Chapter 1

Physics in Flat Spacetime: Geometric Viewpoint

1.1 Overview

This is a book about classical physics. It is a book about the interpretation of physical phenomena on the large scale, when the particle nature of matter and radiation is secondary to the behavior of these particles in bulk, when their statistical as opposed to their individual properties are important and when their inherent graininess can be smoothed over. We shall take a journey, through spacetime and phase space, though statistical and continuum mechanics, through optics and relativity, to comprehend the fundamental laws of classical physics in their own terms and in relation to quantum physics. Through carefully chosen examples, we shall also show how these laws are being applied to important contemporary problems and we shall uncover some deep connections between the fundamental laws and between the practical techniques that are used to apply them.

In order to bring out these connections, we shall adopt a different viewpoint on the laws of physics than that found in most elementary textbooks. In elementary texts, the laws are expressed in terms of quantities (locations in space or spacetime, momenta of particles, etc.) that are measured in some coordinate system or reference frame. For example, Newtonian vectorial quantities (momenta, electric fields, etc.) are triplets of numbers [e.g., $(1.7, 3.9, -4.2)$] representing the vectors' components on the axes of a spatial coordinate system, and relativistic 4-vectors are quadruplets of numbers representing components on the spacetime axes of some reference frame.

By contrast, in this book, we shall express all physical quantities and laws in a geometric form that is independent of any coordinate system. For example, in Newtonian physics, momenta and electric fields will be vectors described as arrows that live in the 3-dimensional, flat Euclidean space of everyday experience. They require no coordinate system at all for their existence or description—though sometimes coordinates will be useful. We shall state physical laws, e.g. the Lorentz force law, as geometric, coordinate-free relationships between these geometric, coordinate free quantities.

By adopting this geometric viewpoint, we shall gain great conceptual power and often also computational power. For example, when we ignore experiment and simply ask what forms

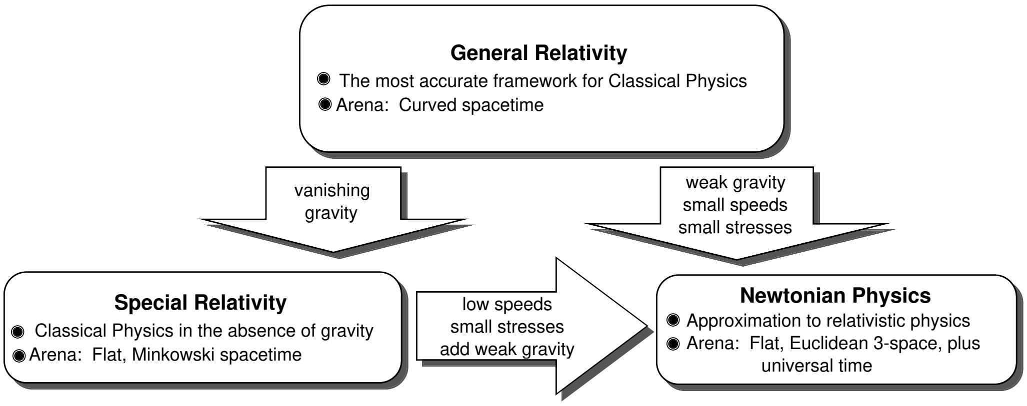


Fig. 1.1: The three frameworks and arenas for the classical laws of physics, and their relationship to each other.

the laws of physics can possibly take (what forms are allowed by the requirement that the laws be geometric), we shall find remarkably little freedom. Coordinate independence strongly constrains the laws (see, e.g., Sec. 1.4 below). This power, together with the elegance of the geometric formulation, suggests that in some deep (ill-understood) sense, Nature's physical laws *are* geometric and have nothing whatsoever to do with coordinates or reference frames.

The mathematical foundation for our geometric viewpoint is *differential geometry* (also often called "tensor analysis" by physicists). This differential geometry can be thought of as an extension of the vector analysis with which all readers should be familiar.

There are three different frameworks for the classical physical laws that scientists use, and correspondingly three different geometric arenas for the laws; cf. Fig. (1.1). *General relativity* is the most accurate classical framework; it formulates the laws as geometric relationships in the arena of *curved 4-dimensional spacetime*. *Special relativity* is the limit of general relativity in the complete absence of gravity; its arena is *flat, 4-dimensional Minkowski spacetime*. *Newtonian physics* is the limit of general relativity when (i) gravity is weak but not necessarily absent, (ii) relative speeds of particles and materials are small compared to the speed of light c, and (iii) all stresses (pressures) are small compared to the total density of mass-energy; its arena is *flat, 3-dimensional Euclidean space* with time separated off and made universal (by contrast with the frame-dependent time of relativity).

In Parts I–V of this book (statistical physics, optics, elasticity theory, fluid mechanics, plasma physics) we shall confine ourselves to the Newtonian and special relativistic formulations of the laws, and accordingly our arenas will be flat Euclidean space and flat Minkowski spacetime. In Part VI we shall extend many of the laws we have studied into the domain of strong gravity (general relativity), i.e., the arena of curved spacetime.

In Parts I and II (statistical physics and optics), in addition to confining ourselves to flat space or flat spacetime, we shall avoid any sophisticated use of curvilinear coordinates; i.e., when using coordinates in nontrivial ways, we shall confine ourselves to Cartesian coordinates in Euclidean space, and Lorentz coordinates in Minkowski spacetime. This chapter is an introduction to all the differential geometric tools that we shall need in these limited arenas.

In Parts III, IV, and V, when studying elasticity theory, fluid mechanics, and plasma physics, we will use curvilinear coordinates in nontrivial ways. As a foundation for them, at the beginning of Part III we will extend our flat-space differential geometric tools to curvilinear coordinate systems (e.g. cylindrical and spherical coordinates). Finally, at the beginning of Part VI, we shall extend our geometric tools to the arena of curved spacetime.

In this chapter we shall alternate back and forth, one section after another, between flat-space differential geometry and the laws of physics, using each to illustrate and illuminate the other. We begin in Sec. 1.2 by recalling the foundational concepts of Newtonian physics and of special relativity. Then in Sec. 1.3 we develop our first set of differential geometric tools: the tools of coordinate-free tensor algebra. In Sec. 1.4 we illustrate our tensor-algebra tools by using them to describe—without any coordinate system or reference frame whatsoever—the kinematics of point particles that move through the Euclidean space of Newtonian physics and through relativity's Minkowski spacetime; the particles are allowed to collide with each other and be accelerated by an electromagnetic field. In Sec. 1.5, we extend the tools of tensor algebra to the domain of Cartesian and Lorentz coordinate systems, and then in Sec. 1.6 we use these extended tensorial tools to restudy the motions, collisions, and electromagnetic accelerations of particles. In Sec. 1.7 we discuss rotations in Euclidean space and Lorentz transformations in Minkowski spacetime, and we develop relativistic spacetime diagrams in some depth and use them to study such relativistic phenomena as length contraction, time dilation, and simultaneity breakdown. In Sec. 1.8 we illustrate the tools we have developed by asking whether the laws of relativity permit a highly advanced civilization to build time machines for traveling backward in time as well as foward. In Sec. 1.9 we develop additional differential geometric tools: directional derivatives, gradients, and the Levi-Civita tensor, and in Sec. 1.10 we use these tools to discuss Maxwell's equations and the geometric nature of electric and magnetic fields. In Sec. 1.11 we develop our final set of geometric tools: volume elements and the integration of tensors over spacetime, and in Sec. 1.12 we use these tools to define the stress tensor of Newtonian physics and relativity's stress-energy tensor, and to formulate very general versions of the conservation of 4-momentum.

1.2 Foundational Concepts

Newtonian Foundational Concepts

The arena for the Newtonian laws is a spacetime composed of the familiar 3-dimensional Euclidean space of everyday experience (which we shall call *3-space*), and a universal time t. Sometimes we shall denote points in 3-space by capital script letters such as $\mathcal{P}$ and $\mathcal{Q}$. These points and the 3-space in which they live require no coordinate system for their definition.

A *scalar* is a single number that we associate with a point, $\mathcal{P}$, in this space. We are interested in scalars that represent physical quantities, e.g., temperature measured on the thermodynamical scale. When a scalar can be associated with all points in some region of space we call it a *scalar field.*

A *vector* in Euclidean 3-space (e.g., the arrow $\Delta\mathbf{x}$ of Fig. 1.2 can be thought of as a straight arrow that reaches from one point, $\mathcal{P}$, to another, $\mathcal{Q}$. Sometimes we shall select one point $\mathcal{O}$ in 3-space as an "origin" and identify all other points, say $\mathcal{Q}$ and $\mathcal{P}$, by their vectorial separations $\mathbf{x}_\mathcal{Q}$ and $\mathbf{x}_\mathcal{P}$ from that origin.

The Euclidean distance $\Delta\sigma$ between two points $\mathcal{P}$ and $\mathcal{Q}$ in 3-space can be measured with a ruler and requires no coordinate system for its definition. (If one does have a coordinate

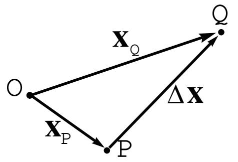


Fig. 1.2: A Euclidean 3-space diagram depicting two points $\mathcal{P}$ and $\mathcal{Q}$, their vectorial separations $\vec{x}_{\mathcal{P}}$ and $\vec{x}_{\mathcal{Q}}$ from the (arbitrarily chosen) origin $\mathcal{O}$, and the vector $\Delta\vec{x} = \vec{x}_{\mathcal{Q}} - \vec{x}_{\mathcal{P}}$ connecting them.

system, it can be computed by the Pythagorean formula.) This distance is also regarded as the length $|\Delta\mathbf{x}|$ of the vector $\Delta\mathbf{x}$ that reaches from $\mathcal{P}$ to $\mathcal{Q}$, and the square of that length is denoted

$$|\Delta\mathbf{x}|^2 \equiv (\Delta\mathbf{x})^2 \equiv (\Delta\sigma)^2 \ . \tag{1.1}$$

Of particular importance is the case when $\mathcal{P}$ and $\mathcal{Q}$ are neighboring points and $\Delta\mathbf{x}$ is a differential quantity $d\mathbf{x}$. We can think of such a vector as residing at $\mathcal{P}$ and if we can associate a vector with every point, then we have a *vector field.* Now the product of a scalar with a vector is still a vector. So if, for example, we consider a single particle or an element of a fluid at two (universal) times, separated by dt, and multiply the displacement of a fluid element $d\mathbf{x}$ by $1/dt$, then we define a new vector, the velocity. Performing this operation at every point defines the velocity field. Similarly, the sum (or difference) of two vectors is also a vector and so taking the difference of two velocity measurements generates the acceleration. Multiplying with a (scalar) mass defines a force; dividing by the charge, can define another vector, the electric field and so on. We can define products of pairs of vectors (e.g., force and displacement) to define a new scalar (e.g., work) or a new vector (e.g., torque). Taking the difference of two scalars or vectors, residing at adjacent points $\mathcal{P}$ and $\mathcal{Q}$ at the same absolute time, defines the standard functions of vector calculus, gradient, divergence. In this fashion, which we trust is quite familiar, and which we shall elucidate and generalize below, we can construct all of the standard scalars and vectors of Newtonian physics. What is important is that these physical quantities also require no coordinate system for their definition. They are geometric objects residing in Euclidean 3-space at a particular time.

It is a fundamental (though often ignored) principle of physics that *the Newtonian physical laws must all be expressible as geometric relationships between these geometric objects and that these relationships do not depend upon any coordinate system or orientation of axes or the time.* We shall return to this principle throughout this book.

A Note on Units

In this text we will be dealing with practical matters and will frequently need to have a quantitative understanding of the magnitude of various physical quantities. This requires us to adopt a particular unit system. We find that the students we teach are about equally divided in preferring cgs/Gaussian units or MKS/SI units. Both of these systems provide a complete and internally consistent set for all of physics and it is a hotly-debated issue as to which of these is the more convenient or aesthetically appealing. We choose not to enter this debate! One's choice of units should not matter and a complete physicist should be able to change from one system to another without thinking. However, when learning new concepts, having to figure out "where the 4 π's go" is a genuine impediment to progress. Our solution to this problem is as follows: We shall use the units that seem most natural for the topic at hand or those which, we judge, constitute the majority usage for the subculture that the topic represents. We shall not pedantically convert cm to m or *vice versa* at every juncture; we trust that the reader can easily make whatever translation is necessary. However, where the equations are actually different, for example as is the case in electromagnetic theory, we shall often provide, in brackets or footnotes, the equivalent equations in the other unit system and enough information for the reader to proceed in his or her preferred scheme.

Special Relativistic Foundational Concepts[1]

Because the nature and geometry of Minkowski spacetime are far less obvious intuitively than those of Euclidean 3-space, we shall need a crutch in our development of the Minkowski foundational concepts. That crutch will be inertial reference frames. We shall use them to develop in turn the following frame-independent Minkowski-spacetime concepts: events, 4-vectors, the principle of relativity, geometrized units, the interval and its invariance, and spacetime diagrams.

An *inertial reference frame* is a (conceptual) three-dimensional latticework of measuring rods and clocks with the following properties: (*i*) The latticework moves freely through spacetime (i.e., no forces act on it), and is attached to gyroscopes so it does not rotate with respect to distant, celestial objects. (*ii*) The measuring rods form an orthogonal lattice and the length intervals marked on them are uniform when compared to, e.g., the wavelength of light emitted by some standard type of atom or molecule; and therefore the rods form an orthonormal, Cartesian coordinate system with the coordinate x measured along one axis, y along another, and z along the third. (*iii*) The clocks are densely packed throughout the latticework so that, ideally, there is a separate clock at every lattice point. (*iv*) The clocks tick uniformly when compared, e.g., to the period of the light emitted by some standard type of atom or molecule; i.e., they are *ideal clocks*. (*v*) The clocks are synchronized by the Einstein synchronization process: If a pulse of light, emitted by one of the clocks, bounces off a mirror attached to another and then returns, the time of bounce t_b as measured by the clock that does the bouncing is the average of the times of emission and reception as measured by the emitting and receiving clock: $t_b = \frac{1}{2}(t_e + t_r)$.[2]

Our second fundamental relativistic concept is the *event*. An event is a precise location in space at a precise moment of time; i.e., a precise location (or "point") in 4-dimensional spacetime. We sometimes will denote events by capital script letters such as $\mathcal{P}$ and $\mathcal{Q}$ — the same notation as for points in Euclidean 3-space; there need be no confusion, since we will avoid dealing with 3-space points and Minkowski-spacetime points simultaneously.

A *4-vector* (also often referred to as a *vector in spacetime*) is a straight arrow $\Delta\vec{x}$ reaching from one event $\mathcal{P}$ to another $\mathcal{Q}$. We often will deal with 4-vectors and ordinary (3-space) vectors simultaneously, so we shall need different notations for them: bold-face Roman font for 3-vectors, $\Delta\mathbf{x}$, and arrowed italic font for 4-vectors, $\Delta\vec{x}$. Sometimes we shall identify an event $\mathcal{P}$ in spacetime by its vectorial separation $\vec{x}_\mathcal{P}$ from some arbitrarily chosen event in spacetime, the "origin" $\mathcal{O}$.

---

[1] For further detail see, e.g., Taylor and Wheeler (1992); pp. 5-29, 51, 53, 54, and 63-70 of Misner, Thorne, and Wheeler (1973), and a forthcoming book by Hartle (2000); and chapter 1 of Schutz (1985).

[2] For a deeper discussion of the nature of ideal clocks and ideal measuring rods see, e.g., pp. 23–29 and 395–399 of Misner, Thorne, and Wheeler (1973).

An inertial reference frame provides us with a coordinate system for spacetime. The coordinates $(x^0, x^1, x^2, x^3) = (t, x, y, z)$ which it associates with an event $\mathcal{P}$ are $\mathcal{P}$'s location (x, y, z) in the frame's latticework of measuring rods, and the time t of $\mathcal{P}$ *as measured by the clock that sits in the lattice at the event's location.* (Many apparent paradoxes in special relativity result from failing to remember that the time t of an event is always measured by a clock that resides at the event, and never by clocks that reside elsewhere in spacetime.)

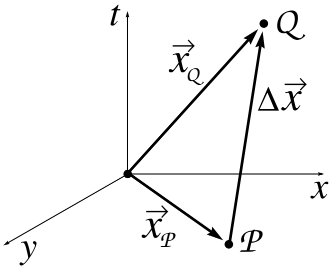


Fig. 1.3: A spacetime diagram depicting two events $\mathcal{P}$ and $\mathcal{Q}$, their vectorial separations $\vec{x}_{\mathcal{P}}$ and $\vec{x}_{\mathcal{Q}}$ from the (arbitrarily chosen) origin, and the vector $\Delta\vec{x} = \vec{x}_{\mathcal{Q}} - \vec{x}_{\mathcal{P}}$ connecting them.

It is useful to depict events on *spacetime diagrams*, in which the time coordinate $t = x^0$ of some inertial frame is plotted upward, and two of the frame's three spatial coordinates, $x = x^1$ and $y = x^2$, are plotted horizontally. Figure Two events $\mathcal{P}$ and $\mathcal{Q}$ are shown there, along with their vectorial separations $\vec{x}_{\mathcal{P}}$ and $\vec{x}_{\mathcal{Q}}$ from the origin and the vector $\Delta\vec{x} = \vec{x}_{\mathcal{Q}} - \vec{x}_{\mathcal{P}}$ that separates them from each other. The coordinates of $\mathcal{P}$ and $\mathcal{Q}$, which are the same as the components of $\vec{x}_{\mathcal{P}}$ and $\vec{x}_{\mathcal{Q}}$ in this coordinate system, are $(t_{\mathcal{P}}, x_{\mathcal{P}}, y_{\mathcal{P}}, z_{\mathcal{P}})$ and $(t_{\mathcal{Q}}, x_{\mathcal{Q}}, y_{\mathcal{Q}}, z_{\mathcal{Q}})$; and correspondingly, the components of $\Delta\vec{x}$ are

$$\begin{aligned} \Delta x^0 &= \Delta t = t_{\mathcal{Q}} - t_{\mathcal{P}}, \quad & \Delta x^1 &= \Delta x = x_{\mathcal{Q}} - x_{\mathcal{P}}, \\ \Delta x^2 &= \Delta y = y_{\mathcal{Q}} - y_{\mathcal{P}}, \quad & \Delta x^3 &= \Delta z = z_{\mathcal{Q}} - z_{\mathcal{P}}\ . \end{aligned} \tag{1.2}$$

We shall denote these components of $\Delta\vec{x}$ more compactly by Δx^α, where the α index (and every other lower case Greek index that we shall encounter) takes on values $t = 0$, $x = 1$, $y = 2$, and $z = 3$. Similarly, in 3-dimensional Euclidean space, we shall denote the Cartesian components $\Delta\mathbf{x}$ of a vector separating two events by Δx^j, where the j (and every other lower case Latin index) takes on the values $x = 1$, $y = 2$, and $z = 3$.

When the physics or geometry of a situation being studied suggests some preferred inertial frame (e.g., the frame in which some piece of experimental apparatus is at rest), then we typically will use as axes for our spacetime diagrams the coordinates of that preferred frame. On the other hand, when our situation provides *no* preferred inertial frame, or when we wish to emphasize a frame-independent viewpoint, we shall use as axes the coordinates of a completely arbitrary inertial frame and we shall think of the spacetime diagram as depicting spacetime in a coordinate-independent, frame-independent way.

The coordinate system (t, x, y, z) provided by an inertial frame is sometimes called an *inertial coordinate system* and sometimes a *Lorentz coordinate system* [because it was Lorentz (1904) who first studied the relationship of one such coordinate system to another, the

Lorentz transformation]. We shall use these terms interchangeably, and we shall also use the term *Lorentz frame* interchangeably with *inertial frame*. A physicist or other intelligent being who resides in a Lorentz frame and makes measurements using its latticework of rods and clocks will be called an *observer*.

Although events are often described by their coordinates in a Lorentz reference frame, and vectors by their components (coordinate differences), it should be obvious that the concepts of an event and a vector need not rely on any coordinate system whatsoever for their definition. For example, the event $\mathcal{P}$ of the birth of Isaac Newton, and the event $\mathcal{Q}$ of the birth of Albert Einstein are readily identified without coordinates. They can be regarded as *points* in spacetime, and their separation vector is the straight arrow reaching through spacetime from $\mathcal{P}$ to $\mathcal{Q}$. Different observers in different inertial frames will attribute different coordinates to each birth and different components to the births' vectorial separation; but all observers can agree that they are talking about the same events $\mathcal{P}$ and $\mathcal{Q}$ in spacetime and the same separation vector $\Delta\vec{x}$. In this sense, $\mathcal{P}$, $\mathcal{Q}$, and $\Delta\vec{x}$ are *frame-independent, geometric objects* (points and arrows) that reside in spacetime.

The *principle of relativity* states that *Every (special relativistic) law of physics must be expressible as a geometric, frame-independent relationship between geometric, frame-independent objects*, i.e. objects such as points in spacetime and vectors, which represent physical quantities such as events and particle momenta.

Since the laws are all geometric (i.e., unrelated to any reference frame), there is no way that they can distinguish one inertial reference frame from any other. This leads to an alternative form of the principle of relativity (one commonly used in elementary textbooks and equivalent to the above): *All the (special relativistic) laws of physics are the same in every inertial reference frame, everywhere in spacetime.* A more operational version of this principle is the following: Give identical instructions for a specific physics experiment to two different observers in two different inertial reference frames at the same or different locations in Minkowski (i.e., gravity-free) spacetime. The experiment must be self-contained, i.e., it must not involve observations of particles or fields that come to the observer from the external universe. For example, an *unacceptable* experiment would be a measurement of the anisotropy of the Universe's cosmic microwave radiation and a computation therefrom of the observer's velocity relative to the radiation's mean rest frame. An *acceptable* experiment would be a measurement of the speed of light using the rods and clocks of the observer's own frame. The principle of relativity says that in this or any other self-contained experiment, the two observers in their two different inertial frames must obtain identically the same experimental results—to within the accuracy of their experimental techniques. Since the experimental results are governed by the (nongravitational) laws of physics, this is equivalent to the statement that all physical laws are the same in the two inertial frames.

Perhaps the most central of special relativistic laws is the one stating that *the speed of light c in vacuum is frame-independent*, i.e., is a constant, independent of the inertial reference frame in which it is measured. It is illustrative to see how this comes about from the laws of electromagnetism (which we assume to be famiiliar) applied in one reference frame. Suppose that we have a large charge Q and a test charge q. There will be a radial electrostatic force and an associated acceleration $a_{\rm es}$ between them, $\propto Qq/\Delta|\mathbf{x}|^2$, when they are separated by a distance $|\Delta\mathbf{x}|$; this force can be measured through their mutual acceleration. Now take a long straight wire, with high resistance, and use it to connect Q to earth and allow the charge to flow along this wire with an initial decay time Δt. The current $I \propto Q/\Delta t$ can then be measured. Place q the same distance $|\Delta \mathbf{x}|$ from the wire and then start it moving with speed v parallel to the wire. There will be a measurable radial electromagnetic acceleration $a_{\rm em}$ acting on q. As the reader can verify, we can use the ratio of these forces to predict the speed of light:

$$c = \left(\frac{2 a_{\rm es} |\Delta \mathbf{x}| v}{a_{\rm em} \Delta t} \right)^{1/2} . \tag{1.3}$$

This (quite impractical) thought experiment demonstrates that, provided one is prepared to trust the laws of electromagnetism and their famous consequence, electromagnetic radiation, then the speed of light is a derivable quantity and all physicists in all reference frames should measure the same value for it in accordance with the principle of relativity. This need not be an additional postulate underlying relativity theory.

The constancy of the speed of light has been verified with nine-digit accuracy in an era when the units of length (centimeters) and the units of time (seconds) were defined independently. More recently, the constancy has become so universally accepted that it has been used to redefine the centimeter (which was hard to measure precisely) in terms of the second (which is much easier to measure with modern technology): The centimeter is related to the second in such a way that the speed of light is precisely $c = 2.9979245 \times 10^{10}$ cm/s [KIP: CHECK]; i.e., one centimeter is the distance traveled by light in $1/(2.9979245 \times 10^{10}$seconds).

Because of this constancy of the light speed, it is permissible when studying special relativity to set c to unity. Doing so is equivalent to the relationship

$$c = 2.9979245 \times 10^{10} \text{cms}^{-1} = 1 \tag{1.4}$$

between seconds and centimeters; i.e., equivalent to

$$1 \text{ second} = 2.9979245 \times 10^{10} \text{ cm} . \tag{1.5}$$

We shall refer to units in which $c = 1$ as *geometrized units*, and we shall adopt them throughout this book, when dealing with relativistic physics, since they make equations look much simpler. Occasionally it will be useful to restore the factors of c to an equation, thereby converting it to ordinary (cgs or mks) units. This restoration is achieved easily using dimensional considerations. For example, the equivalence of mass m and energy E is written in geometrized units as $E = m$. In cgs units E has dimensions ergs = gram cm^2/sec^2, while m has dimensions of grams, so to make $E = m$ dimensionally correct we must multiply the right side by a power of c that has dimensions cm^2/sec^2, i.e. by c^2; thereby we obtain $E = mc^2$.

We turn, next, to another fundamental concept, the *interval* $(\Delta s)^2$ between the two events $\mathcal{P}$ and $\mathcal{Q}$ whose separation vector is $\Delta \vec{x}$. In a specific but arbitrary inertial reference frame, $(\Delta s)^2$ is given by

$$(\Delta s)^2 \equiv -(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = -(\Delta t)^2 + \sum_{i,j} \delta_{ij} \Delta x^i \Delta x^j \; ; \tag{1.6}$$

cf. Eq. (1.2). Here δ_{ij} is the Kronecker delta, (unity if $j = k$; zero otherwise) and the spatial indices i and j are summed over 1, 2, 3. If $(\Delta s)^2 > 0$, the events $\mathcal{P}$ and $\mathcal{Q}$ are said to have a *spacelike* separation; if $(\Delta s)^2 = 0$, their separation is *null* or *lightlike*; and if $(\Delta s)^2 < 0$, their separation is *timelike*. For timelike separations, $(\Delta s)^2 < 0$ implies that Δs is imaginary; to avoid dealing with imaginary numbers, we describe timelike intervals by

$$(\Delta\tau)^2 \equiv -(\Delta s)^2 \; , \tag{1.7}$$

whose square root $\Delta\tau$ is real.

The coordinate separation between $\mathcal{P}$ and $\mathcal{Q}$ depends on one's reference frame; i.e., if $\Delta x^{\alpha'}$ and Δx^{α} are the coordinate separations in two different frames, then $\Delta x^{\alpha'} \neq \Delta x^{\alpha}$. Despite this frame dependence, the principle of relativity forces the interval $(\Delta s)^2$ to be the same in all frames:

$$\begin{aligned}(\Delta s)^2 &= -(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \\ &= -(\Delta t')^2 + (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2\end{aligned} \tag{1.8}$$

[Exercise 1.2].

Because of its frame-invariance, the interval $(\Delta s)^2$ can be regarded as a geometric property of the vector $\Delta\vec{x}$ that reaches from $\mathcal{P}$ to $\mathcal{Q}$; we shall call it the *squared length* $(\Delta\vec{x})^2$ of $\Delta\vec{x}$:

$$(\Delta\vec{x})^2 \equiv (\Delta s)^2 \; . \tag{1.9}$$

This invariant interval between two events is as fundamental to Minkowski spacetime as the Euclidean distance between two points is to flat 3-space. Just as the Euclidean distance gives rise to the geometry of 3-space, as embodied, e.g., in Euclid's axioms, so the interval gives rise to the geometry of spacetime, which we shall be exploring. If this spacetime geometry were as intuitively obvious to humans as is Euclidean geometry, we would not need the crutch of inertial reference frames to arrive at it. Nature (presumably) has no need for such a crutch. To Nature (it seems evident), the geometry of Minkowski spacetime, as embodied in the invariant interval, is among the most fundamental aspects of physical law.

Before we leave this central idea, we should emphasize that vacuum electromagnetic radiation is not the only the only type of wave. In this course, we shall encounter dispersive media, like optical fibers or plasmas, where signals travel slower than c; we shall analyze sound waves and seismic waves where the governing laws do not involve electromagnetism at all. How do these fit into our special relativistic framework? The answer is simple. Each of these waves requires a background medium that is at rest in one particular frame (not necessarily inertial) and the velocity of the wave, specifically the group velocity, is most simply calculated in this frame *from the fundamental laws*. We can then use the kinematic rules of Lorentz transformation to compute the velocity in another frame. However if we had chosen to compute the wave speed in the second frame directly, *using the same fundamental laws*, we would have gotten the same answer, albeit with the expenditure of greater effort. All waves are in full compliance with the principle of relativity. What is special about vacuum electromagnetic waves and, by extension, photons is that no medium (or "ether" as it used to be called) is needed for them to propagate. Their speed is therefore the same in all frames.

This raises an interesting question. What about other waves that do not require a background medium? What about electron de Broglie waves? Here the fundamental wave equation, Schrödinger's or Dirac's, is mathematically different from Maxwell's and contains an important parameter, the electron rest mass. This allows the fundamental laws of relativistic quantum mechanics to be written in a form that is the same in all inertial reference frames and which allows an electron, considered as either a wave or a particle, to travel at a different speed when measured in a different frame.

So, what then about non-electromagnetic waves that do not have an associated rest mass? For a long while, we thought that neutrinos provided a good example, but we now appreciate that they too, like electrons, have rest masses. However, there are particles that have not yet been detected like photinos (the hypothesized, supersymmetric partners to photons) or gravitons (and their associated gravitational waves that we shall discuss in Chapter 26) that are believed to exist without a rest mass (or an ether!), just like photons. Must these travel at the same speed as photons? The answer to this question, according to the principle of relativity, is "yes". The reason is simple. Suppose there were two such waves (or particles) whose governing laws led to different speeds, c and $c' < c$ each the same in all reference frames. If we then move with speed c' in the direction of propagation of the second wave, we would bring it to rest, in conflict with our hypothesis. Therefore all signals, whose governing laws require them to travel with a speed that has no governing parameters must travel with a unique speed which we call "c". The speed of light is more fundamental to relativity than light itself!

****

EXERCISES

Exercise 1.1 *Practice: Geometrized Units*
Convert the following equations from the geometrized units in which they are written to Gaussian cgs units:

(a) The "Planck time" t_P expressed in terms of Newton's gravitation constant G and Planck's constant $\hbar$, $t_P = \sqrt{G\hbar}$.

(b) The Lorentz force law $m d\mathbf{v}/dt = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$.

(c) The expression $\mathbf{p} = \hbar\omega\mathbf{n}$ for the momentum $\mathbf{p}$ of a photon in terms of its angular frequency ω and direction $\mathbf{n}$ of propagation.

How tall are you, in seconds? How old are you, in centimeters?

Exercise 1.2 *Derivation: Invariance of the Interval*
Using a thought experiment, show that the Principle of Relativity, together with the constancy of the speed of light $c = 1$, implies the invariance a timelike interval between two events. [Hint: see Sec. 3.7 of Taylor and Wheeler (1992).]

****

1.3 Tensor Algebra Without a Coordinate System

We now pause in our development of the geometric view of physical law, to introduce, in a coordinate-free way, some fundamental concepts of differential geometry: tensors, the inner product, the metric tensor, the tensor product, and contraction of tensors. In this section we shall allow the space in which the concepts live to be either 4-dimensional Minkowski spacetime, or 3-dimensional Euclidean space; we shall denote its dimensionality by N; and we shall use spacetime's arrowed notation $\vec{A}$ for vectors even though the space might be Euclidean 3-space.

We have already defined a vector $\vec{A}$ as a straight arrow from one point, say $\mathcal{P}$, in our space to another, say $\mathcal{Q}$. Because our space is flat, there is a unique and obvious way to transport such an arrow from one location to another, keeping its length and direction unchanged.[3] Accordingly, we shall regard vectors as unchanged by such transport. This enables us to ignore the issue of where in space a vector actually resides; it is completely determined by its direction and its length.

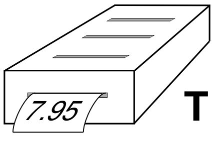


Fig. 1.4: A rank-3 tensor $\mathbf{T}$.

A *rank-n tensor* $\mathbf{T}$ is, by definition, a real-valued, linear function of n vectors. Pictorially we shall regard $\mathbf{T}$ as a box (Fig. 1.4) with n slots in its top, into which are inserted n vectors, and one slot in its end, out of which rolls computer paper with a single real number printed on it: the value that the tensor $\mathbf{T}$ has when evaluated as a function of the n inserted vectors. Notationally we shall denote the tensor by a bold-face sans-serif character $\mathbf{T}$

$$\mathbf{T}(\underbrace{\_,\_,\_,\_}) \ . \tag{1.10}$$

$\nwarrow$ n slots in which to put the vectors

If $\mathbf{T}$ is a rank-3 tensor (has 3 slots) as in Fig. 1.4, then its value on the vectors $\vec{A}, \vec{B}, \vec{C}$ will be denoted $\mathbf{T}(\vec{A}, \vec{B}, \vec{C})$. Linearity of this function can be expressed as

$$\mathbf{T}(e\vec{E} + f\vec{F}, \vec{B}, \vec{C}) = e\mathbf{T}(\vec{E}, \vec{B}, \vec{C}) + f\mathbf{T}(\vec{F}, \vec{B}, \vec{C}) \ , \tag{1.11}$$

where e and f are real numbers, and similarly for the second and third slots.

We have already defined the *squared length* $(\vec{A})^2 \equiv \vec{A}^2$ of a vector $\vec{A}$ as the squared distance (in 3-space) or interval (in spacetime) between the points at its tail and its tip. The *inner product* $\vec{A} \cdot \vec{B}$ of two vectors is defined in terms of the squared length by

$$\vec{A} \cdot \vec{B} \equiv \frac{1}{4}\left[(\vec{A} + \vec{B})^2 - (\vec{A} - \vec{B})^2\right] \ . \tag{1.12}$$

---

[3] This is not so in curved spaces, as we shall see in Part VI.

In Euclidean space this is the standard inner product, familiar from elementary geometry.

Because the inner product $\vec{A}\cdot\vec{B}$ is a linear function of each of its vectors, we can regard it as a tensor of rank 2. When so regarded, the inner product is denoted $\mathbf{g}(\_\,,\_)$ and is called the *metric tensor*. In other words, the metric tensor $\mathbf{g}$ is that linear function of two vectors whose value is given by

$$\mathbf{g}(\vec{A},\vec{B}) \equiv \vec{A}\cdot\vec{B}\ . \tag{1.13}$$

Notice that, because $\vec{A}\cdot\vec{B} = \vec{B}\cdot\vec{A}$, the metric tensor is *symmetric* in its two slots; i.e., one gets the same real number independently of the order in which one inserts the two vectors into the slots:

$$\mathbf{g}(\vec{A},\vec{B}) = \mathbf{g}(\vec{B},\vec{A}) \tag{1.14}$$

With the aid of the inner product, we can regard any vector $\vec{A}$ as a tensor of rank one: The real number that is produced when an arbitrary vector $\vec{C}$ is inserted into $\vec{A}$'s slot is

$$\vec{A}(\vec{C}) \equiv \vec{A}\cdot\vec{C}\ . \tag{1.15}$$

¿From three (or any number of) vectors $\vec{A}$, $\vec{B}$, $\vec{C}$ we can construct a tensor, their *tensor product*, defined as follows:

$$\vec{A}\otimes\vec{B}\otimes\vec{C}(\vec{E},\vec{F},\vec{G}) \equiv \vec{A}(\vec{E})\vec{B}(\vec{F})\vec{C}(\vec{G}) = (\vec{A}\cdot\vec{E})(\vec{B}\cdot\vec{F})(\vec{C}\cdot\vec{G})\ . \tag{1.16}$$

Here the first expression is the notation for the value of the new tensor, $\vec{A}\otimes\vec{B}\otimes\vec{C}$ evaluated on the three vectors $\vec{E}$, $\vec{F}$, $\vec{G}$; the middle expression is the ordinary product of three real numbers, the value of $\vec{A}$ on $\vec{E}$, the value of $\vec{B}$ on $\vec{F}$, and the value of $\vec{C}$ on $\vec{G}$; and the third expression is that same product with the three numbers rewritten as scalar products. Similar definitions can be given (and should be obvious) for the tensor product of any two or more tensors of any rank; for example, if $\mathbf{T}$ has rank 2 and $\mathbf{S}$ has rank 3, then

$$\mathbf{T}\otimes\mathbf{S}(\vec{E},\vec{F},\vec{G},\vec{H},\vec{J}) \equiv \mathbf{T}(\vec{E},\vec{F})\mathbf{S}(\vec{G},\vec{H},\vec{J})\ . \tag{1.17}$$

One last geometric (i.e. frame-independent) concept we shall need is *contraction*. We shall illustrate this concept first by a simple example, then give the general definition. From two vectors $\vec{A}$ and $\vec{B}$ we can construct the tensor product $\vec{A}\otimes\vec{B}$ (a second-rank tensor), and we can also construct the scalar product $\vec{A}\cdot\vec{B}$ (a real number, i.e. a *scalar*, i.e. a *rank-0 tensor*). The process of contraction is the construction of $\vec{A}\cdot\vec{B}$ from $\vec{A}\otimes\vec{B}$

$$\mathrm{contraction}(\vec{A}\otimes\vec{B}) \equiv \vec{A}\cdot\vec{B}\ . \tag{1.18}$$

One can show fairly easily using component techniques (Sec. 1.5 below) that any second-rank tensor $\mathbf{T}$ can be expressed as a sum of tensor products of vectors, $\mathbf{T} = \vec{A}\otimes\vec{B}+\vec{C}\otimes\vec{D}+\ldots$; and correspondingly, it is natural to define the contraction of $\mathbf{T}$ to be $\mathrm{contraction}(\mathbf{T}) = \vec{A}\cdot\vec{B}+\vec{C}\cdot\vec{D}+\ldots$. Note that this contraction process lowers the rank of the tensor by two, from 2 to 0. Similarly, for a tensor of rank n one can construct a tensor of rank $n-2$ by contraction, but in this case one must specify which slots are to be contracted. For example, if T is a third rank tensor, expressible as $\mathsf{T} = \vec{A} \otimes \vec{B} \otimes \vec{C} + \vec{E} \otimes \vec{F} \otimes \vec{G} + \ldots$, then the contraction of T on its first and third slots is the rank-1 tensor (vector)

$$\text{1\&3contraction}(\vec{A} \otimes \vec{B} \otimes \vec{C} + \vec{E} \otimes \vec{F} \otimes \vec{G} + \ldots) \equiv (\vec{A} \cdot \vec{C})\vec{B} + (\vec{E} \cdot \vec{G})\vec{F} + \ldots \; . \tag{1.19}$$

All the concepts developed in this section (vectors, tensors, metric tensor, inner product, tensor product, and contraction of a tensor) can be carried over, with no change whatsoever, into *any* vector space[4] that is endowed with a concept of squared length.

1.4 Particle Kinetics and Lorentz Force Without a Reference Frame

In this section we shall illustrate our geometric viewpoint by formulating the laws of motion for particles, first in Newtonian physics and then in special relativity.

Newtonian Particle Kinetics

In Newtonian physics, a classical particle moves through Euclidean 3-space, as universal time t passes. At time t it is located at some point $\mathbf{x}(t)$ (its *position*). The function $\mathbf{x}(t)$ represents a curve in 3-space, the particle's *trajectory.* The particle's *velocity* $\mathbf{v}(t)$ is the time derivative of its position, its *momentum* $\mathbf{p}(t)$ is the product of its mass m and velocity, and its *acceleration* $\mathbf{a}(t)$ is the time derivative of its velocity

$$\mathbf{v}(t) = d\mathbf{x}/dt \; , \quad \mathbf{p}(t) = m\mathbf{v}(t), \quad \mathbf{a}(t) = d\mathbf{v}/dt = d^2\mathbf{x}/dt^2 . \tag{1.20}$$

Since points in 3-space are geometric objects (defined independently of any coordinate system), so also are the trajectory $\mathbf{x}(t)$, the velocity, the momentum, and the acceleration. (Physically, of course, the velocity has an ambiguity; it depends on one's standard of rest. However, some arbitrary choice of standard of rest has been built into our formalism by our specific choice of the Euclidean 3-space.)

Newton's second law of motion states that the particle's momentum can change only if a force $\mathbf{F}$ acts on it, and that its change is given by

$$d\mathbf{p}/dt = m\mathbf{a} = \mathbf{F} \; . \tag{1.21}$$

If the force is produced by an electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$, then this law of motion takes the familiar Lorentz-force form

$$d\mathbf{p}/dt = \mathbf{E} + \mathbf{v} \times \mathbf{B} \tag{1.22}$$

(here we have used the vector cross product, which will not be introduced formally until Sec. 1.7 below). Obviously, these laws of motion are geometric relationships between geometric objects.

Relativistic Particle Kinetics

---

[4] or, more precisely, any vector space over the real numbers. If the vector space's scalars are complex numbers, as in quantum mechanics, then slight changes are needed.

In special relativity, a particle moves through 4-dimensional spacetime along a curve (its *world line*) which we shall denote, in frame-independent notation, by $\vec{x}(\tau)$. Here τ is time as measured by an ideal clock that the particle carries (the particle's *proper time*), and $\vec{x}$ is the location of the particle in spacetime when its clock reads τ (or, equivalently, the vector from the arbitrary origin to that location).

The particle typically will experience an acceleration as it moves—e.g., an acceleration produced by an external electromagnetic field. This raises the question of how the acceleration affects the ticking rate of the particle's clock. We define the accelerated clock to be *ideal* if its ticking rate is totally unaffected by its acceleration, i.e., if it ticks at the same rate as a freely moving (inertial) ideal clock that is momentarily at rest with respect to it. The builders of inertial guidance systems for airplanes and missiles always try to make their clocks as acceleration-independent, i.e., as ideal, as possible.

We shall refer to the inertial frame in which a particle is momentarily at rest as its *momentarily comoving inertial frame* or *momentary rest frame*. Since the particle's clock is ideal, a tiny interval $\Delta\tau$ of its proper time is equal to the lapse of coordinate time in its momentary rest frame, $\Delta\tau = \Delta t$. Moreover, since the two events $\vec{x}(\tau)$ and $\vec{x}(\tau+\Delta\tau)$ on the clock's world line occur at the same spatial location in its momentary rest frame, $\Delta x^i = 0$ (where $i = 1,\ 2,\ 3$), the invariant interval between those events is $(\Delta s)^2 = -(\Delta t)^2 + \sum_{i,j} \Delta x^i \Delta x^j \delta_{ij} = -(\Delta t)^2 = -(\Delta\tau)^2$. This shows that *the particle's proper time* τ *is equal to the square root of the invariant interval,* $\tau = \sqrt{-s^2}$, *along its world line.*

Figure 1.5 shows the world line of the accelerated particle in a spacetime diagram where the axes are coordinates of an *arbitrary* Lorentz frame. This diagram is intended to emphasize the world line as a frame-independent, geometric object. Also shown in the figure is the particle's 4-*velocity* $\vec{u}$, which (by analogy with the velocity in 3-space) is the time derivative of its position:

$$\vec{u} \equiv d\vec{x}/d\tau \ . \tag{1.23}$$

This derivative is defined by the usual limiting process

$$\frac{d\vec{x}}{d\tau} \equiv \lim_{\Delta\tau\to 0} \frac{\vec{x}(\tau+\Delta\tau) - \vec{x}(\tau)}{\Delta\tau} \ . \tag{1.24}$$

The squared length of the particle's 4-velocity is easily seen to be -1:

$$\vec{u}^2 \equiv \mathbf{g}(\vec{u},\vec{u}) = \frac{d\vec{x}}{d\tau}\cdot\frac{d\vec{x}}{d\tau} = \frac{d\vec{x}\cdot d\vec{x}}{(d\tau)^2} = -1 \ . \tag{1.25}$$

The last equality follows from the fact that $d\vec{x}\cdot d\vec{x}$ is the squared length of $d\vec{x}$ which equals the invariant interval $(\Delta s)^2$ along it, and $(d\tau)^2$ is minus that invariant interval.

The particle's 4-momentum is the product of its 4-velocity and rest mass

$$\vec{p} \equiv m\vec{u} = m d\vec{x}/d\tau \equiv d\vec{x}/d\zeta \ . \tag{1.26}$$

Here the parameter ζ is a renormalized version of proper time,

$$\zeta \equiv \tau/m \ . \tag{1.27}$$

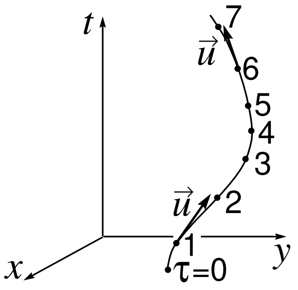


Fig. 1.5: Spacetime diagram showing the world line $\vec{x}(\tau)$ and 4-velocity $\vec{u}$ of an accelerated particle. Note that the 4-velocity is tangent to the world line.

This ζ, and any other renormalized version of proper time with position-independent renormalization factor, are called *affine parameters* for the particle's world line. Expression (1.26), together with the unit length of the 4-velocity $\vec{u}^2 = -1$, implies that the squared length of the 4-momentum is

$$\vec{p}^2 = -m^2 \,. \tag{1.28}$$

In quantum theory a particle is described by a relativistic wave function which, in the geometric optics limit (Chapter 6), has a wave vector $\vec{k}$ that is related to the classical particle's 4-momentum by

$$\vec{k} = \vec{p}/\hbar \,. \tag{1.29}$$

The above formalism is valid only for particles with nonzero rest mass, $m \neq 0$. The corresponding formalism for a *particle with zero rest mass* can be obtained from the above by taking the limit as $m \to 0$ and $d\tau \to 0$ with the quotient $d\zeta = d\tau/m$ held finite. More specifically, the 4-momentum of a zero-rest-mass particle is well defined (and participates in the conservation law to be discussed below), and it is expressible in terms of the particle's affine parameter ζ by Eq. (1.26)

$$\vec{p} = \frac{d\vec{x}}{d\zeta} \,. \tag{1.30}$$

However, the particle's 4-velocity $\vec{u} = \vec{p}/m$ is infinite and thus undefined; and proper time $\tau = m\zeta$ ticks vanishingly slowly along its world line and thus is undefined. Because proper time is the square root of the invariant interval along the world line, the interval between two neighboring points on the world line vanishes identically; and correspondingly *the world line of a zero-rest-mass particle is null.* (By contrast, since $d\tau^2 > 0$ and $ds^2 < 0$ along the world line of a particle with finite rest mass, *the world line of a finite-rest-mass particle is timelike.*)

The 4-momenta of particles are important because of the *law of conservation of 4-momentum* (which, as we shall see in Sec. 1.6, is equivalent to the conservation laws for energy and ordinary momentum): If a number of "initial" particles, named $A = 1, 2, 3, \ldots$ enter a restricted region of spacetime $\mathcal{V}$ and there interact strongly to produce a new set of "final" particles, named $\bar{A} = \bar{1}, \bar{2}, \bar{3}, \ldots$ (Fig. 1.6), then the total 4-momentum of the final particles must be be the same as the total 4-momentum of the initial ones:

$$\sum_{\bar{A}} \vec{p}_{\bar{A}} = \sum_{A} \vec{p}_{A} \; . \tag{1.31}$$

Note that this law of 4-momentum conservation is expressed in frame-independent, geometric language—in accord with Einstein's insistence that all the laws of physics should be so expressible.

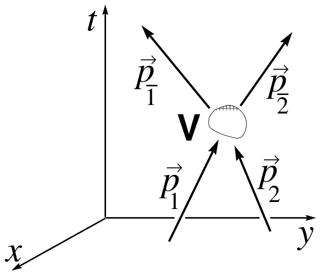


Fig. 1.6: Spacetime diagram depicting the law of 4-momentum conservation for a situation where two particles, numbered 1 and 2 enter an interaction region $\mathcal{V}$ in spacetime, there interact strongly, and produce two new particles, numbered $\bar{1}$ and $\bar{2}$. The sum of the final 4-momenta, $\vec{p}_{\bar{1}} + \vec{p}_{\bar{2}}$, must be equal to the sum of the initial 4-momenta, $\vec{p}_1 + \vec{p}_2$.

If a particle moves freely (no external forces and no collisions with other particles), then its 4-momentum $\vec{p}$ will be conserved along its world line, $d\vec{p}/d\zeta = 0$. Since $\vec{p}$ is tangent to the world line, this means that the direction of the world line never changes; i.e., the free particle moves along a straight line through spacetime. To change the particle's 4-momentum, one must act on it with a 4-force $\vec{F}$,

$$d\vec{p}/d\tau = \vec{F} \; . \tag{1.32}$$

If the particle is a fundamental one (e.g., photon, electron, proton), then the 4-force must leave its rest mass unchanged,

$$0 = dm^2/d\tau = -d\vec{p}^2/d\tau = -2\vec{p}\cdot d\vec{p}/dt = -2\vec{p}\cdot\vec{F} \; ; \tag{1.33}$$

i.e., the 4-force must be orthogonal to the 4-momentum.

As a specific example, consider a fundamental particle with charge q and rest mass $m \neq 0$, interacting with an electromagnetic field. It experiences an electromagnetic 4-force whose relativistic form we shall deduce from simple geometric considerations. The Newtonian version of the electromagnetic force [Eq. (1.22)] is proportional to q and contains one piece (electric) that is independent of velocity $\mathbf{v}$, and a second piece (magnetic) that is linear in $\mathbf{v}$. It is reasonable to expect that, in order to produce this Newtonian limit, the relativistic 4-force will be proportional to q and will be linear in the 4-velocity $\vec{u}$. Linearity means there must exist some second-rank tensor $\mathbf{F}(\_,\_)$ (the "electromagnetic field tensor") such that

$$d\vec{p}/d\tau = \vec{F}(\_) = q\mathbf{F}(\_,\vec{u}) \; . \tag{1.34}$$

Because the 4-force $\vec{F}$ must be orthogonal to the particle's 4-momentum and thence also to its 4-velocity, $\vec{F} \cdot \vec{u} \equiv \vec{F}(\vec{u}) = 0$, expression (1.34) must vanish when $\vec{u}$ is inserted into its empty slots. In other words, for all timelike unit-length vectors $\vec{u}$,

$$\mathbf{F}(\vec{u}, \vec{u}) = 0 \ . \tag{1.35}$$

It is an instructive exercise (Ex. 1.3) to show that this is possible only if **F** is antisymmetric, so the electromagnetic 4-force is

$$d\vec{p}/d\tau = q\mathbf{F}(\_, \vec{u}) \ , \quad \text{where } \mathbf{F}(\vec{A}, \vec{B}) = -\mathbf{F}(\vec{B}, \vec{A}) \text{ for all } \vec{A} \text{ and } \vec{B} \ . \tag{1.36}$$

This is the relativistic form of the Lorentz force law. In Sec. 1.10 below, we shall deduce the relationship of **F** to the electric and magnetic fields, and the relationship of this relativistic Lorentz force to its Newtonian form (1.22).

This discussion of particle kinematics and the electromagnetic force is elegant, but perhaps unfamiliar. In Secs. 1.6 and 1.10 we shall see that it is equivalent to the more elementary (but more complex) formalism based on components of vectors.

****

EXERCISES

Exercise 1.3 *Derivation and Problem: Relativistic Lorentz force law*
Show that Eq. (1.35) can be true for all timelike, unit-length vectors $\vec{u}$ if and only if **F** is antisymmetric. [Hints: (i) Show that the most general second-rank **F** can be written as the sum of a symmetric tensor **S** and an antisymmetric tensor **A**, and that the antisymmetric piece contributes nothing to Eq. (1.35). (ii) Let $\vec{B}$ and $\vec{C}$ be any two vectors such that $\vec{B} + \vec{C}$ and $\vec{B} - \vec{C}$ are both timelike; show that $\mathbf{S}(\vec{B}, \vec{C}) = 0$. (iii) Convince yourself (if necessary using tools developed later in this chapter) that this result, together with the 4-dimensionality of spacetime and the large arbitrariness inherent in the choice of $\vec{A}$ and $\vec{B}$, implies **S** vanishes (i.e., it gives zero when any two vectors are inserted into its slots).]

****

1.5 Component Representation of Tensor Algebra[5]

Euclidean 3-space

In the Euclidean 3-space of Newtonian physics, there is a unique set of orthonormal basis vectors $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\} \equiv \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ associated with any Cartesian coordinate system $\{x, y, z\} \equiv \{x^1, x^2, x^3\}$. The basis vector $\mathbf{e}_j$ points along the x_j coordinate direction, which is orthogonal to all the other coordinate directions, and it has unit length, so

$$\mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk} \ . \tag{1.37}$$

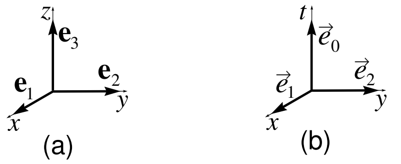


Fig. 1.7: (a) The orthonormal basis vectors $\mathbf{e}_j$ associated with a Euclidean coordinate system in 3-space; (b) the orthonormal basis vectors $\vec{e}_\alpha$ associated with an inertial (Lorentz) reference frame in Minkowski spacetime.

Any vector $\mathbf{A}$ in 3-space can be expanded in terms of this basis,

$$\mathbf{A} = A_j \mathbf{e}_j \ . \tag{1.38}$$

Here and throughout this book, we adopt the Einstein summation convention: repeated indices (in this case j) are to be summed (in this 3-space case over $j = 1, 2, 3$). By virtue of the orthonormality of the basis, the components A_j of $\mathbf{A}$ can be computed as the scalar product

$$A_j = \mathbf{A} \cdot \mathbf{e}_j \ . \tag{1.39}$$

(The proof of this is straightforward: $\mathbf{A} \cdot \mathbf{e}_j = (A_k \mathbf{e}_k) \cdot \mathbf{e}_j = A_k (\mathbf{e}_k \cdot \mathbf{e}_j) = A_k \delta_{kj} = A_j$.)

Any tensor, say the third-rank tensor $\mathbf{T}(\_,\_,\_)$, can be expanded in terms of tensor products of the basis vectors:

$$\mathbf{T} = T_{ijk} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \ . \tag{1.40}$$

The components T_{ijk} of $\mathbf{T}$ can be computed from $\mathbf{T}$ and the basis vectors by the generalization of Eq. (1.39)

$$T_{ijk} = \mathbf{T}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) \ . \tag{1.41}$$

(This equation can be derived using the orthonormality of the basis in the same way as Eq. (1.39) was derived.) As an important example, the components of the metric are $g_{jk} = \mathbf{g}(\mathbf{e}_j, \mathbf{e}_k) = \mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk}$ [where the first equality is the method (1.41) of computing tensor components, the second is the definition (1.13) of the metric, and the third is the orthonormality relation (1.37)]:

$$g_{jk} = \delta_{jk} \quad \text{in any orthonormal basis in 3-space.} \tag{1.42}$$

In Part VI we shall often use bases that are not orthonormal; in such bases, the metric components will not be δ_{jk}.

The components of a tensor product, e.g. $\mathbf{T}(\_,\_,\_) \otimes \mathbf{S}(\_,\_)$, are easily deduced by inserting the basis vectors into the slots [Eq. (1.41)]; they are $\mathbf{T}(\mathbf{e}_i, \mathbf{e}_j \mathbf{e}_k) \otimes \mathbf{S}(\mathbf{e}_l, \mathbf{e}_m) = T_{ijk} S_{lm}$ [cf. Eq. (1.16)]. In words, the components of a tensor product are equal to the ordinary arithmetic product of the components of the individual tensors.

In component notation, the inner product of two vectors and the value of a tensor when vectors are inserted into its slots are given by

$$\mathbf{A} \cdot \mathbf{B} = A_j B_j \ , \quad \mathbf{T}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = T_{ijk} A_i B_j C_k \ , \tag{1.43}$$

---

[5] For a more detailed treatment see, e.g. chapters 2 and 3 of Schutz (1985), or pp. 60–62, 74–89, and 201–203 of Misner, Thorne, and Wheeler (1973).

as one can easily show using previous equations. Finally, the contraction of a tensor [say, the fourth rank tensor $\mathbf{R}(\_,\_,\_,\_)$] on two of its slots [say, the first and third] has components that are easily computed from the tensor's own components:

$$\text{Components of [1\&3contraction of } \mathbf{R}] = R_{ijik} \tag{1.44}$$

Note that R_{ijik} is summed on the i index, so it has only two free indices, j and k, and thus is the component of a second rank tensor, as it must be if it is to represent the contraction of a fourth-rank tensor.

Minkowski spacetime

In Minkowski spacetime, associated with any inertial reference frame there is a Lorentz coordinate system $\{t, x, y, z\} = \{x^0, x^1, x^2, x^3\}$ generated by the frame's rods and clocks, and associated with these coordinates is a set of orthonormal basis vectors $\{\vec{e}_t, \vec{e}_x, \vec{e}_y, \vec{e}_z\} = \{\vec{e}_0, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$; cf. Fig. 1.7. (The reason for putting the indices up on the coordinates but down on the basis vectors will become clear below.) The basis vector $\vec{e}_\alpha$ points along the x^α coordinate direction, which is orthogonal to all the other coordinate directions, and it has squared length -1 for $\alpha = 0$ (vector pointing in timelike direction) and $+1$ for $\alpha = 1, 2, 3$ (spacelike):

$$\vec{e}_\alpha \cdot \vec{e}_\beta = \eta_{\alpha\beta}, \tag{1.45}$$

where $\eta_{\alpha\beta}$ is defined by

$$\eta_{00} = -1\ , \quad \eta_{11} = \eta_{22} = \eta_{33} = 1\ , \quad \eta_{\alpha\beta} = 0 \text{ if } \alpha \neq \beta\ . \tag{1.46}$$

The fact that $\vec{e}_\alpha \cdot \vec{e}_\beta \neq \delta_{\alpha\beta}$ prevents many of the Euclidean-space component-manipulation formulas (1.39)–(1.44) from holding true in Minkowski spacetime. There are two approaches to recovering these formulas. One approach, often used in elementary textbooks [and also used in Goldstein's (1980) *Classical Mechanics* and in the first edition of Jackson's *Classical Electrodynamics*], is to set $x^0 = it$, where $i = \sqrt{-1}$ and correspondingly make the time basis vector be imaginary, so that $\vec{e}_\alpha \cdot \vec{e}_\beta = \delta_{\alpha\beta}$. This approach has severe disadvantages: (i) it hides the true physical geometry of Minkowski spacetime, (ii) it cannot be extended in any reasonable manner to non-orthonormal bases in flat spacetime, and (iii) it cannot be extended in any reasonable manner to the curvilinear coordinates that one must use in general relativity. For this reason, most advanced texts [including the second and third editions of Jackson (1999)] and all general relativity texts take an alternative approach, which we also adopt in this book. This alternative approach requires introducing two different types of components for vectors, and analogously for tensors: *contravariant components* denoted by superscripts, and *covariant components* denoted by subscripts. In Parts I–V of this book we introduce these components only for orthonormal bases; in Part VI we develop a more sophisticated version of them, valid for nonorthonormal bases.

When expanding a vector or tensor in terms of the Minkowski-spacetime basis vectors, one uses its contravariant components:

$$\vec{A} = A^\alpha \vec{e}_\alpha\ , \quad \mathbf{T} = T^{\alpha\beta\gamma} \vec{e}_\alpha \otimes \vec{e}_\beta \otimes \vec{e}_\gamma\ . \tag{1.47}$$

Here and throughout this book, Greek (spacetime) indices are to be summed whenever they are repeated with one up and the other down.

Equations (1.47) can be regarded as definitions of the contravariant components A^α and $T^{\alpha\beta\gamma}$. The covariant components of $\vec{A}$ are defined (when, as in Parts I–V, the basis is orthonormal) by

$$A_\alpha \equiv \eta_{\alpha\beta} A^\beta \ , \quad \text{i.e. } A_0 \equiv -A^0, \quad A_j \equiv +A^j \text{ for } j = 1,2,3 \ . \tag{1.48}$$

Similarly, one can lower any index on a tensor using $\eta_{\alpha\beta}$:

$$\begin{aligned} T^\alpha{}_{\mu\nu} &\equiv \eta_{\mu\beta}\eta_{\nu\gamma}T^{\alpha\beta\gamma} \ , \\ \text{i.e. } T^\alpha{}_{00} = +T^{\alpha 00}, \quad T^\alpha{}_{0j} &= -T^{\alpha 0j}, \quad T^\alpha{}_{j0} = -T^{\alpha j0}, \quad T^\alpha{}_{jk} = +T^{\alpha jk} \ . \end{aligned} \tag{1.49}$$

In words, lowering a temporal index changes the component's sign and lowering a spatial index leaves the component unchanged—and similarly for raising indices.

These definitions give rise to simple formulae for computing a vector's components from the vector itself: By analogy with the Euclidean-space formula $\mathbf{A} \cdot \mathbf{e}_j = A_j$, we compute $\vec{A} \cdot \vec{e}_\alpha = (A^\beta \vec{e}_\beta) \cdot \vec{e}_\alpha = A^\beta \vec{e}_\beta \cdot \vec{e}_\alpha = A^\beta \eta_{\beta\alpha} = A_\alpha$. Thus (and similarly for a tensor)

$$A_\alpha = \vec{A} \cdot \vec{e}_\alpha \ , \quad T_{\alpha\beta\gamma} = \mathbf{T}(\vec{e}_\alpha, \vec{e}_\beta, \vec{e}_\gamma) \ . \tag{1.50}$$

By applying this formula to the metric, and then raising its indices, we obtain for its components in our orthonormal basis

$$g_{\alpha\beta} = \eta_{\alpha\beta} \ , \quad g^\alpha{}_\beta = \delta_{\alpha\beta} \ , \quad g_\alpha{}^\beta = \delta_{\alpha\beta} \ , \quad g^{\alpha\beta} = \eta_{\alpha\beta} \ . \tag{1.51}$$

In other words, the components are nonzero only if the indices are equal, and all nonzero components are $+1$ except $g_{00} = g^{00} = -1$. These metric components enable us to restate the rule (1.48), (1.49) for lowering and raising indices: *Indices are lowered and raised with the components of the metric*

$$A_\alpha = g_{\alpha\beta}A^\beta \ , \quad A^\alpha = g^{\alpha\beta}A_\beta \ , \quad T^\alpha{}_{\mu\nu} \equiv g_{\mu\beta}g_{\nu\gamma}T^{\alpha\beta\gamma} \quad T^{\alpha\beta\gamma} \equiv g^{\beta\mu}g^{\gamma\nu}T^\alpha{}_{\mu\nu} \ . \tag{1.52}$$

These elegant equations have no more content than their predecessors: raising or lowering a spatial index leaves a component unchanged; raising or lowering a temporal index changes the component's sign.

This index notation gives rise to formulas for tensor products, inner products, values of tensors on vectors, and tensor contractions, that are the obvious analogs of those in Euclidean space:

$$[\text{Covariant components of } \mathbf{T}(\_,\_,\_) \otimes \mathbf{S}(\_,\_)] = T^{\alpha\beta\gamma}S^{\delta\epsilon} \ , \tag{1.53}$$

$$\vec{A} \cdot \vec{B} = A^\alpha B_\alpha = A_\alpha B^\alpha \ , \quad \mathbf{T}(\mathbf{A},\mathbf{B},\mathbf{C}) = T_{\alpha\beta\gamma}A^\alpha B^\beta C^\gamma = T^{\alpha\beta\gamma}A_\alpha B_\beta C_\gamma \ , \tag{1.54}$$

$$\begin{aligned} \text{Covariant components of [1\&3contraction of } \mathbf{R}] &= R^\mu{}_{\alpha\mu\beta} \ , \\ \text{Contravariant components of [1\&3contraction of } \mathbf{R}] &= R^{\mu\alpha}{}_\mu{}^\beta \ . \end{aligned} \tag{1.55}$$

Notice the very simple pattern in Eqs. (1.47), (1.50), (1.52)–(1.55), which universally permeates the rules of index gymnastics, a pattern that permits one to reconstruct the rules without any memorization: *Free indices (indices not summed over) must agree in position* *(up versus down) on the two sides of each equation.* In keeping with this pattern, one often regards the two indices in a pair that is summed (one index up and the other down) as "strangling each other" and thereby being destroyed, and one speaks of "lining up the indices" on the two sides of an equation to get them to agree.

Slot-Naming Index Notation

We now pause, in our development of the component version of tensor algebra, for a very important philosophical remark. Consider the rank-2 tensor $\mathbf{F}(\_,\_)$. We can define a new tensor $\mathbf{G}(\_,\_)$ to be the same as $\mathbf{F}$, but with the slots interchanged; i.e., for any two vectors $\vec{A}$ and $\vec{B}$ it is true that $\mathbf{G}(\vec{A},\vec{B}) = \mathbf{F}(\vec{B},\vec{A})$. We need a simple, compact way to indicate that $\mathbf{F}$ and $\mathbf{G}$ are equal except for an interchange of slots. The best way is to give the slots names, say α and β—i.e., to rewrite $\mathbf{F}(\_,\_)$ as $\mathbf{F}(\_{}_\alpha,\_{}_\beta)$ or more conveniently as $F_{\alpha\beta}$; and then to write the relationship between $\mathbf{G}$ and $\mathbf{F}$ as $G_{\alpha\beta} = F_{\beta\alpha}$. *NO!* some readers might object. This notation is indistinguishable from our notation for components on a particular basis. *GOOD!* an astute reader will exclaim. The relation $G_{\alpha\beta} = F_{\beta\alpha}$ in a particular basis is a true statement if and only if "$\mathbf{G} = \mathbf{F}$ with slots interchanged" is true, so why not use the same notation to symbolize both? This, in fact, we shall do. We shall ask our readers to look at any "index equation" such as $G_{\alpha\beta} = F_{\beta\alpha}$ like they would look at an Escher drawing: momentarily think of it as a relationship between components of tensors in a specific basis; then do a quick mind-flip and regard it quite differently, as a relationship between geometric, basis-independent tensors with the indices playing the roles of names of slots. This mind-flip approach to tensor algebra will pay substantial dividends.

As an example of the power of this *slot-naming index notation,* consider the contraction of the first and third slots of a third-rank tensor $\mathbf{T}$. In any basis, where we write $\mathbf{T} = T^{\alpha\beta\gamma}\vec{e}_\alpha \otimes \vec{e}_\beta \otimes \vec{e}_\gamma$, the rule (1.19) for computing the contraction gives 1&3contraction($\mathbf{T}$) $=$ $\vec{e}_\alpha \cdot \vec{e}_\gamma T^{\alpha\beta\gamma}\vec{e}_\beta$, which, since $\vec{e}_\alpha \cdot \vec{e}_\gamma = \eta_{\alpha\gamma}$, gives $T^{\alpha\beta}{}_\alpha\vec{e}_\beta$. This means that 1&3contraction($\mathbf{T}$) has components $T^{\alpha\beta}{}_\alpha$; cf. Eq. (1.55). Correspondingly, in slot-naming index notation we denote 1&3contraction($\mathbf{T}$) by the simple expression $T^{\alpha\beta}{}_\alpha$. We say that the first and third slots are "strangling each other" by the contraction, leaving free only the second slot (named β) and therefore producing a rank-1 tensor (a vector).

By virtue of the "index-lowering" role of the metric, we can also write the contraction as

$$T^{\alpha\beta}{}_\alpha = T^{\alpha\beta\gamma}g_{\alpha\gamma} \ , \tag{1.56}$$

and we can look at this relation from either of two viewpoints: The component viewpoint says that the components of the contraction of $\mathbf{T}$ in any chosen basis are obtained by taking a product of components of $\mathbf{T}$ and of the metric $\mathbf{g}$ and then summing over the appropriate indices. The slot-naming viewpoint says that the contraction of $\mathbf{T}$ can be achieved by taking a tensor product of $\mathbf{T}$ with the metric $\mathbf{g}$ to get $\mathbf{T}\otimes\mathbf{g}(\_,\_,\_,\_,\_)$ (or $T^{\alpha\beta\gamma}g_{\mu\nu}$ in slot-naming index notation), and by then strangling on each other the first and fourth slots [named α in (1.56)], and also strangling on each other the third and fifth slots [named γ in (1.56)].

****

EXERCISES

Exercise 1.4 *Derivation: Component Manipulation Rules*
Derive the component manipulation rules in Eqs. (1.41) and (1.51)–(1.55) of the text. Base your derivations on the definitions which precede those rules in the text. As you proceed, abandon any piece of the exercise when it becomes trivial for you.

Exercise 1.5 *Practice: Numerics of Component Manipulations*
In Minkowski spacetime, in some inertial reference frame, let the components of a vector $\vec{A}$ and a second-rank tensor $\mathbf{T}$ be $A^0 = 1$, $A^1 = 2$, $A^2 = A^3 = 0$; $T^{00} = 3$, $T^{01} = T^{10} = 2$, $T^{11} = -1$, all others vanish. Evaluate $\mathbf{T}(\vec{A}, \vec{A})$ and the components of $\mathbf{T}(\vec{A}, \_)$ and $\vec{A} \otimes \mathbf{T}$.

Exercise 1.6 *Practice: Meaning of Slot Naming Index Notation*
The following expressions and equations are written in slot-naming index notation; convert them to index-free notation: $A_\alpha B^{\beta\gamma}$; $A_\alpha B^{\beta\alpha}$; $S_{\alpha\beta\gamma} = T_{\gamma\beta\alpha}$; $A_\alpha B^\alpha = g_{\mu\nu} A^\mu B^\nu$.

****

1.6 Particle Kinetics in Index Notation and in a Lorentz Frame

As an illustration of the component representation of tensor algebra, let us return to the relativistic, accelerated particle of Fig. 1.5 and, from the frame-independent equations of Sec. 1.4, derive the component description given in elementary textbooks.

We introduce a specific inertial reference frame and associated Lorentz coordinates x^α and basis vectors $\{\vec{e}_\alpha\}$. In this Lorentz frame, the particle's world line $\vec{x}(\tau)$ is represented by its coordinate location $x^\alpha(\tau)$ as a function of its proper time τ. The covariant components of the separation vector $d\vec{x}$ between two neighboring events along the particle's world line are the events' coordinate separations dx^α [Eq. (1.2)—which is why we put the indices up on coordinates]; and correspondingly, the components of the particle's 4-velocity $\vec{u} = d\vec{x}/d\tau$ are

$$u^\alpha = \frac{dx^\alpha}{d\tau} \tag{1.57}$$

(the time derivatives of the particle's spacetime coordinates). Note that Eq. (1.57) implies

$$v^j \equiv \frac{dx^j}{dt} = \frac{dx^j/d\tau}{dt/d\tau} = \frac{u^j}{u^0} \; . \tag{1.58}$$

Here v^j are the components of the *ordinary velocity* as measured in the Lorentz frame. This relation, together with the unit norm of $\vec{u}$, $\vec{u}^2 = g_{\alpha\beta} u^\alpha u^\beta = -(u^0)^2 + \delta_{ij} u^i u^j = -1$, implies that the components of the 4-velocity have the forms familiar from elementary textbooks:

$$u^0 = \gamma \; , \quad u^j = \gamma v^j \; , \quad \text{where} \quad \gamma = \frac{1}{(1 - \delta_{ij} v^i v^j)^{\frac{1}{2}}} \; . \tag{1.59}$$

It is useful to think of v^j as the components of a 3-dimensional vector $\mathbf{v}$, the ordinary velocity, that lives in the 3-dimensional Euclidean space $t = \text{const}$ of the chosen Lorentz

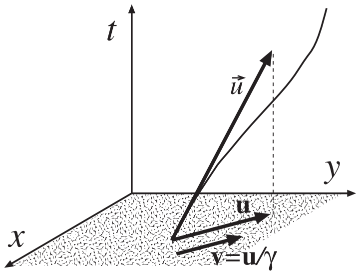


Fig. 1.8: Spacetime diagram in a specific Lorentz frame, showing the frame's 3-space $t = 0$ (stippled region), the 4-velocity $\vec u$ of a particle as it passes through that 3-space (i.e., at time $t = 0$); and two 3-dimensional vectors that lie in the 3-space: the spatial part of the particle's 4-velocity, $\mathbf{u}$, and the particle's ordinary velocity $\mathbf{v}$.

frame. As we shall see below, this 3-space is not well defined until a Lorentz frame has been chosen, and correspondingly, $\mathbf{v}$ relies for its existence on a specific choice of frame. However, once the frame has been chosen, $\mathbf{v}$ can be regarded as a coordinate-independent, basis-independent 3-vector lying in the frame's 3-space $t =$const. Similarly, the spatial part of the 4-velocity $\vec u$ (the part with components u^j in our chosen frame) can be regarded as a 3-vector $\mathbf{u}$ lying in the frame's 3-space; and Eqs. (1.59) become the component versions of the coordinate-independent, basis-independent 3-space relations

$$\mathbf{u} = \gamma \mathbf{v} \ , \quad \gamma = \frac{1}{\sqrt{1-\mathbf{v}^2}} \ . \tag{1.60}$$

Figure 1.8 shows stippled the 3-space $t = 0$ of a specific Lorentz frame, and the 4-velocity $\vec u$ and ordinary velocity $\mathbf{v}$ of a particle as it passes through that 3-space.

The components of the particle's 4-momentum $\vec p$ in our chosen Lorentz frame have special names and special physical significances: The time component of the 4-momentum is the particle's *energy* E as measured in that frame

$$\begin{aligned} E \equiv p^0 &= mu^0 = m\gamma = \frac{m}{\sqrt{1-\mathbf{v}^2}} = \text{ (the particle's energy)} \\ &\simeq m + \frac{1}{2}m\mathbf{v}^2 \quad \text{for } v \equiv |\mathbf{v}| \ll 1 \ . \end{aligned} \tag{1.61}$$

Note that this energy is the sum of the particle's *rest mass-energy* $m = mc^2$ and its *kinetic energy* $m\gamma - m$ (which, for low velocities, reduces to the familiar nonrelativistic kinetic energy $\frac{1}{2}m\mathbf{v}^2$). The spatial components of the 4-momentum, when regarded from the viewpoint of 3-dimensional physics in the 3-space of the chosen Lorentz frame, are the same as the components of the *momentum*, a 3-vector residing in the frame's 3-space:

$$p^j = mu^j = m\gamma v^j = \frac{mv^j}{\sqrt{1-\mathbf{v}^2}} = (j\text{-component of particle's momentum}) \ ; \tag{1.62}$$

or, in basis-independent, 3-dimensional vector notation,

$$\mathbf{p} = m\mathbf{u} = m\gamma\mathbf{v} = \frac{m\mathbf{v}}{\sqrt{1-\mathbf{v}^2}} = \text{ (particle's momentum)} \ . \tag{1.63}$$

For a zero-rest-mass particle, as for one with finite rest mass, we identify the time component of the 4-momentum, in a chosen Lorentz frame, as the particle's energy, and the spatial part as its momentum. Moreover, if—appealing to quantum theory—we regard a zero-rest-mass particle as a quantum associated with a monochromatic wave, then quantum theory tells us that the wave's angular frequency ω as measured in a chosen Lorentz frame will be related to its energy by

$$E \equiv p^0 = \hbar\omega = (\text{particle's energy}) \; ; \tag{1.64}$$

and, since the particle has $\vec{p}^2 = -(p^0)^2 + \mathbf{p}^2 = -m^2 = 0$ (in accord with the lightlike nature of its world line), its momentum as measured in the chosen Lorentz frame will be

$$\mathbf{p} = \hbar\omega\mathbf{n} \; . \tag{1.65}$$

Here $\mathbf{n}$ is the unit 3-vector that points in the direction of travel of the particle, as measured in the chosen Lorentz frame. Eqs. (1.64) and (1.65) are the temporal and spatial components of the geometric, frame-independent relation $\vec{p} = \hbar\vec{k}$ [Eq. (1.29), which is valid for zero-rest-mass particles as well as finite-mass ones].

The introduction of a specific Lorentz frame into spacetime can be said to produce a "3+1" split of every 4-vector into a 3-dimensional vector plus a scalar (a real number). The 3+1 split of a particle's 4-momentum $\vec{p}$ produces its momentum $\mathbf{p}$ plus its energy $E = p^0$; and correspondingly, the 3+1 split of the law of 4-momentum conservation (1.31) produces a law of conservation of momentum plus a law of conservation of energy:

$$\sum_{\bar{A}} \mathbf{p}_{\bar{A}} = \sum_{A} \mathbf{p}_A \; , \quad \sum_{\bar{A}} E_{\bar{A}} = \sum_{A} E_A \; . \tag{1.66}$$

Here the barred quantities are the momenta or energies of the particles entering the interaction region, and the unbarred quantities are the momenta or energies of those leaving; cf. Fig. 1.6.

Because the concept of energy does not even exist until one has chosen a Lorentz frame, and neither does that of momentum, the laws of energy conservation and momentum conservation separately are frame-dependent laws. In this sense they are far less fundamental than their combination, the frame-independent law of 4-momentum conservation.

****

EXERCISES

Exercise 1.7 *Example and Practice: Frame-Independent Expressions for Energy, Momentum, and Velocity*
An observer with 4-velocity $\vec{U}$ measures the properties of a particle with 4-momentum $\vec{p}$.

a. Show that the energy E which the observer measures the particle to have is computable from the frame-independent equation

$$E = -\vec{p}\cdot\vec{U} \; . \tag{1.67}$$

b. Show that the rest mass the observer measures is computable from

$$m^2 = -\vec{p}^2 \ . \tag{1.68}$$

c. Show that the momentum the observer measures has the magnitude

$$|\mathbf{p}| = [(\vec{p}\cdot\vec{U})^2 + \vec{p}\cdot\vec{p}]^{\frac{1}{2}} \ . \tag{1.69}$$

d. Show that the ordinary velocity the observer measures has the magnitude

$$|\mathbf{v}| = \frac{|\mathbf{p}|}{E}, \tag{1.70}$$

where $|\mathbf{p}|$ and E are given by the above frame-independent expressions.

e. Show that the ordinary velocity $\mathbf{v}$, thought of as a 4-vector that happens to lie in the observer's 3-space of constant time is given by

$$\vec{v} = \frac{\vec{p} + (\vec{p}\cdot\vec{U})\vec{U}}{-\vec{p}\cdot\vec{U}} \ . \tag{1.71}$$

Exercise 1.8 *Example: Doppler Shift Derived without Lorentz Transformations*
An atom moving with ordinary velocity $\mathbf{v}$ as measured in some inertial reference frame F emits a photon in a direction $\mathbf{n}$ as measured in F. The photon's energy is later measured, by an observer at rest in F, to be E_F. Let $\vec{U}$ be the emitting atom's 4-velocity and $\vec{p}$ be the photon's 4-momentum. By a computation carried out in frame F, evaluate Eq. (1.67) to obtain the photon energy measured by the emitting atom. Then compute the ratio E_F/E to obtain the standard formula for the photon's Doppler shift in terms of $\mathbf{v}$ and $\mathbf{n}$.

Exercise 1.9 *Problem: Allowed and Forbidden Electron-Photon Reactions*
Show, using spacetime diagrams and/or using frame-independent calculations, that the law of conservation of 4-momentum forbids a photon to be absorbed by an electron, $e + \gamma \rightarrow e$ and also forbids an electron and a positron to annihilate and produce a single photon $e^+ + e^- \rightarrow \gamma$ (in the absence of any other particles to take up some of the 4-momentum); but the annihilation to form two photons, $e^+ + e^- \rightarrow 2\gamma$, is permitted.

****

1.7 Orthogonal and Lorentz Transformations of Bases, and Spacetime Diagrams

Euclidean 3-space

Consider two different Cartesian coordinate systems $\{x, y, z\} \equiv \{x_1, x_2, x_3\}$, and $\{\bar{x}, \bar{y}, \bar{z}\} \equiv \{x_{\bar{1}}, x_{\bar{2}}, x_{\bar{3}}\}$. Denote by $\{\mathbf{e}_i\}$ and $\{\mathbf{e}_{\bar{p}}\}$ the corresponding bases. It must be possible to expand the basis vectors of one basis in terms of those of the other. We shall denote the expansion coefficients by the letter R and shall write

$$\mathbf{e}_i = \mathbf{e}_{\bar{p}} R_{\bar{p}i} \ , \quad \mathbf{e}_{\bar{p}} = \mathbf{e}_i R_{i\bar{p}} \ . \tag{1.72}$$

The quantities $R_{\bar{p}i}$ and $R_{i\bar{p}}$ are *not* the components of a tensor; rather, they are the elements of transformation matrices

$$||R_{\bar{p}i}|| = \left|\left| \begin{array}{ccc} R_{\bar{1}1} & R_{\bar{1}2} & R_{\bar{1}3} \\ R_{\bar{2}1} & R_{\bar{2}2} & R_{\bar{2}3} \\ R_{\bar{3}1} & R_{\bar{3}2} & R_{\bar{3}3} \end{array} \right|\right| \ , \quad ||R_{i\bar{p}}|| = \left|\left| \begin{array}{ccc} R_{1\bar{1}} & R_{1\bar{2}} & R_{1\bar{3}} \\ R_{2\bar{1}} & R_{2\bar{2}} & R_{2\bar{3}} \\ R_{3\bar{1}} & R_{3\bar{2}} & R_{3\bar{3}} \end{array} \right|\right| \ . \tag{1.73}$$

(Here and throughout this book we use double vertical bars to denote matrices.) These two matrices must be the inverse of each other, since one takes us from the barred basis to the unbarred, and the other in the reverse direction, from unbarred to barred:

$$R_{\bar{p}i} R_{i\bar{q}} = \delta_{\bar{p}\bar{q}} \ , \quad R_{i\bar{p}} R_{\bar{p}j} = \delta_{ij} \ . \tag{1.74}$$

The orthonormality requirement for the two bases implies that $\delta_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = (\mathbf{e}_{\bar{p}} R_{\bar{p}i}) \cdot (\mathbf{e}_{\bar{q}} R_{\bar{q}i}) = R_{\bar{p}i} R_{\bar{q}j} (\mathbf{e}_{\bar{p}} \cdot \mathbf{e}_{\bar{q}}) = R_{\bar{p}i} R_{\bar{q}j} \delta_{\bar{p}\bar{q}} = R_{\bar{p}i} R_{\bar{p}j}$. This says that the transpose of $||R_{\bar{p}i}||$ is its inverse—which we have already denoted by $||R_{i\bar{p}}||$; thus,

$$R_{i\bar{p}} = R_{\bar{p}i} \ . \tag{1.75}$$

This property implies that the transformation matrix is orthogonal; i.e., the transformation is a reflection or a rotation [see, e.g., Goldstein (1980)]. Thus (as should be obvious and familiar), the bases associated with any two Euclidean coordinate systems are related by a reflection or rotation. [Note: Eq. (1.75) does *not* say that $||R_{i\bar{p}}||$ is a symmetric matrix; in fact, it typically is not. Rather, (1.75) says that $||R_{i\bar{p}}||$ is the transpose of $||R_{\bar{p}i}||$.]

The fact that a vector $\mathbf{A}$ is a geometric, basis-independent object implies that $\mathbf{A} = A_i \mathbf{e}_i = A_i \mathbf{e}_{\bar{p}} R_{\bar{p}i} = (R_{\bar{p}i} A_i) \mathbf{e}_{\bar{p}} = A_{\bar{p}} \mathbf{e}_{\bar{p}}$; i.e.,

$$A_{\bar{p}} = R_{\bar{p}i} A_i \ , \quad \text{and similarly } A_i = R_{i\bar{p}} A_{\bar{p}} \ ; \tag{1.76}$$

and correspondingly for the components of a tensor

$$T_{\bar{p}\bar{q}\bar{r}} = R_{\bar{p}i} R_{\bar{q}j} R_{\bar{r}k} T_{ijk} \ , \quad T_{ijk} = R_{i\bar{p}} R_{j\bar{q}} R_{k\bar{r}} T_{\bar{p}\bar{q}\bar{r}} \ . \tag{1.77}$$

It is instructive to compare the transformation law (1.76) for the components of a vector with those (1.72) for the bases. To make these laws look natural, we have placed the transformation matrix on the left in the former and on the right in the latter. In Minkowski spacetime, the placement of indices, up or down, will automatically tell us the order.

Minkowski spacetime

Consider two different inertial reference frames in Minkowski spacetime; denote their Lorentz coordinates by $\{x^\alpha\}$ and $\{x^{\bar\mu}\}$ and their bases by $\{\mathbf{e}_\alpha\}$ and $\{\mathbf{e}_{\bar\mu}\}$; and write the transformation from one basis to the other as

$$\vec{e}_\alpha = \vec{e}_{\bar\mu} L^{\bar\mu}{}_\alpha \; , \quad \vec{e}_{\bar\mu} = \vec{e}_\alpha L^\alpha{}_{\bar\mu} \; . \tag{1.78}$$

As in Euclidean 3-space, $L^{\bar\mu}{}_\alpha$ and $L^\alpha{}_{\bar\mu}$ are elements of two different transformation matrices, and since these matrices operate in opposite directions, they must be the inverse of each other:

$$L^{\bar\mu}{}_\alpha L^\alpha{}_{\bar\nu} = \delta^{\bar\mu}{}_{\bar\nu} \; , \quad L^\alpha{}_{\bar\mu} L^{\bar\mu}{}_\beta = \delta^\alpha{}_\beta \; . \tag{1.79}$$

Notice the up/down placement of indices on the elements of the transformation matrices: the first index is always up, and the second is always down. This is just a convenient convention which helps systematize the index shuffling rules in a way that can be easily remembered. Our rules about summing on the same index when up and down, and matching unsummed indices on the two sides of an equation, automatically dictate the matrix to use in each of the transformations (1.78); and similarly for all other equations in this section.

In Euclidean 3-space the orthonormality of the two bases dictated that the transformations must be orthogonal, i.e. must be reflections or rotations. In Minkowski spacetime, orthonormality implies $g_{\alpha\beta} = \vec{e}_\alpha \cdot \vec{e}_\beta = (\vec{e}_{\bar\mu} L^{\bar\mu}{}_\alpha) \cdot (\vec{e}_{\bar\nu} L^{\bar\nu}{}_\beta) = L^{\bar\mu}{}_\alpha L^{\bar\nu}{}_\beta g_{\bar\mu\bar\nu}$; i.e.,

$$g_{\bar\mu\bar\nu} L^{\bar\mu}{}_\alpha L^{\bar\nu}{}_\beta = g_{\alpha\beta} \; , \quad \text{and similarly } g_{\alpha\beta} L^\alpha{}_{\bar\mu} L^\beta{}_{\bar\nu} = g_{\bar\mu\bar\nu} \; . \tag{1.80}$$

Any matrices whose elements satisfy these equations is a *Lorentz transformation.*

¿From the fact that vectors and tensors are geometric, frame-independent objects, one can derive the Minkowski-space analogs of the Euclidean transformation laws for components (1.76), (1.77):

$$A^{\bar\mu} = L^{\bar\mu}{}_\alpha A^\alpha \; , \quad T^{\bar\mu\bar\nu\bar\rho} = L^{\bar\mu}{}_\alpha L^{\bar\nu}{}_\beta L^{\bar\rho}{}_\gamma T^{\alpha\beta\gamma} \; , \qquad \text{and similarly in the opposite direction.} \tag{1.81}$$

Notice that here, as elsewhere, these equations can be constructed by lining up indices in accord with our standard rules.

If (as is conventional) we choose the origins of the two Lorentz coordinate systems to coincide, then the vector $\vec{x}$ extending from the origin to some event $\mathcal{P}$, whose coordinates are x^α and $x^{\bar\alpha}$, has components equal to those coordinates. As a result, the transformation law for the components of $\vec{x}$ becomes the following relationship between the two sets of coordinates:

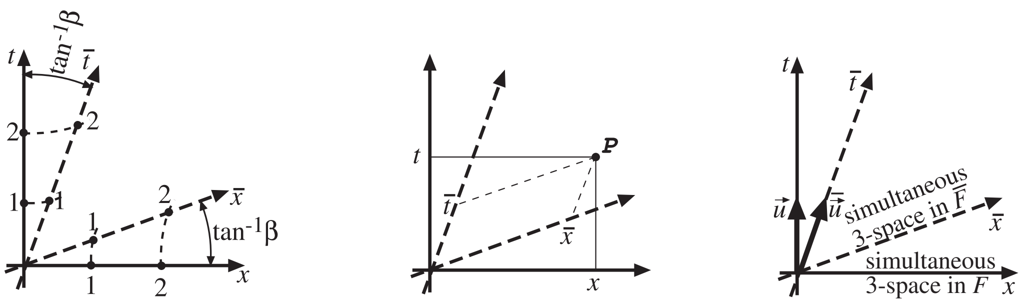
$$x^\alpha = L^\alpha{}_{\bar\mu} x^{\bar\mu} \; , \quad x^{\bar\mu} = L^{\bar\mu}{}_\alpha x^\alpha \; . \tag{1.82}$$

An important specific example of a Lorentz transformation is the following

$$||\, L^\alpha{}_{\bar\mu} \,|| = \left\| \begin{array}{cccc} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right\| \; , \quad ||\, L^{\bar\mu}{}_\alpha \,|| = \left\| \begin{array}{cccc} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right\| \; , \tag{1.83}$$

where β and γ are related by

$$|\beta| < 1 \; , \qquad \gamma \equiv (1 - \beta^2)^{-\frac{1}{2}} \; . \tag{1.84}$$



(a) (b) (c)

Fig. 1.9: Spacetime diagrams illustrating the pure boost (1.85) from one Lorentz reference frame to another.

One can readily verify that these matrices are the inverses of each other and that they satisfy the Lorentz-transformation relation (1.80). These transformation matrices produce the following change of coordinates [Eq. (1.82)]

$$\begin{aligned} t &= \gamma(\bar t + \beta \bar x)\ , \quad & x &= \gamma(\bar x + \beta \bar t)\ , \quad & y &= \bar y\ , \quad & z &= \bar z\ , \\ \bar t &= \gamma(t - \beta x)\ , \quad & \bar x &= \gamma(x - \beta t)\ , \quad & \bar y &= y\ , \quad & \bar z &= z\ . \end{aligned} \tag{1.85}$$

These expressions reveal that any point at rest in the unbarred frame (a point with fixed, time-independent x, y, z) is seen in the barred frame to move along the world line $\bar x = \text{const} - \beta \bar t$, $\bar y = \text{const}$, $\bar z = \text{const}$. In other words, the unbarred frame is seen by observers at rest in the barred frame to move with uniform velocity $\vec v = -\beta \vec e_{\bar x}$, and correspondingly the barred frame is seen by observers at rest in the unbarred frame to move with the opposite uniform velocity $\vec v = +\beta \vec e_x$. This special Lorentz transformation is called a *pure boost* along the x direction.

Figure 1.9 illustrates the pure boost (1.85). Diagram (a) in that figure is a two-dimensional spacetime diagram, with the y- and z-coordinates suppressed, showing the $\bar t$ and $\bar x$ axes of the boosted Lorentz frame $\bar F$ in the t, x Lorentz coordinate system of the unboosted frame F. That the barred axes make angles $\tan^{-1}\beta$ with the unbarred axes, as shown, can be inferred from the Lorentz transformation equation (1.85). Note that invariance of the interval guarantees that the event $\bar x = a$ on the $\bar x$-axis lies at the intersection of that axis with the dashed hyperbola $x^2 - t^2 = a^2$; and similarly, the event $\bar t = a$ on the $\bar t$-axis lies at the intersection of that axis with the hyperbola $t^2 - x^2 = a^2$. As is shown in diagram (b) of the figure, the barred coordinates $\bar t$, $\bar x$ of an event $\mathcal{P}$ can be inferred by projecting from $\mathcal{P}$ onto the $\bar t$- and $\bar x$-axes, with the projection going parallel to the $\bar x$- and $\bar t$- axes respectively. Diagram (c) shows the 4-velocity $\vec u$ of an observer at rest in frame F and that, $\vec{\bar u}$ of an observer in frame $\bar F$. The events which observer F regards as all simultaneous, with time $t = 0$, lie in a 3-space that is orthogonal to $\vec u$ and includes the x-axis. This is the *Euclidean 3-space of reference frame* F and is also sometimes called F's *3-space of simultaneity*. Similarly, the events which observer $\bar F$ regards as all simultaneous, with $\bar t = 0$, live in the 3-space that is orthogonal to $\vec{\bar u}$ and includes the $\bar x$-axis. This is the Euclidean 3-space (3-space of simultaneity) of frame $\bar F$.

Exercise 1.10 uses spacetime diagrams, similar to Fig. 1.9, to deduce a number of important relativistic phenomena, including the contraction of the length of a moving object ("length contraction"), the breakdown of simultaneity as a universally agreed upon concept, and the dilation of the ticking rate of a moving clock ("time dilation"). This exercise is extremely important; every reader who is not already familiar with it should study it.

****

EXERCISES

Exercise 1.10 *Derivation: The Inverse of a Lorentz Boost*
Show that, if the Lorentz coordinates of an inertial frame F are expressed in terms of those of the frame $\bar{F}$ by Eq. (1.85), then the inverse transformation from F to $\bar{F}$ is given by the same equation with the sign of β reversed. Write down the corresponding transformation matrix $L^{\bar{\mu}}{}_{\alpha}$ [analog of Eq. (1.83)].

Exercise 1.11 *Example: Spacetime Diagrams*
Use spacetime diagrams to prove the following:

a. Two events that are simultaneous in one inertial frame are not necessarily simultaneous in another. More specifically, if frame $\bar{F}$ moves with velocity $\vec{v} = \beta\vec{e}_x$ as seen in frame F, where $\beta > 0$, then of two events that are simultaneous in $\bar{F}$ the one farther "back" (with the more negative value of $\bar{x}$) will occur in F before the one farther "forward".

b. Two events that occur at the same spatial location in one inertial frame do not necessarily occur at the same spatial location in another.

c. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are two events with a timelike separation, then there exists an inertial reference frame in which they occur at the same spatial location; and in that frame the time lapse between them is equal to the square root of the negative of their invariant interval, $\Delta t = \Delta\tau \equiv \sqrt{-\Delta s^2}$.

d. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are two events with a spacelike separation, then there exists an inertial reference frame in which they are simultaneous; and in that frame the spatial distance between them is equal to the square root of their invariant interval, $\sqrt{g_{ij}\Delta x^i \Delta x^j} = \Delta s \equiv \sqrt{\Delta s^2}$.

e. If the inertial frame $\bar{F}$ moves with speed β relative to the frame F, then a clock at rest in $\bar{F}$ ticks more slowly as viewed from F than as viewed from $\bar{F}$—more slowly by a factor $\gamma^{-1} = (1-\beta^2)^{\frac{1}{2}}$.

f. If the inertial frame $\bar{F}$ moves with velocity $\vec{v} = \beta\vec{e}_x$ relative to the frame F and the two frames are related by a pure boost, then an object at rest in $\bar{F}$ as studied in F appears shortened by a factor $\gamma^{-1} = (1-\beta^2)^{\frac{1}{2}}$ along the x direction, but its length along the y and z directions is unchanged.

Exercise 1.12 *Example: General Boosts and Rotations*
Show that, if n^j is a 3-dimensional unit vector and β and γ are defined as in Eq. (1.84), then the following is a Lorentz transformation; i.e., it satisfies Eq. (1.80).

$$L^0{}_{\bar 0} = \gamma \ , \quad L^0{}_{\bar j} = L^{\bar j}{}_0 = \beta\gamma n^j \ , \quad L^{\bar j}{}_k = L^{\bar k}{}_j = (\gamma - 1)n^j n^k + \delta^{jk} \ . \tag{1.86}$$

Show, further, that this transformation is a *pure boost along the direction* **n** *with speed* β, and show that the inverse matrix $||L^{\bar\mu}{}_\alpha||$ for this boost is the same as $||L^\alpha{}_{\bar\mu}||$, but with β changed to $-\beta$.

Show that the following is also a Lorentz transformation:

$$|| \, L^\alpha{}_{\bar\mu} \, || = \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & ||R_{i\bar j}|| & \\ 0 & & & \end{array} \right\| \ , \tag{1.87}$$

where $||R_{i\bar j}||$ is a three-dimensional rotation matrix for Euclidean 3-space. Show, further, that this Lorentz transformation rotates the inertial frame's spatial axes (its latticework of measuring rods), while leaving the frame's velocity unchanged; i.e., the new frame is at rest with respect to the old.

The general Lorentz transformation can be expressed as a sequence of pure boosts, pure rotations, and pure inversions (in which one or more of the coordinate axes are reflected through the origin, so $x^\alpha = -x^{\bar\alpha}$).

****

1.8 Time Travel

Time dilation is one facet of a more general phenomenon: Time, as measured by ideal clocks, is a "personal thing," different for different observers who move through spacetime on different world lines. This is well illustrated by the infamous "twins paradox," in which one twin, Methuselah, remains forever at rest in an inertial frame and the other, Florence, makes a spacecraft journey at high speed and then returns to rest beside Methuselah.

The twins' world lines are depicted in Fig. 1.10(a), a spacetime diagram whose axes are those of Methuselah's inertial frame. The time measured by an ideal clock that Methuselah carries is the coordinate time t of his inertial frame; and its total time lapse, from Florence's departure to her return, is $t_{\rm return} - t_{\rm departure} \equiv T_{\rm Methuselah}$. By contrast, the time measured by an ideal clock that Florence carries is the proper time τ, i.e. the square root of the invariant interval (1.9), along her world line; and thus her total time lapse from departure to return is

$$T_{\rm Florence} = \int d\tau = \int \sqrt{dt^2 - \delta_{ij}dx^i dx^j} = \int_0^{T_{\rm Methuselah}} \sqrt{1 - v^2}dt \ . \tag{1.88}$$

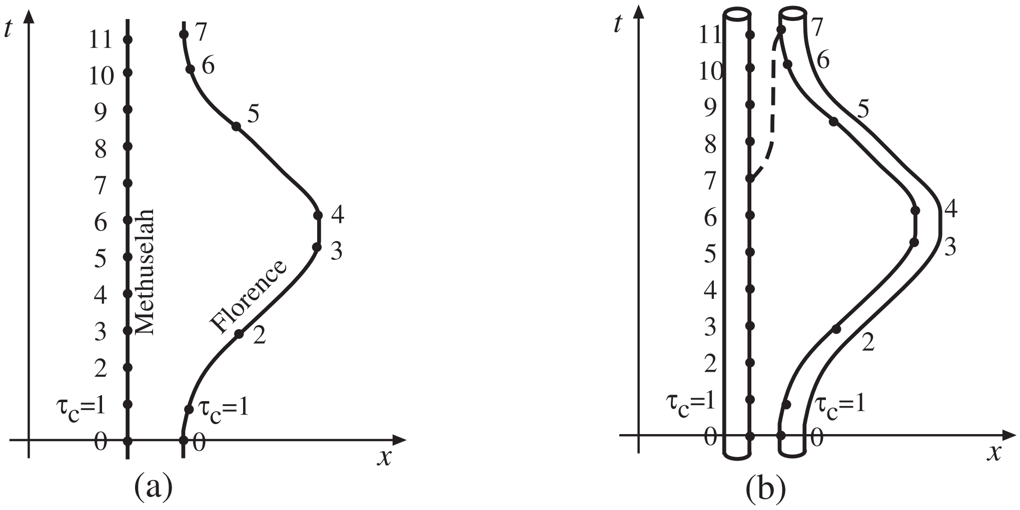


Fig. 1.10: (a) Spacetime diagram depicting the twins paradox. Marked along the two world lines are intervals of proper time as measured by the two twins. (b) Spacetime diagram depicting the motions of the two mouths of a wormhole. Marked along the mouths' world tubes are intervals of proper time τ_c as measured by the single clock that sits on the common mouths.

Here $(t,\ x^i)$ are the time and space coordinates of Methuselah's inertial frame, and v is Florence's ordinary speed, $v = \sqrt{\delta_{ij}(dx^i/dt)(dx^j/dt)}$, relative to Methuselah's frame. Obviously, Eq. (1.88) predicts that $T_{\rm Florence}$ is less than $T_{\rm Methuselah}$. In fact (cf. Exercise 1.13), even if Florence's acceleration is kept no larger than one Earth gravity throughout her trip, and her trip lasts only $T_{\rm Florence} =$ (a few tens of years), $T_{\rm Methuselah}$ can be hundreds or thousands or millions or billions of years.

Does this mean that Methuselah actually "experiences" a far longer time lapse, and actually ages far more than Florence? Yes. The time experienced by humans and the aging of the human body are governed by chemical processes, which in turn are governed by the natural oscillation rates of molecules, which in turn are governed to high accuracy by ideal time (or, equivalently, proper time τ).

In effect, then, Florence's spacecraft has functioned as a time machine to carry her far into Methuselah's future, with only a modest lapse of her own proper time (ideal time; experiential time; aging time).

Is it also possible, at least in principle, for Florence to construct a time machine that carries her into Methuselah's past—and also her own past? At first sight, the answer would seem to be Yes. Figure 1.10(b) shows one possible method, using a *wormhole.* (Another method uses *cosmic strings.*[6])

Wormholes are hypothetical "handles" in the topology of space. A simple model of a wormhole can be obtained by taking a flat 3-dimensional space, removing from it the interiors of two identical spheres, and identifying the spheres' surfaces so that if one enters the surface of one of the spheres, one immediately finds oneself exiting through the surface of the other. When this is done, there is a bit of strongly localized spatial curvature at

---

[6] Gott (1991)

the spheres' common surface, so to analyze such a wormhole properly, one must use general relativity rather than special relativity. In particular, it is the laws of general relativity, combined with the laws of quantum field theory, that tell one how to construct such a wormhole and what kinds of materials (quantum fields) are required to "hold it open" so things can pass through it. Unfortunately, despite considerable effort, theoretical physicists have not yet deduced definitively whether those laws permit such wormholes to exist.[7] On the other hand, assuming such wormholes *can* exist, the following special relativistic analysis shows how one might be used to construct a machine for backward time travel.[8]

The two identified spherical surfaces are called the wormhole's mouths. Ask Methuselah to keep one mouth with himself, forever at rest in his inertial frame, and ask Florence to take the other mouth with herself on her high-speed journey. The two mouths' *world tubes* (analogs of world lines for a 3-dimensional object) then have the forms shown in Fig. 1.10(b). Suppose that a single ideal clock sits on the wormhole's identified mouths, so that from the external Universe one sees it both on Methuselah's wormhole mouth and on Florence's. As seen on Methuselah's mouth, the clock measures his proper time, which is equal to the coordinate time t [see tick marks along the left world tube in Fig. 1.10(b)]. As seen on Florence's mouth, the clock measures her proper time, Eq. (1.88) [see tick marks along the right world tube in Fig. 1.10(b)]. The result should be obvious, if surprising: When Florence returns to rest beside Methuselah, the wormhole has become a time machine. If she travels through the wormhole when the clock reads $\tau_c = 7$, she goes backward in time as seen in Methuselah's (or anyone else's) inertial frame; and then, in fact, traveling along the everywhere timelike, dotted world line, she is able to meet her younger self before she entered the wormhole.

This scenario is profoundly disturbing to most physicists because of the dangers of science-fiction-type paradoxes (e.g., the older Florence might kill her younger self, thereby preventing herself from making the trip through the wormhole and killing herself). Fortunately perhaps, it now seems seems moderately likely (though not certain) that vacuum fluctuations of quantum fields will destroy the wormhole at the moment when its mouths' motion first makes backward time travel possible; and it also seems likely that this mechanism will *always* prevent the construction of backward-travel time machines, no matter what tools one uses for their construction.[9]

****

EXERCISES

Exercise 1.13 *Example: Twins Paradox*
The 4-acceleration of a particle or other object is defined by $\vec{a} \equiv d\vec{u}/d\tau$, where $\vec{u}$ is its 4-velocity and τ is proper time along its world line. Show that, if an observer carries an accelerometer, the magnitude of the acceleration a measured by the accelerometer will always be equal to the magnitude of the observer's 4-acceleration, $a = |\vec{a}| \equiv \sqrt{\vec{a} \cdot \vec{a}}$.

---

[7] See, e.g., Morris and Thorne (1987), and Thorne (1993).

[8] Morris, Thorne, and Yurtsever (1988).

[9] Kim and Thorne (1991), Hawking (1992), Thorne (1993).

In the twins paradox of Fig. 1.10(b), suppose that Florence begins at rest beside Methuselah, then accelerates in Methuselah's x-direction with an acceleration a equal to one Earth gravity, "1g", for a time $T_{\rm Florence}/4$ as measured by her, then accelerates in the $-x$-direction at 1g for a time $T_{\rm Florence}/2$ thereby reversing her motion, and then accelerates in the $+x$-direction at 1g for a time $T_{\rm Florence}/4$ thereby returning to rest beside Methuselah. (This is the type of motion shown in the figure.) What is the resulting relationship between the total time lapse as measured by the two twins? Show that, if $T_{\rm Florence}$ is several tens of years, then $T_{\rm Methuselah}$ can be hundreds or thousands or millions or even billions of years.

Exercise 1.14 *Challenge: Around the World on TWA*
In a long-ago era when TWA flew around the world, J. C. Hafele and R. E. Keating carried out a real live twins' paradox experiment: They synchronized two atomic clocks, and then flew one around the world eastward on TWA, and on a separate trip, around the world westward, while the other clock remained at home at the Naval Research Laboratory near Washington D. C. When the clocks were compared after each trip, they were found to have aged differently. Compute the difference in aging, and compare your result with the experimental data (Hafele and Keating, 1972). Note: The rotation of the Earth is important, but general relativistic effects (notably the gravitational redshift) are less so—though not entirely negligible. Why?

****

1.9 Directional Derivatives, Gradients, Levi-Civita Tensor, Cross Product and Curl

Let us return to the formalism of differential geometry. We shall use the vector notation $\vec{A}$ of Minkowski spacetime, but our discussion will be valid simultaneously for spacetime and for Euclidean 3-space.

Consider a tensor field $\mathbf{T}(\mathcal{P})$ in spacetime or 3-space and a vector $\vec{A}$. We define the *directional derivative* of $\mathbf{T}$ along $\vec{A}$ by the obvious limiting procedure

$$\nabla_{\vec{A}}\mathbf{T} \equiv \lim_{\epsilon\to 0}\frac{1}{\epsilon}[\mathbf{T}(\vec{x}_{\mathcal{P}} + \epsilon\vec{A}) - \mathbf{T}(\vec{x}_{\mathcal{P}})] \tag{1.89}$$

and similarly for the directional derivative of a vector field $\vec{B}(\mathcal{P})$ and a scalar field $\psi(\mathcal{P})$. In this definition we have denoted points, e.g. $\mathcal{P}$, by the vector $\vec{x}_{\mathcal{P}}$ that reaches from some arbitrary origin to the point.

It should not be hard to convince oneself that the directional derivative of any tensor field $\mathbf{T}$ is linear in the vector $\vec{A}$ along which one differentiates. Correspondingly, if $\mathbf{T}$ has rank n (n slots), then there is another tensor field, denoted $\vec{\nabla}\mathbf{T}$, with rank $n+1$, such that

$$\nabla_{\vec{A}}\mathbf{T} = \vec{\nabla}\mathbf{T}(\_\,,\,\_\,,\,\_\,,\,\vec{A})\;. \tag{1.90}$$

Here on the right side the first n slots (3 in the case shown) are left empty, and $\vec{A}$ is put into the last slot (the "differentiation slot"). The quantity $\vec{\nabla}\mathbf{T}$ is called the *gradient* of $\mathbf{T}$. In slot-naming index notation, it is conventional to denote this gradient by $T_{\alpha\beta\gamma;\delta}$, where in general the number of indices preceding the semicolon is the rank of $\mathbf{T}$. Using this notation, the directional derivative of $\mathbf{T}$ along $\vec{A}$ reads [cf. Eq. (1.90)] $T_{\alpha\beta\gamma;\delta}A^\delta$.

It is not hard to show that in any orthonormal (i.e., Cartesian or Lorentz) coordinate system, the components of the gradient are nothing but the partial derivatives of the components of the original tensor,

$$T_{\alpha\beta\gamma;\delta} = \frac{\partial T_{\alpha\beta\gamma}}{\partial x^\delta} \equiv T_{\alpha\beta\gamma,\delta} \; . \tag{1.91}$$

(Here and henceforth all indices that follow a subscript comma represent partial derivatives, e.g., $S_{\alpha,\mu\nu} \equiv \partial^2 S_\alpha/\partial x^\mu \partial x^\nu$.) In a non-Cartesian and non-Lorentz basis, the components of the gradient typically are *not* obtained by simple partial differentiation [i.e. Eq. (1.91) fails] because of twisting and turning and expansion and contraction of the basis vectors as we go from one location to another. In Part III we shall learn how to deal with this by using objects called *connection coefficients*. Until then, however, subscript semicolons and subscript commas can be used interchangeably.

Because the gradient and the directional derivatives are defined by the same standard limiting process as one uses when defining elementary derivatives, they obey the standard Leibniz rule for differentiating products, e.g.

$$\begin{aligned} \nabla_{\vec{A}}(\mathbf{S}\otimes\mathbf{T}) &= (\nabla_{\vec{A}}\mathbf{S})\otimes\mathbf{T} + \mathbf{S}\otimes\nabla_{\vec{A}}\mathbf{T} \; , \\ \text{i.e., } (S^{\alpha\beta}T^{\gamma\delta\epsilon})_{;\mu}A^\mu &= (S^{\alpha\beta}{}_{;\mu}A^\mu)T^{\gamma\delta\epsilon} + S^{\alpha\beta}(T^{\gamma\delta\epsilon}{}_{;\mu}A^\mu) \; ; \end{aligned} \tag{1.92}$$

and

$$\nabla_{\vec{A}}(f\mathbf{T}) = (\nabla_{\vec{A}}f)\mathbf{T} + f\nabla_{\vec{A}}\mathbf{T} \; , \quad \text{i.e., } (fT^{\alpha\beta\gamma})_{;\mu}A^\mu = (f_{,\mu}A^\mu)T^{\alpha\beta\gamma} + fT^{\alpha\beta\gamma}{}_{;\mu}A^\mu \; . \tag{1.93}$$

In an orthonormal basis these relations should be obvious: They follow from the Leibniz rule for partial derivatives.

Because the components $g_{\alpha\beta}$ of the metric tensor are constant in any Lorentz or Cartesian coordinate system, Eq. (1.91) (which is valid in such coordinates) guarantees that $g_{\alpha\beta;\gamma} = 0$; i.e., the metric has vanishing gradient:

$$\vec{\nabla}\mathbf{g} = 0 \; , \quad \text{i.e., } \; g_{\alpha\beta;\mu} = 0 \; . \tag{1.94}$$

¿From the gradient of any vector or tensor we can construct several other important derivatives by contracting on indices: (*i*) Since the gradient $\vec{\nabla}\vec{A}$ of a vector field $\vec{A}$ has two slots, $\vec{\nabla}\vec{A}(\_\;,\_)$, we can strangle (contract) its slots on each other to obtain a scalar field. That scalar field is the *divergence* of $\vec{A}$ and is denoted

$$\vec{\nabla}\cdot\vec{A} \equiv (\text{contraction of } \vec{\nabla}\vec{A}) = A^\alpha{}_{;\alpha} \; . \tag{1.95}$$

(*ii*) Similarly, if $\mathbf{T}$ is a tensor field of rank three, then $T^{\alpha\beta\gamma}{}_{;\gamma}$ is its divergence on its third slot, and $T^{\alpha\beta\gamma}{}_{;\beta}$ is its divergence on its second slot. (*iii*) By taking the double gradient and then contracting on the two gradient slots we obtain, from any tensor field T, a new tensor field with the same rank,

$$\nabla^2 \mathsf{T} \equiv (\vec{\nabla} \cdot \vec{\nabla}) \mathsf{T} \ , \qquad \text{or, in index notation, } T_{\alpha\beta\gamma;\mu}{}^{;\mu} \ . \tag{1.96}$$

In any Euclidean space ∇^2 is called the *Laplacian*; in spacetime it is called the *d'Alembertian*.

The metric tensor is a fundamental property of the space in which it lives; it embodies the inner product and thence the space's notion of distance or interval and thence the space's geometry. In addition to the metric, there is one (and only one) other fundamental tensor that embodies a piece of the space's geometry: the *Levi-Civita tensor* $\boldsymbol{\epsilon}$.

The Levi-Civita tensor has a number of slots equal to the dimensionality N of the space in which it lives, 4 slots in 4-dimensional spacetime and 3 slots in 3-dimensional Euclidean space; and $\boldsymbol{\epsilon}$ is antisymmetric in each and every pair of its slots. These properties turn out to determine $\boldsymbol{\epsilon}$ uniquely up to a multiplicative constant. That constant is fixed by a compatibility relation between $\boldsymbol{\epsilon}$ and the metric g: If $\{\vec{e}_\alpha\}$ is an orthonormal basis (orthonormality being defined with the aid of the metric), and if this basis is right-handed (a new property, not determined by the metric), then

$$\epsilon(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_N) = +1 \ \text{ in a space of } N \text{ dimensions;} \quad \epsilon(\vec{e}_0, \vec{e}_1, \vec{e}_2, \vec{e}_3) = +1 \ \text{ in spacetime.} \tag{1.97}$$

The concept of right handedness should be familiar in Euclidean 2-space or 3-space. In spacetime, the basis is right handed if $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ is right handed and $\vec{e}_0$ points to the future. Equation (1.97) and the antisymmetry of $\boldsymbol{\epsilon}$ imply that in an orthonormal basis, the only nonzero covariant components of $\boldsymbol{\epsilon}$ are

$$\begin{aligned}
\epsilon_{12\ldots N} &= +1 \ , \\
\epsilon_{\alpha\beta\ldots\nu} &= +1 \ \text{ if } \alpha,\ \beta,\ \ldots,\ \nu \text{ is an even permutation of } 1,\ 2,\ \ldots,\ N \\
&= -1 \ \text{ if } \alpha,\ \beta,\ \ldots,\ \nu \text{ is an odd permutation of } 1,\ 2,\ \ldots,\ N \\
&= 0 \ \text{ if } \alpha,\ \beta,\ \ldots,\ \nu \text{ are not all different;}
\end{aligned} \tag{1.98}$$

(In spacetime the indices run from 0 to 3 rather than 1 to $N = 4$.) One can show that these components in one right-handed orthonormal frame imply these same components in all other right-handed orthonormal frames by virtue of the fact that the orthogonal (3-space) and Lorentz (spacetime) transformation matrices have unit determinant.

In 3-dimensional Euclidean space, the Levi-Civita tensor is used to define the cross product:

$$\mathbf{A} \times \mathbf{B} \equiv \boldsymbol{\epsilon}(\_, \mathbf{A}, \mathbf{B}) \quad \text{i.e., in slot-naming index notation, } \epsilon_{ijk} A_j B_k \ ; \tag{1.99}$$

$$\boldsymbol{\nabla} \times \mathbf{A} \equiv (\text{the vector field whose slot-naming index form is } \epsilon_{ijk} A_{k;j}) \ . \tag{1.100}$$

[Equation (1.100) is an example of an expression that is complicated if written in index-free notation; it says that $\boldsymbol{\nabla} \times \mathbf{A}$ is the double contraction of the rank-5 tensor $\boldsymbol{\epsilon} \otimes \nabla \mathbf{A}$ on its second and fifth slots, and on its third and fourth slots.]

Although Eqs. (1.99) and (1.100) look like complicated ways to deal with concepts that most readers regard as familiar and elementary, they have great power. The power comes from the following property of the Levi-Civita tensor [readily derivable from its components (1.98)]:

$$\epsilon_{ijm}\epsilon_{klm} = \delta^{ij}_{kl} \equiv \delta^i_k \delta^j_l - \delta^i_l \delta^j_k \; . \qquad (1.101)$$

Examine the 4-index delta function δ^{ij}_{kl} carefully; it says that either the indices above and below each other must be the same ($i = k$ and $j = l$) with a + sign, or the diagonally related indices must be the same ($i = l$ and $j = k$) with a − sign. With the aid of Eq. (1.101) and the index-notation expressions for the cross product and curl, one can quickly and easily derive a wide variety of useful vector identities; see the very important Exercise 1.15.

****

EXERCISES

Exercise 1.15 *Example and Practice: Vectorial Identities for the Cross Product and Curl*
Here is an example of how to use index notation to derive a vector identity for the double cross product $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})$: In index notation this quantity is $\epsilon_{ijk}A_j(\epsilon_{klm}B_lC_m)$. By permuting the indices on the second ϵ and then invoking Eq. (1.101), we can write this as $\epsilon_{ijk}\epsilon_{lmk}A_jB_lC_m = \delta^{lm}_{ij}A_jB_lC_m$. By then invoking the meaning (1.101) of the 4-index delta function, we bring this into the form $A_jB_iC_j - A_jB_jC_i$, which is the index-notation form of $(\mathbf{A}\cdot\mathbf{C})\mathbf{B}-(\mathbf{A}\cdot\mathbf{B})\mathbf{C}$. Thus, it must be that $\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = (\mathbf{A}\cdot\mathbf{C})\mathbf{B} - (\mathbf{A}\cdot\mathbf{B})\mathbf{C}$.
Use similar techniques to evaluate the following quantities:

a. $\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A})$

b. $(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D})$

c. $(\mathbf{A}\times\mathbf{B})\times(\mathbf{C}\times\mathbf{D})$

****

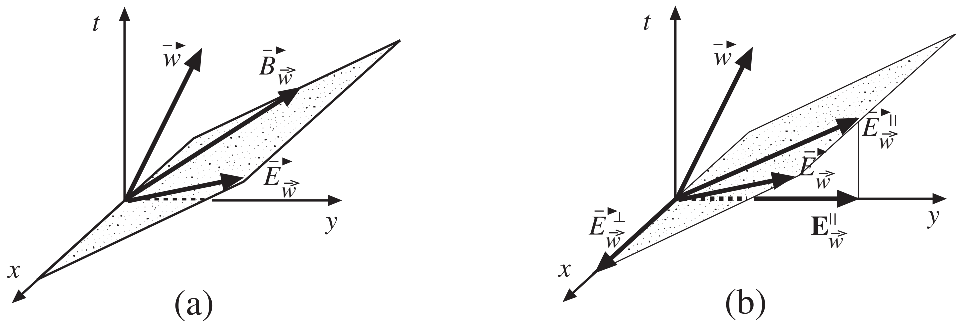
1.10 Nature of Electric and Magnetic Fields; Maxwell's Equations

Now that we have introduced the gradient and the Levi-Civita tensor, we are prepared to study the relationship of the relativistic version of electrodynamics to the nonrelativistic ("Newtonian") version.

Consider a particle with charge q, rest mass m and 4-velocity $\vec{u}$ interacting with an electromagnetic field $\mathsf{F}(\_\,,\_)$. In index notation, the electromagnetic 4-force acting on the particle [Eq. (1.36)] is

$$dp^\alpha/d\tau = qF^{\alpha\beta}u_\beta \; . \qquad (1.102)$$

Let us examine this 4-force in some arbitrary inertial reference frame in which the components of the particle's ordinary velocity are v^j and of its 4-velocity, $u^0 = \gamma$, $u^j = \gamma v^j$ [Eq. (1.59)].



(a) (b)

Fig. 1.11: (a) The electric and magnetic fields measured by an observer with 4-velocity $\vec{w}$, shown as 4-vectors $\vec{E}_{\vec{w}}$ and $\vec{B}_{\vec{w}}$ that lie in the observer's 3-surface of simultaneity (stippled 3-surface orthogonal to $\vec{w}$). (b) Resolution of $\vec{E}_{\vec{w}}$ into pieces parallel and perpendicular to the motion of the observer $\vec{w}$.

Anticipating the connection with the nonrelativistic viewpoint, we introduce the following notation for the contravariant components of the antisymmetric electromagnetic field tensor:

$$F^{0j} = -F^{j0} = E_j \ , \quad F^{ij} = \epsilon_{ijk} B_k \ . \tag{1.103}$$

Inserting these components of F and $\vec{u}$ into Eq. (1.102) and using the relationship $dt/d\tau = u^0 = \gamma$ between t and τ derivatives, we obtain for the components of the 4-force $dp_j/d\tau = \gamma dp_j/dt = q\gamma(E_j + \epsilon_{ijk} v_j B_k)$ and $dp^0/d\tau = \gamma dp^0/dt = \gamma E_j v_j$. Dividing by γ, converting into 3-space index notation, and denoting the particle's energy by $E = p^0$ (not to be confused with the electric field), we bring these into the familiar Lorentz-force form

$$d\mathbf{p}/dt = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \ , \quad dE/dt = \mathbf{v} \cdot \mathbf{E} \ . \tag{1.104}$$

Evidently $\mathbf{E}$ is the electric field and $\mathbf{B}$ the magnetic field as measured in our chosen Lorentz frame.

This may be familiar from standard electrodynamics textbooks, e.g. Jackson (1999). Not so familiar, but quite important, is the following geometric interpretation of the electric and magnetic fields: $\mathbf{E}$ and $\mathbf{B}$ are spatial vectors as measured in the chosen inertial frame. We can also regard these quantities as 4-vectors that lie in the 3-surface of simultaneity $t = \text{const}$ of the chosen frame, i.e. that are orthogonal to the 4-velocity (denote it $\vec{w}$) of the frame's observers (cf. Fig. 1.9). We shall denote this 4-vector version of $\mathbf{E}$ and $\mathbf{B}$ by $\vec{E}_{\vec{w}}$ and $\vec{B}_{\vec{w}}$, where the subscript $\vec{w}$ identifies the 4-velocity of the observers who measure these fields. These fields are depicted in Fig. 1.11

In the rest frame of the observer $\vec{w}$, the components of $\vec{E}_{\vec{w}}$ are $E^0_{\vec{w}} = 0$, $E^j_{\vec{w}} = E_j$ [the E_j appearing in Eqs. (1.103)], and similarly for $\vec{B}_{\vec{w}}$; and the components of $\vec{w}$ are $w^0 = 1$, $w^j = 0$. Therefore, in this frame Eqs. (1.103) can be rewritten as

$$E^{\alpha}_{\vec{w}} = F^{\alpha\beta} w_\beta \ , \quad B^{\beta}_{\vec{w}} = \epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta} w_\alpha \ . \tag{1.105}$$

(To verify this, insert the above components of F and $\vec{w}$ into this equation and, after some algebra, recover Eqs. (1.103) along with $E^0_{\vec{w}} = B^0_{\vec{w}} = 0$.) Equations (1.105) say that in one special reference frame, that of the observer $\vec{w}$, the components of the 4-vectors on the left and on the right are equal. This implies that in every Lorentz frame the components of these 4-vectors will be equal; i.e., it implies that Eqs. (1.105) are true when one regards them as geometric, frame-independent equations written in slot-naming index notation. *These equations enable one to compute the electric and magnetic fields measured by an observer (viewed as 4-vectors in the observer's 3-surface of simultaneity) from the observer's 4-velocity and the electromagnetic field tensor.*

Equations (1.105) embody explicitly the following important fact: The electromagnetic field tensor $\mathbf{F}$ is a geometric, frame-independent quantity. By contrast, the electric and magnetic fields $\vec{E}_{\vec{w}}$ and $\vec{B}_{\vec{w}}$ individually depend for their existence on a specific choice of observer (with 4-velocity $\vec{w}$), i.e., a specific choice of inertial reference frame, i.e., a specific choice of the split of spacetime into a 3-space (the 3-surface of simultaneity orthogonal to the observer's 4-velocity $\vec{w}$) and corresponding time (the Lorentz time of the observer's reference frame). *Only after making such an observer-dependent "3+1 split" of spacetime into space plus time do the electric field and the magnetic field come into existence as separate entities.* Different observers with different 4-velocities $\vec{w}$ make this spacetime split in different ways, thereby resolving the frame-independent $\mathbf{F}$ into different electric and magnetic fields $\vec{E}_{\vec{w}}$ and $\vec{B}_{\vec{w}}$.

By the same procedure as we used to derive Eqs. (1.105), one can derive the inverse relationship, the following expression for the electromagnetic field tensor in terms of the (4-vector) electric and magnetic fields measured by some observer:

$$F^{\alpha\beta} = w^\alpha E^\beta_{\vec{w}} - E^\alpha_{\vec{w}} w^\beta + \epsilon^{\alpha\beta}{}_{\gamma\delta} w^\gamma B^\delta_{\vec{w}} \; . \tag{1.106}$$

Maxwell's equations in geometric, frame-independent form are

$$\begin{aligned} F^{\alpha\beta}{}_{;\beta} &= \begin{cases} 4\pi J^\alpha \text{ in Gaussian units} \\ J^\alpha/\epsilon_o = \mu_o J^\alpha \text{ in SI units} \; . \end{cases} \\ \epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta;\beta} &= 0 \; . \end{aligned} \tag{1.107}$$

(Since we are setting the speed of light to unity, $\epsilon_o = 1/\mu_o$.) Here $\vec{J}$ is the charge-current 4-vector, which in any inertial frame has components

$$J^0 = \rho_e = (\text{charge density}) \; , \quad J^i = j_i = (\text{current density}). \tag{1.108}$$

Exercise 1.18 describes how to think about this charge density and current density as geometric objects determined by the observer's 4-velocity or 3+1 split of spacetime into space plus time. Exercise 1.19 shows how the frame-independent Maxwell equations (1.107) reduce to the more familiar ones in terms of **E** and **B**.

****

EXERCISES

Exercise 1.16 *Derivation and Practice: Reconstruction of* $\mathbf{F}$
Derive Eq. (1.106) by the same method as was used to derive (1.105).

Exercise 1.17 *Challenge: Relationship Between Fields Measured by Different Observers*
In standard electrodynamics textbooks, e.g. Jackson (1999), Lorentz transformations are used to derive the following relationship between the electric and magnetic fields measured by two observers who move relative to each other:

$$\begin{aligned} \mathbf{E}'_{||} &= \mathbf{E}_{||}\ , \quad \mathbf{E}'_{\perp} = \gamma(\mathbf{E}_{\perp} + \mathbf{v}\times\mathbf{B}_{\perp})\ , \\ \mathbf{B}'_{||} &= \mathbf{B}_{||}\ , \quad \mathbf{B}'_{\perp} = \gamma(\mathbf{B}_{\perp} - \mathbf{v}\times\mathbf{E}_{\perp})\ . \end{aligned} \tag{1.109}$$

Here $\mathbf{v}$ is the ordinary velocity of the primed frame as measured in the unprimed frame, the primed fields are measured in the primed frame and unprimed fields in the unprimed frame, $||$ means component parallel to $\mathbf{v}$, and $\perp$ means component perpendicular to $\mathbf{v}$.

Derive Eq. (1.109) from the geometric, frame-independent expression (1.106), without performing any Lorentz transformations. [Hint: Perform your calculation in the primed frame and let $\vec{w}$ be the 4-velocity of the unprimed frame. There will be some trickiness about the meanings of $\mathbf{E}_{||}$ and $\mathbf{B}_{||}$.]

Exercise 1.18 *Problem: 3+1 Split of Charge-Current 4-Vector*
Just as the electric and magnetic fields measured by some observer can be regarded as 4-vectors $\vec{E}_{\vec{w}}$ and $\vec{B}_{\vec{w}}$ that live in the observer's 3-space of simultaneity, so also the charge density and current density that the observer measures can be regarded as a scalar $\rho_{\vec{w}}$ and 4-vector $\vec{j}_{\vec{w}}$ that live in the 3-space of simultaneity. Derive geometric, frame-independent equations for $\rho_{\vec{w}}$ and $\vec{j}_{\vec{w}}$ in terms of $\vec{J}$ and the observer's 4-velocity $\vec{w}$, and a geometric expression for $\vec{J}$ in terms of $\rho_{\vec{w}}$, $\vec{j}_{\vec{w}}$, and $\vec{w}$.

Exercise 1.19 *Problem: Frame-Dependent Version of Maxwell's Equations*
¿From the geometric version of Maxwell's equations (1.107), derive the elementary, frame-dependent version

$$\nabla\cdot\mathbf{E} = \left\{\begin{array}{l} 4\pi\rho_e \text{ in Gaussian units} \\ \rho_e/\epsilon_o \text{ in SI units,} \end{array}\right. \qquad \nabla\times\mathbf{B} - \frac{\partial\mathbf{E}}{\partial t} = \left\{\begin{array}{l} 4\pi\mathbf{j} \text{ in Gaussian units} \\ \mu_o\mathbf{j} \text{ in SI units,} \end{array}\right.$$

$$\nabla\cdot\mathbf{B} = 0\ , \qquad \nabla\times\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t} = 0\ .$$

****

1.11 Volumes, Integration, and the Gauss and Stokes Theorems

The Levi-Civita tensor is the foundation for computing volumes and performing volume integrals in any number of dimensions. In Cartesian coordinates of 2-dimensional Euclidean space, the area (i.e. 2-dimensional volume) of a parallelogram whose sides are $\mathbf{A}$ and $\mathbf{B}$ is

$$\text{2-Volume} = \epsilon_{ab}A_aB_b = A_1B_2 - A_2B_1 = \det\left|\left|\begin{array}{cc} A_1 & B_1 \\ A_2 & B_2 \end{array}\right|\right|\ , \tag{1.110}$$

a relation that should be familiar from elementary geometry. Equally familiar should be the expression for the 3-dimensional volume of a parallelepiped with legs **A**, **B**, and **C**:

$$\text{3-Volume} = \epsilon_{ijk} A_i B_j C_k = \mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \det\left\| \begin{array}{ccc} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{array} \right\| . \tag{1.111}$$

Recall that this volume has a sign: it is positive if $\{\mathbf{A},\, \mathbf{B},\, \mathbf{C},\}$ is a right handed set of vectors and negative if left-handed. The generalization to 4-dimensional spacetime should be obvious: The 4-dimensional parallelepiped whose legs are the four vectors $\vec{A}, \vec{B}, \vec{C}, \vec{D}$ has a 4-dimensional volume given by

$$\text{4-Volume} = \epsilon_{\alpha\beta\gamma\delta} A^\alpha B^\beta C^\gamma D^\delta = \boldsymbol{\epsilon}(\vec{A}, \vec{B}, \vec{C}, \vec{D}) = \det\left\| \begin{array}{cccc} A^0 & B^0 & C^0 & D^0 \\ A^1 & B^1 & C^1 & D^1 \\ A^2 & B^2 & C^2 & D^2 \\ A^3 & B^3 & C^3 & D^3 \end{array} \right\| . \tag{1.112}$$

Note that this 4-volume is positive if the set of vectors $\{\vec{A},\, \vec{B},\, \vec{C},\, \vec{D}\}$ is right-handed and negative if left-handed.

Just as Eqs. (1.101) and (1.110) give us a way to perform area integrals in 2- and 3-dimensional Euclidean space, so Equation (1.112) provides us a way to perform volume integrals over 4-dimensional Minkowski spacetime: To integrate some tensor field $\mathsf{\mathbf{T}}$ over some region $\mathcal{V}$ of spacetime, we need only divide spacetime up into tiny parallelepipeds, multiply the 4-volume $d\Sigma$ of each parallelepiped by the value of $\mathsf{\mathbf{T}}$ at its center, and add. It is not hard to see from Eq. (1.112) that in any right-handed Lorentz coordinate system, the 4-volume of a tiny parallelepiped whose edges are dx^α along the four coordinate axes is $d\Sigma = dt dx dy dz$, and correspondingly the integral of $\mathsf{\mathbf{T}}$ over $\mathcal{V}$ can be expressed as

$$\int_{\mathcal{V}} T^{\alpha\beta\gamma} d\Sigma = \int_{\mathcal{V}} T^{\alpha\beta\gamma} dt dx dy dz\ . \tag{1.113}$$

The analogous expressions in 2- and 3-dimensional Euclidean space should be obvious and familiar.

In Euclidean 3-space, we define the vectorial surface area of a parallelogram with legs **A** and **B** to be

$$\boldsymbol{\Sigma} = \mathbf{A}\times\mathbf{B} = \boldsymbol{\epsilon}(\_\,, \mathbf{A}, \mathbf{B})\ . \tag{1.114}$$

This vectorial surface area has a magnitude equal to the area of the parallelogram and a direction perpendicular to it. Such vectorial surface areas are the foundation for surface integrals in 3-dimensional space, and for the familiar Gauss and Stokes theorems:

$$\int_{\mathcal{V}_3} (\div\mathbf{A}) d\text{Volume} = \int_{\partial\mathcal{V}_3} \mathbf{A}\cdot d\boldsymbol{\Sigma} \tag{1.115}$$

(where $\mathcal{V}_3$ is a 3-dimensional region and $\partial\mathcal{V}_3$ is its two-dimensional boundary),

$$\int_{\mathcal{V}_2} \boldsymbol{\nabla}\times\mathbf{A}\cdot d\boldsymbol{\Sigma} = \int_{\partial\mathcal{V}_2} \mathbf{A}\cdot d\mathbf{l} \tag{1.116}$$

(where $\mathcal{V}_2$ is a 2-dimensional region, $\partial\mathcal{V}_2$ is the 1-dimensional curve that bounds it, and the last integral is a line integral around that curve).

Notice that in Euclidean 3-space, the vectorial surface area $\boldsymbol{\epsilon}(\_, \mathbf{A}, \mathbf{B})$ of the parallelogram with legs $\mathbf{A}$ and $\mathbf{B}$ can be thought of as an object that is waiting for us to insert a third leg $\mathbf{C}$ so as to compute a volume $\boldsymbol{\epsilon}(\mathbf{C}, \mathbf{A}, \mathbf{B})$—the volume of the parallelepiped with legs $\mathbf{C}$, $\mathbf{A}$, and $\mathbf{B}$.

By analogy, in 4-dimensional spacetime any 3-dimensional parallelepiped with legs $\vec{A}, \vec{B}, \vec{C}$ has a vectorial 3-volume $\vec{\Sigma}$ (not to be confused with the scalar 4-volume Σ) defined by

$$\vec{\Sigma}(\_) = \boldsymbol{\epsilon}(\_, \vec{A}, \vec{B}, \vec{C}) \; ; \quad \Sigma_\mu = \epsilon_{\mu\alpha\beta\gamma} A^\alpha B^\beta C^\gamma \; . \tag{1.117}$$

Here we have written the volume vector both in abstract notation and in component notation. This volume vector has one empty slot, ready and waiting for a fourth vector ("leg") to be inserted, so as to compute the 4-volume Σ of a 4-dimensional parallelepiped.

Notice that the volume vector $\vec{\Sigma}$ is orthogonal to each of its three legs (because of the antisymmetry of $\boldsymbol{\epsilon}$), and thus (unless it is null) it can be written as $\vec{\Sigma} = V\vec{n}$ where V is the magnitude of the volume and $\vec{n}$ is the unit normal to the three legs.

Interchanging any two legs of the parallelepiped reverses the 3-volume's sign. Consequently, the 3-volume is characterized not only by its legs but also by the order of its legs, or equally well, in two other ways: (*i*) by the direction of the vector $\vec{\Sigma}$ (reverse the order of the legs, and the direction of $\vec{\Sigma}$ will reverse); and (*ii*) by the *sense* of the 3-volume, defined as follows. Just as a 2-volume (i.e., a segment of a plane) in 3-dimensional space has two sides, so a 3-volume in 4-dimensional spacetime has two sides; cf. Fig. 1.12. Every vector $\vec{D}$ for which $\vec{\Sigma} \cdot \vec{D} > 0$ points out of one side of the 3-volume $\vec{\Sigma}$. We shall call that side the "positive side" of $\vec{\Sigma}$; and we shall call the other side, the one out of which point vectors $\vec{D}$ with $\vec{\Sigma} \cdot \vec{D} < 0$, its "negative side". When something moves through or reaches through or points through the 3-volume from its negative side to its positive side, we say that this thing is moving or reaching or pointing in the "positive sense"; and similarly for "negative sense". The examples shown in Fig. 1.12 should make this more clear.

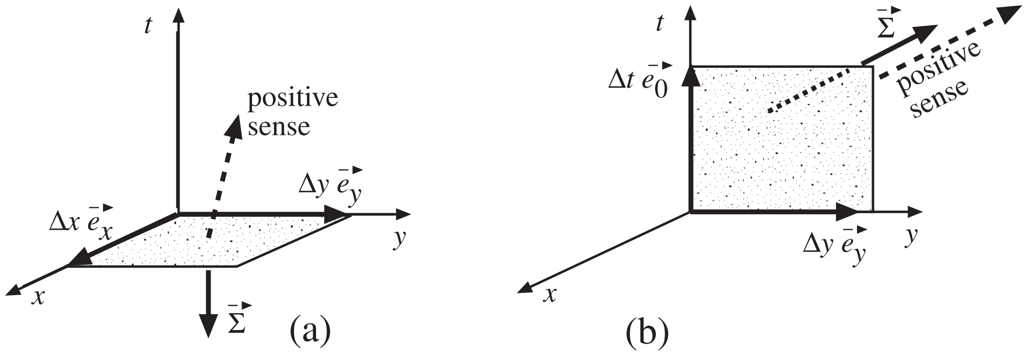


Fig. 1.12: Spacetime diagrams depicting 3-volumes in 4-dimensional spacetime, with one spatial dimension (that along the z-direction) suppressed.

Figure 1.12(a) shows two of the three legs of the volume vector $\vec{\Sigma} = \boldsymbol{\epsilon}(\_, \Delta x\vec{e}_x, \Delta y\vec{e}_y, \Delta z\vec{e}_z)$, where x, y, z are the spatial coordinates of a specific Lorentz frame. It is easy to show that this vector can also be written as $\vec{\Sigma} = -\Delta V \vec{e}_0$, where ΔV is the ordinary volume of the parallelepiped as measured by an observer in the chosen Lorentz frame, $\Delta V = \Delta x \Delta y \Delta z$. Thus, the direction of the vector $\vec{\Sigma}$ is toward the past (direction of decreasing Lorentz time t). From this, and the fact that timelike vectors have negative squared length, it is easy to infer that $\vec{\Sigma} \cdot \vec{D} > 0$ if and only if the vector $\vec{D}$ points out of the "future" side of the 3-volume (the side of increasing Lorentz time t), i.e., the positive side of $\vec{\Sigma}$ is the future side. This means that the vector $\vec{\Sigma}$ points in the negative sense of its own 3-volume.

Figure 1.12(b) shows two of the three legs of the volume vector $\vec{\Sigma} = \boldsymbol{\epsilon}(\_, \Delta t \vec{e}_t, \Delta y \vec{e}_y, \Delta z \vec{e}_z) = -\Delta t \Delta A \vec{e}_x$ (with $\Delta A = \Delta y \Delta z$). In this case, $\vec{\Sigma}$ points in its own positive sense.

This peculiar behavior is completely general: When the normal to a 3-volume is timelike, its volume vector $\vec{\Sigma}$ points in the negative sense; when the normal is spacelike, $\vec{\Sigma}$ points in the positive sense; and—it turns out—when the normal is null, $\vec{\Sigma}$ lies in the 3-volume (parallel to its one null leg) and thus points neither in the positive sense nor the negative.[10]

Note the physical interpretations of the 3-volumes of Fig. 1.12: That in Fig. 1.12(a) is an instantaneous snapshot of an ordinary, spatial, parallelepiped, while that in Fig. 1.12(b) is the 3-dimensional region in spacetime swept out during time Δt by the parallelogram with legs $\Delta y \vec{e}_y$, $\Delta z \vec{e}_z$ and with area $\Delta A = \Delta y \Delta z$.

Just as in 3-dimensional Euclidean space, vectorial surface areas can be used to construct 2-dimensional surface integrals, so also (and in identically the same manner) in 4-dimensional spacetime, vectorial volume elements can be used to construct integrals over 3-dimensional volumes, e.g. $\int_{\mathcal{V}_3} \vec{A} \cdot d\vec{\Sigma}$. More specifically: Let (a, b, c) be (possibly curvilinear) coordinates in the 3-surface $\mathcal{V}_3$, and denote by $\vec{x}(a, b, c)$ the spacetime point $\mathcal{P}$ on $\mathcal{V}_3$ whose coordinate values are (a, b, c). Then $(\partial\vec{x}/\partial a)da$, $(\partial\vec{x}/\partial b)db$, $(\partial\vec{x}/\partial c)dc$ are the vectorial legs of the elementary parallelepiped whose corners are at (a, b, c), $(a+da, b, c)$, $(a, b+db, c)$, etc; and the spacetime components of these vectorial legs are $(\partial x^\alpha/\partial a)da$, $(\partial x^\alpha/\partial b)db$, $(\partial x^\alpha/\partial c)dc$. The 3-volume of this elementary parallelepiped is $d\vec{\Sigma} = \boldsymbol{\epsilon}(\_, (\partial\vec{x}/\partial a)da, (\partial\vec{x}/\partial b)db, (\partial\vec{x}/\partial c)dc)$, which has spacetime components

$$d\Sigma_\mu = \epsilon_{\mu\alpha\beta\gamma} \frac{\partial x^\alpha}{\partial a} \frac{\partial x^\beta}{\partial b} \frac{\partial x^\gamma}{\partial c} dadbdc \,. \tag{1.118}$$

This is the integration element to be used when evaluating

$$\int_{\mathcal{V}_3} \vec{A} \cdot d\nabla\Sigma = \int A^\mu d\Sigma_\mu \,. \tag{1.119}$$

Just as there are Gauss and Stokes theorems for integrals in Euclidean 3-space, so also there are Gauss and Stokes theorems in spacetime. The Gauss theorem has the obvious form

$$\int_{\mathcal{V}_4} (\vec{\nabla} \cdot \vec{A}) d\Sigma = \int_{\partial\mathcal{V}_4} \vec{A} \cdot d\vec{\Sigma} \,, \tag{1.120}$$

where the first integral is over a 4-dimensional region $\mathcal{V}_4$ in spacetime, and the second is over the 3-dimensional boundary of $\mathcal{V}_4$, with the boundary's positive sense pointing outward, away from $\mathcal{V}_4$ (just as in the 3-dimensional case). We shall not write down the 4-dimensional

---

[10] This peculiar behavior gets replaced by a simpler description if one uses one-forms rather than vectors to describe 3-volumes; see, e.g., Box 5.2 of Misner, Thorne, and Wheeler (1973).

Stokes theorem because it is complicated to formulate with the tools we have developed thus far; easy formulation requires the concept of a *differential form,* which we shall not introduce in this book.

****

EXERCISES

Exercise 1.20 *Practice: Evaluation of 3-Surface Integral in Spacetime*
In Minkowski spacetime the set of all events separated from the origin by a timelike interval a^2 is a 3-surface, the hyperboloid $t^2 - x^2 - y^2 - z^2 = a^2$, where $\{t, x, y, z\}$ are Lorentz coordinates of some inertial reference frame. On this hyperboloid introduce coordinates $\{\chi, \theta, \phi\}$ such that

$$t = a\cosh\chi\ , \quad x = a\sinh\chi\sin\theta\cos\phi\ , \quad y = a\sinh\chi\sin\theta\sin\phi;\ , \quad z = a\sinh\chi\cos\theta\ . \tag{1.121}$$

Note that χ is a radial coordinate and (θ, ϕ) are spherical polar coordinates. Denote by $\mathcal{V}_3$ the portion of the hyperboloid with $\chi \leq b$.

a. Verify that for all values of (χ, θ, ϕ), the points (1.121) do lie on the hyperboloid.

b. On a spacetime diagram, draw a picture of $\mathcal{V}_3$, the $\{\chi, \theta, \phi\}$ coordinates, and the elementary volume element (vector field) $d\vec{\Sigma}$.

c. Set $\vec{A} \equiv \vec{e}_0$ (the temporal basis vector), and express $\int_{\mathcal{V}_3} \vec{A} \cdot d\vec{\Sigma}$ as an integral over $\{\chi, \theta, \phi\}$. Evalute the integral.

d. Consider a closed 3-surface consisting of the segment $\mathcal{V}_3$ of the hyperboloid as its top, the hypercylinder $\{x^2 + y^2 + z^2 = a^2\sinh^2 b,\ 0 < t < a\cosh b\}$ as its sides, and the sphere $\{x^2 + y^2 + z^2 \leq a^2\sinh^2 b\ ,\ t = 0\}$ as its bottom. Draw a picture of this closed 3-surface on a spacetime diagram. Use Gauss's theorem, applied to this 3-surface, to show that $\int_{\mathcal{V}_3} \vec{A} \cdot d\vec{\Sigma}$ is the 3-volume of its spherical base.

****

1.12 The Stress-energy Tensor and Conservation of 4-Momentum[11]

We conclude this chapter with a very general formulation of the law of 4-momentum conservation—a formulation that we shall find useful in the next chapter and later in this book.

---

[11] For further detail on the topics of this section, see, e.g., Chapter 5 of Misner, Thorne, and Wheeler (1993).

Suppose that a continuous medium (e.g., a gas) or a continuous field (e.g., electromagnetic waves) flows through spacetime, carrying with it 4-momentum. If the 3-volume is that of Fig. 1.12(a), $\vec{\Sigma} = -\Delta V \vec{e}_0$, then the total 4-momentum carried through $\vec{\Sigma}$ in the positive sense (from past toward future) by the medium or field is readily seen to be the total 4-momentum that an observer in the chosen Lorentz frame measures to lie in $\vec{\Sigma}$ at the moment $t = 0$ of $\vec{\Sigma}$'s brief existence. Similarly, if the 3-volume is that of Fig. 1.12(b), $\vec{\Sigma} = -\Delta A \Delta t \vec{e}_x$, then the total 4-momentum carried through $\vec{\Sigma}$ in the positive sense (from $+x$ toward $-x$) is the 4-momentum that an observer in the chosen Lorentz frame would see cross the area ΔA, from $+x$ toward $-x$, during time Δt.

It is easy to convince oneself on physical grounds that the total 4-momentum carried through these and any other tiny 3-volumes is linear in the 3-volumes. More specifically, if the size of a tiny 3-volume is doubled, then the amount of 4-momentum that flows through it will double; and if a new 3-volume is constructed as the sum of two old 3-volumes, then the total 4-momentum that flows through the new one will be the sum of that which flows through the two old ones. This means that we can define a second-rank *stress-energy tensor* $\mathbf{T}$ as that real-valued linear function with two slots such that, if we insert into the second slot a volume vector $\vec{\Sigma}$ and leave the first slot empty, we will get out the total 4-momentum that flows through $\vec{\Sigma}$ from negative side toward positive:

$$\mathbf{T}(\_, \vec{\Sigma}) = (\text{total 4-momentum } \vec{P} \text{ that flows through } \vec{\Sigma}); \quad \text{i.e., } T^{\alpha\beta}\Sigma_\beta = P^\alpha \, . \tag{1.122}$$

Of course, this stress-energy tensor is different at different locations in spacetime; i.e., it is a tensor field: if one wants to know the 4-momentum which flows through a 3-volume located at an event $\mathcal{P}$, one must do the calculation (1.122) using the value of $\mathbf{T}$ appropriate to that location, $\mathbf{T}(\mathcal{P})$.

¿From this definition of the stress-energy tensor we can read off the physical meanings of its components on a specific, but arbitrary, Lorentz-coordinate basis: Making use of method (1.50) for computing the components of a vector or tensor, we see that in a specific, but arbitrary, Lorentz frame (where $\vec{\Sigma} = -\vec{e}_0$ is a volume vector representing a parallelepiped with unit volume $\Delta V = 1$, at rest in that frame, with its positive sense toward the future):

$$\begin{aligned} -T_{\alpha 0} = \mathbf{T}(\vec{e}_\alpha, -\vec{e}_0) = \vec{P}(\vec{e}_\alpha) &= \begin{pmatrix} \alpha\text{-component of 4-momentum that} \\ \text{flows from past to future across a unit} \\ \text{volume } \Delta V = 1 \text{ in the 3-space } t = \text{const} \end{pmatrix} \\ &= (\alpha\text{-component of density of 4-momentum }) \, . \end{aligned} \tag{1.123}$$

Specializing α to be a time or space component and raising indices, we obtain the specialized versions of (1.123)

$$\begin{aligned} T^{00} &= (\text{energy density as measured in the chosen Lorentz frame}), \\ T^{j0} &= (\text{density of } j\text{-component of momentum in that frame}). \end{aligned} \tag{1.124}$$

Similarly, the αx component of the stress-energy tensor has the meaning

$$T_{\alpha x} \equiv \mathbf{T}(\vec{e}_\alpha, \vec{e}_x) = \begin{pmatrix} \alpha\text{-component of 4-momentum that crosses} \\ \text{a unit area } \Delta y \Delta z = 1 \text{ lying in a surface of} \\ \text{constant } x, \text{ during unit time } \Delta t, \text{ crossing} \\ \text{from the } -x \text{ side toward the } +x \text{ side} \end{pmatrix}$$

$$= \left(\begin{array}{c} \alpha \text{ component of flux of 4-momentum} \\ \text{across a surface lying perpendicular to } \vec{e}_x \end{array} \right) ; \tag{1.125}$$

and the specific forms of this for temporal and spatial α are

$$T^{0x} = \left(\begin{array}{c} \text{energy flux across a surface perpendidular to } \vec{e}_x, \\ \text{from the } -x \text{ side to the } +x \text{ side} \end{array} \right) , \tag{1.126}$$

$$T^{jx} = \left(\begin{array}{c} \text{flux of } j\text{-component of momentum across a surface} \\ \text{perpendicular to } \vec{e}_x, \text{ from the } -x \text{ side to the } +x \text{ side} \end{array} \right) . \tag{1.127}$$

The αy and αz components have the obvious, analogous interpretations.

These interpretations, restated much more briefly, are:

$$\begin{aligned} T^{00} &= (\text{energy density}) \ , \qquad T^{j0} = (\text{momentum density}) \ , \\ T^{0j} &= (\text{energy flux}) \ , \qquad\quad T^{jk} = (\text{stress}) \ . \end{aligned} \tag{1.128}$$

The stress deserves special attention: Corresponding to a specific Lorentz frame there is a specific 3-space of simultaneity $t = \text{const}$, and in that 3-space lives the *stress tensor* $\mathbf{T}$ of 3-dimensional (Newtonian) physics. That Newtonian stress tensor is defined, in analogy with the relativistic stress-energy tensor (1.122), by

$$\mathbf{T}(\_\,, \Sigma) = (\text{total momentum } \mathbf{p} \text{ that flows through the 2-surface } \Sigma \text{ per unit time } t). \tag{1.129}$$

It is straightforward to verify that the components T_{jk} of this stress tensor on the orthonormal basis of a Cartesian coordinate system are identical to the spatial components of the 4-dimensional stress-energy tensor: by analogy with Eq. (1.127)

$$\begin{aligned} T_{jk} = T^{jk} &= \left(\begin{array}{c} j\text{-component of momentum that crosses a unit} \\ \text{area which is perpendicular to } \vec{e}_k, \text{ per unit time,} \\ \text{with the crossing being from } -x^k \text{ to } +x^k \end{array} \right) \\ &= \left(\begin{array}{c} j\text{-component of force per unit area} \\ \text{across a surface perpendicular to } \vec{e}_k \end{array} \right) . \end{aligned} \tag{1.130}$$

As special cases, T_{xx} is the *pressure* in the x-direction, and T^{yx} is the y-directed *shear stress* across a surface of constant x.

Although it is not obvious at first sight, the 4-dimensional stress-energy tensor is symmetric; in index notation (where indices can be thought of as representing the names of slots, or equally well components on an arbitrary basis)

$$T^{\alpha\beta} = T^{\beta\alpha} \ . \tag{1.131}$$

This symmetry can be deduced by a physical argument in a specific, but arbitrary, Lorentz frame: Consider, first, the $x0$ and $0x$ components, i.e., the x-components of momentum density and energy flux. A little thought, symbolized by the following heuristic equation, reveals that they must be equal

$$T^{x0} = \left(\begin{array}{c} \text{momentum} \\ \text{density} \end{array} \right) = \frac{(\Delta E) dx/dt}{\Delta x \Delta y \Delta z} = \frac{\Delta E}{\Delta y \Delta z \Delta t} = \left(\begin{array}{c} \text{energy} \\ \text{flux} \end{array} \right) , \tag{1.132}$$

and similarly for the other space-time and time-space components: $T^{j0} = T^{0j}$. [In Eq. (1.132), in the first expression ΔE is the total energy (or equivalently mass) in the volume $\Delta x \Delta y \Delta z$, $(\Delta E) dx/dt$ is the total momentum, and when divided by the volume we get the momentum density. The third equality is just elementary algebra, and the resulting expression is obviously the energy flux.]

Consider, next, the xy and yx components, i.e., components of the shear stress. One can show by elementary torque arguments (Chap. 10) that, if these components were not equal to each other, then on a tiny cube of material with side L and moment of inertia $\propto L^5$ there would be a net torque $(T^{xy} - T^{yx})L^3$ that induces an angular acceleration $\propto L^{-2}$; this angular acceleration would become infinitely large in the limit $L \to 0$, which is physically ridiculous. Correspondingly, it must be that $T^{xy} = T^{yx}$, and similarly for all other off-diagonal spatial components, $T^{jk} = T^{kj}$.

Since, by the above arguments, $T^{0j} = T^{j0}$ and $T^{jk} = T^{kj}$, all components in our chosen Lorentz frame are symmetric, $T^{\alpha\beta} = T^{\beta\alpha}$. This means that, if we insert arbitrary vectors into the slots of **T** and evaluate the resulting number in our chosen Lorentz frame, we will find

$$\mathbf{T}(\vec{A}, \vec{B}) = T^{\alpha\beta} A_\alpha B_\beta = T^{\beta\alpha} A_\alpha B_\beta = \mathbf{T}(\vec{B}, \vec{A}) \ . \tag{1.133}$$

This shows that **T** is symmetric under interchange of its slots.

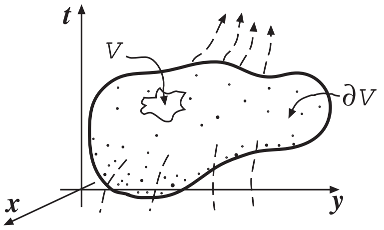


Fig. 1.13: The 4-dimensional region $\mathcal{V}$ in spacetime, and its closed 3-boundary $\partial\mathcal{V}$, used in formulating the law of 4-momentum conservation. The dashed lines symbolize, heuristically, the flow of 4-momentum from past toward future.

With the aid of the stress-energy tensor we can now give an elegant, general, geometric and frame-independent formulation of the conservation of 4-momentum—a generalization of the formulation for particles shown in Fig. 1.6 above: Let $\mathcal{V}$ be a compact, 4-dimensional region of spacetime and denote by $\partial\mathcal{V}$ its boundary, a closed 3-surface in 4-dimensional spacetime (Fig. 1.13). The fields and media present in spacetime carry 4-momentum through $\mathcal{V}$, from the past toward the future. The law of 4-momentum conservation says that all the 4-momentum which enters $\mathcal{V}$ through the past part of its boundary $\partial\mathcal{V}$ must exit through the future part of its boundary. If we choose the positive sense of the boundary's infinitesimal 3-volume $d\vec{\Sigma}$ to point out of $\mathcal{V}$ (toward the past on the bottom boundary and toward the future on the top), then this conservation law can be expressed mathematically as

$$\int_{\partial\mathcal{V}} T^{\alpha\beta} d\Sigma_\beta = 0 \ . \tag{1.134}$$

This *global law of 4-momentum conservation* can be converted into a *local law* with the help of the 4-dimensional Gauss theorem (1.120) (generalized in the obvious way from a vectorial integrand to a tensorial one):

$$\int_{\partial\mathcal{V}} T^{\alpha\beta} d\Sigma_\beta = \int_{\mathcal{V}} T^{\alpha\beta}{}_{;\beta} d\Sigma \ . \tag{1.135}$$

Since the left-hand side vanishes, so must the right-hand side; and in order for this 4-volume integral to vanish for every choice of $\mathcal{V}$, it is necessary that the integrand vanish everywhere in spacetime:

$$T^{\alpha\beta}{}_{;\beta} = 0 \ ; \quad \text{i.e.,} \ \vec{\boldsymbol{\nabla}} \cdot \mathbf{T} = 0 \ . \tag{1.136}$$

In the second, index-free version of this local conservation law, the ambiguity about which slot the divergence is taken on is unimportant, since $\mathbf{T}$ is symmetric in its two slots: $T^{\alpha\beta}{}_{;\beta} = T^{\beta\alpha}{}_{;\beta}$.

In a specific but arbitrary Lorentz frame, the local conservation law (1.136) for 4-momentum has as its temporal and spatial parts

$$\frac{\partial T^{00}}{\partial t} + \frac{\partial T^{0k}}{\partial x^k} = 0 \ , \tag{1.137}$$

i.e., the time derivative of the energy density plus the 3-divergence of the energy flux vanishes; and

$$\frac{\partial T^{j0}}{\partial t} + \frac{\partial T^{jk}}{\partial x^k} = 0 \ , \tag{1.138}$$

i.e., the time derivative of the momentum density plus the 3-divergence of the stress (i.e., of the momentum flux) vanishes. Thus, as one should expect, the geometric, frame-independent law of 4-momentum conservation includes as special cases both the conservation of energy and the conservation of momentum.

As an important example that illustrates the stress-energy tensor, we shall consider a *perfect fluid*. A perfect fluid is a continuous medium whose stress-energy tensor, evaluated in its *local rest frame* (a Lorentz frame where $T^{j0} = T^{0j} = 0$), has the special form

$$T^{00} = \rho \ , \quad T^{jk} = P\delta^{jk} \ . \tag{1.139}$$

Here ρ is a short-hand notation for the energy density (density of total mass-energy, including rest mass) T^{00}, as measured in the local rest frame; and the stress tensor T^{jk} as measured in that frame has the form of an isotropic pressure P, and vanishing shear stress. From this special form of $T^{\alpha\beta}$ in the local rest frame, one can derive the following expression for the stress-energy tensor in terms of the 4-velocity $\vec{u}$ of the local rest frame, i.e., of the fluid itself, the metric tensor of spacetime $\mathbf{g}$, and the rest-frame energy density ρ and pressure P:

$$T^{\alpha\beta} = (\rho + P)u^\alpha u^\beta + P g^{\alpha\beta} \ ; \quad \text{i.e.,} \ \mathbf{T} = (\rho + P)\vec{u} \otimes \vec{u} + P\mathbf{g} \ ; \tag{1.140}$$

see Exercise 1.22, below. In Part III of this book, we shall explore in depth the implications of this stress-energy tensor.

Another example of a stress-energy tensor is that for the electromagnetic field:

$$T^{\alpha\beta} = \frac{1}{4\pi}\left(F^{\alpha\mu}F^{\beta}{}_{\mu} - \frac{1}{4}g^{\alpha\beta}F^{\mu\nu}F_{\mu\nu}\right) \tag{1.141}$$

see Exercise 1.23

****

EXERCISES

Exercise 1.21 *Example: Global Conservation of 4-Momentum in a Lorentz Frame*
Consider the 4-dimensional parallelepiped $\mathcal{V}$ whose legs are $\Delta t\partial/\partial t$, $\Delta x\partial/\partial x$, $\Delta y\partial/\partial y$, $\Delta z\partial/\partial z$, where (t, x, y, z) are the coordinates of some Lorentz frame. The boundary $\partial\mathcal{V}$ of this $\mathcal{V}$ has eight 3-dimensional "faces". Identify these faces, and write the integral $\int_{\partial\mathcal{V}} T^{0\beta}d\Sigma_\beta$ as the sum of contributions from each of them. According to the law of energy conservation, this sum must vanish. Explain the physical interpretation of each of the eight contributions to this energy conservation law.

Exercise 1.22 *Derivation: Stress-Energy Tensor for a Perfect Fluid*
Derive the frame-independent expression (1.140) for the perfect fluid stress-energy tensor from its rest-frame components (1.139).

Exercise 1.23 *Problem: Stress-Energy Tensor for Electromagnetic Field*
Compute from Eq. (1.141) the components of the electromagnetic stress-energy tensor in an inertial reference frame. Your answer should be the expressions given in electrodynamic textbooks:

$$\begin{aligned} T^{00} &= \frac{\mathbf{E}^2+\mathbf{B}^2}{8\pi}\,, \quad T^{0j}\mathbf{e}_j = T^{j0}\mathbf{e}_j = \frac{\mathbf{E}\times\mathbf{B}}{4\pi}\,, \\ T^{jk} &= \frac{1}{8\pi}\left[(\mathbf{E}^2+\mathbf{B}^2)\delta_{jk} - 2(E_jE_k + B_jB_k)\right]\,. \end{aligned} \tag{1.142}$$

****

Bibliography

Einstein, Albert 1905. "Zur Elektrodynamic bewegter Korper," *Annalen der Physik*, **17**, 891–921; English translation in Einstein *et al.* (1923).

Einstein, Albert, Lorentz, Hendrik A., Minkowski, Hermann, and Weyl, Hermann 1923. *The Principle of Relativity*, New York: Dover Publications.

Feynman, Richard P. 1966. *The Character of Physical Law*, M.I.T. Press, Cambridge, Massachusetts.

Goldstein, Herbert 1980. *Classical Mechanics*, New York: Addison Wesley.

Gott, Richard 1991. "Closed Timelike Curves Produced by Pairs of Moving Cosmic Strings—Exact Solutions" *Physical Review Letters*, **66**, 1126-1129.

Hafele, J. C., and Keating, Richard E. 1972a. "Around-the-World Atomic Clocks: Predicted Relativistic Time Gains" *Science*, **177**, 166-168.

Hafele, J. C., and Keating, Richard E. 1972b. "Around-the-World Aotmic Clocks: Observed Relativistic Time Gains"*Science*, **177**, 168-170.

Hartle, J. B. 2000

Hawking, Stephen W. 1992. "The Chronology Protection Conjecture" *Physical Review D*, **46**, 603-611.

Jackson, John David 1999. *Classical Electrodynamics*, New York: Wiley.

Kim, Sung-Won and Thorne, Kip S. 1991. "Do Vacuum Fluctuations Prevent the Creation of Closed Timelike Curves?" *Physical Review D*, **43**, 3929-3947.

Lorentz, Hendrik A. 1904. "Electromagnetic Phenomena in a System Moving with Any Velocity Less than that of Light" *Proceedings of the Academy of Sciences of Amsterdam*, **6**, 809; reprinted in Einstein *et al.* (1923).

Mathews, Jon and Walker, R. L. 1964. *Mathematical Methods of Physics*, New York: Benjamin.

Minkowski, Hermann 1908. "Space and Time," Address to the 80th Assembly of German Natural Scientists and Physicians, at Cologne, 21 September 1908; text published posthumously in *Annalen der Physik*, **47**, 927 (1915); English translation in Einstein *et al.* (1923).

Misner, Charles W., Thorne, Kip S., and Wheeler, John A. 1973. *Gravitation*, San Francisco: Freeman.

Morris, Michael S. and Thorne, Kip S. 1987. "Wormholes in Spacetime and their Use for interstellar Travel—a Tool for Teaching General Relativity" *American Journal of Physics*, **56**, 395-412.

Morris, Michael S., Thorne, Kip S., and Yurtsever, Ulvi 1987. "Wormholes, Time Machines, and the Weak Energy Condition" *Physical Review Letters*, **61**, 1446-1449.

Schutz, Bernard F. 1985. *A First Course in General Relativity*, Cambridge: Cambridge University Press.

Taylor, Edwin F. and Wheeler, John A. 1992. *Spacetime Physics: Introduction to Special Relativity*, San Francisco: Freeman.

Thorne, Kip S. 1993. "Closed Timelike Curves" *General Relativity and Gravitation 1992*, edited by R. J. Gleiser, C. N. Kozameh and O. M. Moreschi, Cambridge University Press, Cambridge, England.

Wald, Robert M. 1984. *General Relativity*, Chicago: University of Chicago Press.