphic. Thus in case of multiple equivalence relations, it is convenient to separate the group structure or the abstract group from particular symmetry groups. That is, we consider only some isomorphic copy of the symmetry groups. For example, a group of translations is isomorphic to $(\mathbb{R}^n, +)$, and if there are translations of points and vectors, the symmetry groups of points and vectors are conveniently and simultaneously described with $(\mathbb{R}^n, +)$. This leads to a mathematical model of symmetry transformations called the group action:

Definition 2.2. Let $(G, \cdot)$ be a group and M a set. A *group action* of G on M is a mapping $f : G \times M \to M$, which satisfies the following axioms:

(1) $f(g \cdot h, p) = f(g, f(h, p))$ for all $g, h \in G$ and $p \in M$.

(2) if e is the identity of G, then $f(e, p) = p$ for all $p \in M$.

When we say "G acts on M," we mean that there is a group action of G on M. To emphasize that a group action is a model for symmetry transformations, we define mappings $f_g : M \to M$ for all $g \in G$ such that $f_g(p) = f(g, p)$ for all $p \in M$. The mappings f_g are symmetry transformations, and with this notation the axioms of group actions are (1) $f_{g \cdot h} = f_g \circ f_h$ and (2) $f_e = id_M$, the identity mapping of M. These clearly show that all the mappings f_g form a group under a composition of mappings. With the concept of group action, we can formalize the notion of symmetry as follows:

Definition 2.3. Let a group G act on a set M by action f and let $\sim$ be an equivalence relation on M. Then we say that M is *G-symmetric with respect to* $\sim$ if $f_g(p) \sim p$ holds for all $p \in M$, $g \in G$.

Finally, we give some definitions of group actions that will be useful later in this thesis.

Definition 2.4. Let $f : G \times M \to M$ be a group action of G on M and p a point of M. Then the subset $\{f_g(p) \in M \mid g \in G\}$ of M is called the *orbit of* p, and it is denoted by Gp.

The orbit of p is thus the set of points where p is mapped by a group action or by symmetry transformations. All the points of M belong to some orbit, and two different orbits have no common points. Thus orbits are equivalence classes of points of M, and each group action induces an equivalence relation for points, the orbit relation. With the concept of orbit, the G-symmetry with respect to $\sim$ means that all the points of an orbit are equivalent with respect to $\sim$. Notice that the orbit relation and $\sim$ are not the same in general: it possible that points from different orbits are equivalent with respect to $\sim$.

The mapping $g \mapsto f_g$ is a structure-preserving mapping from G to the group of symmetry transformations: $g \cdot h \mapsto f_g \circ f_h$ holds for all $g, h \in G$ by the first axiom of the group action. This mapping, in general, need not be an isomorphism: many elements of G may be mapped to the identity mapping of M. However, we consider only the cases where the two groups are isomorphic. Then we identify the mapping f_g with g and use the shorthand notation gp for $f_g(p)$. If the identification is possible, we call the action effective:

Definition 2.5. A group action $f : G \times M \to M$ *effective* if for any two distinct elements g, h of G there is a point p of M such that $f_g(p) \neq f_h(p)$. The action is *free* if for any two distinct $g, h \in G$ and all $p \in M$ we have $f_g(p) \neq f_h(p)$. The action is *transitive* if for any two $p, q \in M$ there exists a $g \in G$ such that $f_g(p) = q$.

Intuitively, if an action is effective, every non-identity element of the group "moves" at least one point of M, whereas a free action is such that all the points are "moved." Notice that a free action is also an effective action. Transitivity of an action means that all the points of M belong to a single orbit.

2.2.2 Instances of symmetry

Because it is about something remaining the same under some transformations, symmetry can be used to characterize invariances and redundancies. Let us first study the case of invariance which is fundamental for physics. For example, we require that the laws of physics be invariant under the displacement of the observer. That is, for example the laws of electromagnetics are the same everywhere on the Earth. Thus it does not matter where we do experiments, because we should always deduce the same laws. This means that the basic laws of electromagnetism should be written with mathematical structures that take this invariance into account. Formally, this means that the equations describing the laws (Maxwell's equations) do not change under some group action on the underlying space where the equations are written. For example, if the space is modeled as Euclidean space E^3, the equations should be invariant (the same), at least, under the translations of the space.

In addition, the conservation laws of physics are closely related to symmetries. For example, if the *action* or the integral of the *Lagrangian* [13] of a mechanical system is symmetric under continuous translations in space and time, these symmetries account for the conservation laws of linear momentum and energy within the system, respectively. In electromagnetism, assuming time as an independent parameter, the conservation of electric charge is related to symmetries under continuous translations in time. In mathematics, *Noether's theorem* [47] [48] is the basic result that connects differentiable symmetries/invariances and conservation laws.

Invariance can also result in redundancies, which can be used to simplify things mathematically. For example, to characterize the object in Figure 2.1, one only needs to specify the positions of one red and one blue square, and then the rest of the figure can be constructed with symmetry transformations. If the top two squares are specified, the symmetry transformation that rotates the figure 180 degrees with respect to its center point will produce the rest of the figure. This is exactly how symmetric electromagnetic BVPs are solved: if a BVP domain (and all the fields defined in that domain) has some invariance (symmetry), the BVP needs to be solved only in a small part of the domain called symmetry cell, and the solution for the whole domain can be constructed with symmetry transformations [6]. The smaller the symmetry cell needed to construct a solution for whole domain, the more we save in time and memory. Continuous symmetries, such as translations and rotations, allow one to reduce the dimension of the problem. Now the symmetry group is so large that only a lower-dimensional subdomain is needed to construct a solution for the whole domain (for details of this topic, dimensional reduction, see chapter 5).

All redundancies do not result of invariances such as the above where objects have repetitions; rather they result from descriptions of physical theory with structures that contain excess degrees of freedom. For example, in classical electromagnetics, vector potentials are not unique even if their curls are the same: two potentials $\mathbf{A}_1$ and $\mathbf{A}_2$ that differ only in a gradient field ∇f, or $\mathbf{A}_2 = \mathbf{A}_1 + \nabla f$, define exactly the same magnetic flux. Thus, in terms of definition 2.3, M is the set of all possible potentials $\mathbf{A}_i$, the equivalence relation $\sim$ for potentials is that they have the same curl or $\mathbf{A}_i \sim \mathbf{A}_j$ if $\nabla \times \mathbf{A}_i = \nabla \times \mathbf{A}_j$, and the group G that acts on M is the group of all transformations of potentials of type $\mathbf{A} \mapsto \mathbf{A} + \nabla f$. This symmetry of physical theory with excess degrees of freedom in the mathematical description is called *gauge symmetry* [58]. Another example of gauge symmetries are units of measurements: the description of a physical system can be given equally well in terms of meters as in inches. That is, it does not matter what units are used, because the descriptions correspond to the same physical system.

One type of symmetry, required of a good physical theory, is *covariance* [51]. The idea behind it is that exactly the same physical situation can be described with multiple different coordinate systems. Because the choice of coordinates is arbitrary, it is clear that all the possible coordinate systems must be treated as equal from the physical point of view. In general, all physics is independent of the choice of coordinates, bases, representations, and such. Thus it must be possible to write the laws of physics in a form that does not separate the coordinate systems. Therefore, covariance is the invariance of the *form of laws* under some change of coordinates. If the form of laws is invariant under the general differentiable change of coordinates, we talk about general covariance. Covariance is thus purely a formal property of a theory and hence physically vacuous. Furthermore, generally covariant objects and equations can be written in a coordinate-free manner; i.e., objects and equations can be written without any reference to coordinates.

The possibility to write physical laws without any coordinates, in fact, makes covariance, in the form defined above, quite meaningless. However, it is possible to extend or generalize the idea: coordinate-free formulations of laws are based on manifolds; i.e., coordinate systems are replaced by manifolds. Then the general differentiable change of coordinates is replaced by diffeomorphisms of manifolds. Consequently, generalized covariance can be defined as an invariance of the form of laws under diffeomorphisms.

Chapter 3

Basic differential geometric structures

This chapter introduces the most important mathematical structures of this thesis. These structures arise from a broad mathematical discipline called differential geometry. As a modeling formalism for electromagnetism, differential geometry differs from vector analysis in significant ways, in particular in that most of its structures necessary for electromagnetic modeling are independent of a metric. Furthermore, its tools are suited for all dimensions. To emphasize the conceptual and structural character differential geometry offers to electromagnetic modeling, we first define most of its structures without coordinates and bases.

Before the definitions of the structures, let us look at the central structures in a totality they form. This is depicted in Figure 3.1, which also shows some of the essential relations between the structures. The figure shows the hierarchy of the structures and also gives motivations for them. The totality we pursue is the analysis on manifolds, which gives us the tools to express Maxwell's equations. In a purely mathematical setting, we could start with a set and just give it some topology and thereby specify an instance of a topological space. In general, we add structures to lower-level structures and thus define instances of higher-level structures. Observe that the analysis on manifolds can be defined without metric. The constitutive equations could also just be given as an extra structure for manifold, however they are related to metric structures of manifold (Riemannian manifold and Hodge). Finally, with a Hodge-operator we can make a connection between the analysis on manifolds (metric-independent) and the metric-dependent vector analysis.

Physical modeling requires that the topology (and the set) are connected to observations. To make a connection between model and observations, we use distance measurements and a mathematical structure called metric space.

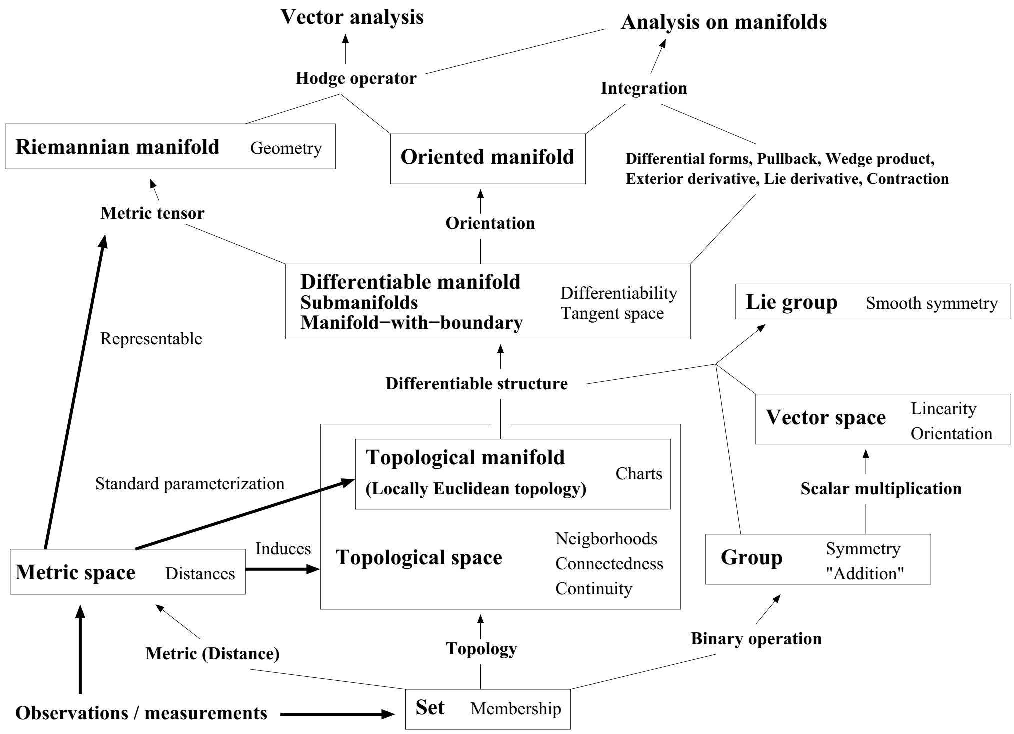

Figure 3.1: Central mathematical structures and their essential relations. Bold texts in boxes are names for the structures and the smaller texts in boxes indicate what motivates the structure or what the structure enables. Texts between boxes with thin arrows are names of the extra structures that the lower-level structures must be endowed to get the higher-level structures. The bold arrows indicate connections between the model and observations.

Each metric space induces a canonical topology for the underlying set as indicated with a bold arrow in Figure 3.1 between the structures. Furthermore, we use manifolds as models for space, and the connection between a manifold and observations is done with distance measurements and so-called standard parameterization (Figure 3.1). Finally, the observation-induced metric is such that it can be represented with a metric tensor (Figure 3.1) and thus Riemannian manifolds model space and its geometry.

3.1 Topological space

Many concepts that relate to space, such as intuitive notions of the connectedness of space, boundary, convergence, and continuity of mappings, are independent of metric concepts such as distance. These are *topological* concepts and depend only on the way space is assembled: in a plain set only the membership of the elements is defined, and it is possible to determine if two elements are identical. However, when the elements are points of a space, intuition calls for the notion of points lying close to each other. Thus there are neighbors and neighborhoods of points, which describe how the space is connected. Neighborhoods of points are described with so-called open sets, which are intuitively sets that have no boundary, or that all the points in an open set are interior points and thus neighbors.

The main subjects in topology are connectedness of sets and continuity of mappings, and they are defined in terms of open sets. A topology for a set M is defined by defining the open sets of M. The open sets constitute a collection of subsets of M, satisfying certain axioms.

Definition 3.1. A *topological space* is a pair $(M, \mathcal{T})$, where M is a set and $\mathcal{T}$ is a collection of subsets of M, called *open sets*, satisfying the following axioms:

(1) the empty set and M belong to $\mathcal{T}$

(2) the union of any collection of sets in T belongs to $\mathcal{T}$

(3) the intersection of any finite collection of sets in $\mathcal{T}$ belongs to $\mathcal{T}$.

The complement $U^c = M \backslash U$ of an open set U is called *closed set*.

Many different topologies can be given for a set. For example, any set M can be endowed with *discrete topology*, in which $\mathcal{T}$ is the collection of all subsets of M. Another topology that can be given for any set M is the *trivial topology*, where $\mathcal{T}$ contains only empty set and the set M itself.

Topology enables us rigorously to describe and define the neighborhoods of points and the connectedness of space:

Definition 3.2. A subset V of a topological space $(M, \mathcal{T})$ is a *neighborhood of point* $p \in V$ if there is an open set U of $\mathcal{T}$ such that $p \in U \subsetneqq V$.

Definition 3.3. A topological space is *connected* if it cannot be divided into two disjoint nonempty closed sets.

The topological space-structure makes it possible to define the continuity of mappings between spaces. Continuous mappings are structure-preserving mappings of a topology, and a topological isomorphism is called homeomorphism.

Definition 3.4. A mapping $f : (M, \mathcal{T}_M) \to (N, \mathcal{T}_N)$ between topological spaces is *continuous* if the inverse image [19] $f^{-1}(U)$ of every open set U of $\mathcal{T}_N$ is an open set of $\mathcal{T}_M$. If f is also bijective and its inverse is continuous, then f is called *homeomorphism.* Topological spaces M and N are called *homeomorphic* if there exists a homeomorphism $f : M \to N$.

Let us next define the subspace topology for subsets of topological spaces. The subspace topology makes a subset a topological space in its own right.

Definition 3.5. Let $A \subset M$ be a subset of a topological space $(M, \mathcal{T})$. The *subspace topology* for A is the topology $\mathcal{T}_A = \{A \cap U \mid U \in \mathcal{T}\}$.

Remark 3.1.1. In the literature, the subspace topology is often called relative topology [11] or induced topology [18].

Topology is defined by open sets. However, in many cases it is not necessary to describe all the open sets but only a subcollection of them such that the other open sets can be constructed from the subcollection. This leads to the concept of basis of a topology.

Definition 3.6. Let $(M, \mathcal{T})$ be a topological space. A subset $\mathcal{B}$ of $\mathcal{T}$ is a *basis* for the topological space $(M, \mathcal{T})$ if every open set of $\mathcal{T}$ can be written as a union of elements of $\mathcal{B}$. A topological space is *second countable* if it has a countable basis.

Remark 3.1.2. A set M is countable if there exists an injective mapping from M to the set of natural numbers $\mathbb{N}$, which are often called counting numbers [1]. Second countable spaces include most "well-behaved" spaces such as Euclidean spaces.

Finally, we define some necessary topological concepts. First, compactness, which makes topological spaces similar in some ways to finite sets:

Definition 3.7. Let $(M, \mathcal{T})$ be a topological space. An *open cover* of M is a collection $\{U_i\}$ of open sets of M such that $M = \cup_i U_i$. A *subcover* of an open cover C of M is a subset of C that is still an open cover of M. M is *compact* if each open cover of M has a finite subcover.

The points of topological space can be distinguished in the topological sense with open sets: two points are indistinguishable if they both always belong or do not belong to a given open set. In other words, points are indistinguishable if they have exactly the same neighborhoods. For example, in the trivial topology, all points are indistinguishable whereas in the discrete topology all points are always distinguishable. Topology thus offers a way to distinguish or separate points, but as the trivial topology shows, not every topology is useful for separation. The ability of a topology to distinguish points should be such that the limits of sequences are unique, which leads to Hausdorff separation [61]:

Definition 3.8. A topological space is a *Hausdorff space* if for every pair of distinct points there exists a pair of disjoint neighborhoods.

3.2 Metric space

The intuitive notion of distance between points is a basic notion of how we perceive the space around us. The corresponding mathematical structure is a metric, which encodes basic qualities of the intuitive notion of distance. Distances have the structure of positive real numbers, or at least every property of distances is also the property of positive real numbers. However, there is no canonical way to relate distances and real numbers without specifying a reference distance (a unit of length). A reference distance is usually specified by some rigid object; consequently, the distance between points corresponds to the number of objects needed to reach a point from another.

Definition 3.9. A *metric space* is a pair (M, dis), where M is a set and $dis : M \times M \to \mathbb{R}$ is a mapping called *distance*, which satisfies the following axioms:

(1) non-negativity: $dis(x, y) \geq 0$ for all $x, y \in M$

(2) identity of indiscernibles: $dis(x, y) = 0$ if and only if $x = y$

(3) symmetry: $dis(x, y) = dis(y, x)$ for all $x, y \in M$

(4) triangle inequality: $dis(x, y) \leq dis(x, z) + dis(z, y)$ for all $x, y, z \in M$.

Example 3.2.1. A basic example of metric spaces is $\mathbb{R}^n$ with its standard metric: let $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ be points of $\mathbb{R}^n$; then the distance between them is given by the formula $\sqrt{(y_1 - x_1)^2 + ... + (y_n - x_n)^2}$.

Next we define isometries, which are the structure-preserving mappings of metric spaces.

Definition 3.10. A mapping $f : (M, dis_M) \to (N, dis_N)$ between metric spaces is an *isometry* if it preserves distances, i.e., if $dis_N(f(x), f(y)) = dis_M(x, y)$ holds for all $x, y \in M$. If f is also bijective, then f is a *metrical isomorphism*. Metric spaces M and N are *isometric* if there exists a metrical isomorphism $f : M \to N$.

Every metric space is also a topological space in a canonical way:

Definition 3.11. Let (M, dis) be a metric space. An *open ball of radius* r *about* p is a subset $B(p, r) = \{q \in M \,|\, dis(p, q) < r\}$ of M. The *metric topology* for (M, dis) is the topology where the set of all open balls is a basis for the topology. In a metric topology, a subset U of M is open if every point of U is contained in some open ball that is contained in U [30].

Remark 3.2.1. Because metric spaces are topological spaces in a canonical way, every property that holds for all topological spaces also holds for metric spaces but not vice versa. Furthermore, a space with a metric topology is always a Hausdorff space [27].

In electromagnetic modeling, distance measurements are important in building our models of space: the points of space we observe are considered distinct only if a nonzero distance between them can be observed. This is reflected in the identity of the indiscernibles axiom of the metric. Thus the points in the set M of the model correspond to observed points. Of course, we usually idealize this by assuming measurements of arbitrary small distances. These measurements give the point set M, but they also equip M with a metric space structure. Then the metric is used to define a topology on M. Finally, the metric of M and the corresponding distance measurements are major tools we use to construct manifolds, which are our preferred models for space.

3.3 Vector space and its orientation

Vector spaces, also known as linear spaces, are very important in physics and mathematics. In differential geometry, they and their structure-preserving mappings, i.e., linear mappings, play a fundamental role. This thesis provides only the basic definitions of this broad subject. The emphasis in this section is more on an additional structure of vector spaces, i.e., orientation.

The orientation of a vector space is a choice between two classes of bases. On the other hand, the intuitive interpretation of orientation in one-dimensional spaces is a positive or right direction to move and is much used in physics and particularly in electromagnetics.

Definition 3.12. Let the elements of a *field* [40] $\mathbb{F}$ be called *scalars*. A *vector space over* $\mathbb{F}$ or $\mathbb{F}$-*vector space* is a set V together with two binary operations,

- *vector addition*: $V \times V \to V$, denoted by $v + w$ for $v, w \in V$, and

- *scalar multiplication*: $\mathbb{F} \times V \to V$, denoted by αv for $\alpha \in \mathbb{F}$, $v \in V$,

satisfying the following axioms for arbitrary elements $v, w \in V$ and scalars $\alpha, \beta \in \mathbb{F}$:

(1) Vector addition establishes a commutative group

(2) Scalar multiplication distributes over the vector and field addition:

  (i) $\alpha(v + w) = \alpha v + \alpha w$
  (ii) $(\alpha + \beta)v = \alpha v + \beta v$

(3) Scalar multiplication is compatible with field multiplication: $\alpha(\beta v) = (\alpha\beta)v$

(4) Scalar multiplication is invariant under the identity of field multiplication: $1v = v$.

Definition 3.13. A mapping $f : (V, +_V) \to (W, +_W)$ between vector spaces over $\mathbb{F}$ is *linear* if it preserves the vector space structure, i.e., if $f(\alpha v +_V \beta w) = \alpha f(v) +_W \beta f(w)$ holds for all $v, w \in V$ and $\alpha, \beta \in \mathbb{F}$. If f is also bijective, then f is a *linear isomorphism*. Vector spaces V and W are *isomorphic* if there exists a linear isomorphism $f : V \to W$.

The elements of a vector space are called vectors. Each vector space has a basis or a set of vectors that allows us to write all the other vectors as a *linear combination* of the basis vectors:

Definition 3.14. A subset $W = \{v_1, ..., v_n\}$ of a vector space V is a *basis* if

(1) W is linearly independent, i.e., $\alpha_1 v_1 + ... + \alpha_n v_n = 0$ holds if and only if $\alpha_1 = ... = \alpha_n = 0$,

(2) W spans V, i.e., for every $w \in V$ there are scalars $\alpha_1, ..., \alpha_n$ such that $w = \alpha_1 v_1 + ... + \alpha_n v_n$ holds.

Notice that a basis is just a set of vectors without any particular order in them. However, often the order of basis vectors is also important, and then we talk about an ordered basis or *frame*. The number of elements in all possible bases of a given vector space is the same, and that number is called the *dimension* of the vector space [42]. Particularly, finite-dimensional vector spaces over $\mathbb{F}$ with the same dimension are always isomorphic [25]. Furthermore, each ordered basis defines a unique isomorphism of an n-dimensional real vector space to $\mathbb{R}^n$. Thus once an ordered basis is selected, every vector can be identified uniquely with an n-tuple of real numbers [25].

The ordered bases of a vector space can be divided into two equivalence classes. Let V be a real vector space, and let $B_1 = (v_1, ..., v_n)$ and $B_2 = (w_1, ..., w_n)$ be two ordered bases for V. Then each w_i can be represented as a linear combination of vectors $(v_1, ..., v_n)$: $w_i = \alpha_{i1} v_1 + ... + \alpha_{in} v_n$. Thus there is a unique linear isomorphism $A : V \to V$, defined by the matrix $A_{ij} = \alpha_{ij}$ that maps the basis B_1 to the basis B_2 (in the above sense). Bases B_1 and B_2 are deemed equivalent if the *determinant* [25] of A is positive. Because the determinant of an isomorphism cannot be zero [35], there are precisely two equivalence classes of ordered bases, and these classes are called orientations. With this equivalence relation, we can define orientation:

Definition 3.15. The two equivalence classes of ordered bases of a real vector space are its *orientations*. An *oriented vector space* is a vector space with one of its orientations.

In other words, orientation is a choice of a "privileged" or "positive" or "direct" equivalence class of ordered bases for a vector space. A basis that belongs to the positive class is called positively oriented, and similarly a basis belonging to the negative class is called negatively oriented. The standard orientation of $\mathbb{R}^n$ means that the standard basis of $\mathbb{R}^n$ is positively oriented, and in the case of $\mathbb{R}^3$ this is considered the right-handed choice.

3.4 Euclidean space

The basic model of space in elementary physics and numerical modeling is the Euclidean space. In general, Euclidean spaces can be of any positive dimension and they are metric spaces where *Euclidean geometry* holds. Formally, the n-dimensional Euclidean space E^n is defined as an n-dimensional affine space with a special metric [7]. Furthermore, E^n and $\mathbb{R}^n$ with its standard metric can be identified isometrically. However, this identification is not canonical but involves an arbitrary choice of perpendicular "coordinate axes" for E^n [5].

Definition 3.16. Let M be a set and V an n-dimensional vector space. M is an *n-dimensional affine space* if there exists a free and transitive group action of V on M.

Notice that because V is a group under the addition of vectors, it makes sense to talk about a group action of a vector space. Now an affine space can be geometrically interpreted as follows: the elements of M are points of space, and if $p \in M$ and $v \in V$, then vp is a point where the point p is translated by v. Thus each vector v induces a translation of the points of M, and because the action is transitive, all points can be translated to all other points. This means that all points belong to the same orbit; therefore, all points in an affine space are equivalent: there is no special point such as the origin in $\mathbb{R}^n$. Thus there is a V-symmetry in the affine space, and the equivalence relation is simply that two points are equivalent if there is a translation between them. A space where all points are equivalent is called a *homogeneous space*, which again emphasizes that there are no special or privileged points [7]. Finally, notice that the addition of vectors of V corresponds to a composition of translations, and that it does not make sense to add points of M.

To make an affine space M a Euclidean space, we need to define a metric d that is compatible with the group action in the sense that each translation is an isometry. That is, a compatible metric dis must satisfy $dis(p, q) = dis(vp, vq)$ for all $p, q \in M$ and $v \in V$. This can be done by defining an *inner product* (see Definition 3.61 or [25][35]) for V, in which case the *norm* [66] of v is the distance between points p and vp.

Definition 3.17. An n-dimensional Euclidean space E^n is an n-dimensional affine space with a metric dis such that the translations are isometries.

Often, however, E^n just means $\mathbb{R}^n$ with its standard metric. This is reasonable when coordinates are needed, but in a purely geometrical setting the affine space model is preferable. Euclidean spaces are important and popular in modeling because they are intuitive and form the stage for *standard calculus* [61], which is a basic tool of physics.

3.5 Manifolds

Manifolds are generalizations of Euclidean spaces in the sense that manifolds can be globally complex though locally they look like an affine space. Manifolds are also locally topologically like Euclidean spaces, but the metrical properties of Euclidean spaces are ignored. This locally Euclidean property makes it possible to cover manifolds locally with coordinates, thus making arithmetic available. Furthermore, the standard calculus of Euclidean spaces can be generalized to manifolds without a metric, i.e., metric properties can be totally separated from differential calculus. Finally, manifolds are here defined abstractly, meaning that they are not embedded in a possibly higher-dimensional Euclidean space, as they are often defined in literature.

Manifolds will be our models for space because they offer the most generic coordinate-free model for space. Particularly, in this thesis, electromagnetics is presented with a formulation that realizes generalized covariance. Moreover, dimensional reduction can be explained in a coordinate- and metric-free manner. Finally, the metric aspects of electromagnetics can be separated from topological and differential aspects.

3.5.1 Topological manifold and charts

Manifolds are defined such that there *exists* a local parameterization with coordinates or tuples of real numbers. In other words, manifolds can be locally covered with coordinate systems; thus the points of the manifold can be parameterized or labeled with coordinates. This is why they are locally like affine spaces and said to be locally Euclidean: every point of a manifold has a neighborhood homeomorphic to a subset of Euclidean space, but globally such a homeomorphism is not necessary. To ensure some nice properties, manifolds are also required to be Hausdorff spaces with a second countable topology.

Definition 3.18. Let $(M, \mathcal{T})$ be a topological space. M is *locally Euclidean of dimension* n if every point has a neighborhood homeomorphic to an open subset of the Euclidean space $\mathbb{R}^n$. A homeomorphism ϕ from a connected open set $U \subset M$ to an open subset of $\mathbb{R}^n$ is called a *chart*. A chart $\phi : U \subset M \to \mathbb{R}^n$ is often denoted by a pair (U, ϕ).

Definition 3.19. A *topological manifold* M of dimension n is a topological space that satisfies the following axioms:

(1) M is locally Euclidean of dimension n

(2) M is Hausdorff

(3) M is second countable.

Remark 3.5.1. The Hausdorff-property makes manifolds more like Euclidean spaces that are Hausdorff spaces, but also it ensures that the limits of sequences are unique. Second countable manifolds are metrizable [62], or a metric exists for the manifold such that its induced metric topology agrees with the topology of the manifold.

Remark 3.5.2. When we in the next section define differentiable manifolds as topological manifolds with an extra structure called differentiable structure, the Hausdorff and second countable properties assure two useful features: a differentiable manifold that is a Hausdorff space with a countable basis can be embedded in a higher-dimensional Euclidean space [64], and it admits a particular system of real-valued functions called *partition of unity* [12].

Example 3.5.1. An example of a two-dimensional topological manifold is a surface such as the sphere. A surface and one of its charts is given in Figure 3.2.

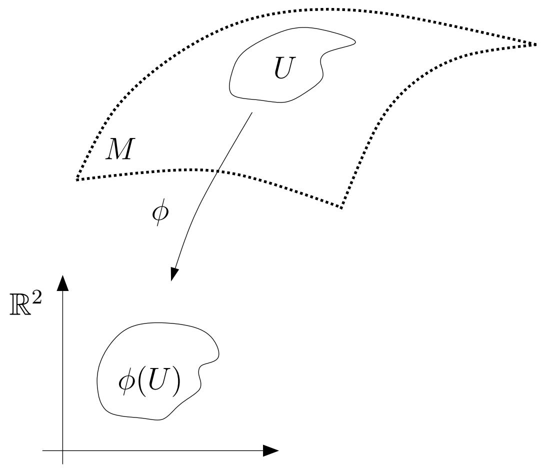

Figure 3.2: Topological manifold and a chart. The figure shows a topological manifold M, which is a surface. U is an open set of M, and ϕ is a chart that maps U to $\mathbb{R}^2$.

Because topological manifolds are topological spaces with extra topological properties (axioms (1)-(3) in the definition), their structure-preserving mappings are homeomorphisms. Moreover, a topological manifold is connected or compact if the topology of the manifold makes it connected or compact, respectively.

The topological manifold M we use to model space is in practice constructed with a chart (or charts). But this chart, in turn, is constructed with the help of the standard metric of $\mathbb{R}^n$ and distance measurements with some *rigid body*: the distance measurements give us the point set M and its metric topology. Of course, M is also a metric space with a metric dis in the sense of Definition 3.9. We then label the points we observe with coordinates or tuples of real numbers such that all the distances measured with a rigid body are the same when calculated with the standard metric of $\mathbb{R}^n$. In other words, the points of M are labeled with coordinates such that the distance $dis(p, q)$ between any two points $p, q \in M$ is the same as $\sqrt{(p_1 - q_2)^2 + ... + (p_n - q_n)^2}$ (the distance between their coordinates). More generally, the distances between points and the corresponding coordinates must be the same only up to some scalar multiple. For example, we may use meters for measurements, but in the chart the calculated distances correspond to inches. Consequently, charts constructed with rigid body measurements are not necessarily isometries though they preserve the shapes. Particularly, the images of every sphere in M under these charts are spheres also in $\mathbb{R}^n$, when measured with the standard metric of $\mathbb{R}^n$. We call these charts *standard parameterizations.* An example of a standard parameterization is given in Figure 3.3.

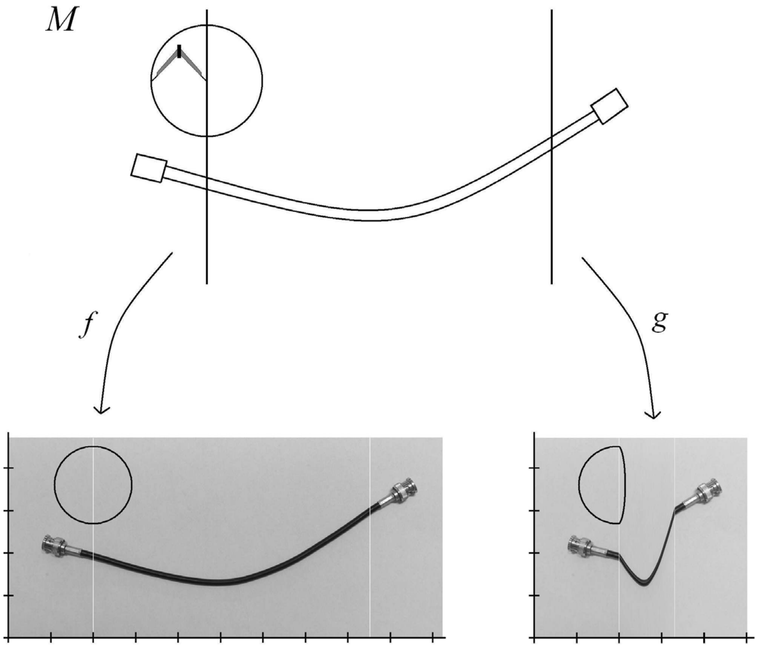

Figure 3.3: Standard parameterization: the line drawing on top refers to a real coaxial cable and represents also a topological manifold M. The rigid body we used in distance measurements gave us a pair of dividers and enabled us to specify spheres. Chart f is a standard parameterization, whereas chart g is not, because all the images of spheres in M are not spheres in $\mathbb{R}^n$ in the sense of the standard metric of $\mathbb{R}^n$.

Definition 3.20. Let M be a topological manifold with a metric space structure given by distance $dis_r : M \times M \to \mathbb{R}$, which corresponds to distance measurements with some rigid body, and let $\phi : U \subset M \to \mathbb{R}^n$ be a chart of M. If there exists $\alpha > 0$ such that $dis_r(p, q) = \alpha\, dis_s(\phi(p), \phi(q))$ holds for all $p, q \in U$, where dis_s is the standard metric of $\mathbb{R}^n$, then ϕ is a *standard parameterization.*

3.5.2 Differentiable manifolds

Because the differentiability of mappings is defined in Euclidean spaces, and because manifolds are locally Euclidean, it is possible to extend the concept of differentiability to mappings between manifolds. At first glance, this seems straightforward: select charts (U, ϕ) of an m-dimensional manifold M and (V, φ) of an n-dimensional manifold N. The (local) representation of a mapping $f : M \to N$ under charts ϕ and φ is a mapping $\varphi \circ f \circ \phi^{-1}$, which maps $\phi(U) \subset \mathbb{R}^m$ to $\varphi(f(U)) \subset \mathbb{R}^n$. Then the mapping f is deemed differentiable in $U \subset M$ if its representation in these charts is differentiable in the classical sense. However, if we use another chart (V, ψ) of N, we can change the charts with the *transition map* or change of chart mapping $\psi \circ \varphi^{-1}$; then the differentiability of f is defined by $\psi \circ \varphi^{-1} \circ \varphi \circ f \circ \phi^{-1} = \psi \circ f \circ \phi^{-1}$. Now this need not be differentiable anymore because transition maps in a topological manifold are homeomorphisms, and thus guaranteed to be only continuous. Thus the differentiability of a function depends on the chosen charts. To make the definition of differentiability independent of the choice of charts, we must restrict the set of admissible charts: all the charts in a set of admissible charts are required to be compatible in the sense that their transition maps are appropriately differentiable.

Definition 3.21. Two charts (U_1, ϕ_1) and (U_2, ϕ_2) of a topological manifold M are *C^r-compatible* if the non-emptiness of $U_1 \cap U_2$ implies that their transition map $\phi_2 \circ \phi_1^{-1}$ is r times continuously differentiable.

Admissible or compatible charts constitute a differentiable structure and a topological manifold together with such a structure is then a differentiable manifold:

Definition 3.22. A *C^r-differentiable structure for M* is a family $\mathcal{D} = \{(U_i, \phi_i)\}$ of the charts of M such that

(1) $\cup_i U_i$ is a cover of M

(2) for any i, j the charts (U_i, ϕ_i) and (U_j, ϕ_j) are C^r-compatible

(3) if (U_i, ϕ_i) is C^r-compatible with every other chart in $\mathcal{D}$, then (U_i, ϕ_i) is itself in $\mathcal{D}$.

Remark 3.5.3. In the above definition, (1) assures that the whole M can be covered with charts of $\mathcal{D}$, (2) assures that all the charts in $\mathcal{D}$ are compatible, and (3) means that $\mathcal{D}$ is a maximal collection of charts with respect to counts (1) and (2).

Definition 3.23. A *differentiable manifold* M of class C^r is a pair $(M, \mathcal{D})$, where M is a topological manifold and $\mathcal{D}$ is a C^r-differentiable structure for M.

Example 3.5.2. A simple example of a differentiable manifold is $\mathbb{R}^n$ with its standard differentiable structure defined as the maximal collection (in the sense of Definition 3.22) containing the chart $(\mathbb{R}^n, i)$, where i is the identity map of $\mathbb{R}^n$. Another example is a finite-dimensional vector space V with its natural differentiable structure: select a basis for V, in which case the basis defines a linear isomorphism from V to $\mathbb{R}^n$. This isomorphism is then a chart that generates the differentiable structure, which is, furthermore, independent of the choice of basis [62].

Example 3.5.3. Product manifolds [62]: Let $(M, \mathcal{D}_M)$ and $(N, \mathcal{D}_N)$ be differentiable manifolds of dimension m and n, respectively. Then $M \times N$ is an $(m+n)$-dimensional differentiable manifold if it is given a differentiable structure that is the maximal collection containing

$$\{(U_i \times V_j, \phi_i \times \varphi_j : U_i \times V_j \longrightarrow \mathbb{R}^m \times \mathbb{R}^n) \mid (U_i, \phi_i) \in \mathcal{D}_M, (V_j, \varphi_j) \in \mathcal{D}_N\}.$$

Notice that it is possible to give many different C^r-differentiable structures on the same topological manifold; therefore, a differentiable structure is truly an additional structure that must be specified [5] [62]. However, if a topological manifold is coverable with a single chart, then given one such chart will uniquely specify the differentiable structure (axiom (3) in the definition). In cases where the manifold cannot be covered with one chart, e.g., the spheres of any dimension, the differentiable structure can be specified by a finite covering of charts called atlas:

Definition 3.24. A *C^r-atlas for M* is a family $\mathcal{A} = \{(U_i, \phi_i)\}_{i \in A}$ of charts of M such that

(1) $\cup_{i \in A} U_i$ is a cover of M

(2) for any $i, j \in A$ the charts (U_i, ϕ_i) and (U_j, ϕ_j) are C^r-compatible.

Remark 3.5.4. A differentiable structure is a maximal atlas.

Two atlases for a topological manifold can define the same differentiable manifold, and when they do, they are called equivalent:

Definition 3.25. C^r-atlases $\mathcal{A}_1 = \{(U_i, \phi_i)\}_{i \in A}$ and $\mathcal{A}_2 = \{(U_j, \phi_j)\}_{j \in B}$ for M are *equivalent*, denoted by $\mathcal{A}_1 \sim \mathcal{A}_2$, if their union $\mathcal{A}_1 \cup \mathcal{A}_2 = \{(U_i, \phi_i)\}_{i \in A \cup B}$ is a C^r-atlas for M.

Any C^r-atlas $\mathcal{A}$ for M uniquely specifies the C^r-differentiable structure $\mathcal{D}$ for M and the equivalence class $[\mathcal{A}]$ of all C^r-atlases is naturally comparable to the differentiable structure $\mathcal{D}$: the atlases in the equivalence class $[\mathcal{A}]$ contain exactly the same charts as the differentiable structure $\mathcal{D}$. From this point on, we do not explicitly mention the class of differentiability unless needed; otherwise it is implicitly assumed to be sufficiently high.

Differentiable mappings between manifolds are the structure-preserving mappings of the differentiable manifold structure, and they are defined next. The isomorphisms are called diffeomorphisms:

Definition 3.26. Let $f : M \to N$ be a continuous mapping between differentiable manifolds $(M, \mathcal{D}_M)$ and $(N, \mathcal{D}_N)$. The mapping f is *differentiable* if its representation $\phi \circ f \circ \psi^{-1}$ for some charts $\psi \in \mathcal{D}_M$ and $\phi \in \mathcal{D}_N$ is differentiable in the classical sense. If f is also bijective such that its inverse is also differentiable, then f is a *diffeomorphism*. Manifolds M and N are *diffeomorphic* if there exists a diffeomorphism $f : M \to N$.

Remark 3.5.5. The definition of differentiability of mappings is independent of the choice of chart used to check the differentiability [34]. This is a direct consequence of the definition of differentiable structure.

Next we describe a useful way to induce a differentiable manifold structure for a topological space from a differentiable manifold via a homeomorphism.

Definition 3.27. Let $f : N \to M$ be a homeomorphism from a topological space N to a differentiable manifold M. Furthermore, let $\mathcal{A} = \{(U_i, \phi_i)\}$ be an atlas of M. Then the *pullback atlas* $f^*\mathcal{A}$ for N is defined by

$$f^*\mathcal{A} = \{(f^{-1}(U_i), \phi_i \circ f)\}.$$

The pullback preserves the equivalence of atlases, i.e., if $\mathcal{A}_1 \sim \mathcal{A}_2$ holds, then $f^*\mathcal{A}_1 \sim f^*\mathcal{A}_2$ also holds. Consequently, the differentiable structure can be also pulled back. A pullback atlas gives a useful and equivalent way to characterize diffeomorphisms and diffeomorphic manifolds:

Proposition 3.1. Let $(M_1, \mathcal{A}_1)$ and $(M_2, \mathcal{A}_2)$ be two differentiable manifolds. If there exists a homeomorphism $f : M_1 \to M_2$ such that $\mathcal{A}_1 = f^*\mathcal{A}_2$, then f is a diffeomorphism and the manifolds are diffeomorphic.

Finally, a remark about embedding a differentiable manifold in a higher-dimensional Euclidean space. According to Whitney's theorem, every n-dimensional differentiable manifold can be smoothly embedded in the Euclidean space R^{2n+1} [64]. Moreover, a historical and still often used way to define manifolds is to consider them special subsets of a (usually) higher-dimensional Euclidean space [32]. However, this approach assumes that there is some space around the manifold; yet if the manifold itself is the space, what is then the space around that space? Thus as physical models of space, manifolds assume their most suitable form when they are defined directly via charts and not as part of some larger space. The direct definition that does not rely on embeddings is called intrinsic definition.

3.5.3 Tangent space

The fields we are going to define on manifolds have some linearity properties. For example, differential forms model electromagnetic fields and assign a linear mapping to each point of a manifold. Linearity requires vector space structures associated with manifolds, and thus far there have been none. Fortunately, the differentiable structure gives a manifold so much smoothness that a vector space can be defined at each point of the manifold that "linearly approximates" the manifold around the points: the vectors can be thought of as infinitesimal displacements of the points. This vector space is called tangent space, and its elements are called tangent vectors. Furthermore, tangent space is defined with the help of charts but in a manner independent of the choice of chart.

The literature gives many equivalent definitions for tangent space and tangent vectors [34][62][58]: the same mathematical structure is defined using different but isomorphic realizations. We use the "geometric definition" [34] and define tangent vectors as equivalence classes of smooth curves on a manifold. The equivalence of the curves intuitively means that at some point they have the same direction and speed. In other words, under the geometric definition, curves can be thought of as trajectories of objects moving on a manifold, and if two curves are equivalent at some point, the objects have the same velocity at that point (Figure 3.4). Notice that this definition is intrinsic to manifolds.

Definition 3.28. Let M be a differentiable manifold. A *curve on M* is a differentiable mapping $c : \mathbb{R} \to M$, and we say that c is a curve through point $p \in M$ if $c(0) = p$ holds.

Definition 3.29. Two curves c_1 and c_2 through point $p \in M$ are *equivalent* if in some chart ϕ the representations $\phi \circ c_i : \mathbb{R} \to \mathbb{R}^n$ of the curves have the same derivative at $p = c_i(0)$, or if $(\phi \circ c_1)'(0) = (\phi \circ c_2)'(0)$ holds.

The set of all equivalence classes of curves through p is denoted by $T_p(M)$. If $dim(M) = n$, then the set $T_p(M)$ can be given the structure of an n-dimensional vector space: let ϕ be a chart containing a neighborhood of p.

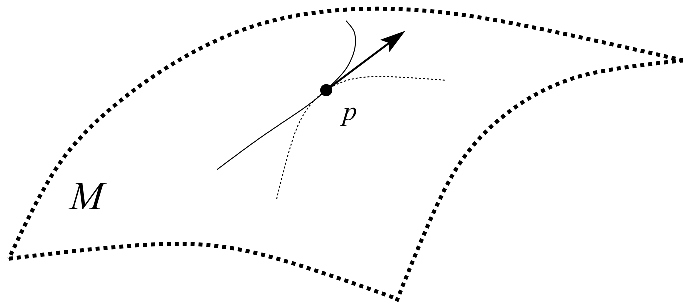

Figure 3.4: Tangent vector. Two equivalent curves on a manifold M that pass through point p. The arrow at p represents the equivalence class, and it can be thought of as the velocity vector at p of objects moving along trajectories defined by the curves.

Then the map $(d\phi)_p : T_p(M) \to \mathbb{R}^n$ defined by $(d\phi)_p([c]) = (\phi \circ c)'(0)$ is bijective, and we can thus transfer the vector space structure of $\mathbb{R}^n$ to $T_p(M)$ by requiring that $(d\phi)_p$ be a linear isomorphism. This induced structure is canonical in the sense that it is independent of the choice of chart.

Next we outline the proof that $(d\phi)_p$ is bijective, and that the induced vector space structure is independent of the choice of chart: $(d\phi)_p$ is injective by the definitions of equivalent curves, and surjectivity can be shown by constructing a suitable curve (class) for each element of $\mathbb{R}^n$. Independence of choice of chart follows from (1) $(\phi \circ c)'(0) = (\phi \circ \varphi^{-1} \circ \varphi \circ c)'(0) = (\phi \circ \varphi^{-1})'(\varphi(c(0))) \cdot (\varphi \circ c)'(0)$, where the last equality is due to the *chain rule* [61], and (2) $\varphi(c_1(0)) = \varphi(c_2(0))$ holds for all curves c_1 and c_2 through p.

Definition 3.30. An equivalence class of $T_p(M)$ is called a *tangent vector at* p, and the vector space $T_p(M)$ of all tangent vectors is called the *tangent space of* p.

Remark 3.5.6. Tangent spaces of different points are not related in any obvious way and it makes no sense to talk about addition of tangent vectors from different tangent spaces. However, because the dimension of all tangent spaces is the same as the manifold, all tangent spaces are isomorphic to each other but, of course, not canonically. The set of all tangent vectors of a manifold M is denoted by $T(M)$.

Remark 3.5.7. Because the tangent spaces of all the points of the Euclidean space $\mathbb{R}^n$ are canonically identified with the vector space $\mathbb{R}^n$, all the tangent spaces are canonically isomorphic.

A vector field on a manifold is the first example of a field over a manifold in which tangent spaces are needed. A vector field on M is a mapping from M to $T(M)$ such that each point $p \in M$ is mapped to one of the tangent vectors of $T_p(M)$ [5][62].

A differentiable mapping $f : M \to N$ between manifolds maps the curves of M to those of N. Thus for each such differentiable mapping, there is a related mapping that maps the tangent vectors of M to those of N. This mapping is called the pushforward of f (Figure 3.5).

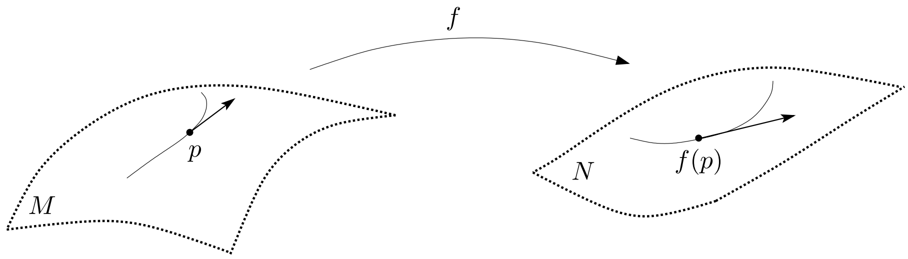

Figure 3.5: Pushforward. A differentiable mapping f maps a manifold M to a manifold N; thus f also maps curves on M to curves on N. Then the pushforward of f at $p \in M$ maps equivalence classes of curves at p (tangent vectors at p) to equivalence classes of curves at $f(p) \in N$.

Definition 3.31. Let $f : M \to N$ be a differentiable mapping between manifolds. Then its *pushforward* is a mapping $f_* : T(M) \to T(N)$ such that at each point $p \in M$ the mapping is the linear map $f_*(p) : T_p(M) \to T_{f(p)}(N)$ such that $f_*(p)([c]) = [f \circ c]$ holds for all $[c] \in T_p(M)$.

Because the pushforward generalizes the concept of *differential* [18] [61] of mappings $\mathbb{R}^m \to \mathbb{R}^n$, the mapping f_* is also called the differential of f. Furthermore, for a real valued mapping over a manifold $f : M \to \mathbb{R}$, we often denote the differential (or push forward) of f by df. The pushforward of a composition of mappings is a composition of pushforwards [5]: If $f : M \to N$ and $g : N \to O$, then $(g \circ f)_* = g_* \circ f_*$.

3.5.4 Oriented manifolds

Oriented manifolds are manifolds in which tangent spaces have a positive and coherent orientation. Orientation is particularly important for the definitions of integration over manifolds and for the so-called Hodge-operator.

Definition 3.32. Let M be an n-dimensional differentiable manifold. A family $\{\mathcal{O}_p\}_{p \in M}$ of orientations of the tangent spaces is *locally coherent* if around every point of M there is an *orientation-preserving chart*, i.e., a chart (U, ϕ) with the property that for every $q \in U$ the pushforward $\phi_*(q)$ takes the orientation $\mathcal{O}_q$ to the usual orientation of $\mathbb{R}^n$.

Locally coherent orientations for a manifold together with the manifold itself define the oriented manifold:

Definition 3.33. An *orientation* of a manifold is a locally coherent family $\{\mathcal{O}_p\}_{p \in M}$ of orientations of tangent spaces. An *oriented manifold* is a pair $(M, \mathcal{O})$, where M is a manifold and $\mathcal{O}$ is an orientation of M.

Lastly, we define the mappings that preserve orientations:

Definition 3.34. A pushforward $f_*(p)$ is *orientation preserving* if its determinant is positive.

Remark 3.5.8. The determinant of a linear map is defined as the determinant of the matrix that represents the linear map, and this is well defined because the result is independent of the choice of representation matrix [25].

Definition 3.35. A diffeomorphism $f : M \to N$ between oriented manifolds is *orientation-preserving* (orientation-reversing) if for every $p \in M$ the pushforward $f_*(p)$ is orientation-preserving (orientation-reversing).

3.5.5 Submanifolds

Subsets of manifolds, particularly lower-dimensional subsets, which are manifolds themselves, are indispensable for describing the physics of electromagnetics. For example, the electromotive force is something that can be measured, and it relates to paths or one-dimensional subsets of space whereas the magnetic flux relates to surfaces or two-dimensional subsets. Moreover, lower-dimensional subsets are central for dimensional reduction. Subsets of a manifold that are also manifolds themselves are called submanifolds (see Figure 3.6). The literature contains many types and definitions of submanifolds, but the following are the most useful for us:

Definition 3.36. Let $f : N \to M$ be a differentiable mapping between manifolds N and M. If the rank of $f_*(p)$ [42] equals $dim(N)$ for each $p \in N$, then f is an *immersion*. If f is an injective immersion, then (N, f) is a *submanifold* of M. If f is an immersion that is also a homeomorphism to its image $f(N)$ with the subspace topology, then f is an *embedding* and (N, f) is *embedded submanifold of* M. If N is a subset of M and if f is the inclusion map of N to M which is also an embedding, then (N, f) is a *regular submanifold of* M.

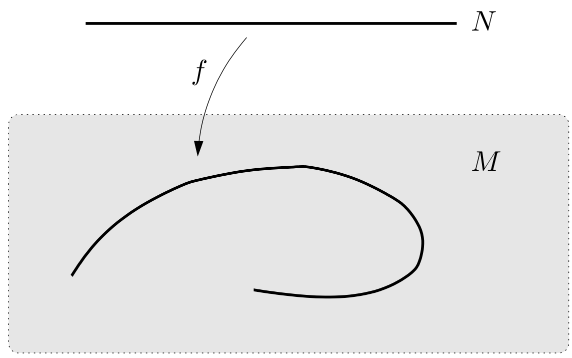

Figure 3.6: Example of a submanifold. A one-dimensional manifold N is embedded to a two-dimensional manifold M with a mapping f. Thus (N, f) is an embedded submanifold of M.

Remark 3.5.9. Technically, a submanifold need not be a subset of a bigger manifold. However, the definitions of submanifolds contain two parts, a manifold N and a mapping f of that manifold to a bigger manifold M. Then N can be thought of as a parameter space for the subset $f(N)$, which is also a manifold. In the literature [5][62], also the term *imbedding* is used instead of *embedding*.

Remark 3.5.10. In case of submanifolds, the differentiable and topological structures for the image $f(N)$ are induced from N via mapping f whereas in case of embedded or regular submanifolds the image $f(N)$ inherits its structures from M. In general this means that for submanifolds the topology of the image $f(N)$ may be finer (i.e. it contains more open sets) than the subspace topology. Thus embedded and regular submanifolds are special cases of submanifolds.

3.5.6 Manifolds-with-boundary

In electromagnetics, engineering problems are often formulated as BVPs. Manifolds are the domains of these BVPs, and, as the name suggests, an essential part of BVPs are boundary values. However, because the manifolds we have defined so far cannot have boundaries, we need an extended definition of manifolds that includes the concept of boundary. Such manifolds are simply called manifolds-with-boundary. They are like n-dimensional manifolds, i.e., locally Euclidean but at their boundary points like half-spaces of the n-dimensional Euclidean space. Thus the only difference in definition with manifolds is that the differentiable structure contains charts of a different kind.

Let us first define the half-spaces H^n that are used in the definition: $H^n = \{(x_1, ..., x_n) \in \mathbb{R}^n | x_n \geq 0\}$ is a subset of $\mathbb{R}^n$, equipped with the subspace topology.

Definition 3.37. An n-dimensional *manifold-with-boundary* is a topological space M that is a Hausdorff space with a countable basis of open sets and a differentiable structure $\mathcal{D}$, defined in the following way: $\mathcal{D} = \{U_i, \phi_i\}$ is a family of charts, where U_i is a connected open set of M, and ϕ_i is a homeomorphic mapping of U_i to an open subset of H^n such that

(1) $\cup_i U_i$ is a cover of M

(2) the charts of $\mathcal{D}$ are compatible

(3) $\mathcal{D}$ is maximal with respect to properties (1) and (2).

Remark 3.5.11. Compatibility of charts is defined as in 3.21, but now the change of chart mappings are mappings from subsets of H^n to H^n.

The boundary of H^n is the set $\partial H^n = \{x = (x_1, ..., x_n) \in \mathbb{R}^n | x_n = 0\}$. With ∂H^n, we can define ∂M, the boundary of a manifold-with-boundary M:

Definition 3.38. Let M be a manifold-with-boundary. A point $p \in M$ is a *boundary point* if it is mapped by a chart around p to a boundary point of H^n. The boundary ∂M of M is the set of all boundary points of M.

Remark 3.5.12. A differentiable manifold given in Definition 3.23 is a special case of manifold-with-boundary but without boundary points.

The boundary ∂M is an $(n-1)$-dimensional submanifold of an n-manifold-with-boundary M. Furthermore, all the concepts of manifolds, such as diffeomorphism and tangent space, generalize directly to manifolds-with-boundary [5][34]. The following result is useful for later purposes:

Lemma 3.1. If $f : M \to N$ is a diffeomorphism between manifolds-with-boundary, its restriction to the boundary ∂M is a diffeomorphism $f|_{\partial M} : \partial M \to \partial N$. [34]

3.6 Lie groups

A Lie group is a group that is also a differentiable manifold such that the two structures are compatible. Because compatibility makes group operations smooth mappings, Lie groups can be characterized as "smooth groups."

In this thesis, Lie groups are needed for dimensional reduction where the symmetry groups are Lie groups: dimensional reduction is based on continuous or smooth symmetries, and this is reflected in the manifold structure of Lie groups.

Definition 3.39. A *Lie group* is a set G, which has both a group structure and a differentiable manifold structure such that the structures are compatible in the sense that the group operations

$$\begin{aligned} (g, h) \in G \times G \mapsto g * h \in G \\ g \in G \mapsto g^{-1} \in G \end{aligned}$$

are smooth mappings.

The dimension of a Lie group is the dimension of the manifold; similarly, a Lie group is connected or compact if the manifold is connected or compact, respectively. Before any examples of Lie groups, we define the structure-preserving mappings and isomorphism of Lie groups:

Definition 3.40. Let $f : M \to N$ be a mapping between Lie groups. f is a *Lie group homomorphism* if f is a group homomorphism that is also a smooth mapping between the manifolds. If f is also a diffeomorphism, then f is a *Lie group isomorphism.*

There are only two connected one-dimensional Lie groups up to isomorphism, [45] and the basic representatives of these classes serve also as good examples of Lie groups: $(\mathbb{R}, +, i)$, i.e., the set of real numbers, where the group structure is the addition of the numbers, and where the manifold structure is given by a chart that is the identity mapping. Henceforth, this Lie group is denoted simply by $\mathbb{R}$, and it is isomorphic to the Lie group of all translations of a vector space in one direction. The basic representative of the other class is S^1, which is the multiplicative group of all complex numbers with the absolute value 1. This Lie group is isomorphic to the Lie group of all rotations around an axis. Notice that $\mathbb{R}$ is a non-compact Lie group whereas S^1 is compact. Geometrically, $\mathbb{R}$ corresponds to the line and S^1 to the circle.

A product of two Lie groups is again a Lie group if the differentiable structure is chosen to be the product structure, and the group structure is the direct product structure [62]. The only connected two-dimensional Lie groups up to isomorphism are the products of the line and circle: $\mathbb{R}^2 = \mathbb{R}\times\mathbb{R}$, $\mathbb{R} \times S^1$, and $S^1 \times S^1$, which geometrically correspond to the plane, cylinder, and torus, respectively [45].

3.7 Foliations of manifolds

Sometimes it is convenient if a manifold can be decomposed into a disjoint union of lower-dimensional submanifolds. For example, it is convenient to decompose a four-dimensional spacetime manifold to 3-dimensional submanifolds, which are spatial spaces at each time moment. Such decompositions are called foliations of a manifold, and the submanifolds are called the leaves of the foliation. The basic property of foliations is that there are special charts such that the images of the leaves are coordinate isovalue (hyper)surfaces (see Figure 3.7). This property gives the manifold a local product structure, or in a sufficiently small neighborhood of every point, the manifold is diffeomorphic to a product manifold.

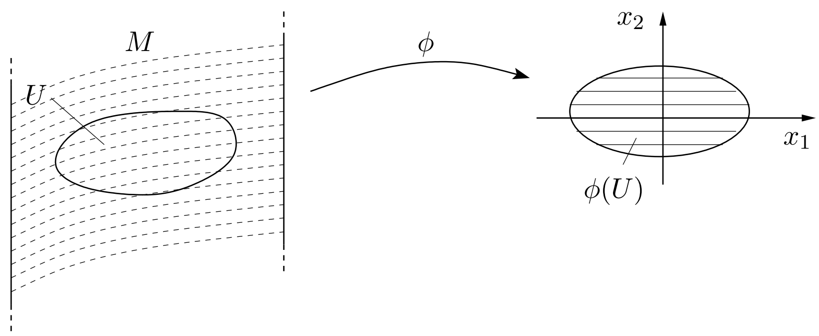

Figure 3.7: Example of a foliation. A 1-dimensional foliation of a 2-manifold M. The dotted lines on M represent the leaves of the foliation and their images under the chart ϕ are the lines $x_2 = constant$.

Definition 3.41. Let M be an n-dimensional manifold. A *p-dimensional foliation* of M is a decomposition of M into a union of disjoint connected subsets $\{L_\alpha\}_{\alpha \in A}$, called the *leaves* of the foliation, with the following property: every point of M has a chart (U, ϕ) such that the components of $U \cap L_\alpha$ are described in the chart by the equations $x_{p+1} = constant$,..., $x_n = constant$.

Remark 3.7.1. Every leaf of $\{L_\alpha\}$ is a p-dimensional regular embedded submanifold of M. Furthermore, the smoothness of a foliation depends on the class of differentiability of the differentiable structure of the manifold [41].

Remark 3.7.2. If a Lie group G acts smoothly on a manifold M, and if its action is free, the orbits of G define a smooth foliation of M; i.e., the orbits are the leaves of the foliations [41].

3.8 Analysis on manifolds

Classical vector analysis is often applied multivariable and multivalue calculus in three-dimensional spaces. Furthermore, most structures are defined or expressed with respect to the metric of the space. In this section, we present a very general multivariable and multivalue—and totally metric-free—calculus based on differentiable manifolds of any dimension. However, instead of vector fields, this calculus employs differential forms. In this thesis, differential forms are used to model electric and magnetic fields.

The calculus enables many metric-free and dimension-independent operators and structures. For example, the differential operators grad, curl, and div are replaced by a single differential operator for differential forms, called the exterior derivative. It is defined for all dimensions and forms of all degrees. Moreover, the so-called pullback of differential forms generalizes the "change of coordinates" for manifolds. A particularly convenient feature of the pullback is its natural compatibility with many essential operators and structures of this general calculus.

Differential forms are fields in a differentiable manifold, which assign to each point an antisymmetric multilinear mapping of tangent vectors. Therefore, the exposition begins with the tangent bundle, which is the set of all tangent vectors with a canonical manifold structure. Finally, Bamberg's and Sternbeg's book [2] is a good mathematical introduction to differential forms and their analysis on $\mathbb{R}^n$ with applications in physics.

3.8.1 Tangent bundle

The set of all tangent vectors $T(M)$ of a differentiable manifold $(M, \mathcal{D})$ can be represented as a disjoint union of tangent spaces indexed with the points of M: $T(M) = \cup_{p \in M} T_p(M)$. There is a natural projection $\pi : T(M) \to M$, which maps tangent vectors to their points or $\pi(v) = p$ for all $v \in T_p(M)$. Notice that the inverse image of p under π is $T_p(M)$. Then with the help of the pushforward and natural projection, the set $T(M)$ can be given a natural differentiable manifold structure induced by $\mathcal{D}$: for each $(U, \phi) \in \mathcal{D}$, we define a chart $\tilde{\phi} : \pi^{-1}(U) \to \mathbb{R}^{2n}$ by $\tilde{\phi}(v) = (\phi(\pi(v)), \phi_*(v))$ for all $v \in \pi^{-1}(U)$. With these charts, $T(M)$ can be shown to be a differentiable manifold with twice the dimension of M [62].

Definition 3.42. The set $T(M)$ together with the natural differentiable structure induced by M is called the *tangent bundle*.

The points of a tangent bundle $T(M)$ can be written as pairs (p, v), where $p \in M$ and $v \in T_p(M)$. A *section* of $T(M)$ is a mapping $\sigma : M \to T(M)$

such that $\pi(\sigma(p)) = p$ for all $p \in M$; i.e., the section selects exactly one tangent vector from each tangent space of M [5]. In the literature, sections are also called cross-sections [58] or lifts [62]. Notice that vector fields on M are sections of $T(M)$. Furthermore, because $T(M)$ and M are differentiable manifolds and sections are mappings between them, *smooth* vector fields can be defined in a coordinate-free way:

Definition 3.43. Let M be a differentiable manifold. A *smooth vector field* on M is a smooth section of the tangent bundle $T(M)$.

3.8.2 Cotangent bundle

The tangent space $T_p(M)$ of a point p is a vector space and has thus a *dual space* [35][40], which in this context is called cotangent space. The elements of the cotangent space of point p are real-valued linear mappings over $T_p(M)$.

Definition 3.44. A *covector at p* is a linear mapping $c : T_p(M) \to \mathbb{R}$.

The set of all covectors at p is denoted by $T_p^*(M)$, and it can be given the structure of vector space. The addition and scalar multiplication are defined elementwise: the sum $c_1 + c_2$ of covectors c_1 and c_2 is defined by $(c_1 + c_2)(v) = c_1(v) + c_2(v)$, and the scalar multiplication αc_1 of c_1 by α is defined by $(\alpha c_1)(v) = \alpha c_1(v)$. With this vector space structure, $T_p^*(M)$ is isomorphic to $T_p(M)$ but *not canonically.* Thus $T_p^*(M)$ has the same dimension as M.

Definition 3.45. With the vector space structure defined above, $T_p^*(M)$ is called the *cotangent space at p.*

The set of all cotangent spaces over M is denoted by $T^*(M)$, and it can be given a bundle structure in a similar fashion as we defined the tangent bundle. The set $T^*(M)$ of a differentiable manifold $(M, \mathcal{D})$ is a disjoint union of cotangent spaces indexed with the points of M: $T^*(M) = \cup_{p \in M} T_p^* M$. There is a natural projection $\pi^* : T^*(M) \to M$, which maps covectors to their points or $\pi^*(c) = p$ for all $c \in T_p^*(M)$. Notice that the inverse image of p under π^* is $T_p^*(M)$. Then with the help of the pushforward and natural projection, the set $T^*(M)$ can be given a natural differentiable manifold structure induced by $\mathcal{D}$: let $(e_1, ..., e_n)$ denote the *canonical basis* [5] of the Euclidean space $\mathbb{R}^n$. Then for each $(U, \phi) \in \mathcal{D}$, we define a chart $\tilde{\phi}^* : (\pi^*)^{-1}(U) \to \mathbb{R}^{2n}$ by $\tilde{\phi}^*(c) = (\phi(\pi^*(c)), c(\phi_*^{-1}(e_1)), .., c(\phi_*^{-1}(e_n))$ for all $c \in (\pi^*)^{-1}(U)$. With these charts, $T^*(M)$ can be shown to be a differentiable manifold with twice the dimension of M [62].

Definition 3.46. The set $T^*(M)$ together with the natural differentiable structure induced by M is called the *cotangent bundle*.

Similarly as for the tangent bundle, the points of a cotangent bundle $T^*(M)$ can be written as pairs (p, c), where $p \in M$ and $c \in T_p^*(M)$. Also *sections* of $T^*(M)$ can be defined as in the case of the tangent bundle: a mapping $\sigma^* : M \to T^*(M)$ such that $\pi^*(\sigma^*(p)) = p$ holds for all $p \in M$ is a section; i.e., the section selects exactly one covector from each cotangent space of M. Thus covector fields on M are sections of $T^*(M)$:

Definition 3.47. Let M be a differentiable manifold. A *smooth covector field* or *one-form* on M is a smooth section of the cotangent bundle $T^*(M)$.

The set of all one-forms over M, or the set of all smooth sections of $T^*(M)$, is denoted by $\Omega^1(M)$. This set can be given the structure of a vector space by pointwise definition: let ω_1 and ω_2 be one-forms over M and $\alpha \in \mathbb{R}$. The sum $\omega_1 + \omega_2$ and the scalar multiplication $\alpha\omega_1$ are defined by the formulae $(\omega_1 + \omega_2)_p(v) = (\omega_1)_p(v) + (\omega_2)_p(v)$ and $(\alpha\omega_1)_p(v) = \alpha(\omega_1)_p(v)$, which hold for all $p \in M$, $v \in T_p(M)$.

3.8.3 Differential forms

A differential form over M is a field that assigns to each point p an anti-symmetric multilinear mapping that maps tangent vectors of $T_p(M)$ to real numbers. The suitability of differential forms as models for electromagnetic fields is well established (see, e.g., [13]). Here we give only the formal definition and in such a way that first forms are defined at one point, or that the definitions of the linear mappings that map tangent vectors to reals are given, and the definition is then extended to the whole manifold.

Before the definition, some preliminary definitions and notations: if V is a vector space, its k-fold Cartesian product by itself is denoted by V^k, e.g., $V^3 = V \times V \times V$. An element v of V^k is a family $v = (v_i | i \in I)$, where the index set I is the set $\{1, 2, ..., k\}$. A *permutation* of I is any bijective mapping $\sigma : I \to I$, and $\sigma(v)$ denotes the family $(v_{\sigma(i)} | i \in I)$. A *transposition* σ is a special kind of permutation that only interchanges two elements; i.e., there exist indices i, j such that $\sigma(i) = j$, $\sigma(j) = i$, and $\sigma(n) = n$ for all other indices n. Clearly all permutations are compositions of transpositions. A permutation is *even (odd)* if it is the composition of an even (odd) number of transpositions. Define *sgn* as the mapping that maps even permutations to 1 and odd permutations to -1. Finally, if $(v_1, ..., v_n)$ is an ordered basis of V, then $(v_{\mu_1}, ..., v_{\mu_k})$ is a k-tuple of the basis vectors $(v_1, ..., v_n)$ such that $1 \leq \mu_1 < \mu_2 < ... < \mu_k \leq n$ (for more on permutations, see [14][25]). With these preliminaries, we can define the properties of linear mappings that are assigned to each point of a manifold:

Definition 3.48. Let V and W be vector spaces over $\mathbb{R}$. A mapping $f : V^k \to W$ is *k-linear* if it is linear on each of the arguments separately. If $f(\sigma(v)) = f(v)$ holds for all $v \in V^k$ and for all permutations $\sigma(v)$, then f is *symmetric*. If $f(\sigma(v)) = sgn(\sigma)f(v)$ holds for all $v \in V^k$ and for all permutations $\sigma(v)$, then f is *antisymmetric*.

The antisymmetry of k-linear mappings has the following important consequence:

Proposition 3.2. Let $f : V^k \to W$ be k-linear and antisymmetric. If $v_1, ..., v_k$ is a set of linearly dependent vectors of V, then $f(v_1, ..., v_k) = 0$. [34]

The set of all antisymmetric k-linear mappings over V is an $\binom{n}{k}$-dimensional vector space if V is n-dimensional [34]; $\binom{n}{k}$ is the *binomial coefficient* [61]. When a k-linear mapping is assigned to a point p of a manifold M, then the role of V^k is played by $(T_p(M))^k$. Thus differential forms are defined as fields of antisymmetric k-linear mappings over a manifold:

Definition 3.49. A *differential k-form* ω on a manifold M is a mapping that assigns to each point $p \in M$ an antisymmetric k-linear mapping $\omega_p : (T_p(M))^k \to \mathbb{R}$.

The smoothness of a differential k-form can be defined in a coordinate-free way [5]: Given any $X_1, ..., X_k$ smooth vector fields on M, a differential k-form ω is smooth if $\omega(X_1, ..., X_k)$, defined by $\omega(X_1, ..., X_k)(p) = \omega_p(X_1(p), ..., X_k(p))$, is a smooth function in M.

We often use a shorthand *k-form* for the *differential k-form*. Notice that zero-forms correspond to real valued mappings on M. Furthermore, if $k > n = dim(M)$, all k-forms are zero. The set of all differential k-forms in M is denoted by $\Omega^k(M)$, and the set of all differential forms of any degree is denoted by $\Omega(M)$. $\Omega^k(M)$ is an infinite dimensional vector space for each k, and the addition and scalar multiplications are given pointwise.

3.8.4 Pullback

The *change of variables* of differential forms is the pullback of a differential form under a differentiable mapping between manifolds. The pullback is a similar induced mapping for differential forms as the pushforward of tangent vectors, but it goes in the opposite direction: if $f : M \to N$ is differentiable, then its pullback f^* is a mapping of the type $\Omega^k(N) \to \Omega^k(M)$. The pullback is the change of variables because if we have a form ω in N, then its pullback $f^*\omega$ is a form defined in M. In other words, ω is a function of the points of N and the tangent vectors of $T(N)$, but its pullback $f^*\omega$ is a function of the points of M and the tangent vectors of $T(M)$. The change of variables is defined by f, which gives the relation between the points of M and N and by the pushforward f_*, which gives the relation between the tangent vectors of $T(M)$ and $T(N)$. Thus the pullback is a generalization of the change of coordinates for differential forms on manifolds.

Definition 3.50. Let $f : M \to N$ be a differentiable mapping. The *pullback* under f is a linear mapping $f^* : \Omega^k(N) \to \Omega^k(M)$ for each $k \geq 0$ such that for all $\omega \in \Omega^k(N)$ the pullback $f^*\omega$ of ω at every $p \in M$ is defined by the formula $(f^*\omega)_p(v_1, ..., v_k) = \omega_{f(p)}(f_*v_1, ..., f_*v_k)$ where $v_1, ..., v_k \in T_p(M)$.

Notice that the dimensions of M and N need not be the same, and that the dimension of M may be higher or lower than the dimension of N. This generalizes the change of coordinates that are always defined between spaces of equal dimension. Furthermore, if f is a change of coordinate mapping, then f^* defines the change of coordinates for ω [5]. Finally, notice that because the pullback goes backwards, the pullback of a composition of mappings has a reverse order [5]: if $f : M \to N$ and $g : N \to O$, then $(g \circ f)^* = f^* \circ g^* : \Omega^k(O) \to \Omega^k(M)$.

With submanifolds, the pullback of the inclusion map offers a way to define restrictions of differential forms to submanifold: let N be a submanifold and a subset of M, and let i denote the inclusion map of N to M. Then the restriction $\omega|_N$ of $\omega \in \Omega(M)$ to N is defined by $\omega|_N = i^*\omega$. Particularly, the restriction $\omega|_{\partial M}$ to the boundary ∂M of a manifold-with-boundary M is an important example of restriction [34], because the boundary values of a BVP are defined with it. Often the restriction $\omega|_N$ is denoted by $t\omega$ and called the *trace* of ω.

3.8.5 Wedge product

Differential geometry has an important operator for differential forms called the wedge product. It is metric-independent and replaces in vector analysis metric-dependent vector field operators' cross-product, dot product, and scalar triple product. The wedge product is defined for all forms in all dimensions. When the set $\Omega(M)$ of all differential forms in M is endowed with the wedge product, it has an algebraic structure called *graded algebra* [40]. Particularly, $\Omega(M)$ together with the wedge product constitutes so-called *ex-*

terior algebra [62] a.k.a. *Grassmann algebra*. The wedge product offers a simple way to construct new (possibly higher degree) forms from old ones.

Definition 3.51. The *wedge product* is a bilinear mapping $\wedge : \Omega(M) \times \Omega(M) \to \Omega(M)$ satisfying the following properties:

(1) the $\wedge$ is associative, i.e., $(\omega \wedge \eta) \wedge \xi = \omega \wedge (\eta \wedge \xi)$.

(2) the $\wedge$ is graded anticommutative, i.e., $(\omega \wedge \eta) = (-1)^{kl}(\eta \wedge \omega)$ for $\omega \in \Omega^k(M)$ and $\eta \in \Omega^l(M)$.

(3) the zero-form $1 : M \to 1$ satisfies $1 \wedge \omega = \omega$ for all $\omega \in \Omega(M)$.

The following proposition states that the wedge product is uniquely defined by above axioms [65]. Furthermore, the proposition shows how to calculate pointwise the values of the wedge product of two differential forms in terms of those forms. For this calculation, we need a special class of permutations of indexes: let $P(k,l)$ denote the set of all such permutations σ of the index set $\{1, ..., k+l\}$ that have the following property: $\sigma(1) < ... < \sigma(k)$ and $\sigma(k+1) < ... < \sigma(k+l)$.

Proposition 3.3. The wedge product given in Definition 3.51 is uniquely defined. Furthermore, if ω and η are a k-form and an l-form, respectively, their wedge product $\omega \wedge \eta$ is a $(k+l)$-form that is defined pointwise by the following formula [34]:

$$(\omega \wedge \eta)_p(v_1, .., v_{k+l}) = \sum_{\sigma \in P(k,l)} sgn(\sigma)\omega_p(v_{\sigma(1)}, .., v_{\sigma(k)})\eta_p(v_{\sigma(k+1)}, .., v_{\sigma(k+l)})$$

Example 3.8.1. If ω and η are one-forms on M, then $\omega \wedge \eta$ is a two-form defined by the formula

$$(\omega \wedge \eta)_p(v_1, v_2) = \omega_p(v_1)\eta_p(v_2) - \omega_p(v_2)\eta_p(v_1).$$

Remark 3.8.1. The wedge product of zero-forms ω and η is the pointwise product: $(\omega \wedge \eta)_p = \omega_p \, \eta_p$.

The wedge product is natural in the sense that it is compatible with the pullback:

Theorem 3.1. The wedge product is compatible with the pullback, i.e., $f^*(\omega \wedge \eta) = f^*\omega \wedge f^*\eta$ holds [34][62].

3.8.6 Exterior derivative

The differential operators gradient, curl, and divergence of vector analysis are defined for scalar and vector fields in a three-dimensional space and are metric-dependent. Differential geometry has a metric-independent differential operator for differential forms, which generalizes the gradient, curl, and divergence into a single operator. It is defined for manifold of any dimension and differential forms of any degree, and the definition has the same form for all these cases. This operator is the exterior derivative, and it has a crucial role because with it we can write Maxwell's equations.

Definition 3.52. The *exterior derivative* is a linear mapping $d : \Omega^k(M) \to \Omega^{k+1}(M)$ for each $k \geq 0$ satisfying the following axioms:

(1) Differential property: for $f \in \Omega^0(M)$, df is the differential of f.

(2) Complex property: $d \circ d = 0$,

(3) Product property: $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k(\omega \wedge d\eta)$ for $\omega \in \Omega^k(M)$.

Proposition 3.4. The exterior derivative is uniquely defined by three axioms of definition 3.52 [5][17][62].

Remark 3.8.2. The exterior derivative is not defined for all forms but only for those that are at least once differentiable.

Remark 3.8.3. Generalized Stokes's theorem for integration of differential forms (section 3.8.9) provides a connection between the boundary of a manifold and the exterior derivative d. The complex property of d is then related to the fact that the boundary of a boundary is empty [46], which corresponds to the complex property of the boundary operator.

The exterior derivative is a natural operator for differential forms in the sense that it is compatible with the pullback:

Theorem 3.2. The exterior derivative is compatible with the pullback, i.e., $d \circ f^* = f^* \circ d$ holds [17][34][62].

Remark 3.8.4. The compatibility of d and the pullback means that differential equations expressed with d, such as $d\omega = \eta$, are generally covariant. In other words, differential equations are invariant under diffeomorphisms. For electromagnetism, this implies that Maxwell's equations are invariant under diffeomorphisms. This is important for symmetry applications and enables definition of the equivalence of electromagnetic BVPs under diffeomorphisms.

The axioms of the exterior derivative shown in Definition 3.52 together with its compatibility with the pullback define the *complex* [40] called *de Rham complex* [34], i.e., the following sequence of linear maps having the properties (1)-(3) of Definition 3.52:

$$0 \longrightarrow \Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \Omega^2(M) \xrightarrow{d} ...$$

Furthermore, the exterior derivative is the only linear map on differential forms that constitute the de Rham complex [34]. The naturality property of d makes the pullback f^* of $f : M \longrightarrow N$ a *chain map* [30] between the de Rham complexes of N and M; i.e., the following diagram commutates:

$$\begin{array}{ccccccccc} 0 \longrightarrow & \Omega^0(N) & \xrightarrow{d} & \Omega^1(N) & \xrightarrow{d} & \Omega^2(N) & \longrightarrow ... \\ & f^*\big\downarrow & & f^*\big\downarrow & & f^*\big\downarrow & \\ 0 \longrightarrow & \Omega^0(M) & \xrightarrow{d} & \Omega^1(M) & \xrightarrow{d} & \Omega^2(M) & \longrightarrow ... \end{array}$$

Maxwell's equations can be written with the exterior derivative without the metric structure of M. In vector analysis, Maxwell's equations are written with the gradient, curl, and divergence, which are the metric counterparts of the exterior derivative. However, because the three depend on the metric structure of M, they cannot yet be defined.

3.8.7 Contraction and extension

In dimensional reduction, we must move from a higher-dimensional manifold to a lower-dimensional submanifold and vice versa. This decreasing and increasing dimension often means that also the degree of a form must be decreased and increased. The coordinate- and metric-free tools for forms that meet these requirements are contraction and extension.

Definition 3.53. Let X be a vector field on a manifold M. Then X defines a linear mapping $i_X : \Omega^k(M) \to \Omega^{k-1}(M)$ for each $k \geq 1$, called *contraction*, which maps a k-form ω to a $(k-1)$-form $i_X\omega$. i_X is defined pointwise by the formula

$$(i_X\omega)_p(v_1, .., v_{k-1}) = \omega_p(X_p, v_1, .., v_{k-1}),$$

which must hold for all $p \in M$ and $v_1, .., v_{k-1} \in T_p(M)$.

Definition 3.54. Let α be a one-form on a manifold M. Then α defines a linear mapping $I_\alpha : \Omega^k(M) \to \Omega^{k+1}(M)$ for each $k \geq 0$, called *extension*, which maps a k-form ω to a $(k+1)$-form $I_\alpha\omega$ defined by $I_\alpha\omega = \alpha \wedge \omega$.

Remark 3.8.5. The contraction of differential forms is often called the interior product, and extension is sometimes called the exterior product. The contraction of differential forms is related to tensor contraction.

Proposition 3.5. The contraction $i_X : \Omega(M) \to \Omega(M)$ has the following properties, which are easily proven [5][62]:

(1) The mapping $X \mapsto i_X\omega$ is a linear map for fixed ω

(2) Product property: $i_X(\omega \wedge \eta) = i_X\omega \wedge \eta + (-1)^k(\omega \wedge i_X\eta)$ for $\omega \in \Omega^k(M)$

(3) Antisymmetry: $i_X(i_Y\omega) = -i_Y(i_X\omega)$

(4) If $Y = fX$ for a scalar field f, then $i_X \circ i_Y = 0$.

Remark 3.8.6. The linearity of i_X and $X \mapsto i_X\omega$ is, in fact, so-called $C^\infty(M)$-linearity: $i_X(\omega + f\eta) = i_X\omega + f(i_X\eta)$, where $f \in C^\infty(M)$ is a smooth scalar field over M. In pointwise terms, this reduces to normal linearity.

Finally a lemma, with several uses later on, about the relationship between contraction and pullback:

Lemma 3.2. Let $f : M \to N$ be a diffeomorphism and let X be a vector field on M. Then $f^* \circ i_{f_*X} = i_X \circ f^*$.

Proof: Let ω be a $(k+1)$-form on N; then

$$\begin{aligned}(f^* i_{f_*X}\omega)_p(v_1, ..., v_k) &= (i_{f_*X}\omega)_p(f_*v_1, ..., f_*v_k)\\ &= \omega_p(f_*X, f_*v_1, ..., f_*v_k)\\ &= (f^*\omega)_p(X, v_1, ..., v_k)\\ &= (i_X f^*\omega)_p(v_1, ..., v_k)\end{aligned}$$

holds for all $p \in M$ and for all $v_1, ..., v_k \in T_p(M)$. $\square$

3.8.8 Lie derivative

The Lie derivative of a differential form is a generalization of the standard directional derivative of functions $\mathbb{R}^m \to \mathbb{R}$ to forms: the direction is specified by a smooth vector field over a manifold in contrast to the standard case in which it is specified by a vector or equivalently by a constant vector field over $\mathbb{R}^m$. The Lie derivative of a k-form is again a k-form, and in the case of zero-forms over $\mathbb{R}^m$, the Lie derivative with respect to a constant vector field reduces to the standard directional derivative. The Lie derivative plays a central role in dimensional reduction, and we use it also to represent the time or temporal derivatives of fields.

Before any formal definitions, we give a geometric characterization of the Lie derivative. Think about a flow of incompressible fluid on some manifold. Its velocity vectors at each point of the manifold form a smooth vector field. As time passes, the fluid particles move along trajectories, and the velocity vectors are tangent to these trajectories. Thus a vector field X induces a curve c for each point p such that the vector X_p is the velocity vector of c at p. These curves are called integral curves, and they fill the manifold. The totality of integral curves induces a 1-parameter group of transformations ϕ_t of the manifold: for each t the mapping ϕ_t maps the fluid particles at time 0 to the points where they are at time t. Now the fluid flow moves a differential form ω at $\phi_0(p) = p$ to some other point $\phi_t(p)$ along the trajectory, and we can pullback ω at $\phi_t(p)$ to the point p, where ω and $\phi_t^*\omega$ can be compared by subtracting ω from $\phi_t^*\omega$. This difference is divided by t, and the limit as t approaches to zero is the value of the Lie derivative of ω with respect to X at point p.

Definition 3.55. Let X be a smooth vector field on a manifold M, and let $c : I \to M$ be a smooth curve on M, where $I \subset \mathbb{R}$ is an open interval. Then c is an *integral curve* of X if the image of the canonical basis vector e of $\mathbb{R}$ under the pushforward of c for each t is the vector $X_{c(t)}$, or $c_*(t)(e) = X_{c(t)}$ holds for all $t \in I$.

If X is a smooth vector field on M, then for each $p \in M$, there exists an integral curve c for X such that $I = (-\epsilon, \epsilon)$ for some $\epsilon > 0$ and $c(0) = p$ [62]. Next, the fluid flows (trajectories) are modeled as group actions of the Lie group $\mathbb{R}$:

Definition 3.56. Let $\varphi : \mathbb{R} \times M \to M$ be a group action of $\mathbb{R}$ on M. φ is a *1-parameter group of transformations* of M if $\varphi_t : M \to M$, where $\varphi_t(p) = \varphi(t, p)$ is a diffeomorphism for each $t \in \mathbb{R}$. Let $I_\epsilon = (-\epsilon, \epsilon) \subset \mathbb{R}$ and U be an open subset of M. A *local 1-parameter group of transformations* of M is a mapping $\varphi : I_\epsilon \times U \to M$ such that

1) φ_t is a diffeomorphism of U onto the open set $\varphi_t(U)$ of M for each $t \in I_\epsilon$

2) if $t, s, t + s \in I_\epsilon$ and if $p, \varphi_s(p) \in U$, then $\varphi_{t+s}(p) = \varphi_t(\varphi_s(p))$.

Remark 3.8.7. A local 1-parameter group of transformations is also known as *flow* [5]. Each 1-parameter group of transformations is a local 1-parameter group of transformations for which $I_\epsilon = \mathbb{R}$ and $U = M$. Furthermore, the set of all transformations φ_t for a local 1-parameter group of transformations is not really a group despite its name. However, it behaves like a group near the identity mapping (or near $0 \in \mathbb{R}$) [37].

Each 1-parameter group of transformations φ induces a vector field X for M: for each $p \in M$, X_p is a vector tangent to the curve $c(t) = \varphi(t,p)$ which is an integral curve of X. Similarly, each local 1-parameter group of transformations induces a smooth vector field for U, and the following proposition also states that the converse is true.

Proposition 3.6. Let X be a smooth vector field on a manifold M. For each point p of M, there exists a neighborhood U of p, a positive number ϵ, and a local 1-parameter group of transformations $\varphi_t : U \to M$, $t \in I_\epsilon$, which induces the given X [37].

Now we have all the tools to define the Lie derivative:

Definition 3.57. Let X be a vector field on M and φ_t a local 1-parameter group of transformations induced by X. The *Lie derivative* of a k-form ω with respect to X is a k-form denoted by $\mathcal{L}_X\omega$, which is defined pointwise for each $p \in M$ by the formula

$$(\mathcal{L}_X\omega)(p) = \lim_{t\to 0} \frac{(\varphi_t^*\omega)(p) - \omega(p)}{t}.$$

The next theorem shows that the Lie derivative can be expressed using the exterior derivative and contraction:

Theorem 3.3. (Cartan's magic formula) $\mathcal{L}_X = d \circ i_X + i_X \circ d$ [62].

3.8.9 Integration on manifolds and Stokes's theorem

With charts, the standard *Lebesgue integration theory* [1][18] of mappings $\mathbb{R}^n \to \mathbb{R}$ can be generalized to integration of differential forms over manifolds: we represent a form as a real valued mapping defined over the codomain of a chart that covers part of the manifold and then apply standard integration. Consequently, the *change of variables formula* [61] for standard integration implies that the integral of the form is independent of the choice of chart. Notice that the integration of differential forms over manifold, as defined here, is independent of the metric of the manifold. The definition of integration we present is from [34].

Definition 3.58. A subset A of an n-manifold is called *measurable* (*set of measure zero*) if it has this property relative to charts, i.e., if for some covering of A by charts (U, ϕ) on M, each $\phi(U \cup A)$ is Lebesgue-measurable (a set of measure zero) in $\mathbb{R}^n$. If A is contained in a chart domain, then A is *small*.

Definition 3.59. Let M be an oriented n-manifold and $\omega \in \Omega^n(M)$. ω is *integrable* if for some decomposition $(A_i)_{i\in\mathbb{N}}$ of M into countably many small measurable subsets and some sequence $(U_i, \phi_i)_{i\in\mathbb{N}}$ of orientation-preserving charts (see Definition 3.32) with $A_i \subset U_i$, the following holds: for every $i \in \mathbb{N}$, the component function

$$a_i = \omega((\phi_i^{-1})_*(e_1), ..., (\phi_i^{-1})_*(e_n)) \circ (\phi_i)^{-1} : \phi_i(U_i) \to \mathbb{R}$$

of ω relative to (U_i, ϕ_i) is Lebesgue-measurable on $\phi_i(A_i)$, and

$$\sum_{i=1}^{\infty} \int_{\phi_i(A_i)} |a_i(x)|dx < \infty.$$

Lemma 3.3. Let M be an oriented n-manifold and ω an integrable n-form. Let $(U_i, \phi_i)_{i\in\mathbb{N}}$ and a_i be as in Definition 3.59. Then the value

$$\sum_{i=1}^{\infty} \int_{\phi_i(A_i)} a_i(x)dx$$

is independent of the decomposition and charts [34].

Definition 3.60. Let M be an oriented n-manifold and ω an integrable n-form. Let $(U_i, \phi_i)_{i\in\mathbb{N}}$ and a_i be as in definition 3.59. The *integral* of ω over M, denoted by the $\int_M \omega$, is the value

$$\sum_{i=1}^{\infty} \int_{\phi_i(A_i)} a_i(x)dx.$$

Remark 3.8.8. The integral of k-forms is defined only over k-dimensional manifolds. Thus, e.g., one-forms and two-forms cannot be integrated over the same manifold, and if one- and two-forms are defined over a three-manifold, then their integral can be defined only over one- and two-dimensional submanifolds, respectively.

Next we provide two important theorems about the integration of differential forms on manifolds. The first is an analog for the change of variables formula for integration on manifolds. The second is the Generalized Stokes's theorem for differential forms, which generalizes and unifies *Stokes's and divergence theorems* of vector analysis and the *fundamental theorem of calculus.* It relates integration, the exterior derivative, and the boundary operator to each other. For more on integration and the theorems, see [34].

Theorem 3.4. Let $f : N \to M$ be an orientation-preserving diffeomorphism between n-manifolds and ω an n-form on M. Then the following holds:

$$\int_M \omega = \int_N f^*\omega.$$

Theorem 3.5. (Generalized Stokes's theorem) Let M be an oriented n-dimensional manifold-with-boundary and $\omega \in \Omega^{n-1}(M)$. Furthermore, let $t\omega$ denote the restriction of ω to the boundary ∂M. Then the following holds:

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

3.9 Metric structure for manifolds

A plain differentiable manifold contains no concepts of distance or angle. These geometric notions must be given with an additional structure called the metric tensor, which is a generalization of the *dot product* of vector analysis. A manifold together with a metric tensor is called a semi-Riemannian manifold.

Because the metric tensor is a separate structure on manifolds, it allows a clear separation between those elements of electromagnetic theory that rest on topology and differentiability from those that are related to a metric. So far, all the structures defined on manifolds are independent of the metric tensor. Particularly, all the items on electromagnetic BVPs, such as fields and Maxwell's equations, except for constitutive equations, can be defined without reference to a metric. The relationship between metric and constitutive equations is studied in chapter 4. The metric-independence of most structures necessary in electromagnetic modeling with differential geometry is in contrast to vector analysis, where the metric is from the beginning embedded in most structures.

The metric tensor enables us to define the so-called Hodge-operator, which is often used to express the constitutive equations [8] [56]. Furthermore, a metric provides a connection between differential forms and the vector fields of vector analysis: with a metric we can represent one-forms and $(n-1)$-forms in an n-manifold as vector fields.

3.9.1 Riemannian and semi-Riemannian manifolds

Riemannian manifolds are differentiable manifolds with an additional structure called the metric tensor. A metric tensor is a field over a manifold that assigns an inner product to each tangent space. The inner product, in turn, is a generalization of the dot product. Thus because the inner product gives a norm for tangent vectors, each tangent space is like the n-dimensional Euclidean space with Euclidean geometry. Because tangent vectors can be interpreted as infinitesimal or virtual displacements of a point, in the same manner the geometry of a tangent space can interpreted as a virtual geometry. Virtual Euclidean geometry can be partly extended to the whole manifold: we can define geometric concepts for the manifold, but the resulting geometry is not necessarily Euclidean. Semi-Riemannian manifolds are generalizations of the Riemannian manifolds in the sense that more general objects than inner products are allowed. A special class of the semi-Riemannian manifolds is the Lorentz manifolds which are used in the general relativity.

Definition 3.61. Let V be a real vector space. A *bilinear form* on V is a mapping $\Phi : V \times V \to \mathbb{R}$ that is linear in each variable. Φ is *symmetric* if $\Phi(v, w) = \Phi(w, v)$ holds for all $v, w \in V$. Φ is *nondegenerate* if there are no nonzero $v \in V$ such that $\Phi(v, w) = 0$ holds for all $w \in V$. Φ is *positive definite* (*negative definite*) if $\Phi(v, v) \geq 0$ ($\Phi(v, v) \leq 0$) holds for all $v \in V$, and $\Phi(v, v) = 0$ holds if and only if $v = 0$. An *inner product* is a symmetric positive definite bilinear form.

If Φ is a symmetric definite (positive or negative) bilinear form on V, then for any subspace W of V the restriction $\Phi|_{W\times W}$ on W is also a symmetric definite bilinear form on W [49]. With the help of the restriction we can define an index of a symmetric bilinear form, which then can be used to define the metric tensor:

Definition 3.62. The *index* of a symmetric bilinear form Φ on V is the largest integer that is the dimension of a subspace $W \subset V$ on which $\Phi|_{W\times W}$ is negative definite.

Remark 3.9.1. If the index of Φ is zero then Φ is an inner product.

Definition 3.63. A *metric tensor* is a smooth choice of a nondegenerate symmetric bilinear form on each tangent space of a manifold such that the index of the forms are the same over the manifold.

The smoothness of the metric tensor can be defined as follows [5]: let X and Y be smooth vector fields on M and m be a metric tensor on M. m is smooth if $m(X, Y)$ is a smooth mapping on M.

Definition 3.64. A *semi-Riemannian manifold* is a pair (M, m), where M is a differentiable manifold, and m is a metric tensor on M. If the index of m is 0 or 1, then (M, m) is a *Riemannian manifold* or a *Lorentz manifold*, respectively.

Example 3.9.1. The standard dot product $\cdot$ of the vector space $\mathbb{R}^n$ is an inner product over $\mathbb{R}^n$. Thus $(\mathbb{R}^n, \cdot)$ is a Riemannian manifold.

It makes sense to speak about the pullback of a metric tensor, and the formula in Definition 3.50 is, in fact, the definition of the pullback of a metric tensor [5]. Thus if $f : M \to N$ is a mapping of a differentiable manifold M to a semi-Riemannian manifold (N, n), then f^*n is a metric tensor for M, and (M, f^*n) is a semi-Riemannian manifold. Particularly, the standard Riemannian structure of a chart ($\mathbb{R}^n$) can always be pulled back to the manifold, making it thus a Riemannian manifold. However, the induced metric tensor in that case is chart-dependent. Furthermore, the contraction of a metric tensor with respect to a vector field is defined as for differential forms.

It is clear from the definition of the Riemannian manifold that Riemannian manifolds are not structurally metric spaces in the sense of Definition 3.9. However, the metric tensor induces canonically a metric space structure for manifolds: the length or the norm $||v||$ of a tangent vector $v \in T_p(M)$ is defined by $||v|| = \sqrt{m_p(v, v)}$. Furthermore, a metric tensor enables comparison of the lengths of tangent vectors of different points. Then with the help of the norm of tangent vectors, we can define the lengths of smooth curves, and the distance between a pair of points is defined as the length of the shortest curve joining the points, i.e., the geodesic. Consequently, every connected Riemannian manifold is also a metric space:

Definition 3.65. Let $c : [a, b] \subset \mathbb{R} \to M$ be a smooth curve on a Riemannian manifold (M, m) such that c is a diffeomorphism from $[a, b]$ to its range. Then its *length* $L(c)$ is defined by

$$L(c) = \int_a^b ||c_*(t)(e)|| \, dt,$$

where e is the canonical basis vector of $\mathbb{R}$.

Remark 3.9.2. The definition of the length of curves is independent of parameterization and depends only on the metric tensor and the range of c [5].

Theorem 3.6. Let M be a connected Riemannian manifold, and let the function $dis : M \times M \to \mathbb{R}$ be defined by the formula

$$dis(p, q) = \inf \{L(c) \mid \text{c is a smooth curve joining } p \text{ and } q\}.$$

Then (M, dis) is a metric space [5].

Remark 3.9.3. The metric topology of (M, dis) agrees with the manifold topology of M with any dis that can be defined on M [5].

Remark 3.9.4. The metric tensor chosen for the manifold may be such that its induced point-to-point metric is the same that was induced by the distance measurements, but it need not be the same.

As a corollary, we have a characterization of isometries, which are the structure-preserving mappings of Riemannian manifolds in terms of metric tensors:

Corollary 3.1. Let $f : M \to N$ be a differentiable mapping between Riemannian manifolds (M, m) and (N, n). If $f^*n = m$, then f is an *isometry.*

Remark 3.9.5. The structure-preserving mappings of semi-Riemannian manifolds, which are also called isometries, can be defined as in the above corollary: Let $f : M \to N$ be a differentiable mapping between semi-Riemannian manifolds (M, m) and (N, n). If $f^*n = m$, then f is an isometry.

Not only distance but also the concept of *orthogonality* derives from the metric tensor: tangent vectors $u, v \in T_p(M)$ are orthogonal if $m_p(u, v) = 0$ holds. This is similar to the traditional dot product $\cdot$, where vectors u and v are said to be orthogonal if $u \cdot v = 0$ holds. Useful concepts related to orthogonality are orthonormal basis and basis field:

Definition 3.66. A basis $\{\partial_1, ..., \partial_n\}$ of a tangent space is *orthonormal* if $m(\partial_i, \partial_j) = \delta_{ij}$ holds, and a basis field is *orthonormal* if its orthonormal in every point. The symbol δ_{ij} is the *Kronecker delta*, defined as follows: $\delta_{ij} = 1$ if $i = j$, otherwise $\delta_{ij} = 0$.

3.9.2 Representation of differential forms as proxy vector fields

In a semi-Riemannian manifold, we can canonically represent differential forms as vector fields and thus give a connection between differential form analysis and vector analysis: vector fields on an n-manifold are mappings that assign to each point of the manifold an element of n-dimensional vector space (the tangent space of the point). On the other hand, one-forms and $(n-1)$-forms on the same manifold are mappings that assign to each point of the manifold an element from a vector space whose dimension is n. Thus these vector spaces are isomorphic, and therefore it is possible to use isomorphisms to *represent* one-forms and $(n-1)$-forms with vector fields. Vectors representing forms are called proxy-vectors. Isomorphisms are not canonical, but if a metric tensor is defined, it provides a canonical isomorphism from a tangent space to the corresponding covector and $(n-1)$-covector spaces.

A vector space V is isomorphic to its dual V^* but has no canonical isomorphism. However, if an inner product Φ is defined on V, then it is possible to define a canonical isomorphism between V and V^*: by the Riesz representation theorem [66], for each $v \in V$, there corresponds a unique element f_v of V^* such that equation $f_v(w) = \Phi(v, w)$ holds for all $w \in V$. On the other hand, all elements of V^* can be defined this way. Thus the vector v together with Φ represents the functional f_v. However, if Φ is changed, the vector representing the functional is changed.

This procedure can be generalized to semi-Riemannian manifolds: let (M, m) be an n-dimensional semi-Riemannian manifold. Then a unique proxy vector field V_m exists for each one-form ω such that $\omega_p(v) = m_p(V_m(p), v)$ holds for all points $p \in M$ and all tangent vectors $v \in T_p(M)$ [16]. In this case, we say that the vector field V_m represents the one-form ω. Again the representation of one-forms with vector fields is not unique but depends on the chosen metric tensor. The isomorphism from one-forms to vector fields defined by the metric tensor is called *sharp*, and it is denoted by $\sharp$ [34]. The inverse of $\sharp$ is called *flat*, and it is denoted by $\flat$ [34].

The representation of $(n-1)$-forms with vector fields requires also a metric tensor, but now the role of the metric tensor remains hidden:

Definition 3.67. Let $(X_1, ..., X_n)$ be a positively oriented orthonormal basis field. An n-form *vol* is a *volume form* if $vol(X_1, ..., X_n) = 1$ holds everywhere.

Remark 3.9.6. The volume form is unique for each metric tensor and thus does not depend on any particular positively oriented orthonormal basis used in its definition [34]. In the literature [5], also the term volume element is used instead of volume form.

With the volume form *vol* we can now represent an $(n-1)$-form ω with the unique proxy vector field V_m that satisfies $vol_p(V_m(p), v_2, ..., v_n) = \omega_p(v_2, ..., v_n)$ for all points $p \in M$ and all tangent vectors $v_2, ..., v_n \in T_p(M)$.

With the contraction, we can define proxy vectors alternatively as follows: let m and *vol* be the metric tensor and the volume form of a semi-Riemannian manifold, respectively. Let ω and η be a one-form and an $(n-1)$-form, respectively. Then their proxy vectors U and V are those that satisfy the following equations:

$$\begin{cases} \omega &= i_U m \\ \eta &= i_V vol. \end{cases}$$

3.9.3 Hodge-operator

When electromagnetic theory is formulated with classical vector analysis, the constitutive equations are relations between vector fields: in the case of linear materials, permittivity, permeability, and conductivity are modeled as scalar or tensor fields that map a vector field to a vector field. That is, the materials are modeled as linear isomorphisms that map vector fields to vector fields. On the other hand, because the forms representing magnetic field (one-form) and magnetic flux density (two-form) are of a different degree, permeability cannot be a scalar field. However, in an n-manifold M the vector spaces $\Omega^k(M)$ and $\Omega^{n-k}(M)$ are isomorphic because the vector space of all antisymmetric k-linear mappings over $T_p(M)$ has the same dimension as the vector space of all antisymmetric $(n-k)$-linear mappings over $T_p(M)$. The isomorphism between $\Omega^k(M)$ and $\Omega^{n-k}(M)$ is not canonical, but on a semi-Riemannian manifold we can define a canonical isomorphism with the so-called Hodge-operator (also known as the star-operator).

Definition 3.68. Let M be an oriented semi-Riemannian n-manifold. The *Hodge-operator* is the unique linear isomorphism $\star : \Omega^k(M) \to \Omega^{n-k}(M)$ for $0 \leq k \leq n$ such that

$$\star(\omega \wedge \alpha) \;=\; i_{\sharp\alpha}(\star\omega)$$

holds for any k-form ω and a one-form α and

$$\star 1 \;=\; vol$$

holds for the zero-form 1 [13].

Proposition 3.7. The Hodge operator has the following properties [34] [62]:

(1) $\star\star = (-1)^{k(n-k)+s} id_{\Omega^k(M)}$, where s is the index of metric tensor

(2) definiteness: $\omega \wedge \star\omega$ is a nonzero n-form that maps bases from the positive class to positive numbers for all $\omega \neq 0$

(3) symmetry: $\omega \wedge \star\eta = \eta \wedge \star\omega$.

Remark 3.9.7. For each metric tensor, there is a corresponding Hodge-operator, and the converse is also true: given a linear isomorphism from one-forms to $(n-1)$-forms with the properties of definiteness and symmetry, it induces a metric tensor [8].

Remark 3.9.8. The properties of definiteness and symmetry of the Hodge operator make it possible to define an inner product $\langle \cdot, \cdot \rangle$ for $\Omega^k(M)$ as follows [62]:

$$\langle \omega, \eta \rangle = \int_M \omega \wedge \star \eta.$$

The Hodge-operator helps establish a relation between the exterior derivative and the differential operators gradient, curl, and divergence: let f be a scalar field and F a vector field. Then the formula for the gradient with any metric tensor is $grad\, f = \sharp(df)$; i.e., $grad\, f$ is a vector field that satisfies $df(V) = m(grad\, f, V)$ for all vector fields V and for the given metric tensor m. Again if the metric tensor is changed, the vector field $grad\, f$ is changed. The formula for the curl is $curl\, F = \sharp[\star\,(d(\flat F))]$, where the vector field F is first flattened or mapped to the one-form $\flat F$ by the isomorphism $\flat$. Then the exterior derivative of $\flat F$ is taken after, which the resulting two-form $d(\flat F)$ is mapped to the one-form $\star\,(d(\flat F))$ using the Hodge operator. Finally, using the sharp operator $\sharp$, the one-form $\star\,(d(\flat F))$ is mapped to the vector field $\sharp(\star\,(d(\flat F)))$. The formula for the divergence is $div\, F = \star(d(\star(\flat F)))$. Again the vector field F is first flattened to the one-form $\flat F$ and then mapped to the two-form $\star(\flat F)$ by the Hodge. The two-form is mapped to the three-form $d(\star(\flat F))$ using the exterior derivative, and finally the Hodge maps this to the scalar field (zero-form) $\star(d(\star(\flat F)))$.

3.10 Calculations in coordinates

In the above, we defined most of the structures without reference to coordinates or bases. This approach emphasized the conceptual characteristics of the structures and the fact that nature is independent of our choices of coordinates and bases. However, when we solve problems numerically with computers, we must, in the end, represent these structures with real numbers. This representation we achieve by choosing coordinates and bases. In this section, we give coordinate and basis representations of some of the structures defined above.

Charts are a convenient way to induce bases for the tangent spaces of a manifold from a chart: in the Euclidean space $\mathbb{R}^n$, the tangent space $T_a(\mathbb{R}^n)$ of a point $a \in \mathbb{R}^n$ is identified with the vector space $\mathbb{R}^n$ in an obvious way. Obvious identification implies that all the tangent spaces of $\mathbb{R}^n$ are canonically isomorphic to each other. Because the vector space $\mathbb{R}^n$ has a canonical (or natural or standard) basis, $(e_1, .., e_n)$, there is a natural basis field for the tangent spaces of the Euclidean space $\mathbb{R}^n$. Now because the codomain $V = \phi(U)$ of every chart ϕ is a manifold itself and a subset of the Euclidean space $\mathbb{R}^n$, a natural basis field exists for the tangent spaces of V. Then the pushforward of ϕ^{-1} induces a basis field called the coordinate frame in the tangent spaces of the manifold.

Definition 3.69. Let ϕ be chart of M, and let $(e_1, .., e_n)$ be the standard ordered basis of $\mathbb{R}^n$. Then the *coordinate frame* of ϕ is the ordered basis field of the tangent spaces of M given by $((\phi^{-1})_* e_1, .., (\phi^{-1})_* e_n)$.

Let $dim(M) = m$ and $dim(N) = n$, and let $f : M \to N$ be a differentiable mapping. The pushforward $f_*(p)$ at p is a linear map $f_*(p) : T_p(M) \to T_{f(p)}(N)$, and if we choose bases for $T_p(M)$ and $T_{f(p)}(N)$, then we can represent $f_*(p)$ as a matrix that maps from $\mathbb{R}^m$ to $\mathbb{R}^n$. If f is represented with charts as $\phi^{-1} \circ f \circ \varphi$, then the pushforward $f_*(p)$ is given by the *Jacobian matrix* [5] of the representation map $\phi^{-1} \circ f \circ \varphi$. In other words, the Jacobian matrix of the representation map is the matrix of $f_*(p)$ given in the coordinate frames $((\phi^{-1})_* e_1, .., (\phi^{-1})_* e_m)$ and $((\varphi^{-1})_* e_1, .., (\varphi^{-1})_* e_n)$.

Next we define bases for $\Omega^1(M)$ and for each $T_p^* M$ using the bases of $T_p(M)$.

Definition 3.70. Let $(X_1, .., X_n)$ be a frame field of $T(M)$. Then its *coframe* is the frame field $(dx_1, ..., dx_n)$ of $\Omega^1(M)$ defined by

$$dx_i(X_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

If the frame $(X_1, .., X_n)$ is a coordinate frame, then $(dx_1, ..., dx_n)$ is called the coordinate coframe [5]. The notation $(dx_1, ..., dx_n)$ for the coordinate coframe comes from the fact that a chart of an n-manifold M can be represented as an n-tuple of coordinate mappings $x_i : M \to \mathbb{R}$ such that x_i gives the ith coordinate of the points under the chart. Then dx_i is the differential of x_i.

Because a k-form is multilinear mapping at each point p, it is completely determined at p if its values are known for some basis vectors of $(T_p(M))^k$. These values are then the components of the multilinear mapping with respect to the basis. To expand the components to the whole manifold, we need basis fields, and coordinate frames offer a convenient way to do this:

Definition 3.71. Let $\omega \in \Omega^k(M)$ hold and let $(X_1, ..., X_n)$ be a coordinate frame field of $T(M)$. Furthermore, let $\tau : \{1, ..., k\} \to \{1, ..., n\}$ be a mapping such that $\tau(1) < \tau(2) < ... < \tau(k)$. In the following, we use the shorthand notation τ_i for $\tau(i)$. The *component functions* of ω with respect to the frame $(X_1, ..., X_n)$ are the functions $\omega_{\tau_1 \dots \tau_k} : M \to \mathbb{R}$ defined by

$$\omega_{\tau_1 \dots \tau_k}(p) = \omega_p(X_{\tau_1}, ..., X_{\tau_k}).$$

The wedge product provides a convenient way to express higher-degree forms as products of one-forms. The standard ordered basis of one-forms induced by a chart is its coordinate coframe $(dx_1, ..., dx_n)$. A basis of two-forms is given by the set $\{dx_1 \wedge dx_2, ..., dx_1 \wedge dx_n, dx_2 \wedge dx_3, ..., dx_2 \wedge dx_n, ..., dx_{n-1} \wedge dx_n\}$, consisting of all possible wedge products of the coframe, excluding the reversed products; i.e., either $dx_i \wedge dx_j$ or $dx_j \wedge dx_i$ is included but not both. Particularly, the standard ordered basis of two-forms in a 3-manifold is $(dx_2 \wedge dx_3, dx_3 \wedge dx_1, dx_1 \wedge dx_2)$. Similarly, we can define bases for k-forms as the set of all possible wedge products of k elements of the coframe, excluding the permuted products. In fact, the component functions of differential forms defined above are exactly the components of the form given in these bases [34]: let $\omega_{\tau_1...\tau_k}(p)$ be the component functions of $\omega \in \Omega^k(M)$ with respect to a coordinate frame $(X_1, ..., X_n)$, and if $(dx_1, ..., dx_n)$ is the coframe, then ω can be expressed as

$$\omega = \sum_{\tau_1<...<\tau_k} \omega_{\tau_1...\tau_k} dx_{\tau_1} \wedge ... \wedge dx_{\tau_k}.$$

There is an explicit formula for the exterior derivative of a k-form based on the coordinate coframe and the component functions:

Proposition 3.8. If a k-form ω is given in a coordinate coframe as $\sum_{\tau_1<...<\tau_k} \omega_{\tau_1...\tau_k} dx_{\tau_1} \wedge ... \wedge dx_{\tau_k}$, then its exterior derivative $d\omega$ is given by the formula

$$d\omega = \sum_{\tau_1<...<\tau_k} d\omega_{\tau_1...\tau_k} \wedge dx_{\tau_1} \wedge ... \wedge dx_{\tau_k}, \tag{3.1}$$

where $d\omega_{\tau_1...\tau_k}$ denotes the differential of the component function $\omega_{\tau_1...\tau_k}$ [34][62].

We can explicitly calculate the Hodge of a k-form in terms of the form itself if we use any orthonormal frame fields, including orthonormal coordinate frames:

Proposition 3.9. If ω is a k-form on a Riemannian manifold M, then its image $\star\omega$ under $\star$ is an $(n-k)$-form such that

$$(\star\omega)(X_{k+1}, ..., X_n) = \omega(X_1, ..., X_k)$$

holds for any orthonormal positively oriented frame field $(X_1, ..., X_n)$.

Chapter 4

Electromagnetic boundary value problems

Electromagnetic BVPs can be formulated with the well established differential geometry approach [13], instead of the traditional vector analysis approach. In this chapter, we show how to formulate electromagnetic BVPs with the tools of differential geometry. The main tools are manifolds-with-boundary, differential forms, the pullback of forms, the wedge product, the exterior derivative, the Lie derivative, and the contraction of forms. This approach and particularly its possibilities for numerical modeling are not well-known to engineers and architects of modeling software.

The domain of an electromagnetic BVP is a model of space, space and time, or even spacetime [57] [63]. Maxwell's equations are normally presented in the form in which space and time are separated so that time is treated as an independent parameter. In this so-called $(3+1)$-decomposition of equations, electromagnetic fields are split into electric and magnetic components. The decomposition is not absolute, but depends on the observer that splits spacetime into space and time.

To demonstrate the generality and flexibility of differential geometric structures, we show in the first three sections below how to derive a *very general* $(3+1)$*-decomposition of Maxwell's equations and constitutive equations*. This derivation is based on an additional structure on a manifold, called the *observer structure*, which is closely related to one-dimensional foliations of the spacetime manifold but is even more general. Observer is defined by a pair consisting of a vector field and a one-form over spacetime. Notice that the general $(3+1)$-decomposition is local and fully coordinate- and metric-free. Furthermore, the usual splitting of spacetime such that time can be treated as an independent parameter is a special case of this general procedure. In addition to $(3+1)$-decompositions, the observer structure constitutes an important tool in the theory of dimensional reduction introduced in chapter 5.

The generality and flexibility of differential geometric structures also allows us to formulate many types of problems, such as static BVPs, initial value problems, and Cauchy problems, in a single setting, simply called BVPs. For example, initial value problems are plain standard BVPs when the initial values are interpreted as a special type boundary values. In fact, in terms of differential geometry, it would be more accurate to say that the initial values are merely boundary values for which we have given a special name and separate treatment. In the fourth section in this chapter, we *formulate a general electromagnetic BVP* based on the general $(3+1)$-decomposition of Maxwell's equations using differential geometry.

The formulation of a general electromagnetic BVP in differential geometry is invariant under diffeomorphisms of the BVP domain; i.e., the formulation is generally covariant. Using general covariance, we can naturally derive the *equivalence of BVPs under diffeomorphisms* in the fifth section. From the physical point of view, equivalent BVPs can be thought to correspond to the same physical situation. Furthermore, if a BVP is formulated with a coordinate system, the change of coordinates procedure for BVPs is an instance of the equivalence of BVPs. Particularly noteworthy is the fact that the *equivalence of BVPs does not depend on a metric*.

The equivalence of BVPs under diffeomorphisms provides insight into efficient solutions of BVPs and particular problems: in chapter 6, we show how several apparently unrelated traditional methods or approaches used to solve problems are, in fact, simply instances of a general method based on the equivalence of BVPs. For example, solutions of open-boundary problems with compact domains and the invisibility cloaking are both instances of a procedure based on equivalent BVPs. Thus the *equivalence of BVPs provides unified theoretical explanations for many traditional solution methods and suggests new ones*: in chapters 6 and 7, we show how the equivalence can be used to speed up parametric modeling. The equivalence of BVPs under diffeomorphisms and the unified aspects it lends to the solution methods of BVPs is a major result of this thesis.

The diffeomorphism-invariance of the differential equations and boundary values is trivial because they are canonically defined for all manifolds-with-boundary. On the other hand, the diffeomorphism-invariance of the constitutive equations is not trivial: the constitutive equations must first be defined using a metric because there is no canonical way to define them in manifolds. However, diffeomorphism-invariance can be defined without a metric such that it depends only on diffeomorphism. Thus if equivalent BVPs are formulated with charts, the material parameters that describe the constitutive equations depend on the charts.

Finally, we study the exact role of the metric in electromagnetic BVPs. We show, as an another major result of this thesis, that the metric of the manifold and the constitutive equations are closely connected, but that in the end, the *role of the metric in formulations of electromagnetic BVPs is restricted to the initial identification of the BVP.*

4.1 Geometric decompositions of fields and the exterior derivative

The general method to split spacetime into space and time is based on the choice of observer defined geometrically by a pair (T, τ), where T is a smooth nonvanishing vector field, and τ is a smooth nonzero one-form such that $\tau(T) = 1$ holds everywhere. The pair (T, τ) defines an additional structure on manifolds called the observer structure, characterizable as a field of local observers. Observer structures endow a manifold with a local product structure: an n-manifold can be locally expressed as the product of a 1-manifold and an $(n-1)$-manifold.

A *holonomic* observer (T, τ) is such that τ is exact, i.e., there exists a zero-form λ such that $\tau = d_M\lambda$ holds. Holonomic observers correspond to $(3+1)$-foliations of a spacetime M and with them we get the usual form of Maxwell's equations. If τ is not exact or (T, τ) does not correspond to any $(3+1)$-foliation, the observer is *nonholonomic*. The pair (T, τ) also defines two complementary projections for fields and the exterior derivative and decomposes them into complementary components, which correspond to the local splitting of the manifold. These decompositions are called geometric decompositions.

4.1.1 Foliations of manifolds and observer structures

Maxwell's equations are defined in a four-dimensional spacetime manifold M, whose points are "events." Usually, a spacetime M is split into a global product $M = M_3 \times \mathbb{R}$, where M_3 is a 3-manifold modeling space and $\mathbb{R}$ models time. When M is a global product, time is treated as an independent parameter. We don't assume M to be a global product of two manifolds; instead, we use observer structures to split the spacetime M locally into a product of space and time: an observer (T, τ) divides M into nonintersecting three-dimensional submanifolds such that each submanifold represents spatial space at some moment of time; i.e., the pair defines a one-dimensional foliation of M.

To define a holonomic (T, τ) exactly, we first need a zero-form that defines a one-dimensional foliation of M or decomposes M into nonintersecting three-dimensional submanifolds. Let us assume that a smooth zero-form λ exists whose level sets form a one-parameter family of three-dimensional hypersurfaces $\lambda = constant$. In other words, the level sets are nonintersecting three-dimensional submanifolds, and M is the disjoint union of these submanifolds. Then the submanifold $\lambda_t = \{p \in M \,|\, \lambda(p) = t\}$ is a leaf of the foliation and can be thought of as a simple model for spatial space at time t. Figure 4.1 gives an example of a foliation of spacetime.

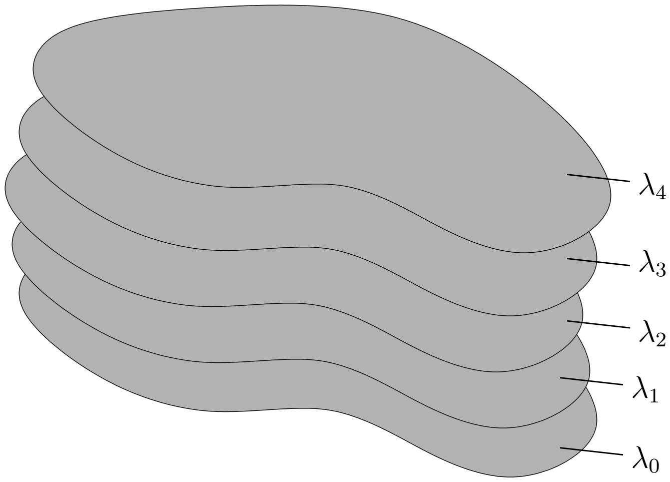

Figure 4.1: Example of a foliation. A (2+1)-foliation of a 3-dimensional space, where the two-dimensional surfaces λ_t are the leaves of the foliation. In the case of a (3+1)-foliation of spacetime, the leaves are three-dimensional.

The smooth one-form τ on M is now defined as the exterior derivative of λ or $\tau = d_M\lambda$. Observe that $d_M\tau = 0$ holds, and that the kernel[1] of τ and the tangent space of the leaves coincide at every point. Then we define T as a smooth vector field on M transversal to the leaves such that $\tau(T) = 1$ holds everywhere. Equivalently, T is such that the Lie derivative of λ with respect to T is one everywhere: because $i_T(d_M\lambda) = i_T\tau = \tau(T) = 1$ holds, it follows that $\mathcal{L}_T\lambda = i_T d_M\lambda = 1$. In summary, T defines the direction of time at each event and τ defines spatial space at every time instant, which in turn are defined by the foliation $\{\lambda_t\}$.

If the observer is nonholonomic, then $d_M\tau \neq 0$ holds, and τ is thus not a closed one-form. We consider nonholonomic observers only with remarks.

[1] τ is a covector at each point of the manifold, and the *kernel* of a covector at a point is the subspace of the tangent space of the point mapped to zero by the covector.

The pair (T, τ) decomposes tangent vectors into spatial and time components: τ decomposes M into nonintersecting submanifolds; thus together with T it also decomposes the tangent spaces of M into a *direct sum* [25][40]: $T_p(M) = span(T_p) \oplus ker(\tau_p)$ holds for all points, where $span(T_p) = \{\alpha T_p \in T_p(M) \,|\, \alpha \in \mathbb{R}\}$ and $ker(\tau_p) = \{v \in T_p(M) \,|\, \tau(v) = 0\}$. Consequently, there is a decomposition of all tangent vectors into two components: the component along the submanifolds, which is the spatial component, and the component parallel to T, which is the time component.

The decomposition of tangent spaces also shows that the foliation gives M locally a product structure: if we take a piece small enough from the manifold, it is diffeomorphic to a subset of the product manifold $\mathbb{R}^3 \times \mathbb{R}$. Notice also that the global decomposition $M = M_3 \times \mathbb{R}$ of a spacetime M corresponds to the situation where M is diffeomorphic to $M_3 \times \mathbb{R}$, and then M is identified with $M_3 \times \mathbb{R}$. Furthermore, notice that because no metric is involved here, *T cannot be assumed to be orthogonal to the submanifolds*.

All the above that was introduced to split spacetime locally into space and time parts can be generalized to other manifolds with a dimension equal to or greater than two. This generalization is important because dimensional reduction, introduced in chapter 5, is based on it. The generalization is as follows:

Definition 4.1. Let M be a differentiable manifold of dimension equal to or greater than two. An *observer structure* for M is a pair (T, τ), where T and τ are a smooth nonzero vector field on M and a smooth one-form on M, respectively, such that $\tau(T) = 1$ holds everywhere. Furthermore, if $\tau = d_M \lambda$ holds for some zero-form λ, then (T, τ) is a *holonomic* observer, otherwise (T, τ) is a *nonholonomic* observer

Open Question 1. What is the exact relationship between traditional observer models, such as one assumed by Newtonian mechanics, and the observer structure presented here? Notice that the observer structures defined above are more general than is used in the general relativity [44], where observers are based on metric tensors with index 1 (Lorentz manifolds). Metric tensor based observers allow the definition of causality structures [49]. What other applications than spacetime splitting to space and time and decompositions of differential equations the observer structure might offer?

4.1.2 Geometric decomposition of fields

An observer (T, τ) induces two *complementary projections* on fields that define the geometric decomposition of the fields: locally a field is decomposed into a component parallel to the vector field T (along the integral curves of T) and into a complementary component along submanifolds defined by τ. These projections play an essential role also in dimensional reduction; therefore, in this section (T, τ) need not relate to time and space, and τ need not be closed, i.e., $d_M \tau \neq 0$ may hold in general. The basic ideas in this and the next sections follow closely [15], where τ is given as the metric dual of T or where $\tau = \flat T$ holds, whereas we avoid involving the metric. Our terminology and notation also differ from those in [15]. Similar ideas about field decompositions are also presented in [26] and [39].

The geometric decomposition of a k-form ω contains two components specific to ω, one a k-form and the other a $(k-1)$-form, and these are called the geometric components of ω. Geometric components are spatial forms (generally called horizontal forms), which means that they map tangent vectors parallel to T to zero. The geometric components of electromagnetic fields are E, D, H, and B, which are all spatial forms. The spatial forms constitute their own exterior algebra under the wedge product.

Complementary projections are defined with extension and contraction.

Proposition 4.1. Let (T, τ) be an observer for a manifold M and $1 \leq k \leq n$. Then $P_T = I_\tau i_T : \Omega^k(M) \to \Omega^k(M)$ and $P_\tau = i_T I_\tau : \Omega^k(M) \to \Omega^k(M)$ are projections for each k. Furthermore, the projections are complementary; or if I denotes the identity mapping of the forms and 0 denotes the zero-valued form, then $P_\tau + P_T = I$ and $P_\tau P_T = P_T P_\tau = 0$ hold.

Proof: $P_T = \tau \wedge i_T$ is a projection if it is *idempotent* or if applying it twice yields the same result as applying it once. Let ω be a k-form, and let P_T be twice applied to it (notice that $i_T \tau = 1$ and $i_T(i_T \omega) = 0$ hold):

$$P_T P_T \omega = \tau \wedge i_T(\tau \wedge i_T \omega)$$
$$= \tau \wedge (i_T \tau \wedge i_T \omega - \tau \wedge i_T(i_T \omega)) = \tau \wedge i_T \omega = P_T \omega.$$

The fact that P_τ is a projection follows from the fact that P_T is a projection:

$$P_\tau \omega = i_T I_\tau \omega = i_T(\tau \wedge \omega)$$
$$= i_T \tau \wedge \omega - \tau \wedge i_T \omega = \omega - \tau \wedge i_T \omega = (I - P_T)\omega.$$

Thus $P_\tau = I - P_T$ and then $P_\tau P_\tau = (I - P_T)(I - P_T) = I - P_T - P_T + P_T^2 = I - P_T - P_T + P_T = I - P_T = P_\tau$. Clearly, $P_\tau + P_T = I$ now holds, and it is trivial to show that $P_\tau P_T = P_T P_\tau = 0$. □

These two projections decompose the forms into complementary parts, and the decomposition is a geometric decomposition:

Definition 4.2. The *geometric decomposition* of $\omega \in \Omega^k(M)$, when $1 \leq k \leq n$, with respect to the observer (T, τ) is

$$\omega = P_\tau \omega + P_T \omega.$$

Remark 4.1.1. The geometric decomposition of ω in terms of contraction and wedge product is $\omega = (\omega - \tau \wedge i_T\omega) + \tau \wedge i_T\omega$.

Remark 4.1.2. The observer (T, τ) need not be holonomic in the definition of geometric decomposition of fields. In fact, this property is not needed at all for geometric decompositions of fields. The τ being exact or not is important only when we decompose the exterior derivative which is used to express Maxwell's equations in spacetime.

Henceforth we adopt the notation $\omega_\tau = P_\tau\omega = \omega - \tau \wedge i_T\omega$, whereby the geometric decomposition is $\omega = \omega_\tau + \tau \wedge i_T\omega$. Because τ is given, the information specifying ω is given by the components ω_τ and $i_T\omega$, which we call the *geometric components* of ω . The definition of ω_τ does not make sense for zero-forms because the contraction is not defined. However, the definition that is in line with the original definition makes use of the part that does make sense: we set $\omega_\tau = \omega$ for the zero-forms. For n-forms on an n-dimensional manifold, the situation is reversed: because $\omega = \tau \wedge i_T\omega$ holds for all n-forms ω, then $\omega_\tau = 0$ holds for all n-forms.

Remark 4.1.3. In the dimensional reduction of BVPs, geometric components are the fields governed by lower-dimensional BVPs.

Because $T_p(M) = span(T_p) \oplus ker(\tau_p)$ holds for all points, every tangent vector $v \in T_p(M)$ decomposes as $v = v_\parallel + v_\perp$, where $v_\parallel \in ker(\tau_p)$ is the component tangent to the hypersurface through p defined by τ, and $v_\perp \in span(T_p)$ is the component parallel to T. We call $v_\parallel$ and $v_\perp$ the *horizontal* and the *vertical* component of v, respectively. The tangent vectors belonging to $ker(\tau)$ are called *horizontal vectors*. In the case of time and space splitting, horizontal vectors are called *spatial vectors*. A submanifold of M whose tangent vectors are all horizontal vectors in $T(M)$ is a *horizontal submanifold*. Next we define the horizontal or spatial forms that have closely connected to horizontal vectors and horizontal submanifolds:

Definition 4.3. A k-form ω is *horizontal* if $P_\tau\omega = \omega$ holds. If $\omega = \omega_\tau + \tau \wedge i_T\omega$, then ω_τ is the *horizontal component* of ω.

The following proposition characterizes horizontal (or spatial) forms.

Proposition 4.2. ω is horizontal if and only if $i_T\omega = 0$ holds.

Proof: If ω is horizontal, then

$$\begin{aligned} i_T\omega = i_T(\omega - \tau \wedge i_T\omega) = i_T\omega - i_T(\tau \wedge i_T\omega) \\ = i_T\omega - i_T\tau \wedge i_T\omega + \tau \wedge i_T(i_T\omega) = i_T\omega - i_T\omega = 0. \end{aligned}$$

On the other hand, if $i_T\omega = 0$ holds, then $\tau \wedge i_T\omega = 0$, and thus $\omega = \omega_\tau$ holds. □

Remark 4.1.4. The geometric components ω_τ and $i_T\omega$ of ω are both horizontal forms. This is crucial for time and space splitting, because geometric components are then spatial forms, as is expected: the fields E, D, H, and B are all spatial forms. The fact that geometric components are horizontal is also crucial for dimensional reduction, because they are the fields to be solved in the lower-dimensional BVP, which is defined on a horizontal submanifold.

If the dimension of M is n, then the horizontal component of a form is exactly the component of the form that restricts to the $(n-1)$-dimensional horizontal submanifolds defined by τ: let $i : N \to M$ be the inclusion map of a horizontal submanifold N to M. Then clearly the pushforward of i maps the tangent vectors of N to the horizontal vectors of M. Then the pullback of i is a restriction of forms to N, and because $i^*\omega(v_1, ..., v_k) = \omega(i_*v_1, ..., i_*v_k)$, where $i_*v_1, ..., i_*v_k$ are always horizontal vectors, it follows that the horizontal component is the component that restricts to N without loss of nontrivial information. Observe that these horizontal submanifolds are regular embedded submanifolds, which in spacetime splitting correspond to space at some moment of time and which is why horizontal forms are called spatial forms.

It is easy to show that the projection P_τ is compatible with the exterior algebra structure of $\Omega(M)$:

Lemma 4.1. The projection P_τ satisfies equations

$$\begin{aligned} P_\tau(\omega + a\eta) &= P_\tau\omega + aP_\tau\eta \quad \omega, \eta \in \Omega^k(M), a \in \mathbb{R} \\ P_\tau(\omega \wedge \eta) &= P_\tau\omega \wedge P_\tau\eta \quad \omega, \eta \in \Omega(M). \end{aligned}$$

Let us denote the set of all horizontal forms on M by $\Omega_h(M)$ and the set of all horizontal k-forms by $\Omega_h^k(M)$. Then Lemma 4.1 shows that $\Omega_h^k(M)$ is a vector subspace of $\Omega^k(M)$ for each k: the sum of two horizontal k-forms is a horizontal k-form, and the scalar multiple of a horizontal k-form is a horizontal k-form. Furthermore, Lemma 4.1 shows that the wedge product of horizontal forms is again a horizontal form. Thus $\Omega_h(M)$ has its own exterior algebra structure under the wedge product.

4.1.3 Geometric decomposition of the exterior derivative

When we apply the projections $P_T = I_\tau i_T$ and $P_\tau = i_T I_\tau$ to the exterior derivative, we produce the geometric decomposition of the exterior derivative. Particularly, with P_τ we can produce the horizontal or spatial exterior derivative, which takes the role of the curl and divergence in the four familiar Maxwell's equations. The horizontal exterior derivative operates nontrivially only to the horizontal components of differential forms, and the horizontal exterior derivative of a form is again a horizontal form. Generally speaking, in an n-manifold, the horizontal exterior derivative reduces to the exterior derivative of the $(n-1)$-manifold, or the horizontal exterior derivative is the exterior derivative of horizontal forms.

Definition 4.4. Let d_M be the exterior derivative of an n-manifold M, where (T, τ) defines an observer structure for M. Then the *horizontal exterior derivative* of d_M with respect to the observer (T, τ) is the operator d_τ on M defined by

$$d_\tau = P_\tau d_M,$$

and the *geometric decomposition of the exterior derivative* d_M is defined by

$$d_M = P_\tau d_M + P_T d_M.$$

Remark 4.1.5. In terms of contraction, wedge product, and d_M the horizontal exterior derivative is given by $d_\tau = d_M - \tau \wedge i_T d_M$ and the geometric decomposition of the exterior derivative is given by $d_M = d_\tau + \tau \wedge i_T d_M$.

Proposition 4.3. The horizontal exterior derivative of a k-form is a horizontal $(k+1)$-form or if $\omega \in \Omega^k(M)$, then $d_\tau \omega \in \Omega_h^{k+1}(M)$.

Proof: $d_\tau \omega$ clearly is a $(k+1)$-form. To show that it is a horizontal form, we use the characterization of Proposition 4.2 or show that $i_T d_\tau \omega = 0$ holds:

$$i_T d_\tau \omega = i_T(d_M \omega - \tau \wedge i_T d_M \omega) = i_T d_M \omega - i_T d_M \omega + \tau \wedge i_T i_T d_M \omega = 0. \quad \square$$

If (T, τ) defines a splitting of spacetime to space and time, then the horizontal exterior derivative is called the *spatial exterior derivative.* The spatial exterior derivative reduces to the exterior derivative of a 3-manifold: let $i : N \to M$ be the inclusion map of a horizontal 3-submanifold N (defined by τ) to M. Then $i^* d_M \omega = i^* d_\tau \omega + i^*(\tau \wedge i_T d_M \omega) = i^* d_\tau \omega$ holds since $i^* \tau = 0$.

It is easy to show that the horizontal exterior derivative d_τ satisfies the product rule for horizontal forms, i.e., the following equation

$$d_\tau(\omega \wedge \eta) = d_\tau\omega \wedge \eta + (-1)^k \omega \wedge d_\tau\eta$$

holds for all $\omega \in \Omega_h^k(M)$ and $\eta \in \Omega_h(M)$. In case of holonomic observers, the complex property $d_\tau d_\tau = 0$ holds for the horizontal exterior derivative d_τ. Then also Stokes's theorem holds for d_τ and horizontal submanifolds: let $\omega \in \Omega_h^k(M)$ and let N be a horizontal (k+1)-submanifold of M. Now, by definition, the restriction of $d_M\omega$ to N is the same as the restriction of $d_\tau\omega$. Thus we have

$$\int_N d_\tau\omega = \int_N d_M\omega.$$

Then because Stokes's theorem holds for d_M, we have Stokes's theorem also for the horizontal forms and the horizontal exterior derivative:

$$\int_N d_\tau\omega = \int_{\partial N} \omega.$$

4.2 Maxwell's equations

With geometric decompositions of fields and the exterior derivative, we can derive a $(3+1)$-decomposition of Maxwell's equations with separate spatial and time derivatives. However, to derive a $(3+1)$-decomposition, we must first introduce Maxwell's equations in a spacetime with no separation of space and time. Hence there is no separation between electric and magnetic fields either, and we must talk about electromagnetic fields. When an observer (T, τ) is introduced, the fields and the exterior derivative get decomposed along with Maxwell's equations in spacetime. Because the decomposition is local, and because time is not an independent parameter, time derivatives are expressed as Lie derivatives with respect to T.

4.2.1 Maxwell's equations in spacetime

In this section, we introduce Maxwell's equations in spacetime. Our goal is not to study this model or justify it, but to show how to derive $(3+1)$-decompositions from it. Thus the model is assumed valid, and the reader may consult [26]. The model can be built from the classical form of Maxwell's equations or based on certain axioms stating, e.g., the conservation of the charge and magnetic flux as in [26].

Let M be a four-dimensional spacetime manifold, i.e., a differentiable manifold, whose points are "events." Let $\mathcal{F}$ be the electromagnetic field two-form, $\mathcal{G}$ the excitation two-form, and $\mathcal{J}$ the source three-form in M. These fields are governed by Maxwell's equations

$$\begin{aligned} d_M\mathcal{F} &= 0 \\ d_M\mathcal{G} &= \mathcal{J}, \end{aligned} \tag{4.1}$$

where d_M denotes the exterior derivative of M. The first equation contains Faraday's law and Gauss's law for the magnetic field. The second equation contains Ampère's law and Gauss's law for the electric field. To peel these laws out, we must decompose the spacetime into a product of space and time.

4.2.2 $(3+1)$-decomposition of Maxwell's equations

Let (T, τ) be an observer defining a local splitting of spacetime into space and time. Now the geometric decompositions of the fields $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{J}$ are

$$\begin{aligned} \mathcal{F} &= \mathcal{F}_\tau + \tau \wedge i_T\mathcal{F} \\ \mathcal{G} &= \mathcal{G}_\tau + \tau \wedge i_T\mathcal{G} \\ \mathcal{J} &= \mathcal{J}_\tau + \tau \wedge i_T\mathcal{J}. \end{aligned}$$

The decomposition of fields $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{J}$ into magnetic and electric parts depends on the choice of observer, i.e., the choice of the pair (T, τ). Therefore, we *rename* the geometric components $\mathcal{F}_\tau$, $i_T\mathcal{F}$, $\mathcal{G}_\tau$, $i_T\mathcal{G}$, $\mathcal{J}_\tau$, and $i_T\mathcal{J}$ as follows:

$$\begin{aligned} \mathcal{F}_\tau &= -B \\ i_T\mathcal{F} &= E \\ \mathcal{G}_\tau &= D \\ i_T\mathcal{G} &= H \\ \mathcal{J}_\tau &= \rho \\ i_T\mathcal{J} &= -J. \end{aligned} \tag{4.2}$$

Then the geometric decompositions are

$$\begin{aligned} \mathcal{F} &= -B + \tau \wedge E \\ \mathcal{G} &= D + \tau \wedge H \\ \mathcal{J} &= \rho - \tau \wedge J. \end{aligned}$$

The derivation of the $(3 + 1)$-decomposition of Maxwell's equation in (4.1) is based on geometric decompositions of the fields (Definition 4.2) and the exterior derivative (Definition 4.4) in the equations of (4.1) under an observer (T, τ). Let us first derive Faraday's law and magnetic Gauss's law from equation $d_M \mathcal{F} = 0$:

$$\begin{aligned} d_M \mathcal{F} &= (P_\tau d_M + P_T d_M)(P_\tau \mathcal{F} + P_T \mathcal{F}) \\ &= P_\tau d_M P_\tau \mathcal{F} + P_\tau d_M P_T \mathcal{F} + P_T d_M P_\tau \mathcal{F} + P_T d_M P_T \mathcal{F}. \end{aligned} \tag{4.3}$$

Now if (T, τ) is holonomic, then the next lemma shows how to simplify the above decomposition of $d_M \mathcal{F}$.

Lemma 4.2. Let M be a manifold with a holonomic observer structure (T, τ), and let $\omega = \omega_\tau + \tau \wedge i_T \omega$ hold. Then the following equations are satisfied:

$$\begin{aligned} P_\tau d_M P_T \omega &= 0 \\ P_T d_M P_\tau \omega &= I_\tau \mathcal{L}_T \omega_\tau \\ P_T d_M P_T \omega &= -I_\tau d_\tau i_T \omega. \end{aligned}$$

Proof: The first and third equation need the assumption of holonomy, i.e., $d_M \tau = 0$ must hold:

$$\begin{aligned} P_\tau d_M P_T \omega &= i_T I_\tau d_M (\tau \wedge i_T \omega) \\ &= i_T I_\tau (d_M \tau \wedge i_T \omega - \tau \wedge d_M i_T \omega) \\ &= -i_T I_\tau (\tau \wedge d_M i_T \omega) \\ &= -i_T (\tau \wedge \tau \wedge d_M i_T \omega) \\ &= 0 \end{aligned}$$

$$\begin{aligned} P_T d_M P_T \omega &= I_\tau i_T d_M I_\tau i_T \omega \\ &= I_\tau i_T (d_M \tau \wedge i_T \omega - \tau \wedge d_M i_T \omega) \\ &= -I_\tau i_T (\tau \wedge d_M i_T \omega) \\ &= -I_\tau (d_M i_T \omega - \tau \wedge i_T d_M i_T \omega) \\ &= -I_\tau d_\tau i_T \omega. \end{aligned}$$

The second equation follows from Cartan's formula and Proposition 4.2:

$$P_T d_M P_\tau \omega = I_\tau i_T d_M \omega_\tau = I_\tau \mathcal{L}_T \omega_\tau. \qquad \square$$

Next we apply the results of the above lemma to equation (4.3) and use definitions 4.3 and 4.4:

$$\begin{aligned} P_\tau d_M P_\tau \mathcal{F} + P_\tau d_M P_T \mathcal{F} + P_T d_M P_\tau \mathcal{F} + P_T d_M P_T \mathcal{F} &= 0 \\ d_\tau \mathcal{F}_\tau + I_\tau (\mathcal{L}_T \mathcal{F}_\tau - d_\tau i_T \mathcal{F}) &= 0. \end{aligned} \tag{4.4}$$

In (4.4) the three-form $d_\tau \mathcal{F}_\tau$ and the two-form $\mathcal{L}_T \mathcal{F}_\tau - d_\tau i_T \mathcal{F}$ are spatial forms. Furthermore, the three-forms $d_\tau \mathcal{F}_\tau$ and $I_\tau(\mathcal{L}_T \mathcal{F}_\tau - d_\tau i_T \mathcal{F})$ are complementary which can be seen by applying the complementary projections P_τ and P_T to equation (4.4):

$$\begin{aligned} P_\tau(d_\tau \mathcal{F}_\tau) + P_\tau(I_\tau(\mathcal{L}_T \mathcal{F}_\tau - d_\tau\, i_T \mathcal{F})) &= d_\tau \mathcal{F}_\tau \\ P_T(d_\tau \mathcal{F}_\tau) + P_T(I_\tau(\mathcal{L}_T \mathcal{F}_\tau - d_\tau\, i_T \mathcal{F})) &= I_\tau(\mathcal{L}_T \mathcal{F}_\tau - d_\tau\, i_T \mathcal{F}). \end{aligned}$$

Thus both $d_\tau \mathcal{F}_\tau$ and $I_\tau(\mathcal{L}_T \mathcal{F}_\tau - d_\tau i_T \mathcal{F})$ must be zero so that equation (4.4) holds. Besides, because $\mathcal{L}_T \mathcal{F}_\tau - d_\tau\, i_T \mathcal{F}$ is a spatial form and τ is nonzero, the three-form $\tau \wedge (\mathcal{L}_T \mathcal{F}_\tau - d_\tau\, i_T \mathcal{F})$ is zero only if $\mathcal{L}_T \mathcal{F}_\tau - d_\tau\, i_T \mathcal{F} = 0$ holds. Hence the following equations hold in M:

$$\begin{aligned} d_\tau \mathcal{F}_\tau &= 0 \\ \mathcal{L}_T \mathcal{F}_\tau - d_\tau\, i_T \mathcal{F} &= 0. \end{aligned}$$

If M is a global product or $M = M_3 \times \mathbb{R}$, and if time is an independent parameter, then with the renamings in equation (4.2) and notation $\frac{\partial}{\partial t}$ for $\mathcal{L}_T$, we recover the familiar form of the laws:

$$\begin{aligned} d_\tau B &= 0 \\ d_\tau E &= -\frac{\partial B}{\partial t}. \end{aligned}$$

Next we derive Ampère's law and electric Gauss's law from equation $d\mathcal{G} = \mathcal{J}$. For this, we use the geometric decompositions of the fields $\mathcal{G}$ and $\mathcal{J}$ and the exterior derivative d_M:

$$\begin{aligned} d_M \mathcal{G} &= \mathcal{J} \\ (P_\tau d_M + P_T d_M)(P_\tau \mathcal{G} + P_T \mathcal{G}) &= P_\tau \mathcal{J} + P_T \mathcal{J} \\ P_\tau d_M P_\tau \mathcal{G} + P_\tau d_M P_T \mathcal{G} + P_T d_M P_\tau \mathcal{G} + P_T d_M P_T \mathcal{G} &= \mathcal{J}_\tau + I_\tau i_T \mathcal{J}. \end{aligned}$$

If we assume holonomic observer and use Lemma 4.2, the last equation above simplifies to the following equation:

$$d_\tau \mathcal{G}_\tau + I_\tau(\mathcal{L}_T \mathcal{G}_\tau - d_\tau i_T \mathcal{G}) = \mathcal{J}_\tau + I_\tau i_T \mathcal{J}.$$

The above equation implies the following two equations:

$$\begin{aligned} d_\tau \mathcal{G}_\tau &= \mathcal{J}_\tau \\ \mathcal{L}_T \mathcal{G}_\tau - d_\tau\, i_T \mathcal{G} &= i_T \mathcal{J}. \end{aligned}$$

Again if time is an independent parameter, then using the renamings and notation $\frac{\partial}{\partial t}$ for $\mathcal{L}_T$, we recover the familiar form of the laws:

$$\begin{aligned} d_\tau D &= \rho \\ d_\tau H &= J + \frac{\partial D}{\partial t}. \end{aligned}$$

Remark 4.2.1. If τ was not exact, in which case we deal with a nonholonomic observer, the above derivations would yield the following equations:

$$\begin{aligned} d_\tau \mathcal{F}_\tau &= -d_\tau \tau \wedge i_T \mathcal{F} \\ d_\tau\, i_T \mathcal{F} &= \mathcal{L}_T \mathcal{F}_\tau + \mathcal{L}_T \tau \wedge i_T \mathcal{F} \\ d_\tau \mathcal{G}_\tau &= \mathcal{J}_\tau - d_\tau \tau \wedge i_T \mathcal{G} \\ d_\tau\, i_T \mathcal{G} &= i_T \mathcal{J} + \mathcal{L}_T \mathcal{G}_\tau + \mathcal{L}_T \tau \wedge i_T \mathcal{G}. \end{aligned}$$

Thus the essential difference between holonomic and nonholonomic systems is that in the latter the states of the system depend on the paths taken to achieve them (e.g. $d_\tau \mathcal{F}_\tau$ depends on $d_\tau \tau$ and $d_\tau\, i_T \mathcal{F}$ depends on $\mathcal{L}_T \tau$). That is, a conservative potential function for electric and magnetic fields is possible only in holonomic systems.

In summary, the (3+1)-decomposition of Maxwell's equations corresponding to a holonomic observer (T, τ) is

$$\begin{aligned} d_\tau E &= -\mathcal{L}_T B \\ d_\tau D &= \rho \\ d_\tau H &= J + \mathcal{L}_T D \\ d_\tau B &= 0. \end{aligned}$$

4.3 Decomposition of constitutive equations

In this section we briefly show how to use the observer-induced projections to derive $(3+1)$-decomposition of constitutive equations. We first introduce a linear constitutive relation between fields $\mathcal{F}$ and $\mathcal{G}$ in spacetime M. The relation is written with a linear operator χ that maps the two-form $\mathcal{F}$ to the two-form $\mathcal{G}$. Then we show how an observer decomposes the operator χ and thereby gives constitutive relations between the geometric components of $\mathcal{F}$ and $\mathcal{G}$. Furthermore, the Ohm's law that connects J to E and B is introduced.

4.3.1 Constitutive equations in spacetime

We only present a suitable relation between fields $\mathcal{F}$ and $\mathcal{G}$ and show how to derive $(3+1)$-decompositions from it. Thus the model is assumed valid, and the reader may consult [26] [56].

Let M be a four-dimensional spacetime manifold, i.e., a differentiable manifold, whose points are "events." Let $\mathcal{F}$ be the electromagnetic field two-form and $\mathcal{G}$ the excitation two-form in M. These fields satisfy the constitutive equation

$$\mathcal{G} = \chi\mathcal{F}, \tag{4.5}$$

where $\chi : \Omega^2(M) \to \Omega^2(M)$ is a linear operator that satisfies the following two axioms:

(1) symmetry: $\omega \wedge \chi\eta = \chi\omega \wedge \eta$ holds for all $\omega, \eta \in \Omega^2(M)$,

(2) closure: $\chi \circ \chi = -I$, where I is the identity mapping of $\Omega^2(M)$.

Observe that these properties (linearity, symmetry, closure) are independent of any metric of spacetime. However, given χ with these properties will induce a metric tensor with index one (Lorentz metric) to spacetime M [56].

4.3.2 $(3+1)$-decomposition of constitutive equation

Let (T, τ) be an observer defining a local splitting of spacetime M into space and time. The decomposition of the operator χ is based on the geometric decompositions of the fields $\mathcal{F}$ and $\mathcal{G}$ in (4.5) and on the linearity of χ:

$$\begin{aligned} \mathcal{G} &= \chi\mathcal{F} \\ P_\tau\mathcal{G} + P_T\mathcal{G} &= \chi P_\tau\mathcal{F} + \chi P_T\mathcal{F}. \end{aligned} \tag{4.6}$$

Now if we apply P_τ to equation (4.6), we get an equation for $P_\tau\mathcal{G}$:

$$\begin{aligned} P_\tau P_\tau\mathcal{G} + P_\tau P_T\mathcal{G} &= (P_\tau\chi P_\tau)(\mathcal{F}) + (P_\tau\chi P_T)(\mathcal{F}) \\ P_\tau\mathcal{G} &= (P_\tau\chi P_\tau)(P_\tau\mathcal{F}) + (P_\tau\chi P_T)(P_T\mathcal{F}). \end{aligned} \tag{4.7}$$

By applying P_T to equation (4.6) we get an equation for $P_T\mathcal{G}$:

$$\begin{aligned} P_T P_\tau\mathcal{G} + P_T P_T\mathcal{G} &= (P_T\chi P_\tau)(\mathcal{F}) + (P_T\chi P_T)(\mathcal{F}) \\ P_T\mathcal{G} &= (P_T\chi P_\tau)(P_\tau\mathcal{F}) + (P_T\chi P_T)(P_T\mathcal{F}). \end{aligned} \tag{4.8}$$

Because of the complementarity of the projections P_τ and P_T, equations (4.7) and (4.8) show that the constitutive equation $\mathcal{G} = \chi\mathcal{F}$ can be written as the following formal matrix equation:

$$\begin{bmatrix} P_\tau\mathcal{G} \\ P_T\mathcal{G} \end{bmatrix} = \begin{bmatrix} P_\tau\chi P_\tau & P_\tau\chi P_T \\ P_T\chi P_\tau & P_T\chi P_T \end{bmatrix} \begin{bmatrix} P_\tau\mathcal{F} \\ P_T\mathcal{F} \end{bmatrix}.$$

Thus χ has the following block decomposition:

$$\chi = \begin{bmatrix} P_\tau\chi P_\tau & P_\tau\chi P_T \\ P_T\chi P_\tau & P_T\chi P_T \end{bmatrix}. \tag{4.9}$$

Let us next look at the decompositions of χ and the constitutive equation $\mathcal{G} = \chi\mathcal{F}$ in terms of the geometric components $\mathcal{G}_\tau$, $i_T\mathcal{G}$, $\mathcal{F}_\tau$, and $i_T\mathcal{F}$. Particularly, we want separate constitutive equations for $\mathcal{G}_\tau$ and $i_T\mathcal{G}$ in terms of $\mathcal{F}_\tau$ and $i_T\mathcal{F}$. By applying the projection P_τ to equation (4.6), we get the following equations for $\mathcal{G}_\tau$:

$$\begin{aligned} P_\tau P_\tau\mathcal{G} + P_\tau P_T\mathcal{G} &= P_\tau\chi P_\tau\mathcal{F} + P_\tau\chi P_T\mathcal{F} \\ P_\tau\mathcal{G} &= P_\tau\chi P_\tau P_\tau\mathcal{F} + P_\tau\chi I_\tau i_T\mathcal{F} \\ \mathcal{G}_\tau &= (P_\tau\chi P_\tau)(\mathcal{F}_\tau) + (P_\tau\chi I_\tau)(i_T\mathcal{F}). \end{aligned} \tag{4.10}$$

Observe that equation (4.10) is equivalent to equation (4.7). Next, we apply the contraction i_T to equation (4.6) to get the equation for $i_T\mathcal{G}$:

$$\begin{aligned} i_T P_\tau\mathcal{G} + i_T P_T\mathcal{G} &= i_T\chi P_\tau\mathcal{F} + i_T\chi P_T\mathcal{F} \\ i_T P_T\mathcal{G} &= i_T\chi P_\tau P_\tau\mathcal{F} + i_T\chi I_\tau i_T\mathcal{F} \\ i_T\mathcal{G} &= (i_T\chi P_\tau)(\mathcal{F}_\tau) + (i_T\chi I_\tau)(i_T\mathcal{F}). \end{aligned} \tag{4.11}$$

Notice that if we apply I_τ to equation (4.11), then it is equivalent to equation (4.8). Thus equations (4.10) and (4.11) define a decomposition of χ which is equivalent to the decomposition in (4.9).

We see that operators $\chi^\tau_\tau = -P_\tau\chi P_\tau$, $\chi^\tau_T = P_\tau\chi I_\tau$, $\chi^T_\tau = -i_T\chi P_\tau$, and $\chi^T_T = i_T\chi I_\tau$ map horizontal forms to horizontal forms. Precisely we have:

$$\begin{aligned} \chi^\tau_\tau &: \Omega^2_h(M) \to \Omega^2_h(M) \\ \chi^\tau_T &: \Omega^1_h(M) \to \Omega^2_h(M) \\ \chi^T_\tau &: \Omega^2_h(M) \to \Omega^1_h(M) \\ \chi^T_T &: \Omega^1_h(M) \to \Omega^1_h(M). \end{aligned}$$

In terms of the renamings in (4.2), the equations (4.10) and (4.11) are as follows:

$$\begin{aligned} D &= \chi^\tau_\tau B + \chi^\tau_T E \\ H &= \chi^T_\tau B + \chi^T_T E. \end{aligned} \tag{4.12}$$

Comparing the above equations to the well-known constitutive equations $D = \epsilon E$ and $H = \nu B$ which hold e.g. in some static cases, we obtain the following relations:

$$\begin{aligned} \epsilon &= \chi_T^\tau \\ \nu &= \chi_\tau^T. \end{aligned} \tag{4.13}$$

The operators χ_τ^τ and χ_T^T describe the so-called magneto-electric media.

Observe that ϵ and ν as well as χ_τ^τ and χ_T^T are observer-dependent. In some cases, but not always, it is possible to choose an observer such that

$$\begin{aligned} D &= \epsilon E \\ H &= \nu B \end{aligned}$$

holds everywhere in M.

If conductors are present, then Ohm's law connects the current J to E and B:

$$J = \sigma_E E + \sigma_B B, \tag{4.14}$$

where σ_E and σ_B are linear mappings $\sigma_E : \Omega_h^1(M) \to \Omega_h^2(M)$ and $\sigma_B : \Omega_h^2(M) \to \Omega_h^2(M)$. Because J, B and E are observer dependent, σ_E and σ_B are also observer dependent. If the observer is attached to the conductors, then a simpler form of Ohm's law holds:

$$J = \sigma_E E.$$

4.4 Electromagnetic BVPs with differential geometry

An electromagnetic BVP is a mathematical model for some physical situation and its electromagnetic fields. A BVP itself consists of a domain with a boundary, differential equations governing the fields over the domain, boundary values of the fields, source fields, and constitutive equations. Furthermore, some global data of the topology of the domain must be specified to fix the homology/cohomology classes and to ensure uniqueness of the solution. When a BVP corresponds to an electromagnetic problem, the domain is a model of space or more generally of space and time or even spacetime. The differential equations are Maxwell's equations, which govern the electric and magnetic fields over the domain. The constitutive equations model the material effects on the fields.

Our aim is to formulate general electromagnetic BVPs using the mathematical structures of differential geometry. These general BVPs can include the full system of Maxwell's equations and, therefore, wave-propagation problems. Thus we include the initial value problems and the Cauchy problems (extended initial value problems that have the initial value and the time derivative of the field given at the initial time). Included are also the mixed problems or the initial-boundary value problems and the Cauchy-boundary value problems. We call these problems by the generic name *boundary value problem* because they all consist of differential equations over a manifold-with-boundary such that the fields are pre-defined at the boundary. Observe that the domain may not be fully bounded by the boundary but it can be partly "open" (e.g., no final time). This term is also arguable because wave-propagation problems can be formulated simply by giving boundary values for the fields: only field values at boundaries, including the initial boundary and thus initial values, must be specified.

In the following, we assume that M is a four-dimensional manifold-with-boundary modeling spacetime. A holonomic observer (T, τ) induces a decomposition of the fields $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{J}$ into electric and magnetic parts, which are governed by Maxwell's equations on M:

$$\begin{aligned} d_\tau E &= -\mathcal{L}_T B \\ d_\tau H &= J + \mathcal{L}_T D \\ d_\tau D &= \rho \\ d_\tau B &= 0. \end{aligned}$$

If the problem is static or time-harmonic, then M is three-dimensional manifold modeling space: a static or time-harmonic BVP has invariance/symmetry with respect to time, and this makes it possible to reduce the dimension of the domain. Of course the differential equations are simplified by this reduction of the dimension (more on this in chapter 5).

The boundary values for the fields E, D, H, and B are usually given only on a part of the boundary ∂M; e.g., the boundary values of magnetic fields are given such that the "tangential" component of H is given for a part of the boundary, and the "normal" component of B is given for the complementary part of the boundary. Thus we assume that the boundary ∂M is a union of two disjoint parts $\partial_1 M$ and $\partial_2 M$. Then let $i^1_\partial : \partial_1 M \to M$ be the inclusion map of the part $\partial_1 M$ of the boundary ∂M to M. Similarly, i^2_∂ is the inclusion map of the part $\partial_2 M$. Then their pullbacks, denoted by t^1 and t^2 and called trace, enable us to define boundary values for the fields E, D, H, and B:

$$\begin{aligned} t^1 E &= e \\ t^1 H &= h \\ t^2 D &= d \\ t^2 B &= b. \end{aligned}$$

Of course, boundary separation need not coincide for magnetic and electric fields. Furthermore, with wave-propagation problems, a need may arise to give more boundary conditions at the initial boundary: if $\partial_1 M$ is the initial boundary, we may need to specify also D and B at $\partial_1 M$ to give sufficient boundary conditions to specify a unique solution. However, since these wider generalities do not add anything essential to the topic, we continue with the above.

In linear materials, the constitutive equations are expressed with linear isomorphisms ϵ, μ, and σ which are observer-dependent mappings from $\Omega^1_h(M)$ to $\Omega^2_h(M)$:

$$\begin{aligned} D &= \epsilon E \\ B &= \mu H \\ J &= \sigma E. \end{aligned}$$

Furthermore, to assure uniqueness of the solution of BVPs, operator ϵ (also μ and σ) must together with the wedge product define a positive energy density as $E \wedge D = E \wedge \epsilon E$. This requires that ϵ be definite in the sense that $E \wedge \epsilon E$ is a nonzero horizontal 3-form that maps direct triplets of horizontal tangent vectors to positive numbers for all $E \neq 0$. Hodge operators are definite linear isomorphisms, which map one-forms to two-forms in three-dimensional Riemannian manifolds (see Definition 3.68 and Proposition 3.7) and that is why they are often used to describe the constitutive relations. Observe that there may not be any holonomic observer such that the constitutive equations have the above simple form, but in general we must consider more general equations (4.13) and (4.14). Finally, in general, operators ϵ, μ, and σ need not be linear (nonlinear materials), in which case they are definite bijective mappings that map one-forms to two-forms.

Specification of differential equations, boundary values, and constitutive equations, even if not contradictory, does not generally guarantee a unique solution to a problem. For instance, consider the domain of a resistor model (Figure 4.2). Inside the domain, the current stationary equations $dE = 0$ and $dJ = 0$ hold as does the constitutive equation $J = \sigma E$. The trace of E is set to zero at the resistor terminals, and the trace of J is set to zero at the

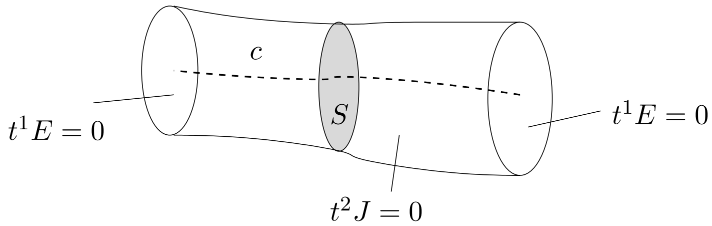

Figure 4.2: Boundary values and relative cohomology classes of a resistor model. To fix the relative cohomology class, we must fix either the value of the integral of E over the curve c or the value of the integral of J over the surface S.

resistor casing. However, this problem is not yet well-posed, because it has multiple solutions. To ensure a unique solution, we must fix either the value of the integral of E over some curve c that connects the terminals or the value of the integral of J over some surface S, through which all the current passes. Thus we must specify either the potential difference between the terminals or the total current through the resistor. Formally, this specifies the *cohomology class* [38] of E or J, and these topological conditions are not local. Furthermore, by de Rham's theorem [24] the cohomology classes can always be specified by fixing proper integrals of the fields. Observe that using a potential often fix the cohomology classes automatically: if we formulate the resistor problem using the scalar potential, we must set the values of the potential at the terminals and therefore we automatically fix the cohomology classes. Because the topology of the domain in the resistor is trivial (no holes), only relative cohomology classes are defined. However, in general with nontrivial topologies also absolute cohomology classes must be considered. In summary, the well-posedness of a BVP requires that we consider cohomology, a topic not discussed in detail in this thesis.

Observe that differential forms and their exterior derivative and boundary values as well as cohomology classes are defined without any Riemannian structure. Thus Maxwell's equations and boundary values are independent of a metric. Though the constitutive equations are often modeled as linear isomorphisms—this does not yet call for a metric—yet to identify the relations for the first time, we need a metric. The exact relation of the constitutive equations to a metric and role of the metric in BVPs are discussed in the rest of the chapter.

4.5 Equivalent formulations of a BVP

We now show that the above formulation of a general electromagnetic BVP with differential geometry is invariant under diffeomorphisms of the domain. This naturally defines an equivalence of BVPs under diffeomorphisms. This equivalence with its many applications is a generalization of the change of coordinates procedure: the role of the coordinate systems is played by manifolds and the change of coordinates mappings is replaced by diffeomorphisms between manifolds. Furthermore, the equivalence is defined *completely without a metric*.

In many applications it is advantageous to formulate BVPs for computers differently from the standard parameterization approach. These applications are all described with the same general theoretical setting: a BVP is formulated in some Riemannian manifold that corresponds to rigid body-measurements, but to gain in numerical solutions, an equivalent problem is posed on another diffeomorphic manifold. Thus the equivalence of BVPs gives a unified explanation for traditional methods of solving BVPs such as change of coordinates, solving open boundary problems with compact domains, and invisibility cloaking. Furthermore, the equivalence of BVPs can generate brand new methods. In chapter 6 and 7, we explain the possibility of accelerating parametric modeling.

Assume that M is a four-dimensional manifold-with-boundary, and (T,τ) is a holonomic observer providing a decomposition of M into space and time. We denote the restriction to the boundary by t_M. Then we assume that the following BVP is formulated on M:

$$\begin{aligned}
d_\tau E &= -\mathcal{L}_T B \\
d_\tau H &= J + \mathcal{L}_T D \\
d_\tau D &= \rho \\
d_\tau B &= 0 \\
t^1_M E &= e \\
t^1_M H &= h \\
t^2_M D &= d \\
t^2_M B &= b \\
D &= \epsilon_M E \\
B &= \mu_M H \\
J &= \sigma_M E.
\end{aligned} \tag{4.15}$$

Next we aim to formulate an *equivalent* BVP on another manifold N which is diffeomorphic to M via mapping $F : N \to M$. Observe that M and

N may be semi-Riemannian manifolds, but that the diffeomorphism F need not be an isometry; i.e., if M and N are semi-Riemannian manifolds, they need not be isomorphic semi-Riemannian manifolds. In fact, this is often exactly what is sought.

The observer (T, τ) and map F induce a pullback observer, a pair (Γ, γ), for N, which decomposes N into space and time:

Definition 4.5. Let $F : N \to M$ be a diffeomorphism and (T, τ) an observer in M. The *pullback observer* on N under F is the observer (Γ, γ) such that $\Gamma = (F^{-1})_* T$ and $\gamma = F^* \tau$.

Observe that $\gamma(\Gamma) = (F^*\tau)(\Gamma) = \tau(F_*\Gamma) = \tau(T) = 1$ holds as it should. The following proposition shows that the holonomy/nonholonomy of the observer is preserved under the pullback F^*:

Proposition 4.4. If (T, τ) is a holonomic observer in M, then the pullback observer (Γ, γ) on N is also holonomic.

Proof: If τ is exact, or $\tau = d_M \lambda$ holds for some zero-form λ, then by the naturality of the exterior derivative, $\gamma = F^*\tau = F^* d_M \lambda = d_N F^* \lambda$ holds for the zero-form λ. Thus γ is also exact. □

Now we have the observers (T, τ) and (Γ, γ) for M and N, respectively. The decomposition of the fields induced by an observer and its pullback observer are compatible with the pullback F^*:

$$
\begin{aligned}
F^*(\omega_\tau) &= (F^*\omega)_\gamma \\
F^*(i_T \omega) &= i_\Gamma F^* \omega \\
F^*(\tau \wedge i_T \omega) &= \gamma \wedge i_\Gamma F^* \omega.
\end{aligned}
$$

Because the corresponding or equivalent fields on N are the pulled-back fields or F^*E, F^*H, F^*D, F^*B, F^*J, and $F^*\rho$, the compatibility of the decompositions and the pullback simplifies things, as we will see.

The essence of formulating an equivalent problem to another manifold is the compatibility of the pullback and the operators d_τ, $\mathcal{L}_T$, t_M, ϵ_M, μ_M, and σ_M. Compatibility is described as an appropriate commutation rule of the pullback F^* with the other operators.

4.5.1 Equivalent differential equations

To derive the equivalent differential equations, we need the following commutation rules for d_τ and $\mathcal{L}_T$:

Theorem 4.1. $F^* \circ d_\tau = d_\gamma \circ F^*$

Proof: Use the naturality of the exterior derivative d_M (see Theorem 3.2) and Lemma 3.2:

$$\begin{aligned}
F^* d_\tau &= F^*(d_M - \tau \wedge i_T d_M) \\
&= F^* d_M - F^*(\tau \wedge i_T d_M) \\
&= d_N F^* - F^*\tau \wedge F^*(i_{F_*\Gamma} d_M) \\
&= d_N F^* - \gamma \wedge i_\Gamma F^* d_M \\
&= d_N F^* - \gamma \wedge i_\Gamma d_N F^* \\
&= d_\gamma F^* \qquad \square
\end{aligned}$$

Theorem 4.2. $F^* \circ \mathcal{L}_T = \mathcal{L}_\Gamma \circ F^*$

Proof: Use the naturality of d_M and Lemma 3.2:

$$\begin{aligned}
F^* \mathcal{L}_T &= F^*(d_M i_T + i_T d_M) \\
&= F^* d_M i_T + F^* i_{F_*\Gamma} d_M \\
&= d_N F^* i_{F_*\Gamma} + i_\Gamma F^* d_M \\
&= d_N i_\Gamma F^* + i_\Gamma d_N F^* \\
&= \mathcal{L}_\Gamma F^* \qquad \square
\end{aligned}$$

Differential equations are derived for the pulled-back fields from the original equations such that both sides of the equations are pulled back to N with F^*:

$$\begin{aligned}
F^*(d_\tau E) &= -F^*(\mathcal{L}_T B) \\
F^*(d_\tau H) &= F^* J + F^*(\mathcal{L}_T D) \\
F^*(d_\tau D) &= F^* \rho \\
F^*(d_\tau B) &= 0.
\end{aligned}$$

Using the above two theorems, we then get differential equations for the pulled-back fields on N:

$$\begin{aligned}
d_\gamma(F^* E) &= -\mathcal{L}_\Gamma F^* B \\
d_\gamma(F^* H) &= F^* J + \mathcal{L}_\Gamma F^* D \\
d_\gamma(F^* D) &= F^* \rho \\
d_\gamma(F^* B) &= 0.
\end{aligned}$$

Observe that the pulled-back fields satisfy the same differential equations as the originals.

Remark 4.5.1. If the chosen observer is nonholonomic, in which case additional terms appear in the differential equations, as shown in remark 4.2.1, then the additional terms could be treated exactly as shown here. Particularly, the terms involving the wedge product are simple because the pullback is naturally compatible with the wedge product.

Remark 4.5.2. If the observer in N were not a pullback observer, the equations for the pulled-back fields would be different because the geometric decompositions of the pulled-back fields and the exterior derivative d_N would not be compatible with the pullback.

4.5.2 Equivalent boundary values

Equivalent boundary values for pulled-back fields are easy to derive. Because the mapping F is diffeomorphic, by Lemma 3.1 the restriction of F to the boundary ∂N is a diffeomorphism $F_\partial : \partial N \to \partial M$. Consequently, we have the following commutation rule:

Theorem 4.3. Let $F : N \to M$ be a diffeomorphism and $F_\partial : \partial N \to \partial M$ the restriction of F to boundaries ∂N and ∂M. Furthermore, let t_N and t_M denote the restrictions of the fields to boundaries ∂N and ∂M. Then ${F_\partial}^* \circ t_M = t_N \circ F^*$ holds.

Proof: Let i_M and i_N be the inclusion mappings of the boundaries ∂M and ∂N to M and N, respectively. Thus $t_M = i_M^*$ and $t_N = i_N^*$ hold. Then the mapping F_∂ satisfies, by definition, the following equation

$$i_M \circ F_\partial = F \circ i_N.$$

Then using the rule $(g \circ f)^* = f^* \circ g^*$, we get the desired result:

$$\begin{aligned} (i_M \circ F_\partial)^* &= (F \circ i_N)^* \\ {F_\partial}^* \circ i_M^* &= i_N^* \circ F^* \\ {F_\partial}^* \circ t_M &= t_N \circ F^* \quad \square \end{aligned}$$

Now if we apply the pullback F_∂^* to the boundary conditions we get

$$\begin{aligned} {F_\partial}^* t_M^1 E &= {F_\partial}^* e \\ {F_\partial}^* t_M^1 H &= {F_\partial}^* h \\ {F_\partial}^* t_M^2 D &= {F_\partial}^* d \\ {F_\partial}^* t_M^2 B &= {F_\partial}^* b. \end{aligned}$$

Then applying the above proposition to the above equations yields boundary conditions for the pulled-back field:

$$\begin{aligned} t^1_N F^* E &= F_\partial{}^* e \\ t^1_N F^* H &= F_\partial{}^* h \\ t^2_N F^* D &= F_\partial{}^* d \\ t^2_N F^* B &= F_\partial{}^* b. \end{aligned}$$

4.5.3 Equivalent constitutive equations

Now we derive a suitable commutation rule for the pullback with operators ϵ_M, μ_M, and σ_M: we need a new operator μ_N such that

$$F^* B = \mu_N F^* H$$

holds. The operator μ_N can be given in terms of the pullback and the original operator μ_M as follows:

$$\mu_N F^* H = F^* B = F^*(\mu_M H).$$

We require that this equation holds for all one-forms H. Thus μ_N must satisfy the following equation:

$$\mu_N \circ F^* = F^* \circ \mu_M. \tag{4.16}$$

This is the commutation for the operator μ.

The operator μ_N differs from the operators d_γ, $\mathcal{L}_\Gamma$, and t_N in a crucial way: its commutation rule is its *definition*, i.e., the operators d_γ, $\mathcal{L}_\Gamma$, and t_N are defined without d_τ, $\mathcal{L}_T$, and t_M, and their commutation rules with the pullback hold as theorems. But μ_N cannot be defined without μ_M, and the commutation rule with the pullback holds by definition. From (4.16), we can produce the equations for ϵ_N, μ_N, and σ_N:

$$\begin{aligned} \epsilon_N &= F_2^* \circ \epsilon_M \circ (F_1^*)^{-1} \\ \mu_N &= F_2^* \circ \mu_M \circ (F_1^*)^{-1} \\ \sigma_N &= F_2^* \circ \sigma_M \circ (F_1^*)^{-1}, \end{aligned} \tag{4.17}$$

where F_1^* and F_2^* denote the pullbacks of one-forms and two-forms, respectively. In general, if M and N are n-manifolds and υ_M maps k-forms to $(n-k)$-forms, then the equivalent operator υ_N is defined by the following equation:

$$\upsilon_N = F_{n-k}^* \circ \upsilon_M \circ (F_k^*)^{-1}. \tag{4.18}$$

Remark 4.5.3. If the constitutive equations are more general, e.g. D depends linearly on E and B as in (4.13), then the equivalent operators are also defined by equation (4.18).

Finally, let us show that if μ_M is definite, then so is μ_N:

$$\begin{aligned} F^*H \wedge \mu_N F^*H &= F^*H \wedge F^*\mu_M (F^*)^{-1} F^*H \\ &= F^*H \wedge F^*\mu_M H \\ &= F^*(H \wedge \mu_M H). \end{aligned}$$

Clearly, $F^*(H \wedge \mu_M H)$ is nonzero for nonzero H, and if F induces an orientation for N from M, then F is orientation-preserving and $F^*(H \wedge \mu_M H)$ maps direct triplets of tangent vectors to positive numbers.

4.5.4 Equivalent BVP

In summary, the BVP on M described in (4.15) can be equivalently expressed as the following BVP on N:

$$\begin{aligned} d_\gamma F^*E &= -\mathcal{L}_\Gamma F^*B \\ d_\gamma F^*H &= F^*J + \mathcal{L}_\Gamma F^*D \\ d_\gamma F^*D &= F^*\rho \\ d_\gamma F^*B &= 0 \\ t^1_N F^*E &= F_\partial{}^* e \\ t^1_N F^*H &= F_\partial{}^* h \\ t^2_N F^*D &= F_\partial{}^* d \\ t^2_N F^*B &= F_\partial{}^* b \\ F^*D &= \epsilon_N F^*E \\ F^*B &= \mu_N F^*H \\ F^*J &= \sigma_N F^*E. \end{aligned} \tag{4.19}$$

Remark 4.5.4. If $N = M$, in which case F is a diffeomorphism from M to itself, then we can talk about invariance of physical laws. The differential equations and the equations determining the boundary condition in (4.19) hold for all diffeomorphism $F : M \to M$. Thus these equations (Maxwell's equations) are invariant under the full diffeomorphism group. On the other hand, the maximal symmetry group under which the constitutive equations are invariant is the Poincarè group, which is the group of all isometries of the Minkowski spacetime [44]. Minkowski spacetime is an example of a Lorentz manifold (see Definition 3.64). Observe that we have defined the equivalence of the constitutive equations under the full diffeomorphism group, but this is only possible by defining a suitable corresponding relation case by case.

The equivalence of BVPs is defined without the assumption that a unique solution exists. However, for the equivalence to be useful, it must preserve the existence of a unique solution, or if a BVP on M has a unique solution, an equivalent BVP on N has also a unique solution: Assume that the generic BVP on M defined in (4.15) has a unique solution. Using the pullback, we have shown (the commutation rules with the pullback) that the pulled-back fields form a solution to the equivalent BVP on N. A solution thus exists for the BVP on N. Conversely, for each solution on N, we can construct a solution on M using the pullback. But because the BVP on M has a unique solution, all the solutions on N must be mapped to the unique solution on M. Furthermore, because the pullback of a diffeomorphism is a linear isomorphism, only one solution on N can be mapped to the unique solution on M. Thus the equivalent BVP on N has a unique solution. In addition, if the generic BVP has multiple solutions or no solutions at all, then the same holds also for the equivalent BVPs.

We have presented the equivalence of BVPs under diffeomorphism but without formal rigor; i.e., formally, we should define the set of all BVPs and then define an equivalence relation for the set of all BVPs. Because the equivalence we have derived is clearly reflexive, symmetric, and transitive, it is, in fact, an equivalence relation. However, full formalization of equivalent BVPs is not pursued in this thesis, though we recognized that such formalization would be highly valuable in understanding BVPs and in designing solver software systems. However, the author does realize that BVPs can be formulated in myriad of ways.

Open Question 2. How to define rigorously the set of all BVPs to make the equivalence of BVPs rigorous?

4.6 Equivalent BVPs: Material parameters and chart

In the previous section, we derived the equivalence of BVPs under diffeomorphisms. Because the codomains of charts of a given manifold are also manifolds, the above description of equivalent BVPs holds also for the ranges of the charts. Furthermore, if the range of the charts are considered coordinate systems, the above procedure expresses the change of coordinates-procedure: F is the change of coordinates mapping from a coordinate system $N \subset \mathbb{R}^n$ to a coordinate system $M \subset \mathbb{R}^n$. The pulled-back fields (F^*H, etc.) are the fields expressed in the new coordinate system N. Similarly, if μ_{M} is the matrix (given in standard bases of forms in $\mathbb{R}^n$) containing the material parameter values for M, then μ_{N} is the matrix containing the material parameter values for the new coordinate system N.

As already explained, the spatial exterior derivative, the Lie derivative, and the trace are canonically defined for each manifold and thus for each chart (i.e., ranges of charts). Therefore, for new chart, we need to define only material parameter values. In practice, material parameters and constitutive equations are related to spatial forms. Hence we now assume that M is a two- or three-dimensional spatial manifold, i.e., the problem is static, time-harmonic, or time is separated independent parameter.

First, M is endowed with a metric structure given by distance measurements with a rigid body. Let f and g be two charts of M such that f is a standard parameterization, where distances are given in some length-unit-system (e.g., meters or inches) so that we know the values of the material parameters in f. Let (dx_1, dx_2, dx_3) be the coframe of the standard frame field in $f(M) \subset \mathbb{R}^3$. Then (dx_1, dx_2, dx_3) is the standard ordered basis of the one-forms in $f(M)$. The corresponding standard ordered basis for two-forms is $(dx_2 \wedge dx_3, dx_3 \wedge dx_1, dx_1 \wedge dx_2)$. The one-form H and the two-form B are then given in the standard bases of the chart f as $H_f = H_1 dx_1 + H_2 dx_2 + H_3 dx_3$ and $B_f = B_1 dx_2 \wedge dx_3 + B_2 dx_3 \wedge dx_1 + B_3 dx_1 \wedge dx_2$. If the constitutive equation $B_f = \mu_f H_f$ holds, it can be written in a component form using matrix formalism as follows:

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \\ \mu_{31} & \mu_{32} & \mu_{33} \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix}. \tag{4.20}$$

Because the matrix of μ_f is given in the standard bases of the standard parameterization f corresponding to some length-unit-system, the matrix is known from the literature such as handbooks, books of tables, and specification sheets of manufacturers.

Let $F = f \circ g^{-1}$ be the change of chart mapping from g to f given in coordinates by

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \mapsto \begin{bmatrix} F_1(y_1, y_2, y_3) \\ F_2(y_1, y_2, y_3) \\ F_3(y_1, y_2, y_3) \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

The pulled-back fields $B_g = F^* B_f$ and $H_g = F^* H_f$ satisfy the constitutive equation $B_g = \mu_g H_g$, and our task now is to solve the matrix of μ_g in terms of the matrix of μ_f. For this, we need the pullback of F as shown in equation (4.17).

Let (dy_1, dy_2, dy_3) and $(dy_2 \wedge dy_3, dy_3 \wedge dy_1, dy_1 \wedge dy_2)$ be the standard ordered bases of one- and two-forms in g corresponding to the standard frame field of g. Now the matrix of F_1^* is given by

$$F_1^* = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} & \frac{\partial x_3}{\partial y_1} \\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} & \frac{\partial x_3}{\partial y_2} \\ \frac{\partial x_1}{\partial y_3} & \frac{\partial x_2}{\partial y_3} & \frac{\partial x_3}{\partial y_3} \end{bmatrix} = J_F^T,$$

where J_F^T is the transpose of the Jacobian matrix of F [5]. The matrix of F_2^* is

$$F_2^* = \begin{bmatrix} \frac{\partial x_2}{\partial y_2}\frac{\partial x_3}{\partial y_3} - \frac{\partial x_2}{\partial y_3}\frac{\partial x_3}{\partial y_2} & \frac{\partial x_3}{\partial y_2}\frac{\partial x_1}{\partial y_3} - \frac{\partial x_3}{\partial y_3}\frac{\partial x_1}{\partial y_2} & \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_3} - \frac{\partial x_1}{\partial y_3}\frac{\partial x_2}{\partial y_2} \\ \frac{\partial x_2}{\partial y_3}\frac{\partial x_3}{\partial y_1} - \frac{\partial x_2}{\partial y_1}\frac{\partial x_3}{\partial y_3} & \frac{\partial x_3}{\partial y_3}\frac{\partial x_1}{\partial y_1} - \frac{\partial x_3}{\partial y_1}\frac{\partial x_1}{\partial y_3} & \frac{\partial x_1}{\partial y_3}\frac{\partial x_2}{\partial y_1} - \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_3} \\ \frac{\partial x_2}{\partial y_1}\frac{\partial x_3}{\partial y_2} - \frac{\partial x_2}{\partial y_2}\frac{\partial x_3}{\partial y_1} & \frac{\partial x_3}{\partial y_1}\frac{\partial x_1}{\partial y_2} - \frac{\partial x_3}{\partial y_2}\frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1} \end{bmatrix}.$$

Equivalently, if J_F is the Jacobian matrix of F, then the matrix of F_2^* is given by the formula

$$F_2^* = |J_F| J_F^{-1}, \tag{4.21}$$

where $|J_F|$ is the determinant of J_F. Thus the matrix μ_g is given by the following equation:

$$\mu_g = |J_F| J_F^{-1} \mu_f J_F^{-T} \tag{4.22}$$

Instead of the mapping $F = f \circ g^{-1}$, we may want to give the mapping $G = F^{-1} = g \circ f^{-1}$, which gives the coordinates of g in terms of the coordinates of f. Because F and G are diffeomorphisms, the Jacobian matrix J_G of G is the inverse of J_F or $J_G = J_F^{-1}$. Then the matrix μ_g is given by the formula

$$\mu_g = \frac{1}{|J_G|} J_G \mu_f J_G^T. \tag{4.23}$$

We can derive the matrix μ_g also for proxy vectors: if Φ_f is the standard metric tensor of f and vol_f the corresponding volume forms, the proxies $\mathbf{H}_f$ and $\mathbf{B}_f$ of H_f and B_f are defined by the following equations (see section 3.9.2):

$$\begin{aligned} i_{\mathbf{H}_f} \Phi_f &= H_f \\ i_{\mathbf{B}_f} vol_f &= B_f. \end{aligned}$$

As expressed by the coordinates of f, the proxies $\mathbf{H}_f$ and $\mathbf{B}_f$ satisfy the constitutive equation $\mathbf{B}_f = \mu_f \mathbf{H}_f$, where μ_f is a linear isomorphism. If $\mathbf{H}_f$ and $\mathbf{B}_f$ are represented as component vectors in the standard frame field of f, then the matrix of μ_f is known from the literature and is the same as in equation (4.20).

To derive an equation for the matrix μ_g, we need to know how to express the proxies $\mathbf{H}_g$ and $\mathbf{B}_g$ in terms of the proxies $\mathbf{H}_f$ and $\mathbf{B}_f$: the proxies $\mathbf{B}_f$ and $\mathbf{B}_g$ satisfy the following equation:

$$B_g = i_{\mathbf{B}_g} vol_g = F^* B_f = F^*(i_{\mathbf{B}_f} vol_f) = i_{F_*^{-1}\mathbf{B}_f} F^* vol_f$$

Because $F^* vol_f = |J_F| vol_g$ holds [5], we get

$$i_{F_*^{-1}\mathbf{B}_f} F^* vol_f = i_{F_*^{-1}\mathbf{B}_f} |J_F| vol_g = i_{|J_F| F_*^{-1}\mathbf{B}_f} vol_g.$$

Thus we must have

$$\mathbf{B}_g \;\; = \;\; |J_F| F_*^{-1} \mathbf{B}_f.$$

Because the pushforward F_* is given in coordinate frames by the Jacobian matrix J_F, the component vectors satisfy

$$\mathbf{B}_g \;\; = \;\; |J_F| J_F^{-1} \mathbf{B}_f.$$

which is the same equation as for two-forms in the 3d case given in equation 4.21. This is not surprising because the standard metric tensor of a chart is such that the coordinate frame is orthonormal, and it follows that the components of $\mathbf{B}$ in any coordinate frame are the same as those of B in the corresponding dual frame. Similarly, the components of $\mathbf{H}$ in any coordinate frame are the same as those of H in the corresponding dual frame. Consequently, the same equation holds also for the component vectors of $\mathbf{H}$ as for those of H:

$$\mathbf{H}_g \;\; = \;\; J_F^T \mathbf{H}_f.$$

We can now derive the matrix μ_g in the constitutive equation $\mathbf{B}_g = \mu_g \mathbf{H}_g$ of the component vectors:

$$\begin{aligned} \mathbf{B}_g &= \mu_g \mathbf{H}_g \\ |J_F| J_F^{-1} \mathbf{B}_f &= \mu_g J_F^T \mathbf{H}_f \\ |J_F| J_F^{-1} \mu_f \mathbf{H}_f &= \mu_g J_F^T \mathbf{H}_f. \end{aligned}$$

This holds for all $\mathbf{H}_f$ if and only if

$$\mu_g \;\; = \;\; |J_F| J_F^{-1} \mu_f J_F^{-T}. \tag{4.24}$$

holds. Because of the orthonormality of the coordinate frames, the result is, of course, the same as for the forms in equation (4.22).

Derivation of (4.24) with proxy vectors and physical arguments for quasi-static cases is presented in [54]: the pullback of one-forms can be thought of as an invariance of virtual work under change of charts. The constitutive equations are related to total system energy ($e_{tot} = \int_M H \wedge \mu H$), which must be invariant under change of charts.

In a 2d case, change of chart formulae are different for forms and their proxy vectors: let (dx_1, dx_2) and (dy_1, dy_2) be the standard basis of one-forms for f and g, respectively. Notice that B must be a one-form so as to have a Hodge-like isomorphism $B = \mu H$. Then the equivalent operator μ_g is given by the equation

$$\mu_g \;=\; F_1^* \circ \mu_f \circ (F_1^*)^{-1}.$$

Thus the matrix of μ_g is given by the formula

$$\mu_g \;=\; J_F^T \mu_f J_F^{-T}.$$

As shown in [54], the matrix of μ_g for proxy vectors in 2d quasi-static cases is given by the formula

$$\mu_g = \frac{1}{|J_G|} J_G \mu_f J_G^T = |J_F| J_F^{-1} \mu_f J_F^{-T}. \tag{4.25}$$

Let us next study why the change of chart formulae differ so much for forms and proxy vectors in 2d. In terms of forms, both B and H are one-forms and have thus the same transformation rules. However, the proxy $\mathbf{H}$ of H is defined with respect to the metric tensor whereas the proxy $\mathbf{B}$ of B is defined with respect to the unit area form. If the metric tensor is denoted by Φ_2 and the corresponding volume form by A, the proxies are defined as follows:

$$\begin{aligned} B &= i_{\mathbf{B}} A \\ H &= i_{\mathbf{H}} \Phi_2. \end{aligned}$$

Thus because proxies are defined differently for H and B, their proxies transform differently under change of chart.

4.7 Metric and electromagnetic BVPs

In the above, equivalence of BVPs was defined without a metric of the manifold. Furthermore, only the constitutive equations have any connection to the metric structure of the manifold. Thus the role of the metric in electromagnetic BVPs is not so restrictive and omnipresent as it appears in formulations based on vector analysis. This section shows that a metric is necessary only for the initial identification of a BVP, after which it can be disregarded. Particularly, we show that although initial identification of the operators ϵ, μ, and σ is done with a specific metric, the operators do not depend on any particular Riemannian structure used in the manifold, but that only their representations with Hodge-operators depend on the Riemannian structure. However, if diffeomorphic manifolds correspond to physically distinct situations, a distinction between the manifolds must be made. This distinction can be made only with a physical reference that exists outside the model: the manifolds are recognize as different by the modeler using external metric, i.e., distance measurements with some rigid body.

4.7.1 Formulation of BVPs in practice

The formulation of a BVP on a manifold M, as in section 4.4, is abstract in the sense that it is devoid of numbers. However, in practice, numbers are needed to represent the objects of the BVP and to apply arithmetic in calculations. To get these numbers, metric and geometry are used to produce a *standard parameterization* f (see Definition 3.20). The metric and geometry of f are features of space that we observe with our sight and rigid body measurements (and time measurement with clocks). Thus *the metric and geometry are part of a systematic process of producing a model of reality.*

By identifying the observed reality with a chart f, we produce a manifold M: f defines the points of M by labeling them with coordinates and also fully defines a differentiable manifold structure for M. Furthermore, the standard metric tensor of the chart f can be pulled back to M thereby to induce a Riemannian structure on M.

In addition, the operators ϵ, μ, and σ, which characterize materials, are described with numbers specific to f: The differential forms on the range of f are represented in the standard bases of $\mathbb{R}^n$. Thus the linear isomorphisms ϵ, μ, and σ are represented as matrices with respect to standard bases. Observe that the numbers found in the literature for the operators ϵ, μ, and σ, are always represented in standard bases and are specific to a class of standard parameterizations that are defined by the same unit of length. To get ϵ, μ, and σ on M, we use coordinate coframes, and in these bases the operators have the same matrices as in f.

Maxwell's equations can be defined directly on M without a metric. The boundary values are first specified with the chart f, after which they can be pulled back to M with f.

4.7.2 Constitutive equations and metric

In the constitutive equations, the operators ϵ, μ, and σ are definite linear operators, and as such, independent of a metric. However, their construction or identification in practice requires a metric: the vector space of covectors on a point of a 3-manifold M is isomorphic to the vector space of two-covectors at the same point. This pointwise isomorphism can be extended to the whole manifold such that the vector spaces of one-forms and two-forms on M are isomorphic. Similarly, the vector space of horizontal one-forms $\Omega_h^1(M_4)$ on a 4-manifold M_4 is isomorphic to that of horizontal two-forms $\Omega_h^2(M_4)$. However, no unique isomorphism exists between them. On the other hand, if a metric tensor m is defined on M, we can define a Hodge-operator $\star_m$ that yields a unique definite isomorphism, such that it can be also used to define energy. With the Hodge-operator $\star_m$, we can represent the linear isomorphism ϵ as a composite of the linear isomorphisms $\epsilon_m : \Omega^2(M) \to \Omega^2(M)$ and $\star_m : \Omega^1(M) \to \Omega^2(M)$. The operator ϵ_m is needed because the metric m is defined globally and usually in a manner independent of the materials occupying the space. Thus the operator ϵ_m characterizes the materials. A similar representation holds for μ and σ, and we have the following decompositions:

$$\begin{aligned} \epsilon &= \epsilon_m \circ \star_m \\ \mu &= \mu_m \circ \star_m \\ \sigma &= \sigma_m \circ \star_m . \end{aligned} \tag{4.26}$$

Notice that both operators ϵ_m and $\star_m$ depend on the metric m whereas ϵ does not: if we change the metric m of M to m'; i.e., if we formulate an equivalent problem for M using the identity mapping of M and change the metric, then clearly the fields, Maxwell's equations, boundary values, and constitutive relations do not change. That is, the *same operator ϵ describes a constitutive relation for both metrics m and m'*. However, the *decomposition* of the operator ϵ is changed by changing the metric. Thus for each definite linear isomorphism ϵ, there exists an equivalence class of pairs of linear operators $\{\epsilon_m, \star_m\}$, where each pair corresponds to some metric tensor: the pairs $\{\epsilon_i, \star_i\}$ and $\{\epsilon_j, \star_j\}$ are equivalent if $\epsilon_i \circ \star_i = \epsilon_j \circ \star_j$ holds.

The above discussion based on the equivalence of BVPs may give an impression that the operator ϵ is completely independent of the metric. However, this is not true, on the contrary: if BVP domains have physically measurable metrical differences, but are topologically the same, the differences between the domains are shown in the operator ϵ. Thus diffeomorphic domains can be physically different, i.e., the domains are recognized as different with an external metric (distance measurements with some rigid body). Furthermore, the initial identification of ϵ is always done with a standard parameterization, which is based on some external metric. However, even though ϵ recognizes metrical differences, the equivalence relation defined above shows that ϵ does not depend on any particular external metric used to identify it.

Even though the operators ϵ, μ, σ do not depend on any particular metric of M, given the operator ϵ (or μ or σ), we can canonically choose a metric for M. The choice of metric is based on the decomposition shown in (4.26): consider a decomposition where ϵ_m is the identity mapping of $\Omega^2(M)$. We then have $\epsilon = \star_m$ with a metric m, whose induced Hodge-operator $\star_m$ is exactly ϵ [8]. This metric is called the *ϵ-metric*. Obviously, it is different for distinct operators ϵ, and in general does not correspond to any rigid body-metric used in the formulation of a BVP. Furthermore, the ϵ-metric is generally defined only locally because the operator ϵ need not be smooth over the whole M (material interfaces).

The above discussion shows that the constitutive equations do not depend on the metric chosen to represent/identify them. However, because they imply a metric, they contain the structure of metric. Thus the constitutive equations are not completely independent of the metric in the same way as Maxwell's equations, but they are independent of the instance of the metric. The decompositions in (4.26) and the above discussion suggests the following useful definition:

Definition 4.6. Let M be an n-dimensional oriented manifold. A definite linear isomorphism $\upsilon : \Omega^k(M) \to \Omega^{n-k}(M)$ is *Hodge-like operator*, if there exists a metric tensor m of M and a linear isomorphism $\upsilon_m : \Omega^{n-k}(M) \to \Omega^{n-k}(M)$ such that $\upsilon = \upsilon_m \circ \star_m$ holds, where $\star_m$ is the Hodge-operator induced by m.

A Hodge-like operator υ is a kind of generalization of Hodge-operator $\star$: both υ and $\star$ are definite linear isomorphisms from $\Omega^k(M)$ to $\Omega^{n-k}(M)$, but υ need not be globally identifiable with any Hodge-operator $\star_m$ in the sense that $\upsilon = \star_m$ holds globally, but they are only identifiable up to a linear isomorphism $\upsilon_m : \Omega^{n-k}(M) \to \Omega^{n-k}(M)$ in the sense that $\upsilon = \upsilon_m \circ \star_m$ holds.

Finally, if the operators ϵ, μ, σ are not linear, but acceptable in the energy sense for describing the constitutive equations, still the above kind of decomposition ($\epsilon = \epsilon_m \star_m$) holds for any metric m. A difference is that the mappings ϵ_m, μ_m, and σ_m are not linear anymore.

4.7.3 Role of metric

The role of metric and geometry in formulating electromagnetic BVPs is to provide tools that help the initial identification of the BVP. Particularly, metric and distance measurements constitute the systematic tools we use to create a connection between model and observations. But for other than providing the connection, geometry is irrelevant in the formulations of electromagnetic BVPs. Furthermore, *because the same physics can be described with different metrics, it follows that physics that is described with the constitutive equations is in the relations itself, not in any particular decompositions based on our choice of metrics.*

Chapter 5

Dimensional reduction of electromagnetic boundary value problems

In this chapter, we develop a symmetry-based theory of the dimensional reduction of electromagnetic BVPs. The theory explains when a BVP can be solved as a lower-dimensional BVP, and how the latter can be formulated. The theory encompasses static and time-harmonic problems as well as 1D- and 2D-problems.

Dimensional reduction is an area where classical vector analysis is not a natural tool: vector analysis is built primarily on three dimensions, and some of the structures have no natural counterparts in other dimensions. This is in striking contrast to the tools of differential geometry, which are inherently independent of dimension and virtually custom-made for the needs of dimensional reduction.

The theory of the dimensional reduction of electromagnetic BVPs presented here uses the conceptual tools of differential geometry. The theory builds exclusively on *symmetry* and is *completely coordinate- and metric-free*. For example, the theory assumes only appropriate invariances and makes no assumptions of some field components being zero in some special coordinate system. The theory includes a *dimensional reduction theorem, which provides a sufficient condition as to when a BVP can be solved as a lower-dimensional BVP*. The observer structure discussed in chapter 4 is an essential tool in the theory, and it is used to formulate lower-dimensional BVPs.

Dimensional reduction is based on symmetries that characterize particular invariances of objects of which a BVP consists: invariances of differential equations, fields governed by the equations, source fields, boundary values, constitutive equations, and cohomology conditions under a differentiable group action of some Lie group G. The terms *group invariant* or G*-invariant* are used for fields, boundary values, or whatever object whose invariance is to be considered. The G-invariance of these objects requires the existence of a suitable group action on the BVP domain, and it is the task of the modeler to recognize and make use of this group action.

Symmetry transformations are diffeomorphic mappings from the domain to itself. For example, the transformations can be translations or rotations of the domain. Notice that the differential equations, including Maxwell's equations, expressed with the exterior derivative are diffeomorphism-invariant and thus automatically invariant under all the group actions we study. The symmetry transformations of the domain induce group actions for the fields by the pullback and pushforward of the transformations. G-invariance of the constitutive equations means that operators ϵ, μ, and σ commute with the pullbacks of the domain's symmetry transformations. Symmetry transformations need not be isometries, which reflects the fact that symmetry and dimensional reduction are independent of the metric. Self-similar antennas such as log-periodic antennas serve as an example of non-isometric symmetries.

Time-harmonic fields, an example of invariant fields, also show that invariance is more flexible than strict constancy: fields at different time instants are equal only up to some complex-valued mapping (in the time-harmonic case, equality is up to the mapping $e^{j\alpha t}$). The real- or complex-valued mapping is denoted by h, and we talk about (G, h)-invariant fields. The theory of dimensional reduction constructed here is based on general (G, h)-invariances.

A BVP is said to be (G, h)*-invariant* if the source fields, boundary values, constitutive equations, and cohomology conditions of the fields are (G, h)-invariant. A major result of the theory is that *if a* (G, h)*-invariant BVP has a unique solution, then the solution is also* (G, h)*-invariant*. Thus an invariance under a group action results in redundancies that can be used to *reduce the size of the domain* of a (G, h)-invariant BVP: because the solution fields are (G, h)-invariant, the fields can be reconstructed over the domain manifold from knowledge of fields over a suitable subdomain. Furthermore, the "larger" the group G, the "smaller" the subdomain is needed to reconstruct fields over the whole domain. The basic requirement of dimensional reduction is that the group G be large enough to make the *dimension* of the subdomain smaller than the domain of the original BVP. This requires that G be a *one- or higher-dimensional Lie group*.

The theory will show that there is a *canonical submanifold to serve as the domain of the lower-dimensional BVP*: symmetry transformations of the domain manifold M divide M into equivalence classes called orbits. If an invariant field is known at some point of an orbit, then by its invariance, the field is known at all other points of the orbit. On the other hand, the domain of the lower-dimensional BVP must be a manifold. Hence a regular submanifold of M that contains exactly one point from each orbit is a suitable domain for the lower-dimensional BVP. The canonical choice for a suitable submanifold is the orbit space or the set of all orbits with a manifold structure diffeomorphic to all those regular submanifolds. Furthermore, there is no canonical metric for the orbit space, which agrees with the fact that the (G, h)-invariance of a BVP is independent of a metric.

The formulation of lower-dimensional BVPs is based on the observer structure. We must choose a G-invariant pair (T, τ) such that T is tangent to the orbits, and τ defines a regular submanifold containing one point from each orbit. Then the observer decomposes the fields, and the geometric components ω_τ and $i_T\omega$ of form ω are the fields to be solved in the lower-dimensional BVP. Differential equations, boundary values, and constitutive equations for the geometric components in the orbit space are then induced from the higher-dimensional BVP. The lower-dimensional BVP depends on the choice of observer though the possibility for *dimensional reduction, of course, does not depend on the choice of observer*.

Finally, even though dimensional reduction is based on symmetry, the structures and concepts needed to explain reductions under continuous symmetries are, in some parts, quite different from those under discrete symmetries [6] such as mirror symmetry. Particularly, major differences appears between continuous and discrete symmetries in the case of differential and constitutive equations. Thus as a whole, reductions under discrete symmetries cannot be generalized straightforwardly to reductions under continuous symmetries, nor are all discrete symmetry results simply special cases resulting from continuous symmetries.

5.1 Group action on a BVP domain

The theory of dimensional reduction is based on a few basic axioms, which restrict possible symmetry groups and their group actions on the BVP domain. This section lays down two such axioms and gives a few examples of group actions. The BVP domain is a manifold-with-boundary M.

Axiom 5.1. The symmetry group G is a Lie group that is a product of connected one-dimensional Lie groups.

For example, G could be $\mathbb{R}$, S^1, $\mathbb{R} \times S^1$, or $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. This axiom is not so restrictive as it may appear at the first glance, because the unconnectedness of the Lie groups would probably not add any practical cases to the theory, and certainly would make the theory much more technical. Furthermore, all connected two-dimensional Lie groups are products of one-dimensional Lie groups (see section 3.6). Notice that it is not necessary to assume a connectedness of G, or that G is a product of connected one-dimensional groups for the next axiom to make sense.

Axiom 5.2. There exists an effective, differentiable group action $F : G \times M \to M$ of a Lie group G on M such that for each $g \in G$ the mapping $F_g : M \to M$, defined as $F_g(p) = F(g, p)$, is a diffeomorphism.

Mappings F_g are symmetry transformations of M. Because the action is effective (see Definition 2.5), the mapping $g \mapsto F_g$ is a group isomorphism from G to the group of symmetry transformations. The isomorphism allows identification of the two groups, and the notation g is used for the transformation F_g. Henceforth, the phrase "group G acts on manifold M" is used to refer to a group action that satisfies axioms 5.1 and 5.2.

Because Axiom 5.2 is somewhat abstract, we now give some concrete examples of group actions that satisfy the assumption.

Example 5.1.1. Domain, which is an infinitely long straight rectangular waveguide (Figure 5.1), has translational symmetry, and its symmetry transformations are translations in the direction of the waveguide. Its orbits are one-dimensional submanifolds, lines parallel to the waveguide. In this example, the group G is $(\mathbb{R}, +)$, and the domain M can be regarded as a subset of $\mathbb{R}^3$ such that the waveguide is oriented along the z-direction. If we denote the points of M as triplets (x, y, z), the group action F and symmetry transformations F_g are as follows:

$$F : (g, (x, y, z)) \mapsto (x, y, z + g)$$
$$F_g : (x, y, z) \mapsto (x, y, z + g).$$

In this example, the action F is free (Definition 2.5).

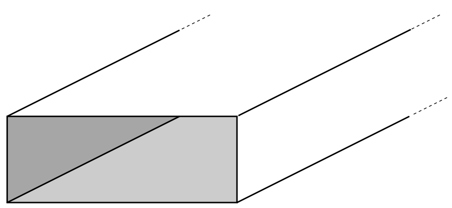

Figure 5.1: Waveguide with translational symmetry.

Before further examples, let us specifically comment on group actions on manifolds. A BVP is not, of course, symmetric with respect to all possible group actions that can be defined on its domain: if the manifold is $\mathbb{R}^3$, as it often is in electromagnetic modeling, clearly there is a multitude of group actions then satisfy Axiom 5.2. However, the space is not empty but contains materials and sources that must be symmetric for the BVP to be symmetric. The symmetry of fields and constitutive equations is described in terms of group actions on manifolds, as defined in the next two sections. Thus a group action on a BVP domain is a mathematical tool used to describe the symmetry of the BVP, and it is the *task of the modeler to recognize a suitable action under which the BVP is symmetric.*

Example 5.1.2. A second example is rotational symmetry in a plane (Figure 5.2). The Lie group is now S^1, and the symmetry transformations are rotations around point p. The action is now effective but not free because the point p is the *fixed point* of the symmetry transformations: the rotations do not "move" the point p but map it to itself. Thus all the orbits are circles centered at p, excluding the orbit Gp, which contains only the point p. The point p is referred to as a singular point, and the orbit Gp is referred to as a singular orbit.

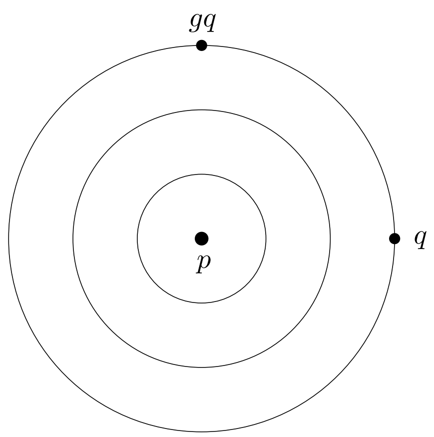

Figure 5.2: Rotational symmetry in a plane. Symmetry transformations are rotations of the points of the plane about point p. Point q and all the other points of the plane except p are rotated 90 degrees counterclockwise by a symmetry transformation g. The circles represent orbits.

Definition 5.1. Let group G act on manifold M. A point p of M is *singular* under the action if p is a fixed point for a symmetry transformation that is not the identity mapping of M. An orbit of G is *singular* under the action if there exist higher-dimensional orbits.

Remark 5.1.1. We will later construct an additional axiom about singular orbits but only when we are forced to do so.

The singular orbit in the second example derives from an *effective but non-free action of a compact Lie group* S^1. However, if the point p is removed from M, the action is free and there are no singular orbits.

Example 5.1.3. A third example of group actions that satisfy Axiom 5.2 is cylindrical symmetry, which also exemplifies group action with singular orbits. Now $G = \mathbb{R} \times S^1$ and the symmetry transformations of $M = \mathbb{R}^3$ are compositions of translations and rotations such that the translations are parallel to the axis of rotation. The orbits are two-dimensional submanifolds, which are cylindrical surfaces except for the axis of rotation, which is a one-dimensional singular orbit. Notice that the points of the rotational axis are not fixed points of translations, but that only the rotations fix the points. Therefore, the definition of a singular point requires only that a symmetry transformation exist that is not the identity mapping.

5.2 Group-invariant fields

This section defines (G, h)-invariant vector fields and differential forms to express the symmetry of BVPs. The group action of G on M induces group action for vector fields via the pushforward of the symmetry transformations of M. For differential forms, the pullbacks of the transformations of M induce a mapping for differential forms, which in general resembles group action, and in the case of Abelian groups is a group action. In the following, $\mathbb{F}$ refers to either $\mathbb{R}$ or $\mathbb{C}$. The motivation for complex-valued mappings h comes from time-harmonic fields, where time-harmonic invariance is simpler and more convenient to define using complex-valued fields. Notice that if h is complex-valued, the differential forms are also complex valued.

Definition 5.2. Let group G act on manifold M and let $h : G \to \mathbb{F}$ be a Lie group homomorphism. Then a vector field X on M is (G, h)-*invariant* if

$$g_* X = h(g) X$$

holds for all $g \in G$. In pointwise terms: $(g_* X)(p) = h(g) X(p)$ holds for all $p \in M$, $g \in G$, or equivalently (by the definition of the pushforward), $g_*(p)(X_p) = h(g) X_{gp}$ holds for all $p \in M$, $g \in G$.

Definition 5.3. Let group G act on manifold M and let $h : G \to \mathbb{F}$ be a Lie group homomorphism. Then a differential form $\omega \in \Omega(M)$ is (G, h)*-invariant* if

$$g^*\omega \quad = \quad h(g)\omega$$

holds for all $g \in G$. In pointwise terms for a k-form: $(g^*\omega)_p(v_1, .., v_k) = h(g)\omega_p(v_1, .., v_k)$ holds for all $p \in M$, $v_1, .., v_k \in T_p(M)$, $g \in G$, or equivalently (by the definition of the pullback), $\omega_{gp}(g_*v_1, .., g_*v_k) = h(g)\omega_p(v_1, .., v_k)$ holds for all $p \in M$, $v_1, .., v_k \in T_p(M)$, $g \in G$.

A very important special case of (G, h)-invariance is $(G, 1)$-invariance, where $1 : G \to \mathbb{F}$ is a mapping with a constant value 1, i.e., every $g \in G$ is mapped to the multiplicative identity of $\mathbb{F}$. In this case, we use the term G*-invariance*. Figure 5.3 gives an example of G- and (G, h)-invariant vector fields under translational and rotational symmetry transformations.

Example 5.2.1. An important example of (G, h)-invariance are the familiar time-harmonic fields: let $M = N \times \mathbb{R}$, where N is a 3-manifold, and the symmetry transformations g are translations in time. A zero-form ω is time-harmonic if the point-wise relation $\omega(p, t + g) = e^{j\alpha g}\omega(p, t)$ holds for all $p \in N$, $t, g \in \mathbb{R}$, and for some fixed $\alpha \in \mathbb{R}$. In general, a k-form ω is time-harmonic if there exists $\alpha \in \mathbb{R}$ such that the relation $g^*\omega = e^{j\alpha g}\omega$ holds for all $g \in \mathbb{R}$.

Remark 5.2.1. It is possible to generalize the concept of (G, h)-invariance of k-forms (and vector fields) to include more general mappings h than just $\mathbb{F}$-valued mappings. For example, all that is required of mappings h is that they be Lie group homomorphisms from G to a Lie group consisting of mappings from $\Omega^k(M)$ to itself. Notice that these general mappings h would not be scalars as are $\mathbb{F}$-valued mappings. However, because the benefits to electromagnetic modeling of this generalization are not clear, it is not pursued here in detail.

Open Question 3. What benefits can we achieve and what new cases can we include in the theory of dimensional reduction, if we allow more general Lie group homomorphisms h than just $\mathbb{F}$-valued mappings?

The symmetry transformations of M induce group actions also for boundary values: the boundary values a of a k-form ω are restrictions of ω to the boundary ∂M, i.e., $t\omega = a$ holds. Because each transformation g is a diffeomorphism, its restriction to the boundary ∂M is a diffeomorphism $g_\partial : \partial M \to \partial M$ (see Lemma 3.1). Thus the (G, h)-invariance of boundary values is defined as follows:

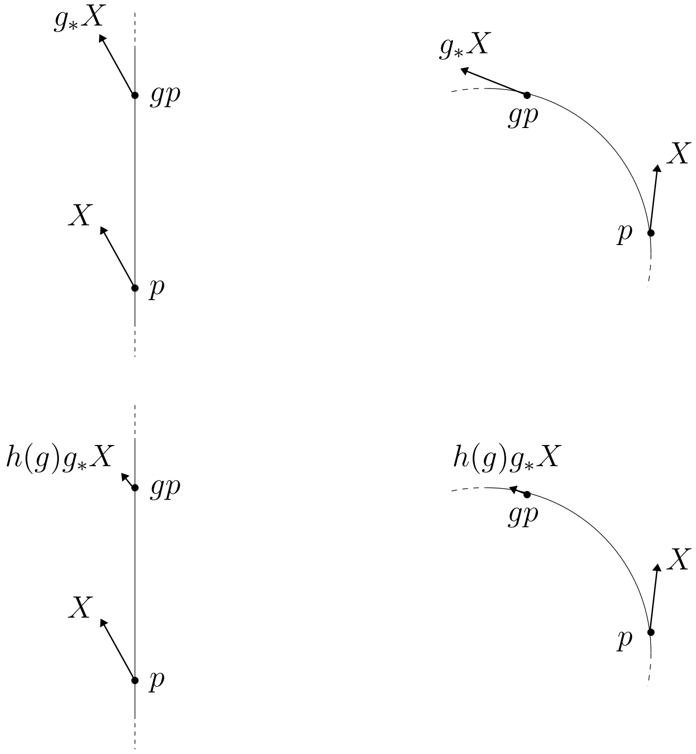

Figure 5.3: Examples of G- and (G, h)-invariant vector fields. Top: A G-invariant vector field under translational and rotational symmetry transformations. The vector at point gp is the pushforward of the vector X at point p under the symmetry transformation g. Bottom: A (G, h)-invariant vector field under the same symmetry transformations.

Definition 5.4. Let group G act on manifold M, and let $h : G \to \mathbb{R}$ be a Lie group homomorphism. Then boundary values $a \in \Omega^k(\partial M)$ for some k-form on M are (G, h)-*invariant* if

$$ {g_\partial}^* a \;=\; h(g) a $$

holds for all $g \in G$.

Boundary values of a (G, h)-invariant form are also (G, h)-invariant:

Proposition 5.1. Let ω be a (G, h)-invariant differential form on M. Then its restriction to the boundary ∂M is also (G, h)-invariant.

Proof: The claim follows from the commutation rule $t \circ g^* = {g_\partial}^* \circ t$, which was proven in Theorem 4.3. □

The definition of (G, h)-invariance of fields is global, but to use it in differential equations, we need to express it locally. Here Axiom 5.1 is required because the (G, h)-invariance of the fields can be expressed locally in an equivalent way with the Lie derivative if the group G is a connected one-dimensional Lie group: the symmetry transformations of a connected one-dimensional Lie group constitute a one-parameter group of transformations. That is, the symmetry transformations can be parameterized with real numbers in a smooth manner such that the parameterization respects the group structures: if $F : G \times M \to M$ is a group action, then $F_{a+b} = F_a \circ F_b$ holds for all $a, b \in \mathbb{R}$. Smooth parameterization with real numbers makes it possible to define the Lie derivative of (G, h)-invariant forms and the derivatives of the $\mathbb{F}$-valued mapping h. Furthermore, the Axiom 5.1 that the Lie group G is a product of connected one-dimensional Lie groups makes it possible to study the group action as separate actions of connected one-dimensional Lie groups, one at a time.

Let us next see how the group action $F : G \times M \to M$ can be represented as a one-parameter group of transformations. If G is not compact, it is isomorphic to $\mathbb{R}$, i.e., there is a Lie group isomorphism $\beta : \mathbb{R} \to G$ such that $\beta(a + b) = \beta(a) \cdot \beta(b)$ holds for all $a, b \in \mathbb{R}$. Let i_M be the identity mapping of M, and let us denote by $\beta \times i_M : \mathbb{R} \times M \to G \times M$ the mapping defined by $(a, p) \mapsto (\beta(a), i_M(p))$ for all $a \in \mathbb{R}$, $p \in M$. With the mapping $\beta \times i_M$ we can represent the action F as a 1-parameter group of transformations φ_β as follows (the subindex β indicates that the representation depends on β):

$$\varphi_\beta = F \circ (\beta \times i_M).$$

If the mappings $(\varphi_\beta)_t : M \to M$ and $F_g : M \to M$ are defined by $(\varphi_\beta)_t(p) = \varphi_\beta(t, p)$ and $F_g(p) = F(g, p)$, then it is easy to show that $(\varphi_\beta)_t = F_{\beta(t)}$ holds for all $t \in \mathbb{R}$. If the group G is compact, it is isomorphic to S^1 via mapping $s : S^1 \to G$. Furthermore, the exponential function gives a Lie group homomorphism from $\mathbb{R}$ to S^1: $a \in \mathbb{R} \mapsto e^{ia} \in S^1$. The mapping $\beta(a) = s(e^{ia})$ is now a Lie group homomorphism from $\mathbb{R}$ to G, which can be used to represent the action F as a 1-parameter group of transformations φ_β. The parameterization β of G also represents the mapping $h : G \to \mathbb{F}$ as a mapping $h_\beta : \mathbb{R} \to \mathbb{F}$ by composition of the mappings, i.e., $h_\beta = h \circ \beta$ holds.

Each 1-parameter group of transformations φ_β induces a smooth nonzero vector field X_β that is everywhere tangent to the orbits of F: if we define $\phi_p : \mathbb{R} \to M$ by $\phi_p(t) = \varphi_\beta(t, p)$, then ϕ_p is a smooth curve through the point p of M. Thus the induced vector field X_β maps point p to the tangent vector $[\phi_p]$ (equivalence class of curves containing the curve ϕ_p). Furthermore, the induced vector field X_β is G-invariant:

Proposition 5.2. If φ_β is a 1-parameter group of transformations of M induced by the action $F : G \times M \to M$ and the Lie group homomorphism $\beta : \mathbb{R} \to G$, then the induced vector field is G-invariant.

Because the induced vector field X_β is everywhere tangent to the orbits, it can be used to define the directional derivative in the direction of the orbits. The Lie derivative (see Definition 3.57) of the field with respect to X_β gives the directional derivative, and the invariance of the field fixes the value of this derivative, as is shown in the next theorem:

Theorem 5.1. Let a one-dimensional connected Lie group G act on a manifold M such that Axiom 5.2 is satisfied. Let $\beta : \mathbb{R} \to G$ be a Lie group homomorphism and X_β the induced vector field. Furthermore, let a field ω be (G, h)-invariant. Then $\mathcal{L}_{X_\beta}\omega = h'_\beta(0)\omega$ holds.

Proof: To show the claim, substitute $(\varphi_\beta)_t = F_{\beta(t)}$ in the definition of the Lie derivative and use the (G, h)-invariance condition:

$$\begin{aligned}
(\mathcal{L}_{X_\beta}\omega) &= \lim_{t\to 0} \frac{((\varphi_\beta)_t^*\omega) - \omega}{t} = \lim_{t\to 0} \frac{(F_{\beta(t)}^*\omega) - \omega}{t} = \lim_{t\to 0} \frac{h(\beta(t))\omega - \omega}{t} \\
&= \lim_{t\to 0} \frac{(h(\beta(t)) - 1)\omega}{t} = \lim_{t\to 0} \frac{h(\beta(t)) - h(\beta(0))}{t}\omega \\
&= \lim_{t\to 0} \frac{h_\beta(t) - h_\beta(0)}{t}\omega = h'_\beta(0)\omega. \quad \square
\end{aligned}$$

Corollary 5.1. If a field ω is G-invariant, then $\mathcal{L}_{X_\beta}\omega = 0$ holds.

Remark 5.2.2. In the case of G-invariance, the vector field X_β need not be induced by any Lie group homomorphism β, but $\mathcal{L}_X\omega = 0$ holds, in fact, for all smooth nonzero vector fields X that are everywhere tangent to the orbits.

Example 5.2.2. The field ω has a time-harmonic invariance if there exists $\alpha \in \mathbb{R}$ such that $g^*\omega = e^{j\alpha g}\omega$ holds for all $g \in \mathbb{R}$. Because the Lie group homomorphism $\beta : \mathbb{R} \to \mathbb{R} = G$ can be chosen to be a trivial mapping $a \mapsto a$, it follows that $h_\beta(g) = e^{j\alpha g}$ and $h'_\beta(0) = j\alpha$ hold. Thus if X_β is the induced vector field (a smooth nonzero vector field everywhere in the direction of time), then $\mathcal{L}_{X_\beta}\omega = j\alpha\omega$ holds.

The equivalent Lie derivative expression for (G, h)-invariance is important because it can be used directly to simplify the differential equations after an observer structure has decomposed them. To simplify the notation, the modeler is assumed implicitly to choose the Lie group homomorphism $\beta :$ $\mathbb{R} \to G$ and denote $h'_\beta(0)\omega$ simply by $h'(0)\omega$, where 0 is the identity element of G. In addition, the induced vector field X_β is simply denoted by X.

Finally, two useful propositions about (G, h)-invariance.

Proposition 5.3. If ω is a (G, h)-invariant k-form, and X is a G-invariant vector field, then $i_X\omega$ is a (G, h)-invariant $(k-1)$-form.

Proof:

$$(g^*(i_X\omega))_p(V_1, .., V_{k-1}) = (i_X\omega)_{gp}(g_*V_1, .., g_*V_{k-1}) = \omega_{gp}(X, g_*V_1, .., g_*V_{k-1})$$
$$= \omega_{gp}(g_*X, g_*V_1, .., g_*V_{k-1}) = h\omega_p(X, V_1, .., V_{k-1}) = h(i_X\omega)_p(V_1, .., V_{k-1}).$$

Because this holds for all $p \in M$, $V_1, .., V_k \in T_p(M)$ and for all $g \in G$, the claim follows. □

Proposition 5.4. If ω is a (G, h)-invariant form, and α is a G-invariant form, then their wedge product $\alpha \wedge \omega$ is also a (G, h)-invariant form.

Proof: The claim follows from the fact that the pullback commutes with the wedge product:

$$g^*(\alpha \wedge \omega) = g^*\alpha \wedge g^*\omega = \alpha \wedge h\omega = h(\alpha \wedge \omega). \quad \square$$

These propositions will be used to establish the (G, h)-invariance of the geometric components of field ω when ω itself is (G, h)-invariant.

5.3 Group-invariant constitutive equations

It is intuitively clear that dimensional reduction requires that the material properties be invariant in some sense. Thus we must define the group invariance of the constitutive equations, which is the topic of this section. The group invariance of the constitutive equations requires a proper invariance for Hodge-like operators (see Definition 4.6), that describe the constitutive equations.

The material properties are modeled with Hodge-like operators ϵ, μ, and σ. If $D = \epsilon E$ holds and if both E and D are to be G-invariant, we have a clear requirement for the Hodge-like operator ϵ: *it should preserve this invariance.* In other words, Hodge-like operators should map a G-invariant E to a G-invariant D: let υ be a Hodge-like operator. Now if ω is G-invariant, or if $g^*\omega = \omega$ holds for all $g \in G$, then $\upsilon\, g^*\omega = \upsilon\, \omega$ holds. On the other hand, if $\upsilon\, \omega$ is G-invariant, then $g^*\upsilon\, \omega = \upsilon\, \omega$ holds. These equations imply that υ maps G-invariant fields to G-invariant fields if the following equation holds:

$$g^*\upsilon\, \omega \;\; = \;\; \upsilon\, g^*\omega \quad \forall g \in G.$$

That is, υ maps G-invariant fields to G-invariant fields if it commutes with the pullbacks of the symmetry transformations.

Definition 5.5. Let group G act on manifold M. A Hodge-like operator υ is G-*invariant* if it commutes with the pullbacks of the symmetry transformations or $\upsilon \circ g^* = g^* \circ \upsilon$ holds for all $g \in G$.

A Hodge-like operator υ is (G, h)-invariant if it maps (G, h)-invariant fields to (G, h)-invariant fields. This requires that the following additional equation hold.

$$h\upsilon_p \, \omega_p \;=\; \upsilon_p \, h \, \omega_p \quad \forall p \in M. \tag{5.1}$$

Because the Hodge-like operator υ is linear, the above equation always holds. Thus a G-invariant Hodge-like operator is automatically also (G, h)-invariant:

Proposition 5.5. If a Hodge-like operator is G-invariant, it is also (G, h)-invariant.

Remark 5.3.1. (G, h)-invariance of Hodge-like operators include so-called anisotropic materials. Thus anisotropy is not a problem for dimensional reduction, but nonlinear materials are problematic (see the next remark and the next section).

Remark 5.3.2. If mappings h are more general than $\mathbb{F}$-valued Lie group homomorphisms, the requirement in (5.1) may not be trivially true but it would, in fact, be an additional requirement for the (G, h)-invariance of Hodge-like operators. If the Hodge-like operators are not linear (nonlinear materials), it seems that only G-invariance is possible: the only Lie group homomorphisms h that satisfy the requirement in (5.1) seems to be the mapping $1 : G \to \mathbb{F}$, which maps all the elements of G to the multiplicative identity of $\mathbb{F}$.

5.4 Unique solution of an invariant BVP is invariant

It is usually assumed that if the sources, boundary values, and constitutive equations are all symmetric "in the same way," the solution fields are also symmetric "in the same way." This section provides a theorem that assures the solution fields to be (G, h)-invariant if a BVP has a unique solution and has (G, h)-invariant sources, boundary values, constitutive equations, and cohomology conditions.

The proof of the (G, h)-invariance of the solution field is based on assumed existence of a unique solution, which requires that we consider cohomology conditions. In practice, the cohomology conditions are given by integrals of solution fields over suitable submanifolds (de Rham's theorem) [24]. Because for the field argument integration is linear, the cohomology condition is a linear operator $\mathcal{H}$ for the fields. By the linearity of $\mathcal{H}$ we get a simple (G, h)-invariance condition for the cohomology condition: $\mathcal{H}(g^*\omega) = \mathcal{H}(h(g)\omega) = h(g)\mathcal{H}(\omega)$. Hence the following definition:

Definition 5.6. The cohomology condition of a field $\omega \in \Omega(M)$ is (G, h)-*invariant* if $\mathcal{H}(g^*\omega) = h(g)\mathcal{H}(\omega)$ holds.

In practice, the (G, h)-invariance of a cohomology condition $\mathcal{H}(\omega)$ means that if the condition is given by integrating ω over a submanifold c, then this integral multiplied by $h(g)$ is the same as integrating ω over the submanifold gc. This follows from the (G, h)-invariance of ω and Theorem 3.4:

$$h(g) \int_c \omega = \int_c g^*\omega = \int_{gc} \omega.$$

Definition 5.7. A BVP is (G, h)-*invariant* if its sources, boundary values, constitutive equations, and cohomology conditions are (G, h)-invariant.

Theorem 5.2. If a (G, h)-invariant electromagnetic BVP has a unique solution, then the solution is (G, h)-invariant.

Proof: Let us study the following generic BVP formulated on an n-dimensional manifold-with-boundary M:

$$\begin{aligned} dC &= Q \\ dK &= L \\ t^1 C &= c \\ t^2 K &= b \\ K &= \upsilon C \\ \mathcal{H}(C) &= e \\ \mathcal{H}(K) &= f. \end{aligned} \tag{5.2}$$

In the above BVP, C is a k-form and K is an $(n-k)$-form. Forms Q and L describe the sources. The boundary ∂M has two complementary parts, and the restrictions of the fields to these parts are denoted by t^1 and t^2. The constitutive equation is given by a Hodge-like operator υ. $\mathcal{H}$ is a linear operator that gives cohomology conditions for the fields C and K, making the solution unique. (In [29] it is shown how to discretize generic BVPs of the above type and their abstract error analysis is also presented.)

In the case of the generic BVP, the existence of a unique solution is equivalent to the following: let us choose $R_C = C_1 - C_2$ and $R_K = K_1 - K_2$, where C_1 and C_2 are fields satisfying the same equations as C, and K_1 and K_2 are fields satisfying the same equations as K. Then R_K and R_C satisfy the following equations:

$$\begin{aligned} dR_C &= 0 \\ dR_K &= 0 \\ t^1 R_C &= 0 \\ t^2 R_K &= 0 \\ R_K &= \upsilon\, R_C \\ \mathcal{H}(R_K) &= 0 \\ \mathcal{H}(R_C) &= 0. \end{aligned} \tag{5.3}$$

By the existence of a unique solution, these equations imply that $R_K = 0$ and $R_C = 0$ hold. Thus $C_1 = C_2$ and $K_1 = K_2$ hold, which is exactly the uniqueness of C and K.

Now the procedure to show that the solution is (G, h)-invariant is the same as in the above case of uniqueness. That is, we must show that the difference fields $g^*C - h(g)C$ and $g^*K - h(g)K$ satisfy the same BVP as described in (5.3) (Remember that C is (G, h)-invariant if $g^*C - h(g)C = 0$ holds for all $g \in G$). Now the sources, boundary values, constitutive equations, and cohomology conditions are assumed to be (G, h)-invariant or that equations

$$\begin{aligned} g^*Q &= h(g)Q \\ g^*L &= h(g)L \\ g^*_\partial c &= h(g)c \\ g^*_\partial b &= h(g)b \\ g^*\upsilon &= \upsilon\, g^* \\ \mathcal{H}(g^*C) &= h(g)\mathcal{H}(C) \\ \mathcal{H}(g^*K) &= h(g)\mathcal{H}(K) \end{aligned}$$

hold for all $g \in G$. Then the difference field $g^*C - h(g)C$ satisfies the following differential equation:

$$\begin{gathered} d(g^*C - h(g)C) = d(g^*C) - d(h(g)C) \\ = g^*(dC) - h(g)(dC) = g^*Q - h(g)Q = 0. \end{gathered}$$

Similar calculation shows that $d(g^*K - h(g)K) = 0$ holds. Moreover, the difference field $g^*C - h(g)C$ satisfies the following boundary values:

$$\begin{gathered} t^1(g^*C - h(g)C) = t^1(g^*C) - t^1(h(g)C) \\ = g^*_\partial(t^1C) - h(g)(t^1C) = g^*_\partial c - h(g)c = 0. \end{gathered}$$

In addition, $t^2(g^*K - h(g)K) = 0$ holds. The difference fields satisfy also the constitutive equation

$$g^*K - h(g)K = g^*\upsilon C - h(g)\upsilon C = \upsilon(g^*C - h(g)C).$$

Finally, the cohomology conditions of the difference fields are zero:

$$\mathcal{H}(g^*K - h(g)K) = \mathcal{H}(g^*K) - \mathcal{H}(h(g)K) = \mathcal{H}(g^*K) - h(g)\mathcal{H}(K) = 0$$

Thus because the difference fields satisfy the BVP in (5.3), the solution fields C and K are (G, h)-invariant. □

Remark 5.4.1. The proof of the above theorem uses the linearity of υ. Thus the above theorem is shown to be valid only for BVPs with linear materials. To include even a strict class of nonlinear operators υ would make the proof much more complicated. Furthermore, only G-invariance is defined for nonlinear operators, see Remark 5.3.2.

Open Question 4. Practical engineering problems include G-invariant BVPs with nonlinear materials. How to prove the above theorem for those cases? Moreover, how to expand the theory of dimensional reduction to include nonlinear materials?

5.5 Orbit space

Dimensional reduction means that a BVP on M can be solved as a lower-dimensional BVP on a lower-dimensional manifold N. N is a submanifold of M, and infinitely many valid submanifolds can be used as domains for the lower-dimensional BVP. Fortunately, a canonical choice exists for a valid submanifold: the orbit space or the set of all orbits of M with a suitable manifold structure. Existence of the orbit space is included in the sufficient conditions for dimensional reduction.

What are then the valid submanifolds of M under the action of a group G? The group invariance of the fields implies that if the values of a field are known at one point of each orbit, then the fields are completely known on the whole M. That is, if the value of a field is known at point p, the value at point q of the same orbit can be constructed by the pullback of the symmetry transformation that maps q to p (see definitions 3.50 and 5.3). Thus a valid submanifold is compatible with the orbits in the sense that it is canonically bijective to the set of all orbits M/G. In other words, a regular submanifold of M is compatible with the orbits if it contains exactly one point from each orbit. On the other hand, invariance does not imply that any regular compatible submanifold is the canonical one; therefore all of them must be equally good choices. Hence all the regular compatible submanifolds must be canonically diffeomorphic to each other: if N_1 and N_2 are valid submanifolds, the diffeomorphism maps each point p of N_1 to point q of N_2, which is in the same orbit as p. This implies that all the compatible regular submanifolds are canonically bijective to M/G; therefore, M/G can be given a unique manifold structure, making it the canonical choice for valid submanifolds.

5.5.1 Manifold structure

We begin constructing the orbit space by defining a topology for a set of all orbits M/G. Let $\pi : M \to M/G$ be the canonical projection, i.e., π maps every point of M to its orbit. With the help of the canonical projection, M/G inherits a natural topological structure that is compatible with the orbits: a subset U of M/G, which is a set of orbits, is open if and only if the points in the orbits of U constitute an open set in M. This topology is the finest topology that makes the mapping π continuous, i.e., the topology contains all the possible subsets of M/G such that the continuity of π is not lost. With this topology M/G is called the orbit space:

Definition 5.8. The set of orbits M/G with a topology is an *orbit space* if the topology is the finest topology that makes the mapping π continuous.

The finest topology that makes the mapping π continuous is often called the quotient topology. Notice that the orbit space, in the above topological sense, always exists. However, it may not be a Hausdorff space, and a manifold structure may thus not exist for [3]. We must assume here that the orbit space is a Hausdorff space. Consequently, only those symmetries that produce an orbit space that is a Hausdorff space are included in our theory of dimensional reduction. The lack of the Hausdorff property is a problem only for noncompact symmetry groups because for differentiable actions of compact Lie groups, the orbit space is always a Hausdorff space [9].

The manifold structure for the orbit space comes from the compatible regular submanifolds of M. The compatibility of regular submanifolds is formally defined using cross-sections:

Definition 5.9. A *cross-section* is a continuous mapping $\kappa : M/G \to M$ such that $\pi \circ \kappa = id_{M/G}$, i.e., $\pi(\kappa(p)) = p$ holds for all $p \in M/G$.

By definition, a cross-section maps each orbit to one of its points, but in a continuous manner. Each cross-section κ is a homeomorphism to its range

$S = \kappa(M/G)$ with the subspace topology because the cross-section is bijective to S, continuous, and its inverse (the restricted projection $\pi|_S$) is continuous. The homeomorphism property makes the image S topologically equivalent to the orbit space. Furthermore, the images of all cross-sections are canonically homeomorphic: the canonical homeomorphism between $\kappa_1(M/G)$ and $\kappa_2(M/G)$ is given by $\kappa_2 \circ \kappa_1^{-1}$.

The orbit space can be given a compatible manifold structure if a cross-section exists whose range is a regular submanifold of M. Observe that the orbit space can have a boundary (even if M does not); therefore, the range of the cross-sections must be regular submanifolds-with-boundary (of course, the boundary may be empty). A submanifold-with-boundary A of M is a manifold-with-boundary A together with a suitable differentiable mapping $A \to M$ (see Definition 3.36). To give the orbit space a compatible manifold structure, we must assume here that such a cross-section exists. Next we define compatible regular submanifolds-with-boundary and call them G-reduced domains.

Definition 5.10. A regular submanifold-with-boundary A of M is a G-*reduced domain* if a cross-section κ exists such that $\kappa(M/G) = A$.

Remark 5.5.1. If the action of G on M is free, then not only do the orbits form a foliation of M, but also any G-reduced domain can be used to produce a foliation: let A be a G-reduced domain. If the set $\{gp \in M \,|\, p \in A\}$ is denoted by A_g, then $\{A_g\}_{g \in G}$ is a foliation of M. If the action is effective but not free, then there are singular points that belong to more than one G-reduced domain and no foliation is possible.

We can now induce a differentiable manifold structure for the orbit space by requiring that a cross-section κ defining a G-reduced domain A be a diffeomorphism. That is, the manifold structure for the orbit space induced from A is a pullback structure defined by κ. Because invariance does not give a canonical choice of a G-reduced domain, all of them must be equally good choices. Formally, this means that all the G-reduced domains must be canonically diffeomorphic: the canonical homeomorphisms $\kappa_j \circ \kappa_i^{-1}$ are diffeomorphisms. Then the pullback manifold structure for the orbit space is independent of the choice of G-reduced domain:

Theorem 5.3. If a G-reduced domain exists and all the G-reduced domains are canonically diffeomorphic, the orbit space can be given a unique pullback manifold structure, i.e., the pullback structure is independent of the choice of G-reduced domain.

Proof: Let the orbit space M/G have a pullback structure induced by κ_1 that defines a G-reduced domain A_1: if $\mathcal{U}_1 = \{(U_i, \phi_i)\}$ is an atlas of A_1, then the corresponding pullback atlas for M/G is $\kappa_1^*\mathcal{U}_1 = \{(\kappa_1^{-1}(U_i), \phi_i \circ \kappa_1)\}$. Let (A_2, κ_2) be another G-reduced domain. Because it is canonically diffeomorphic to A_1, it has the following atlas:

$$\mathcal{U}_2 = (\kappa_1 \circ \kappa_2^{-1})^*\mathcal{U}_1 = \{((\kappa_2 \circ \kappa_1^{-1})(U_i), \phi_i \circ (\kappa_1 \circ \kappa_2^{-1}))\}.$$

The corresponding pullback atlas for the orbit space is

$$\begin{aligned}\kappa_2^*\mathcal{U}_2 &= \{(\kappa_2^{-1}((\kappa_2 \circ \kappa_1^{-1})(U_i)), (\phi_i \circ (\kappa_1 \circ \kappa_2^{-1})) \circ \kappa_2)\} \\ &= \{(\kappa_2^{-1}(\kappa_2(\kappa_1^{-1}(U_i))), (\phi_i \circ (\kappa_1 \circ \kappa_2^{-1} \circ \kappa_2)\} \\ &= \{(\kappa_1^{-1}(U_i), \phi_i \circ \kappa_1)\} = \kappa_1^*\mathcal{U}_1.\end{aligned}$$

Thus because the pullback atlases are the same for arbitrary G-reduced domains A_1 and A_2, the pullback manifold structure for the orbit space is independent of the choice of G-reduced domain. □

The orbit space with its canonical pullback manifold structure is the canonical choice for the domain of lower-dimensional BVPs: because the elements of the orbit space are themselves orbits, it treats all points of each orbit equally, which is in contrast to G-reduced domains, which single out one point from each orbit. That is, compatibility with orbits is trivial for the orbit space. Henceforth, orbit space means the set M/G with a canonical pullback manifold structure. Observe that only those symmetries that produce an orbit space with a canonical pullback manifold structure are included in our theory of dimensional reduction. Thus the existence of a canonical pullback manifold structure for the orbit space is included in the sufficient conditions for dimensional reduction.

The following proposition states that the canonical projection π is a differentiable mapping, in which case its pushforward is defined.

Proposition 5.6. The canonical projection $\pi : M \to M/G$ is a differentiable mapping.

Proof: Let M be an n-dimensional manifold and M/G a k-dimensional manifold. Select a G-reduced domain A and let $p \in A$. There exists a cubic-centered coordinate system (V, ϕ) of M which contains p and a neighborhood $U \subset A$ of p such that the coordinates for the points of U under the chart (V, ϕ) are of form $[x_1, .., x_k, 0, .., 0]$ [62, page 28]. The corresponding pullback coordinates for M/G are now $[x_1, .., x_k]$. Then the canonical projection can be represented in these coordinates as the mapping $[x_1, .., x_k, .., x_n] \mapsto [x_1, .., x_k]$, which is clearly a differentiable mapping. □

Lastly, notice that the orbit space, and thus the G-reduced domains, can be manifolds-with-boundary even if M is not. For example, if $M = \mathbb{R}^3$ and the actions are rotations about the z-axis, then the positive xz-plane ($x \geq 0$) is a G-reduced domain, and the z-axis is its boundary. The boundary consists of orbits with a lower dimension than most of the orbits, making them thus singular orbits. Singular orbits are significant in defining the boundary values of lower-dimensional BVPs.

5.5.2 Metric structure

A metric structure can be induced for the orbit space M/G from M: select any G-reduced domain and endow it with the subspace metric. Then require the corresponding cross-section to be an isometry. Alternatively, the cross-section can be used to pull back the metric tensor of M to M/G. However, this metric for the orbit space depends on the choice of G-reduced domain. For example, consider a translational symmetry (Figure 5.4). A cross-section plane that is everywhere orthogonal to the orbits is a G-reduced domain as is any cross-section plane that is askew with respect to the orbits. Then clearly the distance between any two points on the orthogonal plane is different as the distance between the corresponding points on the askew plane (the

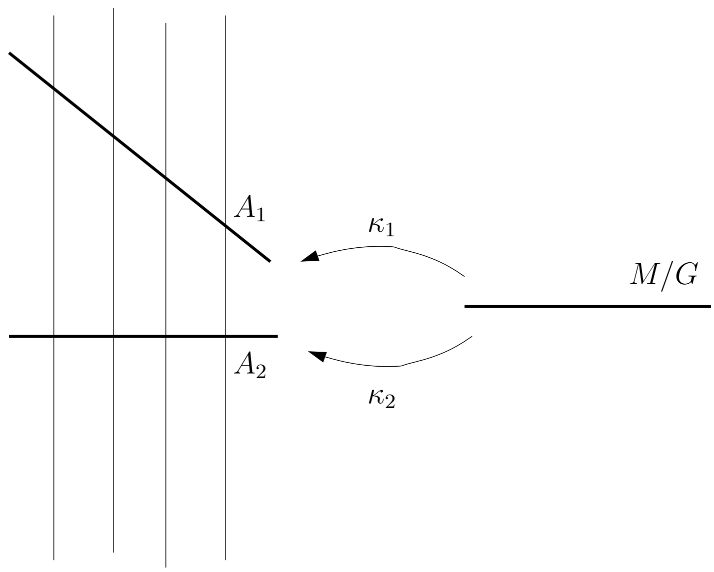

Figure 5.4: Different metrics for the orbit space M/G. Vertical lines represent orbits. A_1 and A_2 are G-reduced domains and κ_1 and κ_2 are the corresponding cross-sections.

corresponding points are those in the same orbits). Thus *there is no canonical choice of metric for the orbit space*.

The symmetry of BVPs has been defined without metrics, i.e., the invariance of fields, boundary values, constitutive equations, and cohomology conditions are defined without any reference to metrics. Of course in some cases the symmetry transformations may be isometries for a suitable choice of metric, but this does not play any role in the definitions of the invariance. Thus *symmetry of BVPs is independent of metrics* and this is in line with the fact that the orbit space does not have a canonical metric. For lower-dimensional BVPs, this means that everything can be defined without a reference to a metric. Even the constitutive equations, the only part of BVPs with some connection to a metric, are directly induced from the original constitutive equation without a metric. Therefore, *dimensional reduction can be explained without any metric in the orbit space*.

5.6 Lower-dimensional BVPs for one-dimensional symmetry groups

In this section, we derive in detail the lower-dimensional BVPs for a symmetry group G, which is a connected one-dimensional Lie group. That is, we study cases where the dimension of a BVP is reduced by one. Such cases comprise static and time-harmonic problems that are results of dimensional reduction under time invariance. In addition, classical static 2d-problems with translational and rotational symmetry are instances of lower-dimensional BVPs for one-dimensional symmetry groups.

A major difference from the usual application of dimensional reduction is that we do not assume any components to be zero in some coordinate system, but rather work directly in the domain manifold and assume only invariance of the BVP. That is, this section starts from a generic BVP being (G, h)-invariant under an action of G such that the action satisfies Axiom 5.2. In addition, we assume that the BVP has a unique solution, in which case Theorem 5.2 guarantees that the solution is (G, h)-invariant. Furthermore, we assume that the orbit space exists. Then we formulate the lower-dimensional, or *reduced BVP* for short, on the canonical lower-dimensional domain, which is the orbit space.

The domain of the generic BVP is a manifold-with-boundary M of dimension n. The generic BVP on M consists of two differential equations, expressed with the exterior derivative d_M, one homogeneous and the other inhomogeneous. The source field is a $(k+1)$-form Q, and the two unknown fields are a k-form C and an $(n-k)$-form K. The constitutive equation between K and C is given by a Hodge-like operator υ, and the boundary values for C and K are given by a k-form c and an $(n-k)$-form b, respectively. The BVP is then:

$$\begin{aligned} d_M C &= Q \\ d_M K &= 0 \\ t^1_M C &= c \\ t^2_M K &= b \\ K &= \upsilon_M C. \end{aligned} \tag{5.4}$$

We begin with geometric decompositions of the fields and the exterior derivative, which are based on a choice of a G-invariant observer structure (T, τ). The observer must be compatible with the invariance in the following sense: τ corresponds to some G-reduced domain A such that $\tau(v) = 0$ holds for all v tangent to A, and T is everywhere tangent to the orbits. The geometric components C_τ, K_τ, $i_T C$, and $i_T K$ are the fields to be solved for the reduced BVP. Notice that in the orbit space, the reduced BVP depends on the choice of observer though the solution of the original higher-dimensional BVP, of course, does not. Furthermore, notice that the geometric components are horizontal fields and can thus be naturally pulled back to the $(n-1)$-dimensional orbit space N. When the geometric components are solved, the geometric decomposition and the invariance of the fields show how to construct the solution for the original BVP in (5.4): $C = C_\tau + \tau \wedge i_T C$ holds in the points of the G-reduced domain A, where C_τ and $i_T C$ are the solutions of the reduced BVP and where τ is given. Then the (G, h)-invariance of C expands the solution to the whole M.

Then we derive "reduced differential equations" for the reduced BVP in the orbit space M/G which we denote here for simplicity by N. The derivation is based on $(3+1)$-decompositions of Maxwell's equations shown in chapter 4. The chosen observer (T, τ) induces $((n-1)+1)$-decompositions of the differential equations in (5.4). Then the invariance of the fields C and K simplifies the equations, eliminating from them the derivatives in the direction of T.

"Reduced boundary values" are derived first for free actions, which have no singular orbits. Then the case of effective but not free action is examined. At the boundary of the orbit space we have points that are not part of the boundary of M. These points are singular orbits, and we now propound an additional axiom that such orbits always reside at the boundary of the orbit space.

The choice of a G-reduced domain (that defines the observer) is arbitrary; in particular, no orthogonality is assumed for the orbits. This forces us to consider a very general type of constitutive equations for the reduced BVPs: we must assume that K_τ and $i_T K$ are both dependent on C_τ and $i_T C$, and we show how this assumption decomposes the original operator υ_M into four "reduced operators," which are not all Hodge-like operators. These derived constitutive equations couple three of the reduced differential equations, but in some cases the equations are simple enough to have two uncoupled reduced BVPs.

5.6.1 Geometric decomposition of fields and the exterior derivative

Geometric decomposition of fields is based on a similar observer structure as that introduced in chapter 4. However, the pair (T, τ) that defines the projections is constructed with help of G-invariance: if the BVP is (G, h)-invariant, the vector field T is induced by the action and by some Lie group homomorphism $\beta : \mathbb{R} \to G$ chosen by the modeler (see section 5.2). The induced vector field T is smooth, nonzero, G-invariant, and everywhere tangent to the orbits. If the BVP is G-invariant, then T can be any nonzero, smooth, G-invariant vector field that is everywhere tangent to the orbits. In the case of free group action, T is everywhere nonzero, but if the action is effective but not free, there are singular orbits that contain only one point. At singular points, T is zero and at every other point of M nonzero. Because these singular orbits are assumed to be at the boundary of the orbit space, they do affect only the boundary values.

To define the one-form τ, we choose any G-reduced domain A. Then τ is a smooth G-invariant one-form on M such that $\tau(v) = 0$ holds for all v in the tangent bundle of A, and $\tau(T) = 1$ holds everywhere except at the singular points where τ is left undefined. Notice that τ is defined completely on A (except at singular points), and G-invariance then extends the definition to the whole M (excluding the singular points). In the case of free action, the one-form τ could also be defined as follows: select a G-reduced domain A, in which case $\{A_g\}_{g \in G}$ is a foliation of M (see Remark 5.5.1). Then τ is such that it returns zero for every vector tangent to some leaves of the foliation, and $\tau(T) = 1$ holds everywhere. Furthermore, because τ is G-invariant, its Lie derivative with respect to T is zero, or $\mathcal{L}_T \tau = 0$ holds.

In singular orbits, T is zero, in which case $i_T \omega = 0$ holds and we set $\omega_\tau = \omega$. Then the geometric decompositions, which always hold at the interior points of M, of the fields K, C, and Q are

$$\begin{aligned} K &= K_\tau + \tau \wedge i_T K \\ C &= C_\tau + \tau \wedge i_T C \\ Q &= Q_\tau + \tau \wedge i_T Q. \end{aligned}$$

The exterior derivative d_M has the following decomposition:

$$d_M = d_\tau + \tau \wedge i_T d_M.$$

Let $\kappa : N = M/G \to M$ be the cross-section that defines the G-reduced domain A and let us denote its pullback by f or $f = \kappa^* : \Omega(M) \to \Omega(N)$. It is then easy to show that the following equations hold:

$$\begin{aligned} f\tau &= 0 \\ f\omega &= f\omega_\tau \\ f \circ d_\tau &= d_N \circ f. \end{aligned} \tag{5.5}$$

The (3+1)-decomposition of differential equations given in section 4.2.2 was based on the property $d_M\tau = 0$. To use the results of section 4.2.2, we must show that $d_M\tau = 0$ holds also for the G-invariant τ defined in this section. For the proof, we need some preliminaries: let $\{X_1, ..., X_{n-1}, T\}$ be a smooth G-invariant basis field for the tangent spaces of M such that at each point $p \in A \subset M$, $\{X_1, ..., X_{n-1}\}$ is a basis of T_pA, and T is a basis of T_pGp (tangent space of the orbit). Then at each $p \in A$, there is a decomposition of the tangent space $T_p(M) = T_pA \bigoplus T_pGp$. We call this basis field on M the geometric basis. The dual basis $\{dx_1, ..., dx_{n-1}, \tau\}$ for one-forms is defined as usual:

$$dx_i(X_j) = \delta_{ij}, \qquad dx_i(T) = 0, \qquad \tau(X_i) = 0, \qquad \tau(T) = 1.$$

The corresponding basis for two-forms and higher-degree forms are then defined as wedge products of the elements of the dual basis; e.g., in a three-dimensional M for two-forms, we have $\{dx_1 \wedge dx_2, dx_1 \wedge \tau, dx_2 \wedge \tau\}$. These bases for differential forms are also called geometric bases. Now we prove the proposition.

Proposition 5.7. If (T, τ) is a G-invariant pair such that T is in the direction of the orbits, then $d_M\tau = 0$ holds.

Proof: The G-invariance of τ is equivalent to $\mathcal{L}_T\tau = 0$, in which case we have $\mathcal{L}_T\tau = d_M(i_T\tau) + i_T(d_M\tau) = d_M 1 + i_T(d_M\tau) = i_T(d_M\tau) = 0$. This means that in the geometric basis the two-form $d_M\tau$ has no terms containing τ. On the other hand, $f\tau = 0$ holds by (5.5); therefore, $f d_M \tau = 0$ holds also because the pullback and the exterior derivative commute. Then on the points of A, each term of $d_M\tau$ in the geometric basis must contain τ or $d_M\tau = 0$ holds. Consequently, the two requirements ($i_T(d_M\tau) = 0$ and $f d_M\tau = 0$) can hold in A at the same time only when $d_M\tau = 0$ holds in A. To show that $d_M\tau = 0$ in M, it is enough to show that $d_M\tau$ is G-invariant: $g^*(d_M\tau) = d_M(g^*\tau) = d_M\tau$ holds for all $g \in G$. Thus $d_M\tau$ is G-invariant, and $d_M\tau = 0$ then holds in M. □

Finally, it is clear that for numerical solution a chart (coordinate system) is needed, and that not all charts are equally convenient for that. Presumably, under a convenient chart the orbits are pure translations (or lines) parallel to one of the coordinate axes, and G-reduced domain A is "orthogonal" to that axis, in the sense that it is a level set of the coordinate defining the direction of the orbits. For example, in a rotationally symmetric case, where the G-reduced domain A is chosen to be a plane everywhere orthogonal to the orbits, it is convenient to choose cylindrical (r, ϕ, z)-coordinates, because under those charts orbits are lines parallel to the ϕ-direction (the angle coordinate) and the G-reduced domain A is a level set of ϕ-coordinate. Under this kind of charts, the vector field T can be chosen to be the coordinate basis vector in the direction of the orbits, and the one-form τ is the corresponding dual basis one-form.

5.6.2 Differential equations for reduced BVPs

The differential equations that hold in M are the original equations for fields K and C:

$$
\begin{aligned}
d_M K &= 0 \\
d_M C &= Q.
\end{aligned}
$$

These equations are decomposed by (T, τ) as follows (see section 4.2.2):

$$
\begin{aligned}
d_\tau(i_T K) &= \mathcal{L}_T K_\tau \\
d_\tau C_\tau &= Q_\tau \\
d_\tau(i_T C) &= i_T Q + \mathcal{L}_T C_\tau \\
d_\tau K_\tau &= 0.
\end{aligned}
\tag{5.6}
$$

The (G, h)-invariance of K and C fixes Lie derivatives for fields K_τ and C_τ: if the generic BVP in (5.4) has a unique solution, the fields K and C are also (G, h)-invariant by Theorem 5.2. Then the next proposition shows that the horizontal components are also (G, h)-invariant:

Proposition 5.8. If $\omega = \omega_\tau + \tau \wedge i_T\omega$ is (G,h)-invariant, and if (T,τ) is G-invariant, then the horizontal component ω_τ is (G,h)-invariant.

Proof: Proposition 5.3 shows that $i_T\omega$ is (G,h)-invariant, and then Proposition 5.4 shows that $\tau \wedge i_T\omega$ is (G,h)-invariant. Now because ω is (G,h)-invariant, it follows that ω_τ must also be (G,h)-invariant. □

Now because K_τ and C_τ are (G,h)-invariant, by Theorem 5.1 they satisfy the following Lie derivative equations:

$$\begin{aligned}\mathcal{L}_T K_\tau &= h'(0)K_\tau \\ \mathcal{L}_T C_\tau &= h'(0)C_\tau.\end{aligned}$$

Remember that $h'(0)K_\tau$ is shorthand for $h'(\beta(0))\beta'(0)K_\tau$, where $\beta : \mathbb{R} \to G$ is a user-defined Lie group homomorphism (see section 5.2). If we cannot assume the existence of unique solution, we must directly assume (G,h)-invariance of fields K and C to derive the above Lie derivative equations for K_τ and C_τ.

With the Lie derivative equations, the decomposed equations in (5.6) are simplified to the following:

$$\begin{aligned}d_\tau(i_T K) &= h'(0)K_\tau \\ d_\tau C_\tau &= Q_\tau \\ d_\tau(i_T C) &= i_T Q + h'(0)C_\tau \\ d_\tau K_\tau &= 0.\end{aligned}$$

Notice that there are no more derivatives in the direction of the orbits. Thus these equations can be pulled back to the orbit space N with f without loss of information:

$$\begin{aligned}f(d_\tau(i_T K)) &= f(h'(0)K_\tau) \\ f(d_\tau C_\tau) &= fQ_\tau \\ f(d_\tau(i_T C)) &= f(i_T Q) + f(h'(0)C_\tau) \\ f(d_\tau K_\tau) &= 0.\end{aligned}$$

Because $fd_\tau = d_N f$ holds, we get the reduced differential equations in the orbit space:

$$\begin{aligned}d_N(f(i_T K)) &= f(h'(0)K_\tau) \\ d_N(fC_\tau) &= fQ_\tau \\ d_N(f(i_T C)) &= f(i_T Q) + f(h'(0)C_\tau) \\ d_N(fK_\tau) &= 0.\end{aligned}$$

In the case of G-invariance, $h'(0) = 0$ holds, and we have the simpler equations

$$\begin{aligned} d_N(f(i_T K)) &= 0 \\ d_N(f C_\tau) &= f Q_\tau \\ d_N(f(i_T C)) &= f(i_T Q) \\ d_N(f K_\tau) &= 0. \end{aligned} \tag{5.7}$$

Notice that because the orbit space N is an $(n-1)$-dimensional manifold, and if C or K is an $(n-1)$-form, then the differential equation above for C_τ or K_τ is trivial. In other words, the exterior derivative of an $(n-1)$-form on an $(n-1)$-dimensional manifold is always zero; therefore, any field C_τ or K_τ satisfies the equation (in the case of C, the equation is also homogeneous because fQ_τ is an n-form on an $(n-1)$-manifold and thereby zero). Thus these trivial equations do not bind the solution in any way. However, we will see in the section 5.6.4 that these fields are bound to fields in the nontrivial equations by the constitutive equations.

Example 5.6.1. In the case of magnetostatics, M is a 3-manifold and K corresponds to two-form B (magnetic flux density), C corresponds to one-form H (magnetic field), and Q corresponds to two-form J (current density). If the BVP is G-invariant, the reduced differential equations in the two-dimensional orbit space N are

$$\begin{aligned} d_N(f(i_T B)) &= 0 \\ d_N(f H_\tau) &= f J_\tau \\ d_N(f(i_T H)) &= f(i_T J) \\ d_N(f B_\tau) &= 0. \end{aligned}$$

Notice that the last equation is trivial. Furthermore, the above equations assume only G-invariance from current J but not that it has any special direction. Usually only the other geometric component of J is assumed to be nonzero, and this makes the other equation for H homogeneous and often very easy to solve. For example, if we have a translational symmetry, and if J lies in the direction of translations, i.e., $J = J_\tau$ holds, then we have the following non-trivial equations to solve:

$$\begin{aligned} d_N(f(i_T B)) &= 0 \\ d_N(f H_\tau) &= f J_\tau \\ d_N(f(i_T H)) &= 0. \end{aligned}$$

If the boundary values for $f(i_T H)$ are homogeneous, the last equation has a trivial solution $f(i_T H) = 0$, which, in fact, is often assumed for dimensional reduction.

5.6.3 Boundary values for reduced BVPs

The boundary values for reduced BVPs are derived for two separate cases. First, the case of free action is covered and then the case of effective but not free action. If the action is effective but not free, there are singular orbits which require separate treatment.

Free action

When group action is free, all orbits are one-dimensional. Then the boundary ∂A of the G-reduced domain A is a regular submanifold of ∂M. Thus the boundary ∂N of the orbit space is an embedded submanifold of ∂M (see Figure 5.5).

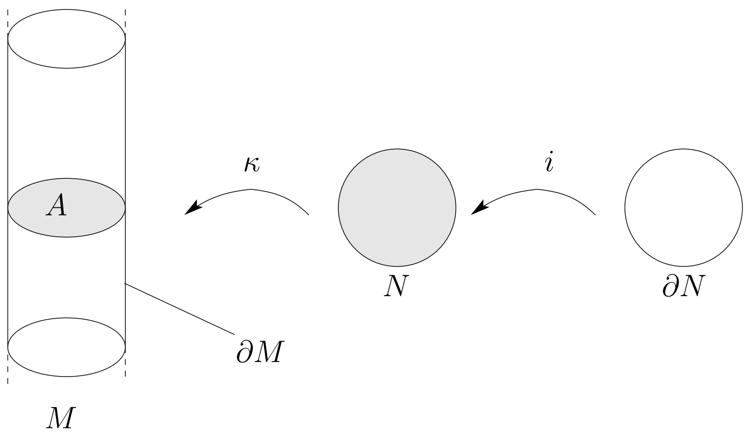

Figure 5.5: Embedding the boundary of the orbit space. M is a 3-manifold, an infinitely long tube with a circular cross-section. A is a G-reduced domain, and the image of the orbit space N under a cross-section κ. i is the inclusion map of ∂N to N. Then $\kappa \circ i$ is the embedding of ∂N to ∂M.

Our task now is to derive boundary values for fields K_τ, $i_T K$, C_τ, and $i_T C$ in the orbit space N from the original boundary values

$$\begin{aligned} t^1_M C &= c \\ t^2_M K &= b. \end{aligned}$$

In the above equation, t^1_M and t^2_M denote the pullbacks of the inclusion maps of the complementary parts of the boundary ∂M. Let $i_{\partial N}$ and $i_{\partial M}$ be the inclusion maps of the boundaries ∂N and ∂M to N and M, respectively, and κ be the cross-section that maps N to A. Then the mapping $\kappa_{\partial N}$ is the induced mapping that maps ∂N to ∂M such that the following commutation

$i_{\partial M} \circ \kappa_{\partial N} = \kappa \circ i_{\partial N}$ holds (see Figure 5.5). Now the fields on ∂M can be pulled back to the boundary ∂N by the pullback of $\kappa_{\partial N}$, which is denoted by $f_{\partial N}$. If the pullbacks of $i_{\partial N}$, $i_{\partial M}$, and κ are denoted by t_N, t_M, and f, respectively, then the following commutation rule holds:

$$f_{\partial N} \circ t_M \quad = \quad t_N \circ f.$$

The vector field T is well-defined over M, and thus also over the subset ∂M. Because $T(p) \in T_p(\partial M)$ holds by the canonical identification of the tangent vectors of $T_p(\partial M)$ and $T_p(M)$ and because we do not want to use unnecessary indices, the restriction of T to the boundary ∂M is not denoted separately but it is understood from the context. Thus the following commutation holds:

$$i_T \circ t_M \quad = \quad t_M \circ i_T.$$

(T at the left-hand side of the above equation is now the restriction of T to the boundary ∂M.)

Now we can express the boundary values in the orbit space: we use the notation t^1_N, t^2_N, $f^1_{\partial N}$, and $f^2_{\partial N}$ for the restrictions and pullbacks defined on the complementary parts of the boundaries ∂N and ∂M. To obtain boundary values for fC_τ and fK_τ, let us first pull back the original boundary values in (5.4) with $f_{\partial N}$ and then use the geometric decompositions and the above commutation rules (notice also that $f\tau = 0$ holds):

$$\begin{aligned} f^1_{\partial N} t^1_M C = t^1_N f C_\tau + t^1_N f(\tau \wedge i_T C) = t^1_N (fC_\tau) = f^1_{\partial N} c \\ f^2_{\partial N} t^2_M K = t^2_N f K_\tau + t^2_N f(\tau \wedge i_T K) = t^2_N (fK_\tau) = f^2_{\partial N} b. \end{aligned}$$

To obtain boundary values for $fi_T C$ and $fi_T K$, let us first contract the original boundary values with respect to T and then pull back the equations with $f_{\partial N}$:

$$\begin{aligned} f^1_{\partial N} i_T t^1_M C = f^1_{\partial N} t^1_M i_T C = t^1_N f(i_T C) = f^1_{\partial N}(i_T c) \\ f^2_{\partial N} i_T t^2_M K = f^2_{\partial N} t^2_M i_T K = t^2_N f(i_T K) = f^2_{\partial N}(i_T b). \end{aligned}$$

In summary, the boundary values for the reduced BVPs are

$$\begin{aligned} t^2_N f K_\tau &= f^2_{\partial N} b \\ t^1_N f C_\tau &= f^1_{\partial N} c \\ t^2_N f(i_T K) &= f^2_{\partial N}(i_T b) \\ t^1_N f(i_T C) &= f^1_{\partial N}(i_T c). \end{aligned}$$

The boundary ∂N of the orbit space is $(n-2)$-dimensional (M is n-dimensional); therefore, all differential forms of degree $(n-1)$ or n are always zero at the boundary. Thus non-homogeneous boundary conditions on ∂M may induce homogeneous conditions on ∂N. This may sound contradictory, but it is not: the non-homogeneous information of the fields is given by the contracted forms, e.g., $i_T b$. These homogeneous equations are trivial in the same sense as some of the above differential equations above. Furthermore, the trivial boundary conditions correspond to fields, which also have trivial differential equations. The way to solve these fields is by substitution for the constitutive equations.

Effective but not free action

When action is effective but not free, there (may) exist singular orbits that contain only one point. For example, in the case of rotational symmetry, where the axis of rotation is part of the domain, the points at the rotational axis are singular orbits whereas the other orbits are circles (see Figure 5.6). To deal with singular orbits, we state the following additional axiom about them:

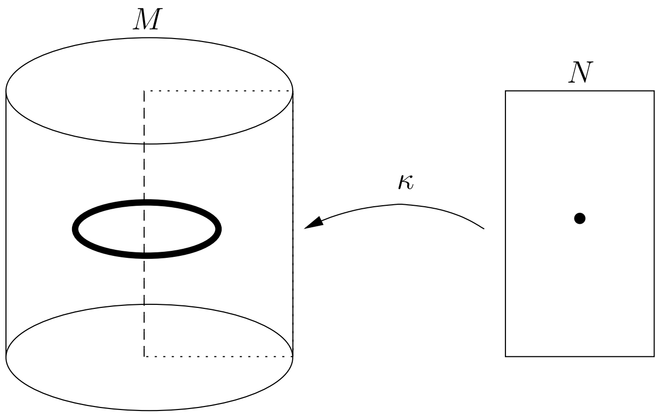

Figure 5.6: Singular orbits. M is a three-dimensional domain that has a rotational symmetry. The thick circle represents a conductor, and the broken line shows the axis of rotation. The points of the axis form singular orbits, which are part of the boundary of the two-dimensional orbit space N. κ is a cross-section that embeds N to M.

Axiom 5.3. Singular orbits always reside at the boundary of the orbit space.

Notice that singular points are not necessarily boundary points of M as the rotational symmetry shows. The axiom restricts the effects of singular orbits to boundary values and is also justified by practical considerations: we are not aware of any practical example where singular orbits are not at the boundary of the orbit space, but we cannot prove that this is always the case for group actions that satisfy Axiom 5.2. However, there is a proposition that provides practical sufficient conditions for singular orbits to reside at the boundary of the orbit space:

Proposition 5.9. Let G be a compact group acting effectively on a connected n-dimensional manifold-with-boundary M. If there is an $(n-2)$-dimensional orbit and none of the dimension $(n-1)$, then G is a Lie group, and the orbit space N of M is a 2-manifold-with-boundary. Furthermore, every point of N corresponding to an orbit of dimension less than $(n-2)$ lies on the boundary curve of N. [10]

As a corollary, we obtain a result that is specific to our situation when the Lie group is one-dimensional:

Corollary 5.2. Let G be a compact one-dimensional Lie group acting effectively on a connected 3-manifold-with-boundary M. If there is a one-dimensional orbit, then the orbit space N is a 2-manifold-with-boundary. Furthermore, every singular orbit lies on the boundary of N.

Open Question 5. Are there group actions that satisfy axioms 5.1 and 5.2 and produce orbit spaces that are manifolds-with-boundary, but where a singular orbit is not at the boundary of the orbit space?

Because of Axiom 5.3, we may assume that the boundary ∂N is a union $\partial N = \partial N_M \cup \partial N_S$, where ∂N_S is the set of all singular orbits whose points are not included in ∂M, and ∂N_M is the part of ∂N that can be embedded in ∂M. The boundary values for ∂N_M are given exactly as in the case of free action; i.e., we need to consider them only at ∂N_S.

At singular points, we had set $\omega_\tau = \omega$ and $i_T\omega = 0$ because the vector field T is always the zero vector at singular points. Thus the contracted components $i_T C$ and $i_T K$ are fixed to zero by the symmetry whereas the horizontal components C_τ and K_τ are not fixed.

To see why symmetry should not fix the fields completely at singular orbits, let us look at an example: the BVP domain is the one shown in Figure 5.6, which has a static current in a circular conductor. We now have a magnetostatic problem. Clearly, the fields H and B are not zero at the points of the rotational axis but nonzero and aligned along the direction of the axis.

Let us now formalize the above result. Let t_{N_S} denote the pullbacks of $\kappa_{\partial N_S} = \kappa \circ i_{\partial N_S}$, where $i_{\partial N_S} : \partial N_S \to N$ is the inclusion map of ∂N_S to N.

Then the boundary values at ∂N_S for the contracted components are trivial:

$$\begin{aligned} t_{N_S} f(i_T C) &= 0 \\ t_{N_S} f(i_T K) &= 0. \end{aligned}$$

This is enough to ensure a unique solution, because the uniqueness of the horizontal components is ensured by the constitutive equations.

Finally, observe that singular sources that are in singular points, are not included in the above theory, because in that case the equations $t_{N_S} i_T C = 0$ and $t_{N_S} i_T K = 0$ may not hold. For example, if there is a charge density ρ along the rotational axis of Figure 5.6, but no charges elsewhere, then $t_{N_S} i_T D$ cannot be zero.

Open Question 6. How to include singular sources in the theory of dimensional reduction using differential forms?

5.6.4 Constitutive equations for reduced BVPs

There are four fields, the geometric components $f(i_T K)$, fK_τ, fC_τ, and $f(i_T C)$, and four differential equations for them. This suggests that there should be two constitutive equations that link geometric components as is the case in Maxwell's equations. However, the constitutive equations can be more complicated than standard ones: even in the case of G-invariance, there may be couplings between three geometric components by a single constitutive equation. This follows from the arbitrary choice of G-reduced domain.

The original constitutive equation is

$$K = \upsilon_M C,$$

where $\upsilon_M : \Omega^k(M) \to \Omega^{n-k}(M)$ is a Hodge-like operator. As pointed out earlier, there is no canonical metric for the orbit space, nor is the invariance dependent on the metric of M. This is why we do not consider metric aspects with constitutive equations and treat υ_M only as a linear operator.

The constitutive equation can be written in terms of geometric decompositions:

$$K_\tau + \tau \wedge i_T K = \upsilon_M(C_\tau) + \upsilon_M(\tau \wedge i_T C). \tag{5.8}$$

From this, we want to induce constitutive equations for fK_τ and $f(i_T K)$ in the orbit space N. That is, we want an equation in which fK_τ is equal to some expression that depends only on fC_τ and $f(i_T C)$ but not on $f(i_T K)$, and then similarly for $f(i_T K)$. The equation for fK_τ follows easily if we notice that $f\tau = 0$. Thus if we pull back the above expression to the orbit space, we have the equation for fK_τ:

$$fK_\tau = f\upsilon_M(C_\tau) + f\upsilon_M(\tau \wedge i_T C). \tag{5.9}$$

The equation for $f(i_T K)$ follows if we notice that $i_T K_\tau = 0$. Thus if we first contract (5.8) with respect to T and then pullback the contracted expression to the orbit space, we have the equation for $f(i_T K)$:

$$f(i_T K) = f i_T \upsilon_M(C_\tau) + f i_T \upsilon_M(\tau \wedge i_T C). \tag{5.10}$$

Equations (5.9) and (5.10) are still in terms of the original n-space operator υ_M, and we want to rewrite them in terms of $(n-1)$-space operators that directly link fK_τ to fC_τ and $f(i_T C)$ in the following way:

$$fK_\tau = \upsilon_\tau^\tau(fC_\tau) + \upsilon_T^\tau(f(i_T C)). \tag{5.11}$$

Similarly for $f(i_T K)$

$$f(i_T K) = \upsilon_\tau^T(fC_\tau) + \upsilon_T^T(f(i_T C)). \tag{5.12}$$

We must determine the four linear operators

$$\begin{aligned} \upsilon_\tau^\tau &: \Omega^k(N) \to \Omega^{n-k}(N) \\ \upsilon_T^\tau &: \Omega^{k-1}(N) \to \Omega^{n-k}(N) \\ \upsilon_\tau^T &: \Omega^k(N) \to \Omega^{n-k-1}(N) \\ \upsilon_T^T &: \Omega^{k-1}(N) \to \Omega^{n-k-1}(N) \end{aligned} \tag{5.13}$$

in terms of υ_M. Notice that because N is $(n-1)$-dimensional, only υ_T^τ and υ_τ^T are Hodge-like operators in the sense that they map between spaces of equal dimensions. Furthermore, because at each point, υ_M maps k-forms on an n-dimensional space M to $(n-k)$-forms on M, it is a mapping from $\binom{n}{k}$-dimensional space to $\binom{n}{n-k}$-dimensional space. Thus it requires $\binom{n}{k} \cdot \binom{n}{k}$ parameters to describe υ_M at each point. On the other hand, the number of parameters required to pointwise describe the reduced operators υ_τ^τ, υ_T^τ, υ_τ^T, and υ_T^T are $\binom{n-1}{k} \cdot \binom{n-1}{n-k}$, $\binom{n-1}{k-1} \cdot \binom{n-1}{n-k}$, $\binom{n-1}{k} \cdot \binom{n-1}{n-1-k}$, and $\binom{n-1}{k-1} \cdot \binom{n-1}{n-1-k}$, respectively. With calculations, we can show that these add up to $\binom{n}{k} \cdot \binom{n}{k}$ parameters. This suggests that υ_M can be "decomposed" to reduced operators to form "blocks" of υ_M.

Let us first define $\upsilon_\tau^\tau : \Omega^k(N) \to \Omega^{n-k}(N)$ by the following equation:

$$\upsilon_\tau^\tau(fC_\tau) = f\upsilon_M(C_\tau) \tag{5.14}$$

In the above equation, f on the left-hand side is the pullback of $(n-k)$-forms from M to N whereas f on the right is the pullback of k-forms from M to N. In the following, the pullback of $(n-k)$-forms is denoted by f_{n-k}; similarly, the pullback of k-forms is denoted by f_k. Now equation (5.14) can be written as

$$(f_{n-k} \circ \upsilon_M)C_\tau \quad = \quad (\upsilon_\tau^\tau \circ f_k)C_\tau, \tag{5.15}$$

where we want to emphasize that both sides of the equation have a composition of linear operators that operate on C_τ. Because we want the above equation to hold for any field C_τ, we have

$$f_{n-k} \circ \upsilon_M \quad = \quad \upsilon_\tau^\tau \circ f_k. \tag{5.16}$$

The rank of the pullback f_k at each point equals to the dimension of N and thus it has a right-inverse denoted by r_k (right-inverse means that the equation $f_k \circ r_k = id_N$ holds). Then we can express υ_τ^τ in terms of υ_M and the pullbacks:

$$\upsilon_\tau^\tau \quad = \quad f_{n-k} \circ \upsilon_M \circ r_k. \tag{5.17}$$

The next case is $\upsilon_T^\tau : \Omega^{k-1}(N) \to \Omega^{n-k}(N)$. By definition, it satisfies the following equation:

$$\upsilon_T^\tau(f(i_T C)) \quad = \quad f\upsilon_M(\tau \wedge i_T C). \tag{5.18}$$

Our goal is again to write this to emphasize the linear operators operating on $i_T C$. Using the extension operator I_τ, equation (5.18) can be written as

$$(f_{n-k} \circ \upsilon_M \circ I_\tau)i_T C \quad = \quad (\upsilon_T^\tau \circ f_{k-1})i_T C. \tag{5.19}$$

Because we want equation (5.19) to hold for any field $i_T C$, we have

$$f_{n-k} \circ \upsilon_M \circ I_\tau \quad = \quad \upsilon_T^\tau \circ f_{k-1}. \tag{5.20}$$

The pullback f_{k-1} has a well-defined right-inverse denoted by r_{k-1}. We can now express υ_T^τ in terms of υ_M, I_τ and the pullbacks:

$$\upsilon_T^\tau \quad = \quad f_{n-k} \circ \upsilon_M \circ I_\tau \circ r_{k-1}. \tag{5.21}$$

The other two operators $\upsilon_\tau^T : \Omega^k(N) \to \Omega^{n-k-1}(N)$ and $\upsilon_T^T : \Omega^{k-1}(N) \to \Omega^{n-k-1}(N)$ are derived in a similar fashion:

$$\begin{aligned} \upsilon_\tau^T \quad &= \quad f_{n-k-1} \circ i_T \circ \upsilon_M \circ r_k \\ \upsilon_T^T \quad &= \quad f_{n-k-1} \circ i_T \circ \upsilon_M \circ I_\tau \circ r_{k-1}. \end{aligned} \tag{5.22}$$

Remark 5.6.1. The reduced constitutive equations could have been derived by using the projections as in section 4.3.1. The projections P_τ and P_T decompose υ_M into four reduced operators in $\Omega(M)$. Then applications of the pullbacks f_{n-k}, f_{n-k-1}, r_{k-1}, and r_k to these reduced operators yield the operators υ_τ^τ, υ_T^τ, υ_τ^T, and υ_T^T.

Let us next look at these equations in terms of matrices. Reduced constitutive equations can be represented as a single equation in formal matrix notation as follows:

$$\begin{bmatrix} f(i_T K) \\ f K_\tau \end{bmatrix} = \begin{bmatrix} \upsilon_\tau^T & \upsilon_T^T \\ \upsilon_\tau^\tau & \upsilon_T^\tau \end{bmatrix} \begin{bmatrix} f C_\tau \\ f(i_T C) \end{bmatrix}.$$

On the other hand, assume that C and K are one- and two-forms, respectively, on a 3-manifold M. Let $\{X, Y, Z\}$ be a geometric basis field such that $\{X, Y\}$ is the basis field of a chosen G-reduced domain A, and Z is in the direction of orbits. Then the dual basis $\{dx, dy, dz\}$ is the geometric basis for one-forms, and $C = C_x dx + C_y dy + C_z dz$ holds for some zero-forms C_x, C_y, and C_z. Let $\{dy \wedge dz, dz \wedge dx, dx \wedge dy\}$ be the geometric basis for two-forms, in which case $K = K_x dy \wedge dz + K_y dz \wedge dx + K_z dx \wedge dy$ holds. Applying the pullback f to the geometric bases, we get the corresponding bases for the orbit space N (in these bases the components are the same in M and N). The pair (Z, dz) defines now the projections, and the geometric components in terms of the geometric bases are

$$\begin{aligned} C_\tau &= C_x dx + C_y dy \\ i_Z C &= C_z \\ K_\tau &= K_z dx \wedge dy \\ i_Z K &= K_y dx - K_x dy. \end{aligned}$$

In the geometric bases, the operator υ_M is given by the following matrix:

$$\upsilon_M = \begin{bmatrix} \upsilon_{xx} & \upsilon_{xy} & \upsilon_{xz} \\ \upsilon_{yx} & \upsilon_{yy} & \upsilon_{yz} \\ \upsilon_{zx} & \upsilon_{zy} & \upsilon_{zz} \end{bmatrix}.$$

It is now easy to deduce the matrices of linear operators υ_τ^τ, υ_Z^τ, υ_τ^Z, and υ_Z^Z in the geometric bases:

$$\upsilon_\tau^\tau = \begin{bmatrix} \upsilon_{zx} & \upsilon_{zy} \end{bmatrix}$$

$$\upsilon_Z^\tau = \begin{bmatrix} \upsilon_{zz} \end{bmatrix}$$

$$\upsilon_\tau^Z = \begin{bmatrix} \upsilon_{xx} & \upsilon_{xy} \\ \upsilon_{yx} & \upsilon_{yy} \end{bmatrix}$$

$$\upsilon_Z^Z = \begin{bmatrix} \upsilon_{xz} \\ \upsilon_{yz} \end{bmatrix}.$$

Thus the matrices of linear operators υ_τ^τ, υ_Z^τ, υ_τ^Z, and υ_Z^Z are blocks of the matrix υ_M.

5.6.5 Summary

We have studied the following type of generic BVP that is posed on an n-dimensional manifold-with-boundary M:

$$\begin{aligned} d_M C &= Q \\ d_M K &= 0 \\ t_M^1 C &= c \\ t_M^2 K &= b \\ K &= \upsilon_M C. \end{aligned}$$

where Q is a $(k+1)$-form describing the sources, C is a k-form, K is an $(n-k)$-form, and υ is a Hodge-like operator.

The BVP is (G, h)-invariant under the effective action of a connected one-dimensional Lie group G. The orbit space is a connected $(n-1)$-dimensional manifold-with-boundary N. The invariance induces a decomposition of the fields when we choose some G-reduced domain A. Then there is a pair (T, τ), where T is a nonzero smooth G-invariant vector field, which is everywhere tangent to the orbits except at the singular points where T is zero, and τ is a smooth G-invariant one-form such that $\tau(v) = 0$ holds for all $v \in T(A)$, and $\tau(T) = 1$ holds everywhere except at the singular points where τ is not defined. The pair (T, τ) defines two complementary projections, which decompose the fields C, K, and Q as follows:

$$\begin{aligned} K &= K_\tau + \tau \wedge i_T K \\ C &= C_\tau + \tau \wedge i_T C \\ Q &= Q_\tau + \tau \wedge i_T Q. \end{aligned}$$

These are the geometric decompositions of the fields induced by the symmetry and the choice of the pair (T, τ). The components C_τ and $i_T C$ are called the geometric components of C, and they are the fields to be solved from the reduced BVPs on the orbit space. To get the geometric components in the orbit space, we use the pullback f of the cross-section that defines the G-reduced domain A. The pullback $f_{\partial N}$, which maps the boundary values of ∂M to ∂N, is used to define the boundary values for the geometric components in the orbit space.

The (G,h)-invariance of the BVP implies the following lower-dimensional BVP on the orbit space:

$$\begin{aligned}
d_N(f(i_T K)) &= f(h'(0)K_\tau) \\
d_N(fC_\tau) &= fQ_\tau \\
d_N(f(i_T C)) &= f(i_T Q) + f(h'(0)C_\tau) \\
d_N(fK_\tau) &= 0 \\
t^2_N fK_\tau &= f^2_{\partial N} b \\
t^1_N fC_\tau &= f^1_{\partial N} c \\
t^2_N f(i_T K) &= f^2_{\partial N}(i_T b) \\
t^1_N f(i_T C) &= f^1_{\partial N}(i_T c) \\
fK_\tau &= \upsilon^\tau_\tau(fC_\tau) + \upsilon^\tau_T(f(i_T C)) \\
f(i_T K) &= \upsilon^T_\tau(fC_\tau) + \upsilon^T_T(f(i_T C)).
\end{aligned}$$

If there are singular orbits at the boundary of the orbit space, then $\partial N = \partial N_M \cup \partial N_S$ holds, and there are two additional boundary conditions imposed at ∂N_S:

$$\begin{aligned}
t_{N_S} f(i_T C) &= 0 \\
t_{N_S} f(i_T K) &= 0.
\end{aligned}$$

The linear operators

$$\begin{aligned}
\upsilon^\tau_\tau &: \Omega^k(N) \to \Omega^{n-k}(N) \\
\upsilon^\tau_T &: \Omega^{k-1}(N) \to \Omega^{n-k}(N) \\
\upsilon^T_\tau &: \Omega^k(N) \to \Omega^{n-k-1}(N) \\
\upsilon^T_T &: \Omega^{k-1}(N) \to \Omega^{n-k-1}(N)
\end{aligned}$$

in the constitutive equations are expressed as follows:

$$\begin{aligned}
\upsilon^\tau_\tau &= f_{n-k} \circ \upsilon_M \circ r_k \\
\upsilon^\tau_T &= f_{n-k} \circ \upsilon_M \circ I_\tau \circ r_{k-1} \\
\upsilon^T_\tau &= f_{n-k-1} \circ i_T \circ \upsilon_M \circ r_k \\
\upsilon^T_T &= f_{n-k-1} \circ i_T \circ \upsilon_M \circ I_\tau \circ r_{k-1},
\end{aligned}$$

where r_k is the right-inverse of the pullback f_k of the k-forms.

If the BVP is G-invariant and if operators υ^τ_τ and υ^T_T are zero, the lower-dimensional BVP breaks up into two separate lower-dimensional BVPs:

$$\begin{aligned}
d_N(f(i_T K)) &= 0 \\
d_N(f C_\tau) &= f Q_\tau \\
t_N^1 f C_\tau &= f_{\partial N}^1 c \\
t_N^2 f(i_T K) &= f_{\partial N}^2 (i_T b) \\
f(i_T K) &= \upsilon_\tau^T (f C_\tau)
\end{aligned}$$

$$\begin{aligned}
d_N(f(i_T C)) &= f(i_T Q) \\
d_N(f K_\tau) &= 0 \\
t_N^1 f(i_T C) &= f_{\partial N}^1 (i_T c) \\
t_N^2 f K_\tau &= f_{\partial N}^2 b \\
f K_\tau &= \upsilon_T^\tau (f(i_T C))
\end{aligned}$$

Furthermore, the other BVP is often quite simple or even trivial (see Example 5.6.1). In fact, because the other problem is often so trivial, its result is just assumed: with suitable coordinate system, the trivial problem effectively says that some component of the solution field is zero.

Open Question 7. Because choice of a G-reduced domain is arbitrary, reduced constitutive equations, in general, connect three of geometric components. Now, with a suitable choice of a G-reduced domain, is it always possible to separate the lower-dimensional BVP of a G-invariant BVP into two separate lower-dimensional BVPs? If the separation with a suitable choice is possible, then how to determine suitable choices?

5.7 Static and time-harmonic electromagnetics

As basic examples of dimensional reduction, we derive here the static and time-harmonic equations, which are consequences of static and time-harmonic invariances. In other words, the static equations are the reduced differential equations of full 4d Maxwell's equations when the fields are invariant with respect to time. Similarly, the time-harmonic equations are reduced differential equations in case of time-harmonic invariance. Furthermore, the reduced constitutive equations are derived.

5.7.1 Orbit space and geometric decompositions

Let M be a four-dimensional spacetime manifold, and let $\mathbb{R}$ act on M such that all the orbits are one-dimensional and in the direction of time. Particularly, let us choose a holonomic observer (T, τ) such that T and τ are both $\mathbb{R}$-invariant, τ defines a foliation of M such that each leaf of the foliation contains all the the "spatial points" at some given "time instant." Let an $\mathbb{R}$-reduced domain A be one of the leaves of the foliations. With these assumptions, the orbit space N is a 3-manifold.

Next we apply this observer to full 4d Maxwell's equations given in section 4.2.1. Thus the fields are $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{J}$, or the electromagnetic field two-form, the excitation two-form, and the source three-form, respectively, and they are governed by Maxwell's equations:

$$\begin{aligned} d_M\mathcal{F} &= 0 \\ d_M\mathcal{G} &= \mathcal{J}. \end{aligned} \tag{5.23}$$

Furthermore, $\mathcal{F}$ and $\mathcal{G}$ are connected by the constitutive equation:

$$\mathcal{G} = \chi\mathcal{F}, \tag{5.24}$$

where χ is a linear isomorphism from $\Omega^2(M)$ to $\Omega^2(M)$ satisfying the axioms of symmetry and closure, see section 4.3.1.

The geometric decomposition of the fields are exactly the same as those in section 4.2.2: $\mathcal{F} = \mathcal{F}_\tau + \tau \wedge i_T\mathcal{F}$, $\mathcal{G} = \mathcal{G}_\tau + \tau \wedge i_T\mathcal{G}$, and $\mathcal{J} = \mathcal{J}_\tau + \tau \wedge i_T\mathcal{J}$. Then we *rename* the geometric components as follows:

$$\begin{aligned} B &= -\mathcal{F}_\tau \\ E &= i_T\mathcal{F} \\ D &= \mathcal{G}_\tau \\ H &= i_T\mathcal{G} \\ \rho &= \mathcal{J}_\tau \\ J &= -i_T\mathcal{J}. \end{aligned} \tag{5.25}$$

Consequently, the geometric decompositions are $\mathcal{F} = -B + \tau \wedge E$, $\mathcal{G} = D + \tau \wedge H$, and $\mathcal{J} = \rho - \tau \wedge J$.

5.7.2 Static and time-harmonic differential equations

As shown in section 4.2.2, the observer (T, τ) decomposes the equations in (5.23) as follows:

$$\begin{aligned} d_\tau \mathcal{F}_\tau &= 0 \\ \mathcal{L}_T \mathcal{F}_\tau - d_\tau\, i_T \mathcal{F} &= 0 \\ d_\tau \mathcal{G}_\tau &= \mathcal{J}_\tau \\ \mathcal{L}_T \mathcal{G}_\tau - d_\tau\, i_T \mathcal{G} &= i_T \mathcal{J} \end{aligned} \tag{5.26}$$

Notice that these equations contain spatial and time derivatives, but no invariance assumptions have yet been made about the fields. Static invariance of the fields $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{J}$ implies that their Lie derivatives with respect to T are zero: $\mathcal{L}_T \mathcal{F} = \mathcal{L}_T \mathcal{G} = 0$ and $\mathcal{L}_T \mathcal{J} = 0$. Then applying the geometric decompositions to these equations yields (see section 5.6.2)

$$\begin{aligned} \mathcal{L}_T \mathcal{F}_\tau = \mathcal{L}_T \mathcal{G}_\tau = \mathcal{L}_T i_T \mathcal{J} &= 0 \\ \mathcal{L}_T i_T \mathcal{F} = \mathcal{L}_T i_T \mathcal{G} &= 0 \\ \mathcal{L}_T \mathcal{J}_\tau &= 0. \end{aligned}$$

Next we substitute these equations for those in (5.26) and obtain the reduced equations for M:

$$\begin{aligned} d_\tau \mathcal{F}_\tau &= 0 \\ d_\tau\, i_T \mathcal{F} &= 0 \\ d_\tau \mathcal{G}_\tau &= \mathcal{J}_\tau \\ d_\tau\, i_T \mathcal{G} &= -i_T \mathcal{J}. \end{aligned}$$

Using renaming gives us the usual static equations, but notice that the domain yet remains a 4-dimensional M:

$$\begin{aligned} d_\tau B &= 0 \\ d_\tau E &= 0 \\ d_\tau D &= \rho \\ d_\tau H &= J. \end{aligned}$$

To transfer these equations to the 3-dimensional orbit space N, we need to pull them back using the cross-section that defines the $\mathbb{R}$-reduced domain A. Let us denote this pullback by f. Then because the pullback and spatial exterior derivative commute, we have the following differential equations in

N:

$$\begin{aligned} d_N(fB) &= 0 \\ d_N(fE) &= 0 \\ d_N(fD) &= f\rho \\ d_N(fH) &= fJ. \end{aligned}$$

Notice that these equations are particular instances of the generic equations shown in (5.7).

The time-harmonic case is similar to the static case, but now we have $(\mathbb{R}, h)$-invariance for fields $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{J}$, where $h : \mathbb{R} \to \mathbb{C}$ such that $h(t) = e^{j\omega t}$ holds for all $t \in \mathbb{R}$, and where ω is the frequency. Now $h'(0) = j\omega$ holds. Substituting this and the renamings for the equations in (5.26), we get the familiar time-harmonic Maxwell's equations in N:

$$\begin{aligned} d_N(fE) &= -j\omega(fB) \\ d_N(fD) &= f\rho \\ d_N(fH) &= fJ + j\omega(fD) \\ d_N(fB) &= 0. \end{aligned}$$

5.7.3 Constitutive equations

In section 5.6.4 it was shown that the reduced constitutive equations in the orbit space are

$$\begin{aligned} f\mathcal{G}_\tau &= \chi_\tau^\tau(f\mathcal{F}_\tau) + \chi_T^\tau(f(i_T\mathcal{F})) \\ f(i_T\mathcal{G}) &= \chi_\tau^T(f\mathcal{F}_\tau) + \chi_T^T(f(i_T\mathcal{F})). \end{aligned}$$

The linear operators

$$\begin{aligned} \chi_\tau^\tau &: \Omega^2(N) \to \Omega^2(N) \\ \chi_T^\tau &: \Omega^1(N) \to \Omega^2(N) \\ \chi_\tau^T &: \Omega^2(N) \to \Omega^1(N) \\ \chi_T^T &: \Omega^1(N) \to \Omega^1(N) \end{aligned}$$

in the constitutive equations are expressed as follows:

$$\begin{aligned} \chi_\tau^\tau &= f_{n-k} \circ \chi_M \circ r_k \\ \chi_T^\tau &= f_{n-k} \circ \chi_M \circ I_\tau \circ r_{k-1} \\ \chi_\tau^T &= f_{n-k-1} \circ i_T \circ \chi_M \circ r_k \\ \chi_T^T &= f_{n-k-1} \circ i_T \circ \chi_M \circ I_\tau \circ r_{k-1}, \end{aligned}$$

where r_k is the right-inverse of the pullback f_k of the k-forms. In terms of renamings, the reduced constitutive equations are

$$\begin{aligned} fD &= -\chi^{\tau}_{\tau}(fB) + \chi^{\tau}_{T}(fE) \\ fH &= -\chi^{T}_{\tau}(fB) + \chi^{T}_{T}(fE). \end{aligned}$$

We recognized in section 4.3.1 that χ^{τ}_{T} equals to ϵ (permittivity) and $-\chi^{T}_{\tau}$ equals to ν (inverse permeability). If a suitable holonomic observer exists such that χ^{τ}_{τ} and χ^{T}_{T} are zero, the reduced constitutive equations are the following familiar equations:

$$\begin{aligned} fD &= \epsilon(fE) \\ fH &= \nu(fB). \end{aligned}$$

5.8 Lower-dimensional BVPs for multi-dimensional symmetry groups

In this section, we derive lower-dimensional BVPs for a symmetry group G, which is the product $G_1 \times ... \times G_k$ of connected one-dimensional Lie groups. Remember that all two-dimensional connected Lie groups are products of one-dimensional connected Lie groups (see section 3.6). For simplicity, we study only the cases where G is the product of two groups, or where $G = G_1 \times G_2$ holds. These cases include, e.g., static one-dimensional problems such as those with a cylindrical symmetry. In addition, all two-dimensional electromagnetic BVPs are results of dimensional reduction by two-dimensional symmetry groups.

Because G is a product of one-dimensional Lie groups, we can use the results in section 5.6: both groups G_1 and G_2 have their own group actions on M, which makes it possible to construct observer structures for both of the actions. We can then produce geometric decompositions of the fields such that we first apply one observers and then the other to the geometric components of the first observer.

Similarly, we derive the reduced differential equations: the action of G_1 on M produces its own orbit space M/G_1, observer structures, and geometric decompositions, and thus its own reduced BVPs on M/G_1. The group action of G_2 on M induces a group action of G_2 on M/G_1. This group action on M/G_1 produces the orbit space $(M/G_1)/G_2$, which is canonically diffeomorphic to the orbit space M/G. Thus the action of G_2 on M/G_1 canonically produces reduced BVPs on M/G. If the group G has more factors, we continue applying the group action of each factor to the recently reduced BVPs until all factors have been applied.

Boundary values and constitutive equations for the geometric components on the orbit space M/G can be derived directly in a fashion similar to the done with one-dimensional symmetry groups.

We study again the same generic BVP shown in (5.4); i.e., the domain M is an n-dimensional manifold-with-boundary, and the following equations hold in M:

$$\begin{aligned} d_M C &= Q \\ d_M K &= 0 \\ t^1_M C &= c \\ t^2_M K &= b \\ K &= \upsilon_M C, \end{aligned} \tag{5.27}$$

where the source field is a $(k+1)$-form Q, and the other two fields are a k-form C and an $(n-k)$-form K. The constitutive equation between K and C is given by a Hodge-like operator υ_M. Our task is to derive a reduced BVP for the $(n-2)$-dimensional orbit space N.

5.8.1 Group action and geometric decompositions

We assume that $G = G_1 \times G_2$ holds, and that G acts as described in Axiom 5.2. Furthermore, we assume that the orbit space $N = M/G$ exists. Thus we have a group action $F : G_1 \times G_2 \times M \to M$, which in this case means that $F(a_1 \circ b_1, a_2 \circ b_2, p) = F(a_1, a_2, F(b_1, b_2, p))$ holds for all $a_1, b_1 \in G_1$ and $a_2, b_2 \in G_2$. Both groups G_1 and G_2 also have their own separate group actions $F_i : G_i \times M \to M$, defined by $F_1(g,p) = F(g, e_2, p)$ and $F_2(g,p) = F(e_1, g, p)$, where e_1 and e_2 are the identity elements of G_1 and G_2, respectively. Both actions F_1 and F_2 must satisfy Axiom 5.2.

The group action F_1 produces the orbit space $N_1 = M/G_1$. Then the group action F_2 on M induces a group action $F^1_2 : G_2 \times N_1 \to N_1$ of G_2 on N_1 such that $F^1_2(g, G_1 p) = G_1 F_2(g,p)$ holds, where $G_1 p$ and $G_1 F_2(g,p)$ are the orbits of p and $F_2(g,p)$, respectively, under the action F_1. We denote by $G_2(G_1 p)$ the orbit of $G_1 p \in N_1$ under the action F^1_2. Then the orbits of F and F^1_2 are canonically identified (bijectively) by the mapping $G_2(G_1 p) \mapsto Gp$. With this mapping, the reduced BVP on $N^1_2 = (M/G_1)/G_2$ can be canonically defined in the orbit space N.

Next, we define the observers related to actions F_1, F_2, and F^1_2. First, we choose a G-reduced domain A, which is an $(n-2)$-dimensional submanifold-with-boundary. Then actions F_2 and F_1 induce G_1- and G_2-reduced domains A_1 and A_2, respectively, such that $A_1 = F_2(G_2, A)$ and $A_2 = F_1(G_1, A)$ hold (see Figure 5.7). Notice that A_1 and A_2 are $(n-1)$-dimensional manifolds.

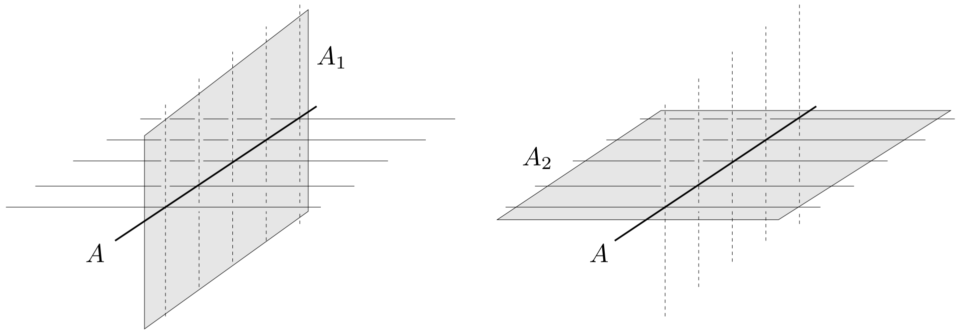

Figure 5.7: Action of a two-dimensional Lie group. The thick line denotes a G-reduced domain A, the vertical lines are orbits of G_2, and the horizontal lines are orbits of G_1. $A_1 = F_2(G_2, A)$ is a G_1-reduced domain, and $A_2 = F_1(G_1, A)$ is a G_2-reduced domain.

Let T and Z be smooth G-invariant vector fields on M everywhere along the orbits of action F_1 and F_2, respectively. Notice that $T_p \in T(A_2)$ and $Z_q \in T(A_1)$ hold for $p \in A_2$ and $q \in A_1$; i.e., T and Z are tangent to A_2 and A_1, respectively. Then we define smooth G-invariant one-forms τ and ζ on M such that $\tau(T) = \zeta(Z) = 1$, $\tau(Z) = \zeta(T) = 0$, and $\tau(V) = \zeta(V) = 0$ hold for all $V \in TA$. Thus (T, τ) is an observer on M related to the action F_1, and (Z, ζ) is an observer on M related to the action F_2. Finally, corresponding to the induced action F_2^1, we define an observer on the orbit space N_1: let $\pi_1 : M \to N_1$ be the canonical projection of F_1 and κ_1 a cross-section such that $\kappa_1(N_1) = A_1$. Then $(Z_1, \zeta_1) = (\pi_{1*}Z, \kappa_1^*\zeta)$ is an observer for the induced action F_2^1.

Now with these observer structures we define the geometric decompositions of the fields. First, we apply (T, τ) to C to produce the following decomposition:

$$C = C_\tau + \tau \wedge i_T C$$

Then we apply the pair (Z, ζ) to the geometric components yielded by (T, τ):

$$\begin{aligned} C &= C_\tau + \tau \wedge i_T C \\ &= (C_\tau)_\zeta + \zeta \wedge i_Z C_\tau + \tau \wedge ((i_T C)_\zeta + \zeta \wedge (i_Z i_T C)) \\ &= (C_\tau)_\zeta + \zeta \wedge i_Z C_\tau + \tau \wedge (i_T C)_\zeta + \tau \wedge \zeta \wedge (i_Z i_T C). \end{aligned}$$

The geometric components of C are $(C_\tau)_\zeta$, $i_Z C_\tau$, $(i_T C)_\zeta$, and $i_Z i_T C$. These and the geometric components of K are the fields to be solved in the orbit space N. To simplify the notation, we use the following notation for the geometric components:

$$\begin{aligned} C_{\tau\zeta} &= (C_\tau)_\zeta \\ C_{\tau Z} &= i_Z C_\tau \\ C_{T\zeta} &= (i_T C)_\zeta \\ C_{TZ} &= i_Z i_T C \end{aligned}$$

Consequently, the geometric decompositions of the fields are

$$\begin{aligned} C &= C_{\tau\zeta} + \zeta \wedge C_{\tau Z} + \tau \wedge C_{T\zeta} + \tau \wedge \zeta \wedge C_{TZ} \\ K &= K_{\tau\zeta} + \zeta \wedge K_{\tau Z} + \tau \wedge K_{T\zeta} + \tau \wedge \zeta \wedge K_{TZ} \\ Q &= Q_{\tau\zeta} + \zeta \wedge Q_{\tau Z} + \tau \wedge Q_{T\zeta} + \tau \wedge \zeta \wedge Q_{TZ}. \end{aligned}$$

Usually, some geometric components become trivial. For example, in a three-dimensional manifold M, the geometric decompositions for zero-, one-, two-, and three-forms, respectively, are

$$\begin{aligned} \phi &= \phi_{\tau\zeta} \\ E &= E_{\tau\zeta} + E_{\tau Z}\,\zeta + E_{T\zeta}\,\tau \\ D &= \zeta \wedge D_{\tau Z} + \tau \wedge D_{T\zeta} + D_{TZ}\,\tau \wedge \zeta \\ \rho &= \tau \wedge \zeta \wedge \rho_{TZ}. \end{aligned}$$

5.8.2 Reduced differential equations

The reduced differential equations of the reduced BVPs in the orbit space N are derived in two steps. First, (T, τ) induces a reduced BVP on N_1, and then this reduced BVP is further reduced by (Z_1, ζ_1). We now have a reduced BVP on N_2^1, which can be canonically identified with N. Notice that this recursive derivation enables exploitation of the results in section 5.6.2.

First reduction

(G, h)-invariance of the fields means that $(g_1, g_2)^*\omega = g^*\omega = h(g)\omega = h((g_1, g_2))\omega$ holds for all $g_1 \in G_1$, $g_2 \in G_2$. If we define $h_1(g) = h(g, e_2)$ and $h_2(g) = h(e_1, g)$, then $h((g_1, g_2)) = h_1(g_1)h_2(g_2)$ holds. Moreover, (G, h)-invariance of the fields implies that they are also (G_1, h_1)- and (G_2, h_2)-invariant. (G_1, h_1)-invariance of fields is equivalent to the following Lie derivative equations (see section 5.6.2):

$$\begin{aligned} \mathcal{L}_T C_\tau &= h_1'(0) C_\tau \\ \mathcal{L}_T K_\tau &= h_1'(0) K_\tau. \end{aligned}$$

From section 5.6.2, we know that (T, τ) and the invariance of the fields (the above Lie derivative equations) yield the following equations in the orbit space N_1:

$$\begin{aligned} d_{N_1}(f_1(i_T K)) &= h_1'(0) f_1 K_\tau \\ d_{N_1}(f_1 C_\tau) &= f_1 Q_\tau \\ d_{N_1}(f_1(i_T C)) &= f_1(i_T Q) + h_1'(0) f_1 C_\tau \\ d_{N_1}(f_1 K_\tau) &= 0, \end{aligned} \tag{5.28}$$

where f_1 is the pullback of the cross-section κ_1 defining the G_1-reduced domain A_1.

Second reduction

The starting point of the second reduction is the reduced BVP in (5.28). To simplify the notation, we use the following renamings:

$$\begin{aligned} \Lambda &= f_1(i_T K) \\ \Psi &= f_1 K_\tau \\ \Sigma &= f_1(i_T C) \\ \Gamma &= f_1 C_\tau \\ \Delta &= f_1(i_T Q) \\ \Theta &= f_1 Q_\tau. \end{aligned} \tag{5.29}$$

Then the BVP in (5.28) reads as

$$\begin{aligned} d_{N_1}\Lambda &= h_1'(0)\Psi \\ d_{N_1}\Gamma &= \Theta \\ d_{N_1}\Sigma &= \Delta + h_1'(0)\Gamma \\ d_{N_1}\Psi &= 0. \end{aligned} \tag{5.30}$$

The F_2-induced observer (Z_1, ζ_1) on the orbit space N_1 decomposes the fields Λ, Ψ, Σ, Γ, Δ, and Θ. (Z_1, ζ_1) also decomposes the equations in (5.30) as follows:

$$\begin{aligned}
d_{\zeta_1}\Lambda_{\zeta_1} &= h_1'(0)\Psi_{\zeta_1} \\
d_{\zeta_1}(i_{Z_1}\Lambda) &= h_1'(0)i_{Z_1}\Psi + \mathcal{L}_{Z_1}\Lambda_{\zeta_1} \\
d_{\zeta_1}\Gamma_{\zeta_1} &= \Theta_{\zeta_1} \\
d_{\zeta_1}(i_{Z_1}\Gamma) &= i_{Z_1}\Theta + \mathcal{L}_{Z_1}\Gamma_{\zeta_1} \\
d_{\zeta_1}\Sigma_{\zeta_1} &= \Delta_{\zeta_1} + h_1'(0)\Gamma_{\zeta_1} \\
d_{\zeta_1}(i_{Z_1}\Sigma) &= i_{Z_1}\Delta + h_1'(0)i_{Z_1}\Gamma + \mathcal{L}_{Z_1}\Sigma_{\zeta_1} \\
d_{\zeta_1}\Psi_{\zeta_1} &= 0 \\
d_{\zeta_1}(i_{Z_1}\Psi) &= \mathcal{L}_{Z_1}\Psi_{\zeta_1}.
\end{aligned}$$

Now the (G_2, h_2)-invariance of the horizontal components Λ_{ζ_1}, Ψ_{ζ_1}, Σ_{ζ_1}, and Γ_{ζ_1} is equivalent to the following Lie derivative equations:

$$\begin{aligned}
\mathcal{L}_{Z_1}\Lambda_{\zeta_1} &= h_2'(0)\Lambda_{\zeta_1} \\
\mathcal{L}_{Z_1}\Sigma_{\zeta_1} &= h_2'(0)\Sigma_{\zeta_1} \\
\mathcal{L}_{Z_1}\Psi_{\zeta_1} &= h_2'(0)\Psi_{\zeta_1} \\
\mathcal{L}_{Z_1}\Gamma_{\zeta_1} &= h_2'(0)\Gamma_{\zeta_1}.
\end{aligned}$$

Let f_2 be the pullback of a cross-section from the orbit space N_2^1 to some G_2-reduced domain R of N_1 corresponding to ζ_1 (by the definition, $\zeta_1(v) = 0$ holds for all $v \in T(R)$). To obtain the final reduced differential equations on N_2^1, we substitute the Lie derivative equations and use the pullback f_2:

$$\begin{aligned}
d_{N_2^1} f_2\Lambda_{\zeta_1} &= h_1'(0)f_2\Psi_{\zeta_1} \\
d_{N_2^1} f_2(i_{Z_1}\Lambda) &= h_1'(0)f_2(i_{Z_1}\Psi) + h_2'(0)f_2\Lambda_{\zeta_1} \\
d_{N_2^1} f_2\Gamma_{\zeta_1} &= f_2\Theta_{\zeta_1} \\
d_{N_2^1} f_2(i_{Z_1}\Gamma) &= f_2(i_{Z_1}\Theta) + h_2'(0)f_2\Gamma_{\zeta_1} \\
d_{N_2^1} f_2\Sigma_{\zeta_1} &= f_2\Delta_{\zeta_1} + h_1'(0)f_2\Gamma_{\zeta_1} \\
d_{N_2^1} f_2(i_{Z_1}\Sigma) &= f_2(i_{Z_1}\Delta) + h_1'(0)f_2(i_{Z_1}\Gamma) + h_2'(0)f_2\Sigma_{\zeta_1} \\
d_{N_2^1} f_2\Psi_{\zeta_1} &= 0 \\
d_{N_2^1} f_2(i_{Z_1}\Psi) &= h_2'(0)f_2\Psi_{\zeta_1}.
\end{aligned}$$

Next, we express the above equations in terms of the original fields (cancel the renamings in (5.29)):

$$
\begin{aligned}
d_{N_2^1} f_2(f_1 i_T K)_{\varsigma_1} &= h_1'(0) f_2(f_1 K_\tau)_{\varsigma_1} \\
d_{N_2^1} f_2 i_{Z_1}(f_1 i_T K) &= h_1'(0) f_2 i_{Z_1}(f_1 K_\tau) + h_2'(0) f_2(f_1 i_T K)_{\varsigma_1} \\
d_{N_2^1} f_2(f_1 C_\tau)_{\varsigma_1} &= f_2(f_1 Q_\tau)_{\varsigma_1} \\
d_{N_2^1} f_2 i_{Z_1}(f_1 C_\tau) &= f_2 i_{Z_1}(f_1 Q_\tau) + h_2'(0) f_2(f_1 C_\tau)_{\varsigma_1} \\
d_{N_2^1} f_2(f_1 i_T C)_{\varsigma_1} &= f_2(f_1 i_T Q)_{\varsigma_1} + h_1'(0) f_2(f_1 C_\tau)_{\varsigma_1} \\
d_{N_2^1} f_2 i_{Z_1}(f_1 i_T C) &= f_2 i_{Z_1}(f_1 i_T Q) + h_1'(0) f_2 i_{Z_1}(f_1 C_\tau) \\
&\quad + h_2'(0) f_2(f_1 i_T C)_{\varsigma_1} \\
d_{N_2^1} f_2(f_1 K_\tau)_{\varsigma_1} &= 0 \\
d_{N_2^1} f_2 i_{Z_1}(f_1 K_\tau) &= h_2'(0) f_2(f_1 K_\tau)_{\varsigma_1}.
\end{aligned} \tag{5.31}
$$

Finally, to express the above equations in terms of the original fields in the orbit space N, we use the canonical identification $Gp \mapsto G_2(G_1 p)$ of N_2^1 and N. Let $f : \Omega(M) \to \Omega(N)$ be the pullback from M to the orbit space N defined by the G-reduced domain A. On the other hand, because $f_1 : \Omega(M) \to \Omega(N_1)$ is the pullback from M to the orbit space N_1 defined by the G_1-reduced domain A_1, and because $f_2 : \Omega(N_1) \to \Omega(N_2^1)$ is the pullback from N_1 to the orbit space N_2^1 defined by A_1 and A_2, it follows that $f_2 \circ f_1 : \Omega(M) \to \Omega(N_2^1)$ is the pullback from M to the orbit space N_2^1 defined by A. Then the canonical identification of N_2^1 and N can be used to canonically identify $f\omega$ and $(f_2 \circ f_1)\omega$. Let us use an example to identify the fields; the field $f_2 i_{Z_1}(f_1 K_\tau)$ is identified with the field $f(i_Z K_\tau) = f K_{\tau Z}$. To see this, let us first apply the pullback f to the definition of $i_Z K_\tau$:

$$
f(i_Z K_\tau) = f(i_Z(K - \tau \wedge i_T K)) = f(i_Z K + \tau \wedge i_Z i_T K)) = f(i_Z K)
$$

Then we perform similar calculations with $f_2 i_{Z_1}(f_1 K_\tau)$. For this, notice that $Z_1 = (\pi_1)_* Z$ and $f_1 = \kappa_1^*$ hold, plus we also need Lemma 3.2:

$$
\begin{aligned}
f_2 i_{Z_1}(f_1 K_\tau) &= f_2 i_{Z_1}(f_1 K - f_1(\tau \wedge i_T K)) = f_2 i_{Z_1} f_1 K \\
&= f_2 i_{(\pi_1)_* Z} \kappa_1^* K = f_2 \kappa_1^* i_{(\kappa_1)_*(\pi_1)_* Z} K = f_2 f_1 i_{Z_{|A_1}} K \\
&= f_2 f_1 i_Z K = (f_2 \circ f_1)(i_Z K).
\end{aligned}
$$

Thus using the canonical identification of f and $(f_2 \circ f_1)$, we see that $f K_{\tau Z}$ and $f_2 i_{Z_1}(f_1 K_\tau)$ have now been identified. With similar calculations we can show the other identifications:

$$
\begin{aligned}
f_2(f_1 i_T K)_{\varsigma_1} &\longleftrightarrow f(i_T K)_\varsigma = f K_{T\varsigma} \\
f_2 i_{Z_1}(f_1 i_T K) &\longleftrightarrow f(i_Z i_T K) = f K_{TZ} \\
f_2(f_1 K_\tau)_{\varsigma_1} &\longleftrightarrow f(K_\tau)_\varsigma = f K_{\tau\varsigma}.
\end{aligned}
$$

Applying the identifications to the equations in (5.31), we deduce the reduced differential equations in N:

$$\begin{aligned}
d_N(fK_{T\zeta}) &= h_1'(0)fK_{\tau\zeta} \\
d_N(fK_{TZ}) &= h_1'(0)fK_{\tau Z} + h_2'(0)fK_{T\zeta} \\
d_N(fC_{\tau\zeta}) &= fQ_{\tau\zeta} \\
d_N(fC_{\tau Z}) &= fQ_{\tau Z} + h_2'(0)fC_{\tau\zeta} \\
d_N(fC_{T\zeta}) &= fQ_{T\zeta} + h_1'(0)fC_{\tau\zeta} \\
d_N(fC_{TZ}) &= fQ_{TZ} + h_1'(0)fC_{\tau Z} + h_2'(0)fC_{T\zeta} \\
d_N(fK_{\tau\zeta}) &= 0 \\
d_N(fK_{\tau Z}) &= h_2'(0)fK_{\tau\zeta}.
\end{aligned} \tag{5.32}$$

As the above shows, we have in general lots of complicated equations. However, often some of them are trivial in the sense that all possible fields satisfy them (the exterior derivative of $(n-2)$-form in a $(n-2)$-manifold is zero). Furthermore, in low dimensions some geometric components are not even defined, whereupon the corresponding differential equations are not defined either. If M is four-dimensional and C and K are two-forms, then they have all the geometric components, and Q has three of the components because $Q_{\tau\zeta}$ is zero. In that case, all the above equations are defined, and assuming $(G,(h_1,h_2))$-invariance, only two of them are homogeneous and two are trivial. Thus the full generality of the above equations cannot even be used in electromagnetics. The following example shows how in common cases the equations are considerably simplified:

Example 5.8.1. Let a magnetostatic BVP on a three-manifold correspond to a situation where we have an infinitely long straight circular conductor with a constant current. Then the problem has a $(\mathbb{R}\times S^1)$-invariance (cylindrical symmetry). Assume that Z defines the translations, in which case the current J is in the direction of Z. Then T defines the rotations. Now the fields H, B, and J have only some of the geometric components:

$$\begin{aligned}
H &= H_{\tau\zeta} + H_{\tau Z}\zeta + H_T\zeta\tau \\
B &= \zeta\wedge B_{\tau Z} + \tau\wedge B_{T\zeta} + B_{TZ}\tau\wedge\zeta \\
J &= \tau\wedge J_{T\zeta}.
\end{aligned}$$

The only equations of (5.32) defined and not trivial (i.e., not satisfied by all fields) are the following three simple differential equations (not partial differential equations because the orbit space is one-dimensional):

$$\begin{aligned} d_N(fB_{TZ}) &= 0 \\ d_N(fH_{\tau Z}) &= fJ_{T\zeta} \\ d_N(fH_{T\zeta}) &= 0. \end{aligned}$$

Let us next introduce the cylindrical coordinates (r, φ, z) such that the conductor is in the z-direction. If $\bar{r}$, $\bar{\varphi}$, and $\bar{z}$ are the coordinate basis vectors, and if dr, $d\varphi$ and dz constitute the corresponding dual basis, then $(\bar{\varphi}, d\varphi)$ and $(\bar{z}, dz)$ are observers compatible with the coordinate system. Now the geometric decompositions and the representations of the fields in the standard coordinate bases are:

$$\begin{aligned} H &= H_{d\varphi\, dz} + H_{d\varphi\, \bar{z}} dz + H_{\bar{\varphi}\, dz} d\varphi \\ &= H_r dr + H_\varphi d\varphi + H_z dz \\ B &= dz \wedge B_{d\varphi\, \bar{z}} + d\varphi \wedge B_{\bar{\varphi}\, dz} + B_{\bar{z}\, \bar{\varphi}} d\varphi \wedge dz \\ &= B_r d\varphi \wedge dz + B_\varphi dz \wedge dr + B_z dr \wedge d\varphi \\ J &= d\varphi \wedge J_{\bar{\varphi}\, dz} \\ &= J_z dr \wedge d\varphi. \end{aligned}$$

The differential equations in the orbit space are now:

$$\begin{aligned} \frac{dB_r}{dr} &= 0 \\ \frac{dH_\varphi}{dr} &= J_z \\ \frac{dH_z}{dr} &= 0. \end{aligned}$$

5.8.3 Reduced boundary values

For the geometric components, boundary values are derived as in the case of one-dimensional symmetry groups. In fact, the same equations

$$\begin{aligned} f_{\partial N} \circ t_M &= t_N \circ f \\ i_T \circ t_M &= t_M \circ i_T \\ i_Z \circ t_M &= t_M \circ i_Z \end{aligned}$$

hold for symmetry groups of any dimension greater than or equal to one. Thus let us apply them to the original boundary values. First, we apply the pullback $f_{\partial N}$ to obtain an equation for $C_{\tau\zeta}$:

$$\begin{array}{rcl} f_{\partial N} t_M C &=& f_{\partial N} c \\ t_N f(C_{\tau\zeta} + \zeta \wedge C_{\tau Z} + \tau \wedge C_{T\zeta} + \tau \wedge \zeta \wedge C_{TZ}) &=& f_{\partial N} c \\ t_N f C_{\tau\zeta} &=& f_{\partial N} c. \end{array}$$

We get the equation for $C_{\tau Z}$ when we first apply the contraction i_Z and then the pullback $f_{\partial N}$:

$$\begin{array}{rcl} f_{\partial N} i_Z t_M C &=& f_{\partial N}(i_Z c) \\ f_{\partial N} t_M i_Z (C_{\tau\zeta} + \zeta \wedge C_{\tau Z} + \tau \wedge C_{T\zeta} + \tau \wedge \zeta \wedge C_{TZ}) &=& f_{\partial N}(i_Z c) \\ t_N f(C_{\tau Z} - \tau \wedge C_{TZ}) &=& f_{\partial N}(i_Z c) \\ t_N f C_{\tau Z} &=& f_{\partial N}(i_Z c). \end{array}$$

If we first apply i_T and then $f_{\partial N}$ to the original boundary value equation, we obtain the equation for $C_{T\zeta}$:

$$\begin{array}{rcl} f_{\partial N} i_T t_M C &=& f_{\partial N}(i_T c) \\ f_{\partial N} t_M i_T (C_{\tau\zeta} + \zeta \wedge C_{\tau Z} + \tau \wedge C_{T\zeta} + \tau \wedge \zeta \wedge C_{TZ}) &=& f_{\partial N}(i_T c) \\ t_N f(C_{T\zeta} + \zeta \wedge C_{TZ}) &=& f_{\partial N}(i_T c) \\ t_N f C_{T\zeta} &=& f_{\partial N}(i_T c). \end{array}$$

Finally, to obtain the equation for C_{TZ}, we apply $f_{\partial N} i_Z i_T$ to the original equation:

$$\begin{array}{rcl} f_{\partial N} i_Z i_T t_M C &=& f_{\partial N}(i_Z i_T c) \\ f_{\partial N} t_M i_Z i_T (C_{\tau\zeta} + \zeta \wedge C_{\tau Z} + \tau \wedge C_{T\zeta} + \tau \wedge \zeta \wedge C_{TZ}) &=& f_{\partial N}(i_Z i_T c) \\ t_N f i_Z (C_{T\zeta} + \zeta \wedge C_{TZ}) &=& f_{\partial N}(i_Z i_T c) \\ t_N f C_{TZ} &=& f_{\partial N}(i_Z i_T c). \end{array}$$

In similar steps, we can derive the equations for the geometric components of K. In summary, the reduced boundary values are

$$\begin{array}{rcl} t_N^1 f C_{\tau\zeta} &=& f_{\partial N}^1 c \\ t_N^1 f C_{\tau Z} &=& f_{\partial N}^1 (i_Z c) \\ t_N^1 f C_{T\zeta} &=& f_{\partial N}^1 (i_T c) \\ t_N^1 f C_{TZ} &=& f_{\partial N}^1 (i_Z i_T c) \\ t_N^2 f K_{\tau\zeta} &=& f_{\partial N}^2 b \\ t_N^2 f K_{\tau Z} &=& f_{\partial N}^2 (i_Z b) \\ t_N^2 f K_{T\zeta} &=& f_{\partial N}^2 (i_T b) \\ t_N^2 f K_{TZ} &=& f_{\partial N}^2 (i_Z i_T b). \end{array}$$

If there are singular orbits such that, e.g., Z is zero on those orbits, then the right-hand terms in above equations containing contractions with respect to Z are zero. Furthermore, since in low dimensions the fields do not have all the geometric components, some of the above equations are not even defined.

5.8.4 Reduced constitutive equations

Reduced constitutive equations are derived like in the case of one-dimensional symmetry groups: first the original constitutive equation $K = \upsilon_M C$ is represented in terms of geometric components. Then we solve the equations for each geometric component of K, after which these equations work as definitions for new operators to be directly defined in the orbit space N.

In terms of geometric components, the constitutive equation $K = \upsilon_M C$ is

$$\begin{aligned} & K_{\tau\zeta} + \zeta \wedge K_{\tau Z} + \tau \wedge K_{T\zeta} + \tau \wedge \zeta \wedge K_{TZ} \\ = & \upsilon_M(C_{\tau\zeta}) + \upsilon_M(\zeta \wedge C_{\tau Z}) + \upsilon_M(\tau \wedge C_{T\zeta}) + \upsilon_M(\tau \wedge \zeta \wedge C_{TZ}). \end{aligned}$$

Because $f\tau = f\zeta = 0$, then simply applying the pullback f to the above equation yields an equation for $K_{\tau\zeta}$:

$$\begin{aligned} fK_{\tau\zeta} &= f\upsilon_M(C_{\tau\zeta}) + f\upsilon_M(\zeta \wedge C_{\tau Z}) \\ &\quad + f\upsilon_M(\tau \wedge C_{T\zeta}) + f\upsilon_M(\tau \wedge \zeta \wedge C_{TZ}). \end{aligned}$$

The desired equation for $K_{\tau Z}$ is derived by first contracting $K = \upsilon_M C$ with respect to Z and then pulling back the result:

$$\begin{aligned} fK_{\tau Z} &= fi_Z\upsilon_M(C_{\tau\zeta}) + fi_Z\upsilon_M(\zeta \wedge C_{\tau Z}) \\ &\quad + fi_Z\upsilon_M(\tau \wedge C_{T\zeta}) + fi_Z\upsilon_M(\tau \wedge \zeta \wedge C_{TZ}). \end{aligned}$$

Similarly, we can derive equations for $K_{T\zeta}$ and K_{TZ}:

$$\begin{aligned} fK_{T\zeta} &= fi_T\upsilon_M(C_{\tau\zeta}) + fi_T\upsilon_M(\zeta \wedge C_{\tau Z}) \\ &\quad + fi_T\upsilon_M(\tau \wedge C_{T\zeta}) + fi_T\upsilon_M(\tau \wedge \zeta \wedge C_{TZ}) \end{aligned}$$

$$\begin{aligned} fK_{TZ} &= fi_Zi_T\upsilon_M(C_{\tau\zeta}) + fi_Zi_T\upsilon_M(\zeta \wedge C_{\tau Z}) \\ &\quad + fi_Zi_T\upsilon_M(\tau \wedge C_{T\zeta}) + fi_Zi_T\upsilon_M(\tau \wedge \zeta \wedge C_{TZ}). \end{aligned}$$

These equations are in terms of the original operator υ, but we want to write them in terms of “reduced operators” that are defined in N and that operate directly on fields $fC_{\tau\zeta}$, $fC_{\tau Z}$, $fC_{T\zeta}$, and fC_{TZ}:

$$\begin{aligned}
fK_{\tau\zeta} &= \upsilon^{\tau\zeta}_{\tau\zeta}(fC_{\tau\zeta}) + \upsilon^{\tau\zeta}_{\tau Z}(fC_{\tau Z}) \\
&\quad + \upsilon^{\tau\zeta}_{T\zeta}(fC_{T\zeta}) + \upsilon^{\tau\zeta}_{TZ}(fC_{TZ}) \\
\\
fK_{\tau Z} &= \upsilon^{\tau Z}_{\tau\zeta}(fC_{\tau\zeta}) + \upsilon^{\tau Z}_{\tau Z}(fC_{\tau Z}) \\
&\quad + \upsilon^{\tau Z}_{T\zeta}(fC_{T\zeta}) + \upsilon^{\tau Z}_{TZ}(fC_{TZ}) \\
\\
fK_{T\zeta} &= \upsilon^{T\zeta}_{\tau\zeta}(fC_{\tau\zeta}) + \upsilon^{T\zeta}_{\tau Z}(fC_{\tau Z}) \\
&\quad + \upsilon^{T\zeta}_{T\zeta}(fC_{T\zeta}) + \upsilon^{T\zeta}_{TZ}(fC_{TZ}) \\
\\
fK_{TZ} &= \upsilon^{TZ}_{\tau\zeta}(fC_{\tau\zeta}) + \upsilon^{TZ}_{\tau Z}(fC_{\tau Z}) \\
&\quad + \upsilon^{TZ}_{T\zeta}(fC_{T\zeta}) + \upsilon^{TZ}_{TZ}(fC_{TZ}).
\end{aligned}$$

Thus we have to define the following 16 reduced operators:

$$\begin{aligned}
\upsilon^{\tau\zeta}_{\tau\zeta} &: \Omega^{k}(N) \to \Omega^{n-k}(N) \\
\upsilon^{\tau\zeta}_{\tau Z} &: \Omega^{k-1}(N) \to \Omega^{n-k}(N) \\
\upsilon^{\tau\zeta}_{T\zeta} &: \Omega^{k-1}(N) \to \Omega^{n-k}(N) \\
\upsilon^{\tau\zeta}_{TZ} &: \Omega^{k-2}(N) \to \Omega^{n-k}(N) \\
\\
\upsilon^{\tau Z}_{\tau\zeta} &: \Omega^{k}(N) \to \Omega^{n-k-1}(N) \\
\upsilon^{\tau Z}_{\tau Z} &: \Omega^{k-1}(N) \to \Omega^{n-k-1}(N) \\
\upsilon^{\tau Z}_{T\zeta} &: \Omega^{k-1}(N) \to \Omega^{n-k-1}(N) \\
\upsilon^{\tau Z}_{TZ} &: \Omega^{k-2}(N) \to \Omega^{n-k-1}(N) \\
\\
\upsilon^{T\zeta}_{\tau\zeta} &: \Omega^{k}(N) \to \Omega^{n-k-1}(N) \\
\upsilon^{T\zeta}_{\tau Z} &: \Omega^{k-1}(N) \to \Omega^{n-k-1}(N) \\
\upsilon^{T\zeta}_{T\zeta} &: \Omega^{k-1}(N) \to \Omega^{n-k-1}(N) \\
\upsilon^{T\zeta}_{TZ} &: \Omega^{k-2}(N) \to \Omega^{n-k-1}(N) \\
\\
\upsilon^{TZ}_{\tau\zeta} &: \Omega^{k}(N) \to \Omega^{n-k-2}(N) \\
\upsilon^{TZ}_{\tau Z} &: \Omega^{k-1}(N) \to \Omega^{n-k-2}(N) \\
\upsilon^{TZ}_{T\zeta} &: \Omega^{k-1}(N) \to \Omega^{n-k-2}(N) \\
\upsilon^{TZ}_{TZ} &: \Omega^{k-2}(N) \to \Omega^{n-k-2}(N).
\end{aligned}$$

Using similar arguments as in section 5.6.4, we can derive the following expressions for the reduced operators (r denotes the right-inverse of f, and the subindexes in f and r denote the degree of forms they operate on):

$$\begin{aligned}
\upsilon^{\tau\zeta}_{\tau\zeta} &= f_{n-k} \circ \upsilon_M \circ r_k \\
\upsilon^{\tau\zeta}_{\tau Z} &= f_{n-k} \circ \upsilon_M \circ I_\zeta \circ r_{k-1} \\
\upsilon^{\tau\zeta}_{T\zeta} &= f_{n-k} \circ \upsilon_M \circ I_\tau \circ r_{k-1} \\
\upsilon^{\tau\zeta}_{TZ} &= f_{n-k} \circ \upsilon_M \circ I_\tau \circ I_\zeta \circ r_{k-2}
\end{aligned}$$

$$\begin{aligned}
\upsilon^{\tau Z}_{\tau\zeta} &= f_{n-k-1} \circ i_Z \circ \upsilon_M \circ r_k \\
\upsilon^{\tau Z}_{\tau Z} &= f_{n-k-1} \circ i_Z \circ \upsilon_M \circ I_\zeta \circ r_{k-1} \\
\upsilon^{\tau Z}_{T\zeta} &= f_{n-k-1} \circ i_Z \circ \upsilon_M \circ I_\tau \circ r_{k-1} \\
\upsilon^{\tau Z}_{TZ} &= f_{n-k-1} \circ i_Z \circ \upsilon_M \circ I_\tau \circ I_\zeta \circ r_{k-2}
\end{aligned}$$

$$\begin{aligned}
\upsilon^{T\zeta}_{\tau\zeta} &= f_{n-k-1} \circ i_T \circ \upsilon_M \circ r_k \\
\upsilon^{T\zeta}_{\tau Z} &= f_{n-k-1} \circ i_T \circ \upsilon_M \circ I_\zeta \circ r_{k-1} \\
\upsilon^{T\zeta}_{T\zeta} &= f_{n-k-1} \circ i_T \circ \upsilon_M \circ I_\tau \circ r_{k-1} \\
\upsilon^{T\zeta}_{TZ} &= f_{n-k-1} \circ i_T \circ \upsilon_M \circ I_\tau \circ I_\zeta \circ r_{k-2}
\end{aligned}$$

$$\begin{aligned}
\upsilon^{TZ}_{\tau\zeta} &= f_{n-k-2} \circ i_Z \circ i_T \circ \upsilon_M \circ r_k \\
\upsilon^{TZ}_{\tau Z} &= f_{n-k-2} \circ i_Z \circ i_T \circ \upsilon_M \circ I_\zeta \circ r_{k-1} \\
\upsilon^{TZ}_{T\zeta} &= f_{n-k-2} \circ i_Z \circ i_T \circ \upsilon_M \circ I_\tau \circ r_{k-1} \\
\upsilon^{TZ}_{TZ} &= f_{n-k-2} \circ i_Z \circ i_T \circ \upsilon_M \circ I_\tau \circ I_\zeta \circ r_{k-2}.
\end{aligned}$$

In formal matrix notation, we can write the reduced constitutive equations into the following single equation:

$$\begin{bmatrix} fK_{\tau\zeta} \\ fK_{\tau Z} \\ fK_{T\zeta} \\ fK_{TZ} \end{bmatrix} = \begin{bmatrix} \upsilon^{\tau\zeta}_{\tau\zeta} & \upsilon^{\tau\zeta}_{\tau Z} & \upsilon^{\tau\zeta}_{T\zeta} & \upsilon^{\tau\zeta}_{TZ} \\ \upsilon^{\tau Z}_{\tau\zeta} & \upsilon^{\tau Z}_{\tau Z} & \upsilon^{\tau Z}_{T\zeta} & \upsilon^{\tau Z}_{TZ} \\ \upsilon^{T\zeta}_{\tau\zeta} & \upsilon^{T\zeta}_{\tau Z} & \upsilon^{T\zeta}_{T\zeta} & \upsilon^{T\zeta}_{TZ} \\ \upsilon^{TZ}_{\tau\zeta} & \upsilon^{TZ}_{\tau Z} & \upsilon^{TZ}_{T\zeta} & \upsilon^{TZ}_{TZ} \end{bmatrix} \begin{bmatrix} fC_{\tau\zeta} \\ fC_{\tau Z} \\ fC_{T\zeta} \\ fC_{TZ} \end{bmatrix}.$$

On the other hand, let C and K be two-forms in a four-dimensional manifold. Let X and Y be vector fields tangent to the leaves of foliation $\{A_g\}_{g\in G}$ and Z and T tangent to the orbits of F_2 and F_1, respectively. Then (X, Y, Z, T) is a basis field for the tangent spaces of M. Its dual basis (dx, dy, ζ, τ) is a geometric basis for one-forms. Let $(dx\wedge dy, \zeta\wedge dx, \zeta\wedge dy, \tau\wedge dx, \tau\wedge dy, \tau\wedge\zeta)$

be the corresponding geometric basis for two-forms. In the geometric basis, C and K have the following expressions:

$$\begin{aligned}
C &= C_{xy} dx \wedge dy + C_{zx} \zeta \wedge dx + C_{zy} \zeta \wedge dy + C_{tx} \tau \wedge dx \\
&\quad + C_{ty} \tau \wedge dy + C_{tz} \tau \wedge \zeta \\
K &= K_{xy} dx \wedge dy + K_{zx} \zeta \wedge dx + K_{zy} \zeta \wedge dy + K_{tx} \tau \wedge dx \\
&\quad + K_{ty} \tau \wedge dy + K_{tz} \tau \wedge \zeta .
\end{aligned}$$

The geometric components of C and K in terms of the geometric basis are

$$\begin{aligned}
C_{\tau\zeta} &= C_{xy} dx \wedge dy \\
C_{\tau Z} &= C_{zx} dx + C_{zy} dy \\
C_{T\zeta} &= C_{tx} dx + C_{ty} dy \\
C_{TZ} &= C_{tz} \\
K_{\tau\zeta} &= K_{xy} dx \wedge dy \\
K_{\tau Z} &= K_{zx} dx + K_{zy} dy \\
K_{T\zeta} &= K_{tx} dx + K_{ty} dy \\
K_{TZ} &= K_{tz} .
\end{aligned}$$

In the geometric bases, the constitutive equation $K = \upsilon_M C$ can be written as follows:

$$\begin{bmatrix} K_{xy} \\ K_{zx} \\ K_{zy} \\ K_{tx} \\ K_{ty} \\ K_{tz} \end{bmatrix} = \begin{bmatrix}
\upsilon^{xy}_{xy} & \upsilon^{xy}_{zx} & \upsilon^{xy}_{zy} & \upsilon^{xy}_{tx} & \upsilon^{xy}_{ty} & \upsilon^{xy}_{tz} \\
\upsilon^{zx}_{xy} & \upsilon^{zx}_{zx} & \upsilon^{zx}_{zy} & \upsilon^{zx}_{tx} & \upsilon^{zx}_{ty} & \upsilon^{zx}_{tz} \\
\upsilon^{zy}_{xy} & \upsilon^{zy}_{zx} & \upsilon^{zy}_{zy} & \upsilon^{zy}_{tx} & \upsilon^{zy}_{ty} & \upsilon^{zy}_{tz} \\
\upsilon^{tx}_{xy} & \upsilon^{tx}_{zx} & \upsilon^{tx}_{zy} & \upsilon^{tx}_{tx} & \upsilon^{tx}_{ty} & \upsilon^{tx}_{tz} \\
\upsilon^{ty}_{xy} & \upsilon^{ty}_{zx} & \upsilon^{ty}_{zy} & \upsilon^{ty}_{tx} & \upsilon^{ty}_{ty} & \upsilon^{ty}_{tz} \\
\upsilon^{tz}_{xy} & \upsilon^{tz}_{zx} & \upsilon^{tz}_{zy} & \upsilon^{tz}_{tx} & \upsilon^{tz}_{ty} & \upsilon^{tz}_{tz}
\end{bmatrix} \begin{bmatrix} C_{xy} \\ C_{zx} \\ C_{zy} \\ C_{tx} \\ C_{ty} \\ C_{tz} \end{bmatrix} .$$

We see now from the above expressions that the matrices of the reduced operators are blocks of the matrix of the original operator υ_M:

$$\begin{gathered}
\upsilon^{\tau\zeta}_{\tau\zeta} = \begin{bmatrix} \upsilon^{xy}_{xy} \end{bmatrix} \quad
\upsilon^{\tau\zeta}_{\tau Z} = \begin{bmatrix} \upsilon^{xy}_{zx} & \upsilon^{xy}_{zy} \end{bmatrix} \quad
\upsilon^{\tau\zeta}_{T\zeta} = \begin{bmatrix} \upsilon^{xy}_{tx} & \upsilon^{xy}_{ty} \end{bmatrix} \quad
\upsilon^{\tau\zeta}_{TZ} = \begin{bmatrix} \upsilon^{xy}_{tz} \end{bmatrix} \\
\upsilon^{\tau Z}_{\tau\zeta} = \begin{bmatrix} \upsilon^{zx}_{xy} \\ \upsilon^{zy}_{xy} \end{bmatrix} \quad
\upsilon^{\tau Z}_{\tau Z} = \begin{bmatrix} \upsilon^{zx}_{zx} & \upsilon^{zx}_{zy} \\ \upsilon^{zy}_{zx} & \upsilon^{zy}_{zy} \end{bmatrix} \quad
\upsilon^{\tau Z}_{T\zeta} = \begin{bmatrix} \upsilon^{zx}_{tx} & \upsilon^{zx}_{ty} \\ \upsilon^{zy}_{tx} & \upsilon^{zy}_{ty} \end{bmatrix} \quad
\upsilon^{\tau Z}_{TZ} = \begin{bmatrix} \upsilon^{zx}_{tz} \\ \upsilon^{zy}_{tz} \end{bmatrix} \\
\upsilon^{T\zeta}_{\tau\zeta} = \begin{bmatrix} \upsilon^{tx}_{xy} \\ \upsilon^{ty}_{xy} \end{bmatrix} \quad
\upsilon^{T\zeta}_{\tau Z} = \begin{bmatrix} \upsilon^{tx}_{zx} & \upsilon^{tx}_{zy} \\ \upsilon^{ty}_{zx} & \upsilon^{ty}_{zy} \end{bmatrix} \quad
\upsilon^{T\zeta}_{T\zeta} = \begin{bmatrix} \upsilon^{tx}_{tx} & \upsilon^{tx}_{ty} \\ \upsilon^{ty}_{tx} & \upsilon^{ty}_{ty} \end{bmatrix} \quad
\upsilon^{T\zeta}_{TZ} = \begin{bmatrix} \upsilon^{tx}_{tz} \\ \upsilon^{ty}_{tz} \end{bmatrix} \\
\upsilon^{TZ}_{\tau\zeta} = \begin{bmatrix} \upsilon^{tz}_{xy} \end{bmatrix} \quad
\upsilon^{TZ}_{\tau Z} = \begin{bmatrix} \upsilon^{tz}_{zx} & \upsilon^{tz}_{zy} \end{bmatrix} \quad
\upsilon^{TZ}_{T\zeta} = \begin{bmatrix} \upsilon^{tz}_{tx} & \upsilon^{tz}_{ty} \end{bmatrix} \quad
\upsilon^{TZ}_{TZ} = \begin{bmatrix} \upsilon^{tz}_{tz} \end{bmatrix} .
\end{gathered}$$

Again, as the above equations show, we have in general lots of complicated equations. However, often a G-reduced domain and a coordinate system (bases) can be selected such that the original matrix of υ_M contains many zeros, in which case most of the reduced operators are zero. Furthermore, in low dimensions some geometric components are not even defined, in which case the corresponding constitutive equations are not defined either. The following examples show how in common cases the reduced constitutive equations are considerably simplified:

Example 5.8.2. Assume a static problem with one-dimensional translational symmetry and let the materials be isotropic. Then we actually have four 2d-problems resulting from a 4d-problem: The time-invariance (static problem) yields separate electrostatic and magnetostatic problems. Then the translational symmetry reduces these further to two separate 2d-problems as can be seen in section 5.6.4. Thus only four nonzero reduced operators can exists, and their matrices are diagonal.

Example 5.8.3. Let us study Example 5.8.1 again. Now because M a is three-dimensional manifold, the geometric components H_{TZ} and $B_{\tau\zeta}$ are not defined. Thus operators $\mu^{\tau\zeta}_{\tau\zeta}$, $\mu^{\tau\zeta}_{\tau Z}$, $\mu^{\tau\zeta}_{T\zeta}$, $\mu^{\tau\zeta}_{TZ}$, $\mu^{\tau Z}_{TZ}$, $\mu^{T\zeta}_{TZ}$, and μ^{TZ}_{TZ} are not defined either. Furthermore, the remaining nine operators are described with a single parameter each.

5.9 Dimensional reduction theorem of electromagnetic BVPs

In conclusion of this chapter and the theory of dimensional reduction of electromagnetic BVPs, we state a *dimensional reduction theorem* that provides sufficient conditions for dimensional reduction.

Dimensional reduction theorem. Let a linear electromagnetic BVP, that has a unique solution, be formulated on an n-dimensional manifold-with-boundary M, and let G be a k-dimensional Lie group ($k < n$) that is a product of connected one-dimensional Lie groups, and let $h : G \to \mathbb{F}$ be a Lie group homomorphism. If G acts effectively and differentiably on M such that

(1) the symmetry transformations of M are diffeomorphisms, and

(2) the sources, boundary values, and cohomology conditions of the fields are (G, h)-invariant, and the constitutive equations are G-invariant, and

(3) there exists a G-reduced domain, and all the G-reduced domains are canonically diffeomorphic, and

(4) all the singular orbits reside at the boundary of the orbit space, and

(5) there exists a G-invariant observer structure for each product factor of G,

then the BVP can be solved as an $(n-k)$-dimensional BVP on the orbit space. Furthermore, the solution of the $(n-k)$-dimensional BVP on the orbit space is unique.

This theorem is proved in the subsections of this chapter: It follows from assumptions (1) and (2) that the electromagnetic fields governed by the BVP are also (G, h)-invariant (see Theorem 5.2). From assumption (3) it follows that the orbit space exists and has a uniquely defined differentiable manifold structure (see Theorem 5.3). Then with a G-invariant observer structure, which exists by assumption (5), and with assumption (4) we can define constructively the reduced BVP on the orbit space from the original BVP (see sections 5.6 and 5.8). If the unique solution of the original BVP is pulled back to the orbit space, it gives a solution to the reduced BVP. On the other hand, every solution of the reduced BVP induces also a solution to the original BVP by (G, h)-invariance. However, because the original BVP has a unique solution, the induced solutions must be the same and hence the same also in the reduced BVP. Consequently, the reduced BVP has a unique solution.

Chapter 6

Applications

The tools of differential geometry can simply and effectively describe the theory of electromagnetism. Besides, they are not only fancy formalism describing a physical *theory* but also offer insight into improving the numerical modeling of *practical engineering* problems. In particular, they offer solutions to mesh generation problems, speed up parametric modeling, and describe simply how to solve open-boundary problems and exploit symmetries. Furthermore, blow-up problems related to axisymmetric problems can be avoided, and fashionable invisibility cloaking can be described in simple terms.

All the solutions presented here to the above problems and tasks are based on using previously defined differential geometric tools. Particularly, excluding dimensional reduction, the solution to various problems is always the same: *formulate an equivalent BVP on a diffeomorphic manifold to overcome the obstacles*. In other words, the equivalence of BVPs under diffeomorphisms provides a *unified approach* to these problems. However, though this approach certainly makes things possible, using it effectively requires thinking beyond traditions and sometimes beyond strong intuition.

6.1 Mesh generation problems

Often a numerical solution to electromagnetic BVPs requires a mesh for the domain. To generate a mesh, the domain is covered with a chart to enable use of arithmetics. However, computers do not use real numbers but rather *finite-precision floating point numbers* [36]. This is not always a problem, but floating points have a feature that can cause trouble: the distance between consecutive floating points increases as the absolute value of the numbers increases. Particularly problematic is the use of standard parameterizations, which involve small details far apart from each other. For example, consider the power line shown in Figure 6.1. Because of the situational symmetry, only one half of the line between the supports is modeled. The length of the lines is in the order of hundreds of meters, and everywhere in the domain the smallest dimensions are in the order of centimeters. In such cases, mesh generation can easily fail or may even be impossible with some particular mesh generators because of the poor accuracy of the floating points. The problem is a serious one because these failures prevent us obtaining any solution for BVPs.

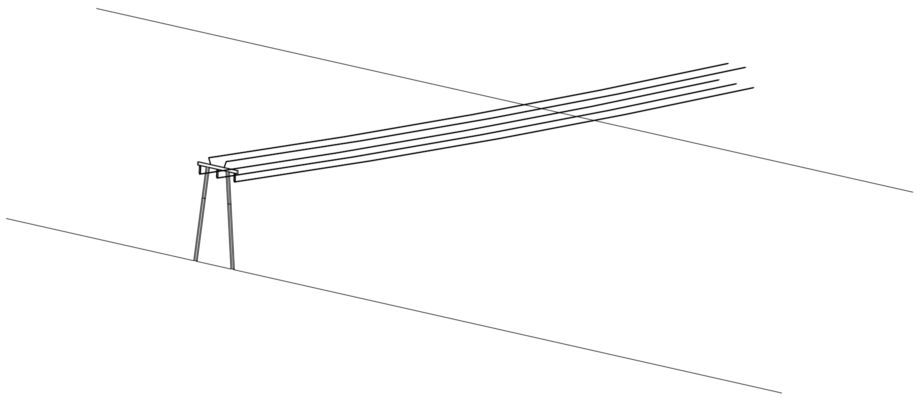

Figure 6.1: Standard parameterization of a power line.

The unified approach to formulating equivalent BVPs can be used to overcome these problems: one solution is to try to use another chart in which detailed parts are closer to each other; i.e., an equivalent BVP is formulated with some nonstandard parameterization as explained in chapter 4. For example, it is much easier to generate a mesh using the chart shown in Figure 6.2 than the one in Figure 6.1.

Using nonstandard parameterizations may have the disadvantage of a poor mesh, yet it does provide some sort of solution. The mesh may be of poor quality because mesh generation software programs assume only standard parameterizations (for more details on mesh quality criteria and nonstandard parameterizations, see [52] [54]).

If the problems cannot be solved reasonably using a single nonstandard parameterization, it is possible to use the full potential of manifolds and to cover the domain with multiple charts to help maintain sufficient floating point accuracy all over the domain (example in Figure 6.3). In the figure, the domain is presented with a standard parameterization and consists of three regions. The figure also shows how the domain can be covered with three

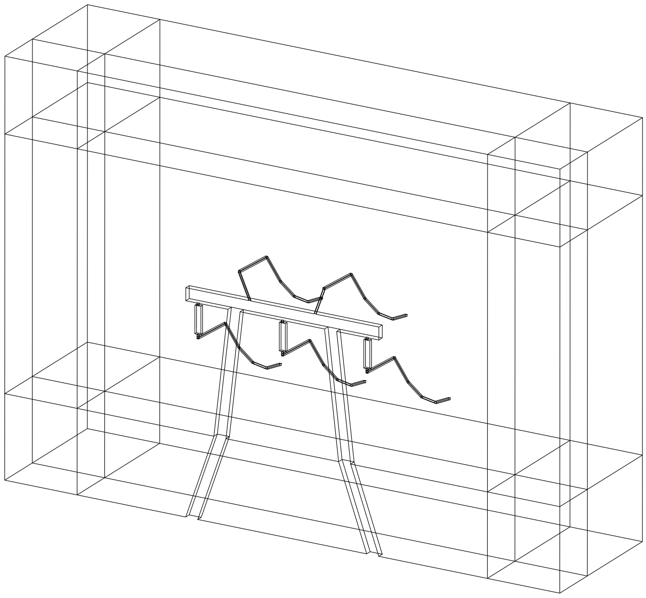

Figure 6.2: Nonstandard parameterization of a power line.

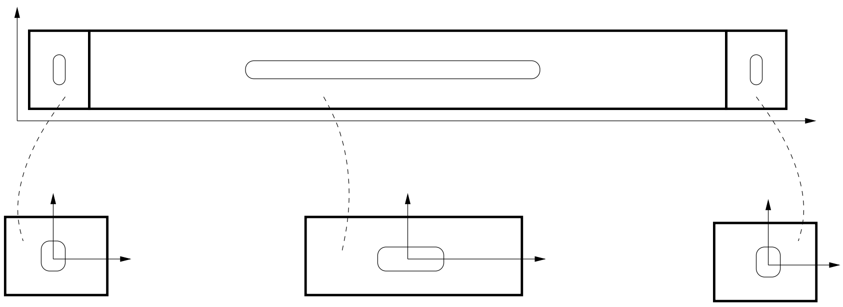

Figure 6.3: Multiple charts. Top, standard parameterization with three rectangular regions; bottom, three charts, each covering one rectangular region.

charts, each covering one of the regions, with the charts overlapping only at the regions' boundaries. Observe that the three charts can overlap only at their boundaries because the manifold can be covered with a single chart (standard parameterization). Moreover, to maintain sufficient floating point accuracy, the origins are moved and the dimensions changed (the three charts are not standard parameterizations). At the moment, most mesh generators do not allow use of multiple charts to cover the domain (for details on mesh generation with multiple charts, see [52] [54]).

6.2 Open-boundary problems

In many electromagnetic BVPs, the problem domain corresponds to the whole or a half space, in which case we have no boundaries, or the domain is only partly bounded by a boundary. These problems are often called open-boundary problems (the domain is still a manifold-with-boundary but possibly with an empty boundary). The fields tend to zero as the distance from the source increases. For a unique solution to these problems, we must give a zero "boundary value" at infinity, the so-called asymptotic condition, which is comparable to boundary values and enforces the fields to vanish as the distance from the source increases without a limit.

Because the fields tend rapidly to zero as the distance from the source increases, the most frequent method to solve these BVPs is to truncate the domain far enough and force the fields to zero at this artificial boundary. In many cases, this is adequate, but then the effects on solution accuracy are hard to estimate. Furthermore, the solution time may be unnecessarily long because a large number of element covers the uninteresting empty space.

Over the years, many other techniques have been proposed. For example, in the so-called "ballooning method" [59] the true distance of the boundary is pushed far away with a thin layer of special elements. Another method couples FEM with analytical solutions, as shown in [60]. In the boundary element method (BEM), only magnetizable regions need to be solved with FEM, after which fields in the rest of the space are solved from boundary integrals [67]. Infinite elements, which employ special decaying basis functions, can also be used, as in [4]. Finally, the transformation method, a.k.a., the shell transformation method, presented in [28] and [33], places the "infinity boundary" at a finite distance with the help of a suitable change of coordinates.

The last two methods proposed above, the infinite elements and the transformation method, are, in fact, based on the same idea of formulating equivalent problems: the user selects a chart where the "infinity boundary" is at a finite distance and formulates the problem with this chart. That is, as shown in Figure 6.4, the interesting part of the domain is bounded by an artificial surface, and the uninteresting empty space outside the surface is scaled down to a finite size in the sense of the standard metric of $\mathbb{R}^n$. Consequently, the domain is bounded, and there is a new boundary corresponding to the "infinity boundary." At the new boundary, the fields are set to zero, and the problem is solved like a normal BVP. In the transformation method, the effects of these changes are modeled in the material parameters, which are changed outside the interesting area according to the change of chart: the problem is first formulated in standard parameterization such as f in Figure

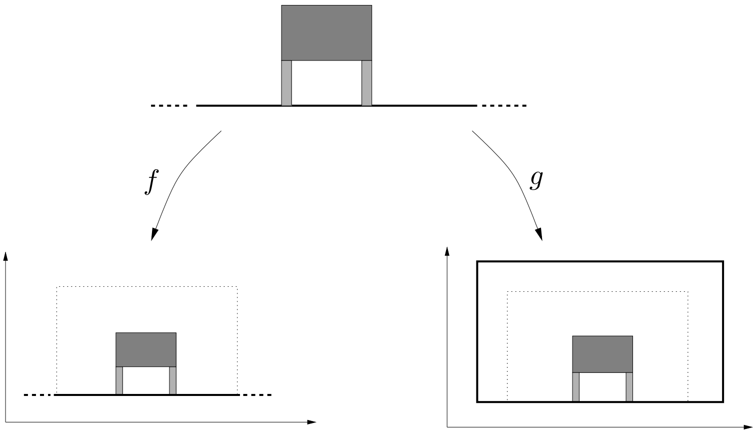

Figure 6.4: Open-boundary problem. The domain is a half space, and the interesting part of the domain contains a device above the ground. The interesting part of the domain is the rectangle shown in the charts f and g by the dotted gray line. The chart f is a standard parameterization; therefore, the domain is not bounded in the sense of the standard metric of $\mathbb{R}^n$. On the other hand, the chart g is such that the space outside the smaller rectangle is between the rectangles and thus the domain is bounded.

6.4, and then an equivalent BVP is formulated on a chart such as g in Figure 6.4. In the case of infinite elements, the problem is solved using charts such as g, but now the effects of the changes are modeled with special basis functions defined outside the interesting part of the domain. The basis functions are such that the result is the same as by using standard basis functions with changed material parameters.

It is important to note that the new boundary corresponding to the "infinity boundary" is not part of the manifold but is a result of a compactification, where the original noncompact manifold M is embedded in a compact manifold N, which differs from M only in those points that correspond to the new boundary. In other words, we have added new points to M to make it compact. This renders the material parameters singular at the new boundary: let the interesting part of the domain be a rectangle R whose width of and height are $2w$ and $2h$, respectively. Then let us make the following change of chart

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \mapsto \begin{bmatrix} a - \frac{1}{x_1 + \frac{1}{a-w} - w} \\ b - \frac{1}{y_1 + \frac{1}{a-h} - h} \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

from a standard parameterization f to a non-standard parameterization g. It maps the points of the upper left quadrant of $\mathbb{R}^2$ outside a rectangle R

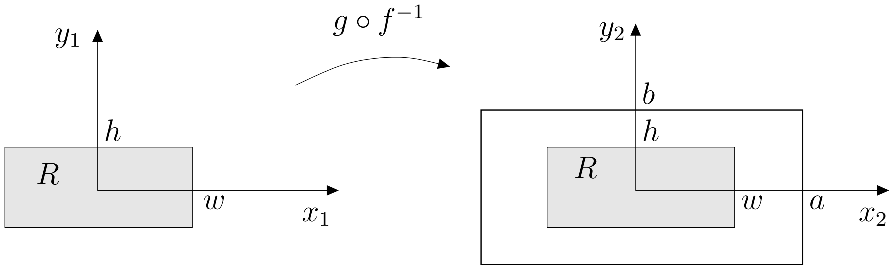

Figure 6.5: Change of charts. Left, a standard parameterization f, in which the codomain is the whole $\mathbb{R}^2$. Right, a nonstandard parameterization g, where the codomain is the region inside the outer rectangle of side lengths $2a$ and $2b$. In both charts the interesting part of the domain corresponds to the rectangle R of side lengths $2w$ and $2h$. The region outside R in f is mapped to the region between the rectangles in g.

centered at the origin to the points between R and a bigger rectangle of side length $2a$ and $2b$ (see Figure 6.5). Let the material parameters in f be described with a diagonal matrix $\epsilon_f = \epsilon\, I$, where ϵ is a real number and I the identity matrix. Then the above change of charts implies that the material parameters in g are given by the matrix

$$\epsilon_g = \epsilon \begin{bmatrix} \frac{(a-x_2)^2}{(b-y_2)^2} & 0 \\ 0 & \frac{(b-y_2)^2}{(a-x_2)^2} \end{bmatrix}.$$

Notice, that the codomain of the chart g has points with coordinates $x_2 = a$ and $y_2 = b$. These points are added points making up the new boundary. Observe that at these points the matrix ϵ_g is singular. However, this need not be a problem at all. For instance, in FEM, the integration over the elements is usually done with the *Gaussian quadrature* [36], which uses only a few *inside* points, where the matrix ϵ_g is never singular. Thus it seems that in practice such compactification works well. However, for a mathematically sound explanation, we need other arguments than those used in the example.

Open Question 8. A compact and noncompact manifold cannot be diffeomorphic because they are not homeomorphic. However, with numerical solution methods, equivalent BVPs can evidently be formulated on nonhomeomorphic manifolds (see the above discussion). Now is it possible rigorously to relax the homeomorphism requirement to define the equivalence of BVPs? Specifically, is it possible to use some kind of compactification and yet maintain equivalence? In other words, is it possible by compactification to change asymptotic conditions on a noncompact manifold to boundary values on a compact manifold?

6.3 Speeding up parametric models

Some BVPs involve object deformations. The shape of an object may change because some force is applied to it, or when the shape is optimized for some engineering goals. For example, magnetostriction changes the shape of ferromagnetic materials when subjected to a magnetic field. An example of shape optimization is the shape of the adjusting shims of an MRI magnet to obtain as homogeneous a field as possible. In addition, some engineering problems involve objects that move with respect to other objects in the problem domain (see Figure 6.6). Often these problems can be parameterized with a few parameters; i.e., a change in shape or movement can be described as changes in the values of some geometric parameters of the domain. Thus these problems can be solved by solving multiple BVPs, each corresponding to some parameter value.

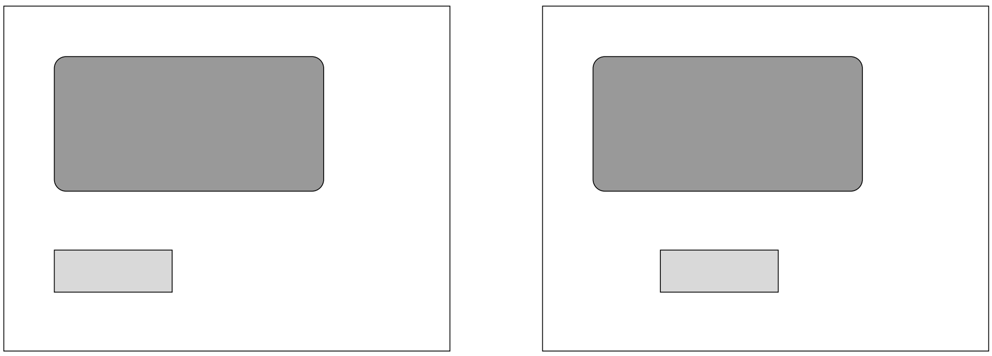

Figure 6.6: Modeling movement. Two BVP domains modeling a movement of a smaller material block with respect to a larger block. Notice that the domains are composed of two material blocks in the air, and as long as the blocks are not in contact, the domains are topologically the same.

Parameterizations have the advantage that multiple BVPs can be formulated in one go and the process can be automated. This saves some time and frees the modeler from formulating all problems separately. However, if one works only with standard parameterizations, there is little to do to speed up the process further.

In a large class of these parametric models, the unified approach to formulating equivalent problems can be used to accelerate the process. We can do it when we realize that often the differences between problem domains corresponding to different parameter values are purely *metrical*: For example, an object is a little longer or the distance of a moving object from other objects has changed. Thus the domains are *topologically the same* or homeomorphic to each other (see Figure 6.6). Moreover, because they are diffeomorphic, BVPs can be formulated in the same differentiable manifold, as explained in chapter 4. Because the same manifold is the domain, we can use the same chart for each BVP and *the same mesh for numerical solutions of all BVPs*.

Metrical differences between BVPs means that they have different Riemannian structures and thus different standard parameterizations. When the same chart is used for all BVPs, metrical differences are taken into account in the constitutive equations. Thus descriptions of BVPs on the chart differ only in material parameters, which can now be parameterized. That is, material parameters are parameterized, and changes in them translate into changes in geometry.

The possibility to use a single mesh can accelerate the solution process in many ways: obviously, time is saved in bypassing multiple mesh generations. Furthermore, only a part of the system matrix needs to be re-assembled because the effects of changes in parameter values are usually limited. In addition, in case of iterative solvers, the same preconditioner and initial guess may be used effectively for multiple parameter values.

In addition to saving time, the possibility of using a single mesh lends reliability to the solution process: for some parameter values, standard parameterization may be troublesome in generating a mesh. Such mesh generation problems can then stall an automated solution process, and in some optimization cases the whole process must be started from the beginning. Furthermore, because the same mesh is used for all parameter values (or at least for multiple parameter values), results can be compared easily and straightforwardly unlike in situations with different meshes for each parameter value. For more details and examples on accelerating parametric models, see [55].

6.4 Invisibility cloaking

The engineering problem with invisibility cloaking is to design and manufacture a material that makes anything inside the cloak invisible [50]. That is, when we use an invisibility cloak, we can see through an object that is cloaked, and it is impossible to observe the object with electromagnetic waves coming from any direction. The cloaking material is designed based on the unified approach of formulating equivalent BVPs: equivalent BVPs correspond to the same physical situation, and we are aiming at a physical situation in which the waves travel without reflections and other disturbances, though the space is not empty but contains a macroscopic region, the cloaked region. We seek to formulate an equivalent BVP on a diffeomorphic space where we know the material parameters and that the waves travel without disturbances. Such a space is empty space with one point removed (see Figure 6.7). We can then formulate the equivalent BVP, including equivalent material parameters, on the space with the cloaked region. The equivalent material parameters are then the material parameters of the desired cloaking material.

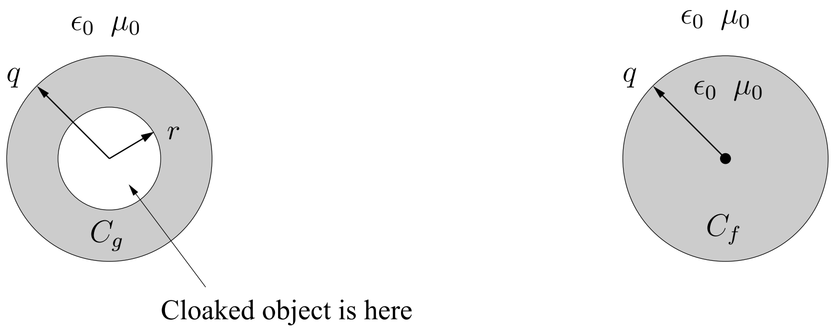

Figure 6.7: Invisibility cloaking. Left, a space with a hole of radius r. The hole is the cloaked region, and the gray region C_g represents the cloaking material. Right, a space with one point (the black dot) removed. The cloaking material region C_g corresponds to C_f, which is the gray circle of the radius q except the center point.

The following is a more detailed description. The goal is a BVP whose domain in standard parameterization has a macroscopic hole in it such that the inside of the hole is the space cloaked. The material parameters of the immediate surroundings of the hole (the cloaking material) are not known. Thus a standard parameterization g of the BVP domain can be chosen to be $\mathbb{R}^3 \setminus B(0,r)$. Furthermore, the cloaking material occupies the space $C_g = B(0,q) \setminus B(0,r)$, where $q > r$. Hence the cloaking material forms a layer with a thickness of $q - r$ around the ball of radius r. The material parameters for the region $\mathbb{R}^3 \setminus B(0,q)$ are empty space parameters, and at this point the material parameters of the region C_g are unknown (situation shown in Figure 6.7).

The equivalent BVP is formulated on a nonstandard parameterization f, where the BVP domain corresponds to $\mathbb{R}^3 \setminus \{0\}$, i.e., to a space with the origin removed (see Figure 6.7). The regions $\mathbb{R}^3 \setminus B(0,q)$ of g and f

correspond to each other via the identity mapping. But the region C_g of g corresponds now to the region $C_f = B(0,q) \setminus \{0\}$ of f via the diffeomorphism $G : C_f \to C_g$ defined by $x \mapsto (\frac{|x|}{q}(q-r)+r)\frac{x}{|x|}$ (for details on this mapping and cloaking analysis, see [23]). This mapping blows out the "point hole" to a hole of radius r in a radial fashion. Furthermore, the mapping G must be continuously extendable to the exterior boundary of C_f, which is not included in C_f, such that G is the identity mapping at the exterior boundary (observe that the exterior boundaries of C_f and C_g are the same subsets of $\mathbb{R}^3$).

The inverse $F = G^{-1}$ is now the mapping from the codomain of g to the codomain of f, where the BVP is fully defined. Then the pullback of F can be used to describe the equivalent material parameters in C_g: if the material parameters in C_f are given by $\mu_f = \mu_0 \star$ and $\epsilon_f = \epsilon_0 \star$, the equivalent material parameters in C_g are (see section 4.5)

$$\begin{aligned} \mu_g &= F^* \mu_f (F^*)^{-1} \\ \epsilon_g &= F^* \epsilon_f (F^*)^{-1}. \end{aligned}$$

If J_G is the Jacobian matrix of mapping G, then the matrix containing the equivalent material parameters is given by the following equation (see section 4.6):

$$\mu_g = \frac{\mu_0}{|J_G|} J_G J_G^T.$$

Notice that the cloaking theory is not primarily about waves but equivalent BVPs. Particularly, cloaking does not depend on frequency because frequency has no role in the diffeomorphism that defines the equivalence. Thus cloaking is possible, in theory, with all frequencies, including the static case where observing is not based on waves. Of course, in practice materials cannot be manufactured that have the same properties for all frequencies.

Cloaking shows that it is impossible, in general, to uniquely find out the interior of some object solely with boundary measurements. Because the material parameters ϵ_g and μ_g are anisotropic, cloaking is an example of a non-uniqueness result in the anisotropic version of the impedance tomography problem [22].

The following discussion introduces the next open question. To keep things simple, we often use only piecewise diffeomorphisms $F : N \to M$ to define the equivalence of BVPs on N and M. In other words, F is a homeomorphism, but its differentiability breaks down on some lower-dimensional subsets of N. For example, in the above, if the mapping G is extended from $C_f = B(0,q) \setminus \{0\}$ to $\mathbb{R}^3 \setminus \{0\}$ such that it is the identity mapping in $\mathbb{R}^3 \setminus B(0,q)$, then it is not differentiable at the points of the exterior boundary of C_f (although it is differentiable in the direction of the exterior boundary). However, it is continuous everywhere and thus a homeomorphism. Observe that in this example there exists suitable mappings that are differentiable everywhere though their representations with elementary functions may be quite complex.

Open Question 9. With numerical solution methods, based on finite approximations, piecewise diffeomorphisms seems to work well in formulating equivalent BVPs. Now is it possible rigorously to relax the diffeomorphisms requirement to define the equivalence of BVPs? And if so, what are the exact conditions that still allow equivalence? (Some sort of boundary conditions may possibly relax the diffeomorphisms requirement.)

6.5 Axisymmetric problems

Axisymmetric problems or rotational symmetric problems are naturally formulated in a cylindrical coordinate system. The change of chart from standard parameterization, which is a Cartesian xyz-coordinate system, to cylindrical $r\phi z$-coordinates yields the r^{-1}-term that appears somewhere in the solution process. This blows up the numerical solution near $r = 0$ [28] [43]. The unified approach to formulating equivalent problems solves this problem: make another change of charts where $r \mapsto r^2$ and formulate the problem with this new chart. Using this chart for solution is often called the rA-method [43].

6.6 Dimensional reduction

Finally, we discuss the benefits of dimensional reduction in numerical modeling and the benefits of using differential geometry to formulate lower-dimensional BVPs. The benefits of dimensional reduction in numerical methods are well-known and significant. First of all, in terms of complexity analysis, the complexity order of the problem can be reduced. In terms of solution time, such reduction can be significant. A second benefit is the reliability of obtaining a solution: it is much harder to generate a mesh for 3d than 2d domains. Furthermore, meshes can be generated much faster for lower-dimensional domains.

These benefits make it worth applying dimensional reduction whenever possible. However, classical vector analysis may be an obstruction in applying dimensional reduction. The main problem is how to formulate lower-dimensional BVPs with vector analysis because vector analysis is inherently three-dimensional, and some of its features have no natural counterparts in other dimensions. Moreover, dimensional reduction is independent of a metric, which further complicates the use of metric-based vector analysis. The inherent three-dimensionality and the omnipresence of the metric makes it cumbersome to apply dimensional reduction; e.g., is it obvious in general cases which are the proxy vectors for geometric components and what are constitutive equations? In the worst case, use of vector analysis can block the recognition of a possibility of dimensional reduction. Particularly, non-isometric symmetries are hard to recognize in the first place, let alone their applications with metric-based vector analysis. Indeed, cases that are hard to perceive as suitable for dimensional reduction, such as helicoidal geometries, are challenging to formulate by vector analysis. An example of this can be seen in [53], where an error escaped the author of this thesis and the three much more experienced co-authors.

In contrast to vector analysis, the tools of differential geometry are well-suited for dimensional reduction. All the main tools of differential geometry needed in dimensional reduction such as differential forms, exterior derivative, contraction, and Lie derivative, are naturally defined for every dimension. Furthermore, these tools are independent of coordinates and metric as is the symmetry on which dimensional reduction is based. These features give clear insight into the subject and turn the application of dimensional reduction into a mechanical procedure.

Chapter 7

Examples

7.1 Parametric models: shape optimization

Our first example is about shape optimization with a single mesh. The goal is to use a C-magnet with adjusting shims to generate as homogeneous a field as possible. Figure 7.1 shows the situation, and the task there is to optimize the shape of the adjusting shims to maximize homogeneity.

We must solve multiple magnetostatic BVPs with only slight changes in the geometry of the domains. In other words, if the BVPs are initially formulated using some rigid-body metric, then the BVPs have slightly different Riemannian structures and standard parameterizations. Because all the domains are diffeomorphic, we can formulate BVPs equivalently to a single differentiable manifold and use only one chart and mesh throughout the calculations. Because the same chart and mesh is used for all BVPs, the only difference between them is in the constitutive equations.

The shapes of the shims are described with a few geometric parameters in the chart chosen for the optimization problem; therefore, the problem can be parameterized with these geometric parameters. Because in the chosen chart the BVPs differ only in their constitutive equations, the representation of the operator μ ($B = \mu H$) as a matrix in the chart is parameterized by geometric parameters. The parameters are optimized with a genetic algorithm, which generates shapes based on an objective function. This example was published also in [55], but with fewer details.

For simplicity, the above problem is solved as a 2d-problem, but of course, the idea can be applied similarly to other dimensions as well. The material parameters, and thus the operators μ, are known in the standard parameterizations corresponding to the different shapes of the shims. Then we formulate equivalent BVPs on the chart chosen for mesh generation and calculations.

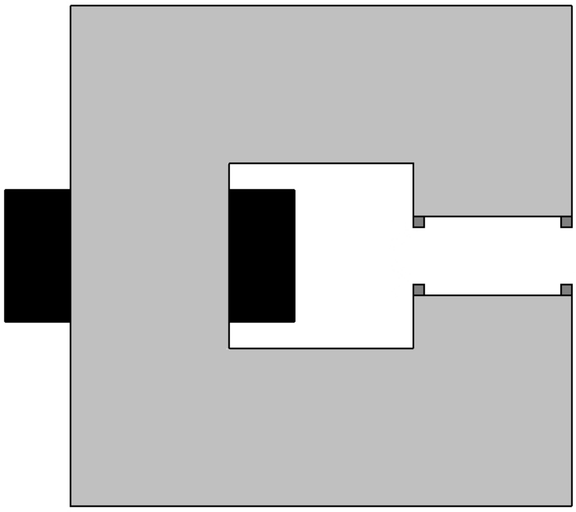

Figure 7.1: C-magnet. The gray area is occupied by a material with high permeability, and the black boxes are the current coils that generate the magnetic field. The small dark gray boxes represent the adjusting shims, made of the same high permeability material. The goal is to optimize the shape of the shims to make the field as homogeneous as possible in the middle of the air gap.

Because changes in the shape of the shims affect only the shims and their surroundings, we want to restrict the extent of the changes to the vicinity of the shims: Figure 7.2 shows how the domain is divided into regions, where the operator μ changes and thereby takes into account the changes in the shape of the shims. Because the outer boundaries of the triangles are fixed, the outside part remains the same for all parameter values. This example has three parameters to optimize the shape of the shims: their height and the width of the inner and outer shims.

Figures 7.3 and 7.4 show interesting parts of the generic standard parameterization and the corresponding part of the chart used in all calculations. The parameters we seek to optimize are the height of the shims $H_0 = f_0 - b = h - g_0$, the width of the inner shims $I_0 = c_0 - a$, and the width of the outer shims $O_0 = e - d_0$. The corresponding parameters in the chart used for calculations are $H = f - b = h - g$, $I = c - a$, and $O = e - d$. In addition, let $L_1 = e - a$ be the width of the air gap, $L_2 = h - b$ the height of the air gap, and $L_3 = b - i$ the height of regions 10-14. Finally, we also define the following parameters: $T_1 = \frac{d-c}{d_0-c_0}$, $T_2 = \frac{g-f}{g_0-f_0}$, $T_3 = \frac{j-c-T_1(j-c_0)}{L_3}$. With these geometric parameters, the change of charts mappings from the generic

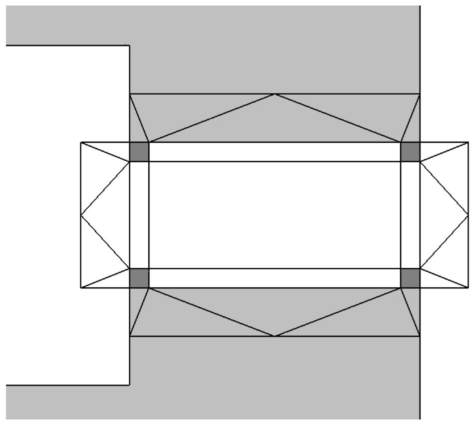

Figure 7.2: Division of the domain into regions. The four dark gray squares represent the shims. The small regions (the triangles and quadrilaterals) account for changes in the geometry.

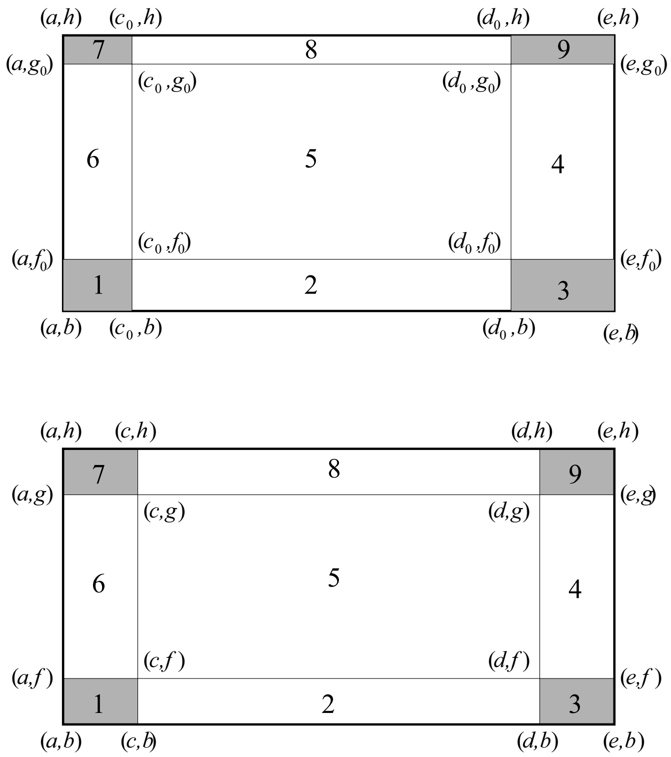

Figure 7.3: Coordinates of the corners of the regions. The dark gray areas represent the shims. Top, a generic standard parameterization; bottom, the chart used for calculations.

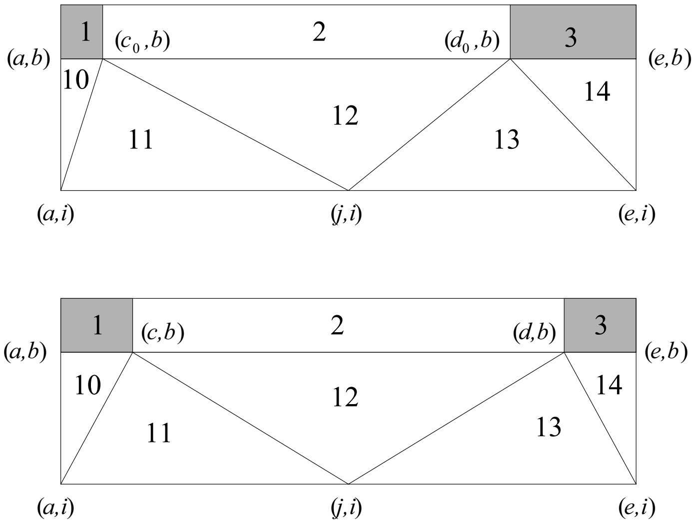

Figure 7.4: Coordinates of the apices of the regions. The dark gray areas represent the shims, and regions 10-14 are part of the C-magnet restricting the effects of changes on their outer boundary. Top, a generic standard parameterization; bottom, the chart used for calculations.

standard parameterization to the chart used for calculations in regions 1-14 are

$$G_1 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} a + \frac{c-a}{c_0-a}(x-a) \\ b + \frac{f-b}{f_0-b}(y-b) \end{bmatrix} = \begin{bmatrix} a + \frac{I}{I_0}(x-a) \\ b + \frac{H}{H_0}(y-b) \end{bmatrix}$$

$$G_2 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} c + \frac{d-c}{d_0-c_0}(x-c_0) \\ b + \frac{f-b}{f_0-b}(y-b) \end{bmatrix} = \begin{bmatrix} c + T_1(x-c_0) \\ b + \frac{H}{H_0}(y-b) \end{bmatrix}$$

$$G_3 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} d + \frac{e-d}{e-d_0}(x-d_0) \\ b + \frac{f-b}{f_0-b}(y-b) \end{bmatrix} = \begin{bmatrix} d + \frac{O}{O_0}(x-d_0) \\ b + \frac{H}{H_0}(y-b) \end{bmatrix}$$

$$G_4 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} d + \frac{e-d}{e-d_0}(x-d_0) \\ f + \frac{g-f}{g_0-f_0}(y-f_0) \end{bmatrix} = \begin{bmatrix} d + \frac{O}{O_0}(x-d_0) \\ f + T_2(y-f_0) \end{bmatrix}$$

$$G_5 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} c + \frac{d-c}{d_0-c_0}(x-c_0) \\ f + \frac{g-f}{g_0-f_0}(y-f_0) \end{bmatrix} = \begin{bmatrix} c + T_1(x-c_0) \\ f + T_2(y-f_0) \end{bmatrix}$$

$$G_6 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} a + \frac{c-a}{c_0-a}(x-a) \\ f + \frac{g-f}{g_0-f_0}(y-f_0) \end{bmatrix} = \begin{bmatrix} a + \frac{I}{I_0}(x-a) \\ f + T_2(y-f_0) \end{bmatrix}$$

$$G_7 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} a + \frac{c-a}{c_0-a}(x-a) \\ g + \frac{h-g}{h-g_0}(y-g_0) \end{bmatrix} = \begin{bmatrix} a + \frac{I}{I_0}(x-a) \\ g + \frac{H}{H_0}(y-g_0) \end{bmatrix}$$

$$G_8 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} c + \frac{d-c}{d_0-c_0}(x-c_0) \\ g + \frac{h-g}{h-g_0}(y-g_0) \end{bmatrix} = \begin{bmatrix} c + T_1(x-c_0) \\ g + \frac{H}{H_0}(y-g_0) \end{bmatrix}$$

$$G_9 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} d + \frac{e-d}{e-d_0}(x-d_0) \\ g + \frac{h-g}{h-g_0}(y-g_0) \end{bmatrix} = \begin{bmatrix} d + \frac{O}{O_0}(x-d_0) \\ g + \frac{H}{H_0}(y-g_0) \end{bmatrix}$$

$$G_{10} : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} a + \frac{c-a}{c_0-a}(x-a) \\ y \end{bmatrix} = \begin{bmatrix} a + \frac{I}{I_0}(x-a) \\ y \end{bmatrix}$$

$$G_{11} : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x + \frac{y-i}{b-i}(c-c_0) \\ y \end{bmatrix} = \begin{bmatrix} x + \frac{I-I_0}{L_3}(y-i) \\ y \end{bmatrix}$$

$$G_{13} : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x + \frac{y-i}{b-i}(d-d_0) \\ y \end{bmatrix} = \begin{bmatrix} x + \frac{O_0-O}{L_3}(y-i) \\ y \end{bmatrix}$$

$$G_{14} : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} e + \frac{d-e}{d_0-e}(x-e) \\ y \end{bmatrix} = \begin{bmatrix} e + \frac{O}{O_0}(x-e) \\ y \end{bmatrix}$$

$$x_0(y) = \frac{y-i}{b-i}(j-c_0) + c_0 \qquad x_1(y) = \frac{y-i}{b-i}(j-c) + c$$

$$G_{12} : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x_1 + \frac{d-c}{d_0-c_0}(x-x_0) \\ y \end{bmatrix} = \begin{bmatrix} x_1 + T_1(x-x_0) \\ y \end{bmatrix}.$$

The regions in Figure 7.2 that are similar to regions 10-14 have mappings similar to mappings $G_{10}-G_{14}$. If the solver software is based on proxy vectors instead of forms, equation (4.25) gives the equivalent operators μ for each region. We only need the Jacobian matrices of the above change of chart mappings. Let μ_s be the permeability of the regions 10-14 in the standard parameterization, then the matrices of the equivalent operators μ for the proxy vectors in the regions 1-14 are

$$\mu_1 = \mu_7 = \mu_s \begin{bmatrix} \frac{IH_0}{I_0H} & 0 \\ 0 & \frac{I_0H}{IH_0} \end{bmatrix}$$

$$\mu_2 = \mu_8 = \mu_0 \begin{bmatrix} T_1\frac{H_0}{H} & 0 \\ 0 & T_1^{-1}\frac{H}{H_0} \end{bmatrix}$$

$$\mu_3 = \mu_9 = \mu_s \begin{bmatrix} \frac{OH_0}{O_0H} & 0 \\ 0 & \frac{HO_0}{H_0O} \end{bmatrix}$$

$$\mu_4 = \mu_0 \begin{bmatrix} T_2^{-1}\frac{O}{O_0} & 0 \\ 0 & T_2\frac{O_0}{O} \end{bmatrix}$$

$$\mu_5 = \mu_0 \begin{bmatrix} \frac{T_1}{T_2} & 0 \\ 0 & \frac{T_2}{T_1} \end{bmatrix}$$

$$\mu_6 = \mu_0 \begin{bmatrix} T_2^{-1}\frac{I}{I_0} & 0 \\ 0 & T_2\frac{I_0}{I} \end{bmatrix}$$

$$\mu_{10} = \mu_s \begin{bmatrix} \frac{I}{I_0} & 0 \\ 0 & \frac{I_0}{I} \end{bmatrix}$$

$$\mu_{11} = \mu_s \begin{bmatrix} 1+(\frac{I-I_0}{L_3})^2 & \frac{I-I_0}{L_3} \\ \frac{I-I_0}{L_3} & 1 \end{bmatrix}$$

$$\mu_{12} = \frac{\mu_s}{T_1} \begin{bmatrix} T_1^2+T_3^2 & T_3 \\ T_3 & 1 \end{bmatrix}$$

$$\mu_{13} = \mu_s \begin{bmatrix} 1+(\frac{O_0-O}{L_3})^2 & \frac{O_0-O}{L_3} \\ \frac{O_0-O}{L_3} & 1 \end{bmatrix}$$

$$\mu_{14} = \mu_s \begin{bmatrix} \frac{O}{O_0} & 0 \\ 0 & \frac{O_0}{O} \end{bmatrix}.$$

We measure the homogeneity of the field in the following way: select a region A of the domain where proxy vector $\mathbf{B}$ of the field B is to be homogeneous and create a homogeneous vector field $\mathbf{b}_0$ of desired direction such that $\int_A \mathbf{b}_0 \cdot \mathbf{b}_0 dv = 1$, where $\cdot$ is the standard inner product of the chart. Our goal is a field $\mathbf{B}$ such that $\mathbf{B} = \alpha \mathbf{b}_0$ holds pointwise. For any field $\mathbf{B}$, consider the identity $\mathbf{B} = \alpha \mathbf{b}_0 + (\mathbf{B} - \alpha \mathbf{b}_0)$. Now to maximize the homogeneity of $\mathbf{B}$, the norm of the deviation $\int_A (\mathbf{B} - \alpha \mathbf{b}_0)^2 dv$ must be minimized. This happens when $\alpha = \int_A \mathbf{B} \cdot \mathbf{b}_0 dv$, in which case the norm of the deviation is $\int_A |\mathbf{B}|^2 dv - \alpha^2$. A genetic algorithm [31] varies the shape of the shims (the three parameters) to optimize this norm.

In our example, the width of the air gap in the C-magnet is 6 cm and its thickness 3 cm. The region A is a square in the middle of the gap with sides of length 1 cm. The region A and the calculated optimum result is shown in Figure 7.5. The optimum shapes calculated by the genetic algorithm are 7.9 mm for the width of the right side shims, 5.6 mm for the left side, and 0.8 mm for the shim height. The BVPs were solved with GetDP [20] and Gmsh [21], and the genetic algorithm ran in MATLAB.

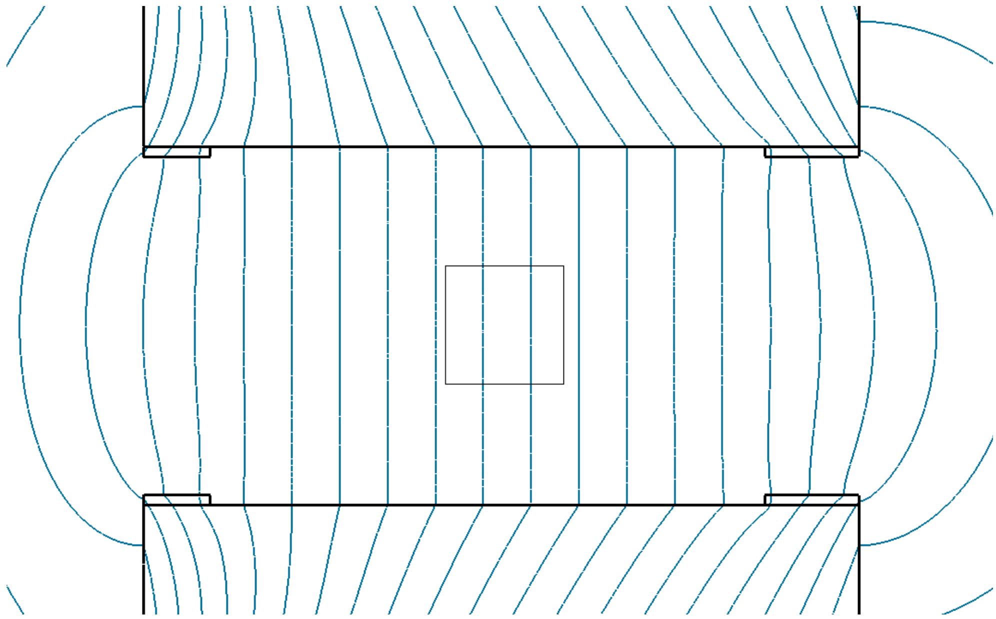

Figure 7.5: Standard parameterization of the C-magnet with optimized shims. Thinner lines stands for flux lines. The square in the middle of the air gap is the region A, where the field should be as homogeneous as possible.

7.2 Dimensional reduction: helicoidal geometries

Our second example is a concrete one of nontrivial dimensional reduction. We have a magnetostatic BVP, which depicts a magnetic field due to helicoidally twisted current wires (Figure 7.6).

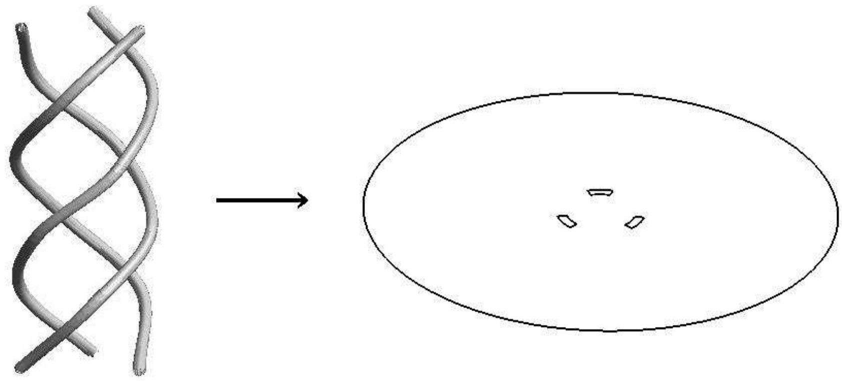

Figure 7.6: Twisted wires. Also shown is a G-reduced domain, which is a plane orthogonal to the direction of translations.

The domain is the whole space, but for simplicity the domain used in calculations is truncated far away from the wires, and fields are set to zero at this artificial boundary. In this manifold-with-boundary M, the following equations hold:

$$\begin{aligned} d_M H &= J \\ d_M B &= 0 \\ B &= \mu H. \end{aligned}$$

The standard way to formulate this problem is to use a standard parameterization f with Cartesian xyz-coordinates, in which case the operator μ is known. The symmetry transformations of the helicoidal action on M are such that the images of the orbits under f are helix curves; i.e., orbits result from combined translations and rotations. Thus the symmetry group G is $\mathbb{R}$, and the action $G \times M \to M$ is free. The boundary values, source J, operator μ, and cohomology conditions are all G-invariant under the helicoidal action. Thus the BVP has also a G-invariant solution under the action. Figure 7.6 shows also a G-reduced domain, which is a cross-plane and can be used to formulate a two-dimensional BVP in the orbit space.

Let A be some G-reduced domain corresponding to smooth G-invariant one-form τ on M and T some smooth nonzero G-invariant vector field on M everywhere along the orbits. Then the lower-dimensional BVP to be solved in the orbit space $N = M/G$ is as follows (see section 5.6.5):

$$\begin{aligned} dH_\tau &= J_\tau \\ d(i_T B) &= 0 \\ d(i_T H) &= 0 \\ dB_\tau &= 0 \\ i_T B &= \mu_\tau^T(H_\tau) + \mu_T^T(i_T H) \\ B_\tau &= \mu_\tau^\tau(H_\tau) + \mu_T^\tau(i_T H). \end{aligned}$$

Because currents are assumed to be in the direction of the orbits, J has only one geometric component. Observe that the equation $dB_\tau = 0$ is trivial, i.e., all two-forms satisfy it.

To solve the above BVP numerically requires that we cover the orbit space with a chart. This is done most conveniently by covering M with some chart, and then by selecting a G-reduced domain described in the chart as a coordinate level set, in which case this level set induces a chart for the orbit space. Figure 7.7 shows a chart f, where the translations are in the z-direction, which is also the axis of rotation (f is called a Euclidean chart because the geometry in f corresponds to rigid-body measurements). Now any plane parallel to the xy-plane (which is a coordinate level set for z-coordinate) is a G-reduced domain and could be used to formulate the reduced problem. Therefore, let us select the xy-plane and denote it by A. However, the standard basis of this chart is not a geometric basis in the sense that Z (the standard basis vector in the z-direction) is not generally in the direction of the orbits. To obtain a geometric basis, we use the chart g with helicoidal coordinates to straighten out the twisting (see Figure 7.7). The uvw-coordinates of g are then given with respect to the xyz-coordinates of f as follows:

$$\begin{cases} u = x\cos(\alpha z) - y\sin(\alpha z) \\ v = x\sin(\alpha z) + y\cos(\alpha z) \\ w = z, \end{cases} \tag{7.1}$$

where α is the twist pitch describing the extent of twisting. Notice that now W (the standard basis vector in the w-direction) is in the direction of the orbits. Now A corresponds to the uv-plane; i.e., A is a particular submanifold of M, but in the two charts its image is the xy-plane and the uv-plane. Furthermore, under the chart g, the orbits (lines parallel to the w-axis) are straight in the w-direction, and the symmetry transformations are pure translations.

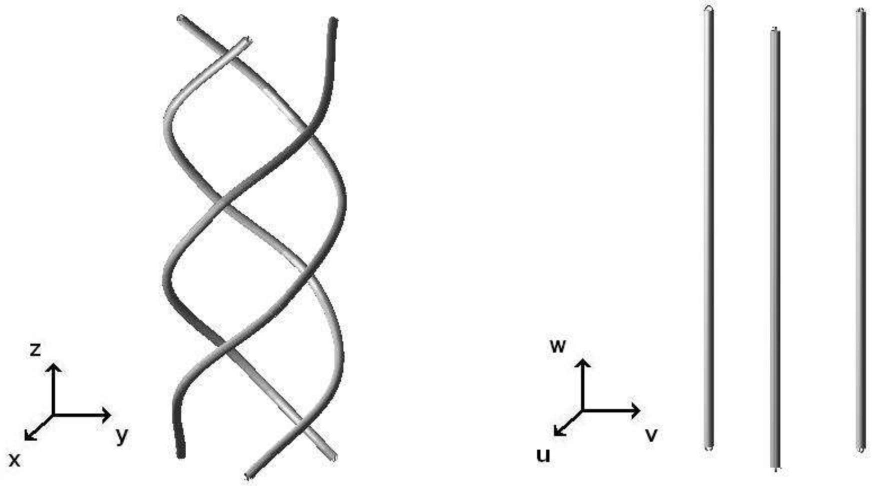

Figure 7.7: Two different charts of the same manifold M. On the left is the Euclidean chart.

Let W be the induced basis vector of the uvw-coordinates in the w-direction and let dw be a one-form satisfying $dw(W) = 1$ and $dw(U) = dw(V) = 0$; i.e. dw is the coordinate differential of the w-coordinate. Then (W, dw) is an observer compatible with the uvw-coordinates and the problem to be solved in the uv-plane, which is a chart for the orbit space, is the following:

$$\begin{aligned} dH_{dw} &= J_{dw} \\ d(i_W B) &= 0 \\ d(i_W H) &= 0 \\ dB_{dw} &= 0 \\ i_W B &= \mu^W_{dw}(H_{dw}) + \mu^W_W(i_W H) \\ B_{dw} &= \mu^{dw}_{dw}(H_{dw}) + \mu^{dw}_W(i_W H). \end{aligned}$$

In the standard basis of the chart f, the matrix of μ is simply $\mu_0 I$, where μ_0 is the permeability of empty space, and I is the identity matrix. By equation (4.25), the matrix of μ in the chart g is

$$\mu_g = \frac{1}{|J_c|} J_c \mu_f J_c^T = \mu_0 \begin{bmatrix} 1+\alpha^2 v^2 & -\alpha^2 uv & -\alpha v \\ -\alpha^2 uv & 1+\alpha^2 u^2 & \alpha u \\ -\alpha v & \alpha u & 1 \end{bmatrix}, \tag{7.2}$$

where J_c is the Jacobian matrix of the change of coordinates in (7.1) when $z = w = 0$. The matrices of the operators μ^W_{dw}, μ^W_W, μ^{dw}_{dw}, and μ^{dw}_W in the standard basis of the uv-plane are the blocks of the matrix of μ_g in (7.2) (see section 5.6.4).

The equation $dB_{dw} = 0$ is trivial, i.e., all two-forms satisfy it. On the other hand, the equation $d(i_W H) = 0$ says that the partial derivatives of the real function $i_W H$ with respect to u- and v-coordinates are everywhere zero. Together with zero boundary conditions this means that $i_W H = 0$. This now simplifies the constitutive equations, and we have the following problem:

$$\begin{aligned} dH_{dw} &= J_{dw} \\ d(i_W B) &= 0 \\ i_W B &= \mu^{W}_{dw}(H_{dw}) \\ B_{dw} &= \mu^{dw}_{dw}(H_{dw}). \end{aligned}$$

Observe that the last equation is only for evaluation.

In the standard basis of the uv-plane, H_{dw} and $i_W B$ are represented with the following component vectors:

$$H_{dw} = \begin{bmatrix} H_u \\ H_v \end{bmatrix} \qquad i_W B = \begin{bmatrix} B_u \\ B_v \end{bmatrix}.$$

Then the matrix of μ^{W}_{dw} is given as the upper left block of the matrix μ_g given in (7.2):

$$\mu^{W}_{dw} = \mu_0 \begin{bmatrix} 1+\alpha^2 v^2 & -\alpha^2 uv \\ -\alpha^2 uv & 1+\alpha^2 u^2 \end{bmatrix}.$$

The solution of the problem gives us components H_{dw} and $i_W B$. Now B_{dw} can be found from the other constitutive equation simply by evaluation:

$$B_{dw} = \mu^{dw}_{dw} H_{dw},$$

where the component vector of B_{dw} and the matrix of μ^{dw}_{dw} are

$$B_{dw} = \begin{bmatrix} B_w \end{bmatrix} \qquad \mu^{dw}_{dw} = \mu_0 \begin{bmatrix} -\alpha v & \alpha u \end{bmatrix}.$$

Thus we have $H = H_{dw}$ and $B = B_{dw} + dw \wedge i_W B$. Notice that H has no component along the orbits, and that in vector notation the following equations hold:

$$H_g = \begin{bmatrix} H_u \\ H_v \\ 0 \end{bmatrix} \qquad B_g = \begin{bmatrix} B_u \\ B_v \\ B_w \end{bmatrix}.$$

To represent these in the chart f, we apply the pullbacks (see section 4.6):

$$H_f = J_c^T H_g = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\alpha v & \alpha u & 1 \end{bmatrix} \begin{bmatrix} H_u \\ H_v \\ 0 \end{bmatrix} = \begin{bmatrix} H_u \\ H_v \\ -\alpha v H_u + \alpha u H_v \end{bmatrix} = \begin{bmatrix} H_x \\ H_y \\ H_z \end{bmatrix}$$

$$B_f = |J_c| J_c^{-1} B_g = \begin{bmatrix} 1 & 0 & \alpha v \\ 0 & 1 & -\alpha u \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} B_u \\ B_v \\ B_w \end{bmatrix} = \begin{bmatrix} B_u + \alpha v B_w \\ B_v - \alpha u B_w \\ B_w \end{bmatrix} = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix}.$$

Thus, as expected, the Euclidean chart f has three nonzero components. These equations also show that $B_f = \mu H_f$ holds for the component vectors B_f and H_f, where μ is a scalar. The solution field B with three-phase current excitation in the Euclidean chart f is shown in figures 7.8 and 7.9.

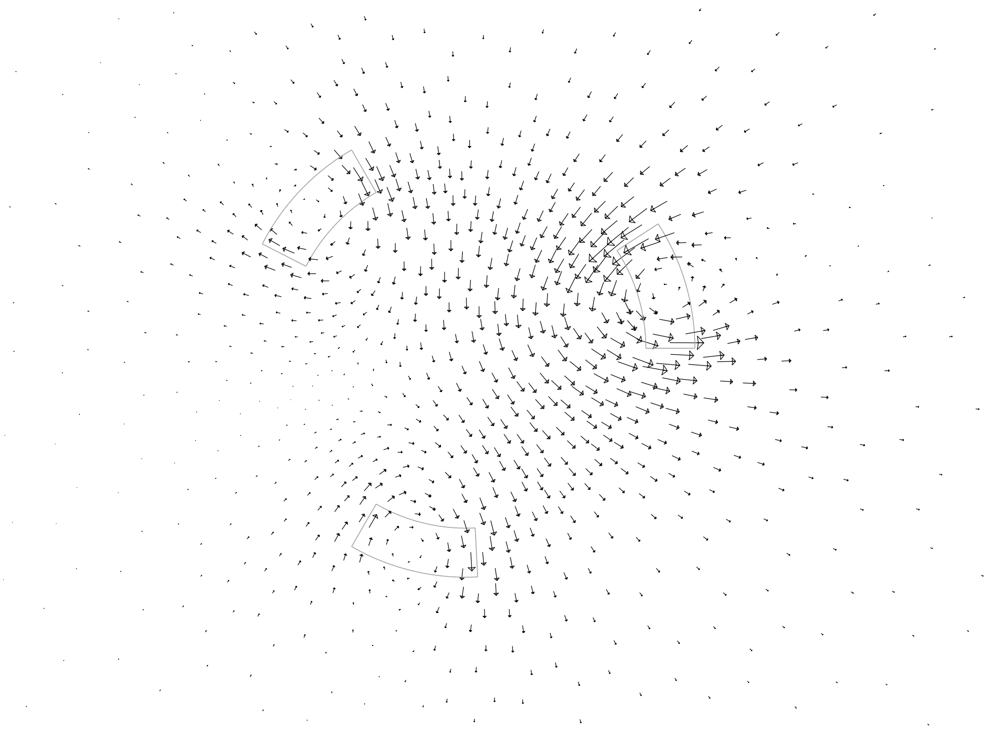

Figure 7.8: Magnetic field B in the xy-plane represented as a proxy vector field.

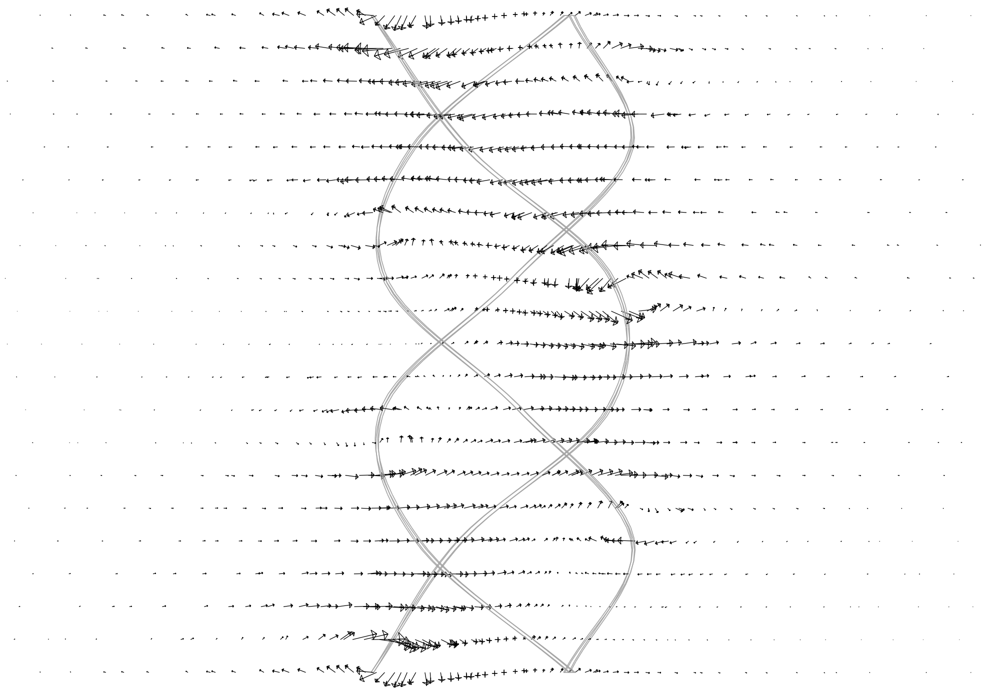

Figure 7.9: Magnetic field B in the yz-plane represented as a proxy vector field.

Chapter 8

Conclusion

This study focused on mathematical structures of differential geometry and investigated their exploitation in the formulation of electromagnetic BVPs. Particularly the usefulness of the structures over the traditional approach based on classical vector analysis was demonstrated with variety of applications.

Differential geometry allows coordinate-free formulation of electromagnetic BVPs, an approach also independent of the choice of metric. Furthermore, the structures are generic to all dimensions. Thus these *structures allow clear separation of coordinates, metric, and dimension from the aspects of electromagnetic BVPs that do not depend on them.* This is in contrast to the structures offered by classical vector analysis: the metric is embedded in most of them, and they are built initially in three-dimensional domains. Thus they must also be adapted to other dimensions.

Mathematical structures of differential geometry provide a *unified setting to formulate general electromagnetic BVPs*, including static and wave-propagation problems (section 4.4). This implies that common aspects of general BVPs can be analyzed in one setting without always having to analyze, as usual, different types of problems separately.

The structures enable formulation of an electromagnetic BVP that is invariant under diffeomorphisms. The formulation is thus generally covariant, which makes it a generalization of covariance under general change of coordinates. Particularly, diffeomorphism-invariance defines naturally an *equivalence of BVPs under diffeomorphisms* (section 4.5). In numerical modeling, this can be exploited in multiple ways: first, the change of coordinates procedure is simple and mechanical. Second, *many traditional and apparently diverse methods and approaches can be explained in a unified manner using the equivalence of BVPs* (chapter 6). These include, e.g., open boundary problems and invisibility cloaking. Furthermore, the equivalence of BVPs proposes a new approach to solve parametric models that include shape optimizations and movement: if the domains corresponding to different parameter values are diffeomorphic, it is possible to formulate all problems with the same chart and thus use the same mesh for all problems.

Because in vector analysis most structures are laid down over a metric, every coordinate system must be treated separately, which makes, e.g., application of general change of coordinates quite challenging. In addition, traditional methods and approaches look apparently different. On the other hand, as a major result of this investigation, we have shown that in formulations of electromagnetic BVPs, a *metric of the space is needed only in the initial identification of BVPs*: together with distance measurements the metric is a tool to establish a connection between model and observations (section 4.7). Furthermore, it does not matter what metric is used in initial identification. Particularly, we have shown that *the constitutive relations do not depend on the chosen metric; only their representations with Hodge-operators do* (section 4.7).

The observer structure for spacetime was introduced to decompose spacetime into space and time (section 4.1). The observer structure is coordinate- and metric-free and can be characterized as a field of local observers. It allows a $(3+1)$-*decomposition of Maxwell's equations and constitutive equations* (sections 4.2 and 4.3). Furthermore, the observer structure was also needed in the theory of dimensional reduction.

Because the tools of differential geometry are generic to all dimensions, they are *natural tools to formulate electromagnetic BVPs of any dimension*: 2d modeling is commonplace in electromagnetics and the tools help bypass some of the oddities of vector analysis. The key to solve problems in lower dimensions is symmetry, though it is seldom discussed in classical treatment of the subject. That is why the concept of symmetry was given a special treatment in this thesis. Particularly, we have shown that the *symmetry of BVPs is independent of coordinates, metric, and dimension* (chapter 5). Another major result of the thesis is the *symmetry-based theory of dimensional reduction of electromagnetic BVPs*, also fully independent of coordinates, metric, and dimension (chapter 5). The theory provides sufficient conditions for a BVP to be solved as a lower-dimensional BVP, and the conditions are stated in the form of a *dimensional reduction theorem* in section 5.9. Furthermore, the theory shows how to formulate lower-dimensional BVPs (sections 5.6 and 5.8).

The generality of the proposed symmetry-based theory of dimensional reduction (not restricted by coordinates, metrics, or dimension) *enables us to find out and apply new symmetries that are not widely known.* As an example of a symmetry that allows dimensional reduction but is not widely known, we gave helicoidal geometries: a complete numerical example of infinitely long helicoidally twisted current wires as a 2d-problem (section 7.2).

Yet significant open questions remain. Especially:

Open Question 1. (page 61) To what extent can the modern mathematical approach of observer structures be exploited in practical electromagnetic design? In other words, are there any other practical applications of observer structures than spacetime splitting and dimensional reduction?

Open Question 2. (page 83) How could a set of all BVPs and the equivalence of BVPs be rigorously defined? The answer would specifically help in designing software systems to solve BVPs from a very wide class of problems.

Open Question 3. (page 99) What benefits can be achieved if the definition of (G, h)-invariance is extended to more general mappings h than just a real- or complex-valued h? This answer could admit totally new kind of symmetries to be used in solutions of BVPs.

Open Question 4. (page 107) How could the theory of dimensional reduction be constructed to include nonlinear constitutive equations? Particularly, how to prove in nonlinear cases that solution fields are G-invariant if the sources, boundary values, constitutive equations, and cohomology conditions are G-invariant? The answer would allow dimensional reduction to be applied to many practical cases.

Open Question 5. (page 122) Are singular orbits always at the boundary of the orbit space? This is now an axiom of the theory, but can it be shown to be a theorem?

Open Question 6. (page 123) How could the theory of dimensional reduction be extended to include singular sources? The answer would add some classical examples to the theory.

Open Question 7. (page 129) In G-invariant cases, the lower-dimensional BVP can decompose into two separate lower-dimensional BVPs. Is it always possible to achieve separation with a suitable choice of observer structure, and how can we determine choices that enable separation? The answer would help simplify the solution process for some BVPs.

Open Question 8. (page 154) Is it possible rigorously to relax the homeomorphism requirement to define the equivalence of BVPs? Particularly, is it possible to change by compactification asymptotic conditions on a noncompact manifold to boundary values on a compact manifold? The answer would help establish a rigorous procedure to deal with some noncompact BVPs.

Open Question 9. (page 159) Is it possible rigorously to relax the diffeomorphism requirement to define the equivalence of BVPs? Specifically, when exactly is it possible to use piecewise diffeomorphism to define the equivalence? The answer would facilitate the application of the equivalence of BVPs in numerical solution methods.

The thesis demonstrated in multiple ways that the mathematical machinery introduced is a suitable and more flexible alternative to the traditional approach. However, only brief mention was made of the technical details of exploiting the machinery in the existing solver software systems. Particularly, no reference was made to the software science aspects of constructing a solver software system based on the introduced mathematical structures. However, this would be an extensive topic itself meriting separate study.

Appendix A

Bibliography classification

This appendix classifies some of the titles in the Bibliography to categories covering important subjects of the thesis.

Mathematics books of central subjects are:

- General topology: [9] [11] [19] [27] [30].
- Manifolds: [5] [12] [34] [62].
- Differential forms and analysis on manifolds: [2] [5] [14] [34] [62].
- General algebra and linear algebra: [19] [25] [32] [35] [40] [42].
- Real analysis: [1] [18] [61].
- Differential geometry: [5] [37].
- Lie groups: [5] [48] [62].
- Symmetry and group actions: [5] [12] [48].

Applied mathematics books and articles by subject are:

- Mathematical physics and differential geometry: [13] [17] [26] [46] [58].
- Observer structures: [15] [26] [39] [56].
- Symmetry and group actions: [6] [7] [13] [53].

Books and articles about electromagnetism are:

- Electromagnetic theory: [13] [26] [51].
- Computational electromagnetism: [6] [7] [24] [28] [29] [33] [38].

Finally some recommendations. Very good introduction to manifolds and differential geometry is Boothby's book [5]. Burke's book [13] is excellent book about applied differential geometry in physics including electromagnetism. Also Frankel's book [17] is a good exposition of applications of differential geometry in physics. Basic mathematical structures of mathematical physics from category theoretical viewpoint with understandable and motivating fashion are presented by Geroch in his book [19]. Olver's book [48] is an extensive introduction to symmetry methods in differential equations. However, his approach to dimensional reduction is not based on differential forms and the exterior derivative as in chapter 5 of this thesis.

Bibliography

[1] C. D. Aliprantis and O. Burkinshaw. *Principles of Real Analysis, Second Edition*. Academic Press Inc., San Diego, California, 1990.

[2] P. Bamberg, S. Sternberg. *A course in mathematics for students of physics*. Cambridge University Press, 1990.

[3] J. Berndt. Lie Group Actions on Manifolds. *http://euclid.ucc.ie/pages/staff/berndt/sophia.pdf*.

[4] P. Bettess. Finite element modelling of exterior electromagnetic problems. *IEEE Trans. Magn.*, vol. 24, No. 1, pp. 238-243, 1988.

[5] W. M. Boothby. *An introduction to differentiable manifolds and Riemannian geometry, revised second edition*. Academic Press, London, San Diego, 2003.

[6] A. Bossavit. Boundary value problems with symmetry, and their approximation by finite elements. *SIAM J. Appl. Math.*, Vol. 53, No. 5, pp. 1352-80, 1993.

[7] A. Bossavit. *Computational electromagnetism*. Academic Press, San Diego California, Chestnut Hill, Massachusetts, 1998.

[8] A. Bossavit. On the notion of anisotropy of constitutive laws: Some implications of the 'Hodge implies metric' result. *COMPEL*, Vol. 20, No. 1, pp. 233–239, 2001.

[9] N. Bourbaki. *Elements of Mathematics. General Topology, Part 1.* Addison-Wesley Publishing Company, Reading, Massachusetts 1966.

[10] G. E. Bredon. Some Theorems on Transformation Groups. *The Annals of Mathematics*, 2nd Ser., Vol. 67, No. 1, pp. 104-118, 1958.

[11] G. E. Bredon. *Topology and Geometry*. Springer, New York, 1993.

[12] F. Brickell and R. S. Clark. *Differentiable Manifolds: An Introduction*. Van Nostrand Reinhold Company ltd, London, 1970.

[13] W. L. Burke. *Applied differential geometry*. Cambridge University Press, New York, 1985.

[14] H. Cartan. *Differential forms*. Dover Publications, Inc., New York, 2006.

[15] M. Fecko. On 3+1 decompositions with respect to an observer field via differential forms. *J. Math. Phys.* Vol. 38, pp. 4542-4560, 1997.

[16] H. Flanders. *Differential Forms with Applications to the Physical Sciences*. Academic Press, New York, 1967.

[17] T. Frankel. *The Geometry of Physics: An Introduction*. Cambridge University Press, 1997.

[18] R. F. Gariepy and W. P. Ziemer. *Modern Real Analysis*. PWS Publishing Company, Boston, Massachusetts, 1995.

[19] R. Geroch. *Mathematical Physics*. The University of Chicago Press, Chicago, London, 1985.

[20] P. Dular and C. Geuzaine. GetDP: a General Environment for the Treatment of Discrete Problems. *http://www.geuz.org/getdp/*.

[21] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *http://www.geuz.org/gmsh/*.

[22] A. Greenleaf, M. Lassas, and G. Uhlmann. Anisotropic conductivities that cannot detected in EIT. *Physiological Measurement* (special issue on Impedance Tomography), Vol. 24, No. 2, 2003, pp. 413-420.

[23] A. Greenleaf, Y. Kurylev, M. Lassas, and G. Uhlmann. Invisibility and Inverse Problems. *Bulletin of the American Mathematical Society*, 46 (2009), 55-97.

[24] P. W. Gross and P. R. Kotiuga. *Electromagnetic Theory and Computation: A Topological Approach*. Cambridge University Press, 2004.

[25] P. R. Halmos. *Finite-Dimensional Vector Spaces*. Springer-Verlag New York Inc., 1987.

[26] F. Hehl and Y. Obukhov. *Foundations of Classical Electrodynamics: Charge, Flux, and Metric*. Birkhäuser, 2003.

[27] M. Henle. *A Combinatorial Introduction to Topology.* Dover Publications, Inc., New York, 1979.

[28] F. Henrotte, B. Meys, H. Hedia, P. Dular, and W. Legros. Finite element modelling with transformation techniques. *IEEE Trans. Magn.*, Vol. 35, No. 3, pp. 1434–1437, 1999.

[29] R. Hiptmair and F. Teixeira (ed.). Discrete Hodge-operators: An algebraic perspective. *Geometric Methods for Computational Electromagnetics*, EMW Publishing, 2001, 32, 247-269.

[30] J. G. Hocking and G. S. Young. *Topology.* Dover Publication Inc., New York, 1988.

[31] C. Houck, J. Joines, and M. Kay. *A Genetic Algorithm for Function Optimization: A Matlab Implementation.* North Carolina State University, Raleigh, NC, technical report NCSU-IE-TR-95-09, 1995.

[32] J. H. Hubbard and B. B. Hubbard. *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach, Second Edition.* Prentice-Hall, Inc. Upper Saddle River, New Jersey, 1999.

[33] J. F. Imhoff, G. Meunier, and J. C. Sabonnadière. Finite element modeling of open boundary problems. *IEEE Trans. Magn.*, Vol. 26, No. 2, pp. 588-591, 1990.

[34] K. Jänich. *Vector analysis.* Springer, New York, 2000.

[35] P. J. Kahn. *Introduction to linear algebra.* Harper & Row, New York, Evanston, London, 1967.

[36] D. Kincaid and W. Cheney. *Numerical Analysis, Second Edition.* Brooks/Cole Publishing Company. Pacific Grove, California, 1996.

[37] S. Kobayashi and K. Nomizu. *Foundations of differential geometry, volume 1.* Interscience Publishers, John Wiley and Sons, 1963.

[38] P. R. Kotiuga. *Hodge Decompositions and Computational Electromagnetics.* PhD thesis, Department of Electrical Engineering, McGill University, 1984.

[39] S. Kurz, B. Auchmann, and B. Flemisch. Dimensional Reduction of Field Problems in a Differential-Form Framework. *COMPEL*, accepted for publication, 2009.

[40] S. Lang. *Algebra, Revised Third Edition.* Springer-Verlag New York, Inc., 2002.

[41] H. Blaine Lawson, Jr. Foliations. *Bull. Amer. Math. Soc.*, Vol. 80, No. 3, pp. 369-418, 1974.

[42] M. Marcus and H. Minc. *Introduction To Linear Algebra.* The Macmillan Company, New York, 1965.

[43] J. B. M. Melissen and J. Simkin. A new coordinate transform for the finite element solution of axisymmetric problems in magnetostatics. *IEEE Trans. Magn.*, Vol. 26, No. 6, pp. 391-394, 1990.

[44] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation.* W. H. Freeman and Company, New York, 1973.

[45] D. Montgomery. What is a Topological Group? *The American Mathematical Monthly*, 52, 6, pp. 302-307, 1945.

[46] C. Nash and S. Sen. *Topology and Geometry for Physicists.* Academic Press, San Diego, London, 1983.

[47] E. Noether. Invariante Variationsprobleme. Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse, 235-257 (1918). English translation “Invariant Variation Problems” available at arxiv.org.

[48] P. J. Olver. *Applications of Lie Groups to Differential Equations, Second Edition.* Springer-Verlag New York inc., 1993.

[49] B. O’Neill. *Semi-Riemannian Geometry With Applications to Relativity.* Academic Press, San Diego, London, 1983.

[50] J. B. Pendry, D. Schurig, and D. R. Smith. Controlling Electromagnetic Fields. *Science*, Vol. 312 (2006), pp. 1780-1782.

[51] E. J. Post. *Formal Structure of Electromagnetics: General Covariance and Electromagnetics.* North Holland, Amsterdam 1962 and Dover, Mineola, New York 1997.

[52] P. Raumonen, S. Suuriniemi, and L. Kettunen. Applications of manifolds: mesh generation. *IET Sci. Meas. Technol.*, Vol. 2, No. 5, pp. 286-294.

[53] P. Raumonen, S. Suuriniemi, T. Tarhasaari, and L. Kettunen. Dimensional Reduction in Electromagnetic Boundary Value Problems. *IEEE Trans. Magn.*, Vol. 44, No. 6, pp. 1146-1149, 2008.

[54] P. Raumonen, S. Suuriniemi, T. Tarhasaari, and L. Kettunen. Manifold and metric in numerical solution of the quasi-static electromagnetic boundary value problems. *www.arxiv.org*, arXiv:0710.1747v1 [math-ph].

[55] P. Raumonen, S. Suuriniemi, and L. Kettunen. Parametric Models in Quasi-static Electromagnetics. *IEEE Trans. Magn.*, Vol. 45, No. 3, 2009.

[56] G. F. Rubilar. Linear pre-metric electrodynamics and deduction of the light cone. *Annalen der Physics*, Vol. 11, No. 10, pp. 717-782, 2002.

[57] E. Schrödinger. *Space-Time Structure*. Cambridge University Press, Cambridge, 1950.

[58] B. Schutz. *Geometrical methods of mathematical physics*. Cambridge University Press, Cambridge, New York, Melbourne, 1980.

[59] P. P. Silvester, D. A. Lowther, C. J. Carpenter, E. A. Wyatt. Exterior finite elements for 2-dimensional field problems with open boundaries. *PROC. IEE*, Vol. 124, No. 12, pp. 1267-1270, 1977.

[60] P. P. Silvester, M. S. Hsieh. Finite element solution of two dimensional exterior field problems. *PROC. IEE*, Vol. 118, pp. 1743-1747, Dec. 1971.

[61] K. R. Stromberg. *An introduction to classical real analysis*. Wadsworth & Brooke/Cole Advanced Books & Software, Pacific Grove, California, 1981.

[62] F. W. Warner. *Foundations of differentiable manifolds and Lie groups*. Springer, New York, 1983.

[63] H. Weyl. *Space Time Matter*. Dover Publications, Inc., New York, fourth ed., 1952.

[64] H. Whitney. Differentiable Manifolds. *The Annals of Mathematics*, Second Series, Vol. 37, No. 3, pp. 645-680, Jul. 1936.

[65] H. Whitney. *Geometric integration theory*. Princeton University Press, Princeton, New Jersey, 1957.

[66] K. Yosida. *Functional Analysis, Sixth Edition*. Springer-Verlag, Berlin, Heidelberg, New York, 1980.

[67] O. C. Zienckiwicz and D. W. Kelly. The coupling of finite element and boundary solution procedures. *Int. J. Numer. Meth. Engng*, No. 11, pp. 355-375, 1977.

Index

(G,h)-invariance
  boundary values, 100
  BVP, 105
  cohomology conditions, 105
  constitutive equations, 104
  fields, 98

Affine space, 20
Asymptotic condition, 152
Atlas, 26

Basis
  ordered, 20
  orthonormal, 51
  topology, 16
  vector space, 19
Bilinear form, 49
Boundary of manifold, 33

Cartan's formula, 46
Chart, 22
  compatible, 25
  orientation-preserving, 31
Coframe, 55
Cohomology class, 76
Continuous mapping, 16
Contraction, 43
Coordinate frame, 55
Cotangent bundle, 38
Cotangent space, 37
Covariance, 12
Covector, 37
Cross-section, 108

Diffeomorphism, 27
  orientation-preserving, 31
Differentiable mapping, 27
Differentiable structure, 25
Differential, 30
Differential form, 39
Dimensional reduction theorem, 148

Embedding, 32
Euclidean space, 21
Extension, 44
Exterior derivative, 42

Foliation, 35
  leaves of foliation, 35

G-invariance, *see* (G,h)-invariance
G-reduced domain, 109
Generalized covariance, 12
Generalized Stokes's theorem, 48
Geometric components, 63
Geometric decomposition
  differential forms, 63
  exterior derivative, 65
Group, 6
Group action, 9
  effective, 10
  free, 10
  transitive, 10
Group isomorphism, 6

Hausdorff space, 17
Hodge-like operator, 90
Hodge-operator, 53
Homeomorphism, 16
Homomorphism, 6

Horizontal
- component, 63
- differential form, 63
- exterior derivative, 65
- submanifold, 63
- vector, 63

Index
- Symmetric bilinear form, 49

Inner product, 49
Integral curve, 45
Isometry, 18, 51
Isomorphism, 6

Lie derivative, 46
Lie group, 34
Lie group homomorphism, 34
Lie group isomorphism, 34
Linear isomorphism, 19
Linear mapping, 19

Manifold
- differentiable, 26
- Lorentz, 50
- oriented, 31
- Riemannian, 50
- semi-Riemannian, 50
- topological, 22
- with-boundary, 33

Metric space, 17
Metric tensor, 49
Metric topology, 18
Metrical isomorphism, 18

Neighborhood, 15

Observer structure, 61
- holonomic, 61
- nonholonomic, 61

One-form, 38
One-parameter group of transformations, 45
Orbit, 9
Orbit space, 108
Orientation, 20
Oriented
- manifold, 31
- vector space, 20

Proxy vector, 51
Pullback
- atlas, 27
- differential form, 40
- metric tensor, 50
- observer, 78

Pushforward, 30

Singular orbit, 98
Singular point, 98
Spatial, *see* Horizontal
Standard parameterization, 24
Structure-preserving mapping, 6
Submanifold, 32
- embedded, 32
- regular, 32

Subspace topology, 16

Tangent bundle, 36
Tangent space, 29
Tangent vector, 29
Time-harmonic invariance, 102
Topological space, 15
- compact, 16
- connected, 15

Trace, 40
Transition map, 25

Vector field, 30
- smooth, 37

Vector space, 19
Volume form, 52

Wedge product, 41