First Year Special Relativity – Lecture 1
Introduction, concepts and classical results

Mitesh Patel, 12th May 2022

1 In this lecture

- Concepts of active and passive transformations of coordinates;
- Physical laws in different coordinate systems;
- Galilean (pre-Relativity) transformations.

2 Introduction

For the material in this lecture, the Galilean transformations are covered in Young and Freedman, Sec. 3.5 and in McCall in Sec. 5.2.

Special Relativity, as everyone knows, is a theory introduced by Albert Einstein. The original publication was in German; the title can be translated as "Electrodynamics of moving bodies". This title indicates that the theory is closely tied to electromagnetism. When it was published in 1905, the theory was originally just called Relativity and was so named because it describes how physical objects behave when moving *relative* to an observer at fixed velocities. Hence, it does not really cover objects which accelerate, so handling these situations can be tricky. Indeed, eleven years after his original theory of Relativity, in 1916 Einstein came up with a more general theory which also included acceleration. The latter is now called General Relativity and is basically the fundamental theory of gravity, while the original theory has been renamed as Special Relativity, as it is a special case, *i.e.* when gravitational fields (and all other forces which cause accelerations) can be neglected, so all objects have fixed velocities.

We will only study Special Relativity in this course. The mathematics is actually quite easy and contains nothing beyond GCSE-level; there is no calculus, for example. The difficulty is understanding the mind-bending concepts. In contrast, General Relativity is mathematically very advanced, which is why it is a fourth year option course and well beyond these lectures.

3 Rotations

As we will see in later lectures, Relativity involves 'transformations' of coordinates. Rotations are another specific case of transformations; they move positions in space from one point to another. Rotations should be familiar to you and they have a lot in common with the relativistic transformations we will meet later. Hence we can introduce some of the concepts involved in Relativity first using rotations. To simplify the mathematics, we will only consider rotations around the z axis, which therefore mix x and y into each other. Also, for simplicity, we will also only consider rotations around the origin $x = y = 0$.

An 'active' rotation is where we consider an object to be rotated in a fixed coordinate system, often called a coordinate 'frame'. This type of rotation corresponds physically to moving the object. For any given point on the object, like a corner, the x and y values change to new values x' and y' according to

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

where the rotation is by angle ϕ defined relative to the x axis as in standard spherical or cylindrical coordinates. Note that because we rotate around the origin, the point $x = y = 0$

transforms to $x' = y' = 0$ for any rotation angle. It is intuitively clear that to rotate back to the starting point, we need to rotate by the same angle but in the opposite direction. To prove this mathematically, we need to invert the 2×2 matrix in the standard way. Firstly, the determinant Δ is given by

$$\Delta = (\cos\phi) \times (\cos\phi) - (\sin\phi) \times (-\sin\phi) = \cos^2\phi + \sin^2\phi = 1$$

and, since $\cos(-\phi) = \cos(\phi)$ while $\sin(-\phi) = -\sin(\phi)$, the matrix equation can be inverted by multiplying on both sides by the inverse matrix to give

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(-\phi) & -\sin(-\phi) \\ \sin(-\phi) & \cos(-\phi) \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

The inverse matrix is seen to be a rotation matrix but with an angle $-\phi$.

There is also a 'passive' rotation, where the coordinate frame is rotated, rather than the object itself. This means the object is completely unchanged. Because the coordinate frame changes, then although the vector position of a corner is unchanged, it appears to have different values for its x and y components in the rotated coordinate frame. It is again intuitively clear that a passive rotation of the coordinate frame by ϕ changes the coordinates x and y in the same way as an active rotation of an object by $-\phi$. Hence, a passive rotation is given by

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Fig. 1 compares the two types of rotation.

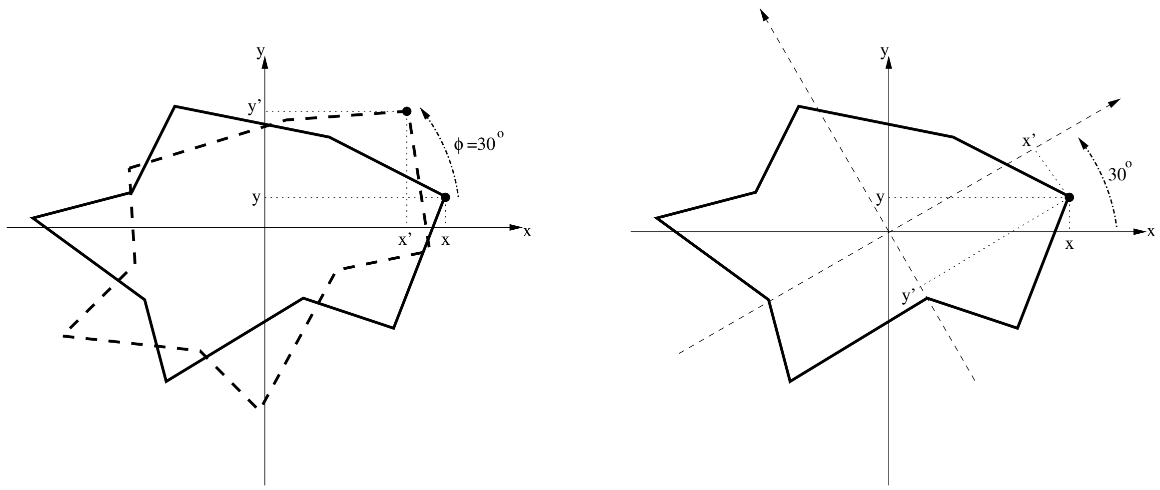

Figure 1: Left: Active rotation by 30° of an object in a fixed coordinate system. Right: Passive rotation by 30° of the coordinate system around a fixed object.

It is a fundamental assumption of physics that space is 'isotropic', which means it is the same in all orientations and has no 'special' direction. Another way to say this is that we believe there is a perfect symmetry in all directions of space.

Consider someone, who we will generally call an 'observer', does an experiment. Afterwards they do an active rotation of the equipment and repeat the experiment. As long as all objects which influence the experiment are included in the move, then the isotropic symmetry of space means that the observer should get the same result the second time. This means all physical laws must still be valid after a rotation. Note, we have one observer and do the experiment twice.

We can consider the same experiment equipment but instead do a passive rotation. To make it effectively identical to the previous case, the passive rotation can be chosen to be the negative angle of the active rotation. Here we do the experiment once, but we have two observers, one in each of the two coordinate frames, with one using the original coordinate frame and another using the rotated frame. It is clear that they will see the same outcome and agree on the result, and hence will deduce the same laws of physics from the experiment. (However, this again only works because the physical laws don't depend on a specific coordinate, e.g. x.) Therefore, when doing these comparisons, we can think in terms of either an active or a passive rotation and as long as there is a symmetry, it doesn't matter which. In Relativity, we will usually think in terms of passive transformations and multiple observers.

There is one more critical concept which arises from the symmetry of space. If physical laws are valid in all frames, then which coordinate frame to use is arbitrary. Any frame is as good as any other frame; no coordinate frame is 'better' than the others. Hence, it is meaningless to think of a 'true' frame which defines $\phi = 0$ because there is no way to define such a frame. Hence, it is impossible for a physics experiment to be able to tell its 'true' ϕ orientation because physical laws work the same for any rotation. Only relative changes in ϕ can have any effect.

4 Invariant and covariant

Let's think about what 'agreement' of the rotated experiment results means in more detail. If the two observers only measure quantities which do not change under rotations, such as mass, temperature, voltage, etc., then it is clear they will write down exactly the same measurements and hence will agree. However, not all physical quantities have this property. They may also measure position, velocity, electric field, etc., which change direction under a rotation and so the two observers will get different values for the three coordinates of each of these quantities.

Given that their values are not identical, how do they check the physical laws are still the same? Clearly, all the quantities in the first group are scalars and the second are vectors. Another word for a scalar is an 'invariant' and this is used more often in Relativity textbooks. For non-native English speakers, in the word 'invariant', the starting 'in' means 'not' and 'variant' (like 'vary') means changable, so an 'invariant' is a quantity that is 'not changable'. Here it means it does not change under rotations. If we only had invariants in physics, then knowing we have the same physical laws for both observers would be trivial. However, it is more complicated; one important example is Newton's second law $\vec{F} = m\vec{a}$, which involves a scalar m but also two vectors $\vec{F}$ and $\vec{a}$. The critical point is that this equation will work in all observer coordinate frames because the two vectors rotate in the same way, *i.e.* both sides of the equation change consistently. Equations like Newton's second law are called 'covariant', where 'co' here means 'together' (e.g. 'cooperate' means 'work together' and 'cohabit' means 'live together'). Hence 'covariant' means both sides of the equation 'change together', *i.e.* in the same way, and hence if Newton's second law holds in one frame, it will hold in all rotation frames. Generalising this, then since we believe all rotation frames are equivalent, all physical laws must be able to be expressed in covariant equations, which have scalars or vectors (or more complicated structures) matching on both sides. Basically, using vectors correctly does this job for us. Fundamentally, scalars and vectors are really *defined* by their transformations under rotations. You already know not to form equations which have different types of quantities on each side, e.g. $\vec{F} = m|\vec{a}|$ must be wrong as it has a vector on one side and a scalar on the other, so it is not covariant.

Note, there are various operations we can do with vectors, such as dot products $\vec{a}.\vec{b}$, including of a vector with itself $\vec{a}.\vec{a} = |\vec{a}|^2$, which give invariants under rotations. Since the vector length (squared) is unchanged when rotated, then by definition, the curve along which each point moves under an active rotation will be the curve with a constant value of length, *i.e.* a circle.

5 Inertial frames and Galilean transformations

There is a fundamental concept underlying Relativity which is equivalent to the rotational symmetry; all physical laws must be valid when transforming from one coordinate frame to another. However, in Relativity, the transformation is not due to a change of angle, but to a change of velocity. Although it is not initially very obvious, the relativistic transformation of coordinates for moving frames is mathematically similar (although clearly not identical) to the mathematics for transformations between rotated frames given above. In rotations, for simplicity we only considered rotations around the z axis. Similarly, when considering inertial frames, we will only consider velocity changes along the x axis.

It is important to remember that Relativity is the special case of frames moving only with fixed velocities. Such coordinate frames are called 'inertial frames' because Newton's first law (which still holds even with Relativity) says the inertia of an object means it will continue to be at rest, or move with a fixed velocity, if no force is applied to it. An inertial coordinate frame is one in which this statement holds so the frame itself must also be at rest or moving with a constant velocity.

We should first see how this works in classical physics, *i.e.* before Relativity was discovered. How do objects appear to observers in different inertial frames? Classically, it seems obvious that if an object has a speed u along the $+x$ axis in one inertial frame, then in a different inertial frame moving with a constant speed v along the $+x$ axis, it will appear to have a speed

$$u' = u - v$$

The position in x of the object in the first frame is related to u by $u = dx/dt$ and similarly in the second frame $u' = dx'/dt$. Hence, integrating the relationship above (and remembering that v is a constant) gives

$$x' = x - vt + C$$

for some constant of integration C. In the same way as we always rotated around the origin, in this course we will always do transformations which make the origins of the two coordinate systems agree at $t = 0$. Hence to get $x' = x = 0$ at $t = 0$ clearly means we take $C = 0$. To make the transformations more similar to what we will see later, they can be written as

$$t' = t, \qquad x' = x - vt, \qquad y' = y, \qquad z' = z$$

These are called the passive Galilean transformations, named after Galileo. (The active transformations simply change the sign of v.) Before the 19^{th} Century, all known physical laws held in all moving frames if the Galilean transformations were used. In particular, Newton's laws are covariant under the Galilean transformations; for instance, energy and momentum will have different values in different inertial frames but in any frame, an observer will measure their values will be conserved.

The Galilean transformations can be contrasted with the passive rotation transformations written out in a similar format

$$t' = t, \qquad x' = x\cos\phi + y\sin\phi, \qquad y' = y\cos\phi - x\sin\phi, \qquad z' = z$$

Critically, note the Galilean transformations do not look particularly like a rotation between x and t because there is no x term in t'. We will discuss the modified transformations used in Relativity in Lecture 4.

The Galilean transformations are intuitive as they correspond to our everyday experience. However, it turns out they are not completely correct. Instead, they are only an approximation to the exact relativistic transformation laws. Many of the conceptual problems in Relativity arise from the fact that our intuition expects the Galilean transformations and so it can initially be hard to understand the consequences of the exact relativistic transformations.

6 Revision questions

- What is the difference between an active and passive rotation?
- In cases where the physical laws don't depend on a specific direction, what do we expect for the results of some experiment after such rotations?
- Is the rotated or original coordinate frame more correct to make calculations in?
- In general, will an observers interpretation of events depend on which frame they observe an experiment from?
- What is a covariant equation?
- What is an inertial frame?
- What are the Galilean transforms and how will energy and momentum transform between inertial frames?