Preface

And the same year I began to think of gravity extending to y^e Orb of the Moon & (having found out how to estimate the force w^{ch} [a] globe revolving within a sphere presses the surface of the sphere) from Kepler's rule of the periodical times of the planets being in sesquilateral proportion to their distance from the center of their Orbs, I deduced that the forces w^{ch} keep the planets in their Orbs must [be] reciprocally as the square of their distances from the centers about which they revolve: & thereby compared the force requisite to keep the Moon in her Orb with the force of gravity at the surface of the Earth, & found them answer [agree] pretty nearly. All this was in the plague years 1665–1666.

Isaac Newton, 1718. Quoted by Westfall ([**84**], p. 109).

The present book has grown from what was originally the very small project of writing an article highlighting three mutually unrelated areas of special and general relativity, namely the twin paradox (§ 6 of Chapter 1), the relativistic Maxwell equations (Chapter 3), and the precession of the orbit of Mercury (§ 7 of Chapter 4). As the writing proceeded, I realized that, while I was keeping the mathematical details to a minimum and omitting all the historical context of the physics, this material would still not be accessible to the audience I had in mind, consisting of people with mid-level undergraduate preparation in mathematics. One would still need to know at least the rudiments of calculus of variations and differential geometry. One thing led to another, and the filling in of those details required two years of work and expanded this work to its present Brobdingnagian size of some 400 pages, plus two additional volumes (posted online) of ancillary material. My vision of the core of the work (Volume 1) remains: it is intended to be a random set of commentaries on certain aspects of relativity. This book is neither a technical introduction to relativity, nor a systematic history of its development, nor yet a professional-quality examination of philosophical issues. The physicists and historians who vetted it for publication pointed out to me a number of areas where my ignorance and brashness led me to make throwaway comments that describe research already performed or in progress. As I shall keep reminding the reader, I am not a specialist in any of these areas; if I make some suggestions that expose my innocence, that is not the worst fate than can befall an author.

I have had two purposes in mind in writing the present work, one pedagogical, the other humanistic. The pedagogical purpose is to present some highlights of the special and general theories of relativity with full mathematical details in a form accessible to advanced undergraduate mathematics, science, and engineering majors (and, of course, any interested person who knows a little university-level mathematics). The humanistic purpose is to reflect on what these theories have meant for the human understanding of the physical world. These two purposes are intertwined, and the material that follows was selected so that both purposes could be pursued together. The book constitutes a meandering journey from one fundamental equation of mechanics to another, from Newton's $\boldsymbol{F} = m\boldsymbol{r}''$ to Einstein's $G_{\mu\nu} + \Lambda g_{\mu\nu} = (8\pi G/c^4)T_{\mu\nu}$. There are very many digressions along the way. The core material in Chapters 5–7 is preceded by three chapters discussing the need for the Lorentz transformation and its properties (Chapter 1), a standard discussion of the mechanics of special relativity (Chapter 2), a side excursion to highlight the fruitful interaction between special relativity and the Maxwell equations (Chapter 3), and a mostly computational chapter (Chapter 4) that shows how general relativity produced its first two major successes, all mathematical justification being postponed. These are followed by the three chapters (Chapters 5–7) of exposition of general relativity, Chapter 5 and Chapter 6 being devoted entirely to the requisite differential geometry. Chapter 7 contains a very brief discussion of nonrotating black holes, followed by an attempt to make the Einstein field equations appear natural. Chapter 8, which brings this work to a close, consists of a chronology of landmark works on physics in general and some reflections on metaphysics from the point of view of an interested nonspecialist.

I hope that physicists will be tolerant of my inexperience in their specialty. In lieu of expertise, what I offer the reader is a selection of topics that I have found fascinating and wish to share with others. To get all of the apologies out of the way at the beginning, let me state again that this is not a systematic exposition of relativity theory, which is an enterprise I leave to the experts. The experts have not been found wanting in this regard. The best example is the 2013 book *Einstein Gravity* by Anthony Zee (see my comments below). Other recent specimens are the books of Narlikar [**62**] and Dray ([**13**], [**14**]). The purely mathematical heart of the present book—Chapters 5 and 6 on curvature—is treated in more detail and more systematically in the recent book of Sternberg [**78**]. The reader is encouraged to look at these books to see how many topics have been omitted here. Generally, these other sources contain fewer computational proofs and more conceptual ones than the present book and offer the additional advantage of having been written by people who *do* have expertise in this area.

The manuscript has been vetted by anonymous reviewers, at least two of whom were physicists. These two were both very kind in their criticism but dissatisfied at many of my omissions, for example the absence of any discussion of the 1923–24 work of Elie Cartan ([**6**], [**7**]), which subsumes my amateur attempt to demonstrate in Chapter 4 that Newtonian mechanics cannot be geometrized in such a way as to account for the precession of a planetary orbit, and the similar absence of any mention of the 1911 work of Willem de Sitter (1872–1934) [**10**], which subsumes my attempt to adapt special relativity to account for that precession. They added that I have also omitted the work of Alexander Friedmann (1888–1925) on curved spaces ([**30**], [**31**]) and that of Georges Lemaître (1894–1966) [**52**] on the so-called Hubble expanding universe. I have also not mentioned the original paper[1] [**45**] of Edwin Hubble (1889–1953) himself. These omissions are, one and all, due to my lack of familiarity with, and/or complete unawareness of, the topics mentioned. I would

[1]Hubble's paper does not mention the earlier work of Lemaître; as Hubble was not careless about citation, one must assume he had not heard of it at the time he published his own work.

not venture to write about subjects in which I am so ill-informed. I emphasize that this book is, as its subtitle indicates, a set of reflections on relativity by a mathematician who hopes someone else might find them as interesting as I do. Nevertheless, I think advanced undergraduates can learn something of value about relativity from reading it.

My title is one I remember seeing on a televised documentary on relativity some decades ago. Since completing the work, I have also chanced upon the 1995 book ([**9**]) by physicist Paul Davies whose title is similar and which discusses relativity and other issues of modern physics. Indeed, a search through almost any online catalog for this title will turn up literally thousands of books whose titles contain the phrase "About Time". Some of them use it in the sense I intend. For others, the word "time" is a synonym for a prison term. The reader will recognize that the title bears a double meaning in all cases. In ordinary speech, "It's about time" expresses vexation that a desired result has been delayed and satisfaction that it has finally arrived. That is *not* the meaning I intend for it here, since I do not flatter myself that the world has been waiting impatiently for the present book. Rather, I have in mind the literal sense that time is precisely what this book is about and even (with some exaggeration) what the theory of relativity itself is about. Poincaré recognized this over a century ago, when he heaped lavish praise on the introduction of what Henrik Lorentz called *local time* and what is now called (following Hermann Minkowski) *proper time.* Even in everyday life, if we want to know the distance to a remote location, we are not usually concerned with the distance itself, but more often with the amount of time it will take us to go there. Our language is full of expressions that describe distance in terms of travel time. In physics, we can do the same thing by prescribing a fixed velocity v_0 and using the equation $d = v_0 t$ to convert a distance d to a time t and vice versa. That technique leads to some interesting geometrized Newtonian astronomy in Chapter 4 and is at the heart of the relativistic approach to planetary motion. The central fact on which relativity is based is that there is an absolute unit of speed that can be used to carry out this conversion, namely $v_0 = c$, the speed of light in a vacuum, a speed that is the same for all observers. All of the equations of special relativity depend on the entanglement of time with a spatial axis of relative motion between two observers and the resulting distinction between observed time and proper time for a moving particle. "It" is literally *about time.*

So much for the title. The subtitle expresses my pedagogical purpose, to which I now turn.

Pedagogical Aims

One of the hardest questions asked by students studying high-school and college-level algebra is, "What is algebra good for?" They know perfectly well that people *never* need to solve even quadratic equations in everyday life, and no one is ever faced with the problem of computing where two trains will meet setting out in opposite directions from Chicago and New York at different times. One can tell them that algebra is the language in which science is written, but the unfortunate truth is that science seldom needs to solve only algebraic equations; much more often, the problems involve differential equations. It is true that one needs to know algebra very well in order to understand differential equations, but that explanation is generally lost on people just beginning the study of the subject. It needs to be pointed out that algebra plays a vital role in the *discovery* of scientific laws. Here are two examples:

- The rule that "distance = speed × time" at constant speed has the same form as the rule that, for a rectangle of constant width and variable length, "area = width × length". It is easy to see geometrically—as the scholars at Merton College, Oxford, seem to have reasoned in the thirteenth century—that if the width (w) happens to be directly proportional to the length (l), that is, $w = kl$ for a constant k, the graph of this relation provides a family of right triangles with legs varying in proportion to each other, and areas A given by the relation we nowadays write as the formula $A = \frac{1}{2}kl^2$. By the analogy that exists between the two relations just cited, the distance (s) fallen when speed is directly proportional to time (t) ought to be $s = \frac{1}{2}gt^2$, where g is the constant of proportionality between speed v and and time t ($v = gt$).[2] Assuming that g is constant near the surface of the Earth, it is the *gravitational acceleration*, which by observation is 9.81 meters per second-squared.

  To extend this example, consider the case of a particle moving around a circle of radius r at uniform speed v. Since it is not moving in a straight line, it has some acceleration, always directed toward the center of the circle. It is intuitively obvious that the magnitude of this acceleration is constant. What is its value? In the absence of any force, the particle would fly off the circle along a tangent line at speed v, and after time t, it would be on a circle of radius $\sqrt{r^2 + (vt)^2}$. In order to stay on the circle, it must "fall" a distance[3]

  $$s = \sqrt{r^2 + (vt)^2} - r = r\left(\sqrt{1 + (vt/r)^2} - 1\right).$$

  Over a small interval of time t, the standard "differential" arguments of calculus show that the right-hand side of this equation is closely approximated by the expression

  $$r \cdot \frac{1}{2}\frac{(vt)^2}{r^2} = \frac{1}{2}\frac{v^2}{r}t^2 .$$

  The relative error in this approximation tends to zero as t tends to 0. Since it is the instantaneous acceleration we are interested in (essentially, the relation that results when $t = 0$), by comparing this expression with the Merton rule, we see that the factor g in the relation $s = \frac{1}{2}gt^2$ corresponds to v^2/r, which is therefore the magnitude of the acceleration. This is what Newton in the quotation above called "the force [with] which a globe revolving within a sphere presses the surface of the sphere".

  To extend the example still more, given the formula v^2/r for the acceleration of a body in motion at speed v around a circle of radius r, Kepler's third law implies, as Newton noted, an inverse square law of gravitational attraction.[4] The implication also goes in the opposite

[2]The analogy between the geometric and mechanical formulas here was explicitly noted and illustrated by the fourteenth-century Bishop of Lisieux Nicole d'Oresme (1323–1382), who laid the groundwork for the analytic geometry of Fermat and Descartes.

[3]Only the radial distance fallen is involved here, since the acceleration has no tangential component.

[4]From Newton's own words, quoted above, this appears to be the reason he believed in an inverse-square law. One can easily imagine other considerations that point to the same conclusion.

direction. An inverse-square law of attraction, together with Newton's second law of motion, implies Kepler's third law, as Newton later proved. (See Chapter 4 below.) To take the simplest case, for a body in circular orbit, the derivation of the inverse-square law is merely a short string of equations:

$$T = k\,r^{\frac{3}{2}} \quad \text{(Kepler's third law, } T = \text{period, } r = \text{radius of the orbit)}\,,$$

$$v = \frac{2\pi r}{T} = \frac{2\pi}{k} r^{-\frac{1}{2}} \quad \text{(speed = distance/time)}\,,$$

$$\frac{v^2}{r} = \left(\frac{2\pi}{k}\right)^2 r^{-2} \quad \text{(the inverse-square law)}\,.$$

Finally, to reap the harvest of this simple mathematics, consider, as Newton claimed to have done, the orbit of the Moon. Its sidereal period T is approximately $27\,\frac{1}{3}$ days,[5] which amounts to 2.361×10^6 seconds. The radius r of the Moon's orbit (approximating it by a circle) is 3.85×10^8 meters. By the formula given above, its acceleration v^2/r is $4\pi^2 r/T^2$, which amounts to 0.002725 meters per second-squared. Since the radius r is almost exactly 60 times the radius of the Earth, this ought to be about equal to the acceleration of gravity at the Earth's surface (9.81 meters per second-squared) divided by $60^2 = 3600$. And indeed, $9.81/3600 = 0.002725$. As Newton said, the two figures "answer pretty nearly"! (Newton wrote these words half a century after the alleged events they describe. Historians do not believe Newton had the law of universal gravitation in 1665. If he had possessed it at that time, the computation he claimed to have made would not have worked, since he believed one degree of a great circle on the Earth's surface was about 60 miles. In fact, it is about 70 miles (111 km). He really did make the computation later, with accurate figures on the size of the Earth provided in a posthumously published work of Jean-Félix Picard (1620–1682).) This astonishingly close agreement with observation is an awe-inspiring example of the power of simple mathematics to reveal the mysteries of the universe.

To summarize, in three brief, elementary arguments using simple geometry, simple algebra, and a tiny bit of calculus, we have set up a plausible physical law that can be applied and tested with astronomical observations and have performed such a test. Our proposed law has passed the test of observation amazingly well.

That is the kind of connection I enjoy making and will be making in these pages. As a mathematician, I am most interested in the contribution that symbolic algebra makes in the process just described. It was the analogy between the relations "length × width = area" and "speed × time = distance", expressed by two symbolic equations of exactly the

---

For example, if the total "gravitational attraction" remains constant as it spreads out from the center of attraction, its "density", which is the attraction on a particle at a given point at distance r, would be "diluted" in proportion to the area of the sphere of that radius, which is to say, in proportion to the square of the distance.

[5]Because the Earth moves about 26 degrees around the Sun during that sidereal period, the Moon must revolve about 390° degrees from one full moon to the next. Hence the synodic period of the Moon is about $29\,\frac{1}{2}$ days.

same algebraic form, that led us (not the Merton scholars) to the Merton rule, and ultimately to the law of falling bodies. It was the algebraic similarity of the formula connecting the distance a heavy body falls in time t with the distance a body in circular motion falls toward the center of the circle in time t that produced the law of acceleration for uniform circular motion.

My second example shows how algebra enables us to reason about the geometry of multi-dimensional spaces too complex to visualize.

- It will be shown in the discussion of curvature in Chapters 5 and 6 that the intuitive geometric idea of projecting the derivative of a tangent vector to a surface in three-dimensional real space $\mathbb{R}^3$ into the tangent plane of the surface leads to the Christoffel symbols, for which an explicit algebraic formula can be given. This algebraic formula can be trivially extrapolated to any finite number of variables by the simple *algebraic* process of extending the range of summation on the indices. In that way, the whole panoply of covariant derivatives, parallel transport, and curvature can be defined in a completely abstract way, without the need for any embedding in an absolute Euclidean space. That is a crucial point, since space-time is the only universe we know, and we have no direct evidence that it is embedded in a space of higher dimension. The geometric, intuitive origin of the Christoffel symbols lies in the familiar territory of surfaces in three-dimensional space. But, by a simple algebraic generalization, replacing 3 by n, we get Christoffel symbols that describe the curvature of manifolds of any dimension.

Humanistic Aims

The special and general theories of relativity are milestones in the progress of theoretical and experimental physics in the late nineteenth and early twentieth centuries. At the same time, these theories have produced a profound paradigm shift in the modes of thought by which scientists order the universe. They were necessary on purely theoretical grounds due to asymmetries in the equations of electromagnetism and on experimental grounds due to the failure of attempts to detect the hypothetical "luminiferous ether" in which light was believed to be an elastic wave. In that respect, they resemble the paradigm shift in chemistry in the eighteenth century, which coincided with the failure of the hypothetical "phlogiston" to show itself to experiment and its replacement by a quantitive theory of oxidation.

After presenting the mathematical skeleton of the theories in Chapters 1–7, I take some time in the final chapter to explore these issues and reflect on the evolution of physics all the way from Aristotle through Einstein. In this chapter, not being a specialist in philosophy, I am more concerned to raise questions for people who may not have thought of them than to propose answers that professional philosophers of science may have thought of already and perhaps even rejected as inadequate.

In order to fulfill my humanistic aims, I have sought breadth and generality rather than profundity and detail. But to avoid sacrificing my pedagogical goals, I have not sought *ultimate* generality. The whole book reflects the tension between these two goals. For example, a minimal exposition of special relativity is given in the first two chapters. That is pedagogically necessary for what follows, but to keep the student's imagination active, I replace some of the standard material on special relativity in the first three chapters with topics that I happen to find interesting. These sections are marked with asterisks to indicate that they can be omitted without loss of continuity. They are the kinds of questions that mathematicians ask—Is the composition of two Lorentz transformations a Lorentz transformation? If so, how can I get its matrix in standard form? Given that this standard form is obtained by "sandwiching" the actual coordinate transformation between two rotations on the spatial portions of the observers' space-times, can we be sure this composition is associative? Can relativistic velocities be made into a group?—and those who are not fascinated by such questions would be well advised to skip these starred sections. I do think, however, that the deduction of the Maxwell curl equations from the divergence equations in Chapter 3 by use of the relativistic transformation of electric and magnetic fields between observers will be interesting to nonmathematicians.

Special Features of This Book

The following is a list of aspects of relativity that are not part of standard physics textbooks, an exception being the recent book of Dray [**13**]. I believe they would be of both pedagogical and humanistic value to university professors teaching this material.

- Let c denote the speed of light.[6] Consider three observers moving away from a common origin, the three pairs having relative speeds u, v, and w (all less than c). With them, we associate three lines in a plane with lengths U, V, and W via transformations $U = k \operatorname{arctanh}(u/c)$, $V = k \operatorname{arctanh}(v/c)$, and $W = k \operatorname{arctanh}(w/c)$, where k is a fixed unit of length. The inverse transformations are $u = c \tanh(U/k)$, $v = c \tanh(V/k)$, and $w = c \tanh(W/k)$. As a consequence of the Lorentz transformation, the lengths U, V, and W will be such that they form a triangle, that is, the sum of the two smaller ones will exceed the largest. If we assign angles to that triangle equal to the angles that each of the three observers will measure between the lines of sight to the other two, its metric properties will be those given by the trigonometric relations in a hyperbolic plane of curvature $-1/k^2$. This fact has been well known for over a century, having been first noted by the Croatian mathematician (of Serbian descent) Vladimir Varićak (1865–1942), who wrote a series of articles on it (see, for example, [**82**]). It became better known after it was discussed in a book [**3**] by the eminent mathematician Emile Borel (1871–1956). It has been discussed in a number of places, most recently in the book of Dray [**13**]. One consequence of this fact is that the simple, very computable, Lorentz transformation can be used to derive all of hyperbolic plane trigonometry, including the formula for the angle of parallelism, horocycles, and equidistant lines, in just a few pages. That derivation is carried out in Appendix 1 (in Volume 2).

  The connection between relativistic velocities and hyperbolic trigonometry is even richer than just indicated. The set of all observable velocities in a fixed plane of reference can naturally be associated

[6] This is now the universal symbol for the speed of light, probably derived from the Latin word for speed, *celeritas*. In his 1905 paper on special relativity, Einstein used the symbol V.

with the points in a disk of radius c. If the relativistic law of cosines is applied to define the first fundamental form on that disk, the result is precisely the Beltrami–Klein model of the hyperbolic plane, named after Eugenio Beltrami (1835–1900) and Felix Klein (1849–1925). The connection between the geometry of relativistic velocities and the geometry of the hyperbolic plane is explained in detail in Appendix 1 (of Volume 2).

In all the discussions of this connection that I have seen, the authors have explained the Lorentz transformation in terms of the geometry of the hyperbolic plane. My aim in the present work is to do the converse: to show how easily the trigonometry of the hyperbolic plane follows from the Lorentz transformation. It seems to me that this would be an excellent way to present non-Euclidean geometry to students, avoiding all the hard computations involved in using the Poincaré disk model or the even more arduous, classical, "bare-fisted" derivation using Saccheri and Lambert quadrilaterals and the fact that the hyperbolic horosphere is a Euclidean plane.

In addition to the pedagogical considerations just mentioned, this connection between relativity and hyperbolic geometry is for me a source of historical wonder: The trigonometry of the hyperbolic plane was worked out by Gauss, Lobachevskii, János Bólyai, and others in the early nineteenth century. There seems to be no logical or physical reason why this purely mathematical creation should have any connection with modern physics, and yet—astonishingly—the Lorentz transformation automatically produces relations among the speeds and directions of three observers in uniform relative motion that mirror *exactly* the relations among the parts of a triangle in the hyperbolic plane. To me, this is an excellent example of what Felix Klein (1849–1925) must have had in mind when, lecturing on the theory of a spinning top at the 1893 World's Columbian Exposition, he referred to a "pre-established harmony" between the physical world and the world of mathematical concepts. Klein was referring to the excellent "fit" between the assumptions of the physical model and what we now call the Cayley–Klein parameters of a rotation, which he introduced in this lecture. He remarked, incidentally, that these parameters allowed the motion to be regarded as occurring in a non-Euclidean space, or as occurring in a Euclidean space, but combined with a strain.

The pre-established harmony described by Klein is nowadays better known as "the unreasonable effectiveness of mathematics" since the appearance of an article [**85**] bearing that title by the late Nobel Prize-winning physicist Eugene Wigner (1902–1995). This article has become famous and is accessible online, for example, at the following website:

`http://www.maths.ed.ac.uk/~aar/papers/wigner.pdf`

For an opposing point of view, see the article by Grattan-Guinness [**35**]. In Chapter 8, we attempt to make this effectiveness appear a bit less unreasonable.

- The relativistic transformation of electric and magnetic fields eliminates an asymmetry in the classical transformation of electric and magnetic fields between two observers in relative motion; and when the relativistic correction is made, it becomes possible to derive Maxwell's two curl equations from the two divergence equations. Although this result is hardly new—I first read about it some 30 years ago—I find that few mathematicians seem to know about it. From the historical point of view, it helps to clarify the main achievement of Einstein's first (1905) paper on the subject. This material forms the subject matter of the optional (starred) Chapter 3.

- In Chapter 4, I compute the relativistic orbit of Mercury and use real data to show that the relativistic portion of the precession of its perihelion is indeed 43 seconds per century. Moreover, if the right-ascension angle $\theta(t)$ at time t is replaced by an angle $\varphi(\tau) = \theta(t)$, where τ is the proper time on Mercury corresponding to time t in the Sun-based coordinates, the orbit is an ellipse in a set of unobservable polar coordinates with φ as the polar angle. (In other words, there is no precession in these coordinates.) This fact illustrates a more general principle: Some classical laws of physics, such as conservation of angular momentum, can be transferred to relativity if observer time is replaced by proper time on the moving object being discussed.
- As already mentioned, my discussion of the precession of the orbit of Mercury is prefaced by two sections aimed at clearing away possible approaches that might have occurred—in fact, did occur—to people reflecting on the problem in the early twentieth century. In both cases (also as already mentioned) I have learned from anonymous reviewers that my thoughts were anticipated and developed in much more detail than I could have managed myself by two prominent mathematical physicists, namely Willem de Sitter and Elie Cartan. Let my explanation of these approaches in my own words serve as a rough approximation to parts of their work. These two explorations, both ultimately unsuccessful as an explanation of planetary motion, nevertheless produce differential equations surprisingly similar to those of the Schwarzschild solution. The "Newtonian geometric" approach, in particular, naturally introduces both the Schwarzschild radius and the elliptic integrals that provide an exact expression of relativistic orbits.
- I have noticed that physics texts tend to avoid the use of elliptic functions, even though their mathematical theory is well known and they have had applications in physics since the eighteenth century—in the study of pendulum motion and the rotation of rigid bodies, for example. There are several places where these functions arise in relativity theory. I use them in Chapter 1 to study a relativistic model of a geocentric universe—essentially giving the mathematical details of the relativistic geometry of a rotating plane, a problem discussed qualitatively by Einstein [**21**]—and in Chapter 4 to give the exact (Schwarzschild) solution of the gravitational equation in the two-body problem.
- In order to put the Lorentz transformation into perspective, I point out that each individual observer uses his own proper time and his own proper space, and these have the same absolute qualities that Newton assigned to a *universal* time and space common to all observers, except that no

material particle can move faster than light relative to that "personal absolute" space. (Just to be clear, two particles in the space *can* move faster than light relative to each other, as judged by an observer at rest in the framework; but neither of them will judge the relative speed to be larger than c.) It is the entangling of the time axis with the spatial axis along the common line of motion when two observers reconcile their coordinates that leads to the bizarre, yet real, phenomena of length and time contraction. A discussion of this transformation provides a healthy caution on the use of vector operations to avoid coordinates. In special relativity theory, vector notation is applicable to the space used by each observer, but is not transferrable between two observers unless a particular set of coordinates is used, since *the Lorentz transformation does not preserve dot and cross products on the spatial portion of space-time*. If Y has velocity $\boldsymbol{u}$ relative to X, and X has velocity $-\boldsymbol{u}$ relative to Y—that is, Y assigns to the velocity of X coordinates that are the negatives of those that X assigns to the velocity of Y—while Z has velocity $\boldsymbol{v}$ relative to Y, vector notation can be used to compute the velocity $\boldsymbol{w}$ that Z has relative to X. But unless $\boldsymbol{u}$ and $\boldsymbol{v}$ are parallel—that is, when expressed in terms of the coordinates used by Y, each is a scalar multiple of the other—this computation cannot be reversed to give the velocity that X has relative to Z through the simple replacement of $\boldsymbol{u}$ by $-\boldsymbol{v}$ and $\boldsymbol{v}$ by $-\boldsymbol{u}$ and computation of the dot products contained in the formula. When X and Y use arbitrary orthonormal coordinate systems, Y will *not* generally assign to the velocity of X the coordinates that are the negatives of those that X assigns to Y. Consequently, these velocities cannot be passed back and forth between different observers using the familiar dot and cross products that each individual observer can use for his own purposes. The observers do not have a common space that they can talk about. Vector notation remains useful for *recording* physical relations, but can be passed from one observer to another only when they both agree to use their common line of motion as the first spatial axis. In other words, one is forced to derive physical relations coordinate-wise, just as Einstein did in 1905, or else adopt a completely new approach to the subject. As this book aims to be elementary, we choose the first of these options.

Other Works on the Subject

As I am not the first person to write a book of this type, let me point out here the main respects in which the present book differs from or resembles others on the subject. This book was begun in 2012, before the publication of the recent book by Anthony Zee ([**88**]), and it was completed before I became aware of the existence of Professor Zee's book. Zee and I have similar pedagogical and humanistic aims. One significant difference is that Zee's book discusses the subject from the standpoint of late twentieth-century physics, a task that would be utterly beyond my ability. Zee's book covers many more topics and in much greater depth than I have done. Except for the Gödel metric, which dates to the 1940s, the present book takes the development of both differential geometry and general relativity only as far as the mid-1920s, and it does not attempt to apply the Einstein field equations to anything outside the simple case of a constant matter density. The most complicated example considered is the Gödel metric. I do explore certain special questions that are more of interest to mathematicians than to physicists.

Outside of Zee's book, perhaps the closest in its pedagogical aim of explaining relativity theory in plain language accessible to undergraduates is the book of the late Richard L. Faber [**27**]. Besides the works of Narlikar, Dray, and Sternberg mentioned above—and, no doubt, dozens of others unknown to me—the subject matter of this book is discussed in the book of Torretti [**81**]. Like Zee's book, Torretti's erudite book is many times more complete and detailed than the present one and contains, as well, detailed historical information on the evolution of Einstein's thought in the years 1913–1915. My pedagogical aim differs from what I perceive to be Torretti's, which appears to be aimed at specialists in the philosophy of science. My goal is to simplify *selected parts* of the theory of relativity, to make them accessible to a person who has studied only the basic three semesters of calculus and two semesters of linear algebra.

Among older works, Bertrand Russell devoted considerable space to an exposition of relativity theory, for example, in [**72**]. As often happens with such books, however, he explained the mechanics of manipulating the formulas without giving much information on the grounds for accepting them as laws of nature. His intended audience also appears to be people well versed in the philosophy of science, and he explores only philosophical issues. He is definitely not writing a textbook for undergraduates.

Given the sheer quantity of writings on this subject—many dozens of works that one can find in almost any library and which are not mentioned here—and my consequent unfamiliarity with most of them, I would not venture to say that anything I have written here is appearing for the first time. Like many other mathematicians, I have too often found that my ideas have also occurred to others, sometimes much earlier. For any passage in this book containing no citation of a source, there are two possible explanations: (1) the ideas contained in the passage are well known and can be found in many standard sources; (2) I have never found the ideas in any source, but have thought them up on my own. In neither case should the absence of a citation be interpreted as a claim of originality on my part.

Background Necessary to Read This Book

The reader is assumed to know the rudiments of advanced calculus,[7] a few techniques for solving differential equations, some linear algebra, and at least the nomenclature of set theory and groups. That is the reason for the word *elementary* in the title. A person with that mathematical background belongs to the audience to whom the present book is addressed. Of course, in order to understand general relativity it is necessary to know something about manifolds and geodesics, and that involves the very simplest parts of the calculus of variations and the existence and uniqueness theorem for the initial-value problem in ordinary differential equations. I have put this material into Volume 2 as Appendices 2, 4, and 5. Appendix 4 grew out of control during the writing, and I was forced to move its topological

[7]This is the name that used to be given to what is nowadays more frequently called elementary real analysis. We actually do not use much of it outside the appendices in Volume 2. Still, it will be helpful if the reader knows the implicit function theorem, Stokes's theorem, the divergence theorem, and the standard criteria for term-wise differentiation and (Riemann) integration of a uniformly convergent sequence.

foundations into Appendix 3. In addition, Appendix 3 contains a large assortment of inessential facts from point-set topology, which may be of intrinsic interest and accustom the reader to thinking in terms of abstract spaces. As for physics, I am assuming that the reader knows the concepts of mass, acceleration, force, momentum, work, and energy in mechanics, along with the basic facts of electricity and magnetism. Many of these concepts, however, are explained further in Chapters 2 and 4 and in Appendices 2 and 5.

Plan of the Work

As this work has grown, I have found it necessary to break it into three volumes.

The first volume consists of eight chapters of material on the main subject of the work, namely special and general relativity. This volume is divided into three parts, as follows: Part 1, which consists of Chapters 1–3, contains the rudiments of the special theory of relativity, explained from a particular point of view, and with some optional excursions into topics that especially intrigue me. Part 2, consisting of Chapters 4–7, is devoted to the general theory of relativity. Part 3, consisting of just the final Chapter 8, is an excursion through the historical and philosophical context of the theory. In it, I attempt to answer metaphysical questions as to the nature of arcane objects such as gravitational fields and explain how, even though we cannot directly experience these things, we can know facts about them and can be justified in regarding them as real. Since I am not a professional philosopher (as will be apparent to those who specialize in this area), my explanations are offered as suggestions to a generally educated reader. Those explanations, elaborated with many examples in this chapter, invoke the harmony between mathematics and the physical world to explain how facts come to be known about unknowable things and use an analogy with the theory of manifolds to explain that the agreement between different methods used to measure physical quantities reinforces the credibility of the theories on which they are based. What I suggest may be a way for broadly educated people to understand approximately what is meant by existence and reality. A high level of precision in this enterprise is, I think, unachievable.

At the end of each chapter, the reader will find a set of problems and questions to help fix the ideas from that chapter. These are almost without exception routine applications of what was discussed in the text, and few of them require any ingenuity to solve. I hope that the reader nevertheless finds them interesting to work. Some of them are intended to make a mathematical or philosophical point.

Sections that represent digressions or material that is not needed for reading what comes after it are marked with an asterisk. The reader who is looking for a minimal introduction to the problem of gravity in free space can omit these sections without losing the thread of the narrative.

Volume 2 arose because the mathematical background involved in Volume 1 is rather large. It consists of six appendices made up of background mathematical material that would have forced digressions if presented in the main narrative of this work. For example, in Chapter 4, on the orbit of Mercury, I have temporarily omitted the mathematical theory that justifies the use of tensor calculus to express the curvature of space-time in order to focus on the computation. The omitted theoretical tensor analysis can be found in Chapters 5 and 6 and Appendices 4 and 6. Taking Volumes 1 and 2 together, the reader should be able to find full mathematical details on all the material. In addition to using these appendices as a way to avoid digressions in the main narrative, I took the opportunity to include among them some topics that I find irresistible. Examples are (1) Euler's wonderful result that a particle moving along a surface but free of tangential acceleration will describe a geodesic on the surface (Appendix 2); (2) Jacobi's elegant last-multiplier principle (Appendix 5), which shows that a system of n ordinary first-order linear differential equations with algebraic coefficients whose divergence vanishes has solutions that can be expressed as quotients of theta functions, provided one can find $n-1$ independent algebraic integrals (functions not identically constant that are constant on the trajectories of a solution); and (3) the whole subject of point-set topology (Appendix 3), which I include as a way for the reader to acquire some practice in visualizing completely abstract spaces, along with the basic facts of the theory, which are used in both real analysis and differential geometry.

Volume 3 contains all the *Mathematica* notebooks that I used to lighten the labor of some of the more complex computations in the book and suggested answers to the exercises in the first two volumes. Eleven *Mathematica* notebooks are referenced in the first volume and one more in the second volume. They are collected as a unit at the beginning of Volume 3. For the convenience of the reader, Volumes 2 and 3 and all twelve *Mathematica* notebooks can be downloaded from my website at the University of Vermont:

`http://www.cems.uvm.edu/~rlcooke/RELATIVITY/`

Acknowledgments

I am grateful to Stephen Wolfram for inventing *Mathematica* and thereby putting enormous computing and graphing power and accuracy in computation into the hands of people of limited patience, who would otherwise face the dreary prospect of spending days or weeks floundering in an attempt to carry out a computation that *Mathematica* enables us to do flawlessly in a fraction of a second. That it will also render a beautiful perspective drawing of the graph of a function of two variables is a delightful bonus.

I wish to thank the anonymous reviewers who vetted the manuscript for publication and made some extremely valuable suggestions, pointing out places where my writing was misleading or revealed too much of my innocence of the full history of this subject. Their advice to tone down the occasional references to controversial political issues was also sound, and I have heeded it.

Special thanks go to my lifelong friend Charles Gillard, with whom I shared philosophical reflections as a teenager when we were both delivering newspapers, at whose wedding I met my wife, and who, now that we are both retired, has been kind enough to read the entire manuscript of the first volume and send me detailed comments, which I have taken into account in the proofreading.

Very special thanks go to my wife Catherine, who endured my getting up at 5:00 AM every morning for two years to work on this project.

Roger Cooke
August 2016