Special Relativity: Quirked Up, but Goated W/ Sauce

Duncan Wilkie

8 June 2022

Foundation

Einstein derived the mathematical formalism of SR from two axioms:

- The laws of physics are the same in any coordinate system with the origin moving at constant velocity (i.e. *there is no absolute rest*)
- The speed of light in the vacuum is constant

Motivation

Why these axioms? Two reasons, as he explains in the eminently readable paper *Zur Elektrodynamik bewegter Körper*:

- The observable characteristics of first mechanics and then electrodynamics just so happened to depend only on the relative motion of bodies.
- Despite the motion of Earth, the velocity of light appears to be isotropic.

Motivation: Electrodynamics

To illustrate Einstein's first point, one needs a crash-course in electromagnetism. The subject is about two vector fields: $\vec{E}$, the electric field produced by charged particles, and $\vec{B}$, the magnetic field produced by moving charges. These fields affect the dynamics of charges by Coulomb's law

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$$

Before Faraday, the subject was only "electrostatics:" the study of situations involving unchanging fields. The known laws were

$$\nabla \cdot \vec{E} = \frac{1}{\epsilon_0}\rho \qquad \text{(Gauss's law)}$$

$$\nabla \times \vec{E} = 0 \qquad (\vec{E} \text{ has a scalar potential})$$

$$\nabla \cdot \vec{B} = 0 \qquad \text{(No magnetic monopoles)}$$

$$\nabla \times \vec{B} = \mu_0 \vec{J} \qquad \text{(Ampère's law)}$$

Motivation: Electrodynamics

Faraday, through pure, brute-force experimentation, deduced that *a changing magnetic field induces an electric field*; mathematically,

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad \text{(Faraday's law)}$$

In terms of the voltage associated with this electric field in a wire loop, this is written

$$\mathcal{E} = -\frac{d\Phi}{dt}$$

Motivation: Electrodynamics

Maxwell noted a mathematical inconsistency in these four fundamental equations. Take the divergence of Ampère's law:

$$\nabla \cdot (\nabla \times \vec{B}) = \mu_0(\nabla \cdot \vec{J})$$

It's an easy proof that for arbitrary C^2 fields $\vec{B}$ that the left side is zero (exercise). However, in almost every case with a changing current, the right side will be nonzero. Noting that by Gauss's law and conservation of charge $\nabla \cdot \vec{J} = -\nabla \cdot \left(\epsilon_0 \frac{\partial \vec{E}}{\partial t}\right)$, Maxwell canceled off this nonzero term by modifying Ampère's law to be

$$\nabla \times \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}$$

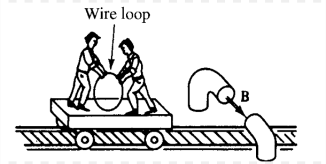
Despite being difficult to observe in the laboratory, this law can be confirmed with high-frequency AC sources (cf. Hertz 1888).

Motivation: Electrodynamics

Maxwell's correction has a very nice interpretation: just as Faraday says *a changing magnetic field induces an electric field*, so *a changing electric field induces a magnetic field*.

Motivation: Electrodynamics

Notice, however, that unlike classical mechanics, it is not at all obvious that a "principle of relativity" applies for electromagnetism.



In the situation pictured above, from the perspective of the ground, the force that creates a current is *purely magnetic*; from the perspective of the train, it is *purely electric*. It *just so happens* that both perspectives yield the same voltage in the loop.

Einstein conjectured that *all physical theories* must give the same results when calculations are performed in inertial reference frames (coordinate systems moving at constant velocities).

Motivation: Electrodynamics

The primitive sources of electric and magnetic forces are point charges and line currents, respectively. These produce electric and magnetic fields (expressed in the natural cylindrical coordinates)

$$\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}\hat{r}$$

and

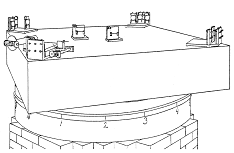
$$\vec{B} = \frac{\mu_0 I}{2\pi r}\hat{\theta}$$

Each law depends on a phenomenological parameter: ϵ_0, the permittivity of free space, and μ_0, the permeability of free space. Qualitatively, these determine "how easily" electric and magnetic fields can travel through the vacuum: if ϵ_0 is lower or μ_0 higher, the fields fall off quicker in the limit $r \to \infty$.

One can derive that for electromagnetic waves $\mu_0\epsilon_0 = \frac{1}{c^2}$; it wouldn't make sense for either μ_0 or ϵ_0 to vary with speed (they're about propagation through the vacuum, after all!), so perhaps c is fixed too.

Motivation: Michelson-Morley

In 1887, Michelson and Morley performed an experiment to try and find the speed of the Earth through "the æther," the conjectured medium through which electromagnetic waves propagate. Presumably, as with any other wave, the observed speed of light waves moving "downstream" from Earth in this medium would be the fastest. Through fancy interferometry, found an upper bound on that speed of $\pm 6\,\mathrm{km/s}$, despite the fact that with respect to the Sun we ought to be moving at $30\,\mathrm{km/s}$.



Thus Einstein conjectured that the speed of light in the vacuum is a universal constant. Modern measurements have gotten the estimate down to $10^{-7}\,\mathrm{km/s}$, and introduced the universal frame candidate of the CMB, with respect to which we move around 10x faster.

Consequences: The Objective

Clearly, saying the speed of light is *always* constant has drastic implications. If one is moving at the speed of light, the light one shines forward moves not at $2c$ but c. Speed is in units of m/s, and so we have two quantities to play with: we can change what "meter" means, i.e. lengths change, or what "seconds" means, i.e. time changes. It'll turn out to require both. In classical mechanics, an event located by coordinates (t, x, y, z) has coordinates (t', x', y', z') with respect to an origin moving with speed v along the x-axis given by

$$x' = x - vt$$

$$y' = y$$

$$z' = z$$

$$t' = t$$

Consequences: The Objective

This is a linear mapping $\mathbb{R}^4 \to \mathbb{R}^4$, and so may be expressed as a matrix with respect to the initial basis as

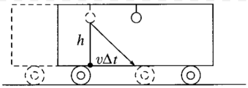
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(Check this by letting it act on the 4-vector (t, x, y, z)!)

Our goal is to find an analogous mapping for which Einstein's axioms hold.

Consequences: Time Dilation

We start by trying to find the time transformation. Consider a light bulb on the roof of a moving train car, pictured below.

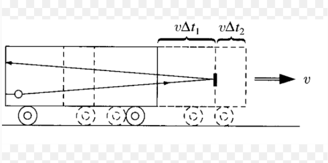


For an observer in the car (coordinates distinguished via priming), the light moves straight down, and the time it takes to strike the ground is simply $\Delta t' = h/c$. For an observer on the ground, however, it travels a longer distance, and so takes the longer time $\Delta t = \sqrt{h^2 + (v\Delta t)^2}/c$. Solving for Δt, one gets $\Delta t = \frac{h}{c}\frac{1}{\sqrt{1-v^2/c^2}}$. We then get the famous time dilation equation

$$\Delta t' = \Delta t\sqrt{1 - v^2/c^2} = \Delta t/\gamma \text{ (defining } \gamma)$$

Consequences: Length Contraction

Now we proceed similarly for length. Consider the same train car, but with the bulb on the back wall and a mirror on the front wall:



What's the travel time of the light? On the train, this is again simply $\Delta t' = 2\frac{\Delta x'}{c}$. On the ground it's more complicated. Letting $\Delta t_1 = \frac{\Delta x + v\Delta t_1}{c} \Leftrightarrow \Delta t_1 = \frac{\Delta x}{c-v}$ be the time to strike the mirror and $\Delta t_2 = \frac{\Delta x - v\Delta t_2}{c} = \frac{\Delta x}{c+v}$ be the time to strike the rear wall, the total time is $\Delta t = \Delta t_1 \Delta t_2 \frac{2\Delta x}{c}\frac{1}{1-v^2/c^2}$. We also have the time dilation formula; substituting these expressions of $\Delta t'$ and Δt in terms of the displacement yields the length contraction formula

$$\Delta x' = \frac{\Delta x}{\sqrt{1 - v^2/c^2}} = \gamma \Delta x$$

Consequences: Lorentz Transformations

Note that $\gamma = \frac{1}{\sqrt{1-v^2/c^2}} < 1$ for $v < c$. In summary, then, *moving clocks run slow* and *moving rulers get shorter*, both by a factor of γ. In principle, this is "all" of special relativity. In analogy to the Galilean transformations, we can write

$$\Lambda = \Lambda^{\nu}_{\mu} \begin{array}{l} t' = \gamma\left(t - \frac{v}{c^2}x\right) \\ x' = \gamma(x - vt) \\ y' = y \\ z' = z \end{array} \Leftrightarrow \begin{pmatrix} \gamma & -\gamma\frac{v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The x' equation follows from length-contracting *and* time-dilating the argument leading to the Galilean formula, and the t' equation may be deduced by plugging the formula for transforming x *from* the primed frame *to* the unprimed frame into that for x'. The matrix can be made Hermitian by taking the first component of the vectors it acts on be ct instead of t.

4-Vectors and the Minkowski Metric

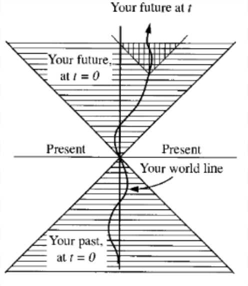
The Lorentz transformation acts on vectors with *four* components. We distinguish between these *spacetime* vectors (by convention with time as the first component) and ordinary space vectors by calling them 4-vectors, and using Greek indices when they're written in expressions. The analog of the dot product on 4-vectors is the same as $\mathbb{R}^4$, but with a negative first component:

$$a^\mu \cdot b^\mu = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3$$

Note the use of superscript indices; this is a critical point, as $a^\mu \neq a_\mu$, differing by the negation of the time component. As we'll see later, this is a metric tensor (the Minkowski metric) on the spacetime manifold.

Light Cones

Just as how $a \cdot a = |a|^2$ on $\mathbb{R}^4$, we can define the “distance” between two intervals in spacetime by the magnitude of their vector difference. An important distinction is that this “distance” can be negative; we can classify separations by the sign of this value. If it’s positive, the separation is spacelike, if it’s negative, timelike; zero, lightlike. These terms are motivated by inspecting so-called **light cones**: the points in the interior of the cone are spacelike, those on its boundary are lightlike, and the exterior is timelike.



Fun Problems

Miscellaneous physical problems

11. (*Special Relativity.*)

(a) Show that the linear transformation $x' = \Lambda x$, with

$$\text{(9.37)} \qquad \Lambda = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

with $|\beta| < 1, \gamma = (1-\beta^2)^{-1/2}$ is a proper Lorentz transformation, and that its inverse is obtained by replacing β by $-\beta$. Show that the spatial (x', y', z') coordinate axes are parallel to the (x, y, z) axes, and that the origin of the primed coordinate system is observed in the unprimed system to move along the positive x axis at velocity β. This Lorentz transformation is called a boost along the positive x axis at velocity β.

(b) (*Time dilation.*) In the primed coordinate system, two events are observed to occur at the same location but separated by the time interval T. Show that in the unprimed system they are separated by the time interval γT.

(c) (*Relativity of simultaneity.*) In the primed coordinate system, two events are observed to occur simultaneously at points on the x' axis a distance L apart. Show that in the unprimed system these events are separated by the time interval $\gamma\beta L$. Which one occurs first? Conversely, suppose that in the unprimed system two events occur at points on the x axis, separated by a distance L and a time interval T with $T^2 - L^2 < 0$. For such "spacelike separated" events, show that β can be chosen to make the events simultaneous in the primed frame.

(d) (*Length contraction.*) A ruler is at rest in the primed system, extending along the x' axis from the origin to $x' = L$. To measure its length, an observer in the unprimed system will note the location of its endpoints at some fixed time (her own t coordinate!) and measure the distance between them. Show that she will obtain L/γ.

Problem 12.9 A Lincoln Continental is twice as long as a VW Beetle, when they are at rest. As the Continental overtakes the VW, going through a speed trap, a (stationary) policeman observes that they both have the same length. The VW is going at half the speed of light. How fast is the Lincoln going? (Leave your answer as a multiple of c.)

Fun Paradoxes and Further Thinking

- Relativity of simultaneity
- Twin paradox
- Ladder paradox
- Shadows move faster than light

Relativistic Mechanics

This is all well and good, but how do we describe the motion of particles now? We define a notion of **proper quantities** (from a mistranslation of the French *propre*) associated with a moving object, analogous to the Newtonian quantities but playing nice with Lorentz transformations ($q' = \Lambda q$). The fundamental proper quantity is proper time τ: the time elapsed from the perspective of the moving object. The proper velocity η is the distance *in the observer frame* per unit time *in the object frame*

$$\vec{\eta} = \frac{d\vec{x}}{d\tau}$$

Proper analogues of all other kinematic and dynamic quantities (momentum, energy, force, etc.) are defined identically.

Relativistic Mechanics: Definition Dump

Proper quantities are Greek; classical, Roman. Quantities that are Roman but are defined in terms of proper quantities are usually referred to as “relativistic” rather than “proper.”

$$\Delta\tau = \Delta t/\gamma$$

$$\vec{\eta} = \frac{d\vec{x}}{d\tau} = \gamma\vec{v}$$

$$\vec{p} = m\eta = \gamma m\vec{v}$$

$$E = p^0 c = \gamma mc^2 \Rightarrow E^2 = p^2c^2 + m^2c^4$$

$$\vec{F} = \frac{d\vec{p}}{dt}$$

$$\vec{\alpha} = \frac{d^2\vec{x}}{d\tau^2}$$

$$\vec{K} = \frac{d\vec{\rho}}{d\tau} = m\vec{\alpha}$$

The convention was abrogated in the last equation to align with the convention for the Minkowski force (which is often called “proper”).

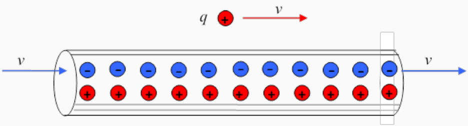
Relativistic Mechanics: Problems

Problem 12.34 In the past, most experiments in particle physics involved stationary targets: one particle (usually a proton or an electron) was accelerated to a high energy E, and collided with a target particle at rest (Fig. 12.29a). Far higher *relative* energies are obtainable (with the same accelerator) if you accelerate *both* particles to energy E, and fire them at each other (Fig. 12.29b). *Classically,* the energy $\bar{E}$ of one particle, relative to the other, is just 4E (why?)—not much of a gain (only a factor of 4). But *relativistically* the gain can be *enormous.* Assuming the two particles have the same mass, m, show that

$$\bar{E} = \frac{2E^2}{mc^2} - mc^2. \tag{12.59}$$

(a) (b)

Magnetism is Literally Just Relativity



Suppose the particle moves to the right at speed $u < v$, the negative charges move to the right at speed v, and the positive charges move to the left with speed v. There ought to be (little to) no net electrical force on the point charge in the frame of the wire, as the charge densities (charge per unit length, in terms of which current is defined) are of equal magnitude. However, what happens in the charge's frame? We must apply Einstein's velocity addition rule, yielding a line charge velocity of

$$v_{\pm} = \frac{v \mp u}{1 \mp vu/c^2}$$

Note in particular $v_{-} > v_{+}$, implying the spaces between the charges in the line are wider in the particle's frame, so *it should feel a net electric force in its frame*.

Magnetism is Literally Just Relativistic Electric Fields in Disguise

Moreover, one can prove using the relativistic dynamics that the magnitude of the electric force in the particle's frame is *identical to the classical expression of the magnetic field on a moving line charge*. The "primitive object" of the usual theory of magnetism are these line currents, so *all* magnetism is attributable to this relativistic effect. This is so deeply profound it bears explicit statement: **magnetic forces are nothing more than electric forces in the moving frame of the charge they affect**. [1]

[1] An introductory text that describes magnetism from this perspective *ab initio* (and more carefully) is Purcell's *Electricity and Magnetism*.

Elegant Closure

Maxwell's equations prettify greatly if one exploits this fact by writing them in a Lorentz-invariant form. We will proceed in much lighter detail due to much greater technicality. Before Einstein, many attempts were made to unify all four of Maxwell's equations into a single one. Most used some sort of tensorial description, to some success, as illustrated by the elegant expression for the electrical force on an arbitrary, time-varying charge configuration in an enclosing volume V due to Maxwell himself:

$$\vec{F} = \int_{\partial V} T_{ij} \cdot d\vec{a} - \epsilon_0 \mu_0 \frac{d}{dt} \int_V \vec{S} d\tau$$

T_{ij} is the so-called "Maxwell stress tensor", an object of a type we'll soon define leading to the final, power-level-9000 form of the laws governing classical electromagnetism.

Elegant Closure: Defining Tensors

Physicists and mathematicians disagree on what tensors are. Unfortunately, physicists' definition is objectively worse by orders of magnitude.

Definition
A tensor is a multilinear functional on arbitrary copies of a vector space or its dual (the vector space of 1-forms), i.e. a map, linear in each argument, of the form $f : V \times \cdots \times V \times V^* \times \cdots V^* \to \mathbb{F}$

Physicists get confused because many of their tensors don't look like they produce scalars. The aforementioned Maxwell stress tensor seems to produce a vector, the traction, but we can always say the tensor also takes as an argument a 1-form over which the output vector is integrated to produce a scalar, or another vector with which the output vector is dotted (canonical isomorphism between V and V^*). In other words, functions that act on vectors to produce a vector are just partial applications of a tensor.

Elegant Closure: The Four-Potential and Field Tensor

The field tensor takes as input a particle's velocity vector and outputs the Minkowski force on the particle due to the magnetic fields, generalizing the electric and magnetic fields into a single object that takes in a 4-velocity and outputs the 4-force:

$$F^{\mu\nu} = \begin{bmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{bmatrix}$$

All the information about the scalar and vector potentials V and $\vec{A}$ (which generate the fields by ∇V and $\nabla \times \vec{A}$, respectively) is contained in the similarly-invariant vector $A^\mu = (V/c, A_x, A_y, A_z)$, which is related to the field tensor by

$$F^{\mu\nu} = \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu}$$

Charge and current densities are described in this form by

$$J^\mu = (c\rho, J_x, J_y, J_z)$$

Elegant Closure: Maxwell's Final Form

The last thing we must introduce is a differential operator: the d'Alembertian, $\Box^2 = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$ is commonly used to describe wave phenomena, as the wave equation can be written $\Box^2 u = 0$. We can now describe all of electromagnetism in a single, two-term partial differential equation:

$$\boxed{\Box^2 A^\mu = -\mu_0 J^\mu}$$

Looking Forward

Relativistic electrodynamics lays the groundwork for general relativity by casting problems in this covariant, coordinate-independent light. This language is readily generalized to that of Hodge operators and differential forms:

$$d \star dA = \mu_0 J$$

similarly contains all the information of Maxwell's equations.

Hints of quantum mechanics pervade it as well: we know photons to be massless and move at the speed of light. The relativistic energy and momentum are *indeterminate* in this case—$0/0$. Therefore, the existence of lightspeed, massless particles that carry momentum and energy is consistent with relativity, but requires *new physics* to explain it. That physics came courtesy Planck, who kicked off quantum mechanics by establishing $E = h\nu$ for photons.

Conclusion: special relativity is unbelievably goated.