The first time I experienced beauty in physics was when I learned how Einstein's special relativity is hidden in the equations of Maxwell's theory of electricity and magnetism. Today, it is my privilege to show this to you. In the end, we will speculate how gravity might fit into this.

4.1 Relativity requires Magnetism

Imagine you had never heard of magnetism. But you happen to know electrostatics and relativity. It is then possible to show that there has to be such a thing as magnetism.

Consider a string of positive charges moving to the right with velocity $v$ and negative charges moving to the left with velocity $-v$:

At a distance $r$, there is a point charge $q$ travelling to the right at speed $u < v$. The charges are close enough together so that we may regard them as continuous line charges with densities $\pm \lambda$ (= charge/length). The net current to the right is

$$I = 2\lambda v.$$  

Because the positive and negative line charges cancel, there is no electrical force on $q$ (in the rest frame of the wire).

Now consider the same situation from the point of view of an observer comoving with the charge $q$ (i.e. an observer travelling to the right with speed $u$):

The velocity of the positive charges is now smaller than the velocity of the negative charges. The Lorentz contraction of the spacing between negative charges is more severe than that between positive charges; therefore, the wire, in this frame, carries a net negative charge! There is now an electric force on the charge. In the box below I show you how to compute this force in terms
of the charge density in the original frame. Transforming this force back to the original frame leads to

\[ F = -quB , \quad \text{where} \quad B = \frac{\mu_0 I}{2\pi r} . \quad (4.1.2) \]

We have derived the Lorentz force using only electrostatics and a sequence of relativistic transformations. I encourage you to look at the details of the calculation. It is neat.

Let us call the original frame \( S \) and the new frame \( S' \). By the Einstein velocity addition rule, the velocities of the positive and negative charges in \( S' \) are

\[ v_{\pm} = \frac{u \mp u}{\sqrt{1 \mp v^2/c^2}} . \quad (4.1.3) \]

If \( \lambda_0 \) is the charge density of the positive line charge in its own rest frame, then

\[ \lambda_{\pm} = \lambda_0 (\gamma_{\pm} - \gamma_{-}) = -2uv \frac{\lambda}{c^2} \frac{1}{\sqrt{1 - u^2/c^2}} . \quad (4.1.4) \]

Of course, \( \lambda_0 \) is not the same as \( \lambda \), but

\[ \lambda = \gamma \lambda_0 , \quad \text{where} \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}} . \quad (4.1.5) \]

With a bit of algebra, you can show that

\[ \gamma_{\pm} = \frac{1 \mp uv/c^2}{\sqrt{1 - u^2/c^2}} \times \gamma . \quad (4.1.6) \]

The net line charge in \( S' \) is

\[ \lambda_{\text{tot}} = \lambda_{+} + \lambda_{-} = \lambda_0 (\gamma_{+} - \gamma_{-}) = -2uv \frac{\lambda}{c^2} \frac{1}{\sqrt{1 - u^2/c^2}} . \quad (4.1.7) \]

This creates an electric field

\[ E' = \frac{1}{\varepsilon_0} \frac{\lambda_{\text{tot}}}{2\pi r} . \quad (4.1.8) \]

Hence, in the frame \( S' \) there is an electrical force on \( q \):

\[ F' = qE' = -\frac{qu}{\sqrt{1 - u^2/c^2}} \times \frac{2\lambda v}{2\pi r} \times \frac{1}{\varepsilon_0 c^2} . \quad (4.1.9) \]

But if there is a force on \( q \) in \( S' \), there must be one in \( S \). In the relativity course last term, you learned how to transform forces between the frames. Since \( q \) is at rest in \( S' \), and \( F' \) is perpendicular to \( u \), the force in \( S \) is given by

\[ F = \sqrt{1 - u^2/c^2} F' = -qu \times \frac{2\lambda v}{2\pi r} \times \frac{1}{\varepsilon_0 c^2} . \quad (4.1.10) \]

The charge is attracted toward the wire by a force that is purely electrical in \( S' \) (where the wire is charged, and \( q \) is at rest), but distinctly non-electrical in \( S \) (where the wire is neutral). Taken together, then, electrostatics and relativity imply the existence of another force. This other force is, of course, the magnetic force. Expressing \( \lambda v \) in terms of the current \( (4.1.1) \), we get

\[ F = -qu \left( \frac{\mu_0 I}{2\pi r} \right) , \quad (4.1.11) \]

where we have defined \( \mu_0 \equiv (\varepsilon_0 c^2)^{-1} \). The term in parentheses is the magnetic field of a long, straight wire, and the force is precisely what we would have obtained by using the Lorentz force law in \( S \),

\[ F = -quB , \quad \text{where} \quad B \equiv \frac{\mu_0 I}{2\pi r} . \quad (4.1.12) \]

We have derived the magnetic force between a current-carrying wire and a moving charge without ever invoking the laws of magnetism.
4. Electrodynamics and Relativity

4.2 Magnetism requires Relativity

4.2.1 Maxwell’s Equations

All of electricity and magnetism is contained in the four Maxwell equations:

\[ \nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0}, \] (M1)

\[ \nabla \cdot \vec{B} = 0, \] (M2)

\[ \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \] (M3)

\[ \nabla \times \vec{B} = \mu_0 \vec{J} + \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}. \] (M4)

where \( \rho \) and \( \vec{J} \) are the total charge and current densities, which satisfy the continuity equation

\[ \frac{\partial \rho}{\partial t} = -\nabla \cdot \vec{J}. \] (C)

The parameters \( \varepsilon_0 \) and \( \mu_0 \) are simply constants that convert units.

I am assuming that you have seen these before\(^1\), but let me remind you briefly of the meaning of each of these equations:

- **Gauss’ Law**
  
  (M1) simply states that “electric charges produce electric fields”. Specifically, the divergence of the electric field \( \vec{E} \) is determined by the charge density \( \rho \) (= charge/volume).

- **No Monopoles**
  
  Unlike electric charges, we have never observed isolated magnetic charges (“monopoles”). Magnetic charges always come in pairs of plus and minus (or North and South). (If you cut a magnet into two pieces, you get two new magnets, each with their own North and South poles.) The net magnetic charge density is therefore zero. (M2) embodies this fact, by having zero in place of \( \rho \) in the equation for the divergence of the magnetic field \( \vec{B} \).

- **Faraday’s Law of Induction**
  
  (M3) describes how a time-varying magnetic field creates (“induces”) an electric field. This is easily demonstrated by moving a magnet through a loop of wire. This induces an electric field, which forces a current through the wire, which might power an attached light bulb.

- **Ampère’s Law**
  
  It was known well before Maxwell that an electric current produces a magnetic field—e.g. Oersted discovered that a magnet feels a force when placed near a wire with a current flowing through it. However, the equation that people used to describe this, \( \nabla \times \vec{B} = \mu_0 \vec{J} \), was wrong. Maxwell pointed out that the original form of Ampère’s law was inconsistent with the conservation of electric charge, eq. (C), (do you see why?). To fix this he added an extra term to (M4). This extra term implies that a time-varying electric field also produces a magnetic field. This led to one of the most important discoveries in the history of physics.

\(^1\)See Example Sheet 3 of your course on ‘Vector Calculus’.
Conservation of Charge

Charges and currents source electric and magnetic fields. But, a current is nothing but moving charges, so it is natural to expect that $\rho$ and $\vec{J}$ are related to each other. This relation is the continuity equation (C). Consider a small volume $V$ (bounded by a surface $S$) containing a total charge $Q$.

The amount of charge moving out of this volume per unit time equals the current flowing through the surface:

$$\frac{\partial}{\partial t} Q = - \int dS \cdot \vec{J}.$$ (4.2.13)

Eq. (C) embodies this locally (i.e. for every point in space),

$$\frac{\partial}{\partial t} Q = \frac{\partial}{\partial t} \int dV \rho = - \int dV \nabla \cdot \vec{J} = - \int dS \cdot \vec{J},$$ (4.2.14)

where we used Stokes’ theorem in the last equality.

Maxwell’s equations are ugly! As we will see, this is because space and time are treated separately, while we now know that Nature is symmetric in time and space. By uncovering this hidden symmetry we will be able to write the four Maxwell equations in terms of a single elegant equation. In the process, we will develop a unified relativistic description of electricity and magnetism.

4.2.2 Let There Be Light!

From (M3) we see that a time-dependent magnetic field, $\vec{B}(t)$, produces an electric field. Similarly, (M4) implies that a time-dependent electric field, $\vec{E}(t)$, creates a magnetic field. So we see that “a changing magnetic field produces an electric field, which produces a magnetic field, which produces an electric field, which ...”. Once we set up a time-dependent electric or magnetic field, it seems to allow for self-sustained solutions that oscillate between electric and magnetic. Since the oscillating fields also propagate through space, we call them electromagnetic waves.

Let us describe these electromagnetic waves more mathematically. We restrict to empty space (i.e. a perfect vacuum with no charges and currents, $\rho = \vec{J} = 0$). The Maxwell’s equations then are

$$\nabla \cdot \vec{E} = 0,$$ (M1’)

$$\nabla \cdot \vec{B} = 0,$$ (M2’)

$$\nabla \times \vec{E} = - \frac{\partial \vec{B}}{\partial t},$$ (M3’)

$$\nabla \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}.$$ (M4’
Now take the curl of the curl equation for the $\vec{E}$-field, i.e. $\nabla \times (\nabla \times \vec{E})$,

$$\nabla \times (\nabla \times \vec{E}) = -\frac{\partial}{\partial t} \nabla \times \vec{B} = -\mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2},$$

(4.2.15)

where we have used $(M4')$ in the final equality. We use a vector identity to manipulate the l.h.s. of this expression,

$$\nabla \times (\nabla \times \vec{E}) = \vec{\nabla} (\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E}$$

(4.2.16)

$$= -\nabla^2 \vec{E},$$

(4.2.17)

where we have used $(M1')$ in the final equality. We get

$$\mu_0 \varepsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2 \vec{E} = 0.$$

(4.2.18)

This partial differential equation is the wave equation for the $\vec{E}$-field. Try substituting

$$\vec{E}(t, x) = \vec{E}_0 \sin \left[ 2\pi \left( \frac{t}{T} - \frac{x}{\lambda} \right) \right],$$

(4.2.20)

where $T$ and $\lambda$ are constants. You will find that this is a solution of (4.2.18), if

$$v \equiv \frac{\lambda}{T} = \frac{1}{\sqrt{\mu_0 \varepsilon_0}},$$

(4.2.21)

where $v$ is the speed of the wave.

It is easy to see that (4.2.20) indeed describes a wave:

First, consider a snapshot of the solution at a fixed time $t = t_0$. The electric field $\vec{E}(t_0, x)$ varies periodically through space:

The solution repeats every distance $\lambda$ along the $x$-axis.

Next, imagine sitting at a fixed point $x = x_0$. The electric field at that point, $\vec{E}(t, x_0)$, oscillates in time:

The solution repeats with a time period $T$. Hence, (4.2.20) indeed describes a wave propagating along the $x$-axis with speed $v = \lambda/T$ given by (4.2.21).

As boring as eq. (4.2.21) looks, it is an absolutely remarkable result! Through the genius of Einstein it led to the special theory of relativity.

---

2Similar manipulations for the curl equation for the $\vec{B}$-field, eq. $(M4')$, give the wave equation for the $\vec{B}$-field

$$\mu_0 \varepsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} - \nabla^2 \vec{B} = 0.$$

(4.2.19)
4.2.3 Racing A Light Beam

Playing with coils and metal plates, you would find that \( \varepsilon_0 = 8.85 \times 10^{-12} \text{C}^2/\text{Nm}^2 \) and \( \mu_0 = 4\pi \times 10^{-7} \text{N/A}^2 \). Eq. (4.2.21) then predicts that all electromagnetic waves propagate with

\[
c \equiv \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 3 \times 10^8 \text{m/s}.
\]  (4.2.22)

But, this is just the speed of light. Electromagnetic waves are light!

Eq. (4.2.22) looks innocent, but notice that I never mentioned the speed of the observer. According to Maxwell’s equation,

\[
\text{the speed of light is independent of the motion of the observer.}
\]

This flies in the face of Newton’s law of velocity addition. The light emitted from a moving spaceship travels at exactly the same speed as the light from a stationary source. This means that you can never catch up with a light ray. No matter how fast you run after a light ray, it will always recede from you at speed \( c \). It would have been easy to dismiss this craziness as a flaw of Maxwell’s theory. However, Einstein dared to accept this strange feature of light as a fundamental principle of Nature and built his theory of relativity around it.

4.2.4 My Time Is Your Space

To give up Newton’s simple law of velocity addition, means to give up on absolute measurements of time and space. In other words, in order for two observers \( O \) and \( O' \) in motion relative to each other to agree on the speed of light, they have to disagree on measurements of time (\( \Delta t \) vs. \( \Delta t' \)) and space (\( \Delta x \) vs. \( \Delta x' \)). However, that disagreement is of a very specific kind, such that everybody agrees on the value of the speed of light

\[
c = \frac{\Delta x}{\Delta t} = \frac{\Delta x'}{\Delta t'}.
\]  (4.2.23)

Eq. (4.2.23) can be rewritten as follows

\[
c^2 \Delta t'^2 - \Delta x'^2 = c^2(\Delta t')^2 - (\Delta x')^2 = 0.
\]  (4.2.24)

In fact, even for objects not moving at the speed of light, observers will disagree on \( \Delta x \) and \( \Delta t \), but will always agree on the combination

\[
\Delta s^2 \equiv c^2 \Delta t^2 - \Delta x^2 = c^2(\Delta t')^2 - (\Delta x')^2.
\]  (4.2.25)

More important than the fact that space and time are relative, is the fact that space and time are related by a special symmetry that guarantees the invariance of \( \Delta s^2 \). This symmetry is called Lorentz symmetry.

As you know, an elegant way to make this symmetry manifest is to combine space and time coordinates into a single four-dimensional spacetime coordinate

\[
t, \vec{x} \implies X^\mu = \begin{pmatrix} \text{time} \\ \vec{x} \end{pmatrix}.
\]  (4.2.26)
**4. Electrodynamics and Relativity**

The “square” of the spacetime four-vector is

\[ s^2 = \eta_{\mu \nu} X^\mu X^\nu, \quad (4.2.27) \]

where I have introduced the *Minkowski metric*,

\[ \eta_{\mu \nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (4.2.28) \]

Eq. (4.2.27) uses the Einstein summation convention, so that repeated indices are summed over. You should remember that the coordinates of different observers are related by a *Lorentz transformation*

\[ X'_\mu = \Lambda^\nu_\mu (v) X^\nu. \quad (4.2.29) \]

You can think of these as “rotations” in the 4d spacetime, that leave \( s^2 = X^\mu X_\mu \) the same. This is analogous to 3d spatial rotations

\[ x'_i = R^i_j (\theta) x_j, \quad (4.2.30) \]

that leave the magnitude of the position 3-vector, \( \ell^2 = x^i x_i \), the same.

**4.3 Relativity unifies Electricity and Magnetism**

I promised you that the Maxwell equations become beautiful when written in a relativistic way. Let’s see how that works.

### 4.3.1 Relativistic Electrodynamics

First, consider Gauss’ law for the magnetic field, eq. (M2),

\[ \nabla \cdot \vec{B} = 0. \quad (4.3.31) \]

Since “div curl = 0”, this is solved automatically if \( \vec{B} \) is written as the curl of a vector field \( \vec{A} \),

\[ \vec{B} = \nabla \times \vec{A}. \quad (4.3.32) \]

Next, we use eq. (4.3.32), to rewrite Faraday’s law for the electric field, eq. (M3), \( \nabla \times \vec{E} = -\dot{\vec{B}} \), as

\[ \nabla \times [\vec{E} + \dot{\vec{A}}] = 0. \quad (4.3.33) \]

Since “curl grad = 0”, we can write \( \vec{E} + \dot{\vec{A}} \) as the gradient of a scalar function. The Maxwell equation (M3) is hence solved if \( \vec{E} \) is written as

\[ \vec{E} = -\nabla \phi - \dot{\vec{A}}. \quad (4.3.34) \]
The choice of vector field $\vec{A}$ is not unique. We can add the gradient of any scalar function $\alpha(t, \vec{x})$ to $\vec{A}$ without changing the magnetic field $\vec{B} = \vec{\nabla} \times \vec{A}$ (since “curl grad = 0”),

$$\vec{A} \rightarrow \vec{A}' = \vec{A} + \vec{\nabla} \alpha .$$

(4.3.35)

The electric field $\vec{E} = -\vec{\nabla} \phi - \vec{A}$ also remains unchanged under the transformation (4.3.35), if $\phi$ simultaneously transforms as

$$\phi \rightarrow \phi' = \phi - \dot{\alpha} .$$

(4.3.36)

We hence have the freedom to pick any scalar $\alpha$ without changing $\vec{E}$ and $\vec{B}$. This is equivalent to the freedom of fixing $\vec{\nabla} \cdot \vec{A}$, since

$$\vec{\nabla} \cdot \vec{A}' = \vec{\nabla} \cdot \vec{A} + \nabla^2 \alpha .$$

(4.3.37)

A particularly useful choice is the so-called Lorenz gauge

$$\vec{\nabla} \cdot \vec{A} = \frac{\dot{\phi}}{c^2} .$$

(4.3.38)

We have shown that the two Maxwell equations (M2) and (M3) are automatically satisfied if the two vector fields $\vec{E}$ and $\vec{B}$ are expressed in terms of a scalar field $\phi$ and a new vector field $\vec{A}$. But, a scalar and a vector can be combined into a four-vector. Just like $t$ and $\vec{x}$ were combined into $X^\mu = (ct, \vec{x})$. Let’s try and define the vector potential

$$\phi, \vec{A} \Rightarrow A^\mu = \begin{pmatrix} \phi/c \\ \vec{A} \end{pmatrix} .$$

(4.3.39)

A single four-vector describes both electric and magnetic fields. As we will see in Lecture 5, in quantum mechanics, $A\mu$ is actually more fundamental than $\vec{E}$ and $\vec{B}$.

The transformations (4.3.35) and (4.3.36) combine into a single equation

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu \alpha ,$$

(4.3.40)

where we have defined the spacetime derivative

$$\partial_\mu \equiv \frac{\partial}{\partial X^\mu} = \left( \frac{1}{c} \frac{\partial}{\partial t} \cdot \frac{\partial}{\partial \vec{x}} \right) .$$

(4.3.41)

The Lorenz gauge condition (4.3.38) becomes

$$\partial^\mu A_\mu = 0 .$$

(4.3.42)

Two Maxwell equations are automatically satisfied, but what about the other two? (M1) and (M4) have source terms: the charge density $\rho$ (a scalar) and the current density $\vec{J}$ (a vector). Again, a scalar and a vector make a four-vector

$$\rho, \vec{J} \Rightarrow J^\mu = \begin{pmatrix} \rho/c^2 \\ \vec{J} \end{pmatrix} .$$

(4.3.43)
As I show in the insert below, expressed in terms of the four-vectors \( A^\mu \) and \( J^\mu \), the Maxwell equations (M1) and (M4), unify into a single pretty equation

\[
\Box A^\mu = \mu_0 J^\mu, \tag{M5}
\]

where \( \Box \) is shorthand for \( \eta^{\mu\nu} \partial_\mu \partial_\nu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \). Notice that \( \Box \) is the same combination that appears in the wave equation (4.2.18). Hence, there will be similar wave solutions for \( A_\mu \).

Substitute eq. (4.3.34) into the Maxwell equation (M1),

\[
\vec{\nabla} \cdot \vec{E} = -\frac{\partial}{\partial t} \vec{\nabla} \cdot \vec{A} - \nabla^2 \phi = \frac{\rho}{\epsilon_0}, \tag{4.3.44}
\]
or

\[
\left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = \frac{\rho}{\epsilon_0} + \frac{\partial}{\partial t} \left( \vec{\nabla} \cdot \vec{A} + \frac{\dot{\phi}}{c^2} \right). \tag{4.3.45}
\]

In the Lorenz gauge (4.3.38), this becomes

\[
\Box \phi = \frac{\rho}{\epsilon_0}. \tag{4.3.46}
\]

Substituting eqs. (4.3.32) and (4.3.34) into (M4) gives

\[
\vec{\nabla} \times \vec{B} = \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \mu_0 \vec{J} + \epsilon_0 \mu_0 (\vec{A} - \vec{\nabla} \phi). \tag{4.3.47}
\]

Using the identity \( \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla} (\vec{\nabla} \cdot \vec{A}) - \nabla^2 \vec{A} \), we get

\[
\left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \vec{A} = \mu_0 \vec{J} - \vec{\nabla} \left( \vec{\nabla} \cdot \vec{A} + \frac{\dot{\phi}}{c^2} \right). \tag{4.3.48}
\]

In the Lorenz gauge (4.3.38), this becomes

\[
\Box \vec{A} = \mu_0 \vec{J}. \tag{4.3.49}
\]

Eqs. (4.3.46) and (4.3.49) combine into eq. (M5).

Let us summarise what we have achieved:

1. We unified electric (\( \vec{E} \)) and magnetic (\( \vec{B} \)) fields into the 4-vector potential \( A_\mu \).

2. We reduced the 4 Maxwell equations to just 1.

### 4.3.2 A Hidden Symmetry

Eq. (M5) is more than just a pretty way of representing all the information of the Maxwell equations. It has Lorentz symmetry! Everything is packaged into 4-vectors without any lose ends. Space and time appear precisely in the symmetric way required by relativity.

The 4-vectors \( A^\mu \) and \( J^\mu \), transform in the same way as \( X^\mu \) under Lorentz transformation, cf. eq. (4.2.29),

\[
A^\mu \rightarrow A'^\mu = \Lambda^\mu_\nu A^\nu, \tag{4.3.50}
\]

\[
J^\mu \rightarrow J'^\mu = \Lambda^\mu_\nu J^\nu. \tag{4.3.51}
\]

\(^3\)Eq. (M5) uses the Lorenz gauge (4.3.38). It looks a little less pretty in a general gauge: \( \Box A^\mu = \mu_0 J^\mu + \partial^\mu (\partial^\nu A_\nu) \). This can be rearranged into \( \partial_\nu F^{\nu \mu} = \mu_0 J^\mu \), where \( F^{\nu \mu} \equiv \partial^\nu A^\mu - \partial^\mu A^\nu \).
The operator $\square$ doesn’t have a free index, so it doesn’t transform. Explicitly,

$$\square \mapsto \square' = \eta_{\mu\nu} \partial^\mu \partial^\nu = \eta_{\mu\nu} \Lambda^\mu_\alpha \Lambda^\nu_\beta \partial^\alpha \partial^\beta = \eta_{\alpha\beta} \partial^\alpha \partial^\beta = \square.$$  \hspace{1cm} (4.3.52)

Hence,

$$\square A^\mu \mapsto \square' A^\mu = \Lambda^\mu_\nu \square A^\nu.$$  \hspace{1cm} (4.3.53)

Combining (4.3.53) and (4.3.51), we see that if (M5) holds in the frame $S'$, then it also holds in the frame $S$,

$$\square A^\mu = \mu_0 J^\nu \iff \square A^\mu = \mu_0 J^\mu.$$  \hspace{1cm} (4.3.54)

This is the defining feature of a theory that is consistent with Einstein’s relativity. Its equations take the same form in all inertial frames.\(^4\)

It is remarkable that Maxwell’s theory automatically was of that form and didn’t need any fixing. This isn’t true for many other theories. Consider, for example, Coulomb’s law

$$\frac{m}{4\pi\varepsilon_0} \frac{d^2 x}{dt^2} = \frac{1}{4\pi\varepsilon_0} \frac{qQ}{x^2}.$$  \hspace{1cm} (4.3.55)

No matter how much massaging you do, you will never be able to write this as an equation involving only four-vectors. It also easy to see that (4.3.55) violates the basic principle of relativity that nothing travels faster than the speed of light. Changing the source at $x = 0$ instantly affects the solution everywhere in space. The famous Coulomb’s law is therefore only an approximation. The exact result follows from Maxwell’s equations.

### 4.4 More Unification?*

We are done with the basic story of the unification of electricity and magnetism through relativity. However, it is tempting to speculate a little further and ask how gravity, the second fundamental force of Nature, could fit into this. Let me emphasize that I am describing this just for fun and it is not clear whether this has anything to do with reality.

#### 4.4.1 Kaluza-Klein Theory

In 1919, Theodor Kaluza send a letter to Einstein. He described to him a proposal for unifying gravity and electromagnetism. Einstein liked the idea.

In this lecture, we have seen that the fundamental field in electromagnetism is the four-vector potential $A_\mu(X)$. Here, I have indicated that $A_\mu$ depends on the spacetime coordinate $X^\mu$. In Lecture 6, we will see that gravity is described by a more complicated object called the metric $g_{\mu\nu}(X)$. This is like the Minkowski metric (4.2.28), except that now all entries of the “matrix” can depend on the spacetime location $X^\mu$. In Lecture 6, we will have much more to say about $g_{\mu\nu}$. For now, just think of $g_{\mu\nu}$ as a $4 \times 4$ matrix whose entries vary through spacetime. Since $g_{\mu\nu}$ is symmetric it has 10 independent components.

It is a fundamental theorem of algebra that

$$15 = 10 + 4 + 1.$$  \hspace{1cm} (4.4.56)

\(^4\)These days we usually reverse the logic. When we come up with new theories, we don’t dare to write down anything that doesn’t have Lorentz symmetry.
Kaluza noticed that 15 is the number of independent components of a $5 \times 5$ symmetric matrix. A $5 \times 5$ symmetric matrix $G_{MN}$ has enough room to fit both the vector potential and the spacetime metric:

$$
\begin{pmatrix}
G_{00} & G_{01} & G_{02} & G_{03} & G_{04} \\
G_{10} & G_{11} & G_{12} & G_{13} & G_{14} \\
G_{20} & G_{21} & G_{22} & G_{23} & G_{24} \\
G_{30} & G_{31} & G_{32} & G_{33} & G_{34} \\
G_{40} & G_{41} & G_{42} & G_{43} & G_{44}
\end{pmatrix}
= 
\begin{pmatrix}
g_{00} & g_{01} & g_{02} & g_{03} & A_0 \\
g_{10} & g_{11} & g_{12} & g_{13} & A_1 \\
g_{20} & g_{21} & g_{22} & g_{23} & A_2 \\
g_{30} & g_{31} & g_{32} & g_{33} & A_3 \\
A_0 & A_1 & A_2 & A_3 & S
\end{pmatrix}
$$

(4.4.57)

In more compact form,

$$
G_{MN} = 
\begin{pmatrix}
g_{\mu\nu} & A_{\mu} \\
A_{\mu} & S
\end{pmatrix}
$$

(4.4.58)

Kaluza went on to speculate about the physical meaning of the object $G_{MN}$. He suggested it may be the metric of a five-dimensional spacetime. More than that: he showed that the equation of gravity in 5D could be split into the equations of Einstein’s theory gravity in 4D and Maxwell’s theory of electromagnetism:

$$
5D \text{ gravity} = 4D (\text{gravity} + \text{electromagnetism})
$$

(4.4.59)

Einstein was intrigued that gravity in five dimensions unifies gravity and electromagnetism in four dimensions.

Kaluza’s theory had an obvious flaw. The world isn’t five-dimensional! Or, is it? In 1926, Oscar Klein pointed out that we wouldn’t notice that the world is five-dimensional if one of the space dimensions, say $x_4$, is curled up into a small circle:

This became known as Kaluza-Klein theory.

### 4.4.2 String Theory

String theory requires extra dimensions of space to be self-consistent. Most versions of the theory have six extra dimensions. They are curled up into a small ball, maybe as small as $10^{-34}$ cm:

---

*In fact, $G_{MN}$ has room to spare for an extra scalar field $S(X)$. We will ignore this detail.*
Now there is enough room to unify all forces of Nature: gravity, electromagnetism, weak and strong nuclear force all become one. Unfortunately, we haven’t yet found ways to test the theory experimentally. It therefore remains an intriguing, but highly speculative, endeavor.