Massachusetts Institute of Technology
Department of Physics
8.033 Fall 2021

Lecture 1
Introduction; Newtonian mechanics and relativity

1.1 Goal of this course

Our goal this semester is to understand Albert Einstein's theories of relativity, and what they tell us about the structure of the laws of physics. We will study the special theory of relativity in great detail for most of the semester. In the last several weeks of the term, we will briefly explore general relativity, focusing on situations in which what we learn from special relativity can be "upgraded" to the general case with relative ease.

More generally, our goal this semester is to think about how we formulate physics in such a way that a deep underlying principle is built into these laws. During the first few weeks of this course, we will motivate how on both theoretical and experimental grounds we came to understand that our universe respects a principle we call *Lorentz symmetry*. We will find that some of the laws of physics we learned previously exactly respect this symmetry, whereas others are approximations (albeit incredibly accurate approximations under the "everyday" conditions that one typically encounters in a laboratory). We will then develop a way of representing physical laws such that they automatically satisfy Lorentz symmetry. You can take this as a kind of "warm-up exercise" for including other symmetries that nature respects, and other principles that we may need to build into physics.

1.2 Newtonian physics and Galilean relativity

Let us begin by examining Newtonian physics. Newton's laws obey **a** law of relativity, though it is not the one that we usually think of when we discuss "relativity" in physics. Rather than Einstein's relativity, Newton's laws respect what we call *Galilean relativity*, following principles that were originally laid out by Galileo. We will begin by examining Galilean relativity in order to see a relativity principle in action, based on laws of physics that we thoroughly know and love ... and to see an interesting shortcoming that Galilean relativity presents as we begin to discuss

To begin this discussion, we need to define some terms:

- **Event**: An "event" is something that happens somewhere at some time. An event is essentially a particular point in both space and time. A key feature is that the event's reality is independent of how we label it, which can depend on our "reference frame."

- **Reference frame**: A system for labeling events in space and time. One can regard a reference frame as essentially a clock and a set of coordinate axes that are tied to a particular observer. For example, the professor, standing in front of the classroom uses a clock on the wall to define time, and takes their position as the origin. They imagine an x axis pointing from their feet toward the back of the classroom; a y axis pointing from the professor's feet to their left, and a z axis pointing from the professor's feet up to the ceiling.

A student, sitting in the front row, sets up a similar reference frame. Also using the wall clock for time, the student likewise defines their position as the origin, imagines an x axis that points from their feet to the *front* of the classroom (they are facing in the opposite direction as the professor, so "forward" for them is opposite of "forward" for the professor), a y axis pointing from their feet to their left, and z axis pointing from their feet to the ceiling.

These two reference frames assign different labels to events, but both are perfectly valid provided they are used consistently.

⌈ASIDE: This is a good point to introduce some notation and definitions. Once we've introduced coordinate axes, it is very useful to define *unit vectors* which point along these axes. We will call the unit vectors along the x, y, and z axes $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ respectively. $\mathbf{e}_x$ is a dimensionless vector of magnitude 1 that is parallel to the x axis; likewise for the other two unit vectors. If more than 1 reference frame is being discussed, we will include some kind of label to distinguish them; e.g., $\mathbf{e}_x^P$ is the x unit vector for the professor's reference frame; $\mathbf{e}_x^S$ is the x unit vector for the student's reference frame.

In our typed-up notes, we will use boldface to denote vectors in 3-dimensional space. This does not work well in chalk, so we will instead use an undertilde (e.g., $\underset{\sim}{e}_x$).⌋

- **Geometric object**: Something with properties that exist independent of the reference frame that we use to describe it.

An example of a geometric object is an event. Suppose that at 2:47 pm on 8 September 2021, a piece of chalks strikes the professor's forehead. All observers note this event. They may assign different labels to it — the professor calls it $t = 14{:}47$ 08-09-2021, $x = 0$, $y = 0$, $z = 1.8$ meters; the student calls it $t = 14{:}47$ 08-09-2021, $x = 3$ meters, $y = -1$ meter, $z = 1.8$ meters — but they agree on the event itself, and their labelings in their reference frames describe the same thing.

Another example of a geometric object is a vector. Consider a meter stick that points from a table at the front of the classroom. All observers agree[1] that it has a length of 1 meter, and all agree that is poking out of the table at some angle. Depending on the reference frames that the different observers use, different observers will use different *representations* of that vector. This is of course fine as long as the different representations are used consistently in describing the physics of the system under study. Figure 1.2 illustrates this concept in a simple 2-dimensional example.

[1]This agreement won't hold up once we move beyond Galilean relativity! Hold that thought for now.

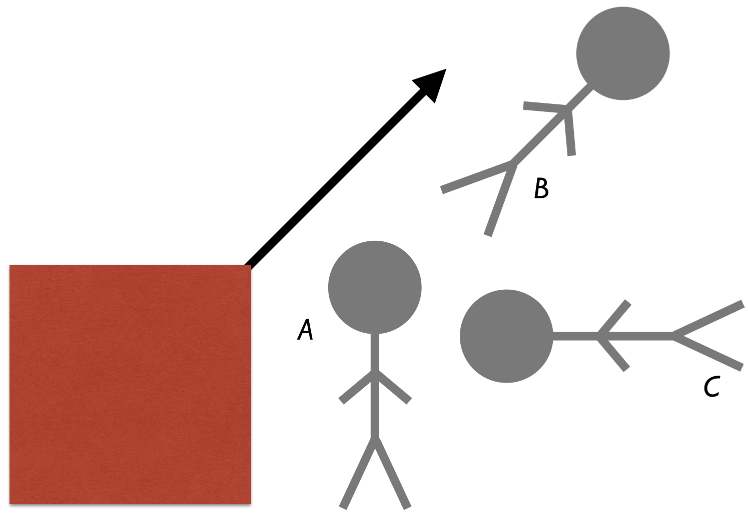

Figure 1: Three observers, three different inertial reference frames. (Imagine they are in a gravity-free environment so that each observer is inertial.) Each agrees that there is a large stick embedded in the table, making a 45° angle with its top and side. Observer A orients it along the direction $(\mathbf{e}^A_x + \mathbf{e}^A_z)/\sqrt{2}$; observer B orients it along $\mathbf{e}^B_z$; and observer C orients it along $(\mathbf{e}^C_x - \mathbf{e}^C_z)/\sqrt{2}$. These different representations are consistent, provided we correctly relate each observer's choice of coordinate axes to those of the other observers.

Much of relativity, whether it is Galileo's or Einstein's, is about making sure that we carefully and consistently describe things in different reference frames, and that we correctly relate the description of quantities according to one reference frame to those quantities according to another reference frame. Geometric objects are excellent tools for describing physics because objects like events, vectors, and tensors (which will be introduced and defined soon) have a meaning that transcends a particular reference frame's representation of that object. It is very important (and useful) to maintain a distinction in your mind between the object (e.g., a particular vector) and the representation of that object (e.g., the components of that vector according to some given frame of reference).

A particularly important kind of reference frame is an *inertial reference frame*, or IRF; we will use this term enough that it is worth abbreviating. This is a frame of reference in which a body's momentum is constant if no forces act on it: It is an *unaccelerated* reference frame. Non-inertial reference frames certainly exist, and require us to introduce what are sometimes called "fictitious" or "non-inertial" forces[2], like the centrifugal force or the Coriolis force.

1.3 Galilean transformations

How do we relate quantities as described in one reference frame to those in another? We can deduce how to do this by thinking about the (hopefully, straightforward) connection between how the two frames relate the representations of geometric objects. Doing so, we build a mathematical "machine" for connecting quantities between two reference frames. In

[2] I prefer "non-inertial" to "fictitious," since "fictitious" sounds quite a bit like "fake." Anyone who has ever crashed a bike taking a turn too fast or hurt their neck on an amusement park ride can tell you that there is nothing at all fictitious about those forces if you happen to be in the non-inertial frame.

Newtonian physics, we call the resulting mathematical machine the *Galilean transformation.*

Let us say IRF C is used by the class to label objects and events; IRF P is used by the professor. The professor and the class are oriented in the same way, so that their x, y, and z axes all point in the same direction. However, the professor is walking across the classroom: the class sees the professor moving with constant speed v in the x direction. How do we relate these two IRFs?

Consider time first. Is there any difference in time according to the class and to the professor? In Newtonian physics, the answer is no: Both the class and the professor get their time from the wall clock, no matter whether they are in motion or not. So we have

$$t_P = t_C \; . \tag{1.1}$$

Their representations of space differ, however. An object at a fixed position in IRF C "falls back" along x in IRF P:

$$\begin{aligned} x_P &= x_C - v t_P \\ &= x_C - v t_C \; , & (1.2) \\ y_P &= y_C \; , & (1.3) \\ z_P &= z_C \; . & (1.4) \end{aligned}$$

(Notice that x_P and x_C coincide when $t_P = t_C = 0$.) The set of four equations, (1.1)–(1.4), relating quantities in P to quantities in C is a Galilean spacetime transformation. It can be neatly written as a matrix equation:

$$\begin{pmatrix} t_P \\ x_P \\ y_P \\ z_P \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} t_C \\ x_C \\ y_C \\ z_C \end{pmatrix} \; . \tag{1.5}$$

Writing out 4×4 matrices like this gets unwieldly; a more compact form is

$$\vec{x}_P = \mathsf{G} \cdot \vec{x}_C \; . \tag{1.6}$$

⌈ASIDE: Here is another good point to introduce some more notation. When studying relativity, it will be useful to make "4-vectors," vectorial quantities with 4 components: the 3 spatial ones that you are probably familiar with from previous coursework, plus 1 more for a time component. In this course, whenever we write an object with an overarrow like $\vec{x}_P$, it refers to such a 4-vector. The same symbol in boldface or with an undertilde ($\mathbf{x}_P$ or $\underset{\sim}{x}_P$) refers to only the spatial components.⌋

How do we invert the relation we just wrote down? That is, given Eq. (1.5) relating quantities in frame C to those in frame P, how do we relate quantities in frame P to those in frame C? On the grounds of physics, this is simple: if C says that P moves with velocity $\mathbf{v} = v\mathbf{e}_x$, then P says that C moves with $\mathbf{v} = -v\mathbf{e}_x$. We quickly deduce that

$$\begin{pmatrix} t_C \\ x_C \\ y_C \\ z_C \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} t_P \\ x_P \\ y_P \\ z_P \end{pmatrix} \; , \tag{1.7}$$

which we will write more compactly as

$$\vec{x}_C = \mathsf{G}' \cdot \vec{x}_P \ . \tag{1.8}$$

It is easy to verify that $\mathsf{G}' = \mathsf{G}^{-1}$, i.e., that G' is the matrix inverse of G. This is an exercise on problem set #1.

Another example of a Galilean transformation: Suppose IRFs C and P are at rest with respect to each other, but are rotated about the z axis, as shown in Figure 1.3.

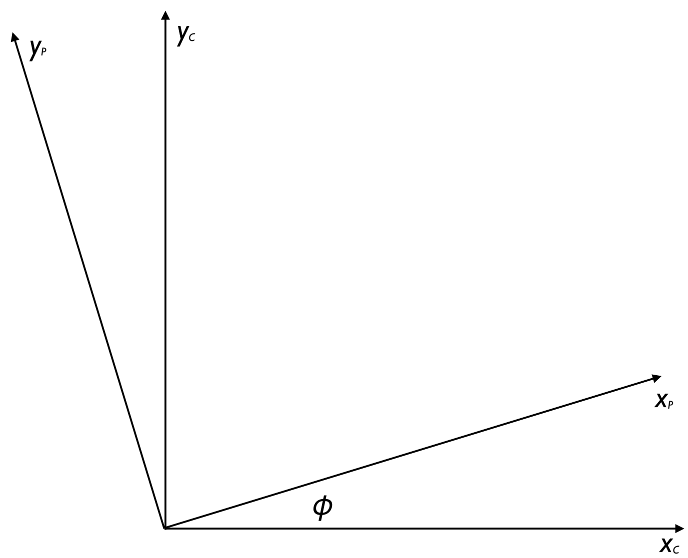

Figure 2: x and y axes for IRFs C and P, related to each by a rotation about the z (which is the same to both frames).

Space and time are related between the two frames with the equations

$$\begin{aligned} t_P &= t_C \ , & (1.9)\\ x_P &= x_C \ \cos\phi + y_C \ \sin\phi \ , & (1.10)\\ y_P &= -x_C \ \sin\phi + y_C \ \cos\phi \ , & (1.11)\\ z_P &= z_C \ . & (1.12) \end{aligned}$$

The spatial part of this transformation is just a simple rotation. We can write this

$$\begin{pmatrix} t_P \\ x_P \\ y_P \\ z_P \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & \sin\phi & 0 \\ 0 & -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} t_C \\ x_C \\ y_C \\ z_C \end{pmatrix} \ , \tag{1.13}$$

or

$$\vec{x}_P = \mathsf{G}_R \cdot \vec{x}_C \ , \tag{1.14}$$

where G_R denotes a Galilean transformation that's a pure rotation.

The examples discussed here are not comprehensive[3], but hopefully gives you the gist of the idea. We will explore these concepts a bit more on the problem sets.

[3]There's one more very simple one that is worth mentioning: A shift of origin. Suppose the professor and the class have the same orientation, but the professor center spaces on their location, and perhaps their watch is set to a different time zone. Then, $\vec{x}_P = \vec{x}_C + \Delta\vec{x}$, where $\Delta\vec{x}$ is a constant offset.

1.4 Transformation of velocities and accelerations

A key feature of the transformations we have been discussing is that they leave inertial frames inertial. To see this, let's go back to our first example, the professor moving with speed v in the x direction as seen by the class and examine how an object's *velocity* transforms under Galilean transformations. Consider an object moving with velocity $\mathbf{u}$, where $u^x = dx/dt$, $u^y = dy/dt$, $u^z = dz/dt$. (Notational note: We will use the letter v or $\mathbf{v}$ to denote the relative velocity of two different frames of reference; we will use the letter u or $\mathbf{u}$ to denote velocity within some specified frame.)

Suppose the class sees some object moving with velocity $\mathbf{u}_C$, with components

$$u_C^x = \frac{dx_C}{dt_C}\ , \qquad u_C^y = \frac{dy_C}{dt_C}\ , \qquad u_C^z = \frac{dz_C}{dt_C}\ . \tag{1.15}$$

What are the components as seen by the professor? Let's apply the Galilean transformation rules and find out:

$$\begin{aligned} u_P^x &= \frac{dx_P}{dt_P} \\ &= \frac{d}{dt_P}\left(x_C - vt_C\right) \\ &= \frac{d}{dt_C}\left(x_C - vt_C\right) \\ &= u_C^x - v\ . \end{aligned} \tag{1.16}$$

By a similar calculation, we see that $u_P^y = u_C^y$, $u_P^z = u_C^z$. This is nothing more than the "normal" velocity transformation that we are familiar with from Newtonian mechanics.

How about accelerations? As usual, we have $\mathbf{a} = d\mathbf{u}/dt$. Imagine that the object is seen by the class to have acceleration $\mathbf{a}$, and use the Galilean transformation to deduce what the professor sees for its acceleration:

$$\begin{aligned} a_P^x &= \frac{du_P^x}{dt_P} \\ &= \frac{d}{dt_P}\left(u_C^x - v\right) \\ &= \frac{d}{dt_C}\left(u_C^x - v\right) \\ &= a_C^x\ . \end{aligned} \tag{1.17}$$

We likewise find $a_P^y = a_C^y$, $a_P^z = a_C^z$: the class and the professor agree on the object's acceleration, at least as long as v is constant in time. As long as the two frames are not accelerated with respect to one another, Galilean transformations take one inertial representation and yield another inertial representation.

1.5 Relativity, covariance, and invariance

You have now been introduced to the workings of Galilean relativity. Having seen a few examples of how it works, this is a good opportunity to carefully define a few terms that we are going to use a lot in this course.

- *Relativity*: We've been using "relativity" quite a bit without actually defining it explicitly. A relativity framework is just a way of transforming observables — particularly the representation of geometric objects which we use to describe important quantities in physics — from the reference frame of one observer to another.

- *Covariance*: We describe a law or principle of physics as covariant if it holds in all frames of reference. For example, as you will explore on a problem set, the law of momentum conservation is covariant in Galilean relativity. This does not mean that observers in all inertial frames agree on the specific value of an object's momentum; indeed, you should be able to convince yourself that you can make that object's momentum take any value at all by changing frames of reference. However, all observers agree that if that body interacts with another body, then the momenta of the two bodies after their interaction is the same as it was before.

- *Invariance*: Some quantities are in fact exactly the same in all frames of reference. The mass of a body is the same to all observers in Galilean relativity; a particular notion of mass is the same to all observers in Einstein's relativity; a body's electric charge is the same to all observers in all forms of relativity. Quantities which are the same in all frames of reference are called *invariants*. Learning when and how to exploit invariance is one of the skills we will practice this term; used well, invariants often make it possible to significantly simplify a calculation.

1.6 Transformation of waves in Galilean relativity

A particularly important example for our discussion is to consider how the representation of a *wave* is affected by a Galilean transformation. Let us first consider waves in general. We imagine there is some field F that propagates through space and has the functional form

$$F = F(x - wt) \; . \tag{1.18}$$

(More generally, we should have $F = F(\mathbf{r} - \mathbf{w}t)$, where $\mathbf{r}$ is a general displacement in three dimensions; we focus on the 1-dimensional limit for simplicity.) The field F depends on the specific physics of the wave under consideration: it could be the pressure of a sound wave, or the height of a water wave, or the displacement from equilibrium of an element of a spring, or ... Suffice it to say that *many* phenomena propagate as waves. The quantity w is the speed with which the wave propagates; its value also depends on the specific physics of the system under consideration.

A wave of this form satisfies the differential equation

$$\frac{\partial^2 F}{\partial t^2} - w^2 \frac{\partial^2 F}{\partial x^2} = 0 \; . \tag{1.19}$$

Equation (1.19) is known as the *wave equation*.

Suppose that t and x in Eq. (1.19) are quantities as measured by the class. How does the wave behave according to the professor? On an upcoming problem set, you will examine this problem by applying a Galilean transformation to the wave equation. You will find that the wave equation changes such that $w \to w - v$. In other words, if the class describes the wave as

$$F = F(x_C - wt_C) , \tag{1.20}$$

then the professor describes the wave as

$$F = F\left(x_P - (w - v)t_P\right) . \tag{1.21}$$

This is as expected given our discussion of how velocities transform in Galilean relativity.

Equations (1.20) and (1.21) have a very important consequence: they tell us that there is a particular, special IRF in which the wave speed is w. This is the "rest frame" of the medium that supports the wave. For example, for a water wave, the wave's speed is as measured in the frame in which the water does not flow.

At the end of the 19th century, everything that we discussed here was quite well understood. In particular, "natural philosophers" (which includes what we more or less think of as physicists today) of this time period had studied many wave phenomena, and all of them were of this form: a wave was a disturbance that propagated in some kind of medium, and the "natural" wave speed corresponded to the rest frame of that medium.

This lasted until Maxwell formulated the equations of electrodynamics that bear his name. Then things got interesting.

Nature and Nature's laws lay hid in night:
God said "Let Newton be!" And all was light.
Alexander Pope (1688 – 1744)

It did not last: The Devil howling "Ho!
Let Einstein be!" restored the status quo.
John Collings Squire (1884 – 1958)