

Week 3: Special Relativity

Newtonian mechanics and Galilean transformations

Newton tried to present his mechanics in an axiomatic form. His *Lex Prima* (or Galilean law of inertia) states: *Each force-less mass point stays at rest or moves on a straight line at constant speed.*

Despite the axiomatic approach, the properties of the space-time structure were assumed by Newton a priori: “Absolute space” and “absolute time.” Distinguishing between straight and curved lines requires an affine structure of space, while measuring velocities relies on a metric structure that allows one to mete distances. Implicitly, Newton assumed an Euclidean structure for space, and thus the distance between two points (x_1, y_1, z_1) and (x_2, y_2, z_2) in a Cartesian coordinate system is

$$\Delta S_{12}^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \quad (1)$$

or for infinitesimal distances

$$dS^2 = dx^2 + dy^2 + dz^2. \quad (2)$$

In a Cartesian *inertial* coordinate system, Newton’s *lex prima* becomes

$$\frac{d^2x}{dt^2} = \frac{d^2y}{dt^2} = \frac{d^2z}{dt^2} = 0. \quad (3)$$

Most often, we call such a coordinate system just an *inertial frame*. Newton’s first law is not just a trivial consequence of its second one, but may be seen as practical definition of those reference frames for which his following laws are valid.

Since the symmetries of Euclidean space, translations $\vec{a}$ and rotations R , leave (2) invariant and thus also (3), all frames connected by $\vec{x}' = R\vec{x} + \vec{a}$ to an inertial frame are inertial frames too. Additionally, proper Galilean transformations $\vec{x}' = \vec{v}t$ connect inertial frames moving with relative speed $\vec{v}$,

$$\vec{x}' = R\vec{x} + \vec{v}t + \vec{a} \quad (4a)$$

$$t' = t \quad (4b)$$

Eq. (4b) corresponds to Newton’s and the “common sense” notation of an absolute time. Taking the time-derivative of Eq.(4a), we obtain the classical addition-law for velocities, $\dot{\vec{x}}' = \dot{\vec{x}} + \vec{v}$.

The *Principle of Relativity* states that *identical experiments performed in different inertial frames give identical results*. Galilean transformation keep (3) invariant, hence Newton’s first law does not allow to distinguish between different inertial frames. Before Einstein, it was thought that it applies only to mechanical experiments. N.B.: Newton himself insisted on the concept of “absolute space.”

Newton's *Lex Secunda*: Observed from an inertial reference frame, the net force on a particle is proportional to the time rate of change of its linear momentum,

$$\vec{F} = \frac{d\vec{p}}{dt} \quad (5)$$

where $\vec{p} = m_I \vec{v}$ and we call m_I the inertial mass.

Newton's law of motion is invariant against general Galilean transformations for the case of a system of particles interacting via central forces (as gravitational or Coulomb forces). If in such a case $\vec{x}(t)$ is a solution of (5), then also $\vec{x}'(t) = R\vec{x}(t) + \vec{v}t + \vec{a}$.

Newtonian gravity

Newton's gravitational law as well as the Coulomb law are examples for an instantaneous action,

$$\vec{F}(\vec{x}) = K \sum_i \frac{\vec{x} - \vec{x}_i}{|\vec{x} - \vec{x}_i|^3}. \quad (6)$$

The force $\vec{F}(\vec{x}, t)$ depends on the distance $\vec{x}(t) - \vec{x}_i(t)$ to all sources i (electric charges or masses) at the same time t , i.e. the force needs no time to be transmitted from $\vec{x}_i$ to $\vec{x}$.

The force law is invariant under changes $\mathbf{x}' = R\mathbf{x} + \mathbf{d}$, if $|\mathbf{x}'| = |\mathbf{x}|$. The translation d could be time-dependent via $d = vt$. The condition that the norm of the vector $\vec{r}$ is invariant, $\vec{x} = \vec{x}'$, restricts the possible transformation matrices R ,

$$\mathbf{x}'^T \mathbf{x}' = \mathbf{x}^T R^T R \mathbf{x} = \mathbf{x}^T \mathbf{x}, \quad (7)$$

as $R^T R = 1$. (T denotes the transposed vector or matrix). Hence rotations R and translations d keep a force law of type (6) invariant. In two dimensions, the rotation matrix R can be chosen as

$$R = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}. \quad (8)$$

The factor K is $-G_N m_g M_g$, where we introduced analog to the electric charge in the Coulomb law the "charge" m_g characterizing the strength of the gravitational force between different particles. Surprisingly, one finds $m_I = m_g$ and we can drop the index.

Since the gravitational field is conservative, we can introduce a potential ϕ via

$$\vec{F} = -m \nabla \phi \quad (9)$$

with

$$\phi(x) = -\frac{GM}{|\vec{x} - \vec{x}'|}. \quad (10)$$

Analog to the electric field $\vec{E} = -\nabla \phi$ we can introduce a gravitational field, $\vec{g} = -\nabla \phi$. We obtain then $\nabla \cdot \vec{g}(\vec{x}) = -4\pi G \rho(\vec{x})$ and as Poisson equation,

$$\Delta \phi(\vec{x}) = -4\pi G \rho(\vec{x}), \quad (11)$$

where ρ is the mass density, $\rho = dm/d^3x$.

Minkowski space

Comparing the non-relativistic Poisson equation $\Delta\phi = 0$ and Maxwell's wave equation, $(\Delta - (1/c^2)\partial_t^2)A_\mu = 0$ suggests that the latter are invariant under transformations that keep the space-time interval

$$\Delta s^2 \equiv \vec{x}^2 - c^2 t^2 = \vec{x}'^2 - c^2 t'^2 \quad (12)$$

invariant. This means that both the time and the space difference between two events measured in the coordinate system K and K' may be different, but that the difference $\Delta s^2 \equiv \vec{x}^2 - c^2 t^2$ is the same for all inertial systems. As a result, time and space should be seen not as separate entities but as components of a four-vector $\mathbf{x}$ with components $x^\mu = (\vec{x}, ct)$, $\mu = 1, 2, 3, 4$, that can be mixed by Lorentz transformations.

The distance of infinitesimally close space-time events is also called line-element ds . It is given by

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2. \quad (13)$$

More precisely, the line-element ds is defined as norm of the displacement vector

$$ds = dx^\mu \mathbf{e}_\mu \quad (14)$$

Choosing as basis the coordinate vectors to $x^\mu = (\vec{x}, ct)$, its components are

$$ds^\mu = dx^\mu = (d\vec{x}, cdt). \quad (15)$$

We compare now our physical requirement on the distance of space-time events, Eq. (13), with the general result for the scalar product of two vectors $\mathbf{a}$ and $\mathbf{b}$. If these vectors have the coordinates a^μ and b^μ in a certain basis $\mathbf{e}_\mu$, then we can write

$$\mathbf{a} \cdot \mathbf{b} = (a^\mu \mathbf{e}_\mu) \cdot (b^\nu \mathbf{e}_\nu) = a^\mu b^\nu (\mathbf{e}_\mu \cdot \mathbf{e}_\nu). \quad (16)$$

Thus we can evaluate the scalar product between any two vectors, if we know the symmetric matrix g composed of the N^2 products of the basis vectors,

$$g_{\mu\nu}(\mathbf{x}) = \mathbf{e}_\mu(\mathbf{x}) \cdot \mathbf{e}_\nu(\mathbf{x}) = g_{\nu\mu}(\mathbf{x}). \quad (17)$$

This symmetric matrix $\mathbf{g}$ is called *metric tensor*.

Applying this now for the displacement vector, we obtain

$$ds^2 = ds \cdot ds = g_{\mu\nu} dx^\mu dx^\nu \stackrel{!}{=} dx^2 + dy^2 + dz^2 - c^2 dt^2. \quad (18)$$

Hence the metric tensor $\mathbf{g}$ becomes for the special case of an Cartesian inertial frame diagonal with elements

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \equiv \eta_{\mu\nu} \quad (19)$$

Einstein's summation convention:

1. Two equal indices, of which one has to be an upper and one a lower index, imply summation. We use Latin letters for indices from one to three, $a = 1, 2, 3$, and Greek letters for indices from one to four, $\alpha = 1, 2, 3, 4$. Thus

$$a_i b^i = \sum_{i=1}^3 a_i b^i \quad \text{and} \quad a_\mu b^\mu = \sum_{\mu=1}^4 a_\mu b^\mu.$$

2. Summation are dummy indices and can be freely exchanged; the free indices of the LHS and RHS of an equation have to agree. Hence

$$8 = a_i b^i = c_{ij} d^{ij} = c_{ik} d^{ik}$$

is okay, while

$$a_\mu = b^\mu \quad \text{or} \quad a^\mu = b^\nu \quad \text{or} \quad a_i^i = c_{ij} d_j^i = w_{ik} s^{ikl}$$

compares apples to oranges.

3. Vector components with lower and upper indices have to be distinguished, because—as we will learn later—they are connected by the metric tensor as follows:

$$a_\mu = g_{\mu\nu} a^\nu.$$

Since $\mathbf{g} \neq \mathbf{1}$, the components differ in general.

4. Form the differential of a function $f(\vec{x})$. Then

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz = \frac{\partial f}{\partial x^i} dx^i \equiv \partial_i dx^i.$$

Thus an upper index in the denominator counts as lower index, and vice versa.

Since the metric g is indefinite, the norm $\sqrt{\mathbf{a} \cdot \mathbf{a}}$ of a vector $\mathbf{a}$ can be

$$\mathbf{a} \cdot \mathbf{a} > 0 \quad \text{space-like}, \quad (20)$$

$$\mathbf{a} \cdot \mathbf{a} = 0 \quad \text{nullvector}, \quad (21)$$

$$\mathbf{a} \cdot \mathbf{a} < 0 \quad \text{time-like}. \quad (22)$$

The cone of all null-vectors starting from a point P is called *lightcone*. The time-like region inside the light-cone consists of two parts, past and future. Only events inside the past light-cone can influence the physics at point P , while P can influence only his future light-cone.

The line $x = x(\sigma)$ describing the position of an observer is called *worldline*. To measure the distance along the worldline it is convenient to introduce the *proper-time* τ ,

$$d\tau^2 = -ds^2/c^2. \quad (23)$$

Thus τ is real with units of time, measuring the time along the worldline. For instance,

$$\tau_{12} = \int_1^2 d\tau = \int_1^2 [dt^2 - (dx^2 + dy^2 + dz^2)/c^2]^{1/2} \quad (24)$$

$$= \int_1^2 dt [1 - (1/c^2)((dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2)]^{1/2} \quad (25)$$

$$= \int_1^2 dt [1 - v^2/c^2]^{1/2} < t_2 - t_1. \quad (26)$$

This relativistic effect of time dilation is often summarized by “A moving clocks runs slower.”

Lorentz transformations

We want to find now the coordinate transformations that keep the space-time interval

$$\Delta s^2 \equiv \vec{x}^2 - c^2 t^2 = \vec{x}'^2 - c^2 t'^2 \quad (27)$$

invariant. This transformations should be similar to the rotation matrices R and are called Lorentz transformations L . If we replace t by it in Δs^2 , the difference between two space-time events becomes the normal euclidean distance. Similarly, the identity $\cos^2 \alpha + \sin^2 \alpha = 1$ for an imaginary angle $\psi = i\alpha$ becomes $\cosh^2 \psi - \sinh^2 \psi = 1$. Thus a close correspondence exists between rotations R in euclidean space that leave $\Delta \vec{x}^2$ invariant and Lorentz transformations L that leave Δs^2 invariant. We try therefore as a guess

$$x = ct' \sinh \psi + x' \cosh \psi, \quad (28)$$

$$y = y', \quad z = z', \quad (29)$$

$$ct = ct' \cosh \psi + x' \sinh \psi. \quad (30)$$

Compared to normal coordinate rotation in euclidean space, the trigonometric functions are replaced by hyperbolic ones.

Consider now in system K the origin of the system K' . Then $x' = 0$ and

$$x = ct' \sinh \psi, \quad (31)$$

$$ct = ct' \cosh \psi. \quad (32)$$

Dividing the two equations gives

$$\frac{x}{ct} = \tanh \psi. \quad (33)$$

Since x/t is the relative velocity of the two systems, we have identified the physical meaning of the imaginary “rotation angle ψ ” as

$$\tanh \psi = \frac{v}{c} \equiv \beta. \quad (34)$$

Using the following identities

$$\cosh \psi = \frac{1}{\sqrt{1 - \tanh^2 \psi}} = \frac{1}{\sqrt{1 - \beta^2}} \equiv \gamma \quad (35)$$

$$\sinh \psi = \frac{\tanh \psi}{\sqrt{1 - \tanh^2 \psi}} = \frac{\beta}{\sqrt{1 - \beta^2}} = \gamma\beta \quad (36)$$

in (28) gives the standard form of the Lorentz transformations,

$$x = \frac{x' + vt'}{\sqrt{1 - (v/c)^2}} = \gamma(x' + \beta ct') \quad (37)$$

$$ct = \frac{ct' + v/cx'}{\sqrt{1 - (v/c)^2}} = \gamma(ct' + \beta x'). \quad (38)$$

The inverse transformation is obtained by inserting $-v$.

Remarks:

- Any other 4-tupel of quantities that transforms as x^μ is called (components of a) four-vector.
- We use from now on natural units $c = \hbar = k = 1$, but keep still the Newton’s constant G_N .

Relativity of simultaneity Consider in a space-time diagram two frames moving with relative speed v . Then the time axes are rotated by the angle $\psi = \operatorname{atanh}(v)$. The same holds for lines of constant time. Thus there are sequences of events $\{A, B, \}$ which one observer (call it 1) sees as $t_A < t_B$, another one (2) as $t_B < t_A$. However, this can happen only if A and B are space-like separated.

The four-velocity

What is the relativistic generalization of the velocity $\vec{v} = d\vec{x}/dt$? The nominator $d\vec{x}$ has already the right behavior to become part of a four-vector, if the denominator would be invariant. We use therefore instead dt the invariant proper time $d\tau = -ds$ and write

$$u^\alpha = \frac{dx^\alpha}{d\tau}. \quad (39)$$

The four-velocity is thus the tangent vector to the world-line of a particle. Written explicitly, we have

$$u^4 = \frac{dt}{d\tau} = \frac{1}{\sqrt{1 - v^2}} = \gamma. \quad (40)$$

and

$$u^i = \frac{dx^i}{d\tau} = \frac{dx^i}{dt} \frac{dt}{d\tau} = \frac{v^i}{\sqrt{1-v^2}} = \gamma v^i. \quad (41)$$

Hence the four-velocity is $u^\alpha = (\gamma\vec{v}, \gamma)$ and its norm is

$$\mathbf{u} \cdot \mathbf{u} = \gamma^2 v^2 - \gamma^2 = -\gamma^2(1 - v^2) = -1. \quad (42)$$

How to guess physical tensors: We may guess just by considering the different number of degrees, which physical 3-dim. quantities should be combined into 4-dim. tensors:

1. A vector $\mathbf{a}$ has $4 = 3 + 1$ components, i.e. combines a vector $\vec{a}$ and a scalar s , that are related by a physical law (e.g. current conservation).
2. An anti-symmetric tensor $\mathbf{F}$ has $6 = 3 + 2 + 1$ components, i.e. combines two vector $\vec{E}$ and $\vec{B}$.
3. A general tensor $\mathbf{F}$ has $16 = 9 + 2 \cdot 3 + 1$ components, i.e. combines from a 3-dim. point of view a 3-dim tensor, 2 vectors and a scalar.
4. A symmetric tensor $\mathbf{F}$ has $10 = 4 + 3 + 2 + 1$ components.

Energy and momentum

After having constructed the four-velocity, the simplest guess for the four-momentum is

$$p^\alpha = mu^\alpha = (\gamma m\vec{v}, \gamma m). \quad (43)$$

For small velocities, $v \ll 1$, we obtain

$$\vec{p} = \left(1 + \frac{v^2}{2} - \dots\right) m\vec{v} \quad (44)$$

$$p^4 = m + \frac{mv^2}{2} - \dots = m + E_{\text{kin,nr}} + \dots \quad (45)$$

Thus we can interpret the components as $p^\alpha = (\vec{p}, E)$. The norm follows with (42) immediately as

$$\mathbf{p} \cdot \mathbf{p} = -m^2. \quad (46)$$

Solving for the energy, we obtain

$$E = \sqrt{m^2 + p^2} \quad (47)$$

including the famous $E = mc^2$ as special case.

We postulate now that in relativistic mechanics Newton's law becomes

$$\mathbf{F} = \dot{\mathbf{p}}. \quad (48)$$

Doppler effect

We consider as last example how the wave-vector of a photon $(\vec{k}, \omega)$ transforms under a Lorentz transformation, Inserting into the transformation law for any four vector the time component $k^4 = \omega$ of the four-vector $k^\mu = (\vec{k}, \omega)$,

$$\omega' = \frac{\omega - V k_x}{[1 - V^2]^{1/2}}, \quad (49)$$

where we denote with ω the observed frequency. Setting $k_x = k \cos \alpha = \frac{\omega}{c} \cos \alpha$, where α the angle between the wave vector and the direction of the movement of the source, we obtain the relativistic Doppler formula

$$\frac{\omega}{\omega'} = \frac{[1 - V^2]^{1/2}}{1 - V \cos \alpha}. \quad (50)$$

For small relative velocities, $v \ll 1$, the quadratic term in the nominator can be neglected,

$$\omega \approx \omega' (1 + V \cos \alpha). \quad (51)$$

Thus one reproduces the non-relativistic Doppler formula where only the radial relative velocity enters. For $\alpha = 0$, photons are emitted in the same direction as the source moves, and the observed frequency ω is “blue-shifted”.