Chapter 1

Electromagnetism and Special Relativity

In the more than hundred years since Einstein introduced Relativity, Special and then General, an enormous amount of experimental evidence has accumulated confirming that nature is relativistic. Among the better-known instances are numerous particle physics experiments where particles near the speed of light are observed and even engineered, global positioning systems relying on the minute time-dilation effect of velocity and gravity, the standard cosmology capable of reproducing the visible universe and its sweeping history so well, and the most recent direct and indirect observations of gravitational waves.

After this remarkable century, General Relativity is an undisputed foundation of modern-day physics. Instead of dwelling on the historical backdrops and scientific thoughts of the pioneers that led to the development of this remarkable framework, we seek to focus on the most economical understanding of relativistic gravity, with an emphasis on the underlying fundamental structures.

Maxwell's electromagnetism offers the fastest route to Relativity. The celebrated Lorentz transformation, which played such a key role in Einstein's Special Relativity, was invented to ensure the invariance of the Maxwell equation under a change of inertial coordinate systems. The conflict between the Maxwell theory with its universal speed of electromagnetic waves and the Galilean kinematics eventually led to Einstein's Special Relativity. Along the way, other intellectual giants of the era such as Henri Poincaré contributed significantly as well.

The influence of Maxwell's electromagnetism goes much further than Special Relativity, as General Relativity itself may be considered as a generalization of the Maxwell theory. The familiar electromagnetic 4-vector potential is a prototype for the affine connection in Riemannian geometry, and the electromagnetic field strength inspires the Riemann curvature in this mathematical comparison. We will often rely on lessons from the Maxwell theory and its Yang–Mills generalizations, to motivate and illustrate the geometric structures behind General Relativity.

Another important heritage of electromagnetism is the notion of the "field." Faraday is responsible for this crucial invention that, instead of merely talking about the forces on particles, treated the field lines as physical reality. The notion became a central concept for modern physics, more so at the quantum level where elementary particles of nature are all deemed to be quantum excitation of such fields. For these reasons, it is proper and natural that we start with a succinct look back at the Maxwell theory.

In this book, we employ Einstein's convention that a pair of repeated indices, one superscript and the other subscript, are by default summed over. Instead of writing

$$\sum_i \dot{\mathbf{x}}^i A_i \ , \qquad \sum_{j,k} \epsilon^{ijk}\partial_j A_k \ , \qquad \sum_\alpha V^\alpha \Lambda_{\alpha\beta} \ ,$$

we will write simply

$$\dot{\mathbf{x}}^i A_i \ , \qquad \epsilon^{ijk}\partial_j A_k \ , \qquad V^\alpha \Lambda_{\alpha\beta} \ ,$$

etc. There are a handful of circumstances, where we must not sum over repeated indices, and we will declare so whenever such need arises.

1.1 Electromagnetism and the Lorentz Transformation

Let us recall the Maxwell equation that govern classical electromagnetism,

$$\vec{\partial} \cdot \vec{E} = \frac{1}{\epsilon_0}\rho_e \ ,$$

$$\vec{\partial} \cdot \vec{B} = 0 \ ,$$

$$\vec{\partial} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \ ,$$

$$\vec{\partial} \times \vec{B} = \mu_0 \vec{j}_e + \underbrace{\epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t}}_{\text{Maxwell}} \ . \tag{1.1}$$

As is well known, Maxwell's contribution, apart from elevating the above set of laws to mathematical precision, is to realize that the last term is needed for self-consistency.

What happens if we drop the last term and work with

$$\vec{\partial} \times \vec{B} = \mu_0 \vec{j}_e \tag{1.2}$$

instead? Imagining a 2-sphere somewhere in the space, we immediately arrive at a contradiction,

$$\begin{aligned} 0 &= \int_{\partial \mathbb{S}^2 = \emptyset} \vec{B} \cdot d\vec{l} = \oint_{\mathbb{S}^2} d\vec{S} \cdot \vec{\partial} \times \vec{B} = \mu_0 \oint_{\mathbb{S}^2} \vec{j}_e \cdot d\vec{S} \\ &= \mu_0 \int_{D^3} dV \left(- \frac{\partial \rho_e}{\partial t} \right) = -\mu_0 \frac{d}{dt} \int_{D^3} dV \rho_e \ , \end{aligned} \tag{1.3}$$

with the 3-ball D^3 enclosed by the 2-sphere $\mathbb{S}^2$, where for the last equality, we used the obvious fact that the net change in the amount of charge inside $\mathbb{S}^2$ should be accountable by the current flowing into and out of the sphere. The net amount of charge inside any finite volume need not be constant in general, as charges can easily flow in and out. Hence, this last equation (1.2) does not make sense in general.

Maxwell realized that something must correct this, and the solution is easy to see from

$$\frac{d}{dt} \left(\int dV \rho_e \right) = \frac{d}{dt} \int \epsilon_0 (\vec{\partial} \cdot \vec{E}) dV = \epsilon_0 \int \frac{\partial \vec{E}}{\partial t} \cdot d\vec{S} \ , \tag{1.4}$$

which requires

$$\vec{\partial} \times \vec{B} = \mu_0 \vec{j}_e + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} \ , \tag{1.5}$$

thus completing the Maxwell equation.

1.1.1 From Maxwell to Lorentz

Before we explore how this implies Special Relativity, let us make a different choice of units for electric charge, which in effect includes the electric permeability ϵ_0 in the definition of charge and current such that

$$\frac{1}{\epsilon_0}\rho_e \quad \rightarrow \quad \rho_e \ , \qquad \frac{1}{\epsilon_0}\vec{j}_e \quad \rightarrow \quad \vec{j}_e \ , \tag{1.6}$$

which leaves only the combination, $\epsilon_0\mu_0$, manifest in the Maxwell equation. This product is identified with the inverse square of the speed c of the electromagnetic wave, as we see in the following equation,

$$c^2 \equiv \frac{1}{\epsilon_0\mu_0} \ , \tag{1.7}$$

which was discovered back in the nineteenth century to be equal to the speed of light, hence the notation c.

Let us take one more step of renaming the magnetic field such that $c\vec{B} \rightarrow \vec{B}$. Combined, these two changes of convention bring the Maxwell equation to a more symmetric form,

$$\begin{aligned}
\vec{\partial}\cdot\vec{E} &= \rho_e \ , \\
\vec{\partial}\cdot\vec{B} &= 0 \ , \\
\vec{\partial}\times\vec{E} + \frac{1}{c}\frac{\partial\vec{B}}{\partial t} &= 0 \ , \\
\vec{\partial}\times\vec{B} - \frac{1}{c}\frac{\partial\vec{E}}{\partial t} &= \frac{1}{c}\vec{j}_e \ ,
\end{aligned} \tag{1.8}$$

which will be more useful later. This choice is called the Heaviside convention.*

In this unit, $\vec{E}$ has the same dimension as $\vec{B}$, such that the Lorentz force law has the form, in the Newtonian limit,

$$\frac{d}{dt}(m\mathbf{v}) = Q\left(\vec{E} + \frac{1}{c}\mathbf{v}\times\vec{B}\right) = q\left(c\vec{E} + \mathbf{v}\times\vec{B}\right) \ , \tag{1.9}$$

*After a further step of rescaling, $\rho_e, \vec{j}_e \ \rightarrow \ 4\pi\rho_e, 4\pi\vec{j}_e$, we obtain the so-called Gaussian unit or CGS unit.

where $\mathbf{v} = d\mathbf{x}/dt$ is the 3-velocity of the charged particle. In this book, we will use the arrow as in $\vec{E}$ and the boldfaced font as in $\mathbf{x}$ for quantities labeled by spatial indices. The unconventional factor of c in front of the last $\vec{E}$ should be noted, which is from another rescaling of Q to $q = Q/c$. Later, we will see how these equations change quite minimally in the relativistic regime. For the particle motion, only the left-hand side will be modified appropriately. More importantly, the structure of the Maxwell equation lifts to General Relativity more or less intact, and a combination of $\vec{E}$ and $\vec{B}$, to be later called F, serves as a prototype for the Riemann curvature.

The idea that light must be a form of electromagnetic oscillations came from combining the equations into

$$\left[\vec{\partial}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\vec{E} = 0 = \left[\vec{\partial}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\vec{B} \ , \tag{1.10}$$

when $\rho_e = 0$, $\vec{j}_e = 0$. Particular types of time-dependent solutions emerge

$$\vec{B},\ \vec{E} \simeq e^{-\mathbb{i}\omega t + \mathbb{i}\mathbf{k}\cdot\mathbf{x}} \ , \tag{1.11}$$

whose exponents must obey

$$\left(\mathbf{k}^2 - \frac{\omega^2}{c^2}\right)\vec{E} = 0 = \left(\mathbf{k}^2 - \frac{\omega^2}{c^2}\right)\vec{B} \tag{1.12}$$

from the combined Maxwell equation.

For plane waves, the wave vector $\mathbf{k}$ and frequency ω must satisfy the following relation:

$$c^2 \underbrace{\mathbf{k}^2}_{\left(\frac{2\pi}{\lambda}\right)^2} = \underbrace{\omega^2}_{\left(\frac{2\pi}{T}\right)^2} \ , \tag{1.13}$$

so that

$$c = \sqrt{\frac{\omega^2}{\mathbf{k}^2}} \tag{1.14}$$

is the speed of the electromagnetic waves, which is strangely enough embedded into the Maxwell equation itself rather than determined by environmental data. The fact that the deduced speed of the electromagnetic waves matched the speed of light c, measured rather precisely by the middle of the nineteenth century by Fizeau and by Foucault, pretty much demanded that light must be a type of electromagnetic wave; back then, light was the known physical phenomena with such an incredulous speed of propagation.

Early on, physicists began to worry about how the usual Galilean transformation law between a pair of inertial frames does not preserve the Maxwell equation. Under a Galilean transformation along x, with the Cartesian coordinates (x, y, z) for $\mathbf{x}$,

$$t' = t \ , \qquad x' = x - vt \tag{1.15}$$

leads to

$$\begin{pmatrix} dt' \\ dx' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \begin{pmatrix} dt \\ dx \end{pmatrix} \ ,$$

$$\begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix} \begin{pmatrix} dt' \\ dx' \end{pmatrix} = \begin{pmatrix} dt \\ dx \end{pmatrix} \ . \tag{1.16}$$

The derivatives transform as

$$\partial_t = \underbrace{\frac{\partial t'}{\partial t}}_{1} \partial_{t'} + \underbrace{\frac{\partial x'}{\partial t}}_{-v} \partial_{x'} \ ,$$

$$\partial_x = \underbrace{\frac{\partial t'}{\partial x}}_{0} \partial_{t'} + \underbrace{\frac{\partial x'}{\partial x}}_{1} \partial_{x'} \ . \tag{1.17}$$

So the differential operator that appears in a straightforward manner in the combined Maxwell equation transforms to a more complex format as shown in the following equation,

$$\begin{aligned} \partial_x^2 - \frac{1}{c^2}\partial_t^2 &= \partial_{x'}^2 - \frac{1}{c^2}(\partial_{t'} - v\partial_{x'})^2 \\ &= \left(1 - \frac{v^2}{c^2}\right)\partial_{x'}^2 + \left(\frac{2v}{c^2}\right)\partial_{t'}\partial_{x'} - \frac{1}{c^2}\partial_{t'}^2 \ . \end{aligned} \tag{1.18}$$

The Maxwell equation would appear different to observers in different inertial frames, which sounds rather counterintuitive. This conflict between the electromagnetism and the Galileo framework signaled the need for a new theory.

The first hint originated from Lorentz, later elevated to a more coherent form by Poincaré, who observed that revised transformation law

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{L} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} , \tag{1.19}$$

with $\beta \equiv v/c$ and $\gamma \equiv 1/\sqrt{1-\beta^2}$, preserves the Maxwell equation.

The transformation matrix obeys

$$\det L = 1 \cdot 1 \cdot \gamma^2 (1-\beta^2) = 1 , \tag{1.20}$$

and has the inverse,

$$L^{-1} = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.21}$$

A pair of coordinate systems related by such a transformation is illustrated in Figure 1.1. In particular, this is a special instance of more general L that preserves an inner product defined by η,

$$\eta \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = L^T \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} L . \tag{1.22}$$

These Lorentz transformations preserve the Maxwell equation.

It should be clear that the combined wave equation (1.10) from the Maxwell equation does not change,

$$-\frac{1}{c^2}\partial_t^2 + \partial_x^2 + \partial_y^2 + \partial_z^2 = -\frac{1}{c^2}\partial_{t'}^2 + \partial_{x'}^2 + \partial_{y'}^2 + \partial_{z'}^2 . \tag{1.23}$$

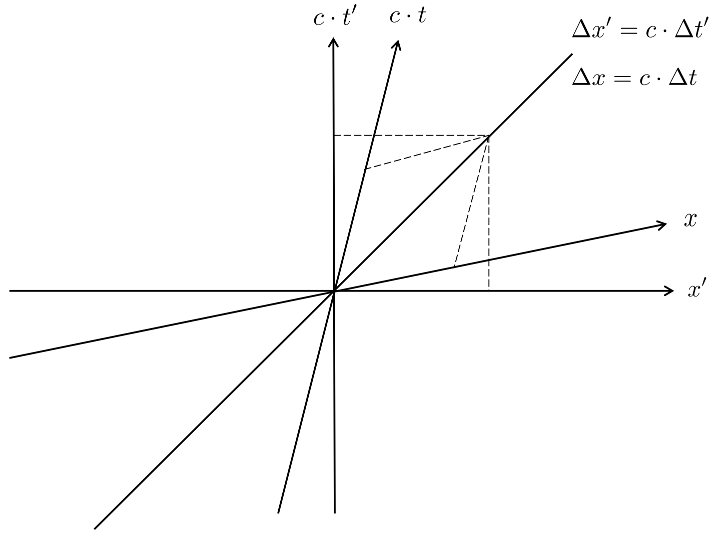

Figure 1.1: A pair of spacetime Cartesian coordinates related by a Lorentz transformation are shown. The diagonal line represents a trajectory of a massless object. A displacement along the latter can be represented by $(\Delta t, \Delta x)$ or by $(\Delta t', \Delta x')$, yet the relation between the temporal and the spatial displacements remains the same either way.

In relation to (1.23) is the following equality between infinitesimal displacements measured in the two different inertial frames,

$$-c^2dt^2 + dx^2 + dy^2 + dz^2 = -c^2dt'^2 + dx'^2 + dy'^2 + dz'^2 \ , \tag{1.24}$$

which eventually led to General Relativity. These invariants were first noted by Poincaré and also by Minkowski after Special Relativity was proposed. The combination of space and time, spanned by (t, x, y, z), is usually referred to as the Minkowski spacetime, while the sign choices here $(-+++)$ are collectively called the Lorentzian signature, as opposed to the more conventional choice $(++++)$ for the would-be purely spatial line element.

This invariance is reminiscent of the more familiar rotational $SO(3)$ group, consisting of matrices O such that

$$O^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} O = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \qquad \det O = 1 \ . \tag{1.25}$$

Collection of all L's such that

$$L^T \eta L = \eta \ , \qquad \det L = 1 \tag{1.26}$$

is the Lorentz group, sometimes denoted as $SO(1,3)$. It was Poincaré who first realized that these transformations form a group.

Let us try a pair of such transformations on the spacetime coordinates, $(x^\mu) = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$, with the Einstein convention that repeated indices are summed over

$$\begin{aligned} x'^\mu &= L^\mu{}_\alpha x^\alpha \ , \\ x''^\lambda &= \tilde{L}^\lambda{}_\mu x'^\mu = (\tilde{L} \cdot L)^\lambda{}_\alpha x^\alpha \ . \end{aligned} \tag{1.27}$$

Again, specializing to frames related by velocities along with x directions,

$$L = \underbrace{\begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\beta = \tanh\alpha} , \qquad \tilde{L} = \underbrace{\begin{pmatrix} \tilde{\gamma} & -\tilde{\gamma}\tilde{\beta} & 0 & 0 \\ -\tilde{\gamma}\tilde{\beta} & \tilde{\gamma} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\tilde{\beta} = \tanh\tilde{\alpha}} , \tag{1.28}$$

we find

$$\tilde{L}L = \begin{pmatrix} \cosh(\alpha + \tilde{\alpha}) & -\sinh(\alpha + \tilde{\alpha}) & 0 & 0 \\ -\sinh(\alpha + \tilde{\alpha}) & \cosh(\alpha + \tilde{\alpha}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{1.29}$$

and thus $[L(\alpha)]^{-1} = L(-\alpha)$. The additive quantity α is called the rapidity.

We will, from now on, combine the time coordinate ct and the spatial coordinate $\mathbf{x} = (x, y, z)$ into a single 4-vector notation x^μ, and sometimes write simply x, meaning the column vector made from these four coordinates. The 4-vector transforms as

$$(x')^\mu \ = L^\mu{}_\alpha x^\alpha \ , \tag{1.30}$$

which induces

$$\frac{\partial}{\partial x^\alpha} = \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial}{\partial x'^\mu} = L^\mu{}_\alpha \frac{\partial}{\partial x'^\mu} \ , \tag{1.31}$$

and the invariance of (1.10) follows immediately from general properties of L. As we move on to curved spacetime later, x^μ's could no longer be considered a vector. Coordinates should be viewed as a set of local functions.

1.1.2 Back to Maxwell

The electromagnetic potentials can be combined to a 4-vector. There are some normalization issues; one self-consistent choice nearest to the usual convention is to use ϕ and $\vec{A}$ such that

$$\begin{aligned} E_i &= \frac{1}{c}\left(-\partial_i \phi - \partial_t A_i\right) \ , \\ B_i &= \epsilon^{ijk}\partial_j A_k \ , \end{aligned} \tag{1.32}$$

where we introduce the unfamiliar factor of $1/c$ for $\vec{E}$ to fit the relativistic notation better.

Since we take $(ct, \mathbf{x}^i) \to (x^\mu)$ as a natural 4-vector, ϕ and A_i combine to a 4-vector as well:

$$(A_\mu) = (\underbrace{-\frac{\phi}{c}}_{A_0}, \vec{A}) \ , \tag{1.33}$$

and take an antisymmetric derivative as

$$\partial_\mu A_\nu - \partial_\nu A_\mu \ . \tag{1.34}$$

With the spatial and temporal indices distinguished, they are

$$\mu = 0, \nu = i : \quad \underbrace{\frac{1}{c}\partial_t A_i + \partial_i \frac{\phi}{c}}_{-E_i} \ ,$$

$$\mu = i, \nu = 0: \quad \underbrace{-\frac{1}{c}\partial_t A_i - \partial_i \frac{\phi}{c}}_{E_i} ,$$

$$\mu = i, \nu = j: \quad \partial_i A_j - \partial_j A_i = \epsilon_{ijk} B_k . \tag{1.35}$$

Combining, we find

$$[F_{\mu\nu}] \equiv [\partial_\mu A_\nu - \partial_\nu A_\mu] = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} . \tag{1.36}$$

The Maxwell equation is naturally written with F as

$$\eta^{\alpha\beta}\partial_\alpha F_{\mu\beta} = j_\mu / c ,$$

$$\epsilon^{\mu\alpha\beta\gamma}\partial_\alpha F_{\beta\gamma} = 0 . \tag{1.37}$$

The charge density and the current are combined into a 4-vector as

$$(j_\mu) \equiv (-\rho_e c, \vec{j}_e) . \tag{1.38}$$

The first gives

$$-\left(\partial_x E_x + \partial_y E_y + \partial_z E_z\right) = -\rho_e ,$$

$$-\frac{1}{c}\frac{\partial E_x}{\partial t} + \partial_y B_z - \partial_z B_y = \frac{1}{c} j_e^x ,$$

$$-\frac{1}{c}\frac{\partial E_y}{\partial t} - \partial_x B_z + \partial_z B_x = \frac{1}{c} j_e^y ,$$

$$-\frac{1}{c}\frac{\partial E_z}{\partial t} + \partial_x B_y - \partial_y B_x = \frac{1}{c} j_e^z , \tag{1.39}$$

whose first line is the Gauss law,

$$\vec{\partial} \cdot \vec{E} = \rho_e , \tag{1.40}$$

while the rest shows the last equation with the Maxwell term,

$$\vec{\partial} \times \vec{B} - \frac{1}{c}\frac{\partial \vec{E}}{\partial t} = \frac{1}{c}\vec{j}_e \ . \tag{1.41}$$

The other, involving no source, is a consequence of F being an antisymmetric derivative of A. For this, recall,

$$\epsilon^{0123} = 1, \quad \epsilon^{\mu\alpha\beta\gamma} = (-1)^n \begin{cases} n{=}0 \text{ for even permutation} \\ \\ n{=}1 \text{ for odd permutation} \end{cases} , \tag{1.42}$$

one of which gives

$$\begin{aligned} 0 &= \epsilon^{0ijk}(\partial_i F_{jk}) \\ &= 2(\partial_x F_{yz} + \partial_y F_{zx} + \partial_z F_{xy}) \\ &= 2(\partial_x B_x + \partial_y B_y + \partial_z B_z) \quad \rightarrow \quad 0 = \vec{\partial}\cdot\vec{B} \ , \end{aligned} \tag{1.43}$$

the magnetic counterpart of the Gauss law, while the rest presents Faraday's induction law as

$$\begin{aligned} \epsilon^{1\alpha\beta\gamma}\partial_\alpha F_{\beta\gamma} &= 2\left(\epsilon^{1023}\partial_0 F_{23} + \epsilon^{1230}\partial_2 F_{30} + \epsilon^{1302}\partial_3 F_{02}\right) \\ &= 2\left(-\frac{1}{c}\frac{\partial B_x}{\partial t} - \frac{\partial E_z}{\partial y} + \frac{\partial E_y}{\partial z}\right) \\ &= -2\left[\frac{1}{c}\frac{\partial B_x}{\partial t} + (\vec{\partial}\times\vec{E})_x\right] \ , \end{aligned} \tag{1.44}$$

etc., combining into Faraday's induction law,

$$\vec{\partial} \times \vec{E} + \frac{1}{c}\frac{\partial \vec{B}}{\partial t} = 0 \ . \tag{1.45}$$

Now, we are ready to see the full invariance of the Maxwell equation and also how $\vec{E}$ and $\vec{B}$ transform under L,

$$\partial_\alpha = L^\mu{}_\alpha \partial'_\mu = \partial'_\mu L^\mu{}_\alpha \qquad \Rightarrow \qquad F_{\alpha\beta} = F'_{\mu\nu} L^\mu{}_\alpha L^\nu{}_\beta \ . \tag{1.46}$$

From this it follows,

$$\eta^{\gamma\beta}\partial_\gamma F_{\alpha\beta} = j_\alpha/c \qquad \rightarrow \qquad \eta^{\mu\nu}\partial'_\mu F'_{\lambda\nu} = j'_\lambda/c \ . \tag{1.47}$$

The same happens with the other half,

$$\epsilon^{\mu\alpha\beta\gamma}\partial_\alpha F_{\beta\gamma} = 0 \qquad \rightarrow \qquad \epsilon^{\mu\alpha\beta\gamma}\partial'_\alpha F'_{\beta\gamma} = 0 \ , \tag{1.48}$$

thanks to

$$\epsilon' = (L \otimes L \otimes L \otimes L)\epsilon = \det(L)\,\epsilon = \epsilon \ . \tag{1.49}$$

The Maxwell equation is invariant under $x' = Lx$ and $A = A'L$, $j = j'L$. The mutually different transformations are easily encoded in the index locations: the upper for x^μ and, much like the partial derivatives, the lower for A_μ and j_μ.

Even though we freely used the vectorial notations $\vec{E}$ and $\vec{B}$ earlier, it should be clear that these are not vectors. In other words, one cannot make a 4-vector out of $\vec{E}$ and $\vec{B}$ by adding time-like components and the antisymmetric "tensor" $F_{\mu\nu}$ is the right underlying object. Nevertheless, the notations E and B shall make a few more appearances down the road.

1.2 Special Relativity

Special Relativity begins by taking the Lorentz transformation seriously, as the proper rule that relates a pair of inertial coordinates based on observers with two distinct velocities. The Lorentz transformation implies interesting and counterintuitive phenomena, such as the Lorentz–FitzGerald contraction, time dilation, and how the old notion of simultaneity becomes ambiguous. Once we accept the fact of life that Maxwell's electromagnetism governs our world, however, none of these should bother us. When we accept the incontrovertible fact that the speed of light looks the same to any inertial observer, none of these consequences should be more surprising.

Instead, we will merely plow ahead with mechanical consequences, eventually bringing us to the most fundamental feature, namely how the speed of light happens to be the limiting speed for our world. This should bring us to a way to understand fundamental forces and particles of nature, in particular gravity. Newton's gravity famously suffers from the conceptual problem of "the Action at a Distance" puzzle, and with the advance of Special Relativity, the problem went well beyond the philosophical issue. By understanding the mechanics of massive relativistic particles, we will come closer to how relativistic gravity should work, which will lead us to General Relativity.

1.2.1 Relativistic Motion

Free nonrelativistic particle extremizes the action,

$$S_{\rm NR} = \int dt\, \frac{1}{2} m\dot{\mathbf{x}}^2 \qquad \Rightarrow \qquad \frac{d}{dt}(m\dot{\mathbf{x}}) = 0 \ , \tag{1.50}$$

and gives us conserved momentum $\mathbf{p} \equiv m\dot{\mathbf{x}}$ and the conserved energy $\mathcal{E} \equiv m\,\dot{\mathbf{x}}^2/2$,

$$\frac{d}{dt}\left(\frac{1}{2} m\,\dot{\mathbf{x}}^2\right) = m\,\ddot{\mathbf{x}}\cdot\dot{\mathbf{x}} = 0 \ . \tag{1.51}$$

How do we change this to a relativistic version?

One cannot use $S_{\rm NR}$ which is clearly not invariant under the Lorentz transformation, L, but a more immediate problem is what to do with the time coordinate. Since t is now on equal footing as $\mathbf{x}$, let us invent a new time parameter $\mathbf{s}$, with which the following looks a bit more natural:

$$\int d\mathbf{s}\, \frac{m}{2}\left[\left(\frac{d\mathbf{x}}{d\mathbf{s}}\right)^2 - c^2\left(\frac{dt}{d\mathbf{s}}\right)^2\right] \ , \tag{1.52}$$

as it is invariant under L.

As it is, this action looks a little odd with its dependence on the choice of a clock "$\mathbf{s}$." Unlike in Newtonian physics, a universal and absolute notion of time no longer exists, suggesting that the choice of the time parameter along a trajectory has no preferred choice. If we used a different clock $\mathbf{s}' = \mathbf{s}'(\mathbf{s})$, the action changes to

$$\int d\mathbf{s}'\left(\frac{d\mathbf{s}'}{d\mathbf{s}}\right)\frac{m}{2}\left[\left(\frac{d\mathbf{x}}{d\mathbf{s}'}\right)^2 - c^2\left(\frac{dt}{d\mathbf{s}'}\right)^2\right] \ , \tag{1.53}$$

which carries an explicit "time" dependence. We will, in fact, later come back to an action that looks remarkably similar to this with a well-motivated physical choice of $\mathbf{s}$. With this in mind, we first wish to find an action that is invariant under any monotonic reparameterization of $\mathbf{s}$.

A simple choice, invariant under $\mathbf{s} \to \mathbf{s}'$ as long as $d\mathbf{s}'/d\mathbf{s} > 0$, is

$$\begin{aligned} S &\equiv -mc \int d\mathbf{s} \sqrt{c^2 \left(\frac{dt}{d\mathbf{s}}\right)^2 - \left(\frac{d\mathbf{x}}{d\mathbf{s}}\right)^2} \\ &\to -mc \int d\mathbf{s}' \sqrt{c^2 \left(\frac{dt}{d\mathbf{s}'}\right)^2 - \left(\frac{d\mathbf{x}}{d\mathbf{s}'}\right)^2} \,. \end{aligned} \tag{1.54}$$

Another physical requirement for the action is that it should reduce to the free non-relativistic one; we wish to match into $\int \frac{1}{2} m\mathbf{v}^2 \, dt$ as $|\mathbf{v}| \ll c$ where $\mathbf{s} = t$ is a natural choice from the Newtonian perspective by expanding S,

$$S = -mc^2 \int dt \sqrt{1 - \left(\frac{1}{c}\frac{d\mathbf{x}}{dt}\right)^2} \simeq mc^2 \int dt \left(-1 + \frac{1}{2}\frac{1}{c^2}\left(\frac{d\mathbf{x}}{dt}\right)^2\right) \,, \tag{1.55}$$

in small $\mathbf{v}^2/c^2$. The first nontrivial part indeed gives us the Newtonian kinetic term.

Starting from this relativistic action, we again find conserved momenta,

$$\begin{aligned} p_0 = \frac{\delta S}{\delta(c\dot{t})} &= -mc \frac{c\dot{t}}{\sqrt{c^2\dot{t}^2 - \dot{\mathbf{x}}^2}} \\ &= -\frac{mc}{\sqrt{1 - \mathbf{v}^2/c^2}} \end{aligned} \tag{1.56}$$

for the time direction and

$$\begin{aligned} \mathbf{p}_i = \frac{\delta S}{\delta \dot{\mathbf{x}}^i} &= -mc \frac{(-\dot{\mathbf{x}}^i)}{\sqrt{c^2\dot{t}^2 - \dot{\mathbf{x}}^2}} \\ &= \frac{m\mathbf{v}^i}{\sqrt{1 - \mathbf{v}^2/c^2}} \end{aligned} \tag{1.57}$$

for spatial directions.

Together, they give us conserved 4-momenta,

$$(p^0, \mathbf{p}) = (-p_0, \mathbf{p}) = \underbrace{\left(\frac{mc}{\sqrt{1 - \mathbf{v}^2/c^2}}, \frac{m\mathbf{v}}{\sqrt{1 - \mathbf{v}^2/c^2}}\right)}_{\text{4-momenta}} \,, \tag{1.58}$$

which obeys

$$p^\mu p_\mu = -p_0^2 + \mathbf{p}_i \mathbf{p}_i = -m^2 c^2 \ . \tag{1.59}$$

The time component gives what we conventionally think of as the energy,

$$\mathcal{E} = -\frac{\delta S}{\delta t} = -c p_0 = \frac{mc^2}{\sqrt{1 - \mathbf{v}^2/c^2}} \ . \tag{1.60}$$

We may write

$$\mathcal{E}^2 - \mathbf{p}^2 c^2 = (mc^2)^2 \tag{1.61}$$

alternatively.

The built-in Lorentz invariance implies that

$$c^2 dt^2 - d\mathbf{x}^2 \tag{1.62}$$

is an invariant quantity. What is the physical meaning of this quantity? Note that

$$c^2 dt^2 - d\mathbf{x}^2 = c^2 dt^2 \left(1 - \frac{\mathbf{v}^2}{c^2} \right) \ . \tag{1.63}$$

If we switch to a coordinate system $(t', \mathbf{x}')$ where the particle is at rest, $d\mathbf{x}'/d\mathbf{s} = 0$,

$$c^2 dt^2 - d\mathbf{x}^2 \quad \rightarrow \quad c^2 (dt')^2 \ . \tag{1.64}$$

So the expression (1.62) essentially measures the square of the infinitesimal time-lapse per unit interval, multiplied by c^2, in the inertial frame tied to the particle in motion. This quantity corresponds to the time measured by a co-moving clock, so we call it the “proper” time, which we will denote by $\boldsymbol{\tau}$.

Therefore, the action S is $-mc^2$ times the proper time accumulated by the particle,

$$S = -mc^2 \int d\boldsymbol{\tau} \ . \tag{1.65}$$

For massive physical objects and in the absence of forces, the actual path achieves the longest possible proper time, given the spacetime’s initial and final points. Note that (1.62) is positive for massive particles and vanishes identically for objects moving at the speed of light. Such trajectories are, respectively, time-like, and light-like. The latter is said to be null. Space-like trajectories, for which the value of (1.62) is negative, are deemed impossible for physical objects.

1.2.2 Proper Time and Twin "Paradox"

In the famous tale of the twin paradox, one brother travels in a rocket, while the other remains on the Earth. The purported paradox relies on the idea that all inertial frames are equivalent, so all objects moving at constant velocity have an equal right to use their own clock and would perceive others moving at different constant velocities to age less rapidly. The puzzles presumes that each would see the other speeding away and then coming back, and envy that his brother is aging less. Yet, at the end of the journey when the first brother came back to the Earth, only one of them must be older. See Figure 1.2.

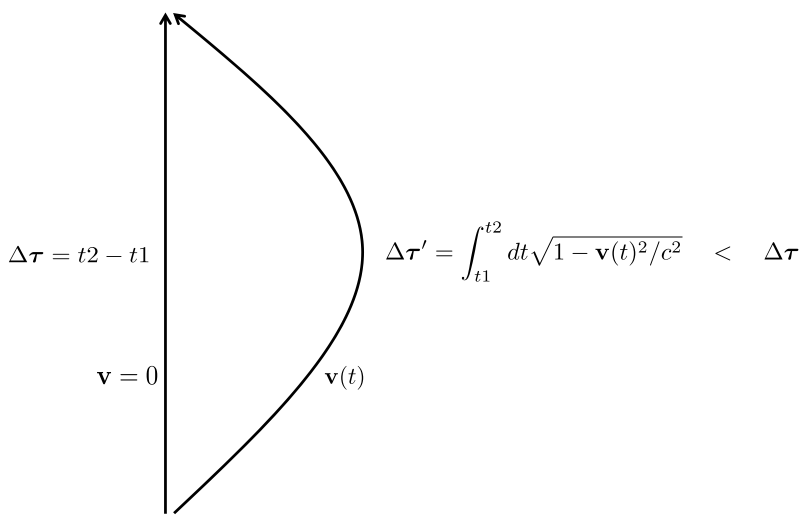

Figure 1.2: An illustration of how the Twin Paradox was never a paradox. The time elapse felt by an observer should be measured by the proper time, not by the coordinate times of anyone's inertial frames.

The popular resolution points out that the brother in the rocket has to experience accelerations and decelerations inevitably. Hence, he was not an inertial observer at times, and this does invalidate the argument based on the equivalence of all inertial observers. This explanation may not sound entirely satisfactory. For a very long journey by one of the twins and with a powerful-enough rocket, one can imagine that acceleration and deceleration may be needed in a relatively short time period. Are we suggesting that something drastic happens during those noninertial periods to the traveling brother?

The answer to the last question is indeed "yes." Regardless of how short the periods of acceleration and deceleration might be, they throw off the naive comparison of coordinate times spent by each brother. We will not go into details here, which one can find such easily from common sources like Wikipedia. However, although not incorrect, these explanations based on accelerations and decelerations do not really address the primary confusion of this story.

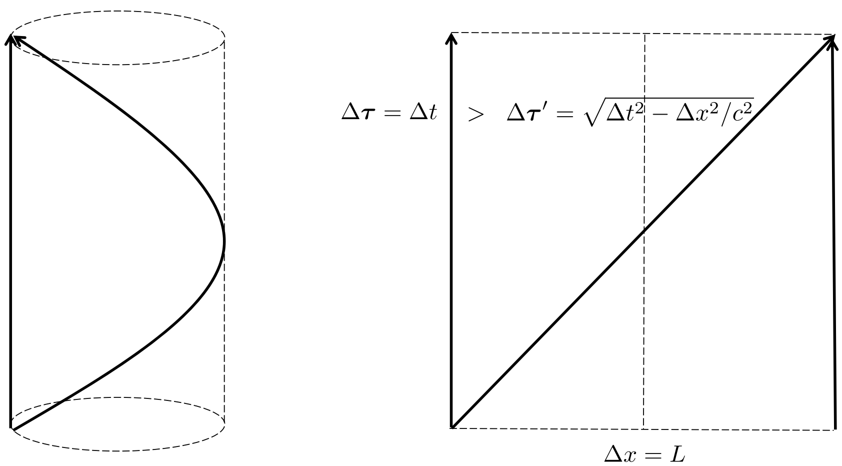

Figure 1.3: A twin "paradox" is recast in a cylinder-like universe whose x direction is periodic. Both are in inertial motions yet experience time differently, so that an acceleration by one of them is never an issue. The reason why one of them ages less should be traced to the inherent geometry itself, or more precisely the "metric" assigned to the spacetime which in turn instructs us how to compute the proper time.

A simple example that illustrates this can be found by imagining a cylinder-like universe, as illustrated in Figure 1.3; the universe is flat, yet its x direction is of a finite length circle by identifying x and $x + L$. In such a universe, the one on the rocket can move at a constant velocity V along the x direction and meet his more stationary brother periodically, without any acceleration or deceleration. Both would happily and justifiably think themselves inertial and stationary, so would both envy the other for aging less? This obvious contradiction cannot hold, so a resolution of the puzzle should have little to do with "inertial" or "noninertial" motions.

The nonsense of the twin paradox came from confusing the roles of coordinate time t and the proper time $\boldsymbol{\tau}$. A person's aging should have nothing to do with the coordinate time choice and depends only on his own proper time. The traveling brother will find the other on the Earth aging faster, simply because after each period Δt between the encounters, his accumulated proper time is the less of the two. Between encounters, the proper time accumulated by the traveling brother is

$$\boldsymbol{\tau}_{\text{traveling}} = \Delta t\sqrt{1 - V^2/c^2} \;, \tag{1.66}$$

while the other one on the Earth must suffer the longer $\boldsymbol{\tau}_{\text{Earth}} = \Delta t = L/V$.

The proper time is designed to be independent of the spacetime coordinate choice, that is, invariant under the Lorentz transformations, so it is not affected by choices of the inertial frame where $\boldsymbol{\tau}$ is computed. The twin "paradox" resulted simply because of easy mistakes committed while translating back and forth between the proper time $\boldsymbol{\tau}$ and the coordinate time t's of various inertial coordinate systems; it is just that if you insist on doing things in a complicated way, against a much simpler alternative, you would be more prone to confusion. A better question the twins could have asked here is why the proper time should take this particular form that favors one brother over the other, but clearly this has little to do with coordinate choices, inertial or not.

1.2.3 Relativistic Motion under Electromagnetic Fields

What if the particle is charged? We start from

$$S = -mc\int d\mathbf{s}\sqrt{c^2\dot{t}^2 - \dot{\mathbf{x}}^2} + q\int \underbrace{\dot{x}^\mu A_\mu}_{=\,(-\dot{t}\phi+\dot{\mathbf{x}}^i A_i)} d\mathbf{s} \tag{1.67}$$

with the electric charge q and the spatial momenta,

$$\mathbf{p}_i = \frac{\delta S}{\delta \dot{\mathbf{x}}^i} = mc\frac{\dot{\mathbf{x}}^i}{\sqrt{c^2\dot{t}^2 - \dot{\mathbf{x}}^2}} + qA_i \;, \tag{1.68}$$

which obeys

$$\frac{d}{d\mathbf{s}}\mathbf{p}_i = \frac{\delta S}{\delta \mathbf{x}^i} = q\left(-\dot{t}\;\partial_i\phi + \dot{\mathbf{x}}^k\partial_i A_k\right) \;, \tag{1.69}$$

or

$$\frac{d}{d\mathbf{s}}\left(mc\frac{\dot{\mathbf{x}}^i}{\sqrt{c^2\dot{t}^2-\dot{\mathbf{x}}^2}}+qA_i\right)=q\left(-\dot{t}\partial_i\phi+\dot{\mathbf{x}}^k\partial_iA_k\right)\ . \tag{1.70}$$

Moving all gauge fields to the right-hand side, and reviving the more conventional velocity $\mathbf{v}^i = \dot{\mathbf{x}}^i/\dot{t}$,

$$\begin{aligned}\frac{d}{d\mathbf{s}}\left(\frac{m\dot{\mathbf{x}}^i}{\sqrt{\dot{t}^2-\dot{\mathbf{x}}^2/c^2}}\right) &= q\left(-\dot{t}\partial_i\phi-\dot{t}\partial_tA_i\right)+q\left(\dot{\mathbf{x}}^k\underbrace{(\partial_iA_k-\partial_kA_i)}_{=\epsilon_{ikl}B_l}\right)\\ &= q\,\dot{t}(-\partial_i\phi-\partial_tA_i)+q\,\dot{t}(\mathbf{v}^k\epsilon_{ikl}B_l)\ ,\end{aligned} \tag{1.71}$$

or

$$\frac{d}{d\mathbf{s}}\left(\frac{m\mathbf{v}^i}{\sqrt{1-\mathbf{v}^2/c^2}}\right)=q\,\dot{t}\left(cE^i+(\mathbf{v}\times B)^i\right)\ , \tag{1.72}$$

which is the same as

$$\frac{d}{dt}\left(\frac{m\mathbf{v}^i}{\sqrt{1-\mathbf{v}^2/c^2}}\right)=q\ \left(cE^i+(\mathbf{v}\times B)^i\right)\ . \tag{1.73}$$

Therefore, a relativistic particle obeys the same Lorentz force law, except $m\mathbf{v}^i \rightarrow m\gamma\mathbf{v}^i$.

What about the time part? The conjugate momentum to t is nothing but the generalization of the Newtonian mechanical energy, modulo a sign,

$$\begin{aligned}p_t=\frac{\delta S}{\delta\dot{t}} &= -mc^3\frac{\dot{t}}{\sqrt{c^2\dot{t}^2-\dot{\mathbf{x}}^2}}-q\,\phi\\ &= -mc^2\gamma-q\,\phi\ ,\end{aligned} \tag{1.74}$$

with the rest mass contribution included. With

$$\frac{\delta S}{\delta t} = q\left(-\dot{t}\,\partial_t\phi+\dot{\mathbf{x}}^i\partial_tA_i\right)\ , \tag{1.75}$$

we have

$$\frac{d}{d\mathbf{s}}\, p_t \;=\; \frac{d}{d\mathbf{s}}\left(-mc^2\gamma\right) - q\,\left(\dot{t}\,\partial_t\phi + \dot{\mathbf{x}}^i\,\partial_i\phi\right) = q\left(-\dot{t}\,\partial_t\phi + \dot{\mathbf{x}}^i\,\partial_t A_i\right)\,, \tag{1.76}$$

or

$$\frac{d}{d\mathbf{s}}\underbrace{(mc^2\gamma)}_{\text{particle energy }\mathcal{E}} \;=\; q\,\dot{\mathbf{x}}^i\underbrace{(-\partial_i\phi - \partial_t A_i)}_{cE_i} \;=\; \underbrace{q\,\dot{\mathbf{x}}^i(cE_i)}_{\text{work done by }\vec{E}}\quad. \tag{1.77}$$

This reproduces the familiar energy conservation law, under the influence of electromagnetic force, except that the particle's energy $\mathcal{E}$ is $mc^2\gamma$ rather than $\frac{1}{2}m\mathbf{v}^2$.