FOUNDATIONS OF CLASSICAL ELECTRODYNAMICS
– Electric charge and magnetic flux conservation and the spacetime relation –

Friedrich W. Hehl & Yuri N. Obukhov

Outprint 15 April 2003

Contents

Preface

In this book we display the fundamental structure underlying classical electrodynamics, i.e., the phenomenological theory of electric and magnetic effects. The book can be used as a textbook for an advanced course in theoretical electrodynamics for physics and mathematics students and, perhaps, for some highly motivated electrical engineering students.

We expect from our readers that they know elementary electrodynamics in the conventional $(1+3)$-dimensional form including Maxwell's equations. Moreover, they should be familiar with linear algebra and elementary analysis, including vector analysis. Some knowledge of differential geometry would help.

Our approach rests on the metric-free integral formulation of the conservation laws of electrodynamics in the tradition of F. Kottler (1922), É. Cartan (1923), and D. van Dantzig (1934), and we stress, in particular, the axiomatic point of view. In this manner we are led to an understanding of why the Maxwell equations have their specific form. We hope that our book can be seen in the classical tradition of the book by E. J. Post (1962) on the *Formal Structure of Electromagnetics* and of the chapter "Charge and Magnetic Flux" of the encyclopedia article on classical field theories by C. Truesdell and R. A. Toupin (1960), including R. A. Toupin's Bressanone lectures (1965); for the exact references see the end of the introduction on page 9.

The manner in which electrodynamics is conventionally presented in physics courses à la R. Feynman (1962), J. D. Jackson (1999), and L. D. Landau & E. M. Lifshitz (1962) is distinctly different, since it is based on a flat spacetime manifold, i.e., on the (rigid) Poincaré group, and on H. A. Lorentz's approach (1916) to Maxwell's theory by means of his theory of electrons. We believe that the approach of this book is appropriate and, in our opinion, even superior for a good understanding of the structure of electrodynamics as a *classical field theory*. In particular, if gravity cannot be neglected, our framework allows for a smooth and trivial transition to the curved (and contorted) spacetime of general relativistic field theories. This is by no means a minor merit when one has to treat magnetic fields of the order of 10^9 tesla in the neighborhood of a neutron star where spacetime is appreciably curved.

Mathematically, integrands in the conservation laws are represented by exterior differential forms. Therefore exterior calculus is the appropriate language in which electrodynamics should be spelled out. Accordingly, we exclusively use this formalism (even in our computer algebra programs which we introduce in Sec. A.1.12). In Part A, and later in Part C, we try to motivate and to supply the necessary mathematical framework. Readers who are familiar with this formalism may want to skip these parts. They could start right away with the physics in Part B and then turn to Part D and Part E.

In Part B four axioms of classical electrodynamics are formulated and the consequences derived. This general framework has to be completed by a specific *electromagnetic spacetime relation* as a fifth axiom. This is done in Part D. The Maxwell–Lorentz theory is then recovered under specific conditions. In Part E, we apply electrodynamics to moving continua, inter alia, which requires a sixth axiom on the formulation of electrodynamics inside matter.

This book grew out of a scientific collaboration with the late *Dermott McCrea* (University College Dublin). Mainly in Part A and Part C, Dermott's handwriting can still be seen in numerous places. There are also some contributions to "our" mathematics from *Wojtek Kopczyński* (Warsaw University). At Cologne University in the summer term of 1991, *Martin Zirnbauer* started to teach the theoretical electrodynamics course by using the calculus of exterior differential forms, and he wrote up successively improved notes to his course. One of the authors (FWH) also taught this course three times, partly based on Zirnbauer's notes. This influenced our way of presenting electrodynamics (and, we believe, also his way). We are very grateful to him for many discussions.

There are many colleagues and friends who helped us in critically reading parts of our book and who made suggestions for improvement or who communicated to us their own ideas on electrodynamics. We are very grateful to all of them: Carl Brans (New Orleans), Jeff Flowers (Teddington), David Hartley (Adelaide), Christian Heinicke (Cologne), Yakov Itin (Jerusalem), Martin Janssen (Cologne), Gerry Kaiser (Glen Allen, Virginia), R. M. Kiehn (formerly Houston), Attay Kovetz (Tel Aviv), Claus Lämmerzahl (Konstanz/Bremen), Bahram Mashhoon (Columbia, Missouri), Eckehard Mielke (Mexico City), Wei-Tou Ni (Hsin-chu), E. Jan Post (Los Angeles), Dirk Pützfeld (Cologne), Guillermo Rubilar (Cologne/Concepción), Yasha Shnir (Cologne), Andrzej Trautman (Warsaw), Arkady Tseytlin (Columbus, Ohio), Wolfgang Weller (Leipzig), and others.

We are particularly grateful to the two reviewers of our book, to Jim Nester (Chung-li) and to an anonym for their numerous good suggestions and for their painstaking work.

We are very obliged to Uwe Essmann (Stuttgart) and to Gary Glatzmaier (Santa Cruz, California) for providing beautiful and instructive images. We are equally grateful to Peter Scherer (Cologne) for his permission to reprint his three comics on computer algebra.

The collaboration with the Birkhäuser people, with Gerry Kaiser and Ann Kostant, was effective and fruitful. We would like to thank Debra Daugherty (Boston) for improving our English.

Please convey critical remarks to our approach or the discovery of mistakes by surface or electronic mail (hehl@thp.uni-koeln.de, yo@thp.uni-koeln.de) or by fax +49-221-470-5159.

This project has been supported by the Alexander von Humboldt Foundation (Bonn), the German Academic Exchange Service DAAD, and the Volkswagen Foundation (Hanover). We are very grateful for the unbureaucratic help of these institutions.

Cologne and Moscow, April 2003

Friedrich W. Hehl & Yuri N. Obukhov

Introduction

Five plus one axioms

In this book we display the structure underlying classical electrodynamics. For this purpose we formulate six axioms: conservation of electric *charge* (first axiom), existence of the Lorentz *force* (second axiom), conservation of magnetic *flux* (third axiom), local *energy-momentum* distribution (fourth axiom), existence of an electromagnetic *spacetime* relation (fifth axiom), and finally, the splitting of the electric current into material and external pieces (sixth axiom).

The axioms expressing the conservation of electric charge and magnetic flux are formulated as integral laws, whereas the axiom for the Lorentz force is represented by a *local* expression basically defining the electromagnetic field strength $F = (E, B)$ as force per unit charge and thereby linking electrodynamics to mechanics; here E is the electric and B the magnetic field strength. Also the energy-momentum distribution is specified as a local law. The fifth axiom, the Maxwell–Lorentz spacetime relation is not as unquestionable as the first four axioms and extensions encompassing dilaton, skewon, and axion fields are discussed and nonlocal and nonlinear alternatives mentioned.

We want to stress the fundamental nature of the *first* axiom. Electric charge conservation is experimentally firmly established. It is valid for single elementary particle processes (like β-decay, $n \rightarrow p^+ + e^- + \overline{\nu}$, for instance, with n as neutron, p as proton, e as electron, and $\overline{\nu}$ as electron antineutrino). In other words, it is a *microscopic* law valid without any known exception.

Accordingly, it is basic to electrodynamics to assume a new type of entity called *electric charge*, carrying a positive or negative sign, with its own physical dimension, independent of the classical fundamental variables mass, length, and time. Furthermore, electric charge is conserved. In an age in which single electrons and (anti)protons are *counted* and caught in traps, this law is so deeply ingrained in our thinking that its explicit formulation as a fundamental law (and not only as a consequence of Maxwell's equations) is often forgotten. We show that this *first axiom* yields the inhomogeneous Maxwell equation together with a definition of the electromagnetic excitation $H = (\mathcal{H}, \mathcal{D})$; here $\mathcal{H}$ is the magnetic excitation ("magnetic field") and $\mathcal{D}$ the electric excitation ("electric displacement"). The excitation H is a microscopic field of an analogous quality as the field strength F. There exist operational definitions of the excitations $\mathcal{D}$ and $\mathcal{H}$ (via Maxwellian double plates or a compensating superconducting wire, respectively).

The *second* axiom for the Lorentz force, as mentioned above, leads to the notion of the field strength and is thereby exhausted. Thus we need further axioms. The only conservation law that can be naturally formulated in terms of the field strength is the conservation of magnetic flux (lines). This *third* axiom has the homogeneous Maxwell equation — that is, Faraday's induction law and the vanishing divergence of the magnetic field strength — as a consequence.

Moreover, with the help of these first three axioms we are led, although not completely uniquely, to the electromagnetic *energy-momentum* current (fourth axiom), which subsumes the energy and momentum densities of the electromagnetic field and their corresponding fluxes, and to the *action* of the electromagnetic field. In this way, the basic structure of electrodynamics is set up, including the complete set of Maxwell's equations. To make this set of electrodynamic equations well determined, we still have to add the fifth axiom.

Magnetic monopoles are alien to the structure of the axiomatics we are using. In our axiomatic framework, a clear *asymmetry* is built in between electricity and magnetism in the sense of Oersted and Ampère wherein magnetic effects are created by moving electric charges. This asymmetry is characteristic for and intrinsic to Maxwell's theory. Therefore the conservation of magnetic *flux* and not that of magnetic charge is postulated as the third axiom.

The existence of a magnetic charge in violation of our third axiom would have far-reaching consequences: First of all, the electromagnetic potential A would not exist. Accordingly, in Hamiltonian mechanics, we would have to give up the coupling of a charged particle to the electromagnetic field via $\Pi = p - e\,A$. Moreover, the second axiom on the Lorentz force could be invalidated since one would have to supplement it with a term carrying the magnetic charge density. By implication, an extension of the fifth axiom on the energy-momentum current would be necessary.

In other words, if ever a *magnetic* monopole[1] were found, our axiomatics would lose its coherence, its compactness, and its plausibility. Or, to formulate it more positively: Not long ago, He [22], Abbott et al. [1], and Kalbfleisch et al. [32] determined experimentally new improved limits for the nonexistence of (Abelian or Dirac) magnetic monopoles. This ever increasing accuracy in the exclusion of magnetic monopoles speaks in favor of the axiomatic approach in Part B.

Topological approach

> *Since the notion of metric is a complicated one, which requires measurements with clocks and scales, generally with* rigid *bodies, which themselves are systems of great complexity, it seems undesirable to take metric as fundamental, particularly for phenomena which are simpler and actually independent of it.*
>
> E. Whittaker (1953)

The distinctive feature of this type of axiomatic approach is that one only needs minimal assumptions about the structure of the spacetime in which these axioms are formulated. For the first four axioms, a 4-dimensional differentiable

[1] Our arguments refer only to Abelian gauge theory. In *non*-Abelian gauge theories the situation is different. There monopoles seem to be a must, at least if a Higgs field is present.

manifold is required that allows for a *foliation* into 3-dimensional hypersurfaces. Thus no *connection* and no *metric* are explicitly introduced in Parts A and B. The Poincaré and the Lorentz groups are totally ignored. Nevertheless, we recover Maxwell's equations already in Part B. This shows that electrodynamics is not as closely related to special relativity theory as is usually supposed.

This minimalistic *topological type* of approach may appear contrived at first look. We should recognize, however, that the metric of spacetime in the framework of general relativity theory represents the gravitational potential and, similarly, the connection of spacetime (in the viable Einstein–Cartan theory of gravity, for example) is intimately linked to gravitational properties of matter. We know that we really live in a curved and, perhaps, contorted spacetime. Consequently our desire should be to formulate the foundations of electrodynamics such that the metric and the connection don't interfere or interfere only in the least possible way. Since we know that the gravitational field permeates all the laboratories in which we make experiments with electricity, we should take care that this ever present field doesn't enter the formulation of the first principles of electrodynamics. In other words, a *clear separation* between pure electrodynamic effects and gravitational effects is desirable and can indeed be achieved by means of the axiomatic approach to be presented in Part B. Eventually, in the spacetime relation (see Part D), the metric does enter.

The power of the topological approach is also clearly indicated by its ability to describe the *phenomenology* (at low frequencies and large distances) of the quantum Hall effect successfully (not, however, its quantization). Insofar as the macroscopic aspects of the quantum Hall effect can be approximately understood in terms of a 2-dimensional electron gas, we can start with $(1 + 2)$-dimensional electrodynamics, the formulation of which is straightforward in our axiomatics. It is then a mere finger exercise to show that in this specific case of $1 + 2$ dimensions there exists a *linear* constitutive law that doesn't require a metric. As a consequence the action is metric-free too. Thus the formulation of the quantum Hall effect by means of a *topological* (Chern–Simons) Lagrangian is imminent in our way of looking at electrodynamics.

Electromagnetic spacetime relation as a fifth axiom

Let us now turn to that domain where the metric does enter the 4-dimensional electrodynamical formalism. When the Maxwellian structure, including the Lorentz force and the action, is set up, it does not represent a concrete physical theory yet. What is missing is the electromagnetic spacetime relation linking the excitation to the field strength, i.e., $\mathcal{D} = \mathcal{D}(E, B)$, $\mathcal{H} = \mathcal{H}(B, E)$, or written 4-dimensionally, $H = H(F)$.

Trying the simplest approach, we assume *locality* and *linearity* between the excitation H and the field strength F, that is, $H = \boldsymbol{\kappa}(F)$ with the linear operator $\boldsymbol{\kappa}$. Together with two more "technical" assumptions, namely that $H = \boldsymbol{\kappa}(F)$ is electric/magnetic *reciprocal* and $\boldsymbol{\kappa}$ *symmetric* (these properties will be discussed in detail in Part D), we are able to *derive* the metric of spacetime from

$H = \boldsymbol{\kappa}(F)$ up to an arbitrary (conformal) factor. Accordingly, the light cone structure of spacetime is a consequence of a linear electromagnetic spacetime relation with the additional properties of reciprocity and symmetry. In this sense, the light cones are derived from electrodynamics.

Electrodynamics doesn't live in a preformed rigid Minkowski spacetime. Rather it has an arbitrary $(1 + 3)$-dimensional spacetime manifold as its habitat which, as soon as a linear spacetime relation with reciprocity and symmetry is supplied, is equipped with local light cones everywhere. With the light cones it is possible to define the Hodge star operator * that maps p-forms to $(4-p)$-forms and, in particular, H to F according to $H \sim {}^\star F$.

Thus, in the end, that property of spacetime that describes its local "constitutive" structure, namely the metric, enters the formalism of electrodynamics and makes it into a complete theory. One merit of our approach is that it doesn't matter whether it is the rigid, i.e., flat, *Minkowski* metric of special relativity or the "flexible" *Riemannian* metric field of general relativity that changes from point to point according to Einstein's field equation. In this way, the traditional discussion of how to translate electrodynamics from special to general relativity loses its sense: The Maxwell equations remain the same, i.e., the exterior derivatives (the "commas" in coordinate language) are kept and are *not* substituted by something "covariant" (the "semicolons"), and the spacetime relation $H = \lambda_0 \, {}^\star F$ looks the same (λ_0 is a suitable factor). However, the Hodge star "feels" the difference in referring either to a constant or to a spacetime-dependent metric, respectively (see [51, 23]).

Our formalism can accommodate generalizations of classical electrodynamics, including those violating Lorentz invariance, simply by suitably modifying the fifth axiom while keeping the first four axioms as indispensable. Then a scalar *dilaton* field $\lambda(x)$, a pseudoscalar *axion* field $\alpha(x)$, and/or a tensorial traceless *skewon* field $\not{S}_i{}^j$ can come up in a most straightforward way. Also the Heisenberg–Euler and the Born–Infeld electrodynamics are prime examples of such possible modifications. In the latter cases, the spacetime relation becomes effectively *non*linear, but it still remains a local expression.

Electrodynamics in matter and the sixth axiom

Eventually, we have to face the problem of formulating electrodynamics inside matter. We codify our corresponding approach in the sixth axiom. The total electric current, entering as the source in the inhomogeneous Maxwell equation, is split into bound charge and free charge. In this way, following Truesdell & Toupin [63] (see also the textbook of Kovetz [34]), we can develop a consistent theory of electrodynamics in matter. For simple cases, we can amend the axioms by a *linear constitutive law*. Since in our approach $(\mathcal{H}, \mathcal{D})$ are microscopic fields, like (E, B), we believe that the conventional theory of electrodynamics inside matter needs to be redesigned.

In order to demonstrate the effectiveness of our formalism, we apply it to the electrodynamics of moving matter, thereby returning to the post-Maxwellian era of the 1880s when a relativistic version of Maxwell's theory had gained momentum. In this context, we discuss and analyse the experiments of Walker & Walker and James and those of Röntgen–Eichenwald and Wilson & Wilson.

List of axioms

1. Conservation of electric charge: (B.1.17).
2. Lorentz force density: (B.2.8).
3. Conservation of magnetic flux: (B.3.1).
4. Localization of energy-momentum: (B.5.7).
5. Maxwell–Lorentz spacetime relation: (D.6.13).
6. Splitting of the electric current in a conserved matter piece and an external piece: (E.3.1) and (E.3.2).

A reminder: Electrodynamics in 3-dimensional Euclidean vector calculus

Before we start to develop electrodynamics in 4-dimensional spacetime in the framework of the calculus of exterior differential forms, it may be useful to remind ourselves of electrodynamics in terms of conventional 3-dimensional Euclidean vector calculus. We begin with the laws obeyed by electric charge and current.

If $\vec{\mathcal{D}} = (\mathcal{D}_x, \mathcal{D}_y, \mathcal{D}_z)$ denotes the *electric excitation* field (historically "electric displacement") and ρ the electric charge density, then the integral version of the *Gauss law*, 'flux of $\mathcal{D}$ through any closed surface' equals 'net charge inside' reads

$$\int\limits_{\partial V} \vec{\mathcal{D}} \cdot d\vec{S} = \int\limits_{V} \rho \, dV \,, \tag{I.1}$$

with $d\vec{S}$ as area and dV as volume element. The *Oersted–Ampère law* with the *magnetic excitation* field $\vec{\mathcal{H}} = (\mathcal{H}_x, \mathcal{H}_y, \mathcal{H}_z)$ (historically "magnetic field") and the electric current density $\vec{j} = (j_x, j_y, j_z)$ is a bit more involved because of the presence of the Maxwellian electric excitation current: The 'circulation of $\vec{\mathcal{H}}$ around any closed contour' equals '$\frac{d}{dt}\Big($flux of $\vec{\mathcal{D}}$ through surface spanned by contour$\Big)$' plus 'flux of $\vec{j}$ through surface' (t =time):

$$\oint\limits_{\partial S} \vec{\mathcal{H}} \cdot d\vec{r} = \frac{d}{dt}\left(\int\limits_{S} \vec{\mathcal{D}} \cdot d\vec{S}\right) + \int\limits_{S} \vec{j} \cdot d\vec{S} \,. \tag{I.2}$$

Here $d\vec{r}$ is the vectorial line element. Here the dot $\cdot$ always denotes the 3-dimensional metric-dependent scalar product, S denotes a 2-dimensional spatial surface, V a 3-dimensional spatial volume, and ∂S and ∂V the respective boundaries. Later we recognize that both, (I.1) and (I.2), can be derived from the *charge conservation law.*

The *homogeneous* Maxwell equations are formulated in terms of the electric field strength $\vec{E} = (E_x, E_y, E_z)$ and the magnetic field strength $\vec{B} = (B_x, B_y, B_z)$. They are defined operationally via the expression of the Lorentz force $\vec{F}$. An electrically charged particle with charge q and velocity $\vec{v}$ experiences the force

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})\,. \tag{I.3}$$

Here the cross $\times$ denotes the 3-dimensional vector product. Then *Faraday's induction law* in its integral version, namely 'circulation of $\vec{E}$ around any closed contour' equals 'minus $\frac{d}{dt}\Big($flux of $\vec{B}$ through surface spanned by contour$\Big)$' reads:

$$\oint\limits_{\partial S} \vec{E} \cdot d\vec{r} = -\frac{d}{dt}\left(\int\limits_{S} \vec{B} \cdot d\vec{S}\right)\,. \tag{I.4}$$

Note the minus sign on its right-hand side, which is chosen according to the Lenz rule (following from energy conservation). Finally, the 'flux of $\vec{B}$ through any closed surface' equals 'zero', that is,

$$\int\limits_{\partial V} \vec{B} \cdot d\vec{S} = 0\,. \tag{I.5}$$

The laws (I.4) and (I.5) are inherently related. Later we formulate the law of *magnetic flux conservation* and (I.4) and (I.5) just turn out to be consequences of it.

Applying the Gauss and the Stokes theorems, the integral form of the Maxwell equations (I.1), (I.2) and (I.4), (I.5) can be transformed into their differential versions:

$$\operatorname{div}\vec{\mathcal{D}} = \rho\,, \qquad \operatorname{curl}\vec{\mathcal{H}} - \dot{\vec{\mathcal{D}}} = \vec{j}\,, \tag{I.6}$$

$$\operatorname{div}\vec{B} = 0\,, \qquad \operatorname{curl}\vec{E} + \dot{\vec{B}} = 0\,. \tag{I.7}$$

Additionally, we have to specify the spacetime relations $\vec{\mathcal{D}} = \varepsilon_0\,\vec{E}$, $\vec{B} = \mu_0\,\vec{\mathcal{H}}$, and if matter is considered, the constitutive laws.

This formulation of electrodynamics by means of 3-dimensional Euclidean vector calculus represents only a preliminary version since the 3-dimensional metric enters the scalar and the vector products and, in particular, the differential operators div $\equiv \vec{\nabla}\cdot$ and curl $\equiv \vec{\nabla}\times$, with $\vec{\nabla}$ as the nabla operator. In the Gauss law (I.1) or $(\mathrm{I.6})_1$, for instance,[2] only *counting procedures* enter,

[2]The subscript 1 refers to the first equation in (I.6).

namely counting of elementary charges inside V (taking care of their sign, of course) and counting of flux lines piercing through a closed surface ∂V. No length or time measurements and thus no metric are involved in such processes, as described in more detail below. Since similar arguments apply also to (I.5) or $(\mathrm{I}.7)_1$, respectively, it should be possible to remove the metric from the Maxwell equations altogether.

On the literature

Basically not too much is new in our book. Probably Part D and Part E are the most original. Most of the material can be found somewhere in the literature. What we do claim, however, is some originality in the *completeness* and in the appropriate *arrangement* of the material, which is fundamental to the structure electrodynamics is based on. Moreover, we try to stress the *phenomena* underlying the axioms chosen and the *operational interpretation* of the quantities introduced. The *explicit derivation* in Part D of the metric of spacetime from pre-metric electrodynamics by means of linearity, reciprocity, and symmetry, although considered earlier mainly by Toupin [61], Schönberg [53], and Jadczyk [28], is new and rests on recent results of Fukui, Gross, Rubilar, and the authors [42, 24, 41, 21, 52]. Also the generalization encompassing the dilaton, the axion, and/or the skewon field opens a new perspective. In Part E the electrodynamics of moving bodies, including the discussions of some classic experiments, contains much new material.

Our *main sources* are the works of Post [46, 47, 48, 49, 50], of Truesdell & Toupin [63], and of Toupin [61]. Historically, the metric-free approach to electrodynamics, based on integral conservation laws, was pioneered by Kottler [33], É. Cartan [10], and van Dantzig [64]. The article of Einstein [16] and the books of Mie [39], Weyl [65], and Sommerfeld [58] should also be consulted on these matters (see as well the recent textbook of Kovetz [34]). A description of the corresponding historical development, with references to the original papers, can be found in Whittaker [66] and, up to about 1900, in the penetrating account of Darrigol [12]. The driving forces and the results of Maxwell in his research on electrodynamics are vividly presented in Everitt's [17] concise biography of Maxwell.

In our book, we consistently use exterior calculus,[3] including de Rham's *odd* (or twisted) differential forms. *Textbooks* on electrodynamics using exterior calculus are scarce. In English, we know only of Ingarden & Jamiołkowski [26], in German of Meetz & Engl [38] and Zirnbauer [67], and in Polish, of Jancewicz [31] (see also [30]). However, as a discipline of mathematical physics, corresponding presentations can be found in Bamberg & Sternberg [4], in Thirring

[3] Baylis [6] also advocates a geometric approach, using Clifford algebras (see also Jancewicz [29]). In such a framework, however, at least the way Baylis does it, the metric of 3-dimensional space is introduced right from the beginning. In this sense, Baylis' Clifford algebra approach is complementary to our metric-free electrodynamics.

[60], and as a short sketch, in Piron [44] (see also [5, 45]). Bamberg & Sternberg are particularly easy to follow and present electrodynamics in a very transparent way. That electrodynamics in the framework of exterior calculus is also in the scope of electrical engineers can be seen from Deschamps [14], Bossavit[8], and Baldomir & Hammond [3].

Presentations of exterior calculus, partly together with applications in physics and electrodynamics, were given amongst many others by Burke [9], Choquet-Bruhat et al. [11], Edelen [15], Flanders [19], Frankel [20], Parrott [43], and Ślebodziński [57]. For differential geometry we refer to the classics of de Rham [13] and Schouten [54, 55] and to Trautman [62].

What else influenced the writing of our book? The *axiomatics* of Bopp [7] is different but related to ours. In the more microphysical axiomatic attempt of Lämmerzahl et al. Maxwell's equations [35] (and the Dirac equation [2]) are deduced from direct experience with electromagnetic (and matter) waves, inter alia. The clear separation of *differential*, *affine*, and *metric* structures of spacetime is nowhere more pronounced than in Schrödinger's [56] *Space-time structure.* A further presentation of electrodynamics in this spirit, somewhat similar to that of Post, has been given by Stachel [59]. Our (1 + 3)-*decomposition* of spacetime is based on the paper by Mielke & Wallner [40]. More recently, Hirst [25] has shown, mainly based on experience with neutron scattering on magnetic structures in solids, that *magnetization* $\mathcal{M}$ is a *microscopic* quantity. This is in accord with our axiomatics which yields the magnetic excitation $\mathcal{H}$ as a microscopic quantity, quite analogously to the field strength B, whereas in conventional texts $\mathcal{M}$ is only defined as a macroscopic *average* over microscopically fluctuating magnetic fields. Clearly, $\mathcal{H}$ and also the electric excitation $\mathcal{D}$, i.e., the electromagnetic excitation $H = (\mathcal{H}, \mathcal{D})$ altogether, ought to be a microscopic field.

• Sections and subsections of the book that can be skipped at a first reading are marked by the symbol $^{\otimes}$.

References

[1] B. Abbott et al. (D0 Collaboration), *A search for heavy pointlike Dirac monopoles*, *Phys. Rev. Lett.* **81** (1998) 524-529.

[2] J. Audretsch and C. Lämmerzahl, *A new constructive axiomatic scheme for the geometry of space-time* In: *Semantical Aspects of Space-Time Geometry.* U. Majer, H.-J. Schmidt, eds. (BI Wissenschaftsverlag: Mannheim, 1994) pp. 21-39.

[3] D. Baldomir and P. Hammond, *Geometry and Electromagnetic Systems* (Clarendon Press: Oxford, 1996).

[4] P. Bamberg and S. Sternberg, *A Course in Mathematics for Students of Physics*, Vol. 2 (Cambridge University Press: Cambridge, 1990).

[5] A.O. Barut, D.J. Moore and C. Piron, *Space-time models from the electromagnetic field*, *Helv. Phys. Acta* **67** (1994) 392-404.

[6] W.E. Baylis, *Electrodynamics. A Modern Geometric Approach* (Birkhäuser: Boston, 1999).

[7] F. Bopp, *Prinzipien der Elektrodynamik*, *Z. Physik* **169** (1962) 45-52.

[8] A. Bossavit, *Differential Geometry for the Student of Numerical Methods in Electromagnetism*, 153 pages, file `DGSNME.pdf` (1991) (see `http://www.lgep.supelec.fr/mse/perso/ab/bossavit.html`).

[9] W.L. Burke, *Applied Differential Geometry* (Cambridge University Press: Cambridge, 1985).

[10] É. Cartan, *On Manifolds with an Affine Connection and the Theory of General Relativity*, English translation of the French original of 1923/24 (Bibliopolis: Napoli, 1986).

[11] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick , *Analysis, Manifolds and Physics*, revised ed. (North-Holland: Amsterdam, 1982).

[12] O. Darrigol, *Electrodynamics from Ampère to Einstein* (Oxford University Press: New York, 2000).

[13] G. de Rham, *Differentiable Manifolds:* Forms, Currents, Harmonic Forms. Transl. from the French original (Springer: Berlin, 1984).

[14] G.A. Deschamps, *Electromagnetics and differential forms*, *Proc. IEEE* **69** (1981) 676-696.

[15] D.G.B. Edelen, *Applied Exterior Calculus* (Wiley: New York, 1985).

[16] A. Einstein, *Eine neue formale Deutung der Maxwellschen Feldgleichungen der Elektrodynamik*, *Sitzungsber. Königl. Preuss. Akad. Wiss. Berlin* (1916) pp. 184-188; see also *The collected papers of Albert Einstein.* Vol.6, A.J. Kox et al., eds. (1996) pp. 263-269.

[17] C.W.F. Everitt, *James Clerk Maxwell. Physicist and Natural Philosopher* (Charles Sribner's Sons: New York, 1975).

[18] R.P. Feynman, R.B. Leighton, and M. Sands, *The Feynman Lectures on Physics,* Vol. 2: Mainly Electromagnetism and Matter (Addison-Wesley: Reading, Mass., 1964).

[19] H. Flanders, *Differential Forms* with Applications to the Physical Sciences. (Academic Press: New York, 1963 and Dover: New York, 1989).

[20] T. Frankel, *The Geometry of Physics* – An Introduction (Cambridge University Press: Cambridge, 1997).

[21] A. Gross and G.F. Rubilar, *On the derivation of the spacetime metric from linear electrodynamics*, *Phys. Lett.* **A285** (2001) 267-272.

[22] Y.D. He, *Search for a Dirac magnetic monopole in high energy nucleus-nucleus collisions*, *Phys. Rev. Lett. 79* (1997) 3134-3137.

[23] F.W. Hehl and Yu.N. Obukhov, How does the electromagnetic field couple to gravity, in particular to metric, nonmetricity, torsion, and curvature? In: *Gyros, Clocks, Interferometers...: Testing Relativistic Gravity in Space.* C. Lämmerzahl et al., eds. Lecture Notes in Physics Vol.562 (Springer: Berlin, 2001) pp. 479-504; see also Los Alamos Eprint Archive gr-qc/0001010.

[24] F.W. Hehl, Yu.N. Obukhov, and G.F. Rubilar, *Spacetime metric from linear electrodynamics II. Ann. Physik (Leipzig)* **9** (2000) Special issue, SI-71–SI-78.

[25] L.L. Hirst, *The microscopic magnetization: concept and application*, Rev. Mod. Phys. **69** (1997) 607-627.

[26] R. Ingarden and A. Jamiołkowski, *Classical Electrodynamics* (Elsevier: Amsterdam, 1985).

[27] J.D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley: New York, 1999).

[28] A.Z. Jadczyk, *Electromagnetic permeability of the vacuum and light-cone structure*, *Bull. Acad. Pol. Sci., Sér. sci. phys. et astr.* **27** (1979) 91-94.

[29] B. Jancewicz, *Multivectors and Clifford Algebra in Electrodynamics* (World Scientific: Singapore, 1989).

[30] B. Jancewicz, *A variable metric electrodynamics. The Coulomb and Biot-Savart laws in anisotropic media*, *Ann. Phys. (NY)* **245** (1996) 227-274.

[31] B. Jancewicz, *Wielkości skierowane w elektrodynamice* (in Polish). Directed Quantities in Electrodynamics. (University of Wrocław Press: Wrocław, 2000); an English version is under preparation.

[32] G.R. Kalbfleisch, K.A. Milton, M.G. Strauss, L. Gamberg, E.H. Smith, and W. Luo, *Improved experimental limits on the production of magnetic monopoles*, *Phys. Rev. Lett.* **85** (2000) 5292-5295.

[33] F. Kottler, *Maxwell'sche Gleichungen und Metrik*, Sitzungsber. Akad. Wien IIa **131** (1922) 119-146.

[34] A. Kovetz, *Electromagnetic Theory* (Oxford University Press: Oxford, 2000).

[35] C. Lämmerzahl and M.P. Haugan, *On the interpretation of Michelson-Morley experiments*, *Phys. Lett.* **A282** (2001) 223-229.

[36] L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields*, Vol.2 of *Course of Theoretical Physics*, transl. from the Russian (Pergamon: Oxford, 1962).

[37] H.A. Lorentz, *The Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat.* 2nd ed. (Teubner: Leipzig, 1916).

[38] K. Meetz and W.L. Engl, *Elektromagnetische Felder:* Mathematische und physikalische Grundlagen, Anwendungen in Physik und Technik (Springer: Berlin, 1980).

[39] G. Mie, *Lehrbuch der Elektrizität und des Magnetismus,* 2nd ed. (Enke: Stuttgart 1941).

[40] E.W. Mielke ad R.P. Wallner, *Mass and spin of double dual solutions in Poincaré gauge theory*, Nuovo Cimento **101** (1988) 607-623, erratum **B102** (1988) 555.

[41] Yu.N. Obukhov, T. Fukui, and G.F. Rubilar, *Wave propagation in linear electrodynamics*, *Phys. Rev.* **D62** (2000) 044050, 5 pages.

[42] Yu.N. Obukhov and F.W. Hehl, *Space-time metric from linear electrodynamics*, *Phys. Lett.* **B458** (1999) 466-470.

[43] S. Parrott, *Relativistic Electrodynamics and Differential Geometry* (Springer: New York, 1987).

[44] C. Piron, *Électrodynamique et optique.* Course given by C. Piron. Notes edited by E. Pittet (University of Geneva, 1975).

[45] C. Piron and D.J. Moore, *New aspects of field theory*, Turk. J. Phys. **19** (1995) 202-216.

[46] E.J. Post, *Formal Structure of Electromagnetics: General Covariance and Electromagnetics* (North Holland: Amsterdam, 1962, and Dover: Mineola, New York, 1997).

[47] E.J. Post, *The constitutive map and some of its ramifications*, Annals of Physics (NY) **71** (1972) 497-518.

[48] E.J. Post, *Kottler-Cartan-van Dantzig (KCD) and noninertial systems*, Found. Phys. **9** (1979) 619-640.

[49] E.J. Post, *Physical dimension and covariance*, Found. Phys. **12** (1982) 169-195.

[50] E.J. Post, *Quantum Reprogramming — Ensembles and Single Systems: A Two-Tier Approach to Quantum Mechanics* (Kluwer: Dordrecht, 1995).

[51] R.A. Puntigam, C. Lämmerzahl and F.W. Hehl, *Maxwell's theory on a post-Riemannian spacetime and the equivalence principle*, Class. Quantum Grav. **14** (1997) 1347-1356.

[52] G.F. Rubilar, Yu.N. Obukhov and F.W. Hehl, *General covariant Fresnel equation and the emergence of the light cone structure in pre-metric electrodynamics*, *Int. J. Mod. Phys.* **D11** (2002) 1227-1242.

[53] M. Schönberg, *Electromagnetism and gravitation*, Rivista Brasileira de Fisica **1** (1971) 91-122.

[54] J.A. Schouten, *Ricci-Calculus*, 2nd ed. (Springer: Berlin, 1954).

[55] J.A. Schouten, *Tensor Analysis for Physicists,* 2nd ed. reprinted (Dover: Mineola, New York 1989).

[56] E. Schrödinger, *Space-Time Structure* (Cambridge University Press: Cambridge, 1954).

[57] W. Ślebodziński, *Exterior Forms and Their Applications.* Revised translation from the French (PWN–Polish Scientific Publishers: Warszawa, 1970).

[58] A. Sommerfeld, *Elektrodynamik.* Vorlesungen über Theoretische Physik, Band 3 (Dieterich'sche Verlagsbuchhandlung: Wiesbaden, 1948). English translation: A. Sommerfeld, *Electrodynamics,* Vol. 3 of Lectures in Theoretical Physics (Academic Press: New York, 1952).

[59] J. Stachel, *The generally covariant form of Maxwell's equations*, in: *J.C. Maxwell, the Sesquicentennial Symposium.* M.S. Berger, ed. (Elsevier: Amsterdam, 1984) pp. 23-37.

[60] W. Thirring, *Classical Mathematical Physics: Dynamical Systems and Field Theories*, 3rd ed. (Springer: New York, 1997).

[61] R.A. Toupin, *Elasticity and electro-magnetics*, in: *Non-Linear Continuum Theories, C.I.M.E. Conference, Bressanone, Italy 1965.* C. Truesdell and G. Grioli coordinators. Pp.203-342.

[62] A. Trautman, *Differential Geometry for Physicists*, Stony Brook Lectures (Bibliopolis: Napoli, 1984).

[63] C. Truesdell and R.A. Toupin, *The classical field theories*, in: *Handbuch der Physik*, Vol. III/1, S. Flügge ed. (Springer: Berlin, 1960) pp. 226-793.

[64] D. van Dantzig, *The fundamental equations of electromagnetism, independent of metrical geometry*, *Proc. Cambridge Phil. Soc.* **30** (1934) 421-427.

[65] H. Weyl, *Raum, Zeit, Materie*, Vorlesungen über Allgemeine Relativitätstheorie, 8th ed. (Springer: Berlin, 1993). Engl. translation of the 4th ed.: *Space-Time-Matter* (Dover: New York, 1952).

[66] E. Whittaker, *A History of the Theories of Aether and Electricity.* 2 volumes, reprinted (Humanities Press: New York, 1973).

[67] M.R. Zirnbauer, *Elektrodynamik.* Tex-script July 1998 (Springer: Berlin, to be published).

Part A

Mathematics: Some exterior calculus

We will assume that the $\mathbf{J}$ operator is defined in $M^6(\mathbb{C})$ by the same formula as $\mathbf{J}(\omega)$ in M^6. In other words, $\mathbf{J}$ remains a *real* linear operator in $M^6(\mathbb{C})$, i.e. for every complex 2-form $\omega \in M^6(\mathbb{C})$ one has $\overline{\mathbf{J}(\omega)} = \mathbf{J}(\overline{\omega})$. The eigenvalue problem for the operator $\mathbf{J}(\omega_\lambda) = \lambda\,\omega_\lambda$ is meaningful only in the complexified space $M^6(\mathbb{C})$ because, in view of the property (A.1.107), the eigenvalues are $\lambda = \pm i$. Each of these two eigenvalues has multiplicity three, which follows from the reality of $\mathbf{J}$. Note that the 6×6 matrix of the $\mathbf{J}$ operator has six eigenvectors, but the number of eigenvectors with eigenvalue $+i$ is equal to the number of eigenvectors with eigenvalue $-i$ because they are complex conjugate to each other. Indeed, if $\mathbf{J}(\omega) = i\omega$, then the conjugation yields $\overline{\mathbf{J}(\omega)} = \mathbf{J}(\overline{\omega}) = -i\overline{\omega}$.

Let us denote the 3-dimensional subspaces of $M^6(\mathbb{C})$ that correspond to the eigenvalues $+i$ and $-i$ by

$$\begin{aligned} \overset{\text{(s)}}{M} \;&:=\; \left\{\omega \in M^6(\mathbb{C}) \mid \mathbf{J}(\omega) = i\omega\right\}, \\ \overset{\text{(a)}}{M} \;&:=\; \left\{\varphi \in M^6(\mathbb{C}) \mid \mathbf{J}(\varphi) = -i\varphi\right\}, \end{aligned} \tag{A.1.111}$$

respectively. Evidently, $\overline{\overset{\text{(s)}}{M}} = \overset{\text{(a)}}{M}$. Therefore, we can restrict our attention only to the self-dual subspace. This will be assumed in our derivations from now on.

Accordingly, every form ω can be decomposed into a self-dual and an anti-self-dual piece[4],

$$\omega = \overset{\text{(s)}}{\omega} + \overset{\text{(a)}}{\omega}\,, \tag{A.1.112}$$

with

$$\overset{\text{(s)}}{\omega} = \frac{1}{2}\left[\omega - i\mathbf{J}(\omega)\right], \qquad \overset{\text{(a)}}{\omega} = \frac{1}{2}\left[\omega + i\mathbf{J}(\omega)\right]. \tag{A.1.113}$$

It can be checked that $\mathbf{J}(\overset{\text{(s)}}{\omega}) = +i\,\overset{\text{(s)}}{\omega}$ and $\mathbf{J}(\overset{\text{(a)}}{\omega}) = -i\,\overset{\text{(a)}}{\omega}$.

A.1.12 Computer algebra

Also in electrodynamics, research usually requires the application of computers. Besides numerical methods and visualization techniques, the manipulation of formulas by means of "computer algebra" systems is nearly a must. By no means are these methods confined to pure algebra. Differentiations and integrations, for example, can also be executed with the help of computer algebra tools.

"If we do work on the *foundations* of classical electrodynamics, we can dispense with computer algebra," some true fundamentalists will claim. Is this really true? Well later, in Chap. D.2, we analyze the Fresnel equation; we couldn't have done it to the extent we did without using an efficient computer algebra

[4] A discussion of the use of self dual and anti-self dual 2-forms in general relativity can be found in Kopczyński and Trautman [13], e.g..

Figure A.1.5: "Here is the new Reduce-update on a hard disk."

system. Thus, our fundamentalist is well advised to learn some computer algebra. Accordingly, in addition to introducing some mathematical tools in exterior calculus, we mention computer algebra systems like *Reduce*[5], *Maple*[6], and *Mathematica*[7] and specifically explain how to apply the Reduce package *Excalc*[8] to the exterior forms that occur in electrodynamics.

In practical work in solving problems by means of computer algebra, it is our experience that it is best to have access to different computer algebra systems. Even though in the course of time good features of one system "migrated" to other systems, still, for a certain specified purpose one system may be better suited than another one — and for different purposes these may be different systems. There does not exist as yet *the* optimal system for all purposes. There-

[5] Hearn [8] created this Lisp-based system. For introductions to Reduce, see Toussaint [36], Grozin [6], MacCallum and Wright [15], or Winkelmann and Hehl [38]; in the latter text you can learn how to get hold of a Reduce system for your computer. Reduce as applied to general-relativistic field theories is described, for example, by McCrea [16] and by Socorro et al. [28]. In our presentation, we partly follow the lectures of Toussaint [36].

[6] Maple, written in C, was created by a group at the University of Waterloo, Canada. A good introduction is given by Char et al. [3].

[7] Wolfram (see [39]) created the C-based Mathematica software package which is in very widespread use.

[8] Schrufer@Schrüfer [25, 26] is the creator of that package (cf. also [27]). Excalc is applied to Maxwell's theory by Puntigam et al. [22].

fore, it is not a rare occasion that we have to feed the results of a calculation by means of one system as input into another system.

For computations in electrodynamics, relativity, and gravitation, we keep the three general-purpose computer algebra systems: Reduce, Maple, and Mathematica. Other systems are available[9]. Our workhorse for corresponding calculations in exterior calculus is the Reduce package Excalc, but also in the *MathTensor* package[10] of Mathematica exterior calculus is implemented. For the manipulation of tensors we use the following packages: In Reduce the library of McCrea[11] and *GRG*[12], in Maple *GRTensorII*[13], and in Mathematica, besides MathTensor, the *Cartan* package[14].

Computer algebra systems are almost exclusively interactive systems nowadays. If one is installed on your computer, you can usually call the system by typing its name or an abbreviation thereof, i.e., '`reduce`', '`maple`', or '`math`', and then hitting the return key, or clicking on the corresponding icon. In the case of 'reduce', the system introduces itself and issues a '`1:`'. It waits for your first command. A command is a statement, usually some sort of expression, a part of a formula or a formula, followed by a terminator[15]. The latter, in Reduce, is a semicolon `;` if you want to see the answer of the system, and otherwise a dollar sign `$`. Reduce is case insensitive, i.e., the lowercase letter `a` is not distinguished from the uppercase letter `A`.

Formulating Reduce input

As an input statement to Reduce, we type in a certain legitimately formed expression. This means that, with the help of some operators, we compose formulas according to well-defined rules. Most of the built-in operators of Reduce, like the arithmetic operators `+` (plus), `-` (minus), `*` (times), `/` (divided by), `**` (to the power of)[16] are self-explanatory. They are so-called *infix operators* since they are positioned *in* between their arguments. By means of them we can construct combined expressions of the type $(x+y)^2$ or $x^3 \sin x$, which in Reduce read `(x+y)**2` and `x**3*sin(x)`, respectively.

If the command

---

[9] In the review of Hartley [7] possible alternative systems are discussed (see also Heinicke et al. [11]).

[10] Parker and Christensen [21] created this package; for a simple application see Tsantilis et al. [37].

[11] See McCrea's lectures [16].

[12] The GRG system, created by Zhytnikov [40], and the $\mathrm{GRG}_{\mathrm{EC}}$ system of Tertychniy [35, 34, 20] grew from the same root; for an application of $\mathrm{GRG}_{\mathrm{EC}}$ to the Einstein–Maxwell equations, see [33].

[13] See the documentation of Musgrave et al. [19]. Maple applications to the Einstein–Maxwell system are covered in the lectures of McLenaghan [17].

[14] Soleng [30] is the creator of 'Cartan'.

[15] Has nothing to do with Arnold Schwarzenegger!

[16] Usually one takes the circumflex for exponentiation. However, in the Excalc package this operator is redefined and used as the wedge symbol for exterior multiplication.

	Switch	description if switch is on	example
*	allfac	factorize simple factors	$2x+2 \to 2(x+1)$
	div	divide by the denominator	$(x^2+2)/x \to x+2/x$
*	exp	expand all expressions	$(x+1)(x-1) \to x^2-1$
*	mcd	make (common) denominator	$x+x^{-1} \to (x^2+1)/x$
*	lcm	cancel least common multiples	
	gcd	cancel greatest common divisor	
	rat	display as polynomial in `factor`	$\frac{x+1}{x} \to 1+x^{-1}$
*	ratpri	display rationals as fraction	$1/x \to \frac{1}{x}$
*	pri	dominates `allfac,div,rat,revpri`	
	revpri	display polynom. in opposite order	$x^2+x+1 \to 1+x+x^2$
	rounded	calculate with floats	$1/3 \to 0.333333333333$
	complex	simplify complex expressions	$1/i \to -i$
	nero	don't display zero results	$0 \to$
*	nat	display in Reduce input format	$\frac{x^2}{3} \to x{*}{*}2/3$
	msg	suppress messages	
	fort	display in Fortran format	
	tex	display in TeX format	

Table A.1.2: Switches for Reduce's reformulation rules. Those marked with `*` are turned on by default; the other ones are off.

```
(x+y)**2;
```

is executed, you will get the expanded form $x^2+2xy+y^2$. There is a so-called switch `exp` in Reduce that is usually switched on. You can switch it off by the command

```
off exp;
```

Type in again

```
(x+y)**2;
```

Now you will find that Reduce doesn't do anything and gives the expression back as it received it. With `on exp;` you can go back to the original status.

Using the switches is a typical way to influence Reduce's way of how to evaluate an expression. A partial list of switches is presented in a table on the next page.

Let us give some more examples of expressions with infix operators:

```
(u+v)*(y-x)/8
(a>b) and (c<d)
```

Here we have the logical and relational operators `and`, > (greater than), < (less than). Widely used are also the infix operators:

```
neq   >=   <=   or   not   :=
```

The operator `neq` means not equal. The assignment operator `:=` assigns the value of the expression on its right-hand side (its second argument) to the identifier on its left-hand side (its first argument). In Reduce, logical (or Boolean) expressions have only the truth values `t` (true) or `nil` (false). They are only allowed within certain statements (namely in `if`, `while`, `repeat`, and `let` statements) and in so-called rule lists.

A *prefix operator* stands in front of its argument(s). The arguments are enclosed by parentheses and separated by commas:

```
cos(x)
int(cos(x),x)
factorial(8)
```

In ordinary notation, the second statement reads $\int \cos x \, dx$.

The following mathematical functions are built-in as prefix operators:

```
sin       sind      cos       cosd      tan       tand
cot       cotd      sec       secd      csc       cscd
asin      asind     acos      acosd     atan      atand
acot      acotd     asec      asecd     acsc      acscd
sinh      cosh      tanh      coth      sech      csch
asinh     acosh     atanh     acoth     asech     acsch
sqrt      exp       ln        log       log10     logb
dilog     erf       expint    cbrt      abs       hypot
factorial
```

Identifiers ending with `d` indicate that this operator expects its argument to be expressed in degrees. `log` stands for the natural logarithm and is equivalent to `ln`. `logb` is the logarithm in base n; accordingly n must be specified as a second argument of `logb`. `hypot` calculates the hypotenuse according to $\mathrm{hypot}(x,y) = \sqrt{x^2+y^2}$. `csc` is the cosecant, `dilog` the Euler dilogarithm with $\mathrm{dilog}(z) = -\int_0^z \log(1-\zeta)/\zeta \, d\zeta$, `erf` the Gaussian error function with $\mathrm{erf}(x) = 2/\sqrt{\pi} \int_0^x e^{-t^2} dt$, `abs` the absolute value function, `expint` the exponential

Part B

Axioms of classical electrodynamics

In Part B we put phenomenological classical electrodynamics into such a form that the underlying physical facts are clearly visible. We recognize that the conservation of *electric charge* and of *magnetic flux* are the two main experimentally well-founded axioms of electrodynamics. To formulate them we use exterior calculus because it is the appropriate mathematical framework for handling fields the integrals of which — here charge and flux — possess an invariant meaning.

The densities of the electric charge and electric current are assumed to be phenomenologically specified. These quantities will not be resolved any further and will be considered as fundamental for classical electrodynamics.

B.1
Electric charge conservation

. . . it is now discovered and demonstrated, both here and in Europe, that the Electrical Fire is a real Element, or Species of Matter, not created *by the Friction, but* collected *only.*

Benjamin Franklin (1747)[1]

B.1.1 Counting charges. Absolute and relative dimension

Progress in semiconductor technology has enabled the fabrication of structures so small that they can contain just one mobile electron. By varying controllably the number of electrons in these 'artificial atoms' and measuring the energy required to add successive electrons, one can conduct atom physics exper-

[1] See Heilbron [16], p. 330. On the same page Heilbron states: "Although Franklin did not 'discover' conservation, he was unquestionably the first to exploit the concept fruitfully. Its full utility appeared in his classic analysis of the condenser." Note that the discovery of charge conservation preceded the discovery of the Coulomb law (1785) by more than 40 years. This historical sequence is reflected in our axiomatics. However, our reason is not an historical but a conceptual one: Charge conservation should come first and the Maxwell equations should be formulated so as to be compatible with that law.

iments in a regime that is inaccessible to experiments on real atoms.

R. C. Ashoori (1996)[2]

Phenomenologically speaking, electromagnetism has two types of sources: The electric charge density ρ and the electric current density $\vec{j}$. The electric current density can be understood, with respect to some reference system, as moving electric charge density.

In this section we give a *heuristic* discussion of charge conservation that will be used, in the next section, as a motivation for the formulation of a first axiom for electrodynamics.

Imagine a 3-dimensional simply connected region Ω_3 that is enclosed by the 2-dimensional boundary $\partial\Omega_3$ (see Fig. B.1.1). In the region Ω_3, there are elementary particles with charge $\pm e$ (e = elementary electric charge) and quarks with charges $\pm\frac{1}{3}e$ and $\pm\frac{2}{3}e$, respectively, which all move with some velocity. It is assumed in electrodynamics that we can attribute to the 3-region Ω_3 at any time in a certain reference system a well-defined net charge Q with the *dimension*[3] of charge q (in SI units, coulomb, abbreviated C):

$$Q = \int_{\Omega_3} \rho\,, \qquad [Q] = q\,, \quad [\rho] = q\,. \tag{B.1.1}$$

Here $[Q]$ should be read as "dimension of Q" and, analogously, $[\rho]$ as "dimension of ρ." We call $[Q]$ and $[\rho]$ also *absolute* dimension of Q and ρ, respectively.

The integrand ρ is called the electric charge density 3-form. It assigns to the volume 3-form in an arbitrary coordinate system ($a, b, \ldots = 1, 2, 3$)

$$dx^a \wedge dx^b \wedge dx^c\,, \tag{B.1.2}$$

a scalar-valued charge

$$\rho = \frac{1}{3!}\,\rho_{abc}\, dx^a \wedge dx^b \wedge dx^c\,, \qquad \rho_{abc} = \rho_{[abc]}\,, \tag{B.1.3}$$

which, as a scalar, can be added up in any coordinate system to yield Q. Thus, as already noted, even the charge *density* ρ carries the same absolute dimension as the net charge Q.

In spatial spherical coordinates $x^a = (r, \theta, \phi)$, for instance, the coordinates carry different dimensions: The x^1 or r-coordinate has the dimension of a length,

[2]See his article [1]. For a review on the counting of single electrons, see Devoret and Grabert [5] and, for more recent developments, Fujiwara and Takahashi [12].

[3]We base our dimensional analysis on the work of Dorgelo and Schouten (see Schouten [45], pp. 126–138 and Post [37], pp. 23–46). A guide of how to use the International System of Units (SI) has been provided by Nelson [34]; see also the references therein.

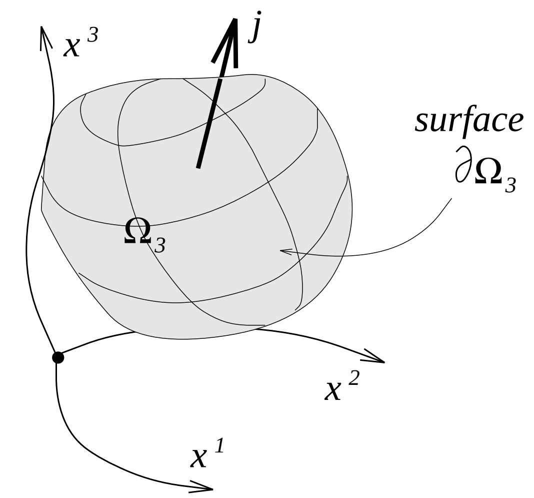

Figure B.1.1: Charge conservation in 3-dimensional space. The electric current density 2-form j flows out of the volume Ω_3.

whereas $x^2 = \theta$ and $x^3 = \phi$ are dimensionless. Accordingly, the different coordinate components ρ_{abc} of ρ have different dimensions that cannot be measured straightforwardly in an experiment. Thus these components of ρ are unsuitable for a general definition of the relative (also called physical) dimension of a quantity.

Therefore we introduce an arbitrary 3-dimensional local coframe $\vartheta^{\hat{a}}$ with $\hat{a}, \hat{b}, \ldots = \hat{1}, \hat{2}, \hat{3}$, and its dual frame $e_{\hat{b}}$ with $e_{\hat{b}} \rfloor \vartheta^{\hat{a}} = \delta^{\hat{a}}_{\hat{b}}$. We mark the numbers $\hat{1}, \hat{2}, \hat{3}$ by a circumflex in order to be able to identify them as being related to *an*holonomic 3D indices. A physicist needs such a (co)frame to anchor his or her apparatus. We assign to each of the three 1-forms $\vartheta^{\hat{a}}$ the *dimension of a length* ℓ and to the corresponding vectors $e_{\hat{b}}$ the dimension of ℓ^{-1}:

$$\begin{aligned} \left[\vartheta^{\hat{a}}\right] &= \ell\,, \qquad \text{for } \hat{a} = \hat{1}, \hat{2}, \hat{3}\,, &\text{(B.1.4)}\\ \left[e_{\hat{b}}\right] &= \ell^{-1}\,, \qquad \text{for } \hat{b} = \hat{1}, \hat{2}, \hat{3}\,. &\text{(B.1.5)} \end{aligned}$$

Length is here understood as a *segment*, that is, as part of a straight line between two points. Accordingly, length is synonymous with a 1-dimensional extension in affine geometry; it is not, however, the distance in the sense of Euclidean geometry. In other words, length as dimension does *not* presuppose the existence of a metric; it is a pre-metric concept.

Now we can decompose the charge density 3-form with respect to the coframe $\vartheta^{\hat{a}}$,

$$\rho = \frac{1}{3!}\,\rho_{\hat{a}\hat{b}\hat{c}}\,\vartheta^{\hat{a}} \wedge \vartheta^{\hat{b}} \wedge \vartheta^{\hat{c}}\,, \quad \text{rel. dim.} := [\rho_{\hat{a}\hat{b}\hat{c}}] = q\,\ell^{-3}\,. \qquad \text{(B.1.6)}$$

The dimensions of all *an*holonomic components $\rho_{\hat{a}\hat{b}\hat{c}}$ of the charge density ρ are the same. We call $[\rho_{\hat{a}\hat{b}\hat{c}}]$ the *relative dimension* of ρ. Accordingly, the absolute dimension of ρ is *charge*, i.e., $[\rho] = q$, whereas the relative dimension of ρ is *charge/(length)*^{3}, i.e., $[\rho_{\hat{a}\hat{b}\hat{c}}] = q\,\ell^{-3}$.

In the *hypothesis of locality*[4] it is assumed that the measuring apparatus in the coframe $\vartheta^{\hat{a}}$, even if the latter is accelerated, measures the anholonomic components of a physical quantity, such as the components $\rho_{\hat{a}\hat{b}\hat{c}}$, exactly as in a momentarily comoving inertial frame of reference. In the special case of the measurment of time, Einstein spoke about the clock hypothesis.

If we assume, as suggested by experience, that the electric charge Q has *no intrinsic screw-sense*, then the sign of Q does not depend on the orientation in space. Accordingly, the charge density ρ is represented by a *twisted* 3-form; for the definition of twisted quantities, see the end of Sec. A.1.3.

Provided the coordinates x^a are given in Ω_3, we can determine dx^a and the volume 3-form (B.1.2). There is no need to use a metric nor a connection, the properties of a 'bare' differential manifold (continuum) are sufficient for the definition of (B.1.1). This can also be recognized as follows:

The net charge Q in (B.1.1) can be determined by counting the charged elementary particles inside $\partial\Omega_3$ and adding up their elementary electric charges. Nowadays one catches single electrons in traps. Thus the counting of electrons is an experimentally feasible procedure, not only a thought experiment devised by a theoretician. This consideration shows that it is not necessary to use a distance concept nor a length standard in Ω_3 for the determination of Q. Only 'counting procedures' are required and a way to delimit an arbitrary finite volume Ω_3 of 3-dimensional space by a boundary $\partial\Omega_3$ and to know what is inside $\partial\Omega_3$ and what outside. Accordingly, ρ is the *prototype* of a charge density with absolute dimension $[\rho] = q$ and relative dimension $[\rho_{abc}] = q\,\ell^{-3}$. It becomes the conventional charge density, that is charge per *scaled* unit volume, if a unit of distance (m in SI-units) is introduced additionally. Then, in SI-units, we have $[\rho_{\hat{a}\hat{b}\hat{c}}] = \mathrm{C\,m^{-3}}$.

Out of the region Ω_3, crossing its bounding surface $\partial\Omega_3$, there will, in general, flow a net electric current, see Fig.B.1.1,

$$\mathcal{J} = \oint_{\partial\Omega_3} j\,, \tag{B.1.7}$$

with absolute dimension $q\,t^{-1}$ (t = time), which must not depend on the orientation of space either. The integrand j, the twisted electric current density

[4]The formulation of Mashhoon [30] reads: "...the hypothesis of locality — i.e., the presumed equivalence of an accelerated observer with a momentarily comoving inertial observer — underlies the standard relativistic formalism by relating the measurements of an accelerated observer to those of an inertial observer." The two observers are assumed to be otherwise identical. That is, the two observers are copies of one observer: One copy is inertial and the other accelerated; this is the only difference between them. The limitations of the hypothesis of locality are discussed in [31].

2-form with the same absolute dimension $q\,t^{-1} \overset{\text{SI}}{=} \text{C}\,\text{s}^{-1}$ = ampere = A, here s = second, assigns to the area element 2-form

$$dx^a \wedge dx^b \tag{B.1.8}$$

a scalar-valued charge current

$$j = \frac{1}{2!}\, j_{ab}\, dx^a \wedge dx^b\,, \qquad j_{ab} = j_{[ab]}\,. \tag{B.1.9}$$

The postulate of electric charge conservation requires

$$\frac{dQ}{dt} = -\mathcal{J}\,, \tag{B.1.10}$$

provided the area 2-form $dx^a \wedge dx^b$ is directed in such a way that the outflow is counted positively in (B.1.7). The time variable t, provisionally introduced here, does not need to possess a scale or a unit. It can be called a "smooth causal time" in the sense of parameterizing a future-directed curve in the spacetime manifold with t as a monotone increasing and sufficiently smooth variable. Substitution of (B.1.1) and (B.1.7) into (B.1.10) yields an integral form of charge conservation:

$$\frac{d}{dt}\int\limits_{\Omega_3} \rho + \oint\limits_{\partial\Omega_3} j = 0\,. \tag{B.1.11}$$

By applying the 3-dimensional Stokes theorem, the differential version turns out to be

$$\frac{\partial\rho}{\partial t} + d\,j = 0\,. \tag{B.1.12}$$

Let us put (B.1.11) into a 4-dimensional form. For this purpose we integrate (B.1.11) over a certain time interval from t_1 to t_2 (see Fig. B.1.2). Note that this figure depicts the same physical situation as in Fig. B.1.1:

$$\begin{aligned} &\int\limits_{t_1}^{t_2} dt\,\frac{d}{dt}\int\limits_{\Omega_3} \rho \quad + \quad \int\limits_{t_1}^{t_2} dt \wedge \oint\limits_{\partial\Omega_3} j \\ = &\int\limits_{t=t_2,\,\Omega_3} \rho - \int\limits_{t=t_1,\,\Omega_3} \rho \quad + \int\limits_{[t_1,t_2]\times\partial\Omega_3} dt \wedge j = 0\,. \end{aligned} \tag{B.1.13}$$

Obviously we are integrating over a 3-dimensional boundary of a compact piece of the 4-dimensional spacetime. If we introduce in four dimensions the twisted 3-form

$$J := -j \wedge dt + \rho\,, \qquad \text{with} \qquad [J] = q\,, \tag{B.1.14}$$

then the integral can be written as a 4-dimensional boundary integral,

$$\oint\limits_{\partial\Omega_4} J = 0\,, \tag{B.1.15}$$

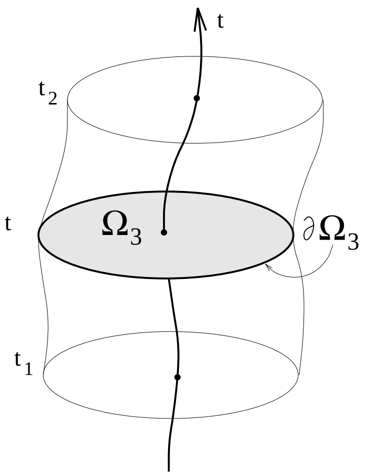

Figure B.1.2: Charge conservation in 4-dimensional spacetime. The 3-dimensional "cross section" Ω_3 sweeps out a 4-dimensional volume Ω_4 on its way from $t = t_1$ to $t = t_2$.

where $\Omega_4 = [t_1, t_2] \times \Omega_3$. The twisted 3-form J of the electric charge-current density with absolute dimension q plays the central role as source of the electromagnetic field.

B.1.2 Spacetime and the first axiom

Motivated by the integral form of charge conservation (B.1.15), we can now turn to an axiomatic approach of electrodynamics. First we will formulate a set of minimal assumptions that we shall need for defining an appropriate spacetime manifold.

Let spacetime be given as a 4-dimensional *connected, Hausdorff, orientable, and paracompact* differentiable manifold X_4. This manifold is "bare," that is, it carries neither a connection nor a metric so far. We assume, however, the conventional continuity and differentiability requirements of physics. To recall, a topological space X is Hausdorff when for any two points $p_1 \neq p_2 \in X$ one can find open sets $p_1 \in U_1 \subset X$, $p_2 \in U_2 \subset X$, such that $U_1 \cap U_2 = \emptyset$. An X is connected when any two points can be connected by a continuous curve. Finally, a connected Hausdorff manifold is paracompact when X can be covered by a countable number of coordinate charts. The (smooth) coordinates in arbitrary charts will be called x^i, with $i, j, k, \ldots = 0, 1, 2, 3$. The vector basis (frame) of the tangent space will be called e_α and the 1-form basis (coframe) of the cotangent space ϑ^α with (anholonomic) indices $\alpha, \beta, \gamma, \ldots = \hat{0}, \hat{1}, \hat{2}, \hat{3}$. On the X_4 we can

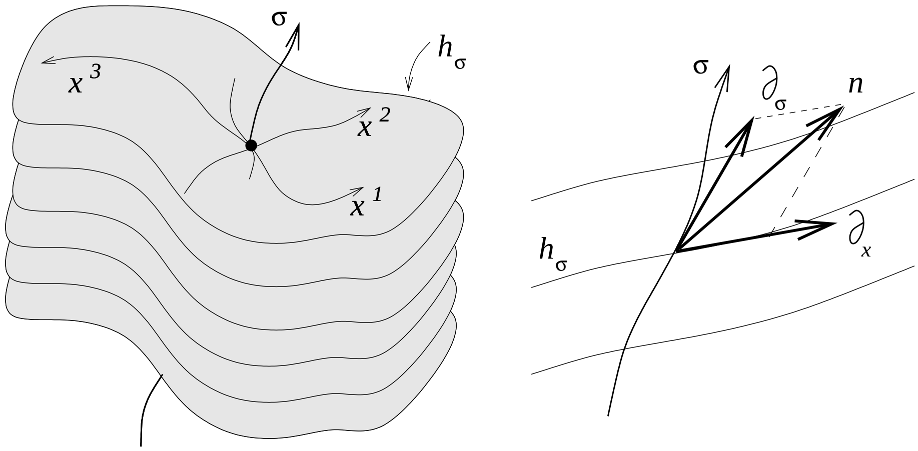

Figure B.1.3: Spacetime and its (1 + 3)-foliation.

define twisted and ordinary untwisted tensor-valued differential forms. In order to avoid possible violations of causality, we will, as usual, consider only *non-compact* spacetime manifolds X_4.

The X_4 with the described topological properties is known to possess a (1+3)-foliation (see Fig. B.1.3); i.e., there exists a set of nonintersecting 3-dimensional hypersurfaces h_σ that can be parameterized by a monotone increasing (would-be time) variable σ with the dimension of time: $[\sigma] = t$. Although at this stage we do not introduce any metric on X_4, it is well known that the existence of a $(1 + 3)$-foliation is closely related to the existence of pseudo-Riemannian structures.

Among the vector fields *transverse* to the foliation, we choose a vector field n (not to be confused with the dimension n of a manifold discussed in Part A) normalized by the condition

$$n \rfloor d\sigma = \pounds_n \sigma = 1 \,, \qquad [n] = [\sigma]^{-1} = t^{-1} \,. \tag{B.1.16}$$

Physically, the folia h_σ of constant σ represent a simple model of a "3-space," while the function σ serves as a "time" variable. The vector field n is usually interpreted as a congruence of observer worldlines. In Sec. E.4.1, this rather formal mathematical construction becomes a full-fledged physical tool when the metric is introduced.

Now we are in a position to formulate our first axiom. We require the existence of a *twisted charge-current density 3-form* J with the absolute dimension of charge q, i.e., $[J] = q \overset{\text{SI}}{=} \text{C}$ which, if integrated over an arbitrary *closed* 3-dimensional submanifold $C_3 \subset X_4$, obeys

$$\oint_{C_3} J = 0 \,, \qquad \partial C_3 = 0 \qquad \text{(first axiom)} \,. \tag{B.1.17}$$

We recall, a manifold is closed if it is compact and has no boundary. In particular, the 3-dimensional boundary $C_3 = \partial\Omega_4$ of an arbitrary 4-dimensional region Ω_4 is a closed manifold. However, in general, not every closed 3-manifold is a 3-boundary of some spacetime region.

This is the first axiom of electrodynamics. It has a firm phenomenological basis.

B.1.3 Electromagnetic excitation H

Since (B.1.17) holds for an arbitrary 3-dimensional boundary C_3, we can choose $C_3 = \partial\Omega_4$. Then, by Stokes' theorem, one finds

$$\int_{\Omega_4} dJ = 0\,. \tag{B.1.18}$$

Since Ω_4 can be chosen arbitrarily, the electric current turns out to be a closed form:

$$dJ = 0\,. \tag{B.1.19}$$

This is, in four dimensions, the differential version of charge conservation.

Having proved that J is a closed 3-form, we now recognize (B.1.17) as the statement that all periods of the current J vanish. Then, by de Rham's first theorem (see (A.3.42)), the current is also an exact form:

$$J = dH\,. \tag{B.1.20}$$

This is the inhomogeneous Maxwell equation. The twisted electromagnetic excitation 2-form H has the absolute dimension of charge q, i.e., $[H] = q$.

The excitation H in this setup appears as a potential of the electric current. It is determined only up to a closed 2-form

$$H \longrightarrow H + \psi\,, \qquad d\psi = 0\,. \tag{B.1.21}$$

In Sec. B.3.4, however, we discuss how a *unique* excitation field H is selected from the multitude of H's occurring in (B.1.21) by a very weak assumption: If the field strength F, to be defined below via the second axiom in (B.2.8), vanishes, then the excitation H vanishes too. In this context we recognize that the electric and magnetic pieces of H can be *measured* by means of an ideal electric conductor and a superconductor of type II, respectively.

Charge conservation, as formulated in the first axiom, is experimentally verified in all microscopic experiments, in particular in those of high-energy elementary particle physics. Therefore the excitation H represents a *microscopic* field as well. Charge conservation is *not* only valid as a macroscopic average. Similarly, the excitation is not only a quantity that shows up in macrophysics, it rather is a microscopic field, too, analogous to the electromagnetic field strength F to be introduced below.

B.1.4 Time-space decomposition of the inhomogeneous Maxwell equation

Time is nature's way of keeping everything from happening at once.

Anonymous

Given the spacetime foliation, we can decompose any exterior form in "time" and "space" pieces. With respect to the fixed vector field n, normalized by (B.1.16), we define, for a p-form Ψ, the part *longitudinal* to the vector n by

$$^{\perp}\Psi := d\sigma \wedge \Psi_{\perp}\,, \qquad\qquad \Psi_{\perp} := n \rfloor \Psi\,, \tag{B.1.22}$$

and the part *transversal* to the vector n by

$$\underline{\Psi} := (1 \;-\; {}^{\perp})\Psi = n \rfloor (d\sigma \wedge \Psi)\,, \qquad\qquad n \rfloor \underline{\Psi} \equiv 0\,. \tag{B.1.23}$$

Then, we finally have

$$\Psi = {}^{\perp}\Psi + \underline{\Psi} = d\sigma \wedge \Psi_{\perp} + \underline{\Psi}\,, \tag{B.1.24}$$

with the absolute dimensions

$$[\Psi_{\perp}] = [\Psi]\, t^{-1}\,, \qquad [\underline{\Psi}] = [\Psi]\,. \tag{B.1.25}$$

Thus the projection operators "$^{\perp}$" and "_" form a complete set. Furthermore, every 0-form is transversal, whereas every 4-form is longitudinal, that is, for $p = 0$, we find $\Psi = \underline{\Psi}$ and, for $p = 4$, $\Psi = {}^{\perp}\Psi$.

In order to apply this decomposition to field theory, the following rules for exterior multiplication can be derived from (B.1.22) and (B.1.23),

$$^{\perp}(\Psi \wedge \Phi) = {}^{\perp}\Psi \wedge \underline{\Phi} + \underline{\Psi} \wedge {}^{\perp}\Phi \;= d\sigma \wedge \left(\Psi_{\perp} \wedge \underline{\Phi} + (-1)^p \underline{\Psi} \wedge \Phi_{\perp}\right), \tag{B.1.26}$$

$$\underline{\Psi \wedge \Phi} = \underline{\Psi} \wedge \underline{\Phi}\,, \tag{B.1.27}$$

if Ψ is a p-form. We introduce the 3-dimensional exterior derivative as the transversal part of d, that is, according to (B.1.23), as $\underline{d} := n \rfloor (d\sigma \wedge d)$. Then the exterior derivative of a p-form decomposes as follows:

$$^{\perp}(d\Psi) = d\sigma \wedge (\pounds_n \underline{\Psi} - \underline{d}\Psi_{\perp}), \qquad\qquad \underline{d\Psi} = \underline{d}\,\underline{\Psi}. \tag{B.1.28}$$

According to (A.2.51), the *Lie derivative* of a p-form along a vector field ξ can be written as

$$\pounds_{\xi}\Psi := \xi \rfloor d\Psi + d(\xi \rfloor \Psi). \tag{B.1.29}$$

Notice that the Lie derivative along the foliation vector field n commutes with the projection operators as well as with the exterior derivative, i.e.,

$$\pounds_n {}^{\perp}\Psi = {}^{\perp}(\pounds_n \Psi), \quad \pounds_n \underline{\Psi} = \underline{\pounds_n \Psi}, \quad \pounds_n d\Psi = d(n \rfloor d\Psi) = d(\pounds_n \Psi). \tag{B.1.30}$$

These rules also imply that

$$\pounds_n \underline{d}\,\underline{\Psi} = \underline{d}\,\pounds_n \underline{\Psi}. \tag{B.1.31}$$

The Lie derivative of the transversal piece $\underline{\Psi}$ of a form with respect to the vector n is abbreviated by a dot,

$$\underline{\dot{\Psi}} := \pounds_n \underline{\Psi}\,, \tag{B.1.32}$$

since this turns out to be the time derivative of the corresponding quantity.

In order to make this decomposition formalism more transparent, it is instructive to consider natural (co)frames on a 3-dimensional hypersurface h_σ of constant σ (see Fig. B.1.3). Let x^a be local coordinates on h_σ with $a = 1, 2, 3$. The differentials dx^a are *not* transversal to n in general. Indeed, in terms of the local spacetime coordinates (σ, x^a), the normalization (B.1.16) allows for a vector field of the structure $n = \partial_\sigma + n^a\,\partial_a$ with some (in general, nonvanishing) functions n^a. We use (B.1.22) and (B.1.23) and find a nontrivial longitudinal piece ${}^\perp(dx^a) = n^a\,d\sigma$, whereas the transversal piece reads $\underline{dx}^a = dx^a - n^a\,d\sigma$. Obviously $d\sigma$ is purely longitudinal. Hence it is convenient to choose, at an arbitrary point of spacetime, a basis of the cotangent space

$$(\vartheta^{\hat{0}}, \vartheta^a) = (d\sigma, \underline{dx}^a)\,, \qquad [\vartheta^{\hat{0}}] = t\,,\ [\vartheta^a] = \ell\,, \tag{B.1.33}$$

which is compatible with the foliation given. The coframe (B.1.33) manifestly spans the longitudinal and transversal subspaces of the cotangent space. The corresponding basis of the tangent space reads

$$(e_{\hat{0}}, e_a) = (n, \partial_a)\,, \qquad [e_{\hat{0}}] = t^{-1},\ [e_a] = \ell^{-1}\,. \tag{B.1.34}$$

This coframe and this frame are anholonomic, in general.

The general decomposition scheme can be applied to the inhomogeneous Maxwell equation (B.1.20). First, we decompose its left-hand side. Then the current reads

$$J = {}^\perp J + \underline{J} = -j \wedge d\sigma + \rho =: (j, \rho)\,, \tag{B.1.35}$$

with

$$j := -J_\perp \qquad \text{and} \qquad \rho := \underline{J}\,. \tag{B.1.36}$$

The minus sign is chosen in conformity with (B.1.14). Since $[J] = q$, we have for the absolute and the relative dimensions, respectively,

$$\begin{aligned} [j] &= q\,t^{-1} \overset{\text{SI}}{=} \mathrm{C\,s^{-1}} = \mathrm{A}, \quad [\rho] = q \overset{\text{SI}}{=} \mathrm{C}, \\ [j_{\hat{a}\hat{b}}] &= [j]\,\ell^{-2} \overset{\text{SI}}{=} \mathrm{A\,m^{-2}}, \quad [\rho_{\hat{a}\hat{b}\hat{c}}] = [\rho]\,\ell^{-3} \overset{\text{SI}}{=} \mathrm{C\,m^{-3}}. \end{aligned} \tag{B.1.37}$$

In three dimensions, we recover the twisted current density 2-form j (see (B.1.9)) and the twisted charge density 3-form ρ (see (B.1.3)). Now charge conservation (B.1.19) can easily be decomposed, too. We substitute (B.1.36) into (B.1.28) for $\Psi = J$. Then, with the abbreviation (B.1.32), we recover (B.1.12):

$$\dot{\rho} + \underline{d}\, j = 0\,. \qquad \text{(B.1.38)}$$

This, at the same time, gives an exact meaning to the time derivative, which we treated somewhat sloppily in Sec. B.1.1. Note that $\underline{d}\,\rho = 0$, as a 4-form in three dimensions, represents an identity with no additional information.

Before we turn to the right-hand side of (B.1.20), we decompose the excitations according to[5]

$$H = {}^{\perp}H + \underline{H} = -\mathcal{H} \wedge d\sigma + \mathcal{D} =: (\mathcal{H}, \mathcal{D}) \qquad \text{(B.1.39)}$$

with the twisted magnetic excitation 1-form $\mathcal{H}$ and the twisted electric excitation 2-form $\mathcal{D}$:

$$\mathcal{H} := H_{\perp} \qquad \text{and} \qquad \mathcal{D} := \underline{H}\,. \qquad \text{(B.1.40)}$$

The signs in these two definitions are discussed in section B.4.2. The absolute and relative dimensions of $\mathcal{H}$ and $\mathcal{D}$ are, respectively,

$$\begin{aligned} [\mathcal{H}] &= qt^{-1} \stackrel{\text{SI}}{=} \text{A}, & [\mathcal{D}] &= q \stackrel{\text{SI}}{=} \text{C}, \\ [\mathcal{H}_{\hat{a}}] &= [\mathcal{H}]\,\ell^{-1} \stackrel{\text{SI}}{=} \text{A}\,\text{m}^{-1}, & [\mathcal{D}_{\hat{a}\hat{b}}] &= [\mathcal{D}]\,\ell^{-2} \stackrel{\text{SI}}{=} \text{C}\,\text{m}^{-2}. \end{aligned} \qquad \text{(B.1.41)}$$

Everything is now prepared: The longitudinal part of (B.1.20) reads

$${}^{\perp}J = d\sigma \wedge J_{\perp} = {}^{\perp}(dH) = d\sigma \wedge \left(\pounds_n \underline{H} - \underline{d} H_{\perp}\right) \qquad \text{(B.1.42)}$$

and the transversal part

$$\underline{J} = \underline{dH} = \underline{d}\,\underline{H}\,. \qquad \text{(B.1.43)}$$

By means of (B.1.35), (B.1.36), and (B.1.39), the last two equations can be rewritten as

$$\underline{d}\,\mathcal{H} - \dot{\mathcal{D}} = j \qquad \text{(B.1.44)}$$

and

$$\underline{d}\,\mathcal{D} = \rho\,, \qquad \text{(B.1.45)}$$

respectively. We find it remarkable that all we need to recover the inhomogeneous set of the Maxwell equations (B.1.44), (B.1.45) (see also their vector versions in (I.6) of the introduction) is electric charge conservation in the form (B.1.17) for arbitrary periods — and nothing more. Incidentally, the boundary conditions for $\mathcal{H}$ and $\mathcal{D}$, if required, can also be derived from (B.1.17). Because of the dot, formula (B.1.44) represents an evolution equation, whereas (B.1.45) constitutes a constraint on the initial distributions of $\mathcal{D}$ and ρ.

---

[5]The historical name of $\mathcal{H}$ is "magnetic field" and that of $\mathcal{D}$ "electric displacement." The ISO names (International Organization for Standardization) are "magnetic field" for $\mathcal{H}$ and "electric flux density" for $\mathcal{D}$. As we can see, the ISO names are in conflict with a reasonable 4D interpretation; that is, the 4D quantity H remains without a name. The ISO names are based on the difference between a 1-form, here $\mathcal{H}$ (field), and a 2-form, here $\mathcal{D}$ (flux density).

B.2
Lorentz force density

B.2.1 Electromagnetic field strength F

By now we have exhausted the information contained in the axiom of charge conservation. We have to introduce new concepts in order to complete the fundamental structure of electrodynamics. Whereas the excitation $H = (\mathcal{H}, \mathcal{D})$ is linked to the charge current $J = (j, \rho)$, the electric and magnetic field strengths are usually introduced as forces acting on unit charges q at rest or in motion, respectively.

Let us start with a *heuristic* discussion. We turn first to classical mechanics. Therein the force $\mathcal{F}$ has to be a covector since this is the way it is defined in Lagrangian mechanics, namely $\mathcal{F}_a \sim \partial L/\partial x^a$, with the Lagrangian L. In the purely electric case, the force $\mathcal{F}$ acting on a test charge q at rest in spatial components reads:

$$\mathcal{F}_a \sim q\, E_a\,. \tag{B.2.1}$$

Thereby we can *define* the electric field strength E in 3-dimensional space. The electric field strength E has three independent components exactly as the electric excitation $\mathcal{D}$.

In order to link up mechanics with the rudimentary electrodynamics of the first axiom, we consider a delta-function-like test charge current $J = (j, \rho)$ centered around some point with coordinates x^i. Generalizing (B.2.1), the simplest 4-dimensional ansatz for defining the electromagnetic field strength reads:

$$\text{force} \sim \text{field strength} \times \text{charge current}\,. \tag{B.2.2}$$

Also in four dimensions, the *force* $\sim \partial L/\partial x^\alpha$ is represented by the components f_α of a *covector*. Accordingly, the ansatz (B.2.2) can be made more precise:

$$f_\alpha = F_\alpha \wedge J\,. \tag{B.2.3}$$

The components f_α represent a twisted form since they change sign under spatial reflections. The Lagrangian 4-form has the absolute dimension of an action, h. Correspondingly the force, which is understood as the (variational) derivative of the action with respect to the spacetime coordinates, has the dimension

$$[f_\alpha] = h\,[\vartheta^\alpha]^{-1} = (h\,t^{-1}, h\,\ell^{-1}) = (\text{energy}, \text{momentum})\,. \tag{B.2.4}$$

A force cannot possess more than four independent components. Since the current J in (B.2.3) is a twisted 3-form, the only possibility seems to be that f_α is a covector-valued 4-form (recall the general definition given in Sec. A.1.3, example 3). In other words, the covector index transforms under the 4-dimensional linear group as the frame does, namely $f_{\alpha'} = L_{\alpha'}{}^{\alpha} f_\alpha$ (see (A.1.5)). Then f_α has four components, indeed, and assigns to any 4-volume element the components of a covector:

$$f_\alpha = \frac{1}{4!}\, f_{ijkl\,\alpha}\, dx^i \wedge dx^j \wedge dx^k \wedge dx^l\,, \quad f_{ijkl\,\alpha} = f_{[ijkl]\,\alpha}\,. \tag{B.2.5}$$

The anholonomic components of f_α, namely $f_{\mu\nu\rho\sigma\,\alpha}$, determine the relative dimension of f_α. Since $f_{\mu\nu\rho\sigma\,\alpha} = f_{[\mu\nu\rho\sigma]\,\alpha}$, one of the indices μ,ν,ρ,σ has to be a time index and the remaining three space indices. Therefore, we find (W = watt, N = newton),

$$\begin{aligned} [f_{\mu\nu\rho\sigma\,\alpha}] &= h\,[\vartheta^\alpha]^{-1}\,t^{-1}\,\ell^{-3} = (h\,t^{-2}\,\ell^{-3}\,,\, h\,t^{-1}\,\ell^{-4}) \\ &= (\text{power density}\,,\,\text{force density}) \\ &\overset{\text{SI}}{=} (\mathrm{W\,m^{-3}},\, \mathrm{N\,m^{-3}})\,. \end{aligned} \tag{B.2.6}$$

Thus we have identified the f_α in (B.2.3) indeed as 4-force density, the spatial components of which carry the relative dimension of $f\ell^{-3}$. This shows once more that the assignment of the dimensions in (B.2.4) was correct. As a consequence, the field F_α in (B.2.3) turns out to be an untwisted covector-valued 1-form $F_\alpha = F_{\alpha\beta}\,\vartheta^\beta$ with the coframe ϑ^β. The covector index of F_α transforms as $F_{\alpha'} = L_{\alpha'}{}^{\alpha} F_\alpha$, whereas, as a 1-form, we have $F_\alpha = F_{\alpha i}\,dx^i = F_{\alpha\beta}\,\vartheta^\beta$.

As such, $F_{\alpha\beta}$ would have 16 independent components in general. However, we know already that the electric field strength E has 3 independent components. If we expect the analogous to be true for the magnetic field strength B, then $F_{\alpha\beta}$ should carry 6 independent components at most. Then, it is near at hand to assume that $F_{\alpha\beta}$ is antisymmetric: $F_{\alpha\beta} = -F_{\beta\alpha}$. Accordingly, the electromagnetic field strength turns out to be a 2-form $F = \frac{1}{2}\,F_{\alpha\beta}\,\vartheta^\alpha \wedge \vartheta^\beta$ and $F_\alpha = e_\alpha \rfloor F = F_{\alpha\beta}\,\vartheta^\beta$ with the frame e_α. For the absolute dimension of F this yields, if we use (B.2.3), (B.2.4), and $[J] = q$,

$$\begin{aligned} [F] &= [F_\alpha]\,[e_\beta]^{-1} = [f_\alpha]\,[J]^{-1}\,[e_\beta]^{-1} = h\,q^{-1}\,[\vartheta^\alpha]^{-1}\,[e_\beta]^{-1} \\ &= h\,q^{-1} \overset{\text{SI}}{=} \mathrm{J\,s\,C^{-1}} = \mathrm{V\,s} =: \mathrm{Wb}\,. \end{aligned} \tag{B.2.7}$$

In SI, we have the abbreviations J = joule (not to be mixed up with the current 3-form J), V = volt, and Wb = weber.

Finally, the force equation (B.2.3) for a test charge current reads $f_\alpha = (e_\alpha \rfloor F) \wedge J$. Accordingly, we defined the field strength F as the incorporation of the possible forces acting in an electromagnetic field on test charges, thereby relating the mechanical notion of a force to the electromagnetic state around electric charges and currents. In the first axiom, we consider the *active* role of charge that creates the excitation field; here we study its *passive* role, namely which forces act on it in an electromagnetic environment. This concludes our heuristic considerations and we now turn to our new axiom.

B.2.2 Second axiom relating mechanics and electrodynamics

We have then

$$f_\alpha = (e_\alpha \rfloor F) \wedge J \qquad \text{(second axiom)}. \tag{B.2.8}$$

The *untwisted electromagnetic field strength 2-form* F carries the absolute dimension of a magnetic flux: $[F] = h\, q^{-1} \overset{\text{SI}}{=} \text{Wb}$. This equation for the Lorentz force density yields an *operational* definition of the electromagnetic field strength F and represents our second axiom of electrodynamics. Observe that (B.2.8), like (B.1.17), is an equation that is free from the metric and the connection; it is defined on any 4-dimensional differentiable manifold. Since $F \wedge J$ is a 5-form, (B.2.8) can alternatively be written as $f_\alpha = -F \wedge (e_\alpha \rfloor J)$.

A decomposition of the 2-form F into "time" and "space" pieces according to

$$F = {}^{\perp}F + \underline{F} = E \wedge d\sigma + B =: (E, B)\,, \tag{B.2.9}$$

yields, in three dimensions, the untwisted electric field strength 1-form E and the untwisted magnetic field strength[1] 2-form B,

$$E := -F_\perp \qquad \text{and} \qquad B := \underline{F}\,. \tag{B.2.10}$$

The choices of the signs in these definitions are related to the Lenz rule, inter alia. They are discussed in section B.4.2. The absolute and relative dimensions

[1]The historical names are "electric field" for E and "magnetic induction" for B. The ISO names "electric field" and "magnetic flux density," respectively, again don't merge to a 4D name. Note also that the ISO names don't recognize the difference between a twisted and an untwisted form. The 1-forms $\mathcal{H}$ and E are both called "field" and $\mathcal{D}$ and B "flux density," as if they were quantities of the same type. As a matter of fact, the four quantities $\mathcal{H}$, $\mathcal{D}$, E, B, constituting the electromagnetic field, obey four distinctly different transformation laws under diffeomorphisms (coordinate transformations).

of E and B are, respectively,

$$\begin{aligned} [E] &= hq^{-1}t^{-1} \overset{\text{SI}}{=} \mathrm{V}\,, \qquad [B] = hq^{-1} \overset{\text{SI}}{=} \mathrm{Wb}\,, \\ [E_{\hat{a}}] &= [E]\,\ell^{-1} \overset{\text{SI}}{=} \mathrm{V\,m^{-1}}\,, \quad [B_{\hat{a}\hat{b}}] = [B]\,\ell^{-2} \overset{\text{SI}}{=} \mathrm{Wb\,m^{-2}} =: \mathrm{T}\,, \end{aligned} \tag{B.2.11}$$

where T abbreviates tesla. We can decompose the 1-form E and the 2-form B as follows: $E = E_a\, dx^a$ and $B = \frac{1}{2}\, B_{ab}\, dx^a \wedge dx^b = \mathcal{B}^c\, \hat{\epsilon}_c$. For the 2-form $\hat{\epsilon}_c$ in three dimensions, one should compare with (A.1.65). The components E_a and $\mathcal{B}^a = \frac{1}{2}\, \epsilon^{abc} B_{bc}$ correspond to the conventional Cartesian components of the electric and the magnetic field vectors $\vec{E}$ and $\vec{B}$, respectively.

Clearly then, the *electric line tension* (electromotive force or voltage) $\int_{\Omega_1} E$ and the *magnetic flux* $\int_{\Omega_2} B$ must play a decisive role in Maxwell's theory. Hence, as building blocks for laws governing the electric and magnetic field strengths, we have only the electric line tension and the magnetic flux at our disposal.

The Lorentz force density f_α, as a 4-form, that is, as a form of maximal rank ("top dimensional form"), is purely longitudinal with respect to the normal vector n. Thus only its longitudinal piece $(f_\alpha)_\perp$, a covector-valued 3-form, survives. It turns out to be

$$\begin{aligned} k_\alpha &:= (f_\alpha)_\perp = n \rfloor \left[(e_\alpha \rfloor F) \wedge J\right] = (e_\alpha \rfloor E) \wedge J + (e_\alpha \rfloor F) \wedge j \\ &= \rho \wedge (e_\alpha \rfloor E) + j \wedge (e_\alpha \rfloor B) - j \wedge E \wedge (e_\alpha \rfloor d\sigma)\,. \end{aligned} \tag{B.2.12}$$

If we now display its time and space components, we find

$$k_{\hat{0}} = -j \wedge E\,, \qquad [k_{\hat{0}}] = ht^{-2} \overset{\text{SI}}{=} \mathrm{W}\,, \tag{B.2.13}$$

$$k_a = \rho \wedge (e_a \rfloor E) + j \wedge (e_a \rfloor B)\,, \quad [k_a] = h\ell^{-1}t^{-1} \overset{\text{SI}}{=} \mathrm{N}\,. \tag{B.2.14}$$

The zero in $k_{\hat{0}}$ carries a circumflex since it is an anholonomic index. The relative dimensions pick up an additional factor of $\ell^{-3} \overset{\text{SI}}{=} \mathrm{m}^{-3}$. The time component $k_{\hat{0}}$ represents the *electric power density*, the space components k_a the 3-dimensional Lorentz force density.

Let us come back to the Lorentz force density f_α. We consider now a *point particle* with net electric charge q and the velocity vector $u = u^\alpha e_\alpha$. Its electric current J is a 3-form. Hence we define the flow 3-form of the particle as

$$U := u \rfloor \hat{\epsilon} = u^\alpha e_\alpha \rfloor \hat{\epsilon} = u^\alpha \hat{\epsilon}_\alpha \qquad \text{or} \qquad u = {}^\diamond(\vartheta^\alpha \wedge U)\, e_\alpha\,, \tag{B.2.15}$$

with the volume 4-form $\hat{\epsilon}$ and the 3-form $\hat{\epsilon}_\alpha$ of (A.1.74). Clearly, the flow 3-form U has the same number of independent components as the velocity vector u. Then the electric current of the point particle reads

$$J = q\,U = q\,u \rfloor \hat{\epsilon}\,. \tag{B.2.16}$$

We substitute this into the second axiom (B.2.8),

$$\begin{aligned} f_\alpha &= q(e_\alpha \rfloor F) \wedge u \rfloor \hat{\epsilon} \\ &= -q\, u \rfloor \left[(e_\alpha \rfloor F) \wedge \hat{\epsilon}\right] + q(u \rfloor e_\alpha \rfloor F)\, \hat{\epsilon}\,. \end{aligned} \tag{B.2.17}$$

Since $\hat{\epsilon}$ is a 4-form, the expression in the square bracket is a 5-form and vanishes. Then the antisymmetry of the interior product yields:

$$f_\alpha = (-q\, e_\alpha \rfloor u \rfloor F)\, \hat{\epsilon} =: \mathcal{F}_\alpha\, \hat{\epsilon}\,. \tag{B.2.18}$$

Consequently, for the components $\mathcal{F}_\alpha$ of the 1-form $\mathcal{F}$, we have

$$\mathcal{F}_\alpha = {}^\diamond f_\alpha = -q\, u^\beta e_\alpha \rfloor e_\beta \rfloor F = -q\, u^\beta F_{\beta\alpha} = q\, F_{\alpha\beta} u^\beta\,, \tag{B.2.19}$$

whereas the 1-form itself reads

$$\mathcal{F} = \mathcal{F}_\alpha \vartheta^\alpha = -q\, u \rfloor F\,, \qquad \text{with} \qquad u \rfloor \mathcal{F} = 0\,. \tag{B.2.20}$$

This is the point particle version of the Lorentz force, which is in accord with our starting point in (B.2.1). In contrast, we chose as a second axiom (B.2.8) the field-theoretical version of the Lorentz force, namely its density, since only in that version does there appear the current 3-form J of the first axiom.

If there is an electromagnetic field configuration such that the Lorentz force density vanishes ($f_\alpha = 0$), we call it a *force-free electromagnetic* field:

$$(e_\alpha \rfloor F) \wedge dH = 0\,. \tag{B.2.21}$$

Here we have substituted the inhomogeneous Maxwell equation.

In plasma physics such configurations play a decisive role if restricted purely to the magnetic field. We (1 + 3)-decompose (B.2.21). A look at (B.2.13) and (B.2.14) shows that for the magnetic field only the space components (B.2.14) of the Lorentz force density matter:

$$k_a = \underline{d}\mathcal{D} \wedge (e_a \rfloor E) - \dot{\mathcal{D}} \wedge (e_a \rfloor B) + \underline{d}\mathcal{H} \wedge (e_a \rfloor B)\,. \tag{B.2.22}$$

If the electric excitation and its time derivative vanish ($\mathcal{D} = 0$, $\dot{\mathcal{D}} = 0$) — clearly a frame dependent, i.e., noncovariant statement — then we find for the *force-free magnetic* field[2]

$$(e_a \rfloor B) \wedge \underline{d}\mathcal{H} = 0 \tag{B.2.23}$$

or, in components,

$$\mathcal{B}^b\, \partial_{[a} \mathcal{H}_{b]} = 0\,. \tag{B.2.24}$$

[2]See Lundquist [26] and Lüst & Schlüter [27]. However, they as well as later authors put $B = \mu_0\, {}^\star\mathcal{H}$ right away and start with $\vec{\mathcal{H}} \times \nabla \times \vec{\mathcal{H}} = 0$.

B.2.3 $^{\otimes}$The first three invariants of the electromagnetic field

The first axiom supplied us with the fields J (twisted 3-form) and H (twisted 2-form) and the second axiom, additionally, with F (untwisted 2-form). Algebraically we can construct therefrom the so-called first invariant of the electromagnetic field,

$$I_1 := F \wedge H\,, \qquad [I_1] = h\,. \tag{B.2.25}$$

It is a twisted 4-form or, equivalently, a scalar density of weight $+1$ with one independent component. Clearly, I_1 could qualify as a Lagrange 4-form: It is a twisted form and it has the appropriate dimension.

Furthermore, a second invariant can be assembled,

$$I_2 := F \wedge F\,, \qquad [I_2] = (h/q)^2\,, \tag{B.2.26}$$

an untwisted 4-form with the somewhat strange dimension of the square of a magnetic flux, and a third one,

$$I_3 := H \wedge H\,, \qquad [I_3] = q^2\,, \tag{B.2.27}$$

equally an untwisted form.

In order to get some insight into the meaning of these 4-forms, we substitute the (1 + 3)-decompositions (B.1.39) and (B.2.9):

$$\begin{aligned} I_1 = F \wedge H &= d\sigma \wedge (B \wedge \mathcal{H} - E \wedge \mathcal{D})\,, && \text{(B.2.28)}\\ I_2 = F \wedge F &= -2\, d\sigma \wedge B \wedge E, && \text{(B.2.29)}\\ I_3 = H \wedge H &= 2\, d\sigma \wedge \mathcal{H} \wedge \mathcal{D}\,. && \text{(B.2.30)} \end{aligned}$$

If for an electromagnetic field configuration the first invariant vanishes, then we find,

$$I_1 = 0 \qquad \text{or} \qquad \frac{1}{2} B \wedge \mathcal{H} = \frac{1}{2} E \wedge \mathcal{D}\,. \tag{B.2.31}$$

As we will see in (B.5.61), this means that the magnetic energy density equals the electric energy density. Similarly, for the second invariant, we have

$$I_2 = 0 \qquad \text{or} \qquad B \wedge E = 0\,. \tag{B.2.32}$$

Then the electric field strength 1-form can be called "parallel" to the magnetic field strength 2-form. Analogously, for the vanishing third invariant,

$$I_3 = 0 \qquad \text{or} \qquad \mathcal{H} \wedge \mathcal{D} = 0\,, \tag{B.2.33}$$

the excitations are "parallel" to each other. It should be understood that the characterizations $I_1 = 0$ and so on are (diffeomorphism and frame) invariant statements about electromagnetic field configurations.

This is as far as we can go by algebraic manipulations. If we differentiate, we can construct dJ, dH, dF. The first two expressions are known, $dJ = 0$ and $dH = J$, whereas the last one, dF, is left open so far. We turn to it in the next chapter.

Let us collect our results in a table:

Table B.2.1: Invariants of the electromagnetic field in 4D.

Invariant	**dimension**	**un-/twisted**	**name**
$I_1 = F \wedge H$	h	twisted	$\sim$ Lagrangian
$I_2 = F \wedge F$	$(h/q)^2$	untwisted	Chern 4-form
$I_3 = H \wedge H$	q^2	untwisted	...
$I_4 = A \wedge J$	h	twisted	coupling term in Lagr.

The fourth invariant I_4 and the names are explained in Sec. B.3.3. The I_k's are 4-forms, respectively, with one independent component each. We call them invariants. With the diamond operator $^\diamond$, the dual with respect to the Levi-Civita epsilon (see (A.1.78) at the end of Sec. A.1.9), we can attach to each 4-form I_k a (metric-free) scalar *density* $^\diamond I_k$.

B.3
Magnetic flux conservation

B.3.1 Third axiom

The spacetime manifold, which underlies our consideration and which has been defined at the beginning of Sec. B.1.2, is equipped with the property of an orientation. Then we can integrate the untwisted 2-form F in four dimensions over a 2-dimensional surface. Since F is a 2-form, the simplest invariant statement that comes to mind reads

$$\oint_{C_2} F = 0\,, \qquad \partial C_2 = 0 \qquad \text{(third axiom)}\,, \tag{B.3.1}$$

for any closed 2-dimensional submanifold $C_2 \subset X_4$. Indeed, this is the axiom we are looking for.

It is straightforward to find supporting evidence for (B.3.1) by using the decomposition (B.2.9). Faraday's induction law results from (B.3.1) if one chooses the 2-dimensional surface as $C_2' = \partial\Omega'{}_3$ with $\Omega'{}_3 = [\sigma_0, \sigma] \times \Omega_2'$, where Ω_2' is represented in Fig. B.3.1 by a line in h_{σ_0}; i.e., it is transversal to the vector field n:

$$\oint_{\partial\Omega_2'} E + \frac{d}{d\sigma}\int_{\Omega_2'} B = 0\,. \tag{B.3.2}$$

What is usually called the law of the absence of magnetic charge we find by again choosing the 2-dimensional surface C_2 in (B.3.1) as the boundary of a

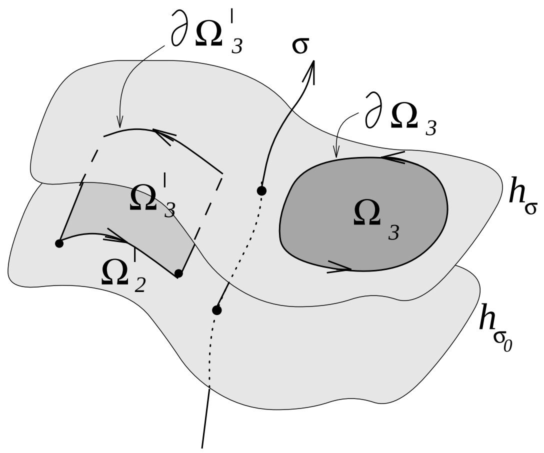

Figure B.3.1: Different 2-dimensional periods of the magnetic flux integral. The 3-dimensional integration region Ω_3' lies transversal to the hypersurface h_σ whereas Ω_3 lies in it.

3-dimensional submanifold Ω_3 that lies in one of the folia h_σ (see Fig. B.3.1):

$$\oint_{\partial\Omega_3 \subset h_\sigma} B = 0\,. \tag{B.3.3}$$

The proofs are analogous to those given in (B.1.13). They are given below on a differential level.

Magnetic flux $\int_{\Omega_2} B$ and its conservation are of central importance to electrodynamics. At low temperatures, certain materials can become superconducting; i.e., they lose their electrical resistance. At the same time, if they are exposed to an external (sufficiently weak) magnetic field, the magnetic field is expelled from their interior except for a thin layer at their surface (Meissner–Ochsenfeld effect). In the case of a superconductor of type II (niobium, for example), provided the external field is higher than a certain critical value, *quantized* magnetic flux lines carrying a flux quantum of[1] $\Phi_0 := \mathrm{h}/(2e) \overset{\mathrm{SI}}{\approx} 2.068 \times 10^{-15}\,\mathrm{Wb}$ can penetrate the surface of the superconductor and build up a triangular lattice, an Abrikosov lattice. A cross section of such a flux-line lattice is depicted in Fig. B.3.2; a schematic view is provided in Fig. B.3.3. In high-temperature superconductors half-integer flux quanta have also been observed.[2] What is important for us is that, at least under certain circumstances, single quantized flux lines can be *counted* and that they behave like a *conserved* quantity; i.e., they migrate but are not spontaneously created nor destroyed. The counting

[1]Here h is Planck's constant and e the elementary charge.
[2]See the discussion of Tsuei and Kirtley [50].

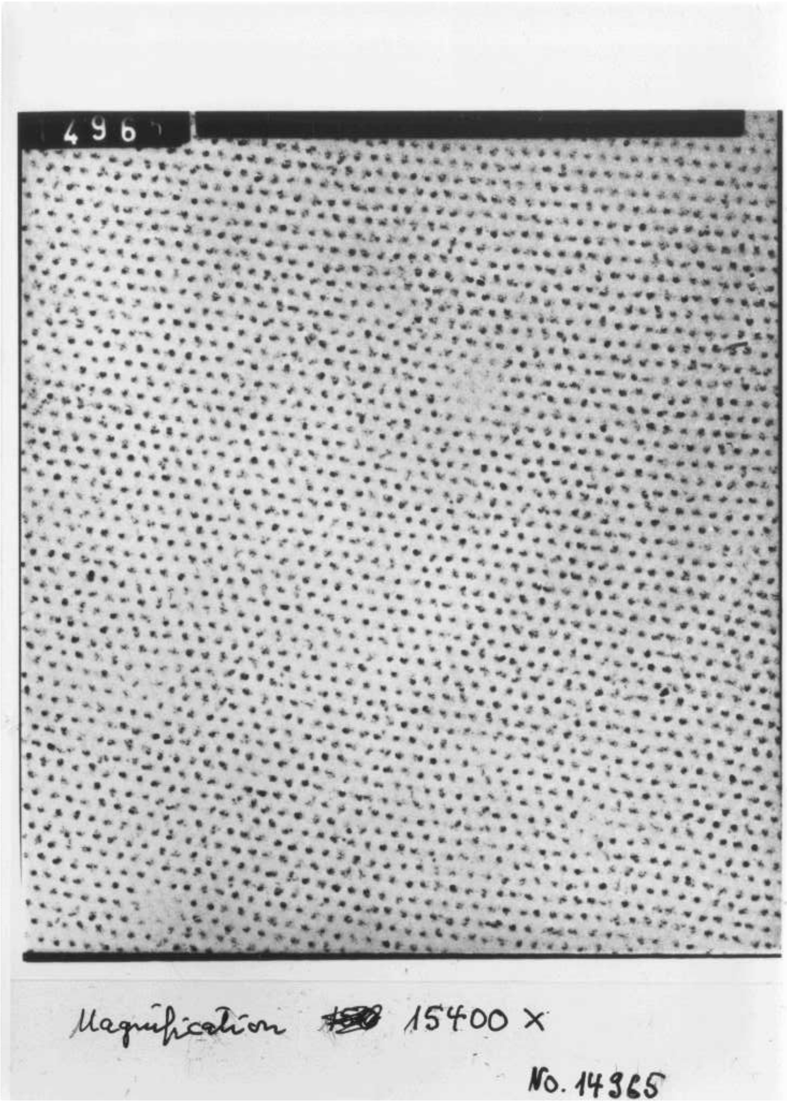

Figure B.3.2: Direct observation of individual flux lines in type II superconductors according to Essmann & Träuble [7, 8]. The image shown here belongs to a small superconducting Niobium disc (diameter 4 mm, thickness 1 mm) which, at a temperature of 1.2 K, was exposed to an external magnetic excitation of $\mathcal{H} = 78\,\mathrm{kA/m}$. At the surface of the disc the flux lines were decorated by small ferromagnetic particles that were fixed by a replica technique. Eventually the replica was observed by means of an electron microscope. The parameter of the flux-line lattice was 170 nm (courtesy of U. Essmann).

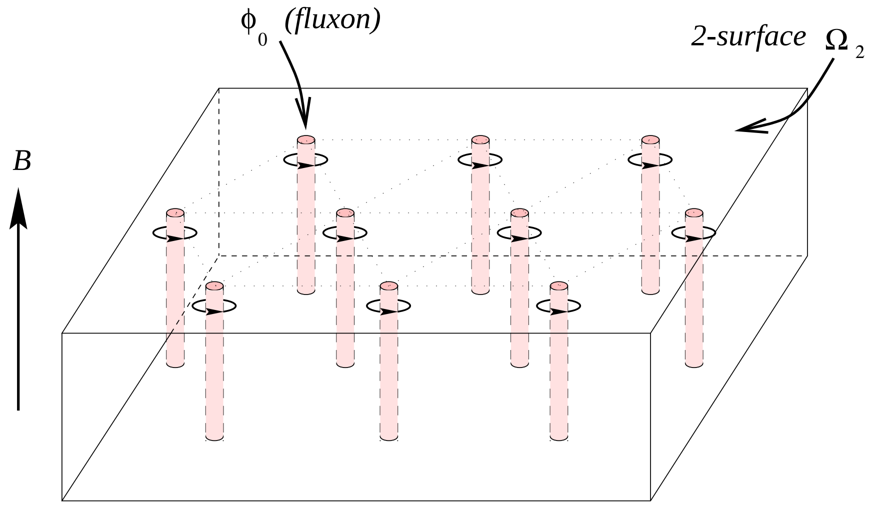

Figure B.3.3: Sketch of an Abrikosov lattice in a type II superconductor in 3-dimensional space.

argument supports the view that magnetic flux is determined in a metric-free way, the migration argument, even if this is somewhat indirect, that the flux is conserved. The computer simulation of the magnetic field of the Earth in Fig. B.3.4 may give one an intuitive feeling that the magnetic lines of force, which can be understood as unquantized magnetic flux lines, are controlled by the induction law and, at the same time, are close to our visual perception of magnetic field configurations.

B.3.2 Electromagnetic potential

In (B.3.1) we specialize to the case when C_2 is a 2-boundary $C_2 = \partial\Omega_3$ of an arbitrary 3-dimensional domain Ω_3. Then, by Stokes' theorem, we find

$$\int_{\Omega_3} dF = 0\,. \tag{B.3.4}$$

Since Ω_3 can be chosen arbitrarily, the electromagnetic field strength turns out to be a closed form:

$$dF = 0\,. \tag{B.3.5}$$

This is the 4-dimensional version of the set of homogeneous Maxwell equations.

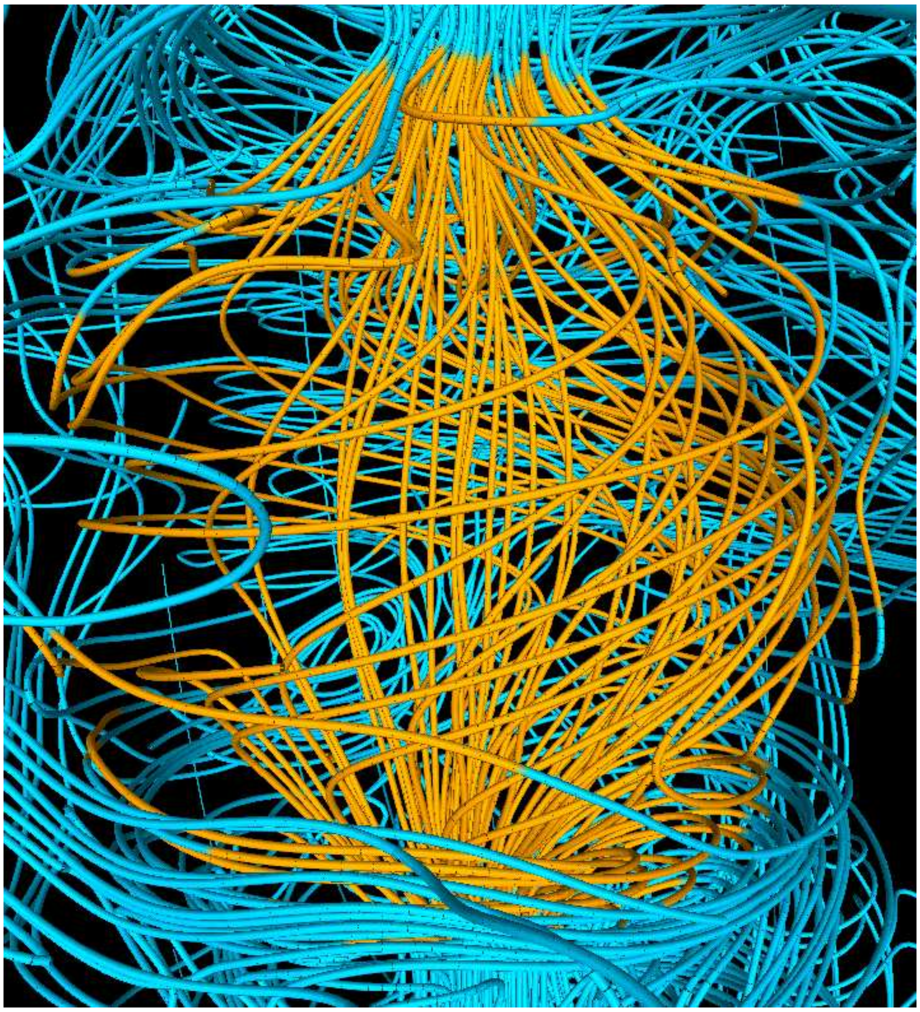

Figure B.3.4: Figure of G. A. Glatzmaier: Snapshot of magnetic lines of force in the core of our computer-simulated Earth. Lines are in gold (blue) where they are inside (outside) the inner core. The axis of rotation is vertical in this image. The field is directed inward at the inner core north pole (top) and outward at the south pole (bottom); the maximum magnetic field strength is about 30 mT (see [14] and [43]). In color you can look at such pictures at the website http://www.es.ucsc.edu/~glatz/geodynamo.html.

The axiom (B.3.1) now shows that all periods of F are zero. Consequently the field strength is an exact form

$$F = d\,A\,. \tag{B.3.6}$$

Equation (B.3.5) is implied by (B.3.6) because of $dd = 0$. However, the inverse statement that (B.3.5) implies (B.3.6) does not hold in a global manner unless the conditions for the first de Rham theorem are met.

The untwisted electromagnetic potential 1-form A has the dimension of $h\,q^{-1}$. Decomposed in "time" and "space" pieces, it reads

$$A = -\varphi\, d\sigma + \mathcal{A}\,, \tag{B.3.7}$$

with

$$\varphi := -A_{\perp} \qquad \text{and} \qquad \mathcal{A} := \underline{A}\,. \tag{B.3.8}$$

Here we recover the familiar 3-dimensional scalar and covector potentials φ and $\mathcal{A}$, respectively. The potential is only determined up to a closed 1-form

$$A \longrightarrow A + \chi\,, \qquad d\,\chi = 0\,, \tag{B.3.9}$$

a fact that has far-reaching consequences for the quantization of the electromagnetic field. We decompose (B.3.6) and find straightforwardly

$$^{\perp}F = {}^{\perp}(dA) \qquad \text{or} \qquad E = -\underline{d}\varphi - \dot{\mathcal{A}} \tag{B.3.10}$$

and

$$\underline{F} = \underline{dA} \qquad \text{or} \qquad B = \underline{d}\mathcal{A}\,. \tag{B.3.11}$$

A decomposition of (B.3.5) by means of (B.1.28) and (B.2.10) yields the homogeneous set of Maxwell's equations,

$$^{\perp}(dF) = d\sigma \wedge \left(\pounds_n \underline{F} - \underline{d}F_{\perp}\right) = 0\,, \qquad \underline{dF} = 0\,, \tag{B.3.12}$$

or in 3-dimensional notation,

$$\underline{d}\,E + \dot{B} = 0 \tag{B.3.13}$$

and

$$\underline{d}\,B = 0\,, \tag{B.3.14}$$

respectively. If we integrate these equations over a 2- or 3-dimensional volume and apply Stokes' theorem, we find (B.3.2) and (B.3.3). Again, by analogy with the inhomogeneous equations, (B.3.13) and (B.3.14) represent an evolution equation and a constraint, respectively.

B.3.3 $^{\otimes}$Abelian Chern–Simons and Kiehn 3-forms

For the electromagnetic theory, we now have the 1-form A as a new building block. This has an immediate consequence for the second invariant, namely

$$I_2 = (dA) \wedge F = d(A \wedge F) = d(A \wedge dA)\,, \tag{B.3.15}$$

since $dF = 0$. In other words, I_2, the so-called Abelian *Chern* 4-form, is an *exact* form and, accordingly, cannot be used as a nontrivial Lagrangian. We read off from (B.3.15) the untwisted Abelian *Chern–Simons* 3-form

$$C_A := A \wedge F\,, \qquad [C_A] = (h/q)^2\,, \tag{B.3.16}$$

which has a topological meaning in that it is related to the winding number of the so-called $U(1)$ vector bundle of electromagnetism. Also in all other dimensions with $n \geq 3$, the Abelian Chern–Simons form is represented by a 3-form. If we $(1+3)$-decompose C_A, we find

$$C_A = \mathcal{A} \wedge B + d\sigma \wedge (\varphi\, B + \mathcal{A} \wedge E)\,, \tag{B.3.17}$$

which includes the so-called magnetic helicity density[3] $\mathcal{A} \wedge B$, which is also a 3-form (however in 3D) with one independent component. Obviously,

$$dC_A = F \wedge F\,. \tag{B.3.18}$$

Consequently, even though C_A is not gauge invariant, its differential dC_A is gauge invariant.

What we just did to the second invariant, we can now implement in an analogous way for the first invariant:

$$I_1 = (dA) \wedge H = d(A \wedge H) + A \wedge J\,. \tag{B.3.19}$$

We define the *twisted Kiehn* 3-form[4]

$$K := A \wedge H\,, \qquad [K] = h\,. \tag{B.3.20}$$

It carries the dimension of an action. In n dimensions, it would be an $(n-1)$-form — in sharp contrast to the Abelian Chern–Simons 3-form. If we decompose K into 1 + 3, we find

$$K = \mathcal{A} \wedge \mathcal{D} + d\sigma \wedge (-\varphi\, \mathcal{D} + \mathcal{A} \wedge \mathcal{H})\,. \tag{B.3.21}$$

The 2-form $\mathcal{A} \wedge \mathcal{H}$ is the purely magnetic piece of K in 3D. By differentiating K, or by rearranging (B.3.19), we find

$$dK = F \wedge H - A \wedge J\,. \tag{B.3.22}$$

---

[3]The *magnetic helicity* is given by the integral $\int_V \mathcal{A} \wedge B$ (see Moffatt [33], Marsh [28, 29], and Rañada and Trueba[40, 41, 49]).

[4]See Kiehn and Pierce [24, 22, 23].

Even though the Kiehn 3-form changes under a gauge transformation $A \to A + d\psi$, its exterior derivative dK, provided we are in the free-field region with $J = 0$, is gauge *in*variant, as can be seen in (B.3.22).

From (B.3.22), we can read off the new twisted electromagnetic interaction 4-form

$$I_4 := A \wedge J\,, \qquad [I_4] = h\,. \tag{B.3.23}$$

It decomposes according to

$$A \wedge J = d\sigma \wedge (\varphi\,\rho + \mathcal{A} \wedge j)\;. \tag{B.3.24}$$

Because $dJ = 0$, it is gauge invariant up to an exact form:

$$I_4 \quad \longrightarrow \quad I_4 + (d\psi) \wedge J = I_4 + d(\psi \wedge J)\,. \tag{B.3.25}$$

Accordingly I_4, like I_1, also qualifies as a piece of an electrodynamic Lagrangian 4-form since it is twisted, of dimension h, and gauge invariant (up to an irrelevant exact form). Note, however, that dK in (B.3.22), as an exact form, cannot feature as a Lagrangian in 4D even though both pieces on the right-hand side of (B.3.22) independently have the correct behavior. In other words, the relative factor beween I_1 and I_4 is inappropriate for a Lagrange 4-form.

Let us put our results together. If we start with the 2-form H and differentiate and multiply, then we find the sequence $(H, dH, H \wedge H)$. We cannot go any further since forms with a rank $p > 4$ will vanish identically. Similarly, for the 1-form A, we have $(A, dA, A \wedge dA, dA \wedge dA)$ and, if we mix both, the sequence $(A \wedge H, dA \wedge H, A \wedge dH)$. Note that $A \wedge A \equiv 0$ since A is a 1-form. In this way, we create the new 3-forms K (see (B.3.20)) and C_A (see (B.3.16)):

Table B.3.1: Three-forms of the electromagnetic field in 4D.

3-form	**dimension**	**un-/twisted**	**name**
$K = A \wedge H$	h	twisted	Kiehn
$C_A = A \wedge F$	$(h/q)^2$	untwisted	Chern–Simons
J	q	untwisted	electric current

By the same token, the four different invariants I_k of the electromagnetic field arise as 4-forms (see Table B.2.1, p. 123).

B.3.4 Measuring the excitation

The electric excitation $\mathcal{D}$ can be operationally defined using the Gauss law (B.1.45). Put at some point P in free space two small, thin, and electrically conducting (metal) plates of arbitrary shape with insulating handles ("Maxwellian

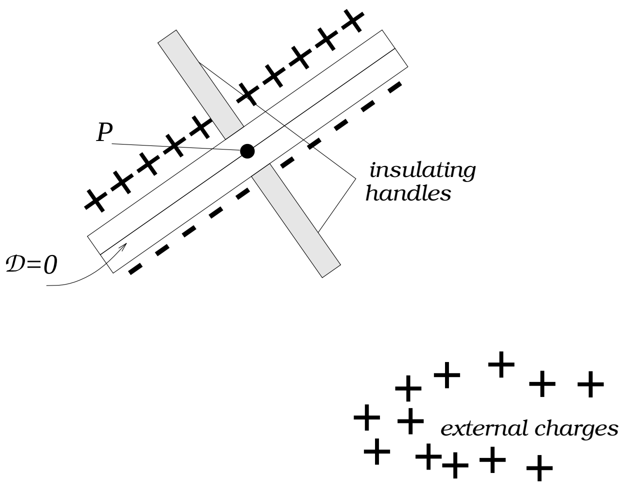

Figure B.3.5: Measurement of $\mathcal{D}$ at a point P.

double plates") (see Fig. B.3.5). Suppose we choose local coordinates (x^1, x^2, x^3) in the neighborhood of P such that $P = (0, 0, 0)$. The vectors ∂_a, with $a = 1, 2, 3$, span the tangent space at P. Press the plates together at P and orient them in such a way that their common boundary is given by the equation $x^3 = 0$ (and thus (x^1, x^2) are the local coordinates on each plate's surface). Separate the plates and measure the net charge $\mathcal{Q}$ and the area S of that plate for which the vector ∂_3 points outwards from its boundary surface. Determine experimentally $\lim_{S\to 0} \mathcal{Q}/S$ as well as possible. Then

$$\mathcal{D}_{12} = \lim_{S\to 0} \frac{\mathcal{Q}}{S}\,. \tag{B.3.26}$$

Technically, one can realize the limiting process by constructing the plates in the form of parallelograms with the sides $a\,\partial_1$ and $a\,\partial_2$. Then (B.3.26) is replaced by

$$\mathcal{D}_{12} = \lim_{a\to 0} \frac{\mathcal{Q}}{a^2}\,. \tag{B.3.27}$$

Repeat this measurement with the two other possible orientations of the plates (be careful to choose one of the two plates in accordance with the orientation prescription above). Then, similarly, one finds $\mathcal{D}_{23}$ and $\mathcal{D}_{31}$. Thus finally the electric excitation is measured:

$$\mathcal{D} = \frac{1}{2}\mathcal{D}_{ab} dx^a \wedge dx^b = \mathcal{D}_{12} dx^1 \wedge dx^2 + \mathcal{D}_{23} dx^2 \wedge dx^3 + \mathcal{D}_{31} dx^3 \wedge dx^1\,. \tag{B.3.28}$$

The dimension of $\mathcal{D}$ is that of a *charge*, i.e. $[\mathcal{D}] = q$; for its components we have $[\mathcal{D}_{ab}] = q\,\ell^{-2}$.

Let us prove the correctness of this prescription. The double plates are assumed to be *ideal conductors*. If an external field is applied to an ideal conductor, a redistribution of the carriers of the electric charges (usually electrons) takes place in its interior by means of flowing electric currents such that the Lorentz force acting on each "electron" finally vanishes. This was found in experiments and later used to define an ideal conductor. In the end, no internal currents are flowing any more and, relative to the conductor, the charge distribution is at rest. Therefore the vanishing 3D Lorentz force density (B.2.14) implies that the electric field E inside the conductor has to vanish as well:

$$E\Big|_{\text{inside}} = 0. \tag{B.3.29}$$

The vanishing of E defines a unique electrodynamical state in the conductor. In particular, a certain electric charge distribution is specified. If we assume that (B.3.29) implies

$$\mathcal{D}\Big|_{\text{inside}} = 0, \tag{B.3.30}$$

then this charge distribution, as we will see in a moment, determines $\mathcal{D}$ uniquely: We can read off from the charge distribution the electric excitation.[5] That (B.3.29) has (B.3.30) as a consequence is a rudiment of the spacetime relation that links F and H and that makes the excitation field $\mathcal{D}$ unique. For that reason we have postponed this discussion until now, since knowledge of the notion of the field strength E is necessary for it. By (B.3.30), we *selected* from the allowed class (B.1.21) of the excitations that $\mathcal{D}$ that is measured by the double plates.

In a more general setting, let us consider the electromagnetic excitation $\mathcal{D}$ near a 2-dimensional boundary surface $S \subset h_\sigma$ that separates two parts of space filled with two different types of matter. Choose a point $P \in S$. Introduce local coordinates (x^1, x^2, x^3) in the neighborhood of P in such a way that $P = (0,0,0)$ while (x^1, x^2) are the coordinates on S (see Fig. B.3.6). Let V be a 3-dimensional domain that is half (denoted V_1) in one medium and half (V_2) in another, with the two halves $V_{1,2}$ spanned by the triples of vectors $(\partial_1, \partial_2, a\partial_3)$ and $(\partial_1, \partial_2, -a\partial_3)$, respectively. For definiteness, we assume that ∂_3 points from medium 1 to medium 2. Now we consider the Gauss law (B.1.45) in this domain. Integrate (B.1.45) over V and take the limit $a \to 0$. The result is

$$\int\limits_{\Delta S_{12}} \mathcal{D}_{(2)} - \int\limits_{\Delta S_{12}} \mathcal{D}_{(1)} = \lim_{a\to 0} \int\limits_{V} \rho\,, \tag{B.3.31}$$

where $\mathcal{D}_{(1)}$ and $\mathcal{D}_{(2)}$ denote the values of the electric excitations in the medium 1 and 2, respectively, while ΔS_{12} is a piece of S spanned by (∂_1, ∂_2).

Despite the fact that the volume V clearly goes to zero, the right-hand side of (B.3.31) is nontrivial when there is a *surface charge* exactly on the boundary

[5] The only really appropriate discussion of the definitions of $\mathcal{D}$ and E that we know of is given by Pohl [36].

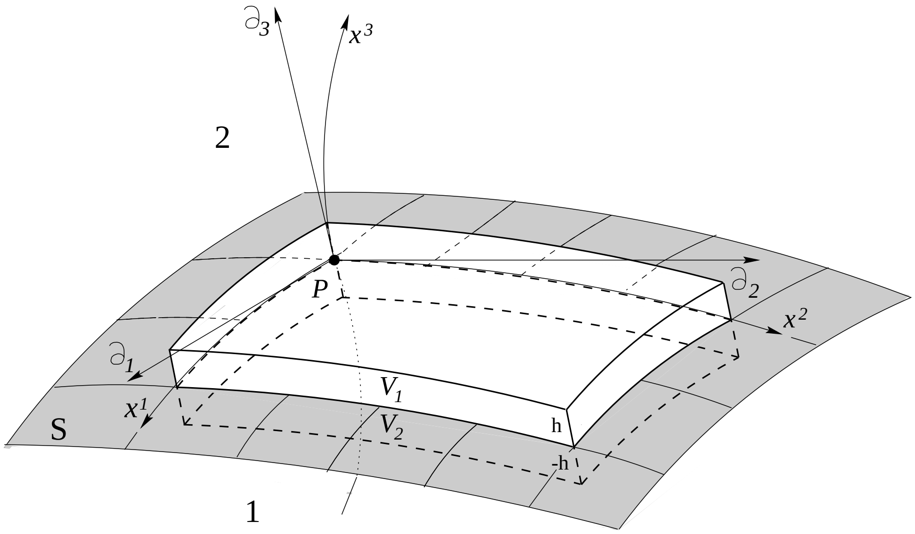

Figure B.3.6: Electric excitation on the boundary between two media.

S. Mathematically, in local coordinates (x^1, x^2, x^3), this can be described by the δ-function structure of the charge density:

$$\rho = \rho_{123}\, dx^1 \wedge dx^2 \wedge dx^3 = \delta(x^3)\, \widetilde{\rho}_{12}\, dx^1 \wedge dx^2 \wedge dx^3 \,. \tag{B.3.32}$$

Substituting (B.3.32) into (B.3.31), we then find

$$\int\limits_{\Delta S_{12}} \mathcal{D}_{(2)} - \int\limits_{\Delta S_{12}} \mathcal{D}_{(1)} = \int\limits_{\Delta S_{12}} \widetilde{\rho}\,, \tag{B.3.33}$$

where $\widetilde{\rho} = \widetilde{\rho}_{12}\, dx^1 \wedge dx^2$ is the *2-form of the electric surface charge density.* Since P and ΔS_{12} are arbitrary, we conclude that on the separating surface S between two media, the electric excitation satisfies

$$\mathcal{D}_{(2)}\big|_S - \mathcal{D}_{(1)}\big|_S = \widetilde{\rho}. \tag{B.3.34}$$

Returning to the measurement process with plates, we have $\mathcal{D}_{(1)} = 0$ inside an ideal conductor, and (B.3.34) justifies the definition of the electric excitation as the charge density on the surface of the plate (B.3.26). Thereby we recognize that the electric excitation $\mathcal{D}$ *is,* by its very definition, the ability to separate charges on (ideally conducting) double plates.

A similar result holds for the magnetic excitation $\mathcal{H}$. Analogously to the small 3-dimensional domain V, let us consider a 2-dimensional domain Σ that is half (Σ_1) in one medium and half (Σ_2) in another one, with the two halves $\Sigma_{1,2}$

spanned by the vectors $(l, a\partial_3)$ and $(l, -a\partial_3)$, respectively. Here $l = l^1\partial_1 + l^2\partial_2$ is an arbitrary vector tangent to S at P. Integrating the Maxwell equation (B.1.44) over Σ, and taking the limit $a \to 0$, we find

$$\int_l \mathcal{H}_{(2)} - \int_l \mathcal{H}_{(1)} = \lim_{a\to 0} \int_\Sigma j. \tag{B.3.35}$$

The second term on the left-hand side of (B.1.44) has a zero limit because $\dot{\mathcal{D}}$ is continuous and finite in the domain of a loop that contracts to zero. However, the right-hand side of (B.3.35) produces a nontrivial result when there are surface currents flowing exactly on the boundary surface S. Analogously to (B.3.32), this is described by

$$j = j_{13}\, dx^1 \wedge dx^3 + j_{23}\, dx^2 \wedge dx^3 = \delta(x^3)\, \widetilde{j}_a\, dx^a \wedge dx^3. \tag{B.3.36}$$

Substituting (B.3.36) into (B.3.35), we find

$$\int_l \mathcal{H}_{(2)} - \int_l \mathcal{H}_{(1)} = \int_l \widetilde{j}, \tag{B.3.37}$$

and, since l is arbitrary, eventually:

$$\mathcal{H}_{(2)}\Big|_S - \mathcal{H}_{(1)}\Big|_S = \widetilde{j}. \tag{B.3.38}$$

Thus we see that on the boundary surface between the two media the magnetic excitation is directly related to the *1-form of the electric surface current density* $\widetilde{j}$.

To measure $\mathcal{H}$, following a suggestion by Zirnbauer,[6] we can use the Meissner effect. Because of this effect, the magnetic field B is driven out of the superconductor. Take a thin superconducting wire (since the wire is pretty cold, perhaps around 10 K, "taking" is not to be understood too literally) and put it at the point P where you want to measure $\mathcal{H}$ (see Fig.B.3.7(a)). Because of the Meissner effect, the magnetic field B and, if we assume analogously to the electric case that $B = 0$ implies $\mathcal{H} = 0$, the excitation $\mathcal{H}$ is expelled from the superconducting region (apart from a thin surface layer of some 10 nm, where $\mathcal{H}$ can penetrate). According to the Oersted–Ampère law, (B.1.44) $\underline{d}\,\mathcal{H} = j$ (we assume quasi-stationarity in order to be permitted to forget about $\dot{\mathcal{D}}$). The compensation of $\mathcal{H}$ at P can be achieved by surface currents $\widetilde{j}$ flowing around the superconducting wire. These induced surface currents $\widetilde{j}$ are of such a type that they generate an $\mathcal{H}_{\text{ind}}$ that compensates the $\mathcal{H}$ to be measured: $\mathcal{H}_{\text{ind}} + \mathcal{H} = 0$. We have to change the angular orientation such that we eventually find the

[6] See his lecture notes [55]; compare also Ingarden and Jamiołkowski [17].

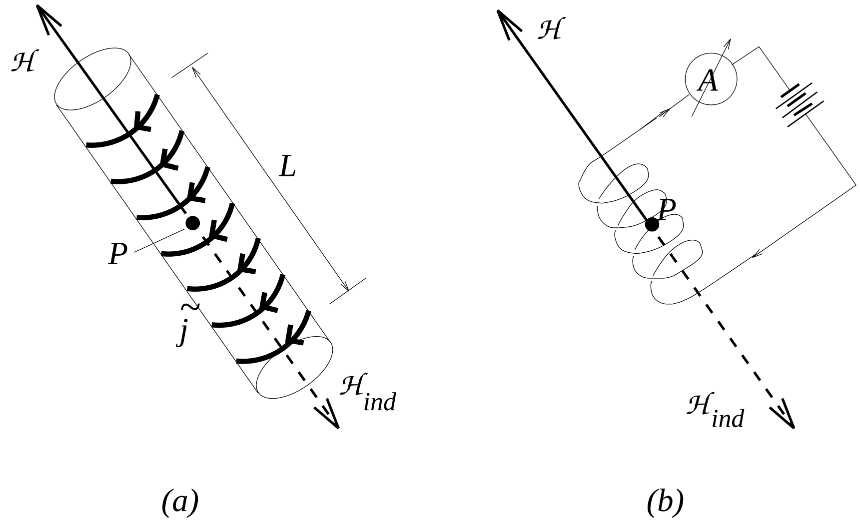

Figure B.3.7: Measurement of $\mathcal{H}$ at a point P. We denote the surface currents by $\tilde{j}$.

maximal current $\tilde{j}_{\max}$. Then, if the x^3-coordinate axis is chosen tangentially to the "maximal orientation," we have

$$\mathcal{H}_3 = \lim_{L\to 0} \frac{\tilde{j}_{\max}}{L}\,, \tag{B.3.39}$$

where L is the length parallel to the wire axis transverse to which the current has been measured. Accordingly, we find

$$\mathcal{H} = \mathcal{H}_3\, dx^3\,. \tag{B.3.40}$$

Clearly the dimension of $\mathcal{H}$ is that of an *electric current*: $[\mathcal{H}] = q\,t^{-1} \overset{\text{SI}}{=} \text{A}$ and $[\mathcal{H}_{\hat{a}}] = q\,\ell^{-1}\,t^{-1} \overset{\text{SI}}{=} \text{A/m}$ ("ampere-turn per meter").

This is admittedly a thought experiment so far. However, the surface currents are experimentally known to exist.[7] Furthermore, with nanotechnological tools it should be possible to realize this experiment. It is conceivable that the surface currents can be measured by means of the Coulomb drag.[8]

Alternatively, one can take a real small *test coil* at P, orient it suitably, and read off the corresponding maximal current at a galvanometer as soon as the

[7] See the corresponding recent magneto-optical studies of Jooss et al. [21].

[8] The idea is the following (M. Zirnbauer, private communication): If in a first thin film a current flows (this represents the surface of the superconductor), then via the Coulomb interaction, it can drag along with it in a separate second conducting thin film the charge carriers which can then be measured (outside the superconductor). This Coulomb drag, which is known to exist (see [44]), requires, however, the breaking of translational invariance.

effective $\mathcal{H}$ vanishes (see Fig. B.3.7(b)). Multiply this current with the winding number of the coil and find $\tilde{\jmath}_{\mathrm{max}}$. Divide by the length of the test coil and you are back to (B.3.39). Whether $\mathcal{H}$ is really compensated for, you can check with a magnetic needle which, in the field-free region, should be in an indifferent equilibrium state. Note that today one has magnetic "needles" of nanometer size.

In this way, we can build up the excitation $H = \mathcal{D} - \mathcal{H} \wedge d\sigma$ as a measurable electromagnetic quantity in its own right. The excitation H, together with the field strength F, we call the "electromagnetic field."

B.4

Basic classical electrodynamics summarized, example

B.4.1 Integral version and Maxwell's equations

We are now in a position to summarize the fundamental structure of electrodynamics in a few lines. According to (B.1.17), (B.2.8), and (B.3.1), the three axioms on a connected, Hausdorff, paracompact, and oriented spacetime read, for any C_3 and C_2 with $\partial C_3 = 0$ and $\partial C_2 = 0$:

$$\oint_{C_3} J = 0\,, \qquad f_\alpha = (e_\alpha \rfloor F) \wedge J\,, \qquad \oint_{C_2} F = 0\,. \tag{B.4.1}$$

The first axiom governs matter and its conserved electric charge, the second axiom links the notion of that charge and the concept of a mechanical force to an operational definition of the electromagnetic field strength. The third axiom determines the flux of the field strength as sourcefree.

In Part C we learn that a metric of spacetime brings in the temporal and spatial distance concepts and a linear connection the inertial guidance field (and thereby parallel displacement). Since the metric of spacetime represents Einstein's *gravitational* potential (and the linear connection is also related to gravitational properties), the three axioms (B.4.1) of electrodynamics are not contaminated by gravitational properties, in contrast to what happens in the usual textbook approach to electrodynamics. A curved metric or a non-flat linear connection do not affect (B.4.1), since these geometric objects don't enter the axioms. Up to now, we could do without a metric and without a connection. Yet we do have the basic Maxwellian structure already at our disposal. Consequently, the structure of electrodynamics that has emerged so far has nothing to do with Poincaré or Lorentz covariance. The transformations involved are diffeomorphisms and frame transformations alone.

If one desires to generalize special relativity to general relativity theory or to the Einstein–Cartan theory of gravity (a viable alternative to Einstein's theory formulated in a non-Riemannian spacetime; see again Part C), then the Maxwellian structure in (B.4.1) is untouched by it. As long as a 4-dimensional connected, Hausdorff, paracompact and oriented differentiable manifold is used as the spacetime, the axioms in (B.4.1) stay covariant and remain the same. In particular, arbitrary frames, holonomic and anholonomic ones, can be used for the evaluation of (B.4.1).

For what reasons are charge conservation $(B.4.1)_1$ and flux conservation $(B.4.1)_3$ such "stable" natural laws? Apparently since the elementary "portions" of charge $e/3$ and of flux $\pi\hbar/e$ (with $e =$ elementary charge and $\hbar =$ (reduced) Planck constant) are additive 4-dimensional scalars that are constant in time and space. In other words,[1]

$$e = \text{const}\,, \quad \hbar = \text{const}\,, \qquad e, \hbar \ \text{ are 4D-scalars}\,, \tag{B.4.2}$$

guarantees that electric charge and magnetic flux are conserved quantities in classical electrodynamics. Therefore (B.4.2) secures the validity of the first and third axioms. A time dependence of e and/or $\hbar$, for example, would dismantle the Maxwellian framework.

According to (B.4.1), the more specialized differential version of electrodynamics (skipping the boundary conditions) reads as follows:

$$\begin{array}{rclcr} dJ &=& 0\,, \qquad f_\alpha = (e_\alpha \rfloor F)\wedge J\,, & & dF = 0\,, \\ J &=& dH\,, & & F = dA\,. \end{array} \tag{B.4.3}$$

This is the structure that we defined by means of our axioms so far. But, in fact, we know a bit more: Because of the existence of conductors and superconductors, we can *measure* the excitation H. Thus, even if H emerges as a kind of potential for the electric current, it is more than that: It is measurable. This is in clear contrast to the potential A that is not measurable. Thus we have to take the gauge invariance under the substitution

$$A' = A + \chi\,, \quad \text{with} \quad d\chi = 0\,, \tag{B.4.4}$$

very seriously, since measurable quantities must be gauge invariant.

The system (B.4.3) can be straightforwardly translated into the Excalc language: We denote the electromagnetic potential A by `pot1`. The "1" we wrote in order to remember better that the potential is a 1-form. The field strength F is written as `farad2`, i.e., as the Faraday 2-form. The excitation H is named `excit2`, and the left-hand sides of the homogeneous and the inhomogeneous Maxwell equations are called `maxhom3` and `maxinh3`, respectively. Then we need

[1] See the corresponding discussion of Peres [35].

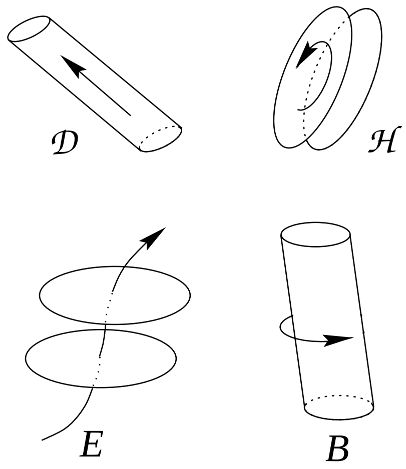

Figure B.4.1: Faraday–Schouten pictograms of the electromagnetic field in 3-dimensional space. The images of 1-forms are represented by two neighboring planes. The nearer the planes, the stronger the 1-form is. The 2-forms are pictured as flux tubes. The thinner the tubes, the stronger the flow. The difference between a twisted and an untwisted form accounts for the two different types of 1- and 2-forms, respectively.

the electric current density `curr3` and the left-hand side of the continuity equation `cont4`. For the first axiom, we have

```
pform cont4=4, curr3=3, maxinh3=3, excit2=2$

cont4      := d curr3;
maxinh3    := d excit2;
```

and for the second and third axiom,

```
pform force4(a)=4, maxhom3=3, farad2=2, pot1=1$
                             % has to be preceded by a coframe statement
frame e$
farad2      :=  d pot1;
force4(-a)  := (e(-a) _|farad2)^curr3;
maxhom3     :=  d farad2;
```

These program bits and pieces, which look almost trivial, will be integrated into a complete and executable Maxwell sample program after we have learned about the energy-momentum distribution of the electromagnetic field and about its action.

The physical interpretation of equations (B.4.3) can be found via the (1 + 3)-decomposition that we had derived earlier as

$$
\begin{aligned}
J &= -j \wedge d\sigma + \rho =: (j, \rho)\,, && \text{(B.4.5)}\\
H &= -\mathcal{H} \wedge d\sigma + \mathcal{D} =: (\mathcal{H}, \mathcal{D})\,, && \text{(B.4.6)}\\
F &= E \wedge d\sigma + B =: (E, B)\,, && \text{(B.4.7)}\\
A &= -\varphi\, d\sigma + \mathcal{A} =: (\varphi, \mathcal{A})\,, && \text{(B.4.8)}
\end{aligned}
$$

(see (B.1.35), (B.1.39), (B.2.9), and (B.3.7), respectively). The signs in (B.4.5) to (B.4.8) will be discussed in the next section. A 3-dimensional visualization of the electromagnetic field is given in Fig. B.4.1.

We first concentrate on equations that contain only measurable quantities, namely the Maxwell equations. For their (1 + 3)-decomposition we found (see (B.1.44), (B.1.45) and (B.3.13), (B.3.14)),

$$
dH = J \quad \begin{cases} \underline{d}\,\mathcal{D} = \rho & \text{(1 constraint eq.)}\,, \\ \dot{\mathcal{D}} = \underline{d}\,\mathcal{H} - j & \text{(3 time evol. eqs.)}\,, \end{cases} \tag{B.4.9}
$$

$$
dF = 0 \quad \begin{cases} \underline{d}\,B = 0 & \text{(1 constraint eq.)}\,, \\ \dot{B} = -\underline{d}\,E & \text{(3 time evol. eqs.)}\,. \end{cases} \tag{B.4.10}
$$

Accordingly, we have $2 \times 3 = 6$ time evolution equations for the $2 \times 6 = 12$ variables $(\mathcal{D}, B, \mathcal{H}, E)$ of the electromagnetic field. Thus the Maxwellian structure in (B.4.3) is underdetermined. We need, in addition, an electromagnetic spacetime relation that expresses the excitation $H = (\mathcal{H}, \mathcal{D})$ in terms of the field strength $F = (E, B)$, i.e., $H = H[F]$. For classical electrodynamics, this functional becomes the Maxwell–Lorentz spacetime relation that we discuss in Part D.

Whereas, on the unquantized level, the Maxwellian structure in (B.4.1) or (B.4.3) is believed to be of universal validity, the spacetime relation is more of an ad hoc nature and amenable to corrections. The "vacuum" can have different spacetime relations depending on whether we take it with or without vacuum polarization. We will come back to this question in Chapters D.6 and E.2.

@@@ *Tables (put inside the* **front cover of the book***)*

Table II. SI units of the electromagnetic field and its source: C = coulomb, A = ampere, Wb = weber, V = volt, T = tesla, m = meter, s = second. The units oersted and gauss are phased out and do not exist any longer in SI. In the second column we find the units related to the absolute and in the third column those related to the relative dimensions.

Field	**SI unit of field**	**SI unit of components of field**
ρ	C	$\mathrm{C/m^3}$
j	$\mathrm{A = C/s}$	$\mathrm{A/m^2 = C/(s\,m^2)}$
$\mathcal{D}$	C	$\mathrm{C/m^2}$
$\mathcal{H}$	$\mathrm{A = C/s}$	$\mathrm{A/m = C/(s\,m)}$ $(\to 4\pi \times 10^{-3}\,\mathrm{oersted})$
E	$\mathrm{Wb/s = V}$	$\mathrm{V/m = Wb/(s\,m)}$
B	Wb	$\mathrm{Wb/m^2 = T}$ $(\to 10^4\,\mathrm{gauss})$

@@@ Tables (put inside the **front cover of the book***)*

Table I. The electromagnetic field and its source.

Field	name	mathematical object	independent components	related to	reflec-tion	absolute dimension
ρ	electric charge	twisted 3-form	ρ_{123}	volume	$-\rho$	q = electric charge
j	electric current	twisted 2-form	j_{23}, j_{31}, j_{12}	area	$-j$	q/t
$\mathcal{D}$	electric excitation	twisted 2-form	$\mathcal{D}_{23}, \mathcal{D}_{31}, \mathcal{D}_{12}$	area	$-\mathcal{D}$	q
$\mathcal{H}$	magnetic excitation	twisted 1-form	$\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$	line	$-\mathcal{H}$	q/t
E	electric field strength	untwisted 1-form	E_1, E_2, E_3	line	E	Φ/t
B	magnetic field strength	untwisted 2-form	B_{23}, B_{31}, B_{12}	area	B	Φ = magnetic flux

B.4.2 $^{\otimes}$Lenz and anti-Lenz rule

In generally covariant form, the Maxwell equations read $dH = J$ and $dF = 0$. However, their physical interpretation requires, via (B.4.5)–(B.4.7), an $(1+3)$-decomposition. We discussed the operational interpretation of the different quantities in ordinary 3-space, and now we turn to the signs involved.[2] Generally, we could write the 4-current as

$$J = i_{\mathrm{T}}\, j \wedge d\sigma + i_{\mathrm{S}}\, \rho\,, \tag{B.4.11}$$

where we introduced the factors i_{T} and i_{S} with values from $\{+1,-1\}$. Any factor the absolute value of which differs from 1 is assumed to be absorbed in the corresponding form. Hence (B.4.11) is the most general decomposition relative to a given foliation.

We differentiate (B.4.11) and use $\underline{d}\,\rho = 0$. Then the first axiom yields

$$d\,J = i_{\mathrm{T}}\, d\,j \wedge d\sigma + i_{\mathrm{S}}\, d\,\rho = d\sigma \wedge \left(-i_{\mathrm{T}}\, \underline{d}\,j + i_{\mathrm{S}}\,\dot{\rho}\right) = 0\,. \tag{B.4.12}$$

We know that ρ is the charge density and j the current density. With the usual conventions on orientation, namely that the outward direction from an enclosed 3D volume is positive etc., both quantities fulfill the continuity equation (B.1.12). Thus, $i_{\mathrm{S}} = -i_{\mathrm{T}}$ and

$$J = i_{\mathrm{T}}\left(j \wedge d\sigma - \rho\right)\,. \tag{B.4.13}$$

Let us now turn to the inhomogeneous Maxwell equation. For the excitation we have

$$H = h_{\mathrm{T}}\, \mathcal{H} \wedge d\sigma + h_{\mathrm{S}}\, \mathcal{D}\,, \tag{B.4.14}$$

where we introduced again the sign factors h_{T} and h_{S} with values from $\{+1,-1\}$. We differentiate (B.4.14)

$$\begin{aligned} dH &= h_{\mathrm{T}}\, \underline{d}\mathcal{H} \wedge d\sigma + h_{\mathrm{S}}\, \underline{d}\mathcal{D} + h_{\mathrm{S}}\, d\sigma \wedge \dot{\mathcal{D}} \\ &= d\sigma \wedge \left(h_{\mathrm{T}}\, \underline{d}\,\mathcal{H} + h_{\mathrm{S}}\, \dot{\mathcal{D}}\right) + h_{\mathrm{S}}\, \underline{d}\mathcal{D}\,. \end{aligned} \tag{B.4.15}$$

Accordingly, the inhomogeneous Maxwell equation $dH = J$ (see (B.4.9)) with the source (B.4.13) decomposes as

$$h_{\mathrm{T}}\, \underline{d}\,\mathcal{H} + h_{\mathrm{S}}\, \dot{\mathcal{D}} = i_{\mathrm{T}}\, j\,, \qquad h_{\mathrm{S}}\, \underline{d}\,\mathcal{D} = -i_{\mathrm{T}}\, \rho\,. \tag{B.4.16}$$

The conventional definition of the sign of the electric excitation is that the (positive) electric charge ρ is its *source*: $\underline{d}\,\mathcal{D} = \rho$. Consequently, $h_{\mathrm{S}} = -i_{\mathrm{T}}$ and $(\mathrm{B.4.16})_1$, the Oersted-Ampère-Maxwell law, reads

$$h_{\mathrm{T}}\, \underline{d}\,\mathcal{H} = -h_{\mathrm{S}}\left(j + \dot{\mathcal{D}}\right)\,. \tag{B.4.17}$$

[2] We follow here the paper of Itin et al. [19], see also the references given there.

The contributions of j and $\dot{\mathcal{D}}$ add up correctly to the the Maxwell current $j+\dot{\mathcal{D}}$. In analogy to $\mathcal{D}$, the magnetic excitation $\mathcal{H}$ is defined such as to have $j+\dot{\mathcal{D}}$ as its source:

$$\underline{d}\,\mathcal{H} = j + \dot{\mathcal{D}}\,, \qquad h_{\mathrm{T}} = -h_{\mathrm{S}} = i_{\mathrm{T}}\,. \tag{B.4.18}$$

Accordingly,

$$H = i_{\mathrm{T}}\,(\mathcal{H}\wedge d\sigma - \mathcal{D})\ . \tag{B.4.19}$$

We have not yet made a convention for the sign of the charge. If we manipulate with a cat's skin and a rod of amber, we can take the conventional choice. This amounts to $i_{\mathrm{T}} = -1$. Then the decompositions (B.4.13) and (B.4.19) coincide with (B.4.5) and (B.4.6), respectively.

Finally, we address the field strength:

$$F = f_{\mathrm{T}}\,E\wedge d\sigma + f_{\mathrm{S}}\,B = f_{\mathrm{T}}\,(E\wedge d\sigma + f_{\Lambda}\,B)\ , \qquad f_{\Lambda} := \frac{f_{\mathrm{S}}}{f_{\mathrm{T}}}\,. \tag{B.4.20}$$

The Lenz factor f_{Λ} can be either $+1$ or -1. The overall sign of F is again chosen by a convention in connection with the sign of the Lorentz force relative to a positive charge. Conventionally, $f_{\mathrm{T}} = 1$.

The homogeneous Maxwell equation $dF = 0$ (see (B.4.10)) decomposes according to

$$\underline{d}\,E + f_{\Lambda}\,\dot{B} = 0\,, \qquad \underline{d}\,B = 0\,. \tag{B.4.21}$$

Thus, the only sign left open is the one in Faraday's induction law $(\text{B.4.21})_1$. The choice $f_{\Lambda} = +1$ is equivalent to the *Lenz rule*: The induced voltage $\int E$ is directed such that it acts opposite to the $\dot{B}$ that creates it. The (unphysical) choice $f_{\Lambda} = -1$ we call the anti-Lenz rule.

Our choice in (B.4.7) corresponds to the Lenz rule. We will justify this later. Therefore, we have derived all signs in (B.4.5)–(B.4.7) with the exception of that of the Lenz factor f_{Λ}. Since the Lenz rule follows from energy considerations, we have to postpone further discussion until we have an axiom about the electromagnetic energy distribution. At the end of section B.5.3, we will resume our discussion.

B.4.3 $^{\otimes}$Jump conditions for electromagnetic excitation and field strength

Equations (B.3.34) and (B.3.38) represent the so-called jump (or continuity) conditions for the components of the electromagnetic excitation. In this section we give a more convenient fomulation of these conditions. Namely, let us consider on a 3-dimensional slice h_σ of spacetime (see Fig. B.1.3) an arbitrary 2-dimensional surface S, the points of which are defined by the parametric equations

$$x^a = x^a(\xi^{\dot 1}, \xi^{\dot 2})\,, \quad a = 1,2,3\,. \tag{B.4.22}$$

Here $\xi^A = (\xi^{\dot 1}, \xi^{\dot 2})$ are the two parameters specifying the position on S. We denote the corresponding indices $A, B, \ldots = \dot 1, \dot 2$ by a dot in order to distinguish them from the other indices. We assume that this surface is not moving; i.e., its form and position are the same in every $\sigma =$ it const hypersurface. We introduce the 1-form density ν *normal* to the surface S,

$$\begin{aligned}\nu \;&:=\; \left(\frac{\partial x^2}{\partial \xi^{\dot 1}}\frac{\partial x^3}{\partial \xi^{\dot 2}} - \frac{\partial x^2}{\partial \xi^{\dot 2}}\frac{\partial x^3}{\partial \xi^{\dot 1}}\right) dx^1 + \left(\frac{\partial x^3}{\partial \xi^{\dot 1}}\frac{\partial x^1}{\partial \xi^{\dot 2}} - \frac{\partial x^3}{\partial \xi^{\dot 2}}\frac{\partial x^1}{\partial \xi^{\dot 1}}\right) dx^2 \\ &+ \left(\frac{\partial x^1}{\partial \xi^{\dot 1}}\frac{\partial x^2}{\partial \xi^{\dot 2}} - \frac{\partial x^1}{\partial \xi^{\dot 2}}\frac{\partial x^2}{\partial \xi^{\dot 1}}\right) dx^3, \end{aligned} \tag{B.4.23}$$

and the two vectors *tangential* to S,

$$\tau_A := \frac{\partial x^a}{\partial \xi^A}\,\partial_a\,, \qquad A = \dot 1, \dot 2. \tag{B.4.24}$$

Accordingly, we have the two conditions $\tau_A \rfloor \nu = 0$. The surface S divides the whole slice h_σ into two halves. We denote them by the subscripts $_{(1)}$ and $_{(2)}$, respectively.

Then, by repeating the limiting process above for the integrals near S,, we find, instead of (B.3.34) and (B.3.38), the jump conditions

$$\left(\mathcal{D}_{(2)} - \mathcal{D}_{(1)} - \widetilde{\rho}\right)\Big|_{\rm S} \wedge \nu \;=\; 0, \tag{B.4.25}$$

$$\tau_A \rfloor \left(\mathcal{H}_{(2)} - \mathcal{H}_{(1)} - \widetilde{j}\right)\Big|_{\rm S} \;=\; 0. \tag{B.4.26}$$

Here, $\mathcal{D}_{(1)}$ and $\mathcal{H}_{(1)}$ are the excitation forms in the first half and $\mathcal{D}_{(2)}$ and $\mathcal{H}_{(2)}$ in the second half of the 3D space h_σ.

One can immediately verify that the analysis in Sec. B.3.4 of the operational determination of the electromagnetic excitation was carried out for the special case of a surface S defined by the equations $x^1 = \xi^{\dot 1}$, $x^2 = \xi^{\dot 2}$, $x^3 = 0$ (then $\nu = dx^3$ and $\tau_A = \partial_A$, $A = 1, 2$). It should be noted though that the formulas (B.3.34) and (B.3.38) are not less general than (B.4.25) and (B.4.26) since the local coordinates can always be chosen in such a way that (B.4.25), (B.4.26) reduces to (B.3.34), (B.3.38). However, in practical applications, the use of (B.4.25), (B.4.26) turns out to be more convenient.

In an analogous way, one can derive the jump conditions for the components of the electromagnetic field strength. Starting from the homogeneous Maxwell equations (B.3.13) and (B.3.14) and considering their integral form near S, we obtain

$$\left(B_{(2)} - B_{(1)}\right)\Big|_{\rm S} \wedge \nu \;=\; 0, \tag{B.4.27}$$

$$\tau_A \rfloor \left(E_{(2)} - E_{(1)}\right)\Big|_{\rm S} \;=\; 0. \tag{B.4.28}$$

As above, $B_{(1)}$ and $E_{(1)}$ are the field strength forms in the first half and $E_{(2)}$ and $E_{(2)}$ in the second half of the 3D space h_σ.

The homogeneous Maxwell equation does not contain charge and current sources. Thus, equations (B.4.27), (B.4.28) describe the continuity of the tangential piece of the magnetic field and of the tangential part of the electric field across the boundary between the two domains of space. In (B.4.25) and (B.4.26), the components of the electromagnetic excitation are not continuous in general, with the surface charge and current densities $\widetilde{\rho}$, $\widetilde{j}$ defining the corresponding discontinuities.

B.4.4 Arbitrary local noninertial frame: Maxwell's equations in components

In (B.4.9) and (B.4.10), we displayed the Maxwell equations in terms of geometrical objects in a coordinate and frame invariant way. Sometimes it is necessary, however, to introduce locally arbitrary (co)frames of reference that are noninertial, that is, accelerated in general. Then the components of the electromagnetic field with respect to a coframe ϑ^α, the physical components emerge and the Maxwell equations can be expressed in terms of these physical components.

Let the current, the excitation, and the field strength be decomposed according to

$$J = \frac{1}{3!} J_{\alpha\beta\gamma}\, \vartheta^\alpha \wedge \vartheta^\beta \wedge \vartheta^\gamma \,, \tag{B.4.29}$$

$$H = \frac{1}{2} H_{\alpha\beta}\, \vartheta^\alpha \wedge \vartheta^\beta \,, \qquad F = \frac{1}{2} F_{\alpha\beta}\, \vartheta^\alpha \wedge \vartheta^\beta \,, \tag{B.4.30}$$

respectively. We substitute these expressions into the Maxwell equations. Then the coframe needs to be differentiated. As a shorthand notation, we introduce the anholonomicity 2-form $C^\alpha := d\vartheta^\alpha = \frac{1}{2} C_{\mu\nu}{}^\alpha\, \vartheta^\mu \wedge \vartheta^\nu$ (see (A.2.35)). Then we straightforwardly find:

$$\partial_{[\alpha} H_{\beta\gamma]} - C_{[\alpha\beta}{}^\delta H_{\gamma]\delta} = \frac{1}{3}\, J_{\alpha\beta\gamma} \,, \quad \partial_{[\alpha} F_{\beta\gamma]} - C_{[\alpha\beta}{}^\delta F_{\gamma]\delta} = 0 \,. \tag{B.4.31}$$

If we use the (metric-free) Levi-Civita tensor density $\epsilon^{\alpha\beta\gamma\delta}$, then the excitation and the current,

$$\check{\mathcal{H}}^{\alpha\beta} := \frac{1}{2!}\, \epsilon^{\alpha\beta\gamma\delta}\, H_{\gamma\delta} \,, \qquad \check{\mathcal{J}}^\alpha := \frac{1}{3!}\, \epsilon^{\alpha\beta\gamma\delta}\, J_{\beta\gamma\delta} \,, \tag{B.4.32}$$

are represented as densities. Accordingly, Maxwell's equations (B.4.3) in components read alternatively

$$\partial_\beta \check{\mathcal{H}}^{\alpha\beta} + C_{\beta\gamma}{}^\alpha\, \check{\mathcal{H}}^{\beta\gamma} = \check{\mathcal{J}}^\alpha \,, \quad \partial_{[\alpha} F_{\beta\gamma]} - C_{[\alpha\beta}{}^\delta F_{\gamma]\delta} = 0 \,. \tag{B.4.33}$$

The terms with the C's emerge if Maxwell's equations are referred to an arbitrary local (co)frame. Only if we restrict ourselves to natural (or coordinate) frames is $C = 0$, and Maxwell's equations display their conventional form.[3]

This representation of electrodynamics can be used in special or in general relativity. If one desires to employ a laboratory frame of reference, then this is the way to do it: The object of anholonomicity in the lab frame has to be calculated. By substituting it into (B.4.31) or (B.4.33), we find the Maxwell equations in terms of the components $F_{\alpha\beta}$ and so on of the electromagnetic field quantities with respect to the lab frame — and these are the quantities one observes in the laboratory. Therefore the $F_{\alpha\beta}$ and so on are called *physical components.* As soon as one starts from (B.4.3), the derivation of the sets (B.4.31) or (B.4.33) is an elementary exercise. Many discussions of the Maxwell equations within special relativity in noninertial frames could be appreciably shortened by using this formalism.

B.4.5 $^{\otimes}$Electrodynamics in flatland: 2-dimensional electron gas and quantum Hall effect

Our formulation of electrodynamics can be generalized straightforwardly to arbitrary dimensions n. If we assume again the charge conservation law as a first axiom, then the rank of the electric current must be $n-1$. The force density in mechanics, in accordance with its definition $\sim \partial L/\partial x^i$ within the Lagrange formalism, should remain a covector-valued n-form. Hence we keep the second axiom in its original form. Accordingly, the field strength F is again a 2-form:

$$\oint\limits_{C_{n-1}} J = 0\,, \qquad f_\alpha = (e_\alpha \rfloor F) \wedge J\,, \qquad \oint\limits_{C_2} F = 0\,. \tag{B.4.34}$$

This may seem like an academic exercise. However, at least for $n = 3$, there exists an application: Since the middle of the 1960s, experimentalists were able to create a *2-dimensional electron gas* (2DEG) in suitable transistors at sufficiently low temperatures and to position the 2DEG in a strong external transversal magnetic field. Under such circumstances, the electrons can only move in a plane transverse to B and one space dimension can be suppressed.

[3]To formulate electrodynamics in accelerated systems by means of tensor analysis, see J. Van Bladel [51]. New experiments in rotating frames (with ring lasers) can be found in Stedman [47].

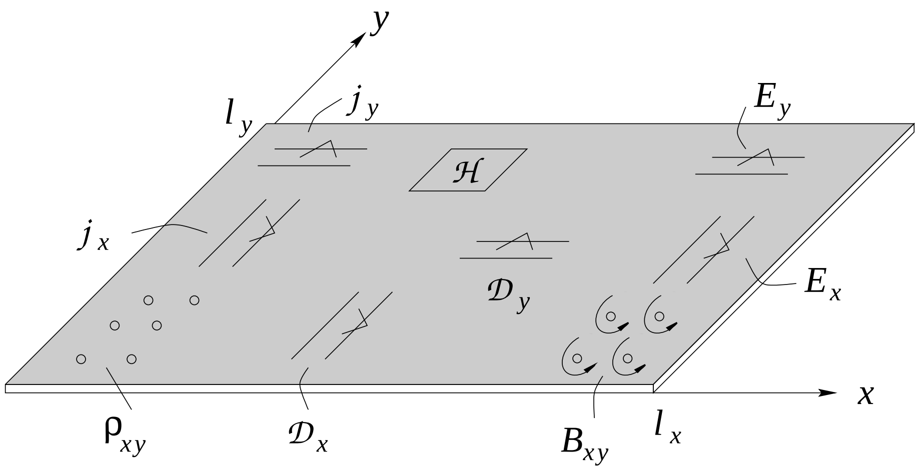

Figure B.4.2: Electrodynamics in $1+2$ dimensions: The arsenal of electromagnetic quantities in flatland. The sources are the charge density ρ_{xy} and the current density (j_x, j_y) (see (B.4.40)). The excitations $(\mathcal{D}_x, \mathcal{D}_y)$ and $\mathcal{H}$ (twisted scalar) are somewhat unusual (see (B.4.41)). The double plates for measuring $\mathcal{D}$, e.g., become double wires. The magnetic field has only one independent component B_{xy}, whereas the electric field has two, namely (E_x, E_y) (see (B.4.42)).

Electrodynamics in $1+2$ dimensions

In electrodynamics with one time and two space dimensions, we have from the first and third axioms (in arbitrary coordinates),

$$d\overset{(2)}{J} = 0\,, \qquad \overset{(2)}{J} = d\overset{(1)}{H}\,, \tag{B.4.35}$$

and

$$d\overset{(2)}{F} = 0\,, \qquad \overset{(2)}{F} = d\overset{(1)}{A}\,, \tag{B.4.36}$$

respectively, where we indicated the rank of the forms explicitly for better transparency. Here, the remarkable feature is that field strength F and current J carry the same rank; this is only possible for $n = 3$ spacetime dimensions. Moreover, the current J, the excitation H, and the field strength F all have the same number of independent components, namely three.

Now we $(1+2)$-decompose the current and the electromagnetic field:

$$\text{twisted 2-form:} \qquad \overset{(2)}{J} = -\,j\wedge d\sigma + \rho\,, \tag{B.4.37}$$

$$\text{twisted 1-form:} \qquad \overset{(1)}{H} = -\mathcal{H}\,d\sigma + \mathcal{D}\,, \tag{B.4.38}$$

$$\text{untwisted 2-form:} \qquad \overset{(2)}{F} = E\wedge d\sigma + B\,. \tag{B.4.39}$$

Accordingly, in the space of the 2DEG, we have (again in arbitrary coordinates)

$$\begin{aligned} j &= j_1\,dx^1 + j_2\,dx^2, & \rho &= \rho_{12}\,dx^1\wedge dx^2, & \text{(B.4.40)}\\ \mathcal{H}&, & \mathcal{D} &= \mathcal{D}_1\,dx^1 + \mathcal{D}_2\,dx^2, & \text{(B.4.41)}\\ E &= E_1\,dx^1 + E_2\,dx^2, & B &= B_{12}\,dx^1\wedge dx^2. & \text{(B.4.42)}\end{aligned}$$

We recognize the rather degenerate nature of such a system. The magnetic field B, for example, has only one independent component B_{12}. Such a configuration is visualized in Fig. B.4.2 by using rectilinear coordinates; they need not be Cartesian coordinates.

Charge conservation (B.4.35) in decomposed form and in components reads,

$$\underline{d}\,j + \dot\rho = 0\,, \qquad \partial_1 j_2 - \partial_2 j_1 + \dot\rho_{12} = 0\,. \tag{B.4.43}$$

The $(1+2)$-decomposed Maxwell equations look exactly as in (B.4.9) and (B.4.10). We also express them in components:

$$\begin{aligned} \underline{d}\,\mathcal{D} &= \rho\,, & \partial_1\mathcal{D}_2 - \partial_2\mathcal{D}_1 &= \rho_{12}\,, & \text{(B.4.44)}\\ \underline{d}\,\mathcal{H} - \dot{\mathcal{D}} &= j\,, & \partial_1\mathcal{H} - \dot{\mathcal{D}}_1 &= j_1\,, & \text{(B.4.45)}\\ && \partial_2\mathcal{H} - \dot{\mathcal{D}}_2 &= j_2\,, & \text{(B.4.46)}\end{aligned}$$

and

$$\begin{aligned} \underline{d}\,E + \dot B &= 0\,, \qquad \partial_1 E_2 - \partial_2 E_1 + \dot B_{12} = 0\,, & \text{(B.4.47)}\\ \underline{d}\,B &\equiv 0\,. & \text{(B.4.48)}\end{aligned}$$

We assume an infinite extension of flatland. If that cannot be assumed as a valid approximation, one has to allow for line currents at the boundary of flatland ("edge currents") in order to fulfill the Maxwell equations.

In our formulation the Maxwell equations don't depend on the metric. Thus, instead of the plane, as in Fig. B.4.2, we could have drawn an *arbitrary* 2-dimensional manifold, a surface of a cylinder or of a sphere, for example; the Maxwell equations (B.4.44) to (B.4.48) would still be valid since they had been derived for arbitrary (curvilinear) coordinates and arbitrary frames.

Before we can apply this formalism to the quantum Hall effect (QHE), we first remind ourselves of the classical Hall effect (of 1879).

Hall effect *(excerpt from the literature,[4] conventional formalism)*

This subsection on the Hall effect and the next one on QHE are excerpts from the literature. Thus equations (B.4.49) to (B.4.56) are written in the conventional notation, and they refer to Cartesian coordinates and may contain the metric of 3-dimensional space. However, our subsequent presentation of QHE,

[4] See Landau & Lifshitz [25], pp. 96–98 or Raith [39], p. 502.

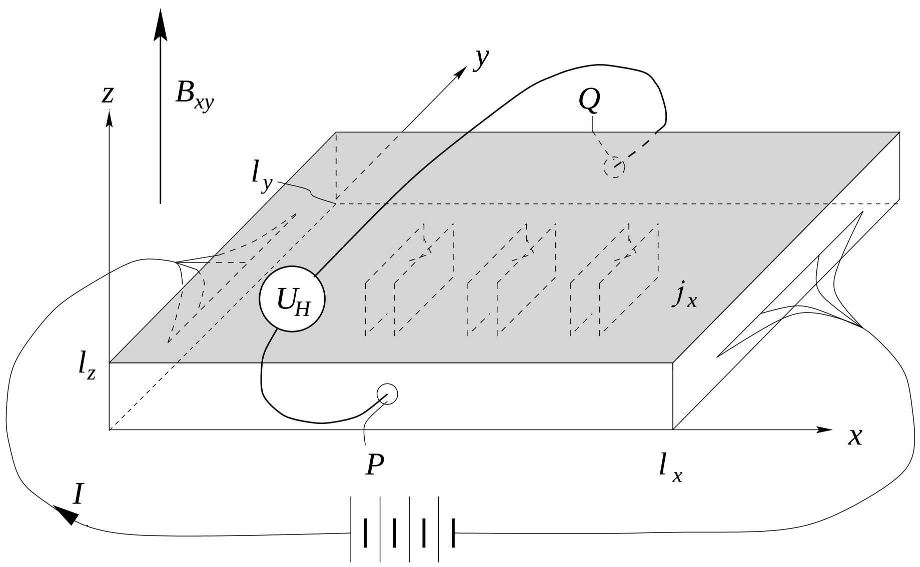

Figure B.4.3: Hall effect (schematic): The current density j in the conducting plate is affected by the external constant magnetic field B (in the figure only symbolized by one arrow) such as to create the Hall voltage U_{H}.

which starts with (B.4.57) and which contains all relevant features of a phenomenological description of QHE, is strictly metricfree.

We connect the two yz-faces of a (semi)conducting plate of volume $l_x \times l_y \times l_z$ with a battery (see Fig. B.4.3). A current I will flow and in the plate the current density is j_x. Transverse to the current, between the contacts P and Q, there exists no voltage. However, if we apply a constant magnetic field B along the z-axis, then the current j is deflected by the Lorentz force and the Hall voltage U_{H} occurs which, according to experiment, turns out to be

$$U_{\mathrm{H}} = R_{\mathrm{H}}\, I = A_{\mathrm{H}}\, \frac{BI}{l_z}\,, \qquad \text{with} \qquad [A_{\mathrm{H}}] = \frac{l^3}{q}\,. \tag{B.4.49}$$

R_{H} is called the Hall resistance and A_{H} the Hall constant. We divide U_{H} by l_y. Because $E_y = U_{\mathrm{H}}/l_y$, we find

$$E_y = A_{\mathrm{H}}\, B_{xy}\, \frac{I_x}{l_y\, l_z} = A_{\mathrm{H}}\, B_{xy}\, j_x\,. \tag{B.4.50}$$

Let us stress that the classical Hall effect is a volume (or bulk) effect. It is to be described in the framework of ordinary $(1+3)$-dimensional Maxwellian electrodynamics.

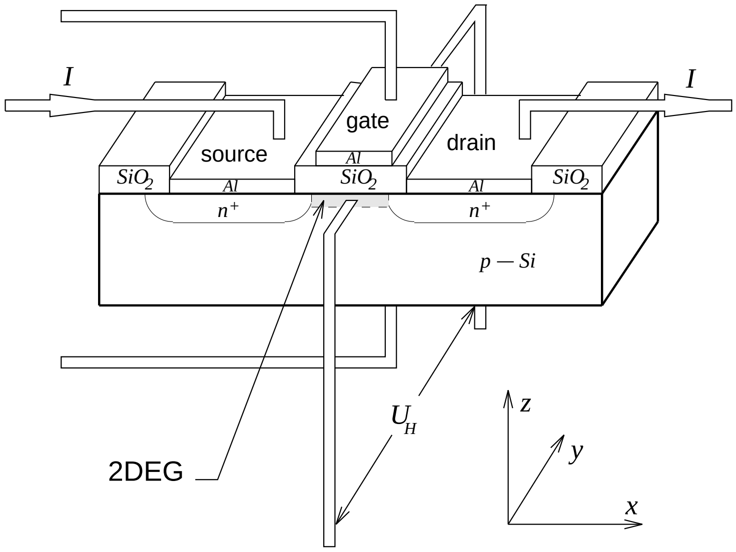

Figure B.4.4: A Mosfet with a 2-dimensional electron gas (2DEG) layer between a semiconductor (Si) and an insulator (SiO_2). Adapted from Braun [3]. In 1980, with such a transistor, von Klitzing et al. performed the original experiment on QHE.

Quantum Hall effect *(excerpt from the literature,[5] conventional formalism)*

A prerequisite for the discovery in 1980 of QHE were the advances in transistor technology. Since the 1960s one was able to assemble 2-dimensional electron gas layers in certain types of transistors, such as in a **m**etal-**o**xide-**s**emiconductor **f**ield **e**ffect **t**ransistor[6] (Mosfet) (see a schematic view of a Mosfet in Fig. B.4.4). The electron layer is only about 50 nanometers thick, whereas its lateral extension may go up to the millimeter region.

In the quantum Hall regime, we have very low temperatures (between 25 mK and 500 mK) and very high magnetic fields (between 5 T and 15 T). Then the conducting electrons of the specimen, because of a (quantum mechanical) excitation gap, cannot move in the z-direction, they are confined to the xy-plane. Thus an almost ideal *2-dimensional electron gas* (2DEG) is constituted.

The Hall conductance (= 1/resistance) exhibits very well-defined plateaus at integral (and, in the fractional QHE, at rational) multiples of the fundamental

---

[5] See, for example, von Klitzing [52], Braun [3], Chakraborty and Pietiläinen [4], Janssen et al. [20], Yoshioka [54], and references given there.

[6] A fairly detailed description can be found in Raith [39], pp. 579-582.

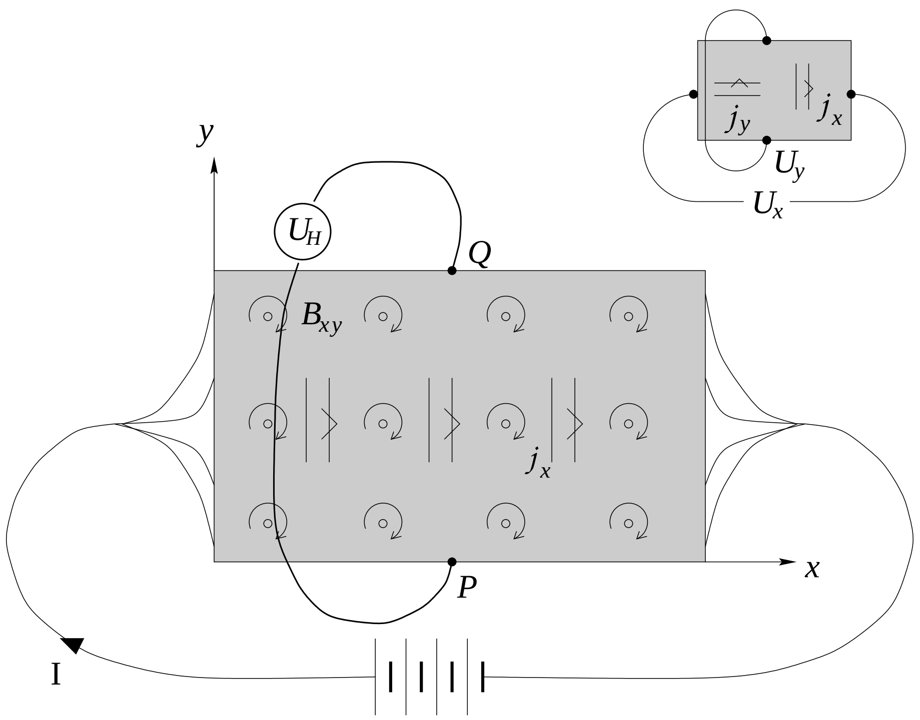

Figure B.4.5: Schematic view of a quantum Hall experiment with a 2-dimensional electron gas (2DEG). The current density j_x in the 2DEG is exposed to a strong transverse magnetic field B_{xy}. The Hall voltage U_{H} can be measured in the transverse direction to j_x between P and Q. In the inset we denoted the longitudinal voltage with U_x and the transversal one with $U_y (= U_{\mathrm{H}})$.

conductance[7] of $e^2/\mathrm{h} \overset{\mathrm{SI}}{=} 1/(25\,812.807\,\Omega)$, where e is the elementary charge and h Planck's constant. Therefore this effect is instrumental in precision experiments for measuring, in conjunction with the Josephson effect, e and h very accurately. We concentrate here on the *integer* QHE.

Turning to Fig. B.4.5, we consider the rectangle in the xy-plane with side lengths l_x and l_y, respectively. The quantum Hall effect is observed in such a 2-dimensional system of electrons subject to a strong uniform transverse magnetic field $\vec{B}$. The configuration is similar to that of the classical Hall effect (see Fig. B.4.3), but the system is cooled down to a uniform temperature of about of 0.1 K. The *Hall resistance* R_{H} is defined by the ratio of the Hall voltage U_y and

[7]Recently, the conductance of a *single* H_2-molecule has been measured (see Smit et al. [46]). The value $2e^2/h$ was found, which is consistent with the fundamental conductance of QHE.

the electric current I_x in the x-direction: $R_{\mathrm{H}} = U_y/I_x$. Longitudinally, we have the ordinary dissipative Ohm resistance $R_{\mathrm{L}} = U_x/I_x$. For fixed values of the magnetic field B and area charge density of electrons $n_{\mathrm{e}}\,e$, the Hall resistance R_{H} is a constant. Phenomenologically, QHE can be described by means of a *linear* tensorial Ohm–Hall law as a constitutive relation. Thus, for an *isotropic* material,

$$U_x = R_{\mathrm{L}}\,I_x - R_{\mathrm{H}}\,I_y\,, \tag{B.4.51}$$

$$U_y = R_{\mathrm{H}}\,I_x + R_{\mathrm{L}}\,I_y\,. \tag{B.4.52}$$

If we introduce the electric field $E_x = U_x/l_x\,,\ E_y = U_y/l_y$ and the 2D current densities $j_x = I_x/l_y\,,\ j_y = I_y/l_x$, we find

$$\vec{E} = \boldsymbol{\rho}\,\vec{j} \quad \text{or} \quad \vec{j} = \boldsymbol{\sigma}\vec{E}\,, \qquad \text{with} \qquad \boldsymbol{\rho} = \boldsymbol{\sigma}^{-1} = \begin{pmatrix} R_{\mathrm{L}}\frac{l_y}{l_x} & -R_{\mathrm{H}} \\ R_{\mathrm{H}} & R_{\mathrm{L}}\frac{l_x}{l_y} \end{pmatrix}, \tag{B.4.53}$$

where $\boldsymbol{\rho}$, the specific resistivity, and $\boldsymbol{\sigma}$, the specific conductivity, are represented by second-rank tensors.

With a classical electron model for the conductivity, the Hall resistance can be calculated to be $R_{\mathrm{H}} = B/(n_{\mathrm{e}}\,e)$, where B is the only component of the magnetic field in 2-space. The magnetic flux quantum for an electron is $\mathrm{h}/e = 2\Phi_0$, with $\Phi_0 \overset{\mathrm{SI}}{=} 2.07 \times 10^{-15}$ Wb. With the fundamental resistance we can rewrite the Hall resistance with the dimensionless *filling factor* ν as

$$R_{\mathrm{H}} = \frac{B}{n_{\mathrm{e}}\,e} = \frac{\mathrm{h}}{e^2}\,\nu^{-1}\,, \qquad \text{where} \qquad \nu := \frac{\frac{\mathrm{h}}{e}\,n_{\mathrm{e}}}{B}\,. \tag{B.4.54}$$

Observe that ν^{-1} measures the amount of magnetic flux (in flux units) per electron.

Thus classically, if we increase the magnetic field, keeping n_{e} (i.e., the gate voltage) fixed, we would expect a strictly linear increase of the Hall resistance. Surprisingly, however, we find the following (see Fig. B.4.6):

1. The Hall R_{H} has *plateaus* at rational heights. The plateaus at integer height occur with a high accuracy: $R_{\mathrm{H}} = (\mathrm{h}/e^2)/i$, for $i = 1, 2, \ldots$ (for holes we would get a negative sign).

2. When (ν, R_{H}) belongs to a plateau, the longitudinal resistance, R_{L}, *vanishes* to a good approximation.

In accordance with these results, we choose ν such that R_{L} is zero and σ_{H} constant at a plateau for one particular system, namely

$$\boldsymbol{\rho} = R_{\mathrm{H}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \text{or} \qquad \boldsymbol{\sigma} = \sigma_{\mathrm{H}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{B.4.55}$$

This yields

$$\vec{j} = \sigma_{\mathrm{H}}\,\epsilon\,\vec{E} \qquad \text{with} \qquad \epsilon := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{B.4.56}$$

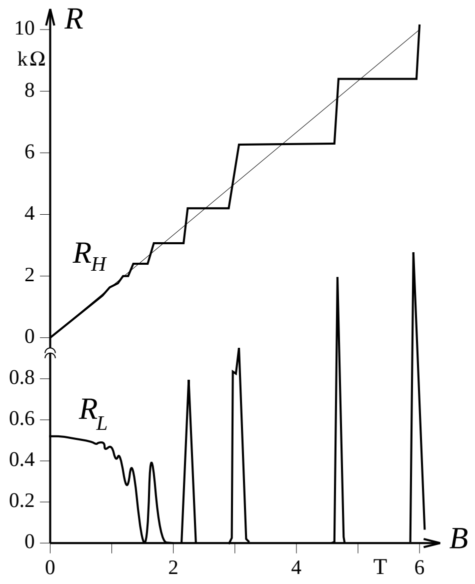

Figure B.4.6: The Hall resistance R_{H} at a temperature of 8 mK as a function of the magnetic field B (adapted from Ebert et al. [6]).

a phenomenological law valid at low frequencies and large distances. We will see in the next subsection that we can express this constitutive law in a generally covariant form. This shows that the laws governing QHE are entirely independent of the particular geometry (metric) under consideration.

Covariant description of the phenomenology at the plateaus

For a (1 + 2)-dimensional covariant formulation,[8] let us turn back to the first and the third axioms in (B.4.35) and (B.4.36), respectively. In (B.4.56), $\vec{j}$ is linearly related to $\vec{E}$. Accordingly, we will also assume *linearity* in $1+2$ dimensions between the 2-forms J and F,

$$J_{ij} = -\,\sigma_{ij}{}^{kl}\,F_{kl}\;, \tag{B.4.57}$$

with the Hall resistance $\sigma_{ij}{}^{kl}$. The minus sign is chosen because we want to recover (B.4.56) eventually. Since $J_{ij} = -J_{ji}$ and $F_{kl} = -F_{lk}$, we have

$$\sigma_{ij}{}^{kl} = -\sigma_{ji}{}^{kl} = -\sigma_{ij}{}^{lk}\;. \tag{B.4.58}$$

[8] This can be found in the work of Fröhlich et al. [11, 9, 10] (see also Avron [2], Richter and Seiler [42], and references given there). An early article on QHE and topology is that of Post [38], which seems to have been largely overlooked.

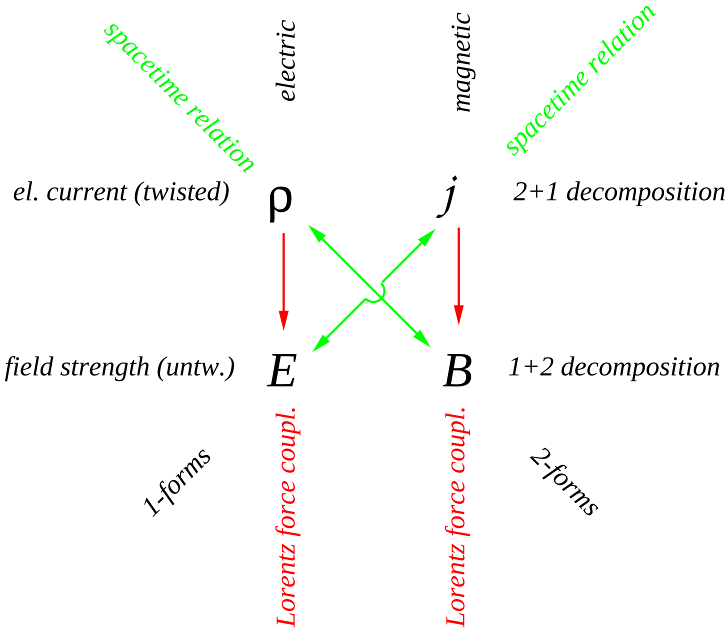

Figure B.4.7: Interrelationship between current and field strength in the Chern–Simons electrodynamics of QHE: In the *horizontal* direction, the 2-dimensional space part of a quantity is linked with its 1-dimensional time part to a 3-dimensional spacetime quantity. Lines in the *vertical* direction connect a pair of quantities that contribute to the Lorentz force density. And a *diagonal* line represents a spacetime relation.

Therefore the simplest $(1+2)$-covariant ansatz is found by assuming additionally *isotropy* in the $1+2$ dimensions, that is,

$$\sigma_{ij}{}^{kl} = \sigma_{\rm H}\, \delta^{kl}_{ij}\,, \tag{B.4.59}$$

with the generalized Kronecker delta δ^{kl}_{ij}. Substituted into (B.4.57), this yields

$$\boxed{J = -\sigma_{\rm H}\, F}\,, \qquad [\sigma_{\rm H}] = q^2/h = 1/\mathit{resistance}\,. \tag{B.4.60}$$

The Hall resistance $\sigma_{\rm H}$ is a twisted scalar (or pseudoscalar). Of course, we can only hope that such a simple ansatz is valid provided the Mosfet is isotropic with respect to its Hall conductance in the plane of the Mosfet. Because of $(\text{B.4.35})_1$ and $(\text{B.4.36})_1$, we find quite generally that the Hall conductance must be *constant* in space and time:

$$\sigma_{\rm H} = \mathit{const}\,. \tag{B.4.61}$$

We should be aware that something remarkable happened here. We have *not* introduced a spacetime relation à la $H \sim F$, as is done in $(1+3)$-dimensional electrodynamics (see Part D). Since for $n = 3$, H and F have still the same number of independent components, this would be possible. However, it is the ansatz $J \sim F$ that leads to a successful description of the phenomenology of QHE. One can consider (B.4.60) as a spacetime relation for the $(1 + 2)$-dimensional quantum Hall regime (see Fig. B.4.7). Of course, the integer or fractional numbers of σ_{H} cannot be derived from a classical theory. Nevertheless, classical $(1 + 2)$-dimensional electrodynamics immediately suggests a relation of the type (B.4.60), a relation that is free of any metric. In other words, the $(1 + 2)$-dimensional electrodynamics of QHE, as a classical theory, is metricfree. Equation (B.4.60) transforms the Maxwell equations $(\mathrm{B.4.35})_2$ and $(\mathrm{B.4.36})_2$ into a complete system of partial differential equations that can be integrated. At the same time, it also yields an explicit relation between H and F, namely

$$dH = -\sigma_{\mathrm{H}}\, F \qquad \text{or} \qquad dH = -\sigma_{\mathrm{H}}\, dA\,. \tag{B.4.62}$$

The last equation can be integrated. We find

$$H = -\sigma_{\mathrm{H}}\, A\,, \tag{B.4.63}$$

since the potential 1-form A is only defined up to a gauge transformation anyway. The metricfree *differential* relation $(\mathrm{B.4.62})_1$ between excitation H and field strength F is certainly not what we would have expected from classical Maxwellian electrodynamics. It represents, for $n = 3$, a totally new type of (Chern–Simons) electrodynamics.

We compare (B.4.60) with (B.4.37) and (B.4.39) to find

$$j = \sigma_{\mathrm{H}}\, E \qquad \text{and} \qquad \rho = -\sigma_{\mathrm{H}}\, B\,. \tag{B.4.64}$$

In (B.4.56), the current is represented as a vector. Therefore we introduce the vector density $\check{j} := {}^\diamond j$ by means of the diamond operator of (A.1.78) or, in components, $\check{j}^a = \epsilon^{ab}\, j_a/2$. Then $(a, b = 1, 2)$,

$$\check{j}^a = \frac{1}{2}\,\epsilon^{ab}\, j_b = \left(\sigma_{\mathrm{H}}\, \epsilon^{ab}\right)\, E_a\,; \tag{B.4.65}$$

that is, we recover (B.4.56), thereby verifying the ansatz (B.4.60). For the charge density $\check{\rho} := {}^\diamond\rho = \epsilon^{ab}\,\rho_{ab}/2$, we find

$$\check{\rho} = -\sigma_{\mathrm{H}}\, B_{12}\,, \qquad \text{with} \qquad \partial_1 \check{j}^1 + \partial_2 \check{j}^2 + \dot{\check{\rho}} = 0\,. \tag{B.4.66}$$

As mentioned, (B.4.57) and the subsequent equations in this subsection are all metricfree. In particular, this is true for (B.4.65). Nevertheless, we recover (B.4.56) from it. In our derivation the *crystal structure* does not enter. The assumptions of linearity and isotropy were sufficient. The (crystalline) material, in which QHE takes place, does define orthogonality by suitable crystal axes. However, this information is not needed for the derivation of (B.4.64).

Moreover, we really would need a theory of electrodynamics inside matter for the discussion of QHE. Such questions are treated in Part E. In spite of our naïve ansatz (B.4.60) in the realm of $(1+2)$-dimensional vacuum electrodynamics, we get the (long wavelength limit of the) phenomenology of QHE correct and we can post-dict the topological nature of QHE. This is not a minor achievement, and we consider it as a definite proof of the relevance of our way of formulating electrodynamics. Our guess is that the isotropy assumption in (B.4.59) is a sort of effective macroscopic average already. And probably for that reason the ansatz (B.4.60) works.

Preview: 3D Lagrangians

We introduce Lagrangians in Sec. B.5.4. Nevertheless, let us conclude this section on electrodynamics in flatland with a few remarks on the appropriate Lagrangian for QHE.

The Lagrangian in $(1+2)$-dimensional electrodynamics has to be a twisted 3-form with the dimension h of an action. Obviously, the first invariant of (B.2.25) qualifies,

$$I_1 = F \wedge H = -\frac{1}{\sigma_{\rm H}}\, J \wedge H = \frac{1}{\sigma_{\rm H}}\, H \wedge d\,H\,. \tag{B.4.67}$$

But also the Chern–Simons 3-form of (B.3.16), if multiplied by the twisted scalar $\sigma_{\rm H}$, has the right characteristics,

$$\begin{aligned}\sigma_{\rm H}\, C_A &= \sigma_{\rm H}\, A \wedge F = \sigma_{\rm H}\, A \wedge d\,A \\ &= -A \wedge d\,H = d(A \wedge H) - F \wedge H = -F \wedge H + d\,K\,,\end{aligned} \tag{B.4.68}$$

with the Kiehn 3-form $K = A \wedge H$. Therefore, both candidate Lagrangians are intimately linked,

$$\sigma_{\rm H}\, C_A = -I_1 + d\,K\,. \tag{B.4.69}$$

In other words, they yield the same Lagrangian since they only differ by an irrelevant exact form.

Let us define then the Lagrange 3-form

$$L_{\rm 3D} := \frac{1}{2\sigma_{\rm H}}\, H \wedge d\,H - A \wedge d\,H\,. \tag{B.4.70}$$

By means of its Euler–Lagrange equation,

$$d\left(\frac{\partial L}{\partial(d\,H)}\right) + \frac{\partial L}{\partial H} = 0 \qquad \text{or} \qquad \frac{1}{\sigma_{\rm H}}\, d\,H - d\,A = 0\,, \tag{B.4.71}$$

we can recover the 3D spacetime relation (B.4.60) that we started from. Alternatively, we can consider J as the external field in the Lagrangian

$$L_{\rm 3D'} := -\frac{\sigma_{\rm H}}{2}\, A \wedge d\,A - A \wedge J\,. \tag{B.4.72}$$

A similar computation as in (B.4.71) yields directly (B.4.60).

B.5
Electromagnetic energy-momentum current and action

B.5.1 Fourth axiom: localization of energy-momentum

> *Minkowski's "greatest discovery was that at any point in the electromagnetic field* in vacuo *there exists a tensor of rank 2 of outstanding physical importance. ...* each component of the tensor $E_p{}^q$ has a physical interpretation, *which in every case had been discovered many years before Minkowski showed that these 16 components constitute a tensor of rank 2. The tensor $E_p{}^q$ is called the* energy tensor *of the electromagnetic field."*
>
> Edmund Whittaker (1953)

Let us consider the Lorentz force density $f_\alpha = (e_\alpha \rfloor F) \wedge J$ in (B.4.3). If we want to derive the energy-momentum law for electrodynamics, we have to try to express f_α as an *exact* form. Then energy-momentum is a kind of generalized potential for the Lorentz force density, namely $f_\alpha \sim d\,\Sigma_\alpha$. For that purpose, we start from f_α. We substitute $J = dH$ (inhomogeneous Maxwell equation) and subtract out a term with H and F exchanged and multiplied by a constant factor a:

$$f_\alpha = (e_\alpha \rfloor F) \wedge dH - a\,(e_\alpha \rfloor H) \wedge dF\,. \qquad \text{(B.5.1)}$$

Because of $dF = 0$ (homogeneous Maxwell equation), the subtracted term vanishes. The factor a will be left open for the moment. Note that we need a nonvanishing current $J \neq 0$ for our derivation to be sensible.

We partially integrate both terms in (B.5.1):

$$f_\alpha = d\big[a\,F\wedge(e_\alpha \rfloor H) - H\wedge(e_\alpha \rfloor F)\big] - a\,F\wedge d(e_\alpha \rfloor H) + H\wedge d(e_\alpha \rfloor F)\,.$$

The first term already has the desired form. We recall the main formula for the Lie derivative of an arbitrary form Φ, namely $\pounds_{e_\alpha}\Phi = d(e_\alpha \rfloor \Phi) + e_\alpha \rfloor (d\Phi)$ (see (A.2.51)). This allows us to transform the second part of (B.5.2):

$$\begin{aligned} f_\alpha = d\big[a\,F\wedge(e_\alpha \rfloor H) \quad &- \quad H\wedge(e_\alpha \rfloor F)\big] \\ -a\,F\wedge(\pounds_{e_\alpha}H) \quad &+ \quad H\wedge(\pounds_{e_\alpha}F) \\ +a\,F\wedge e_\alpha \rfloor (dH) \quad &- \quad H\wedge e_\alpha \rfloor (dF)\,. \end{aligned} \tag{B.5.2}$$

The last line can be rewritten as

$$+a\,e_\alpha \rfloor [F\wedge dH] - a\,(e_\alpha \rfloor F)\wedge dH - e_\alpha \rfloor [H\wedge dF] + (e_\alpha \rfloor H)\wedge dF\,. \tag{B.5.3}$$

As 5-forms, the expressions in the square brackets vanish. Two terms remain, and we find

$$\begin{aligned} f_\alpha = d\big[a\,F\wedge(e_\alpha \rfloor H) \quad &- \quad H\wedge(e_\alpha \rfloor F)\big] \\ -a\,F\wedge(\pounds_{e_\alpha}H) \quad &+ \quad H\wedge(\pounds_{e_\alpha}F) \\ -a\,(e_\alpha \rfloor F)\wedge dH \quad &+ \quad (e_\alpha \rfloor H)\wedge dF\,. \end{aligned} \tag{B.5.4}$$

Because $dF = 0$, the third line adds up to $-af_\alpha$. Accordingly,

$$\begin{aligned} (1+a)\,f_\alpha = d\big[a\,F\wedge(e_\alpha \rfloor H) \quad &- \quad H\wedge(e_\alpha \rfloor F)\big] \\ -a\,F\wedge(\pounds_{e_\alpha}H) \quad &+ \quad H\wedge(\pounds_{e_\alpha}F)\,. \end{aligned} \tag{B.5.5}$$

Now we have to make up our minds about the choice of the factor a. With $a = -1$, the left-hand side vanishes and we find a mathematical identity. A real conservation law is only obtained when, eventually, the second line vanishes. In other words, here we need an a posteriori argument, i.e., we have to take some information from experience. For $a = 0$, the second line does not vanish. However, for $a = 1$, we can hope that the first term in the second line compensates the second term if somehow $H \sim F$. In fact, under "ordinary circumstances," to be explored below, the two terms in the second line do compensate each other for $a = 1$. Therefore we postulate this choice and find

$$f_\alpha = (e_\alpha \rfloor F)\wedge J = d\,{}^{\mathrm{k}}\Sigma_\alpha + X_\alpha\,. \tag{B.5.6}$$

Here the *kinematic energy-momentum* 3-form of the electromagnetic field, a central result of this section, reads

$${}^{\mathrm{k}}\Sigma_\alpha := \frac{1}{2}\,[F\wedge(e_\alpha \rfloor H) - H\wedge(e_\alpha \rfloor F)] \qquad \text{(fourth axiom)}\,, \tag{B.5.7}$$

and the remaining force density 4-form turns out to be

$$X_\alpha := -\frac{1}{2}\,(F\wedge \pounds_{e_\alpha}H - H\wedge \pounds_{e_\alpha}F)\;. \tag{B.5.8}$$

The absolute dimension of ${}^{\mathrm{k}}\Sigma_\alpha$ as well as of X_α is h/ℓ.

Our derivation of (B.5.6) doesn't lead to a unique definition of ${}^{\mathrm{k}}\Sigma_\alpha$. The addition of any closed 3-form would be possible,

$$ {}^{\mathrm{k}}\Sigma'_\alpha := {}^{\mathrm{k}}\Sigma_\alpha + Y_\alpha\,, \qquad \text{with} \qquad dY_\alpha = 0\,, \tag{B.5.9} $$

such that

$$ f_\alpha = d\,{}^{\mathrm{k}}\Sigma'_\alpha + X_\alpha\,. \tag{B.5.10} $$

In particular, Y_α could be exact: $Y_\alpha = dZ_\alpha$. The 2-form Z_α has the same dimension as ${}^{\mathrm{k}}\Sigma_\alpha$. It seems impossible to build up Z_α exclusively in terms of the quantities e_α, H, F in an algebraic way. Therefore, $Y_\alpha = 0$ appears to be the most natural choice. Thus, by the fourth axiom we postulate that ${}^{\mathrm{k}}\Sigma_\alpha$ *in (B.5.7) represents the energy-momentum current* that correctly localizes the energy-momentum distribution of the electromagnetic field in spacetime. We call it the *kinematic* energy-momentum current since we didn't find it by a dynamic principle, which we will not formulate before (B.5.89), but rather by means of some sort of kinematic arguments.

The current ${}^{\mathrm{k}}\Sigma_\alpha$ can also be rewritten by applying the anti-Leibniz rule for $e_\alpha \rfloor$ either in the first or second term on the right-hand side of (B.5.7). With the 4-form

$$ \Lambda := -\frac{1}{2}\, F \wedge H\,, \tag{B.5.11} $$

we find

$$ {}^{\mathrm{k}}\Sigma_\alpha = \quad e_\alpha \rfloor \Lambda + F \wedge (e_\alpha \rfloor H) = -e_\alpha \rfloor \Lambda - H \wedge (e_\alpha \rfloor F)\,. \tag{B.5.12} $$

Let us come back to (B.5.6). Of course, $d\,{}^{\mathrm{k}}\Sigma_\alpha$ and ${}^{\mathrm{k}}\Sigma_\alpha$ cannot be integrated since they are covector-valued forms. However, if we have a suitable vector field $\xi = \xi^\alpha e_\alpha$ available, we can transvect its components ξ^α with ${}^{\mathrm{k}}\Sigma_\alpha$. Accordingly, we define the "charge" density

$$ \mathcal{Q} := \xi^\alpha\, {}^{\mathrm{k}}\Sigma_\alpha = \frac{1}{2}\left[F \wedge (\xi \rfloor H) - H \wedge (\xi \rfloor F)\right]\,, \tag{B.5.13} $$

which should *not* be mixed up with the notion of an *electric* charge. In flat spacetime, for example, with t as time coordinate, a suitable (Killing) vector would be $\xi = \partial_t$ and $\mathcal{Q} = \xi^\alpha\, {}^{\mathrm{k}}\Sigma_\alpha = {}^{\mathrm{k}}\Sigma_t$ would signify the energy density. Generally, the integrals $\int d\mathcal{Q}$ and $\int \mathcal{Q}$ are well defined.

Let us now compute $d\mathcal{Q}$. We transvect (B.5.6) with the vector components ξ^α. Then,

$$ \begin{aligned} \xi^\alpha f_\alpha &= (\xi \rfloor F) \wedge J = \xi^\alpha\, d\,{}^{\mathrm{k}}\Sigma_\alpha + \xi^\alpha X_\alpha \\ &= d\left(\xi^\alpha\, {}^{\mathrm{k}}\Sigma_\alpha\right) - {}^{\mathrm{k}}\Sigma_\alpha \wedge d\,\xi^\alpha + \xi^\alpha X_\alpha \end{aligned} \tag{B.5.14} $$

or

$$ d\,\mathcal{Q} = (\xi \rfloor F) \wedge J + {}^{\mathrm{k}}\Sigma_\alpha \wedge d\,\xi^\alpha - \xi^\alpha X_\alpha\,. \tag{B.5.15} $$

If we evaluate the last term by substituting (B.5.8), we have to remember the rule for the multiplication of a vector by a scalar in a Lie derivative (see Sec. A.2.10):

$$\pounds_{fu}\,\omega = f\,\pounds_u\,\omega + df \wedge (u \rfloor \omega)\,. \tag{B.5.16}$$

We substitute $f = \xi^\alpha$, $u = e_\alpha$, $\omega = H$ and reorder:

$$\xi^\alpha \pounds_{e_\alpha} H = \pounds_\xi\, H - d\xi^\alpha \wedge (e_\alpha \rfloor H)\,. \tag{B.5.17}$$

An analogous formula is valid for the 2-form F. Then (B.5.15) becomes

$$d\,\mathcal{Q} = (\xi \rfloor F) \wedge J + \frac{1}{2}\,(F \wedge \pounds_\xi H - H \wedge \pounds_\xi F) \tag{B.5.18}$$

or, in holonomic components, with $\check{\mathcal{Q}}^i := \epsilon^{ijkl} Q_{jkl}/6$, $\check{\mathcal{J}}^i := \epsilon^{ijkl} J_{jkl}/6$, and $\check{\mathcal{H}}^{ij} := \epsilon^{ijkl} H_{kl}/2$,

$$\partial_i \check{\mathcal{Q}}^i = \xi^k F_{kl}\, \check{\mathcal{J}}^l + \frac{1}{4}\left(F_{kl}\, \pounds_\xi \check{\mathcal{H}}^{kl} - \check{\mathcal{H}}^{kl}\, \pounds_\xi F_{kl}\right)\,. \tag{B.5.19}$$

Under which conditions is $\mathcal{Q}$ conserved, i.e., $d\,\mathcal{Q} = 0$? We have to assume vacuum, i.e., $J = 0$. Moreover, the conditions $\pounds_\xi H = \alpha(x) H$ *and* $\pounds_\xi F = \alpha(x) F$, with an arbitrary function $\alpha(x)$, are sufficient for the vanishing of the expression within the last two parentheses. We will come back to this type of question at the end of this section under *Preview.*

B.5.2 Properties of the energy-momentum current, electric/magnetic reciprocity

${}^{\rm k}\Sigma_\alpha$ is trace free

The energy-momentum current ${}^{\rm k}\Sigma_\alpha$ is a 3-form. We can blow it up to a 4-form according to $\vartheta^\beta \wedge {}^{\rm k}\Sigma_\alpha$. Since it still has sixteen components, we haven't lost any information. If we recall that for any p-form Φ we have $\vartheta^\alpha \wedge (e_\alpha \rfloor \Phi) = p\Phi$, we immediately recognize from (B.5.7) that

$$\vartheta^\alpha \wedge {}^{\rm k}\Sigma_\alpha = 0\,, \tag{B.5.20}$$

which amounts to *one* equation. This property — the vanishing of the "trace" of ${}^{\rm k}\Sigma_\alpha$ — is connected with the fact that the electromagnetic field (the "photon") carries no mass and the theory is thus invariant under dilations. Why we call it the trace of the energy-momentum will become clear below (see (B.5.38)).

${}^{\rm k}\Sigma_\alpha$ is electric/magnetic reciprocal

Furthermore, we can observe another property of ${}^{\rm k}\Sigma_\alpha$. It is remarkable how symmetric H and F enter (B.5.7). This was achieved by our choice of $a = 1$.