Foreword

Richard P. McNitt

Engineers/scientists are oft exposed to the concept that the half-life of one's useful scientific knowledge is of the order of a decade, that those not *keeping up* will quickly be left behind. The contents of this book are proof, to me, that—if anything—a decade may well overstate the length of scientific half-life in the realm of electromagnetic/material interactions. I vividly recall a graduate course in physics (taught by a well-known physicist) four decades ago where we were exposed, more or less in passing, to some *second-order* effects: Peltier, Seebeck, and Thompson. As most of the students were engineers with strong interest in devices, some suggestions for utilization were quickly put forward—for instance, a possible direct-current refrigerator—but the professor noted that the effects were "small" for extant materials and thus unsatisfactory for such use. Similarly, we were given some rudimentary information about liquid crystals (possible thermometers), and birefringence of some strained materials, the latter an area of active research at that time for photoelasticians. As to electromagnetic waves, the "Maxwell equations of the time" were considered adequate, providing solutions in free space, isotropic homogenous materials, and waveguides. Most mathematical requirements were satisfied by the utilization of linear partial differential equations with constant coefficients.

This book presents many aspects of what is essentially a brave new engineering/scientist world, presenting major findings of the last decade, current research activity, speculations, and suggestions for future attack. There are powerful and general (we certainly had not been exposed to "Diffeo(4)", or to Monte

Carlo simulations) mathematical methods presented, *generalized* Maxwell's equations are suggested and then utilized to resolve complex situations, a plethora of new *effects* are described and explained, and rather exotic materials and materials systems are presented. Instead of being confined to existing materials and materials systems, engineers are now able to work with materials scientists to design systems (composites, thin films, etc.) they need even to the nanolevel. One author points out that the cause-and-effect orientation of the past is now integrated into a systems approach that has a goal-and-means orientation. Further, in the new realm of nanomaterials, quantum effects also come to the fore.

It is illuminating to list just some of the named *effects* (some of which are *third-* and *fourth-order*, but of increasing engineering significance) listed in this tome, effects that are available for exploitation by the informed and contemporary engineer: Faraday (rotation), Fresnel-Fizeau, Kerr, Matteuci, Mockels, Sagnac, Villari, Voigt (Cotton-Mouton), and Wiedemann. Topics such as natural optical rotation, electro-magnetic- and piezo-toroidics, magnetoelectric, magnetoimpedance, paramagnetoelectric, piezomagnetoelectric, whistler waves, and others are examined. Concepts such as excitons, light-assisted tunneling, photonic crystals, spatial solitons, semiconductor quantum wells, superlattices (metamaterials), and negative phase-velocity materials are considered, particularly as to how they will be effective in new materials with names such as Permalloy, Terfenol-D, and Yablonovite. Faced with such evidence of so many things that were not generally known four decades ago, it is tempting to validate a decade half-life, and to acknowledge that many engineers who thought they were well trained, had only mastered $(1/2)^4 =$ one sixteenth of the information presented in this book!

The broad and exceptionally well-explained contents should be of significant value to three very different populations:

(i) The engineering scientist whose formal education occurred some decades ago and who wishes to be brought *up-to-date*. ... The chapters are comprehensive, well written, and informative. Although not condescending, each chapter starts with fundamentals and completely develops the appropriate theory.

(ii) Those currently active in the very broad arena, who wish a compact yet comprehensive overview of the field as well as of those works that would be complementary to their own. ... This book should prove to be of real value in expanding the scope of their individual researches.

(iii) Graduate students. ... The authors should be commended for expressing their visions as to what remains to *be done*, what is important, and the possible modes of attack. This book should prove to be an excellent source of thesis problems as well as a *map* to achieve the desired solution.

As one who has been in all three of these groups (in reverse order), I found this book to be a treasure trove. . . . I trust you will too.

In an old Pennsylvania Dutch saying, I sign myself as one who is

Old too soon, smart too late.

Separating Field and Constitutive Equations in Electromagnetic Theory

Evert J. Post

Abstract
This essay is an outline of techniques to separate out constitutive functions in order to isolate the purest field law statement that is independent of constitutive specifics. A reorganization of this kind is a near-necessity when dealing with complex electromagnetic mediums. If so desired, it creates the possibility of simultaneously and coherently treating anisotropy, the Fresnel–Fizeau effect, the Sagnac effect, natural optical rotation and Faraday rotation. Aside from these applied aspects, the efforts required to proceed in orderly fashion open up worthwhile perspectives and insights from a purely theoretical angle. The tool of general covariant description is found to have a perceptive potential that goes way beyond the SR(3) group.

1 The beginnings

From the very beginning, Maxwell's inception of the concepts of macroscopic electromagnetic theory was in a form that left little to be desired; so what could be the concern of an essay on the stages of its evolution? The answer to this somewhat rhetorical question can thus be expected to indicate that the conceptual basis of Maxwell's formalism remains largely intact. Yet, special attention is to be devoted to what might be called questions on how we mathematically express the findings of Maxwell. In short, we face matters of mathematical engineering.

With the field equations in conceptually good shape, the task of tackling complicated mediums reminds us of the fact that the properties of a medium are determined by a completely separate set of equations that delineate the physical behavior of that medium. However, more than before in the case of simple mediums, attention needs to focus again on what exactly is the structure of the field equations that can be taken to apply in general. The final objective is then to combine that form with constitutive information solely specifying a particular medium.

As an early example of mixing field and constitutive functions, consider electromagnetic theory relating to wave propagation in matter-free space (i.e., vacuum). The equations then assume the form

$$\operatorname{div}\mathbf{H} = 0, \qquad \operatorname{curl}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}, \tag{1}$$

and

$$\operatorname{div}\mathbf{E} = 0, \qquad \operatorname{curl}\mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}, \tag{2}$$

where **E** and **H** are electric and magnetic fields, respectively, whereas the speed of light c is clearly a constitutive parameter. The elimination of either **E** or **H** reveals both fields satisfying the d'Alembertian wave equation. This result brought optics into the realm of electromagnetic theory and was a dramatic revelation of 19th-century physics [1].

An inspection of this early form of the Maxwell field equations for vacuum does not give much of an inkling about the existence of a set of constitutive equations for the vacuum. The vacuum properties are here represented by the free-space light velocity c, which during the 19th century had been identified as the ratio of electric- and magnetic-based units. In this mixed unit system, **E** and **H** were taken to have the same physical dimension.

It follows from this treatment of the vacuum situation that Maxwell treated empty space as some sort of nonmedium. The constitutive equations of vacuum were, so to say, built into the field equations. For a material medium exhibiting electric and magnetic polarizabilities, 19th-century physics introduced two new field quantities known as electric displacement **D** and magnetic induction **B**. In a simple isotropic medium, **B** and **D** were related to **E** and **H** through the isotropic constitutive equations

$$\mathbf{D} = \epsilon \mathbf{E}, \qquad \mathbf{B} = \mu \mathbf{H}, \tag{3}$$

in which ϵ and μ are merely numbers to describe medium properties; hence, **E**, **D**, **H** and **B** all have the same physical dimension. The field equations then take the following form:

$$\operatorname{div} \mathbf{B} = 0, \qquad \operatorname{curl} \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{4}$$

$$\operatorname{div} \mathbf{D} = 0, \qquad \operatorname{curl} \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t}. \tag{5}$$

Equations (3)–(5) led to another landmark conclusion that the speed of propagation u in a material medium would be smaller than c according to $u^2 = c^2/\epsilon\mu$.

For material mediums, (3)–(5) give an effective separation between field equations and constitutive equations. However, the presence of c in the field equations (4) and (5) testifies to the fact that matter-free space still remains in an exceptional position. It was engineering rather than the aesthetics of mathematical physics that would lead to the removal of this remaining defect in the beauty of the mathematical representation of electromagnetic theory.

2 Georgi's rationalization

The mixed-unit system of physics became highly impractical as electrification began to spread around the world. Around 1900, Giorgi proposed to adapt the electromagnetic units of Maxwell field theory to the practical units used in engineering. To that end, he introduced a set of constitutive equations for vacuum as follows:

$$\mathbf{D} = \epsilon_0 \mathbf{E}, \qquad \mathbf{B} = \mu_0 \mathbf{H}. \tag{6}$$

As $c^2 = 1/\epsilon_0\mu_0$, the field equations (4) and (5) accordingly were transformed into

$$\operatorname{div} \mathbf{B} = 0, \qquad \operatorname{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{7}$$

and

$$\operatorname{div} \mathbf{D} = 0, \qquad \operatorname{curl} \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}. \tag{8}$$

The physics establishment fought the Giorgi system for a long time, their main argument being that ϵ_0 and μ_0 were mere numerical artifacts undeserving of official status in their serene domain of physics. It would take a quarter century before the Giorgi system received a modicum of international recognition. It is noted parenthetically that there are recent reports that $\epsilon_0\mu_0 = 1/c^2$ may not be a universal constant [2]. The ratio of μ_0 and ϵ_0 acquired significance as a measure of free-space impedance, which is now increasingly regarded as a universal constant. So, the establishment was only half right with their numerical argument, which means they had no scientific leg to stand on and really deserved a failing mark for aesthetics of scientific description. The upshot is that the constitutive equations of simple isotropic material mediums now assume the form

$$\mathbf{D} = \epsilon_0\epsilon\mathbf{E}, \qquad \mathbf{B} = \mu_0\mu\mathbf{H}; \tag{9}$$

yet the field equations are still given as (7) and (8).

After the rationalization of units and the ensuing separation of field and constitutive equations by Giorgi, the first rumblings of relativity could be felt. Establishment physics was at that time still committed to its mixed unit system. So, when Minkowski boldly exposed the world of physics to the challenge of envisioning space and time as parameters of a manifold of four dimensions [3], c was still present in the Maxwell field equations. This was a somewhat disturbing element interfering with all that beautiful mathematical symmetry. While the Giorgi rationalization was already around to remove that defect, it would take a few more decades for full international recognition.

The Giorgi rationalization was either not known in mathematical physics circles or those who knew about it were unwilling to accept a proposition that, to them, seemed motivated by mere engineering considerations. The aftermath of this ill-advised professional chauvinism is still with us today. The reader may open many a book on quantum field theory and will find the convenient substitution $c = 1$ and sometimes, even worse, an imaginary time $x_0 = ict$. Some people just felt that spacetime should have a positive definite metric. These measures were motivated by a false aesthetics. With hindsight, they merely contributed to the drift of theory away from physical reality.

Additionally, the Giorgi rationalization was not popular among the creators of the special and general theories of relativity either. Minkowski and even Einstein were somewhat guilty of condoning the aforementioned convenient substitution for c [3]. It was forced on them by an old system of mixed units. Yet by the same token, as initiators of a new way of thinking, they might have expected others to help smoothen out conflicts ensuing from unsuitable earlier traditions.

Unfortunately, these *ad hoc* adaptations were made to gain a phony idealized representation of a new discipline. They remain a serious factor that has made many physicists uncomfortable with the formalisms of relativity. Without really discounting the whole structure and major consequences thereof, it gives rise to a feeling that only the initiated know what to do and when.

Here we are confronted with organizing an overview of the behavior of complex mediums. By necessity, this includes the discussion of moving mediums and accelerated frames of reference. Such a program cannot be done well without taking full advantage of the spacetime methods of description as opened up by Minkowski. However, after taking note of the difficulties created by *ad hoc* procedures invited by the mixed system of physical units, we proceed to implement a clean separation between field equations and constitutive equations such as achieved by the Georgi rationalization.

3 Georgi version of Minkowski electrodynamics

Let us now proceed to recast (7) and (8) in the Minkowski spacetime form. Since medium specifics require detail in terms of tensor components, the field equations likewise call for a tensorial form. The first two Maxwell equations (7) then combine into the following single equation, which contains an explicit version of a generalized curl or exterior derivative:

$$\partial_{[\lambda} F_{\nu\kappa]} = 0. \tag{10}$$

The subscripted indexes equal 0, 1, 2, and 3—with 0 as the time label and 1, 2, and 3 as space labels. Equations (8) are also combined into a single equation

$$\partial_\nu \mathfrak{G}^{\lambda\nu} = \mathfrak{C}^\lambda, \tag{11}$$

which involves a generalized divergence but can be equivalently expressed using the exterior derivative as follows:

$$\partial_{[\lambda} \tilde{G}_{\nu\kappa]} = \tilde{C}_{\lambda\nu\kappa}. \tag{12}$$

The index brackets [ ] indicate summation over even and odd permutations of the enclosed subscripted/superscripted indexes: even permutations get a plus sign, odd permutations get a negative sign. The components of $\mathfrak{G}^{\lambda\nu}$ and $\tilde{G}_{\nu\kappa}$ are related through the antisymmetric unit tensors; and so are $\mathfrak{C}^\nu$ and $\tilde{C}_{\lambda\nu\kappa}$. The Gothic symbols and tilde markings refer to transformation specifics that are essential to appreciate later the natural metric-free general invariant nature of the Minkowski rendition. More details are provided in Section 4.

At this point some readers may well be discouraged by what seems an excess of unexplained notational engineering. Yet, the truth is that dealing with complex mediums needs disciplined organization in order not be swamped by disorder later on.

Therefore, in the spirit of persistence we now proceed to componentwise identify F and $\mathfrak{G}$ with the usual electromagnetic fields according to the following matrix equivalences:

$$\begin{pmatrix} F_{00} & F_{01} & F_{02} & F_{03} \\ F_{10} & F_{11} & F_{12} & F_{13} \\ F_{20} & F_{21} & F_{22} & F_{23} \\ F_{30} & F_{31} & F_{32} & F_{33} \end{pmatrix} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}, \tag{13}$$

and

$$\begin{pmatrix} \mathfrak{G}^{00} & \mathfrak{G}^{01} & \mathfrak{G}^{02} & \mathfrak{G}^{03} \\ \mathfrak{G}^{10} & \mathfrak{G}^{11} & \mathfrak{G}^{12} & \mathfrak{G}^{13} \\ \mathfrak{G}^{20} & \mathfrak{G}^{21} & \mathfrak{G}^{22} & \mathfrak{G}^{23} \\ \mathfrak{G}^{30} & \mathfrak{G}^{31} & \mathfrak{G}^{32} & \mathfrak{G}^{33} \end{pmatrix} = \begin{pmatrix} 0 & D_1 & D_2 & D_3 \\ -D_1 & 0 & H_3 & -H_2 \\ -D_2 & -H_3 & 0 & H_1 \\ -D_3 & H_2 & -H_1 & 0 \end{pmatrix}. \tag{14}$$

For example, $\partial_{[0}F_{12]} = \partial_0 F_{12} + \partial_2 F_{01} + \partial_1 F_{20} = \partial B_3/\partial t + \partial E_1/\partial y - \partial E_2/\partial x = 0$ is the z component of $\operatorname{curl}\mathbf{E} = -\partial\mathbf{B}/\partial t$, while, similarly, $\partial_{[1}F_{23]} = 0$ is the same as $\operatorname{div}\mathbf{B} = 0$. With patience and persistence, one can make similar identifications for $\tilde{G}_{\nu\kappa}$ as well as $\mathfrak{C}^{\nu}$ and $\tilde{C}_{\lambda\nu\kappa}$.

These chores are not higher mathematics; instead, they are part of learning what may be seen as a revised language. While learning a new language, it is not wise to start with a *patois* that lacks the sophistication of a well-developed medium of communication. In electromagnetic theory, too many have settled for a *patois*. In the following we adhere, as much as possible, to notational stipulations adopted by Schouten [4] to give due attention to cited distinctions in transformation behavior.

Many readers may have seen these transcriptions and even have worked through them, and then could have asked: "What good is all of this?" Electromagnetic theory has been cast in so many different renditions that one cannot help wondering what is their benefit compared with the usual rendition in terms of vectors [5]? To compare the respective virtues, it is perhaps first necessary to be thoroughly aware of the far-reaching restrictions to which the system of vector analysis is subject.

The system of vector analysis identifies polar and axial vectors, which means that it is restricted to three spatial dimensions without reflections. Its invariance group is SR(3), which was adequate for early applications in mediums with rotational and inversion symmetries. Yet difficulties arise already with mediums that do not have a center of symmetry: say, solutions of optically active sugars [6-7].

An isotropic, optically active medium supports left– and right–circularly polarized plane waves with distinct propagation velocities [6]-[8]. The latter relate under reflection, yet reflections have no place in vector analysis. Physically, natural optical activity is due to a (dispersive) cross-coupling between magnetic and electric fields, which is a coupling between polar and axial vectors that happen to be identified in vector analysis. The early constitutive relations (3), (6) and (9) do not provide for such an option—not even for complex values of ϵ and μ.

An inspection of the great textbooks of the past (e.g., the 1934 edition of Max Born's *Optik* [9]) shows that few correctly state the absence of a center of symmetry as a key point in the description of optical activity [10]. Hence, in the past, the alternative has been frequently one of taking recourse to *ad hoc* phenomenological descriptions that could have no counterpart in the microphysical developments of solid state physics.

If an electric-magnetic cross-coupling already hints at a constitutive spacetime description, an electromagnetic description of the Fresnel–Fizeau and the Sagnac effects [8], [12] further exacerbates the limitations of vector renditions of electromagnetic theory. To avoid a recurring need for *ad hoc* procedures, a full-fledged spacetime description of field and constitutive equations truly becomes a *sine qua non*, no matter what type of objections one may have against the theory of relativity.

Confronted with the specifics of any of the foregoing phenomenons, a lonely researcher well trained in the vector renditions of electromagnetic theory can hardly be expected to be strongly motivated to recast that machinery. The Minkowski renditions available to him suffer a mix of field and constitutive parameters in the field equations, as is clear from the presence of c in (1) and (2). Those mixed units of the past made it impossible to accomplish a clean separation between field equations and constitutive equations. In the end, the researcher may throw up his/her hands in despair and decide not to get involved any further with spacetime projects. From that moment onwards, he/she may prejudice others against such formalisms.

Perhaps, a last-resort effort is still justified in overcoming such negativism. Let it be known that, taking advantage of the Giorgi rationalization of physical units, one finds that the Minkowski rendition does lead to a clean separation between field and constitutive equations. So, let (10)–(14) be testimony to this functional separation between these field equations and the constitutive equations that are to come.

An inspection of the modern textbook literature may show that today the Giorgi rationalization of units has been accepted, yet it does not mean there is a consistent habit of keeping constitutive parameters out of field equations. The widely celebrated *Feynman Lectures* [13] repeatedly state the Maxwell equations in forms wherein the parameters ϵ_0 and μ_0 explicitly occur. This testifies to a lingering belief that ϵ_0 and μ_0 are mere numerical factors, which is a faulty conviction of the past when the establishment was still fighting the Giorgi rationalization.

It is reiterated that ϵ_0 and μ_0 are components of the free-space constitutive tensor. They do not change under the SR(3) invariance group of vector analysis, but they do transform under the much wider group of transformations permitted by the Minkowski rendition—and that is the group needed to account for spacetime motions. In other words, modern textbooks still dwell, at least in this respect, on the remnants of the era of fighting Giorgi rationalization. Habits, once acquired, pass on from generation to generation.

All of this proves there still are no firm convictions universally held as to what are field equations and what are constitutive equations. A principle due to Neumann[1] [11] requires constitutive equations to be numerically invariant under the symmetry group of the medium under consideration. Hence, in order not to prejudicially affect the symmetry of mediums, one likes the field equations to remain invariant in form under at least all the invariance groups of all mediums to which the field equations are to be applied.

[1] Karl Gottfried Neumann (1832–1925), not to be confused with John von Neumann (of quantum mechanics fame and a computer pioneer in the 20th century), was a German mathematician whose name is immortalized in Neumann functions and Neumann boundary conditions. References to the Neumann principle are almost completely absent from current textbooks.

The Diffeo(4)-invariant Minkowski rendition of the Maxwell equations [10]-[14] meets that requirement well. The vector rendition of Maxwell equations meets that requirement only for SR(3) and its subgroups, which incidentally excludes isotropic, optically active mediums [6-7]. Since reflections and inversions are not contained in SR(3), we see here how unnecessary restrictive conditions of the past have been inviting *ad hoc* methodology.

The question now is: Where do we go from here? In light of the unusual degree of resistance encountered in implementing revisions of methodology, what are the chances of bringing the Minkowski rendition of electromagnetic theory to the fruition it deserves? Let me mention some of its principal features:

1. The Giorgi version of the Minkowski rendition of electromagnetic theory meets the criterion of separation between field equations and constitutive equations.
2. A further detailed mathematical specification shows that the field equations preserve their form under general differential spacetime transformations totally independent of the metric. This remarkable Diffeo(4) property identified by Kottler [15], Cartan [16] and van Dantzig [17] implies that these macroscopically established equations retain validity in the microscopic domain.
3. The fields thus defined in electromagnetic theory are integrands of either scalar- or pseudoscalar-valued cyclic[2] integrals that exhibit direct physical relevance, macroscopically as well as microscopically.
4. Since electromagnetic theory first involved the use of topological concepts (e.g., enclosing and linking) as essential features, its mathematical methods were later developed into a general procedure to access the topology of field configurations: a mathematical theory known as *de Rham cohomology* [19]. Topological properties remain invariant under *general differentiable deformations*, and so the term diffeomorphism Diffeo(4) has gained its natural place in electromagnetic theory.
5. The linear constitutive equation of matter-free space is a fourth-rank tensor (i.e., four sub/superscripted indexes are attached to it), which is a concomitant of the metric tensor invoking the free-space impedance. It has the same index symmetries as the Riemann–Christoffel tensor [14].

These are the main highlights of what may be called an *extended Giorgi version of Minkowski electrodynamics*. It recognizes the metric-free Diffeo(4) invariance in addition to the perspectives of de Rham cohomology, which are very fundamental for physical content, especially quantization. Thus, it shows a ramified and very constructive interaction with various domains of physics and mathematics. However, this wider scope manifests a greatly reduced tolerance for a good deal of somewhat wild and *ad hoc* experimentation in description that has been going on under the SR(3) umbrella of vector analysis.

[2]Cycles are integration domains of zero boundary: a 1-cycle is the topological equivalent of a circle (i.e., a closed loop), a 2-cycle is the topological equivalent of a spherical surface, a 3-cycle could relate to a closed physical universe, and so on, [18].