1 This writeup

Contains material of first few lectures: Some introductory comments and concepts, preliminary concepts and techniques, an overview (partial listing with brief discussions) of the topics to be considered in this module.

We will consider one of Galileo's thought experiments, and have Einstein's special relativity forced upon us.

2 What is relativity about?

Relativity is about observing physical phenomena from different frames of reference, in particular when the frames are at motion with respect to each other.

2.1 Changing perspectives

Part of what we study as relativity is just what the word sounds like — how things appear different, depending on your reference frame.

However, the underlying principle is *invariance*, i.e., that physical laws should not depend on reference frame. It could have been called *theory of invariance* instead of *theory of relativity*.

We are all aware that how we see things depends on our perspective, our frame of reference. Popular wisdom says that things look different from the top than from the bottom: The general manager and the lowest-ranked worker have utterly different perceptions about the company they work for.

In this semester, we will study the physics analog of differing perceptions due to differing points of view. A 'frame of reference' will have a more precise meaning for us than in the 'social-science' example above. When we describe space with coordinates (x, y, z), we have chosen an origin and three directions; this is our reference frame. (It also means that we have chosen units of distance, i.e., a scale factor, but we will usually take this for granted.) You can mostly think about a reference frame as a coordinate system. Your description of space depends on the coordinate system you have chosen. Your description is 'relative' to a choice of frame.

What if you decide to (or are forced to) work with a different coordinate system? One in which the same point is described with coordinate system (x', y', z')? How are the two descriptions of the same point related? Relativity is concerned with transformations between such descriptions. The different coordinate systems ('frames') might be displaced or rotated with respect to each other, but more importantly, we will consider frames that are at motion relative to each other. In addition to the spatial coordinates, we will also be forced to consider changes of the time measured from two frames.

2.2 Why is it called *special* relativity?

The name "special" relativity means that we will specialize to cases where the two frames do not have any relatve acceleration.

Having studied Newtonian physics (classical mechanics) previously, it will be straightforward for us to write down the relationship between how an event (happening at some instant at a defined location) is measured from one frame and how it is measured relative to another, moving at constant velocity with respect to the first. These are known as the Galilean transformations. We will then find that these transformations are inconsistent with electrodynamics. To fix this problem, we will have to correct the Galilean transformations. This is what Einstein did in his famous 1905 paper, thus introducing special relativity.

Reference frames with relative acceleration are considered seriously in general relativity, which Einstein developed in the decade after 1905. Consideration of accelerated frames leads to an understanding of gravity. We will deal with acceleration occassionally during our study of SR, but in a limited manner.

2.3 Moved, turned, flipped, boosted

In the next few sections, we will consider describing space (and, including time, describing spacetime) from two different coordinate systems. The simplest case is when the two frames have axes pointing in the same directions but have separate origins — one frame is *translated* with respect to another.

Then we will consider one frame *rotated* with respect to another. Rotation turns out to be deeply related to special relativity — we will have to keep returning to frame rotations several times later in the semester.

We will then consider *reflections* of coordinate axes, which turn right-handed coordinate systems to left-handed ones.

We can also think of one frame moving at a certain velocity with respect to another. Such a transformation is called a *boost*; one frame is boosted with respect to the other.

The term **boost** is not so far from the sense it carries in non-scientific (everyday) English. A boost will mean a transformation from one frame to another which is moving at constant velocity relative to the first. Considering boosts, we will be faced with an incompatibility between Newtonian

(Galilean) mechanics and the theory of electromagnetism. The effort to resolve this inconsistency is what leads to special relativity.

The term 'boost' is used in physics also outside of relativity. Those of you who have studied quantum mechanics might know (or can check) that, if the single particle wavefunction $\psi(\vec{r})$ has momentum expectation value $\vec{q}$, then the wavefunction $\psi(\vec{r})e^{i\vec{k}\cdot\vec{r}}$ has momentum expectation value $\vec{q}+\hbar\vec{k}$. The factor $e^{i\vec{k}\cdot\vec{r}}$ has the effect of **boosting** a wavefunction.

2.4 Inertial frames

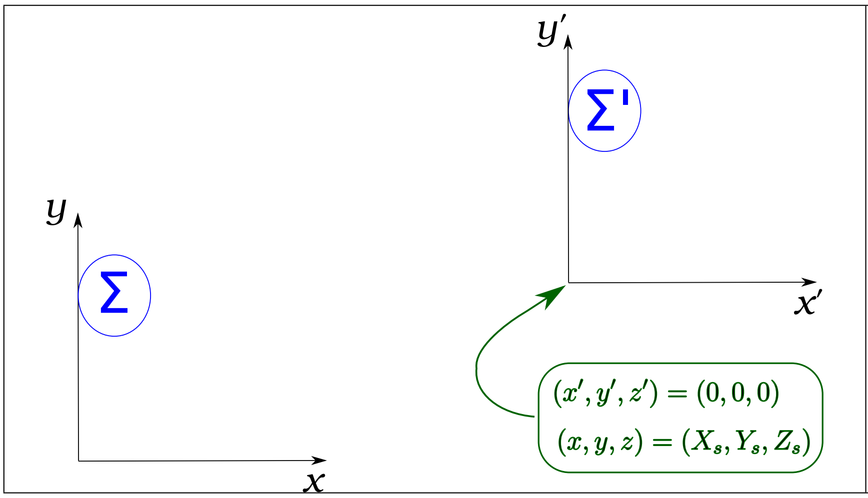
The reference frames we will deal with will be **inertial frames**.

- Inertial frames are those frames in which there are **no fictitious forces**. In other words, there are no apparent forces due to the acceleration of the frame itself.
- Examples of fictitious forces: centrifugal forces, Coriolis forces, the forward force that you 'feel' when your car is decelerating.
- If a frame is an intertial frame, than any other inertial frame is either at rest with respect to the first, or moving at constant velocity with respect to the first.
- In other words, if Σ and Σ' are inertial frames, then Σ' may be rotated or translated relative to Σ, or it may be moving at constant velocity wit respect to Σ (boosted relative to Σ.)

  However, Σ' may NOT be accelerating or rotating with respect to Σ.

3 Translations

We consider two frames Σ and Σ' which are related by translation; the axes point in the same directions, but are not aligned because the origins are displaced from one another. The origin of Σ' is at location (X_s, Y_s, Z_s) when measured in the Σ frame, although they are of course at location $(0,0,0)$ when measured in the Σ' frame. The Σ' frame is translated by (X_s, Y_s, Z_s) compared to the Σ frame, or equivalently, the Σ frame is translated by $(-X_s, -Y_s, -Z_s)$ relative to the Σ' frame. (The subscript s stands for 'shift'.)



Frame Σ' is displaced (translated) relative to frame Σ.

3.1 Nicely linear, but not homogenous

The same point in space is now measured from the two frames, and found to have coordinates (x, y, z) and (x', y', z'). You can show geometrically that they would be related by

$$x' = x - X_s, \quad y' = y - Y_s, \quad z' = z - Z_s.$$

(If not completely obvious, please draw the coordinates of a point in both frames.)

The transformation is linear, and so one might want to write it as a matrix equation:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} X_s \\ Y_s \\ Z_s \end{pmatrix}$$

The transformation is so simple that the transformation matrix is a unit matrix. However, it is not described by the (trivial) transformation matrix alone — an extra additive vector term is required to describe the translation. In other words, the linear set of transformation equations is not *homogeneous*.

This is a bit unpleasant — it would be nicer to be able to describe transformations through a matrix alone. In fact, in most of this book we will consider homogeneous transformations by fixing translations to be zero.

3.2 Including time

We might also want to compare time measurements done in the two frames. Assuming that the clocks used for the two frames are synchronized, it seems utterly sensible and intuitive that the times measured for some event in the two frames are the same:

$$t' = t.$$

We can put the transformation of time into our matrix equation:

$$\begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} - \begin{pmatrix} X_s \\ Y_s \\ Z_s \\ 0 \end{pmatrix}$$

You might have heard or read — somewhere in popular culture — that time is the "fourth dimension". The reason is that events used to be commonly defined like this by an earlier generation of physicists and relativity textbooks — by attaching the time variable at the end of three spatial variables. However, popular culture is a little behind in this case — there has been a cultural shift in physics convention. The majority of textbooks (but not all), especially more advanced textbooks, now place the time variable *before* the spatial variables, so that it is actually the "zeroth dimension" and not the fourth dimension:

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0 \\ X_s \\ Y_s \\ Z_s \end{pmatrix}$$

It may seem unfortunate that the fourth or zeroth variable has different units (dimensions) compared to the other, spatial variables. This somewhat offensive feature can be corrected by multiplying t with a constant having the dimensions of speed. In special relativity, there is a fundamental speed — the speed of light — which plays this corrective role. We will add this factor in later chapters.

We have considered shifts in the spatial origin, but (by having the clocks of the two frames synchronized) avoided a shift in time. If the clocks are not synchronized but run at identical rates, we would have $t' = t + T_s$,

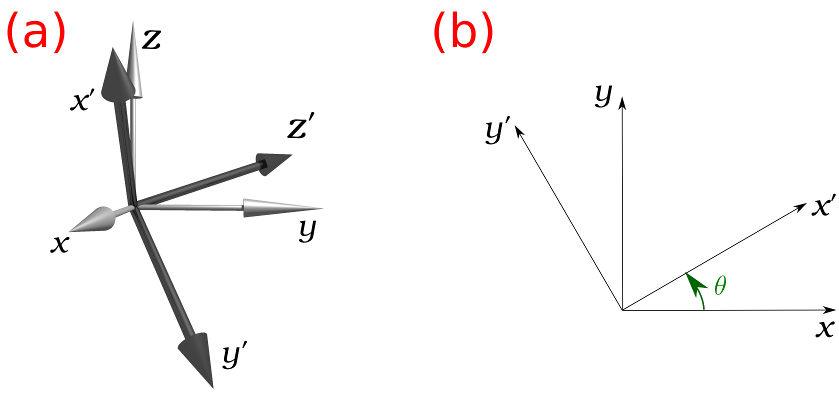


Figure 1: (a) Frame Σ' is rotated with respect to frame Σ. The two frames share the same origin. (b) Special case of 2D rotation. The z and z' axes (not shown) coincide — they both point out of the plane.

and the shift vector would be changed:

$$-\begin{pmatrix} 0 \\ X_s \\ Y_s \\ Z_s \end{pmatrix} \longrightarrow -\begin{pmatrix} T_s \\ X_s \\ Y_s \\ Z_s \end{pmatrix}$$

4 Rotation

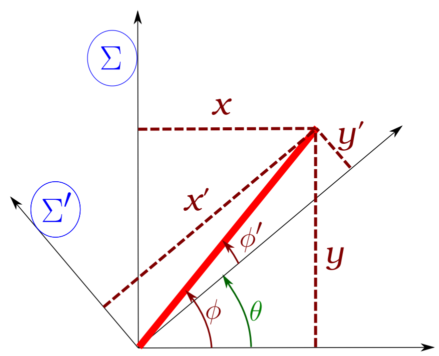
We now consider the case of one frame rotated with respect to the other. Rotation will turn out to be an important part of Lorentz transformations, so it is a good idea to be familiar with it. An example rotation of a coordinate frame is shown in Figure 1(a). This is a general case of 3D rotation, with all three axes of frame Σ' (x', y', z') pointing in different directions from the three axes of frame Σ (x, y, z).

One could describe the general 3D rotation using three angles, e.g., the Euler angles. Since the algebra gets cumbersome, we also consider the simpler case of 2D rotations: we keep the z axis unchanged, and consider rotations about the z-axis, i.e., rotations of the x-y plane, as shown in Figure 1(b). The transformation can then be described by a single parameter, θ, representing the angle of counterclockwise rotation taking Σ to Σ'.

4.1 Rotation matrices

We now consider the coordinates of a point (components of a displacement vector) in the two frames, (x, y, z) and (x', y', z').

Here is a derivation of the 2D rotation formula, using the polar coordinates of the point in the two frames, (r, ϕ) and (r', ϕ'). Since the origins of the two frames coincide and the point itself is unchanged by the transformation, $r = r'$: both r and r' represent the length of the thick line in the figure. Using $x = r\cos\phi$ and $y = r\sin\phi$, we get $x' = r'\cos\phi' = r\cos(\phi - \theta) = r\cos\phi\cos\theta + r\sin\phi\sin\theta = x\cos\theta + y\sin\theta$. The expression for y' is an exercise.



From the geometry: $\phi' = \phi - \theta$.

4.1.1 2D rotations

We first work out the simpler case of rotations around the z axis (2D rotations). From the geometry, it is clear that the z coordinate of a point is the same in the two frames. (As the z and z' axes coincide and the origins are the same, a displacement vector has the same projection onto z as onto z'.) From the example drawing in the table, it is also clear that the x, y components of a point can be very different from the x', y' components. The transformation between the two descriptions of the same point is

$$x' = x\cos\theta + y\sin\theta \quad (1)$$
$$y' = -x\sin\theta + y\cos\theta \quad (2)$$
$$z' = z \quad (3)$$

This is a rather important transformation equation that you will meet and use throughout your physics studies. Unless you are not familiar with it, you are well advised to derive and re-derive this until it is thoroughly digested. It can be derived in various ways: two derivations are discussed in the table and another in Chapter 2 (page *).

Frame Σ' is obtained by rotating frame Σ by angle θ. We could equally well think of frame Σ as being obtained by rotating frame Σ' by angle $-\theta$. Hence we should obtain the coordinates (x', y', z') from the coordinates (x, y, z) by using the same transformation but replacing $\theta \to -\theta$. Thus we must have

$$x = x'\cos\theta - y'\sin\theta \quad (4)$$
$$y = +x'\sin\theta + y'\cos\theta \quad (5)$$
$$z = z' \quad (6)$$

Are equations (1) and (4) compatible? Starting from (1) and solving for (x', y', z'), one ends up with (4). It is strongly recommended that you work this out.

The transformation equations (1) can be written as a matrix equation:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_z(\theta) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

The notation $R_z()$ indicates that this is a rotation around the z axis. Of course, if it is agreed that only rotations around the z axis are being considered, so that the z coordinate remains unchanged in the transformation, then one may omit the z variable and write the transformation as

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

The transformation is pleasantly linear. Since the origin is not shifted, it is also *homogenous*, i.e., there is no additive constant, in contrast to what we saw for translations. The transformation is thus described by a single matrix, $R_z(\theta)$. Once the axis of rotation is fixed, this is a one-parameter family of 2×2 matrices $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, the parameter being the angle θ. This set of matrices is known as $SO(2)$, and is one of the fundamental matrix classes in physics.

4.1.2 Successive rotations $\equiv$ multiplying matrices

One expects that a rotation by angle θ followed by a rotation by angle $-\theta$ (around the same axis) returns the coordinate frame to it's original state. Representing the displacement vector (position coordinates) of a point with resepect to frame Σ as

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

we infer that $R_z(-\theta)\left[R_z(\theta)X\right]$ should be equal to X, as the final frame after the two rotations is identical to Σ. Thus

$$X = R_z(-\theta)\left[R_z(\theta)X\right] = \left[R_z(-\theta)R_z(\theta)\right]X$$

where we have used the associativity of matrix multiplication, $A(BC) = (AB)C$. The only way for X to be equal to $\left[R_z(-\theta)R_z(\theta)\right]X$ is to have

$$R_z(-\theta)R_z(+\theta) = \mathbb{I}$$

where $\mathbb{I}$ is the identity matrix. Thus the inverse of $R_z(\theta)$ should be $R_z(-\theta)$.

We can check that this is true by taking the matrices $R_z(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ and $R_z(-\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ and multiplying them explicitly, to get the 2×2 unit matrix. We have suppressed the z coordinate here, but you get the same result if you retain the z coordinate and mutliply; the result is the 3×3 unit matrix.

We have learnt an important lesson here: the result of applying two rotations successively is a rotation represented by multiplying the matrices representing the individual rotations. The order of matrices may seem unintuitive — the rotation to be applied first appears to the right.

Continuing with 2D rotations around the same axis (say the z axis), let us now consider two successive rotations with arbitrary angles. The result will be independent of the order of the two rotations. (If this is not obvious, take a solid object and, pretending it is a coordinate frame, try rotating successively around the same axis.) We also expect that the result is a rotation around the same axis by an angle which is the sum of the two angles.

Both these features are expressed neatly using matrix representations of rotations:

$$R_z(\theta)R_z(\phi) \;=\; R_z(\phi)R_z(\theta) \;=\; R_z(\theta+\phi).$$

In other words, any two R_z matrices *commute*, and their product is an R_z matrix with the sum of angles as argument. You should check this by multiplying the explicit matrix expressions for $R_z(\theta)$ and $R_z(\phi)$, and using the trigonometric identities for $\sin(\theta+\phi)$ and $\cos(\theta+\phi)$.

> Two matrices P and Q are said to **commute** if $PQ = QP$. Commutation (or lack thereof) is important also in quantum mechanics (for operators) and in group theory (for group elements).

4.1.3 3D rotations

A general rotation in 3D can always be described by a succession of 2D rotations around one of the axes. E.g., any 3D rotation can be parametrized by Euler angles (ϕ, θ, ψ) such that the full rotation is obtained by a rotation by ϕ around the z axis, followed by a rotation by θ around the new x axis, followed by a rotation by ψ around the new z axis. The effect of such a rotation on the coordinates of a point is

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \;=\; R_z(\psi)R_x(\theta)R_z(\phi) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Thus a general 3D rotation is also a linear transformation, described by a 3×3 matrix

$$\begin{aligned} \mathcal{R} &= R_z(\psi) R_x(\theta) R_z(\phi) \\ &= \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Note the order in which the matrices appear: the first transformation to be applied appears on the right, not on the left. The result of the matrix multiplication is cumbersome to write out explicitly and not particularly enlightening.

Unlike two rotations around the same axis, two arbitrary 3D rotations do not generally commute, for example, you can show by explicit matrix multiplication that the rotations $R_z(\theta)$ and $R_x(\phi)$ do not commute. If this result is not physically obvious, take a non-symmetric object and try two rotations around different axes in both orders.

4.1.4 Including time

Of course, in special relativity we will eventually want to include time. For static frame rotations, the time measurements in the two frames are identical. Regarding time as the zeroth dimension as we did in the previous section, a rotation can be represented as

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathcal{R} & \\ 0 & & & \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

where $\mathcal{R}$ is a 3×3 orthogonal rotation matrix. As an example, a rotation around the z axis would be represented by the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We will have to include rotations into relativistic spacetime (Lorentz) transformations, so this is an actual preview of Lorentz transformations.

4.2 Orthogonality and length invariance

Rotation matrices are **orthogonal**, meaning that the transpose of a rotation matrix is its inverse: $\mathcal{R}^{\mathrm{T}}\mathcal{R} = \mathcal{R}\mathcal{R}^{\mathrm{T}} = \mathbb{I}$. You can show this explicitly for the trigonometric expressions above for, e.g., $R_z(\theta)$ and $R_z(\phi)$. Since an arbitrary rotation is the product of three such 2D rotations and the product of orthogonal matrices is orthogonal (please show yourself the last statement), any rotation matrix is orthogonal.

Physically, orthogonality of rotation matrices reflects the fact that the length of the displacement vector stays unchanged (invariant) under rotation. We have already used the invariance of the distance between a point and the origin, in noting that the polar coordinates (r, ϕ) and (r, ϕ') for a given point have the same radial value in the two frames. The length of the vector *is invariant*, or it is *an invariant*. (The word invariant can be either an adjective or a noun.)

Defining the displacement (position) vector as

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

the length-squared of this vector is

$$x^2 + y^2 + z^2 \; = \; \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \; = \; X^{\mathrm{T}} X.$$

Invariance implies that $X' = \mathcal{R}X$ has the same length as X, i.e.,

$$X^{\mathrm{T}}X \; = \; (X')^{\mathrm{T}}X' \; = \; (\mathcal{R}X)^{\mathrm{T}}(\mathcal{R}X) \; = \; \left(X^{\mathrm{T}}\mathcal{R}^{\mathrm{T}}\right)(\mathcal{R}X) \; = \; X^{\mathrm{T}}\left(\mathcal{R}^{\mathrm{T}}\mathcal{R}\right)X$$

Since $X^{\mathrm{T}}X = X^{\mathrm{T}}\left(\mathcal{R}^{\mathrm{T}}\mathcal{R}\right)X$ for *any* vector X, we must have

$$\mathcal{R}^{\mathrm{T}}\mathcal{R} = \mathbb{I} \qquad \boxed{\text{Rotation matrices are orthogonal}}$$

The orthogonality property, representing the invariance of ordinary Euclidean lengths, will have close analogs in special relativity theory, where the invariance of a similar quantity will lead to a property of Lorentz transformations similar to orthogonality. Stay tuned!

4.2.1 Unit Determinant

Orthogonal matrices have determinant $+1$ or -1.

If A is an orthogonal matrix, then $1 = \det(\mathbb{I}) = \det(A^{\mathsf{T}}A) = \det(A^{\mathsf{T}})\det(A) = \det(A)\det(A) = \det(A)^2$, so that $\det(A) = \pm 1$.

Since rotation matrices are orthogonal, their determinant needs to have unit magnitude. Negative determinants are only relevant when there is a

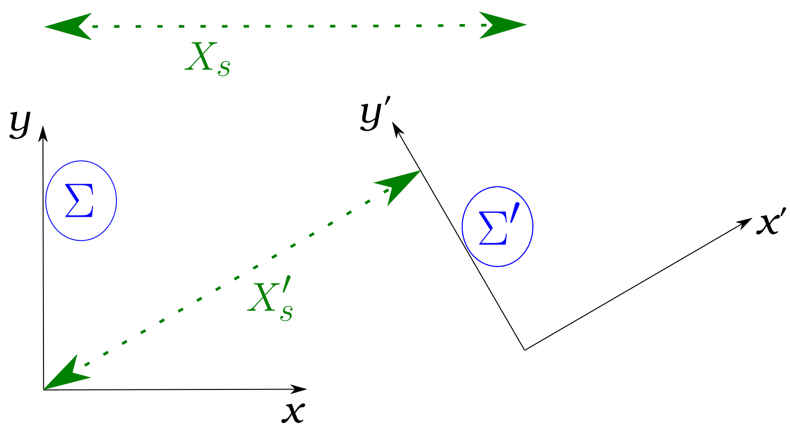


Figure 2: Frame Σ' is rotated **and** translated relative to frame Σ.

reflection of the coordinate frame (next section) in addition to rotation, so for pure rotations $\det(\mathcal{R}) = 1$.

Pure rotations (without reflection) are known as *proper* rotations, i.e., proper rotations have positive determinant $+1$

4.3 Active or passive?

Euler angles are often treated in classical mechanics modules. In that context, they are used more commonly to describe the *actual* rotation of a rigid body. Transformation equations describing such physical rotations (or other translations/reflections) of an object are called active transformations. This is to distinguish from frame transformations, where the transformation is between descriptions of the same object in two different frames. The object itself does not physically change under this transformation, and exists independently of the choice of reference frame.

We have considered, and will be considering, passive transformations only. The point whose coordinates we are transforming does not change, and exists independently of our choice of reference frame.

4.4 Moved as well as turned

What if the frame is translated as well as rotated?

If the frame is first translated so that the origin is at point (X_s, Y_S, Z_s) relative to the initial frame Σ, and then subjected to a rotation described by matrix $\mathcal{R}$, then the coordinates of a point (x, y, z) are changed to

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathcal{R} \begin{pmatrix} x - X_s \\ y - Y_s \\ z - Z_s \end{pmatrix}$$

Alternatively, the transformation might be described as a rotation performed first, and then a translation by (X'_s, Y'_S, Z'_s):

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathcal{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} X'_s \\ Y'_s \\ Z'_s \end{pmatrix}$$

The two descriptions of the amount of translation, (X_s, Y_S, Z_s) in the original directions and (X'_s, Y'_S, Z'_s) in the rotated directions, are related to each other by the same rotation matrix $\mathcal{R}$.

These transformations are linear but not homogeneous, and are thus not described by a single 3×3 matrix. One can construct 4×4 matrix to describe such transformations. The corresponding non-homogeneous transformations in relativity are known as Poincaré transformations; we will treat them near the end of semester.

Unlike rotations preserving the origin, the displacement vector of a point will not have the same magnitude in the transformed frame. In other words, the displacement from the origin is not an invariant. However, if we consider the displacement between two different physical points, its magnitude is an invariant. If we consider positions $\vec{r}_1 = (x_1, y_1, z_1)$ and $\vec{r}_2 = (x_2, y_2, z_2)$, then neither $|\vec{r}_1|$ nor $|\vec{r}_2|$ are preserved under the transformation, as the origin is shifted. However, the magnitude

$$|\vec{r}_1 - \vec{r}_2| = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$$

is invariant under the transformation considered.

5 Reflection

After having moved and turned reference frames, it's time to flip them.

Reflection will turn a right-handed frame into a left-handed frame. If the unit vectors in the original right-handed frame Σ are denoted $\hat{e}_x$, $\hat{e}_y$, $\hat{e}_z$, then the right-handed orientation of the three orthogonal axes implies $\hat{e}_x \times \hat{e}_y = \hat{e}_z$. The unit vectors of the frame Σ', obtained after reflection, will obey $\hat{e}'_x \times \hat{e}'_y = -\hat{e}'_z$ because this Σ' has a left-handed coordinate system.

The simplest example is to reflect around one of the three planes formed by the axes, so that two of the axes remain unchanged, and one coordinate changes sign. In the example shown in Figure 3(a), a point with coordinates (x, y, z) in frame Σ will have coordinates $(x', y', z') = (x, y, -z)$ in the reflected frame. The transformation is

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} .$$

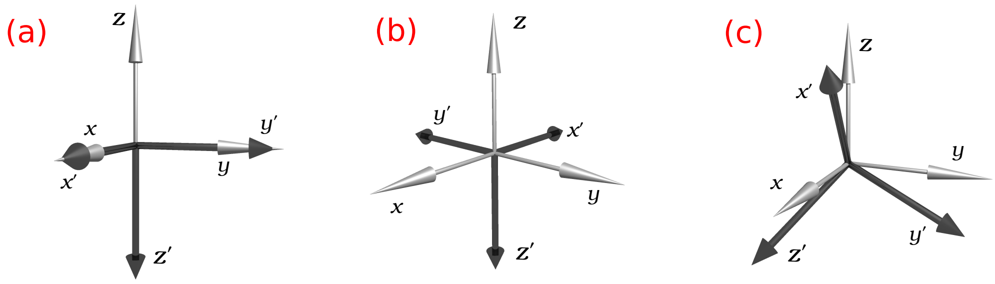


Figure 3: (a) Frame Σ' is obtained by reflecting frame Σ around the x-y plane. (b) Now all three axes of frame Σ' point opposite to the corresponding axis of frame Σ. (c) A reflection plus a rotation.

In the transformation shown in Figure 3(b), all three axes of the reflected frame Σ' are directed opposite to the original frame Σ. Such a transformation can be generated by three successive reflections around the three planes, or alternatively, a reflection around the x-y plane followed by a rotation around the z or z' axis by angle π. The transformation equation is

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = P \begin{pmatrix} x \\ y \\ z \end{pmatrix} .$$

The transformation matrix that flips signs of all spatial coordinates is known as the parity transformation. Note that both the transformation matrices above have determinant -1.

Figure 3(c) shows a general rotation combined with a reflection. or with an odd number of reflections. The reason we know that an odd number of reflections must be involved is that the resulting frame is left-handed. We could represent this transformation as

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = P\mathcal{R}_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathcal{R}_B P \begin{pmatrix} x \\ y \\ z \end{pmatrix} .$$

i.e., a pure rotation followed by reflection(s), or reflection(s) followed by a pure rotation. The exact rotation describing the transformation depends on whether the reflection(s) are performed first or last, i.e., $\mathcal{R}_A \neq \mathcal{R}_B$.

5.1 Flipping is improper

A general rotation combined with a handedness-changing reflection is known as an *improper* rotation.

Since this class of transformations does not shift the origin, the magnitude of the displacement vector of a physical point remains invariant. Invariance of the displacement magnitude, as we have seen, forces the transformation matrices to have determinant $+1$ or -1. Transformations with determinant $+1$ represent pure rotations or *proper* rotations. Those with determinant -1 are parity-flipping (handedness-changing) *improper* rotations, and include an odd number of reflections.

If $\mathcal{R}$ has determinant $+1$, it is clear that $P\mathcal{R}$ or $\mathcal{R}P$ has determinant -1.

Generally, considering this type of transformation is not helpful in the study of physical rotations, so one often restricts to the set of proper rotations. So why are we bothering with them at all? The reason is that it is sometimes considered convenient to *define* the class of rotations as those which satisfy $\mathcal{R}^{\mathrm{T}}\mathcal{R} = \mathbb{I}$, i.e., the set of orthogonal matrices. But this class of transformations include reflections. Rotations in this general sense also includes flips. It then makes sense to divide rotations into two groups, parity-preserving and parity-violating.

A similar issue will arise in the study of relativistic (Lorentz) transformations.

6 Boosts

6.1 Discussion: "Are we moving?" and "Is physics invariant?" and inertial frames

Without visual clues, can we tell whether our world (our 'frame') is moving?

After having studied Newtonian physics for some time, it is perhaps intuitive to you that, if you are in a frame moving at constant velocity, the laws of mechanics are independent of the speed of motion of the frame. You only expect to 'feel' the motion of your frame if the frame is accelerating. We will refer to this basic idea as the *principle of relativity*. The idea is commonly attributed to Galileo. Based on this idea, the Galilean/Newtonian transformations for boosts will be developed in the next section. We will discuss the idea qualitatively further below.

The principle of relativity was discussed by Galileo Galilei in his 1632 book *Dialogue Concerning the Two Chief World Systems*. He discusses a situation (a thought experiment) in which the experience of fish in a fishbowl (and flies and butterflies and humans) within the hold of a moving ship is compared with the experience of humans and animals in a stationary ship. A variant is shown in Figure 4.

The principle might be regarded as an extension of Newton's law, which says that an object moving with constant velocity does not change it's velocity unless a force is applied on it. An object stationary within the moving vehicle (e.g., the fishbowl) is moving at constant velocity (relative to the ground/house); hence it will not change its state of motion unless forced. In other words, within the moving vehicle, it will appear as if the

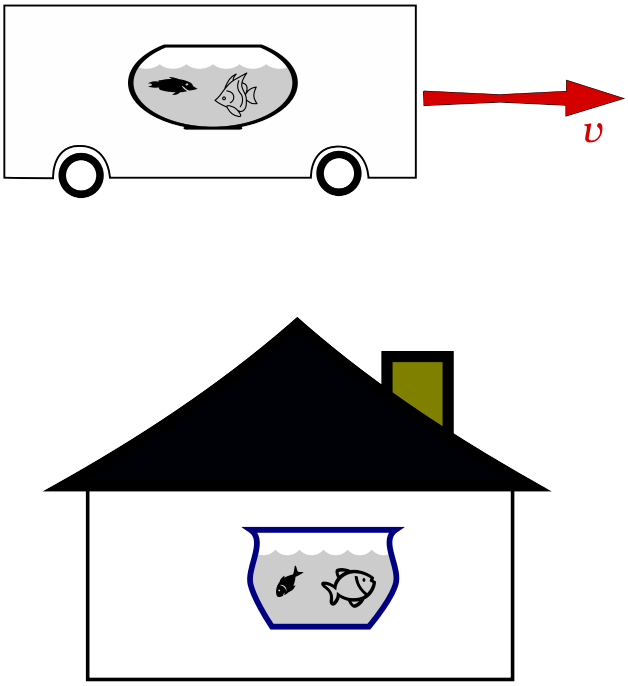


Figure 4: Are the natural laws experienced in a constant-velocity vehicle different from those experienced in a stationary building?
If the relative velocity is truly constant, is there a way of telling whether my frame is moving, without looking out through a window?
If the velocity is perfectly constant, do the fish in the vehicle feel the motion of the vehicle?

fishbowl remains stationary, and feels no force, i.e., does not feel the motion of the vehicle.

Do we usually know when our vehicle is moving? Yes, but according to the principle of relativity that is due to additional clues, not due to steady (constant-velocity) motion. For example, when in a train compartment or riding in the back of a truck, we might be able to look outside and see the landscape moving. Even we don't look outside, we hear the roar of the engine and might feel the vibrations caused by the engine. Even if these were negligible, the truck or train velocity is never quite completely constant — the tracks and roads are rarely completely straight and never completely smooth. We feel deviations from steady motion, not the steady motion itself. This is also true for extremely high speeds. When I sit in an airplane, I am aware of the noise and vibration of the engine, but if I don't look outside I might not know if we are already flying or still waiting to take off. (This happens to me often — I doze off shortly after being seated in an airplane and after waking up may not know the state of motion.) The only time one is physically aware of airplane motion is when the motion is not smooth and the velocity is changing (turbulence, take-off, bad landing).

The discussion above is tricky because of the need to separate different ways of awareness of motion. However, there is one common instance of high-speed motion that we can readily all agree that we do not 'feel' — the motion of earth as it hurtles around the sun, or the motion of our part of earth as it spins around the earth axis. These motions are complicated by the fact that they include some acceleration. The reason we do not feel these accelerations are because they appear as minuscule perturbations to the acceleration due to earth's gravity. We will omit detailed discussions of acceleration and gravity. (There will be enough of that if you ever get to studying general relativity.) The point made here is that our well-being is not affected by the fact that we are rushing through space at high speed. The Galilean principle of relativity does seem to be consistent with this everyday experience.

We seem to be converging toward the idea that, as far as physical laws are concerned, frames of references moving at constant velocity with respect to each other are equivalent. But is one of these frames special or more important than the rest? Thinking of the two fishbowls in Figure 4, we might be biased toward thinking that a frame fixed with respect to the surface of the earth is somehow special, as most of our personal experiences are gained at low speeds relative to the earth surface. However, a little empathy reminds us that for someone living on another planet for an extended period, a frame fixed with respect to the surface of that planet will seem most natural. So the most reasonable supposition seems to be that all constant-velocity frames are equally valid in terms of physical laws and none of them are special. The viewpoint of the fish in the house (that the vehicle carrying other fish is whizzing by rightwards with speed v) is no more 'correct' than the viewpoint of the fish in the vehicle (that the house and the landscape are all moving leftwards with speed v).

Whose viewpoint is the correct viewpoint?

The idea that all constant-velocity frames are equally important or correct takes some getting used to. Traveling by train, many of us have had the following experience, which makes the idea a bit more palatable. Sitting near a train window at or near a station, one's view through the window can be filled by the train on the neighbouring set of tracks. Looking up from a book or waking up from a nap, I have seen a neighboring train move relative to me, and I was unable to tell whether it was my train moving, or the other train moving, or both. The confusion is usually resolved when one catches a glimpse of things fixed to the earth surface (the platform, buildings etc), e.g., by looking through the other window. I seem to be unable to feel whether I am in the stationary (relative to the earth surface) or moving train. There seems to be nothing special, as far as the laws of mechanics are concerned, about frames fixed relative to the earth surface.

For our ordinary modes of transport on earth (car, train, plane or boat), the experience of the principle of relativity is complicated by issues like 'looking out' and engine noise and turbulence; and our discussion above has been complicated accordingly. So, we will often talk about observers traveling is spacecrafts or rockets. It is not difficult to imagine a spacecraft moving at great speeds through space without a running engine — according to Newton's law, no engine is needed if there is no friction. Also, if there are two space ships (far from earth) traveling at different velocities, we have no natural bias toward one frame over the other, so the idea that all constant-velocity frames are equivalent is easier to accept.

The extended discussion above suggests that the laws of *mechanics* are the same in all constant-velocity frames, i.e., that the laws of mechanics are invariant under boosts. If you throw a ball while in a constant-velocity train compartment, it will move along a parabolic trajectory. Forces will cause change of momentum, and angular momentum will change in response to torques, exactly as if you were doing these experiments on the ground. We have not really *proved* that mechanics works the same way in constant-velocity frames, just like we never proved Newton's laws. However, the principle of relativity for mechanics seems just as reasonable as Newton's laws — in fact, seems like a minor extension of Newton's laws.

What about electricity & magnetism?

Now, in Galileo's time, electromagnetism was not very well understood. Are the laws of electromagnetism also equally valid in all constant-velocity frames? If we did Ampere's and Faraday's experiments in a high-velocity airplane, would we infer the same laws? Do like (unlike) charges repel (attract) the same way on a spacecraft traveling rapidly from earth to Neptune, as they do on a spacecraft floating idly on the way? In a stationary laboratory, we know that a changing magnetic field creates a curly electric field (induction, Faraday's law, $\nabla \times \vec{E} = -\partial_t \vec{B}$). Does the same happen in a laboratory on a high-speed train? In other words, are the laws of electromagnetism also invariant under boosts?

Inferring the invariance or non-invariance of electromagnetism from our experience is more tricky. Most of us have not performed electromagnetic experiments while moving at high speeds. But it would seem a reasonable guess that if the laws of mechanics are invariant under boosts, then so are the laws of electromagnetism. Clearly, this deserves further thought.

6.2 Inertial frames, acceleration and gravity

I discussed constant-velocity frames, taking some care to separate effects of steady (constant-velocity) motion from effects of acceleration of frames. I now try to make more precise the notion of constant-velocity frames, which we will call inertial frames.

One reason to avoid relative acceleration of frames is that we do not want to deal with 'artificial' of 'fictitious' forces that are only due to the acceleration of frames. This is the backward acceleration we feel relative to our vehicle when the vehicle accelerates forward. Another example is the centrifugal force: this is the force you feel if you are attached to a rotating frame. The Coriolis force is another fictitious force due to rotations of the frame, which you might learn about in a semester on (relatively advanced) classcial mechanics, or if you ever delve into physical oceanography, or to explain Foucault's pendulum.

This suggests the following definition: an inertial frame is one in which stationary or constant-velocity objects are not subjected to acceleration due to the location or motion of the frame. A particle in motion in such a frame will continue at constant velocity unless acted upon by a non-fictitious force. Of course, the notion of an inertial frame is an idealization (like many ideas in physics). A frame attached to a spaceship with its engine turned off, far away from planetary/stellar/astronomical bodies, is a good approximation to an inertial frame.

In our previous discussion, we talked about frames attached to earth or to trucks/trains/airplanes near earth. To think of these as inertial frames, we have to neglect the intrinsic and orbital rotations of earth, which is reasonable as these cause barely perceptible fictitious forces. However, there is a more serious problem — due to gravity, objects feel a downward acceleration in any one of these frames. One way to deal with this problem is to only compare frames with the same gravity, and then concentrate motions only in the directions perpendicular to the gravity direction. As far as motion in the horizontal directions are concerned, we can ignore gravity. We are hiding some problems 'under the rug' here. First, it is not *a priori* obvious that we can apply whatever we derive about inertial frames selectively to the directions perpendicular to the downward direction. Second, in general relativity the distinction between gravitational and fictitious forces is questioned and actually abandoned. However, we will be able to build a consistent theory of special relativity (transformation between inertial frames) which works also in the presence of gravity.

Having defined inertial frames, we will proceed with the assumption that there are an infinite number of inertial frames, moving at constant velocity relative to each other.

6.3 Events and spacetime

A most basic concept in relativity is that of an *event*.

In everyday usage, an event is something happening at a particular location at a particular time. We use the word in a sharper sense: an event is specified by a position – a *point* in three-dimensional (3D) space, and an *instant* – a sharp value of the time variable.

$$\text{An event:} \qquad (\mathbf{r}, t) \quad \text{or} \quad (\vec{r}, t)$$

I will denote ordinary (3D) vectors either with bodface ($\mathbf{r}$) or with an arrow on top ($\vec{r}$). Now a 3D vector (sometimes to be called a 3-vector) is represented by a 3-tuple, a collection of 3 real numbers. Hence an event is thus represented by a 4-tuple, a collection of 4 real numbers. Events are examples of what will be later called 4-vectors.

$$(x, y, z, t) \qquad \text{Later also written as:} \qquad (ct, x, y, z) = (x^0, x^1, x^2, x^3)$$

Here (x, y, z) are the Euclidean coordinates of the spatial location ($\mathbf{r}$) of the event. As discussed previously, we will mostly use time as the zeroth rather than the fourth dimension, and multiplying time by a constant c (with units of speed) makes the four dimensions have the same units.

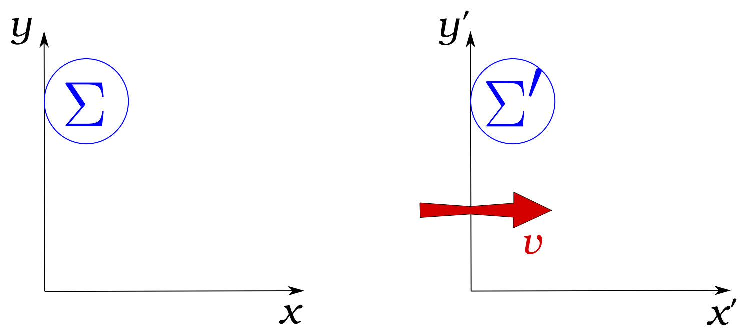
Just as the collection of all points (x, y, z) is called *space*, so the collection of all events, (ct, x, y, z) is called *spacetime*. Spacetime, i.e., the collection of all events, is also known as *Minkowski space*.

The notation (x^0, x^1, x^2, x^3) uses superscripts rather than subscripts to index the dimensions.

6.4 The 'standard' boost and homogeneity

We will repeatedly consider two inertial frames (Σ and Σ') in relative motion, thinking about how observers in the two frames measure the same event(s). It is therefore convenient to consider a 'standard' orientation of the relative velocity and of the axes of the two frames.

We choose the two sets of reference axes to be aligned at some point of time, i.e. the x', y' and z' axes are aligned at some moment respectively to the x, y and z axes respectively. The direction of relative motion is taken to be the common x and x' direction. At the moment when the reference axes are aligned, the clocks of the two frames are synchronized and the time is set to be $t = 0$ and $t' = 0$.



The origin of Σ is at $(0, 0, 0)$ as measured from the Σ frame, and $(-vt', 0, 0)$ from the Σ' frame. The origin of Σ' is $(+vt, 0, 0)$ from Σ frame; $(0, 0, 0)$ from Σ' frame.

This configuration of two inertial frames will be referred to as the "standard configuration", and the boost from Σ to Σ' will be referred to as the "standard boost". This configuration/boost is widely used in most textbooks on special relativity; hence it is 'standard'. Occassionally one also sees texts using relative motion in the common z direction.

It is important that we have taken the origins of the two frames to coincide at the common zero time $t = t' = 0$. This will ensure that we can express our transformations (whether Galilean transformations or the corrected Lorentz transformations) as 4×4 matrices, i.e., that the transformations are homogeneous. Recall that we had a similar issue with spatial (3D) transformations. The Galilean and Lorentz transformations will be four-dimensional analogs of rotation *without any translation*.

When we consider successive boosts in different directions, it will not be possible to have the relative velocities all in the same direction, but we will still take all the origins to be coincident at the common zero time.

7 The Galilean transformations

We will now write down the boost transformations as they were understood before Einstein, i.e., the Galilean transformations.

Consider an event measured to be (ct, x, y, z) from Σ and (ct', x', y', z') from Σ'. For the sake of homogeneity, we only consider cases where the clocks on the two frames are synchronized to be $t = t' = 0$ at the moment when the origins coincide.

In Newtonian or Galilean physics, time is an absolute concept whose flow does not depend on the observer; therefore the time measurements in the two frames remain synchronized: $t' = t$.

7.1 Standard boost

We first specialize to the standard boost: the relative motion is in the common x, x' direction. At the instant $t = t'$, the origin of frame Σ' is at location $(+vt, 0, 0)$ relative to Σ. Hence the transformation of coordinates is

$$x' = x - vt = x - \left(\frac{v}{c}\right) ct \qquad y' = y \qquad z' = z \tag{7}$$

Putting these all together, the transformation (written as a matrix equation) is

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v/c & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{8}$$

You can of course also find (ct, x, y, z) in terms of (ct', x', y', z'), e.g., by algebraically inverting Eq. (7). You could also invert the transformation matrix in Eq.(8), which is equivalent but a bit of overkill in terms of effort.

Explicitly inverting matrices is clumsy and generally avoided whenever possible. Also for numerical (computer) calculations, inverting matrices is algorithmically expensive and rarely necessary.

Avoiding the algebra, the inverse transformation can be obtained by replacing $v \to -v$, as frame Σ moves with velocity $-v$ relative to frame Σ':

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ v/c & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} \tag{9}$$

You should check that multiplying the GT matrix in Eq. (8) and the inverse GT matrix in Eq. (9), in either order, gives you the 4×4 unit matrix.

7.2 Beyond the standard GT — the general GT

Imagine a textbook choosing the standard configuration to involve relative motion in the common z, z' direction. The GT matrix is then

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v/c & 0 & 0 & 1 \end{pmatrix}.$$

More generally, if the relative velocity of frame Σ' is $\vec{v} = (v_x, v_y, v_z)$ with respect to frame Σ (while the directions remain the same), then the transformation matrix is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_x/c & 1 & 0 & 0 \\ -v_y/c & 0 & 1 & 0 \\ -v_z/c & 0 & 0 & 1 \end{pmatrix}.$$

The most general GT while preserving homogeneity is to allow the frame Σ' to be rotated with respect to frame Σ as well as in relative motion. Their origins are coincident at the common zero time, but the axes are not. The general GT matrix is

$$\begin{pmatrix} 1 & 0 \quad 0 \quad 0 \\ -v_x/c & \\ -v_y/c & \mathcal{R} \\ -v_z/c & \end{pmatrix}$$

where $\mathcal{R}$ is an orthogonal 3×3 matrix, representing a rotation (and possibly a reflection, if we want to include improper twists). Here v_x, v_y, v_z are the velocity components in the rotated directions.

The common feature of the GT matrices for these different cases is the $(1, 0, 0, 0)$ on the top row. This represents the pre-Einstein notion that time is absolute ($ct' = ct$).

The factor c is used here to give the time component the same units as the spatial components. It's not really necessary here, but might be useful to get used to, because it appears naturally when we consider Lorentz transformations.
Admittedly it makes things look complicated. You can set $c = 1$ if you want. Then you interpret as transformations of (t, x, y, z) rather than of (ct, x, y, z). Or as time being measured in units that make $c = 1$.

7.3 Adding velocities by multiplying matrices

If an object moves with velocity $\vec{u}$ relative to frame Σ, what is its velocity relative to frame Σ'? For simplicity, let us first concentrate on the velocity being in the same direction as the relative velocity between the frames: in standard configuration, $\vec{v} = (v, 0, 0)$ and $\vec{u} = (u, 0, 0)$.

The answer ($u' = u - v$) is probably intuitively obvious to you, but we will derive it carefully. This is because the answer will be wholly non-intuitive when we move from Galilean transformations to Lorentz transformations.

Consider (ct, x, y, z) now to be not just a single event, but a family of events: with the spatial part (x, y, z) being the location of a particle at time t. Since the particle moves with velocity $(u, 0, 0)$, we have $x = ut + x_0$. The same interpretation is applied to (ct', x', y', z'), so that $x' = u't' + x'_0 = u't + x'_0$.

The GT tells us that $x' = x - vt$. Taking the time derivative with respect to t (or equivalently t'),

$$\frac{dx'}{dt} = \frac{dx}{dt} - v \qquad \Longrightarrow \qquad u' = u - v$$

This can also be derived in terms of the GT matrices. Imagine a frame $\tilde{\Sigma}$ attached to the particle and in standard configuration with the frames Σ and Σ', i.e., the axes of $\tilde{\Sigma}$ are aligned with the axes of Σ and Σ' at the common zero time. The transformation of events in going from Σ' to $\tilde{\Sigma}$ is given by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -u'/c & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

as $\tilde{\Sigma}$ moves in the common x', $\tilde{x}$ direction with speed u', relative to Σ'. Meanwhile the transformation in going from Σ to Σ' is that appearing in Eq. (8). The transformation from Σ to $\tilde{\Sigma}$ is thus obtained by combining the two boosts:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -u'/c & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v/c & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -(u'+v)/c & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The composition of two Galiliean boosts, unsurprisingly, is also a Galilean boost, with velocity given by the sum of the two. frame $\tilde{\Sigma}$, and hence the particle, moves with speed $u = u' + v$ relative to frame Σ.

This principle of velocity addition can be stated simply in terms of frames without reference to a moving particle: consider frame Σ' moving with speed v_1 relative to frame Σ in the common x, x' direction, and frame $\tilde{\Sigma}$ moving with speed v_2 relative to frame Σ' in the common x', $\tilde{x}$ direction. Then by matrix multiplication one finds, as above, that $\tilde{\Sigma}$ is moving with speed $v_1 + v_2$ relative to frame Σ in the common x, $\tilde{x}$ direction. Since the transverse directions play no role, we can suppress them and write

$$\begin{pmatrix} 1 & 0 \\ -v_1/c & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -v_2/c & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -(v_1 + v_2)/c & 1 \end{pmatrix}.$$

Addition of velocity, of course, works also for velocities not in the same direction. Consider frame Σ' moving with speed $\vec{v_1} = (v_{1x}, v_{1y}, v_{1z})$ relative to frame Σ, and frame $\tilde{\Sigma}$ moving with speed $\vec{v_2} = (v_{2x}, v_{2y}, v_{2z})$ relative to frame Σ'. Composition of the two Galilean boosts gives

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_{1x}/c & 1 & 0 & 0 \\ -v_{1y}/c & 0 & 1 & 0 \\ -v_{1z}/c & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_{2x}/c & 1 & 0 & 0 \\ -v_{2y}/c & 0 & 1 & 0 \\ -v_{2z}/c & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -(v_{1x} + v_{2x})/c & 1 & 0 & 0 \\ -(v_{1y} + v_{2y})/c & 0 & 1 & 0 \\ -(v_{1z} + v_{2z})/c & 0 & 0 & 1 \end{pmatrix},$$

i.e., a GT with velocity $\vec{v_1} + \vec{v_2}$. The Galilean transformations are thus consistent with Euclidean vector addition, which is no surprise.

We have represented addition of velocities in terms of matrix multiplication. Now matrix multiplication is not generally commutative, but in this case it has to be, as the addition of ordinary vectors (such as velocity) is commutative. You should check explicitly that each of the matrix multiplication operations in this subsection give the same answer if performed in the reverse order.

7.4 Consistent with Newtonian mechanics ☺

The principle of relativity implies that the laws of mechanics should look the same in every inertial frame. Let us examine whether the basic law of Newtonian mechanics, $F = ma$, is invariant under the Galilean transformations.

If a particle has velocity $\vec{u}$ relative to Σ and $\vec{u}'$ relative to Σ', then we have found

$$\vec{u}' = \vec{u} - \vec{v}$$

where $\vec{v}$ is the constant velocity of Σ' sith respect to Σ. Therefore the acceleration of the particle is the same in both frames:

$$\frac{d}{dt}\vec{u}' = \frac{d}{dt}\vec{u} - \frac{d}{dt}\vec{v} = \frac{d}{dt}\vec{u}$$

Therefore, Newton's law, Force = mass $\times$ acceleration, holds in one inertial frame if it holds in any other. We are assuming that the mass of particles does not depend on frame, and by definition the physical force is the same in all inertial frames.

Thus the Galilean transformations are consistent with Newtonian mechanics.

7.5 Electrodynamics causes problems ☹

Unfortunately, it turns out that the GT are completely incompatible with electromagnetism.

7.5.1 Speed of light

In particular, if a light pulse moves in the common direction x, x' and its speed is measured to be c from frame Σ, then the speed will be $c - v$ measured from frame Σ', according to GT.

Recall: light appears as traveling wave solutions from Maxwell's equations: in MKS units, the speed of light is

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$$

The constants ϵ_0 and μ_0 are related to electrostatics and magnetostatics. If one did electrostatic experiments on frame Σ and on frame Σ', e.g., in two trains moving at different speeds, it's reasonable to expect the same electrostatic phenomena, i.e., the same values of ϵ_0, and similarly the same values of μ_0, independent of which intertial frame the experiments are performed in. However, this implies that the speed of light is the same in all, which contradicts the GT.

Of course, maybe it's possible that electromagnetic experiments depend on the initial frame, and that the speed of light really does depend on the frame. This was assumed by some in the late 19th century, and was settled (negatively) by the famous Michelson-Morley experiment: the speed of light does NOT depend on the inertial frame.

7.5.2 The look of Maxwell's equations

If the equations of electromagnetism (Maxwell's equations and the resulting wave equation) are transformed according to the GT, their forms change.

(We will not show this in this class, because it is tedious to work out.)

We really do not expect the form of fundamental equations of nature to depend on the frame of reference. The fact that Maxwell's equations are not invariant under Galilean transformations, is what led Einstein to propose corrections to the GT.

8 Einstein's postulates

From the previous section, it appears that either Electromagnetism must be frame-dependent, or the GT's must be corrected. Faced with this choice in 1905, Einstein 'fixed' the GT's such that the laws of nature remain frame-invariant. This leads to the Lorentz transformations.

8.1 The two postulates

Einstein based his work on the following two postulates:

> 1. First postulate (principle of relativity): The laws of physics are the same in all inertial frames of reference.
>
> 2. Second postulate (invariance of c): The speed of light in free space has the same value c in all inertial frames of reference.

The "laws of physics" include electromagnetic phenomena, conservation of momenta and energy, etc.

"Light" here means not just visible light, but all electromagnetic radiation, including, e.g., X-rays and gamma-rays. The speed of EM radiation in vacuum is independent of everything, including the speed of the source from which it is emitted.

There is some redundancy in the listing of the postulates, in that the constancy of the speed of light might be regarded as a 'law of physics', since $c = 1/\sqrt{\epsilon_0 \mu_0}$ and the electromagnetic constants are results of physics experiments. Thus the second postulate may be regarded as being contained in the first.

The development of special relativity is physics, not mathematics, even though we call these "postulates". Various additional assumptions, usually implicit, are involved in the development of the subsequent theory.

9 The 'standard' Lorentz boost

Frames Σ' and Σ are coincident at time $t = t' = 0$ and their relative motion is in the common x, x' direction.

Eventually we will refer to relativistic transformations as Lorentz transformations. Lorentz transformations in general can include relative rotations between the two frames in addition to a boost in any direction. The standard configuration contains two frames with a relative boost only in the common x, x' direction. We refer to this transformation as the standard Lorentz boost.

For the standard configuration (representing a pure boost in the x direction), the Lorentz transformation between time and space coordinates is

$$
\begin{aligned}
ct' &= \gamma_v \left(ct - \frac{v}{c} x \right) && (10) \\
x' &= \gamma_v \left(x - \frac{v}{c} ct \right) && (11) \\
y' &= y && (12) \\
z' &= z && (13)
\end{aligned}
$$

We have introduced the all-important Lorentz factor

$$\gamma_v = \frac{1}{\sqrt{1 - v^2/c^2}}$$

with which we will become very familiar as we progress through this book (semester).