(Very) basic introduction to special relativity

Math 3560 Groups and symmetry, Fall 2014

Raul Gomez

Cornell University

Newton's reference frames

Figure : Isaac Newton

Newton's three laws

1. Law of inertia An object either is at rest or moves at a constant velocity, unless acted upon by an external force.

Newton's three laws

1. Law of inertia An object either is at rest or moves at a constant velocity, unless acted upon by an external force.
2. Law of motion The acceleration of a body is directly proportional to the force acting on the body, and inversely proportional to its mass. In mathematical terms,

$$\mathbf{F} = m\mathbf{a}.$$

Newton's three laws

1. Law of inertia An object either is at rest or moves at a constant velocity, unless acted upon by an external force.
2. Law of motion The acceleration of a body is directly proportional to the force acting on the body, and inversely proportional to its mass. In mathematical terms,

$$\mathbf{F} = m\mathbf{a}.$$

3. Law of action-reaction For every action there is an equal and opposite reaction.

About Newton's first law

Newton's first law says that there is a reference frame (that is there is a system of coordinates) on which Newton's first law holds.

About Newton's first law

Newton's first law says that there is a reference frame (that is there is a system of coordinates) on which Newton's first law holds.

Any such a frame is called an inertial frame.

About Newton's first law

Newton's first law says that there is a reference frame (that is there is a system of coordinates) on which Newton's first law holds.

Any such a frame is called an inertial frame.

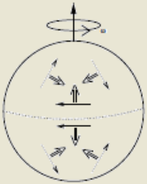
Although this statement may seem obvious, it is not clear at all how to choose such a reference frame in the universe.

Example

If we use a fixed point on earth to define a reference frame, then we would obtain an example of a non inertial frame.

Example

If we use a fixed point on earth to define a reference frame, then we would obtain an example of a non inertial frame. In this frame the centrifugal force and the Coriolis force (which is responsible for the formation of hurricanes) violate Newton's first law.



Systems of particles

Now assume that we have a system of N particles with position $\mathbf{r}_i$ and mass m_i.

Systems of particles

Now assume that we have a system of N particles with position $\mathbf{r}_i$ and mass m_i.

Then Newton's second law reads:

$$\mathbf{F}_i = m_i \, \mathbf{a}_i.$$

Systems of particles

Now assume that we have a system of N particles with position $\mathbf{r}_i$ and mass m_i.

Then Newton's second law reads:

$$\mathbf{F}_i = m_i \mathbf{a}_i.$$

Decompose $\mathbf{F}_i$ as

$$\mathbf{F}_i = \sum_{j \neq i} \mathbf{F}_{ij} + \mathbf{F}_i^{\text{ext}},$$

where $\mathbf{F}_{ij}$ is the force on the i-th particle due to the j-th particle.

Decompose $\mathbf{F}_i$ as

$$\mathbf{F}_i = \sum_{j \neq i} \mathbf{F}_{ij} + \mathbf{F}_i^{\text{ext}},$$

where $\mathbf{F}_{ij}$ is the force on the i-th particle due to the j-th particle.

Decompose $\mathbf{F}_i$ as

$$\mathbf{F}_i = \sum_{j \neq i} \mathbf{F}_{ij} + \mathbf{F}_i^{\text{ext}},$$

where $\mathbf{F}_{ij}$ is the force on the i-th particle due to the j-th particle.

Suming over all the i's we get

$$\begin{aligned} \sum_i F_i &= \sum_{i,j,j \neq i} \mathbf{F}_{ij} + \sum_i \mathbf{F}_i^{\text{ext}} \\ &= \sum_{i<j} (\mathbf{F}_{ij} + \mathbf{F}_{ji}) + \sum_i \mathbf{F}_i^{\text{ext}}. \end{aligned}$$

Decompose $\mathbf{F}_i$ as

$$\mathbf{F}_i = \sum_{j \neq i} \mathbf{F}_{ij} + \mathbf{F}_i^{\text{ext}},$$

where $\mathbf{F}_{ij}$ is the force on the i-th particle due to the j-th particle.

Suming over all the i's we get

$$\begin{aligned} \sum_i F_i &= \sum_{i,j,j \neq i} \mathbf{F}_{ij} + \sum_i \mathbf{F}_i^{\text{ext}} \\ &= \sum_{i<j} (\mathbf{F}_{ij} + \mathbf{F}_{ji}) + \sum_i \mathbf{F}_i^{\text{ext}}. \end{aligned}$$

But according to Newton's third law, $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$, hence

$$\sum_i \mathbf{F}_i = \sum_i \mathbf{F}_i^{\text{ext}} =: \mathbf{F}^{\text{ext}},$$

where $\mathbf{F}^{\text{ext}}$ is the total exterior force acting on the system.

We define the total mass of the system to be

$$M = \sum_i m_i,$$

and the center of mass to be

$$\mathbf{R} = \frac{\sum_i m_i \mathbf{r}_i}{M}.$$

We define the total mass of the system to be

$$M = \sum_i m_i,$$

and the center of mass to be

$$\mathbf{R} = \frac{\sum_i m_i \mathbf{r}_i}{M}.$$

Then, if $\mathbf{A}$ is the second derivative of $\mathbf{R}$ with respect to time, we have that

$$\mathbf{F}^{\text{ext}} = M\mathbf{A}.$$

We define the total mass of the system to be

$$M = \sum_i m_i,$$

and the center of mass to be

$$\mathbf{R} = \frac{\sum_i m_i \mathbf{r}_i}{M}.$$

Then, if $\mathbf{A}$ is the second derivative of $\mathbf{R}$ with respect to time, we have that

$$\mathbf{F}^{\text{ext}} = M\mathbf{A}.$$

In a system where there are no external forces (like the universe) this equation says that fixing the origin at the center of mass gives you an inertial reference frame.

Example

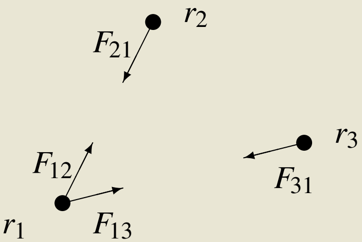
Let

$$\mathbf{r}_1 = (0,0) \quad \mathbf{r}_2 = (3,6) \quad \mathbf{r}_3 = (8,2),$$

and assume that

$$\mathbf{F}_{12} = (1,2) \quad \mathbf{F}_{13} = (2,1/2) \quad \mathbf{F}_{23} = (0,0).$$

The we obtain the following picture:



Absolute vs. relative point of view

Figure : Gottfried Leibniz

Classical Electromagnetism

Figure : James Clerk Maxwell

Maxwell equations (1862)

Let $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{J}$ be the electric field, magnetic field and current density, respectively, and let ρ be a charge distribution in $\mathbb{R}^3$.

1. Gauss law.

$$\nabla \cdot \mathbf{E} = \rho / \varepsilon_0.$$

2. Gauss law for magnetism.

$$\nabla \cdot \mathbf{B} = 0.$$

3. Fadaray Law.

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

4. Ampère circulation law (with Maxwell correction.)

$$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{J} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right).$$

where ε_0 and μ_0 are the electric and magnetic constants.

Maxwell correction to Faraday's law is critical in that it predicts the existance of electromagnetic waves.

Maxwell correction to Faraday's law is critical in that it predicts the existance of electromagnetic waves.

More precisely, if $\rho = 0$, then Maxwell equations are equivalent to the system of equations:

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} + \frac{\partial^2 \mathbf{E}}{\partial y^2} + \frac{\partial^2 \mathbf{E}}{\partial z^2} - \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0$$

and

$$\frac{\partial^2 \mathbf{B}}{\partial x^2} + \frac{\partial^2 \mathbf{B}}{\partial y^2} + \frac{\partial^2 \mathbf{B}}{\partial z^2} - \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0,$$

plus a condition relating $\mathbf{E}$ and $\mathbf{B}$.

Electromagnetic waves

Maxwell was able to solve his equations in the vacuum, and predicted the existance of electromagnetic waves.

Electromagnetic waves

Maxwell was able to solve his equations in the vacuum, and predicted the existance of electromagnetic waves.

In particular, he was able to calculate the speed at which his predicted electromagnetic waves should travel.

Electromagnetic waves

Maxwell was able to solve his equations in the vacuum, and predicted the existance of electromagnetic waves.

In particular, he was able to calculate the speed at which his predicted electromagnetic waves should travel.

The value Maxwell obtained, using the known values at the time of the constants ε_0 and μ_0, was essentially the same as the speed of light as measured at that time.

Electromagnetic waves

Maxwell was able to solve his equations in the vacuum, and predicted the existance of electromagnetic waves.

In particular, he was able to calculate the speed at which his predicted electromagnetic waves should travel.

The value Maxwell obtained, using the known values at the time of the constants ε_0 and μ_0, was essentially the same as the speed of light as measured at that time.

Based on this observation, Maxwell predicted that light was just an example of an electromagnetic wave.

Maxwell predictions created considerable problems ammong the physics community of the time. The main issues were:

Maxwell predictions created considerable problems ammong the physics community of the time. The main issues were:

1. On earth, waves always propagate through a medium, they can't exist in the vacuum.

Maxwell predictions created considerable problems ammong the physics community of the time. The main issues were:

1. On earth, waves always propagate through a medium, they can't exist in the vacuum.
2. Maxwell calculated the speed of the electromagnetic waves, but this was the speed with respect to what exactly?

Maxwell predictions created considerable problems ammong the physics community of the time. The main issues were:

1. On earth, waves always propagate through a medium, they can't exist in the vacuum.
2. Maxwell calculated the speed of the electromagnetic waves, but this was the speed with respect to what exactly?

Trying to kill two birds with one stone, they proposed the existance of an essentially indetectible substance called luminiferous aether.

Maxwell predictions created considerable problems ammong the physics community of the time. The main issues were:

1. On earth, waves always propagate through a medium, they can't exist in the vacuum.
2. Maxwell calculated the speed of the electromagnetic waves, but this was the speed with respect to what exactly?

Trying to kill two birds with one stone, they proposed the existance of an essentially indetectible substance called luminiferous aether.

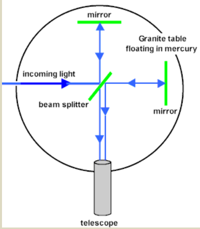
This was supposed to be the medium through which electromagnetic waves propagated. The speed calculated by Maxwell was then the speed of light with respect to this medium.

The Michelson Morley experiment

So in 1887, Albert Michelson and Edwartd Morley designed an experiment to measure the velocity of the earth with respect to the surrounding aether.

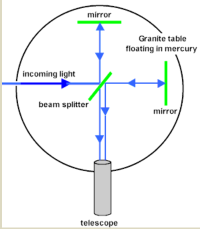
The Michelson Morley experiment

So in 1887, Albert Michelson and Edwartd Morley designed an experiment to measure the velocity of the earth with respect to the surrounding aether.



The Michelson Morley experiment

So in 1887, Albert Michelson and Edwartd Morley designed an experiment to measure the velocity of the earth with respect to the surrounding aether.



The result however, was totally unexpected. The speed of light was the same in every direction!

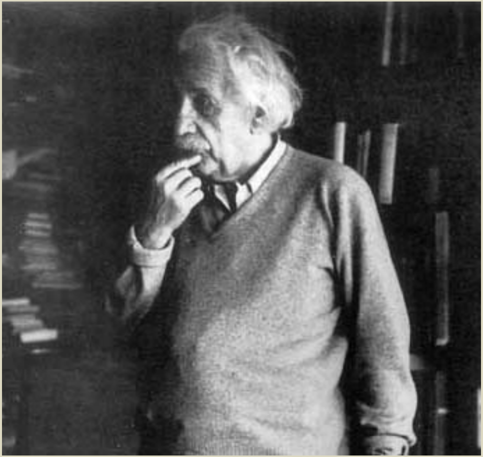


Figure : Albert Einstein

Space-time coordinates

To simplify calculations, from now on we will only consider universes with one spatial dimension.

Space-time coordinates

To simplify calculations, from now on we will only consider universes with one spatial dimension.

That is, we will consider events in the universe by a space coordinate and a time coordinate.

Space-time coordinates

To simplify calculations, from now on we will only consider universes with one spatial dimension.

That is, we will consider events in the universe by a space coordinate and a time coordinate.

Observe that if $r(t)$ represents the position of a particle at time t, then we can describe its position as a point in $\mathbb{R}^2$ using the coordinates $(t, r(t))$.

The Minkowski metric

Figure : Hermann Minkowski

The Minkowski metric

Is it possible to define a concept of distance between different events in the universe?

The Minkowski metric

Is it possible to define a concept of distance between different events in the universe?

For example, we can say:"The homecoming fireworks were lighted up last Friday at Schoellkopf Field."

The Minkowski metric

Is it possible to define a concept of distance between different events in the universe?

For example, we can say:"The homecoming fireworks were lighted up last Friday at Schoellkopf Field."

To express this event in coordinates we would need three spacial coordinates and one temporal coordinate.

The Minkowski metric

Is it possible to define a concept of distance between different events in the universe?

For example, we can say:"The homecoming fireworks were lighted up last Friday at Schoellkopf Field."

To express this event in coordinates we would need three spacial coordinates and one temporal coordinate.

But if we ask about the distance between this moment and the homecoming event, we only need two numbers: Schoellkopf Field is half a mile from here, and it's been four days since Friday.

The Minkowski metric

Is it possible to define a concept of distance between different events in the universe?

For example, we can say:"The homecoming fireworks were lighted up last Friday at Schoellkopf Field."

To express this event in coordinates we would need three spacial coordinates and one temporal coordinate.

But if we ask about the distance between this moment and the homecoming event, we only need two numbers: Schoellkopf Field is half a mile from here, and it's been four days since Friday.

Can we combine this 2 numbers to get a notion of "distance" between this two events?

It may seem natural to simply consider the time as another coordinate and use the Pythagorean theorem to compute the “distance” between events.

It may seem natural to simply consider the time as another coordinate and use the Pythagorean theorem to compute the "distance" between events.

Unfortunately this is not very useful because the laws of physics are not invariant under the action of $\mathrm{SO}(4, \mathbb{R})$.

It may seem natural to simply consider the time as another coordinate and use the Pythagorean theorem to compute the "distance" between events.

Unfortunately this is not very useful because the laws of physics are not invariant under the action of $SO(4, \mathbb{R})$.

For example, if we write Maxwell equations and then we transform everything using an element of $SO(4, \mathbb{R})$ we get an equation that looks very different.

This is clear if we look at the wave equation in one dimension:

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0.$$

This is clear if we look at the wave equation in one dimension:

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0.$$

Looking at this equation suggest a different way of calculating the "distance" between 2 events.

This is clear if we look at the wave equation in one dimension:

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0.$$

Looking at this equation suggest a different way of calculating the "distance" between 2 events.

Definition

Let $v_1 = (t_1, x_1)$ and $v_2 = (t_2, x_2)$ be two vectors in $\mathbb{R}^2$. We define the Minkowski product of v_1 and v_2 by

$$v_1 \circ v_2 = t_1 t_2 - x_1 x_2.$$

Definition

Let $v_1 = (t_1, x_1)$ and $v_2 = (t_2, x_2)$ be two vectors in $\mathbb{R}^2$. We define the Minkowski product of v_1 and v_2 by

$$v_1 \circ v_2 = t_1 t_2 - x_1 x_2.$$

Definition

Let $v_1 = (t_1, x_1)$ and $v_2 = (t_2, x_2)$ be two vectors in $\mathbb{R}^2$. We define the Minkowski product of v_1 and v_2 by

$$v_1 \circ v_2 = t_1 t_2 - x_1 x_2.$$

Using this inner product, we can define a new "norm" (or rather its square) by

$$N(v)^2 = v \circ v,$$

Definition

Let $v_1 = (t_1, x_1)$ and $v_2 = (t_2, x_2)$ be two vectors in $\mathbb{R}^2$. We define the Minkowski product of v_1 and v_2 by

$$v_1 \circ v_2 = t_1 t_2 - x_1 x_2.$$

Using this inner product, we can define a new "norm" (or rather its square) by

$$N(v)^2 = v \circ v,$$

and using this norm we can define the Minkowski "metric"

$$d(v_1, v_2)^2 = N(v_1 - v_2)^2.$$

and using this norm we can define the Minkowski “metric”

$$d(v_1, v_2)^2 = N(v_1 - v_2)^2.$$

and using this norm we can define the Minkowski "metric"

$$d(v_1, v_2)^2 = N(v_1 - v_2)^2.$$

The name "metric" here is a little misleading because the number I'm getting on the right may very well be negative, in which case, taking the square root is not really well defined.

and using this norm we can define the Minkowski "metric"

$$d(v_1, v_2)^2 = N(v_1 - v_2)^2.$$

The name "metric" here is a little misleading because the number I'm getting on the right may very well be negative, in which case, taking the square root is not really well defined.

However the Minkowski product of a vector in $\mathbb{R}^2$ with itself is well defined and is a very useful quantity as we will see soon.

Special relativity

The point of the Minkowski inner product is that Maxwell equations (and in general the equations of physics) are preserved under the action of $\mathrm{SO}(1,1)$.

Special relativity

The point of the Minkowski inner product is that Maxwell equations (and in general the equations of physics) are preserved under the action of $\mathrm{SO}(1,1)$.

That is, I can use $\mathrm{SO}(1,1)$ to change my reference frame and still get the same physics laws.

Special relativity

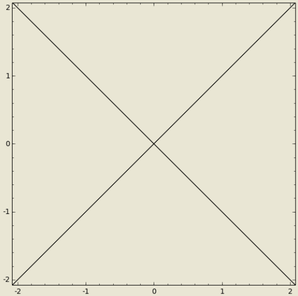
The point of the Minkowski inner product is that Maxwell equations (and in general the equations of physics) are preserved under the action of $\mathrm{SO}(1,1)$.

That is, I can use $\mathrm{SO}(1,1)$ to change my reference frame and still get the same physics laws.

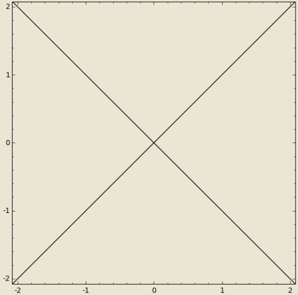
As we will see shortly, this is extremely useful.

Let's start by considering the points in $\mathbb{R}^2$ such that $v \circ v = 0$.

Let's start by considering the points in $\mathbb{R}^2$ such that $v \circ v = 0$.



Let's start by considering the points in $\mathbb{R}^2$ such that $v \circ v = 0$.



This are called the light rays, precisely because they describe the trajectory of light coming up from the origin.

The rest of the plane gets divided into two regions, the time-like directions:

$$\{v \in \mathbb{R}^2 \,|\, v \circ v > 0\},$$

and the space-like directions:

$$\{v \in \mathbb{R}^2 \,|\, v \circ v < 0\}.$$

Observe now that the set

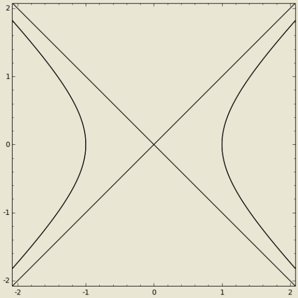
$$\{v \in \mathbb{R}^2 \,|\, v \circ v = 1\},$$

describes an hyperbola. This are the points that, in a sense, are at "distance" 1 from the origin.

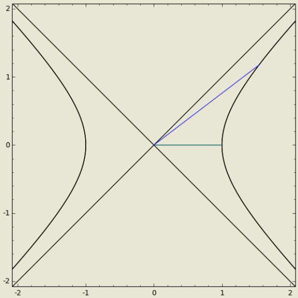
Observe now that the set

$$\{v \in \mathbb{R}^2 \,|\, v \circ v = 1\},$$

describes an hyperbola. This are the points that, in a sense, are at “distance” 1 from the origin.



Now we will consider two particles, r_1 and r_2 traveling at constant speed from the origin and whose graph in $\mathbb{R}^2$ are given by the parametric equations: $t \mapsto (t,0)$, $t \mapsto (t\cosh\alpha, t\sinh\alpha)$.



If we transform this picture using the element

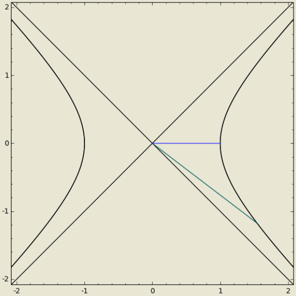
$$g(-\alpha) = \begin{bmatrix} \cosh\alpha & -\sinh\alpha \\ -\sinh\alpha & \cosh\alpha \end{bmatrix},$$

we obtain:

If we transform this picture using the element

$$g(-\alpha) = \begin{bmatrix} \cosh\alpha & -\sinh\alpha \\ -\sinh\alpha & \cosh\alpha \end{bmatrix},$$

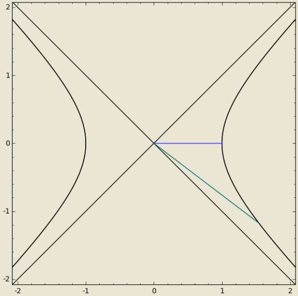
we obtain:



If we transform this picture using the element

$$g(-\alpha) = \begin{bmatrix} \cosh\alpha & -\sinh\alpha \\ -\sinh\alpha & \cosh\alpha \end{bmatrix},$$

we obtain:

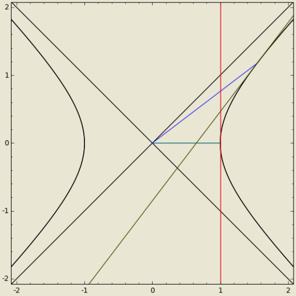


Observe that this means that for both observers their relative velocities are the same, and the speed of light is the same with respect to both observers!

How can this be possible?

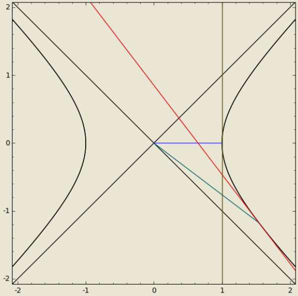
How can this be possible?
This is because the events that occur at the “same time” depend on the observer!

How can this be possible?
This is because the events that occur at the “same time” depend on the observer!



How can this be possible?

This is because the events that occur at the “same time” depend on the observer!



We will try to ilustrate this phenomenom with the following example:

We will try to ilustrate this phenomenom with the following example:

Suppose that we have an observer that is traveling inside a ship, and we have a second observer watching him through a window in the coast.

We will try to ilustrate this phenomenom with the following example:

Suppose that we have an observer that is traveling inside a ship, and we have a second observer watching him through a window in the coast.

There is a lamp in the middle of the room, and observer 1 turns the lamp on.

We will try to ilustrate this phenomenom with the following example:

Suppose that we have an observer that is traveling inside a ship, and we have a second observer watching him through a window in the coast.

There is a lamp in the middle of the room, and observer 1 turns the lamp on.

For the first observer the light will reach all the walls in the room at the same time.

We will try to ilustrate this phenomenom with the following example:

Suppose that we have an observer that is traveling inside a ship, and we have a second observer watching him through a window in the coast.

There is a lamp in the middle of the room, and observer 1 turns the lamp on.

For the first observer the light will reach all the walls in the room at the same time.

However, the observer at the coast will see the light arriving to the different walls at different times.