Chapter 1

Introduction

Sooner or later, any discussion of basic electromagnetic theory is certain to come to the issue of how best to categorize the vectors **B**, the magnetic induction, and **H**, the magnetic field strength. Polar or axial is the central issue [2]. From an elementary physical perspective, taking **B** as an axial vector seems appropriate since, as far as we know, all magnetism originates from currents (see Appendix 14.2). From a mathematical standpoint, however, taking **H** as a polar (or true) vector seems a better fit with an integral equation such as Ampere's law, $\int \mathbf{H} \cdot d\mathbf{l} = \mu_0 I$, particularly in relation to the subject of differential forms [3–5]. But taking the view that **B** can be one sort of vector while **H** is another seems to be at odds with an equation such as $\mathbf{B} = \mu \mathbf{H}$ in which the equality implies that they should be of the same character. A separate formal operator is required in order to get around this problem, for example, by writing $\mathbf{B} = \mu * \mathbf{H}$ where $*$ converts a true vector to an axial one and vice versa, but for most people, any need for this is generally ignored.

Geometric algebra provides a means of avoiding such ambiguities by allowing the existence of entities that go beyond vectors and scalars. In 3D, the additional entities include the bivector and the pseudoscalar. Here the magnetic field is represented by a bivector, which cannot be confused with a vector because it is quite a different kind of entity. Multiplication with a pseudoscalar, however, conveniently turns the one into the other. But the new entities are far from arbitrary constructs that have simply been chosen for this purpose, they are in fact generated inherently by allowing a proper form of multiplication between vectors, not just dot and cross products.

Using geometric algebra, Maxwell's equations and the Lorentz force are expressed in remarkably succinct forms. Since different types of entities, for example vectors and bivectors, can be combined by addition, the field quantities, the sources, and the differential operators can all be represented in a way that goes quite beyond simply piecing together matrices. While multiplication between entities of different sorts is also allowed, the rules all stem from the one simple concept of vector multiplication, the geometric product. Multiplication of a vector by itself results in a scalar, which provides the basis for a metric. Inner and outer products are very simply extracted from the geometric product of two vectors, the inner

Understanding Geometric Algebra for Electromagnetic Theory, First Edition. John W. Arthur.

product being the scalar part of the result whereas the outer product is the bivector part. Given that the product of a vector with itself is a scalar, inner and outer products are directly related to the ideas of parallel and perpendicular. The bivector therefore represents the product of two perpendicular vectors and has the specific geometric interpretation of a directed area. The pseudoscalar in 3D corresponds to a trivector, the product of three vectors, and can be taken to represent a volume. This hierarchy gives rise to the notion that a geometric algebra is a graded algebra, scalars being of grade 0, vectors grade 1, bivectors grade 2, and so on. Crucially, objects of different grades may be added together to form a general form of object known as a multivector. Just how this is possible will be explained in due course, but an example is $t+\mathbf{r}$, which has pretty much the same meaning as writing $(t,\mathbf{r})$ in normal vector algebra where, for example, this is the way we would normally write the time and position parameters of some given variable, for example, $\mathbf{E}(t,\mathbf{r})$. Why not $\mathbf{E}(t+\mathbf{r})$?

In the geometric algebras we shall be dealing with, pseudoscalars always have a negative square, a property that leads to complex numbers being superfluous. It is also found that the inner and outer products may generally be considered to be step-down and step-up operations, respectively. Provided the results are nonzero, the inner product of an object of grade n with another of grade m creates an object of grade $|m-n|$, whereas their outer product creates an object of grade $m+n$.

In addition to the novel algebraic features of geometric algebra, we also find that it is easy to turn it to calculus. In fact, the vector derivative $\boldsymbol{\nabla}$ provides all the functions of gradient, divergence, and curl in a unified manner, for, as the name suggests, it behaves like other vectors and so we can operate not only on scalars but also on vectors and objects of any other grade. While the inner and outer products with $\boldsymbol{\nabla}$ relate respectively to divergence and curl, the salient point is that we can use $\boldsymbol{\nabla}$ as a *complete entity* rather than in separate parts. Although the time derivative still requires to be dealt with separately by means of the usual scalar operator ∂_t, this no longer needs to stand entirely on its own, for just as we can combine time and position in the multivector form $t+\mathbf{r}$, we can do the same with the scalar and vector derivatives, in particular by combining the time and space derivatives in the form $\partial_t+\boldsymbol{\nabla}$. The everyday tools of electromagnetic theory are based on standard vector analysis in which time and space are treated on separate 1D and 3D footings, but here we have a more unified approach, which though not quite 4D may be appropriately enough referred to as (3+1)D where

$$(3+1)\text{D} = 3\text{D (space)} + 1\text{D (time)}$$

For example, we can write novel-looking equations such as $(\partial_t+\boldsymbol{\nabla})(t+\mathbf{r})=4+\mathbf{v}$ and $(\partial_t+\boldsymbol{\nabla})\mathbf{r}^2=2(\mathbf{r}+\mathbf{r}\cdot\mathbf{v})$, and many more besides, but we do have to be careful about what such equations might mean. Note however that (3+1)D properly refers to the physical model, where vectors represent space and scalars represent time, whereas the geometric algebra itself is 3D and should be strictly referred to as such. For the same reason, (3+1)D cannot be equated to 4D—time is treated as a scalar here, whereas it would properly require to be a vector in order to contribute a fourth dimension. When we do opt for a full 4D treatment, however, this is found to provide a very elegant and fully relativistic representation of spacetime. This has even more significance for the representation of electromagnetic theory because it unravels the basic mystery as to the existence of the magnetic field. It simply arises from a proper treatment of Coulomb's law so that there is no separate mechanism by which a moving charge produces a magnetic field. In fact, this was one of the revolutionary claims put forward in 1905 by Albert Einstein (see, e.g., Reference 2).

The aim of this work is to give some insight into the application of geometric algebra to electromagnetic theory for a readership that is much more familiar with the traditional methods pursued by the great majority of textbooks on the subject to date. It is our primary intention to focus on understanding the basic concepts and results of geometric algebra without attempting to cover the subject in any more mathematical detail than is strictly necessary. For example, although quaternions and the relationship between a 2D geometric algebra and complex numbers are important subjects, we discuss them only by way of background information as they are not actually essential to our main purpose.

We have also tried to avoid indulging in mathematics for its own sake. For example, we do not take the axiomatic approach; rather, we try to make use of existing ideas, extending them as and when necessary. Wherever it helps to do so, we draw on the intuitive notions of parallel and perpendicular, often using the symbols $\perp$ and // as subscripts to highlight objects to which these attributes apply. On the whole, the approach is also practical with the emphasis being on physical insight and understanding, particularly when there is an opportunity to shed light on the powerful way in which geometric algebra deals with the fundamental problems in electromagnetic theory.

The reader will find that there are some excellent articles that give a fairly simple and clear introduction to basic geometric algebra, for example, in Hestenes [6, 7] and in the introductory pages of Gull et al. [8], but in general, the literature on its application to electromagnetic theory tends to be either limited to a brief sketch or to be too advanced for all but the serious student who has some experience of the subject. The aim of this work is therefore to make things easier for the novice by filling out the bare bones of the subject with amply detailed explanations and derivations. Later on, however, we consider the electromagnetic field of an accelerating point charge. While this may be seen as an advanced problem, it is worked out in detail for the benefit of those readers who feel it would be worth the effort to follow it through. Indeed, geometric algebra allows the problem to be set up in a very straightforward and elegant way, leaving only the mechanics of working through the process of differentiation and setting up the result in the observer's rest frame.

Even if the reader is unlikely to adopt geometric algebra for routine use, some grasp of its rather unfamiliar and thought-provoking ideas will undoubtedly provide a better appreciation of the fundamentals of electromagnetics as a whole. Hopefully, any reader whose interest in the subject is awakened will be sufficiently encouraged to tackle it in greater depth by further reading within the cited references. It is only necessary to have a mind that is open to some initially strange ideas.

We start with a brief examination of geometric algebra itself and then go on to take a particular look at it in (3+1)D, which we may also refer to as the Newtonian world inasmuch as it describes the everyday intuitive world where time and space are totally distinct and special relativity does not feature. In his *Principia* of 1687, Newton summarized precisely this view of space and time which was to hold fast for over two centuries: “I will not define time, space, place and motion, as being well known to all” [9]. We then embark on finding out how to apply it to the foundations of basic electromagnetics, after which we briefly review what has been achieved by the process of restating the traditional description of the subject in terms of geometric algebra—what has been gained, what if anything has been lost, what it does not achieve, and what more it would be useful to achieve. This then leads to exploring the way in which the basic principles may be extended by moving to a 4D non-Euclidean space referred to as spacetime, in which time is treated as a vector in an equivalent but apparently somewhat devious manner to spatial vectors in that its square has the opposite sign. The concept of spacetime was originated by Hermann Minkowski in a lecture given in 1909, the year of his death. “Raum und Zeit,” the title of the lecture, literally means “space and time” whereas the modern form, spacetime, or space-time, came later. After covering the basics of this new geometric algebra, we learn how it relates to our ordinary (3+1)D world and in particular what must be the appropriate form for the spacetime vector derivative.

Once we have established the requisite toolset of the spacetime geometric algebra, we turn once again to the basic electromagnetic problems and show that not only are the results more elegant but also the physical insight gained is much greater. This is a further illustration of the power of geometric algebra and of the profound effect that mathematical tools in general can have on our perception of the workings of nature. It would be a mistake to have the preconception that the spacetime approach is difficult and not worth the effort; in fact, the reverse is true. Admittedly, many relativity textbooks and courses may give rise to such apprehensions. Even if the reader is resolutely against engaging in a little special relativity, they need not worry since the spacetime approach may simply be taken at face value without appealing to relativity. Only a few simple notions need to be accepted:

- Time can be treated as a vector.
- The time vector of any reference frame depends in a very simple way on its velocity.
- The square of a vector may be positive, zero, or negative.
- An observer sees spacetime objects projected into (3+1)D by a simple operation known as a spacetime split that depends only on the time vector of the chosen reference frame.

Again we draw on the notions of parallel and perpendicular, and, as a further aid, we also introduce a notation whereby underscoring with a tilde, ~, indicates that any vector marked in this way is orthogonal to some established time vector. That is to say, given $\boldsymbol{t}$ as the time vector, we can express any vector $\boldsymbol{u}$ in the form $u_t\boldsymbol{t} + \underset{\sim}{\boldsymbol{u}}$

where $u_t\boldsymbol{t}\,//\,\boldsymbol{t}$ and $\underset{\sim}{\boldsymbol{u}} \perp \boldsymbol{t}$. As a result, $\underset{\sim}{\boldsymbol{u}}$ may be interpreted as being a purely spatial vector. This has many advantages, an obvious one being that $\boldsymbol{u} = u_t\boldsymbol{t} + u_x\boldsymbol{x} + u_y\boldsymbol{y} + u_z\boldsymbol{z}$ can be written more simply as $\boldsymbol{u} = u_t\boldsymbol{t} + \underset{\sim}{\boldsymbol{u}}$.

It is quite probable that many readers may wish to skip the two chapters that mainly cover themes from special relativity, but it is also just as probable that they will refer to them later on if and when they feel the need to look "under the lid" and investigate how spacetime works on a physical level. This may be a useful approach for those readers who do initially skip these chapters but later decide to tackle the radiated field of an accelerating charge. On the other hand, those intrepid readers who wish from the outset to embark on the full exposé will probably find it best to read through the chapters and sections in sequence, even if this means skimming from time to time. With or without the benefit of special relativity, it is to be hoped that all readers should be about to put geometric algebra into practice for themselves and to appreciate the major themes of this work:

- The electric and magnetic fields are not separate things, they have a common origin.
- The equations governing them are unified by geometric algebra.
- In general, they are also simplified and rendered in a very compact form.
- This compactness is due to the ability of geometric algebra to encode objects of different grades within a single multivector expression.
- The grade structure is the instrument by which we may "unpack" these multivector expressions and equations into a more traditional form.
- Coulomb's law + spacetime = Σ classical electromagnetic theory.

While SI units are used throughout this book, in the later stages we introduce a convention used by several authors in which constants such as c, ε_0, and μ_0 are suppressed. This is often seen in the literature where natural units have been used so that $c = 1$. As a result, the equations look tidy and their essential structures are clearer. However, this need not be taken as a departure from SI into a different set of units; in our case, it is just a simple device that promotes the key points by abstracting superfluous detail. Restoring their conventional forms complete with the usual constants is fairly straightforward.

We use the familiar bold erect typeface for 3D vectors; for example, the normal choice of orthonormal basis vectors is $\mathbf{x}, \mathbf{y}, \mathbf{z}$, whereas for spacetime, we switch to bold italic simply to make it easier to distinguish the two when they are side by side. The usual spacetime basis is therefore taken as $\boldsymbol{t}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$. Vectors may be expressed in component form as $u_x\boldsymbol{x} + u_y\boldsymbol{y}\ldots$ and so on. While this is a departure from the notation typically seen in the literature, in our view, many readers will be more comfortable with this rather than having to deal with the use of indexed basis elements and the same typeface for scalars and vectors alike. When indexed basis elements are required, we use $\mathbf{e}_x, \mathbf{e}_y\ldots$ to mean the same thing as $\mathbf{x}, \mathbf{y}\ldots$ and so on. This is no different from using numerical indices since, after all, indices are only labels. This makes it possible to use the summation sign, for example,

$\mathbf{u} = \sum_{k=x,y,z} u_k \mathbf{e}_k$. Other notational forms and symbols are kept to a minimum, being introduced as and when necessary either because they are standard and need to be learned or simply because they fulfill a specific purpose such as readability; for example, we use the less familiar but neater form ∂_u rather than $\partial/\partial u$ for derivatives. Finally, the glossary of Appendix 14.1 provides an *aide memoire* to key terms and notation, and the other appendices provide a little more detail on some issues that, though of interest, would have simply been a digression in the main text.