1.2 Motivation: Frame-Invariant Equations, and The Equivalence Principle

Last time we mentioned that the laws of physics are the same in every reference frame. This intuitive idea is, really, the one true motivation of relativity. However we only mentioned it in passing. It led, through Maxwell's equations, to the constancy of speed of light, which in turn led to the necessity to downgrade time from a parameter to a coordinate. However, the statement that laws of physics are the same in every reference frame should be far more central in our study of physics, rather than just a tool to get to the constancy of speed of light.

Before addressing the universality of laws of physics, however, we need to sort out something first: exactly what a "reference frame" is. Laws of physics are, after all, written as equations. Hence our job is to find equations describing particles' motion that are valid in every reference frame. It is therefore crucial that we understand how to deal with reference frames mathematically.

1 But Mathematically, What Exactly Is A Reference Frame?

Imagine you want to describe the motion of a particle on a garden-variety two-dimensional plane (i.e. the usual Cartesian $\mathbb{R}^2$). In fact, since we're focusing on reference frames rather than motion, let's strip away some unnecessary details and assume that is particle is just sitting there at rest. Therefore what we really want to do is to describe the position of a particle on $\mathbb{R}^2$.

Figure 1: Here it is: a particle at rest on a garden-variety $\mathbb{R}^2$ plane.

To describe its position, we need to set up a reference frame, i.e. we need a position to call zero (usually this zero position is called the "origin"), we need two directions to call x and y, and we need to decide how long the unit distance is [1]. Nothing too fancy here.

[1]For example, you could choose a frame so that $x = 1$ is 1 centimeter away from $x = 0$, or you can choose 1 meter, or you can choose 24 lightyears. I personally love the number 24: the square of every prime number (apart from 2 and 3) is one more than a multiple of 24.

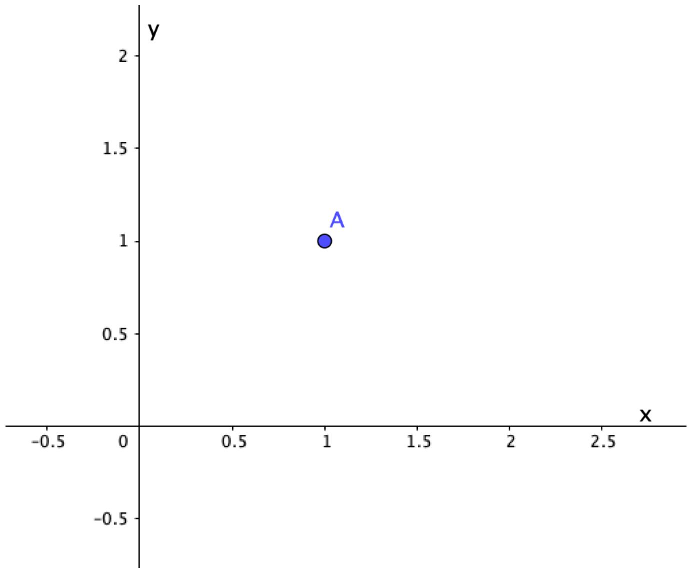

Figure 2: Setting up a reference frame. Position of particle is $(x, y) = (1, 1)$.

But that's not the only possible way to set up the reference frame. For example, we could have chosen another point as the origin.

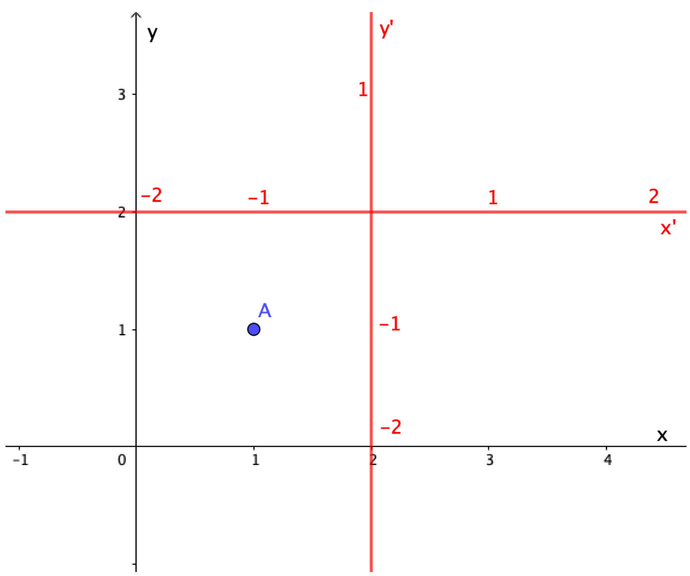

Figure 3: Setting up another reference frame with the same directions as x and y, the same unit length, but whose origin is at $(x, y) = (2, 2)$. Position of particle is $(x', y') = (-1, -1)$.

It's obvious that if a particle's position in the x-y reference frame is (x, y), then its position in the x'-y' reference frame is $(x', y') = (x - 2, y - 2)$.

This is, in fact, a functional relationship between the old coordinate variables x, y and the new coordinate variables x', y':

$$\begin{cases} x' = x - 2 \\ y' = y - 2 \end{cases}$$

Explicitly, if you have the values for x, y that describe the position of a particle in one frame, then the above functions give you the values for x', y' that describe the position of a particle in the other frame.

The other two ways we had for setting up different reference frames also correspond to simple functional relationships. If the unit distance is 1 centimeter in x'-y' frame instead of the 1 meter in x-y frame, then the point $(x, y) = (1, 1)$ would be described in the primed frame as $(x', y') = (100, 100)$. [2] The functional relationship is thus

$$\begin{cases} x' = 100x \\ y' = 100y \end{cases}$$

Rotations work similarly. I'll assume that you know this already. This specific rotation in $\mathbb{R}^2$ won't really concern us, so you don't need to memorize who is the sine and who is the cosine. All that matters is the realization that switching to a rotated reference frame is, like switching to any other kind of reference frames, nothing more than a simple functional relationship between the old and new coordinate variables:

$$\begin{cases} x' = x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta \end{cases}$$

Let's box this realization:

Switching to a different reference frame means a functional relationship between the old and new coordinate variables.

Or, actually, physicists use the word "transform" instead of "switch", so let's adopt the same wording:

Transforming to a different reference frame means a functional relationship between the old and new coordinate variables.

[2] (1,1) in unprimed frame means 1 unprimed unit distance, which is 1 meter away from the origin (in both x and y). If the unit distance were 1 centimeter, then the point would be 100 unit distances, which is 100 centimeters, away from the origin.

Perhaps the most famous coordinate transform is the polar coordinate:

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

It's a coordinate transform: again, given the values for r, θ that describe the position of a particle in one coordinate system, the above functions give you the values for x, y that describe the position of the particle in the other coordinate system.

We know that laws of physics are the same in every reference frame; this means the equations that we call "laws of physics" need to hold true in an arbitrary coordinate system [3]. Well, here's a law of physics: *free particles move on straight lines* (Newton's first law). Let's try to see whether the equations we use to describe this law hold true in all coordinate systems.

In Cartesian $\mathbb{R}^2$, the differential equation that describes straight lines is

$$\frac{d^2y}{dx^2} = 0$$

Or in more familiar terms, $y = ax + b$.

Since the laws of physics shouldn't favor one coordinate system over another, all coordinate variables are on the same footing. Thus, the coordinate variables r, θ should take the exact same places in the equations representing the laws of physics as the places held by the coordinate variables x, y. Therefore, we would expect something like

$$\frac{d^2r}{d\theta^2} = 0$$

or $r = a\theta + b$ to describe straight lines in polar coordinates, if $\dfrac{d^2y}{dx^2} = 0$ were indeed an equation for the law of physics "free particles travel on straight lines". Unfortunately, it's not that simple. Let's choose a simple combination $a = 1, b = 0$, i.e. $r = \theta$. Can you try to imagine how it looks like? Here it is:

[3] I'm using the words "reference frame" and "coordinate system" somewhat interchangeably now, and it's ok: we know that transforming to a different reference frame is, mathematically, just defining a couple of new coordinate variables in terms of the old coordinate variables functionally.

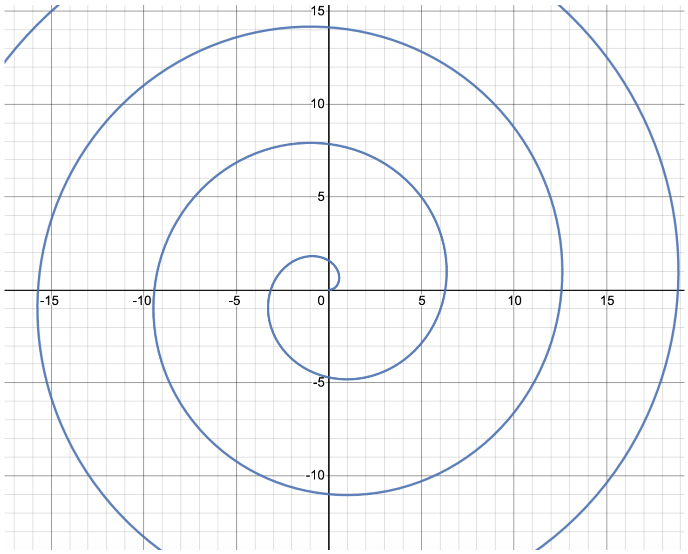

Figure 4: The curve $r = \theta$ in polar coordinates. Definitely not straight lines!

In fact, if we were to parametrize the particle's path using time t as the parameter (let's forget about how time should be a coordinate and just use Newtonian time as parameter for simplicity here), then free particles travel on

$$\begin{cases} \dfrac{d^2x(t)}{dt^2} = 0 \\ \\ \dfrac{d^2y(t)}{dt^2} = 0 \end{cases}$$

in Cartesian coordinates, i.e. the nice and familiar straight lines, but in polar coordinates they travel on the following monstrosity [4]:

$$\begin{cases} \dfrac{d^2r(t)}{dt^2} = r\left(\dfrac{d\theta(t)}{dt}\right)^2 \\ \\ \dfrac{d^2\theta(t)}{dt^2} = -\dfrac{2}{r}\dfrac{dr}{dt}\dfrac{d\theta}{dt} \end{cases}$$

Obviously, the naive $\dfrac{d^2(\text{coordinate variable})}{dt^2} = 0$ is not a good equation for the law of physics "free particles travel on straight lines", since it doesn't hold true in every coordinate system.

[4]These are the differential equations for a straight line in polar coordinates. Don't worry, we'll learn how to derive it later.

If we want to find an equation that encodes the law of physics "free particles travel on straight lines", then it couldn't be some particular differential equation. It has to be a differential equation such that, when we somehow plug in Cartesian coordinates, polar coordinates, or really ANY coordinates on $\mathbb{R}^2$ into that equation, it gives us the correct equation for straight lines *in that coordinate*. The form of that overarching equation cannot change as we transform to different kinds of coordinates.

In other words, **the equations for the laws of physics need to be invariant under coordinate transforms**. Keeping that in mind, let's include time as a coordinate, and inspect one of the world's favorite "law of physics": the famous $F = ma$.

2 Is F=ma a law of physics?

Let's simplify matters by considering a one-dimensional world, i.e. the Cartesian $\mathbb{R}^1$. Let's invite back our two good friends, Alice on a stationary platform, and Bob on a moving train. Alice's stationary frame will be unprimed, and Bob's moving frame will be primed. Let's also add in a free particle. For simplicity, the free particle is simply sitting at $x = L$ at rest in Alice's stationary frame (not moving is a kind of moving on a straight line: it's moving with a velocity of zero).

Also, to get ourselves familiar with the idea that time should be a coordinate instead of a parameter, let's use 2-dimensional spacetime reference frames, i.e. t-x for Alice, and t'-x' for Bob.

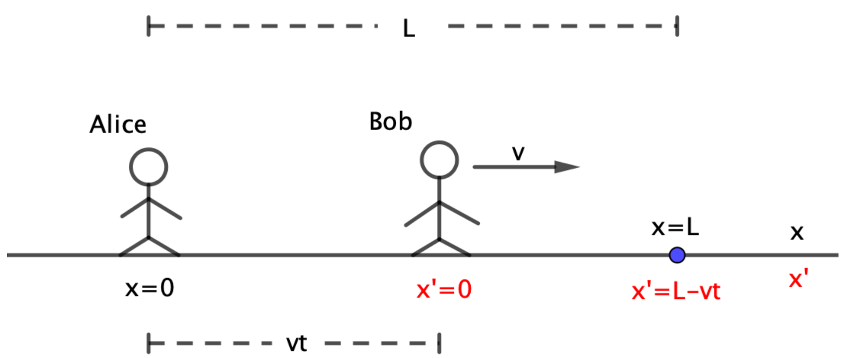

Figure 5: Newtonian spatial space of the setup. In Alice's frame, the free particle is staying at rest at $x = L$, and Bob is moving to the right with constant velocity v. When Bob passes Alice, they start their two clocks together, i.e. the moment when $x = 0$ and $x' = 0$ coincides, $t = t' = 0$.

Let's assume Newtonian time, i.e. the two clocks carried by Alice and Bob tick off at the same speed. Since the two clocks are started from zero together and they tick off at the same rate, their reading will always agree, so $t' = t$. From the picture, it is clear that when Alice's clock reads t, in her frame Bob will be displaced by vt from her, so the particle will be displaced by $L - vt$ from Bob.

We can now write down the coordinate transform between Alice's stationary frame (unprimed) and Bob's train frame (primed), i.e. a function from the unprimed coordinate variables to the primed coordinate variables:

$$\begin{cases} t' = t \\ x' = x - vt \end{cases}$$

Exercise:
1. Draw the particle's path in spacetime, in both Alice's and Bob's coordinates. As a convention, when drawing spacetime, the spatial dimension (so x) is the horizontal axis, and the time dimension (so t) is the vertical axis.

2. Write down the parametric equations for the particle's path in both Alice's and Bob's coordinates. Use the Newtonian time t_{param} as the parameter. I'll do one for you: since the time coordinate t is just the parameter of the path t_{param} (both are the Newtonian time), so for Alice, $t(t_{param}) = t_{param}$.

When doing the above exercise, first do them for Alice's coordinates, then plug the results for Alice's coordinates into the coordinate transform to obtain the results in Bob's coordinates. Make sure to check whether the result in Bob's coordinates make sense.

Answer:

Alice:

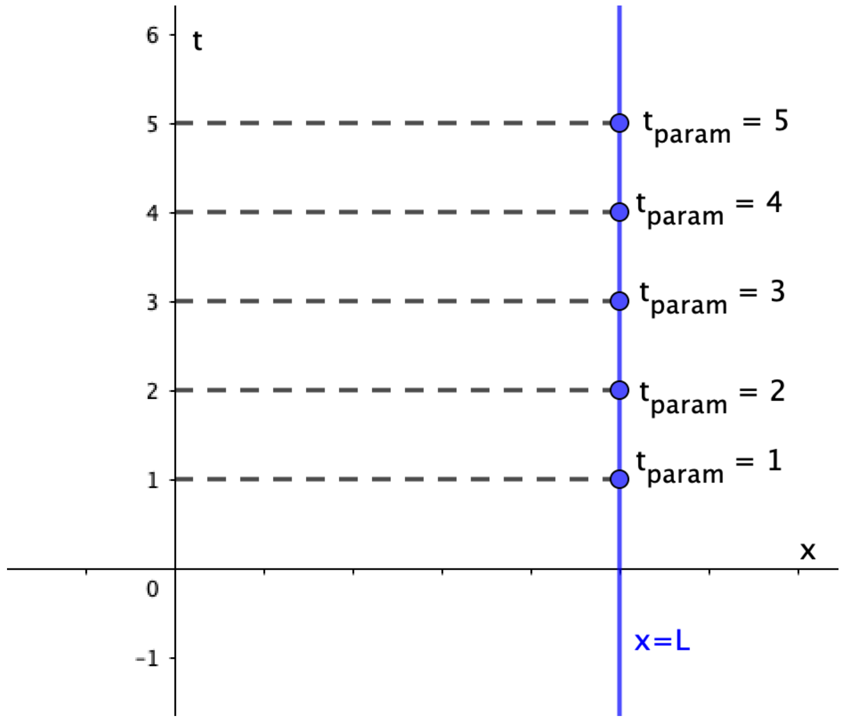

Figure 6: Particle's path in Alice's spacetime coordinates. It is sitting at rest at $x = L$.

Particle's path in Alice's frame is

$$\begin{cases} t(t_{param}) = t_{param} \\ x(t_{param}) = L \end{cases}$$

* Don't feel uncomfortable with $t(t_{param}) = t_{param}$. We are simply parameterizing a curve using one of the coordinates as the parameter. The subscript is just for explicitly reminding ourselves that it is a parameter. It's like calling the curve $y = x^2$ by $(x(x), y(x)) = (x, x^2)$, and then renaming the parameter to $(x(x_{param}), y(x_{param})) = (x_{param}, x^2_{param})$. It's a parameter; it doesn't matter what we call it.

Bob:

Plugging

$$\begin{cases} t(t_{param}) = t_{param} \\ x(t_{param}) = L \end{cases}$$

into

$$\begin{cases} t' = t \\ x' = x - vt \end{cases}$$

gives us

$$\begin{cases} t'(t_{param}) = t(t_{param}) = t_{param} \\ x'(t_{param}) = x(t_{param}) - vt(t_{param}) = L - vt_{param} \end{cases}$$

We can then graph the parametric equation:

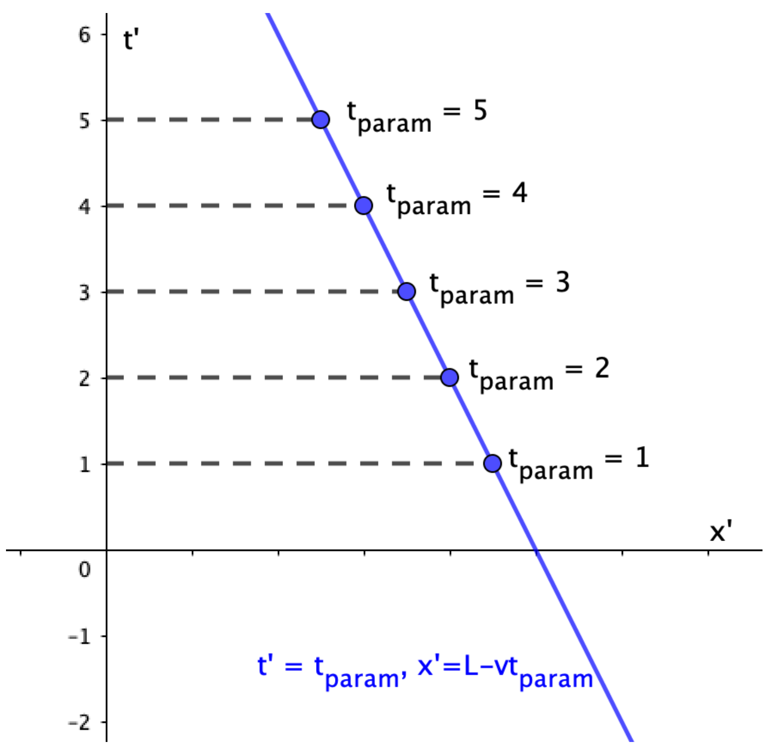

Figure 7: Particle's path in Bob's spacetime coordinates. It's moving to the left with constant velocity v, as expected.

Let's now check $F = ma$ in both frames. Since it's a free particle, so $F = 0$. Therefore we can throw away the mass and simplify it to $a = 0$.

Acceleration is $\dfrac{d^2x}{dt_{param}^2}$ (or $\dfrac{d^2x}{dt^2}$, it doesn't make a difference here because the coordinate t and the parameter t_{param} are the same, both are the Newtonian time, but we'll stick to the better practice of defining velocities and accelerations with respect to parameters).

In Alice's frame:

$$\begin{cases} t(t_{param}) = t_{param} \\ x(t_{param}) = L \end{cases}$$

$$\therefore a = \frac{d^2x}{dt_{param}^2} = \frac{d}{dt_{param}}\left(\frac{dx}{dt_{param}}\right) = \frac{d}{dt_{param}}\left(\frac{dL}{dt_{param}}\right) = 0$$

since L is constant.

In Bob's frame:

$$\begin{cases} t'(t_{param}) = t_{param} \\ x'(t_{param}) = L - vt_{param} \end{cases}$$

$$\therefore a' = \frac{d^2x'}{dt_{param}^2} = \frac{d}{dt_{param}}\left(\frac{dx'}{dt_{param}}\right) = \frac{d}{dt_{param}}\left(\frac{d(L - vt_{param})}{dt_{param}}\right) = \frac{d}{dt_{param}}\Big(-v\Big) = 0$$

since Bob's velocity v is constant.

You will get the same answer if you use $a = \dfrac{d^2x}{dt^2}$ and $a' = \dfrac{d^2x'}{dt'^2}$.

So $F = ma$ holds in both frames. But what about when Bob's train is accelerating?

Imagine that instead of moving with a constant velocity v, Bob's train is moving with a constant *acceleration* a_0.

The particle's path in Alice's coordinates is still

$$\begin{cases} t(t_{param}) = t_{param} \\ x(t_{param}) = L \end{cases}$$

but in Bob's frame it's now [5]

$$\begin{cases} t'(t_{param}) = t_{param} \\ \\ x'(t_{param}) = L - \dfrac{1}{2}a_0 t_{param}^2 \end{cases}$$

The acceleration of the particle in Bob's frame is

$$a' = \frac{d^2x'}{dt_{param}^2} = \frac{d}{dt_{param}}\left(\frac{dx'}{dt_{param}}\right) = \frac{d}{dt_{param}}\left(\frac{d(L - \frac{1}{2}a_0 t_{param}^2)}{dt_{param}}\right) = \frac{d}{dt_{param}}\left(-a_0 t_{param}\right) = -a_0$$

which is not zero, as $F = ma$ would suggest for a free particle!

So, since $F = ma$ doesn't hold true in all reference frames, it is unfortunately **not** a law of physics. Bummer.

[5] Reminder: uniform accelerating trains cover distances as $\frac{1}{2}at^2$.

3 Curved coordinates

But from our exploration, some trails of clue can be found. Specifically, the equations only seem to break down when the coordinates are curved.

In our $\mathbb{R}^2$, equations for straight lines were valid for the usual Cartesian coordinates x, y. Although we didn't explicitly check them, but under translations

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x-2 \\ y-2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

stretches and/or compressions

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 100x \\ 100y \end{pmatrix} = \begin{pmatrix} 100 & 0 \\ 0 & 100 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and rotations

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\sin\theta \\ -x\sin\theta + y\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

the equations for straight lines are all the same: $\dfrac{d^2(\text{coordinate variable})}{dt^2} = 0$. This is because those transformations are linear [6], hence the second derivative of the coordinate variables vanishes.

However, a non-linear (or "curved") coordinate transform like polar coordinates doesn't enjoy this property. The second derivative of a non-linear function can, in general, be anything. Correspondingly, the equation for straight lines in polar coordinates was different.

[6] Reminder that a transformation is linear if it can be written as a matrix.

For Alice and Bob, when Bob was moving with a constant velocity, their coordinate transform was also linear:

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} t \\ x - vt \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so $F = ma$ holds for both frames. However, when Bob's frame isn't linear (i.e. "curved") with respect to Alice's frame, i.e. when train is accelerating and

$$\begin{cases} t' = t \\ \\ x' = x - \dfrac{1}{2} a_0 t^2 \end{cases}$$

then $F = ma$ breaks down.

In general, **coordinates can do whatever they like**. There's nothing wrong with curved coordinates: people use polar coordinates and spherical coordinates (in $\mathbb{R}^3$) all the time. But as we have seen, many garden-variety Newtonian equations do not hold true in curved coordinates. Hence they can't be called "laws of physics" by our standards.

To write down equations for laws of physics, we have to find new equations that hold true in both straight coordinates, and curved coordinates as well. This is the central question in general relativity. We will devote the entirety of the first section to this question.

4 The Equivalence Principle

There's a separate observation we can make from the accelerating train exploration. Notice that in Bob's frame, the particle is traveling with an acceleration of $-a_0$. This smells like a uniform gravitational field, like how at the surface of the Earth everything travels with an acceleration of g downwards on top of whatever it was doing on its own. [7]

Einstein used elevators instead of trains: in an elevator accelerating to the right with some constant acceleration a_0, according to an observer in the elevator, everything will seem to be accelerating to the left with a_0. In effect, there is a uniform "gravity" pulling everything to the left.

This is the famous **Equivalence Principle**.

> **The Equivalence Principle:**
> Uniform gravity is indistinguishable from uniform acceleration.

But as we know, non-uniform gravitational fields do exist. Gravitational fields around clumps of mass, like the fields around the sun or around some black hole, is certainly not uniform. It points radially, and its strength diminishes with the (square of the) distance. For these gravitational fields, there's no way to mimic them with an accelerating reference frame.

[7] Some flat-Earthers explain gravity with this: the apparent gravity we experience everyday is an illusion, and in reality it's the flat earth accelerating upwards with g. It's interesting that there are people who understand the relationship between gravity and acceleration and at the same time believe that the Earth is flat.

You might have heard that black holes curve spacetime in scifi movies. Indeed, when it comes to general relativity, there seems to be a great deal of hype around curvature. The relationship between matter and spacetime is best summarized by John Wheeler: "**spacetime tells matter how to move; matter tells spacetime how to curve**".

We will eventually end up with an equation that tells us, given the structure of a spacetime, how free particles travel. This equation, called the geodesic equation, is the "spacetime tells matter how to move":

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 0$$

And we will have another equation that tells us, given a mass, how spacetime responds to it. This equation, called the Einstein field equation, is the "matter tells spacetime how to curve":

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}$$

So yes, around a clump of mass, spacetime is indeed curved. This curvature is drastically different from the fictitious gravity from the equivalence principle. In the equivalence principle, there are no gravitational sources like suns or black holes anywhere. The spacetime there is flat. There is an apparent gravity (i.e. an apparent curvature of spacetime) because **we have chosen curved coordinates** (i.e. accelerating frame). Had we chosen a set of linear coordinates, for example a non-accelerating reference frame, that apparent fictitious gravity will go away [8].

By contrast, the spacetime around a black hole or a clump of mass is intrinsically curved; there are no linear coordinates we can choose such that we can eliminate the gravity. This point is so important that I want to stress it as much as possible:

On a flat spacetime (one without any gravitational source), we can choose either flat coordinates or curved coordinates.

On a curved spacetime (one with a gravitational source), no global flat coordinates are possible.

[8] This is the good old Newton's first law: if you are in a train moving with constant velocity, you feel just as if you are on a stationary platform.

The important takeaway is that there are two kinds of curvature: the curvature of the coordinates we assign, and the curvature of spacetime itself. The equivalence principle is a set of curved coordinates on a flat spacetime; around a black hole it's intrinsically curved spacetime. Lightclocks moving across each other at constant speed is flat coordinates on flat spacetime.

It's like the following difference: on the garden-variety plane $\mathbb{R}^2$, we can use either flat coordinates (Cartesian), or curved coordinates (polar). This is the curvature of coordinates. No matter how we curve our coordinates, the space itself (which is $\mathbb{R}^2$) will still be flat. By contrast, a two-dimensional surface like the surface of a sphere is intrinsically curved, regardless of what coordinates we use on it.

The two kinds of curvature play different roles in general relativity. Curvature of spacetime is an object of study: we want to study how spacetime curves in the presence of some mass (the Einstein field equations), or for that matter, we might want to study how to detect if a given spacetime is curved at all. However, curvature of coordinates is not something we study. It is more of a principle: we want the laws of physics to obey the principle that laws of physics should be the same in every reference frame. This is a constraint on our equations: we want equations that are invariant after we do an arbitrary coordinate transform. The geodesic equation and the Einstein field equation of course will obey this constraint.

It is important that you don't get the two kinds of curvature mixed up. As we said, we will devote the entirety of the first section to finding out a frame-invariant equation that describes how free particles move in spacetime. After that, we will use that equation on flat spacetime, but with both flat (non-accelerating) and curved (accelerating) coordinates. We will then see how that equation will give rise to the fictitious gravity in the curved coordinates, as described by the equivalence principle.