

**SPECIAL RELATIVITY.  
MATH2410**

**KOMISSAROV S.S**

2012



# Contents

<b>Contents</b>	<b>2</b>
<b>1 Space and Time in Newtonian Physics</b>	<b>9</b>
1.1 Space	9
1.1.1 Einstein summation rule	11
1.2 Time	12
1.3 Galilean relativity	12
1.4 Newtonian Mechanics	13
1.5 Galilean transformation	13
1.6 The lack of speed limit	15
1.7 Light	15
1.8 Advanced material: Maxwell equations, electromagnetic waves, and Galilean invariance	17
1.8.1 Maxwell equations	17
1.8.2 Some relevant results from vector calculus	17
1.8.3 Wave equation in electromagnetism	18
1.8.4 Plane waves	18
1.8.5 Wave equation is not Galilean invariant	19
<b>2 Basic Special Relativity</b>	<b>21</b>
2.1 Einstein's postulates	21
2.2 Einstein's thought experiments	21
2.2.1 Experiment 1. Relativity of simultaneity	21
2.2.2 Experiment 2. Time dilation	22
2.2.3 Experiment 3. Length contraction	24
2.2.4 Synchronization of clocks	26
2.3 Lorentz transformation	28
2.3.1 Derivation	28
2.3.2 Newtonian limit	29
2.4 Relativistic velocity "addition"	30
2.4.1 One-dimensional velocity "addition"	30
2.4.2 Three-dimensional velocity "addition"	31
2.5 Aberration of light	31
2.6 Doppler effect	33
2.6.1 Transverse Doppler effect	33
2.6.2 Radial Doppler effect	34
2.6.3 General case	35
<b>3 Space-time</b>	<b>37</b>
3.1 Minkowski diagrams	37
3.2 Space-time	38
3.3 Light cone	40
3.4 Causal structure of space-time	41
3.5 Types of space-time intervals	43

3.6	Vectors . . . . .	44
3.6.1	Definition . . . . .	44
3.6.2	Operations of addition and multiplication . . . . .	46
3.6.3	Coordinate transformation . . . . .	46
3.6.4	Infinitesimal displacement vectors . . . . .	47
3.7	Tensors . . . . .	48
3.7.1	Definition . . . . .	48
3.7.2	Components of tensors . . . . .	48
3.7.3	Coordinate transformation . . . . .	48
3.8	Metric tensor . . . . .	49
3.8.1	Definition . . . . .	49
3.8.2	Classification of space-time vectors and other results specific to space-time . .	50
<b>4</b>	<b>Relativistic particle mechanics</b>	<b>53</b>
4.1	Tensor equations and the Principle of Relativity . . . . .	53
4.2	4-velocity and 4-momentum . . . . .	55
4.3	Energy-momentum conservation . . . . .	57
4.4	Photons . . . . .	59
4.5	Particle collisions . . . . .	60
4.5.1	Nuclear recoil . . . . .	60
4.5.2	Absorption of neutrons . . . . .	62
4.6	4-acceleration and 4-force . . . . .	62

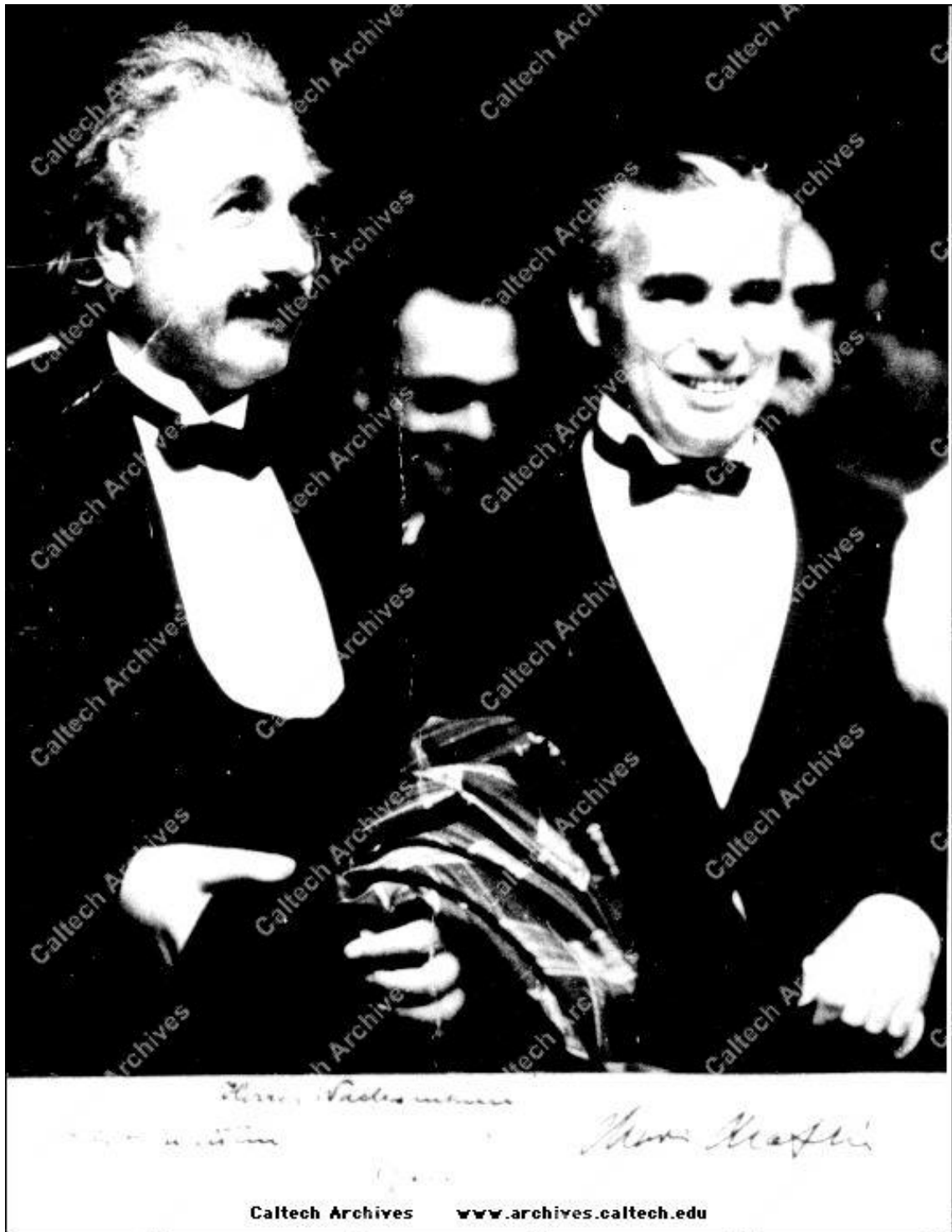


Figure 1: Charlie Chaplin: "They cheer me because they all understand me, and they cheer you because no one can understand you."



Figure 2: Arthur Stanley Eddington, the great English astrophysicist. From the conversation that took place in the lobby of The Royal Society: Silverstein - "... only three scientists in the world understand theory of relativity. I was told that you are one of them." Eddington - "Emm ... ."

Silverstein - "Don't be so modest, Eddington!" Eddington - "On the contrary. I am just wondering who this third person might be."

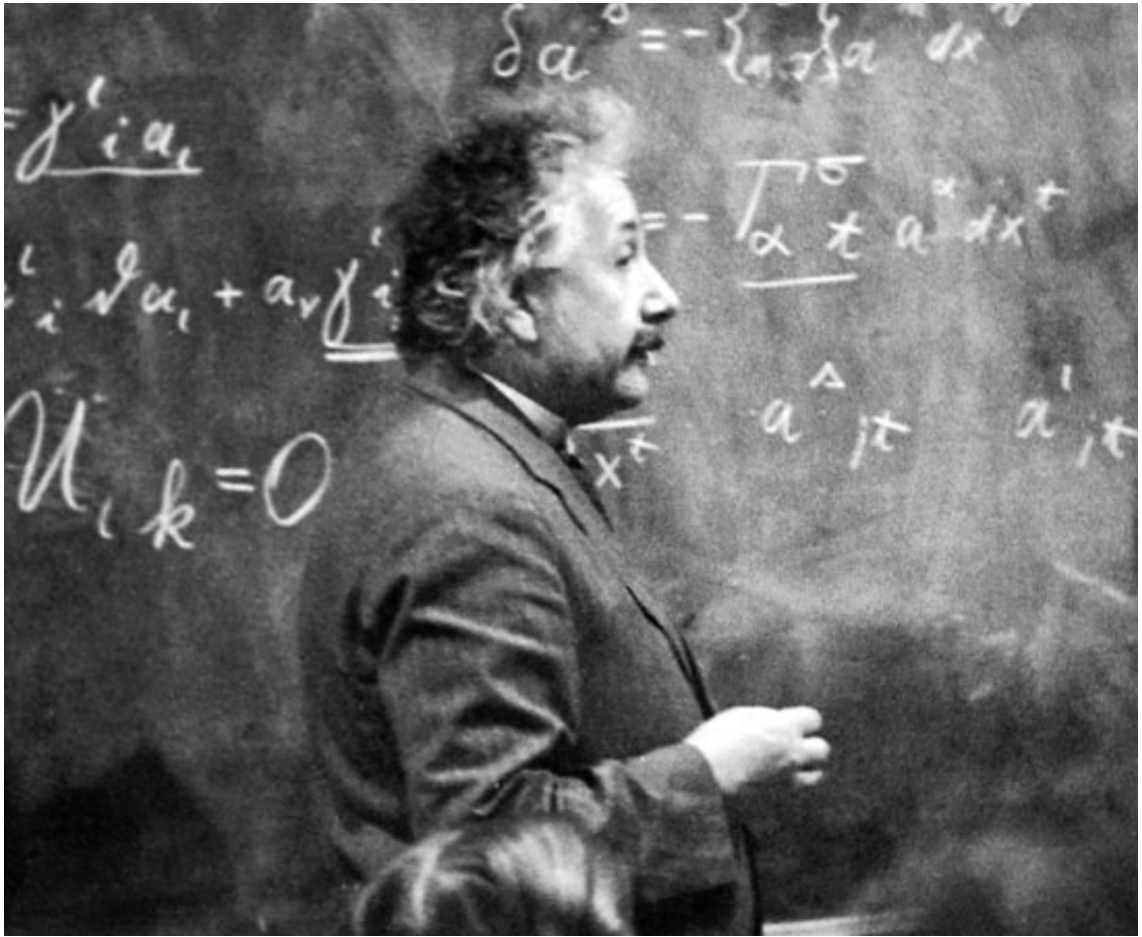


Figure 3: Einstein: “Since the mathematicians took over the theory of relativity I do no longer understand it.”





# Chapter 1

## Space and Time in Newtonian Physics

### 1.1 Space

The abstract notion of physical space reflects the properties of physical objects to have sizes and physical events to be located at different places relative to each other. In Newtonian physics, the physical space was considered as a fundamental component of the world around us, which exists by itself independently of other physical bodies and normal matter of any kind. It was assumed that 1) one could “interact” with this space and unambiguously determine the motion of objects in this space, in addition to the easily observed motion of physical bodies relative to each other, 2) that one may introduce points of this space, and determine at which point any particular event took place. The actual ways of doing this remained mysterious though. It was often thought that the space is filled with a primordial substance, called “ether” or “plenum”, which can be detected one way or another, and that “atoms” of ether correspond to points of physical space and that motion relative to these atoms is the motion in physical space. This idea of physical space was often called “the absolute space” and the motion in this space “the absolute motion”.

There also was a consensus that the best mathematical model for the absolute space was the 3-dimensional Euclidean space. By definition, in such space one can construct cuboids, rectangular parallelepipeds, such that the lengths of their edges,  $a, b$ , and  $c$ , and the diagonal  $l$  satisfy the following equation

$$l^2 = a^2 + b^2 + c^2, \tag{1.1}$$

no matter how big the cuboid is. This was strongly supported by the results of practical geometry.

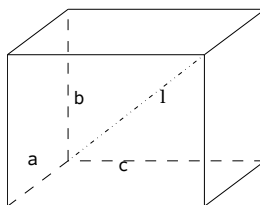


Figure 1.1: A cuboid

Given this property one can construct a set of Cartesian coordinates,  $\{x^1, x^2, x^3\}$  (the same meaning as  $\{x, y, z\}$ ). These coordinates are distances between the origin and the point along the

coordinate axes.

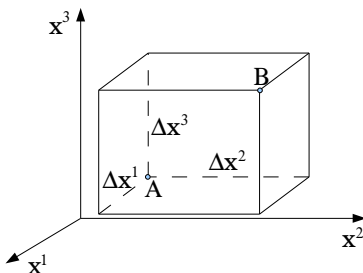


Figure 1.2: Cartesian coordinates

In Cartesian coordinates, the distance between point A and point B with coordinates  $\{x_a^1, x_a^2, x_a^3\}$  and  $\{x_b^1, x_b^2, x_b^3\}$  respectively is

$$\Delta l_{ab}^2 = (\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2, \quad (1.2)$$

where  $\Delta x^i = x_a^i - x_b^i$ .

For infinitesimally close points this becomes

$$dl^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2, \quad (1.3)$$

where  $dx^i$  are infinitesimally small differences between Cartesian coordinates of these points. This equation allows us to find distances along curved lines by means of integration.

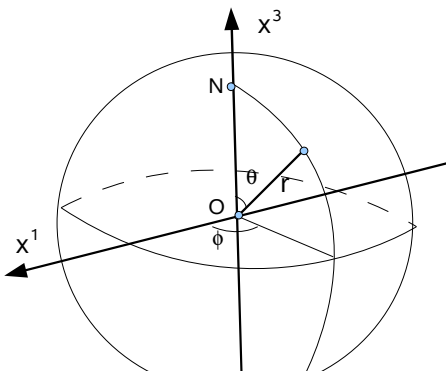


Figure 1.3: Spherical coordinates

Other types of coordinates can also be used in Euclidean space. One example is the spherical coordinates,  $\{r, \theta, \phi\}$ , defined via

$$x^1 = r \sin \theta \cos \phi; \quad (1.4)$$

$$x^2 = r \sin \theta \sin \phi; \quad (1.5)$$

$$x^3 = r \cos \theta. \quad (1.6)$$

Here  $r$  is the distance from the origin,  $\theta \in [0, \pi]$  is the polar angle, and  $\phi \in [0, 2\pi)$  is the azimuthal angle. The coordinate lines of these coordinates are not straight lines but curves. Such coordinate systems are called curvilinear. The coordinate lines of spherical coordinates are perpendicular to

each other at every point. Such coordinate systems are called orthogonal (There are non-orthogonal curvilinear coordinates). In spherical coordinates, the distance between infinitesimally close points is

$$dl^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2. \quad (1.7)$$

Since, the coefficients of  $d\theta^2$  and  $d\phi^2$  vary in space

$$\Delta l_{ab}^2 \neq (\Delta r)^2 + r^2 \Delta\theta^2 + r^2 \sin^2 \theta \Delta\phi^2 \quad (1.8)$$

for the distance between points with finite separations  $\Delta r, \Delta\theta, \Delta\phi$ . In order to find this distance one has to integrate

$$l_{ab} = \int_a^b dl \quad (1.9)$$

along the line connecting these points. For example the circumference of a circle of radius  $r_0$  is

$$\Delta l = \int dl = 2 \int_0^\pi r_0 d\theta = 2\pi r_0. \quad (1.10)$$

(Notice that we selected such coordinates that the circle is centered on the origin,  $r = r_0$ , and it is in a meridional plane,  $\phi = \text{const}$ . As the result, along the circle  $dl = r_0 d\theta$ .)

In the generic case of curvilinear coordinates, the distance between infinitesimally close points is given by the positive-definite quadratic form

$$dl^2 = \sum_{i=1}^3 \sum_{j=1}^3 g_{ij} dx^i dx^j, \quad (1.11)$$

where  $g_{ij} = g_{ji}$  are coefficients that are functions of coordinates. Such quadratic forms are called *metric forms*.  $g_{ij}$  are in fact the components of so-called *metric tensor* in the coordinate basis of utilised coordinates. It is easy to see that in a Cartesian basis

$$g_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (1.12)$$

Not all positive definite metric forms correspond to Euclidean space. If there does not exist a coordinate transformation which reduces a given metric form to that of Eq.1.3 then the space with such metric form is not Euclidean. This mathematical result is utilized in General Relativity.

### 1.1.1 Einstein summation rule

In the modern mathematical formulation of the Theory of Relativity it is important to distinguish between upper and lower indexes as the index position determines the mathematical nature of the indexed quantity. E.g. a single upper index indicates a vector (or a contravariant vector), as in  $b^i$ , whereas  $b_i$  stands for a mathematical object of a different type, though uniquely related to the vector  $b$ , the so-called one-form (or covariant vector). This applies to coordinates, as the Cartesian coordinates can be interpreted as components of the radius-vector. Our syllabus is rather limited and we will not be able to explore the difference between covariant and contravariant vectors in full, as well as the difference between various types of tensors. However, we will keep using this modern notation.

The Einstein summation rule is a convention on the notation for summation over indexes. Namely, any index appearing once as a lower index and once as an upper index of the same indexed object or in the product of a number of indexed objects stands for summation over all allowed values of this index. Such index is called a dummy, or summation index. Indexes which are not dummy are called free indexes. Their role is to give a correct equation for any allowed index value, e.g. 1,2, and 3 in three-dimensional space.

According to this rule we can rewrite Eq.1.11 in a more concise form:

$$dl^2 = g_{ij} dx^i dx^j. \quad (1.13)$$

This rule allows to simplify expressions involving multiple summations. Here are some more examples:

1.  $a_i b^i$  stands for  $\sum_{i=1}^n a_i b^i$ ; here  $i$  is a dummy index;
2. In  $a_i b_i$   $i$  is a free index.
3.  $a_i b^{kij}$  stands for  $\sum_{i=1}^n a_i b^{kij}$ ; here  $k$  and  $j$  are free indexes and  $i$  is a dummy index;
4.  $a^i \frac{\partial f}{\partial x^i}$  stands for  $\sum_{i=1}^n a^i \frac{\partial f}{\partial x^i}$ . Notice that index  $i$  in the partial derivative  $\frac{\partial}{\partial x^i}$  is treated as a lower index.

## 1.2 Time

The notion of time reflects our everyday-life observation that all events can be placed in a particular order reflecting their causal connection. In this order, event A appears before event B if A caused B or could cause B. This causal order seems to be completely independent on the individual analysing these events. This notion also reflects the obvious fact that one event can last longer than another one. In Newtonian physics, time was considered as a kind of fundamental ever going process, presumably periodic, so that one can compare the rate of this process to rates of all other processes. Although the nature of this process remained mysterious it was assumed that all other periodic processes, like the Earth rotation, reflected it. Given the fundamental nature of time it would be natural to assume that this process occurs in ether.

This understanding of time has lead naturally to the absolute meaning of simultaneity. That is one could unambiguously decide whether two events were simultaneous or not. Similarly, physical events could be placed in only one particular order, so that if event A precedes event B according to the observation of some observer, this has to be the same for all other observers, unless a mistake is made. Similarly, any event could be described by only one duration, when the same unit of time is used to quantify it. These are the reasons for the time of Newtonian physics to be called the “absolute time”. There is only one time for everyone.

Both in theoretical and practical terms, a unique temporal order of events could only be established if there existed signals propagating with infinite speed. In this case, when an event occurs in a remote place everyone can become aware of it instantaneously by means of such “super-signals”. Then all events immediately divide into three groups with respect to this event: (i) The events simultaneous with it – they occur at the same instant as the arrival of the super-signal generated by the event; (ii) The events preceding it – they occur before the arrival of this super-signal and could not be caused by it. But they could have caused the original event; (iii) The events following it – they occur after the arrival of this super-signal and can be caused by it. But they cannot cause the original event.

If, however, there are no such super-signals, things become highly complicated as one needs to know not only the distances to the events but also the motion of the observer and how exactly the signals propagated through the space separating the observer and these events. Newtonian physics assumes that such infinite speed signals do exist and they play fundamental role in interaction between physical bodies. This is how in the Newtonian theory of gravity, the gravity force depends only on the current location of the interacting masses.

## 1.3 Galilean relativity

Galileo, who is regarded to be the first true natural scientist, made a simple observation which turned out to have far reaching consequences for modern physics. He noticed that it was difficult to

tell whether a ship was anchored or coasting at sea by means of mechanical experiments carried out on board of this ship.

It is easy to determine where a body is moving through air – in the case of motion, it experiences the air resistance, the drag and lift forces. But here we are dealing only with a motion relative to air. What about the motion relative to the absolute space and the interaction with ether? If such an interaction occurred then one could measure the “absolute motion”. Galileo’s observation tells us that this must be at least a rather weak interaction. No other mechanical experiment, made after Galileo, has been able to detect such an interaction. Newtonian mechanics states this fact in its First Law as: “The motion of a physical body which does not interact with other bodies remains unchanged. It moves with constant speed along straight line.” This means that one cannot determine the motion through absolute space by mechanical means. Only the relative motion between physical bodies can be determined this way.

## 1.4 Newtonian Mechanics

Newton (1643-1727) founded the classical mechanics - a basic set of mathematical laws of motion based on the ideas of space and time described above. One of the key notions he introduced is the notion of an *inertial reference frame*.

A reference frame is a solid coordinate grid (usually Cartesian but not always), used to quantitatively describe the physical motion. In general, such a frame can be in arbitrary motion in space and one can introduce many different reference frames. However, not all reference frames are equally convenient. Some of them are much more convenient than others as the motion of bodies that are not interacting with other bodies is particularly simple in such frames, namely they move with zero acceleration,

$$\underline{a} = \frac{d\underline{w}}{dt} = 0, \quad (1.14)$$

where  $\underline{w} = d\underline{r}/dt$  is the body velocity. Such frames are called *inertial frames*. We stress that existence of such frames is a basic assumption (postulate) of the theory. It is called *the first law of Newtonian mechanics*.

If the motion of a body as measured in inertial frame is in fact accelerated then it is subject to interaction with other bodies. Mathematically such interaction is described by a force vector  $\underline{f}$ . The acceleration is then given by *the second law of Newtonian mechanics*

$$m\underline{a} = \underline{f} \quad (1.15)$$

where  $m$  is a scalar quantity called the inertial mass of the body (this is the property describing body’s ability to resist action of external forces). Each kind of interaction should be described by additional laws determining the force vector as a function of other parameters (e.g the law of gravity).

*The third law of Newtonian mechanics* deals with binary interactions, or interactions involving only two bodies, say A and B. It states that

$$\underline{f}_a = -\underline{f}_b, \quad (1.16)$$

where  $\underline{f}_a$  is the force acting on body A and  $\underline{f}_b$  is the force acting on body B.

## 1.5 Galilean transformation

Consider two Cartesian frames,  $S$  and  $\tilde{S}$ , with coordinates  $\{x^i\}$  and  $\{\tilde{x}^i\}$  respectively. Assume that (i) their corresponding axes are parallel, (ii) their origins coincide at time  $t = 0$ , (iii) frame  $\tilde{S}$  is moving relative to  $S$  along the  $x^1$  axis with constant speed  $v$ , as shown in Figure 1.4. This will be called the *standard configuration*.

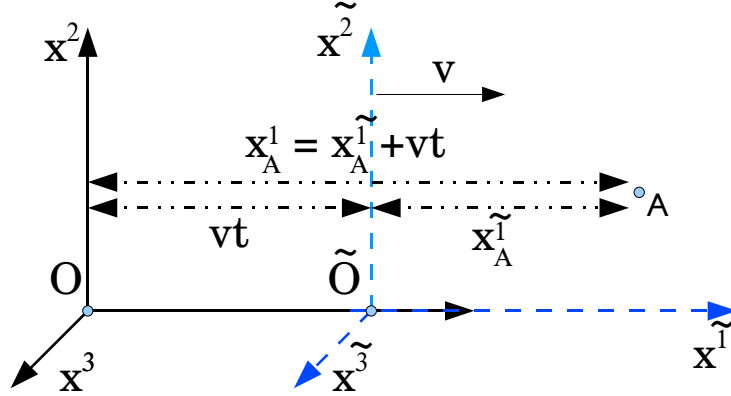


Figure 1.4: Measuring the  $x^1$  coordinate of event A in two reference frames in the standard configuration. For simplicity, we show the case where this event occurs in the plane  $x^3 = 0$ .

The *Galilean transformation* relates the coordinates of events as measured in both frames. Given the absolute nature of time in Newtonian physics, it is the same for both frames. So this may look over-elaborate if we write

$$t = \tilde{t}. \quad (1.17)$$

However, this makes sense if we wish to stress that both fiducial observers, who ride these frames and make the measurements, use clocks made to the same standard and synchronised with each other.

Next let us consider the spatial coordinates of some event A. The  $x^2$  coordinate of this event is its distance in the absolute space from the plane given by the equation  $x^2 = 0$ . Similarly, the  $\tilde{x}^2$  coordinate of this event is its distance in the absolute space from the plane given by the equation  $\tilde{x}^2 = 0$ . Since these two planes coincide both these distances are the same and hence  $x^2 = \tilde{x}^2$ . Similarly, we conclude that  $x^3 = \tilde{x}^3$ . As to the remaining coordinate, the distance between planes  $x^1 = 0$  and  $\tilde{x}^1 = 0$  at the time of the event is  $vt$ , and hence  $x^1 = \tilde{x}^1 + vt$  (see Figure 1.4). Summarising,

$$\begin{aligned} x^1 &= \tilde{x}^1 + vt, \\ x^2 &= \tilde{x}^2, \\ x^3 &= \tilde{x}^3. \end{aligned} \quad (1.18)$$

This is the Galilean transformation for the standard configuration. If we allow the frame  $\tilde{S}$  to move in arbitrary direction with velocity  $v^i$  then a more general result follows,

$$x^i = \tilde{x}^i + v^i t. \quad (1.19)$$

From this we derive the following two important conclusions:

- The velocity transformation law:

$$w^i = \tilde{w}^i + v^i, \quad (1.20)$$

where  $w^i = dx^i/dt$  and  $\tilde{w}^i = d\tilde{x}^i/d\tilde{t}$  are the velocities of a body as measured in frames  $S$  and  $\tilde{S}$  respectively.

- The acceleration transformation law:

$$a^i = \tilde{a}^{\tilde{i}}, \quad (1.21)$$

where  $a^i = d^2x^i/dt^2$  and  $\tilde{a}^{\tilde{i}} = d^2x^{\tilde{i}}/dt^2$  are the accelerations of a body as measured in frames  $S$  and  $\tilde{S}$  respectively. Thus, in both frames the acceleration is exactly the same.

These results are as fundamental as it gets, because they follow directly from the notions of absolute space and absolute time.

Equation 1.21 tells us that the body acceleration is the same in all inertial frames and so must be the force producing this acceleration. Thus, the laws of Newtonian mechanics are the same (or *invariant*) in all inertial frames. This is exactly what was discovered by Galileo and is now known as *the Galilean principle of relativity*. In particular, if the acceleration measured in the frame  $S$  vanishes then it also vanishes in the frame  $\tilde{S}$ . This tells us that there exist infinitely many inertial frames and they all move relative to each other with constant speed. Only one of these frames is at rest in the absolute space (here we do not differentiate between frames with different orientation of their axes or/and locations of their origins) but we cannot tell which one. This makes the absolute space a very elusive if not ghostly “object”.

## 1.6 The lack of speed limit

Is there any speed limit a physical body can have in Newtonian mechanics? The answer to this question is No. To see this consider a particle of mass  $m$  under the action of constant force  $f$ . According to the second law of Newton its speed then grows linearly,

$$w = w_0 + \frac{f}{m}t,$$

without a limit.

This conclusion also agrees with the Galilean principle of relativity. Indeed, suppose there is a maximum allowed speed, say  $w_{max}$ . According to this principle it must be the same for all inertial frames. Now consider a body moving with such a speed to the right of the frame  $\tilde{S}$ . This frame can also move with speed  $w_{max}$  relative to the frame  $S$ . Then according to the Galilean velocity addition this body moves relative to frame  $S$  with speed  $2w_{max}$ . This contradicts to our assumption that there exist a speed limit, and hence this assumption has to be discarded.

## 1.7 Light

The nature of light was a big mystery in Newtonian physics and a subject of heated debates between scientists. One point of view was that light is made by waves propagating in ether, by analogy with sound which is made by waves in air. The speed of light waves was a subject of great interest to scientists. The most natural expectation for waves in ether is to have infinite speed. Indeed, waves with infinite speed fit nicely the concept of absolute time, and if such waves exist then there is no more natural medium for such wave as the ether of absolute space. However, the light turned out to have finite speed. Dutch astronomer Roemer noticed that the motion of Jupiter’s moons had systematic variation, which could be easily explained only if one assumed that light had finite, though very large, speed. Since then, many other measurements have been made which all agree on the value for the speed of light

$$c \simeq 3 \times 10^{10} \text{ cm/s.}$$

The development of mathematical theory of electromagnetism resulted in the notions of electric and magnetic fields, which exist around electrically charged bodies. These fields do not manifest themselves in any other ways but via forces acting on other electrically charged bodies. Attempts to describe the properties of these fields mathematically resulted in Maxwell’s equations, which agreed with experiments most perfectly.

What is the nature of electric and magnetic field? They could just reflect some internal properties of matter, like air, surrounding the electrically charged bodies. Indeed, it was found that the electric and magnetic fields depended on the chemical and physical state of surrounding matter. However, the experiments clearly indicated that the electromagnetic fields could also happily exist in vacuum (empty space). This fact prompted suggestions that in electromagnetism we are dealing with ether.

Analysis of Maxwell equations shows that electric and magnetic fields change via waves propagating with finite speed. In vacuum the speed of these waves is the same in all direction and equal to the known speed of light! When this had been discovered, Maxwell immediately interpreted light as electromagnetic waves or ether waves. Since according to the Galilean transformation the result of any speed measurement depends on the selection of inertial frame, the fact that Maxwell equations yielded a single speed could only mean that they are valid only in one particular frame, the rest frame of ether and absolute space. On the other hand, the fact that the astronomical observations and laboratory experiments did not find any variation of the speed of light as well seemed to indicate that Earth was almost at rest in the absolute space.

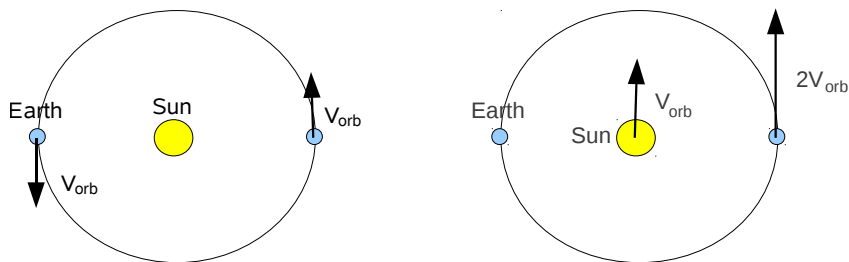


Figure 1.5: Left panel: Earth's velocity relative to the Sun at two opposite points of its orbit. Right panel: Earth's absolute velocity at two opposite points of its orbit and the Sun absolute velocity, assuming that at the left point the Earth velocity vanishes.

However, Newtonian mechanics clearly shows that Earth cannot be exactly at rest in absolute space all the time. Indeed, it orbits the Sun and even if at one point of this orbit the speed of Earth's absolute motion is exactly zero it must be nonzero at all other points, reaching the maximum value equal to twice the orbital speed at the opposite point of the orbit. This simple argument shows that during one calendar year the speed of light should show variation of the order of the Earth orbital speed and that the speed of light should be different in different directions by at least the orbital speed. Provided the speed measurements are sufficiently accurate we must be able to see these effects. American physicists Mickelson and Morley were first to design experiments of such accuracy (by the year 1887) and to everyone's amazement and disbelief their results were negative. Within their experimental errors, the speed of light was the same in all directions all the time! Since then, the accuracy of experiments has improved dramatically but the result is still the same, clearly indicating shortcomings of Newtonian physics with its absolute space and time. Moreover, no object has shown speed exceeding the speed of light. In his ground-breaking work "On the electrodynamics of moving bodies", published in 1905, Albert Einstein paved way to new physics with completely new ideas on the nature of physical space and time, the Theory of Relativity, which accommodates these remarkable experimental findings.



## 1.8 Advanced material: Maxwell equations, electromagnetic waves, and Galilean invariance

### 1.8.1 Maxwell equations

Maxwell (1831-1879) completed the mathematical theory of electrodynamics. After his work, the evolution of electromagnetic field in vacuum is described by

$$\underline{\nabla} \cdot \underline{B} = 0, \quad (1.22)$$

$$\frac{1}{c} \frac{\partial \underline{B}}{\partial t} + \underline{\nabla} \times \underline{E} = 0, \quad (1.23)$$

$$\underline{\nabla} \cdot \underline{E} = 0, \quad (1.24)$$

$$-\frac{1}{c} \frac{\partial \underline{E}}{\partial t} + \underline{\nabla} \times \underline{B} = 0, \quad (1.25)$$

where  $c$  is a constant with dimension of speed. Laboratory experiments with electromagnetic materials allowed to measure this constant – it turned out to be equal to the speed of light!

Comment: Later, the works by Planck(1858-1947) and Einstein(1879-1955) lead to the conclusion that electromagnetic energy is emitted, absorbed, and propagate in discrete quantities, or photons. Thus, Newton's ideas have been partially confirmed as well. Such particle-wave duality is a common property of micro-particles that is accounted for in quantum theory.

### 1.8.2 Some relevant results from vector calculus

Notation:  $\{x^k\}$  - Cartesian coordinates ( $k = 1, 2, 3$ );  $\{\hat{e}_k\}$  are the unit vectors along the  $x^k$  axes;  $\underline{r} = x^k \hat{e}_k$  is the position vector (radius vector) of the point with coordinates  $\{x^k\}$ ;  $\underline{A}(\underline{r}) = A^k(\underline{r}) \hat{e}_k$  is a vector field in Euclidean space (vector function);  $A^k$  are the components of vector  $\underline{A}$  in the basis  $\{\hat{e}_k\}$ ;  $f(\underline{r})$  is a scalar field in Euclidean space (scalar function).

The divergence of vector field  $\underline{A}$  is defined as

$$\underline{\nabla} \cdot \underline{A} = \frac{\partial A^k}{\partial x^k}. \quad (1.26)$$

(Notice use of Einstein summation convention in this equation!). This is a scalar field. One can think of  $\underline{\nabla}$  as a vector with components  $\partial/\partial x^k$  and consider  $\underline{\nabla} \cdot \underline{A}$  as a scalar product of  $\underline{\nabla}$  and  $\underline{A}$ . The curl of vector field  $\underline{A}$  is defined via the determinant rule for vector product

$$\underline{\nabla} \times \underline{A} = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ \frac{\partial}{\partial x^1} & \frac{\partial}{\partial x^2} & \frac{\partial}{\partial x^3} \\ A^1 & A^2 & A^3 \end{vmatrix} = \quad (1.27)$$

$$= \hat{e}_1 \left( \frac{\partial A^3}{\partial x^2} - \frac{\partial A^2}{\partial x^3} \right) + \hat{e}_2 \left( \frac{\partial A^1}{\partial x^3} - \frac{\partial A^3}{\partial x^1} \right) + \hat{e}_3 \left( \frac{\partial A^2}{\partial x^1} - \frac{\partial A^1}{\partial x^2} \right). \quad (1.28)$$

This is a vector field. The gradient of a scalar field is defined as

$$\underline{\nabla} f = \frac{\partial f}{\partial x^1} \hat{e}_1 + \frac{\partial f}{\partial x^2} \hat{e}_2 + \frac{\partial f}{\partial x^3} \hat{e}_3. \quad (1.29)$$

This is a vector field. The Laplacian of scalar field  $f$  is defined as

$$\nabla^2 f = \underline{\nabla} \cdot \underline{\nabla} f. \quad (1.30)$$

From this definition and eqs.(1.26,1.29) one finds that

$$\nabla^2 f = \frac{\partial^2 f}{\partial (x^1)^2} + \frac{\partial^2 f}{\partial (x^2)^2} + \frac{\partial^2 f}{\partial (x^3)^2}. \quad (1.31)$$

This is a scalar field. The Laplacian of vector field  $\underline{A}$  is defined as

$$\nabla^2 \underline{A} = \hat{e}_k \nabla^2 A^k. \quad (1.32)$$

(Notice use of Einstein summation convention in this equation!). This is a vector field. The following vector identity is very handy

$$\underline{\nabla} \times (\underline{\nabla} \times \underline{A}) = \underline{\nabla}(\underline{\nabla} \cdot \underline{A}) - \nabla^2 \underline{A}. \quad (1.33)$$

### 1.8.3 Wave equation in electromagnetism

Apply  $\underline{\nabla} \times$  to eq.1.23 to obtain

$$\frac{1}{c} \underline{\nabla} \times \frac{\partial \underline{B}}{\partial t} + \underline{\nabla} \times (\underline{\nabla} \times \underline{E}) = 0.$$

Using eq.1.33 and commutation of partial derivatives this reduces to

$$\frac{1}{c} \frac{\partial}{\partial t} \underline{\nabla} \times \underline{B} + \underline{\nabla}(\underline{\nabla} \cdot \underline{E}) - \nabla^2 \underline{E} = 0.$$

The second term vanishes due to eq.1.24 and eq.1.25 allows us to replace  $\underline{\nabla} \times \underline{B}$  with  $\frac{\partial \underline{E}}{\partial t}$  in the first term. This gives us the final result

$$\frac{1}{c^2} \frac{\partial^2 \underline{E}}{\partial t^2} - \nabla^2 \underline{E} = 0. \quad (1.34)$$

In a similar fashion one can show that

$$\frac{1}{c^2} \frac{\partial^2 \underline{B}}{\partial t^2} - \nabla^2 \underline{B} = 0. \quad (1.35)$$

These are examples of the canonical *wave equation*

$$\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi = 0, \quad (1.36)$$

where  $\psi(\underline{r}, t)$  is some function of space and time.

### 1.8.4 Plane waves

Look for solutions of eq.1.36 that depend only on  $t$  and  $x^1$ . Then

$$\frac{\partial \psi}{\partial x^2} = \frac{\partial \psi}{\partial x^3} = 0$$

and eq.1.36 reduces to

$$\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \frac{\partial^2 \psi}{\partial (x^1)^2} = 0. \quad (1.37)$$

This is a one-dimensional wave equation. It is easy to verify by substitution that it has solutions of the form

$$\psi_{\pm}(t, x^1) = f(x^1 \pm ct), \quad (1.38)$$

where  $f(x)$  is an arbitrary twice differentiable function.  $\psi_+(t, x^1) = f(x^1 + ct)$  describes waves propagating with speed  $c$  in the negative direction of the  $x^1$  axis and  $\psi_-(t, x^1) = f(x^1 - ct)$  describes waves propagating with speed  $c$  in the positive direction of the  $x^1$  axis. Thus, equations (1.34,1.35), tell us straight away that Maxwell equations imply electromagnetic waves propagating with speed  $c$ , the speed of light.

### 1.8.5 Wave equation is not Galilean invariant

Can the electromagnetic phenomena be used to determine the absolute motion, that is motion relative to the absolute space. If like the equations of Newtonian mechanics the equations of electrodynamics are the same in all inertial frames then they cannot. Thus, it is important to see how the Maxwell equations transform under the Galilean transformation. However, it is sufficient to consider only the wave equation, eq.1.36, which is a derivative of Maxwell's equations. For simplicity sake, one can deal only with its one-dimensional version, eq.1.37. Denoting  $x^1$  as simply  $x$  we have

$$-\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} + \frac{\partial^2 \psi}{\partial x^2} = 0. \quad (1.39)$$

The Galilean transformation reads

$$\tilde{x} = x - vt.$$

It is easy to see that in new variables,  $\{t, \tilde{x}\}$ , equation 1.39 becomes

$$-\frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} + \frac{2v}{c^2} \frac{\partial^2 \psi}{\partial t \partial \tilde{x}} + \left(1 - \frac{v^2}{c^2}\right) \frac{\partial^2 \psi}{\partial \tilde{x}^2} = 0. \quad (1.40)$$

Since eq.1.40 has a different form compared to eq.1.39 we conclude that *the wave equation, and hence the Maxwell equations, are not invariant under Galilean transformation! One should be able to detect motion relative to the absolute space!*

In order to elucidate this result consider the wave solutions of eq.1.40. Direct substitution shows that this equation is satisfied by

$$\Psi_{\pm}(t, \tilde{x}) = f(\tilde{x} + a_{\pm}t), \quad (1.41)$$

where  $f(x)$  is again an arbitrary twice differentiable function and

$$a_{\pm} = v \pm c. \quad (1.42)$$

These are waves propagating with speeds  $a_{\pm}$ . In fact,  $\Psi_{+}(t, \tilde{x}) = f(\tilde{x} + (v + c)t)$  describes wave propagating with speed

$$\tilde{c} = -v - c$$

and  $\Psi_{-}(t, \tilde{x}) = f(\tilde{x} + (v - c)t)$  describes wave propagating with speed

$$\tilde{c} = -v + c$$

Comparing these results with eq.1.20 shows us that what we have got here is simply the Galilean velocity transformation for electromagnetic waves. Thus, we arrive to the following conclusions

1. Electromagnetic waves can propagate with speed  $c$  in all directions only in one very special inertial frame, namely the frame that is in rest in absolute space. In any other frame it will be different in different directions, as dictated by the Galilean transformation, and equal to  $c$  only in the directions normal to the frame velocity relative to the absolute space.
2. The Maxwell equations are not general. They hold only in the frame at rest in the absolute space.

*In spite of looking very convincing these conclusions however do not comply with physical experiments which show beyond any doubt that in all frames the electromagnetic waves propagate with the same speed  $c$  in all directions! These experimental results show that the Galilean transformation is not that general as thought before Einstein, and hence the notions of space and time as described of Newtonian physics are not correct.*



## Chapter 2

# Basic Special Relativity

### 2.1 Einstein's postulates

Paper “On the electrodynamics of moving bodies” by Einstein (1905).

- *Postulate 1 (Principle of Relativity):*

“All physical laws are the same (invariant) in all inertial frames”.

- *Postulate 2:*

“The speed of light (in vacuum) is the same in all inertial frames”. It is the same in magnitude and does not depend on direction of propagation.

Postulate 1 implies that no physical experiment can be used to measure the absolute motion. In other words the notions of absolute motion and absolute space become redundant. As far as physics is concerned the absolute space does not exist!

Postulate 2 is fully consistent with Postulate 1 (and hence may be considered as a derivative of Postulate 1). Indeed, if Maxwell's equations (1.22-1.25) are the same in all inertial frames then the electromagnetic waves propagate with the same speed which is given by the constant  $c$  in these equations. Since Postulate 2 is in conflict with the Galilean velocity addition it shows that the very basic properties of physical space and time have to be reconsidered. In the next section we carry out a number of very simple “thought experiments” which show how dramatic the required modifications are.

### 2.2 Einstein's thought experiments

In these experiments we assume that when the speed of the same light signal is measured in any inertial frame the result is always the same, namely  $c$ .

#### 2.2.1 Experiment 1. Relativity of simultaneity

A carriage is moving with speed  $v$  past a platform. A conductor, who stands in the middle of the carriage, sends simultaneously two light pulses in the opposite directions along the track. Both passengers and the crowd waiting on the platform observe how the pulses hit the ends of the carriage.

The carriage passengers agree that both pulses reach the ends simultaneously as they propagate with the same speed and have to cover the same distance (see left panel of figure 2.1). If  $\tilde{L}$  is the carriage length as measured by its passengers then the required time is

$$\tilde{t}_{left} = \tilde{t}_{right} = \tilde{L}/2c.$$

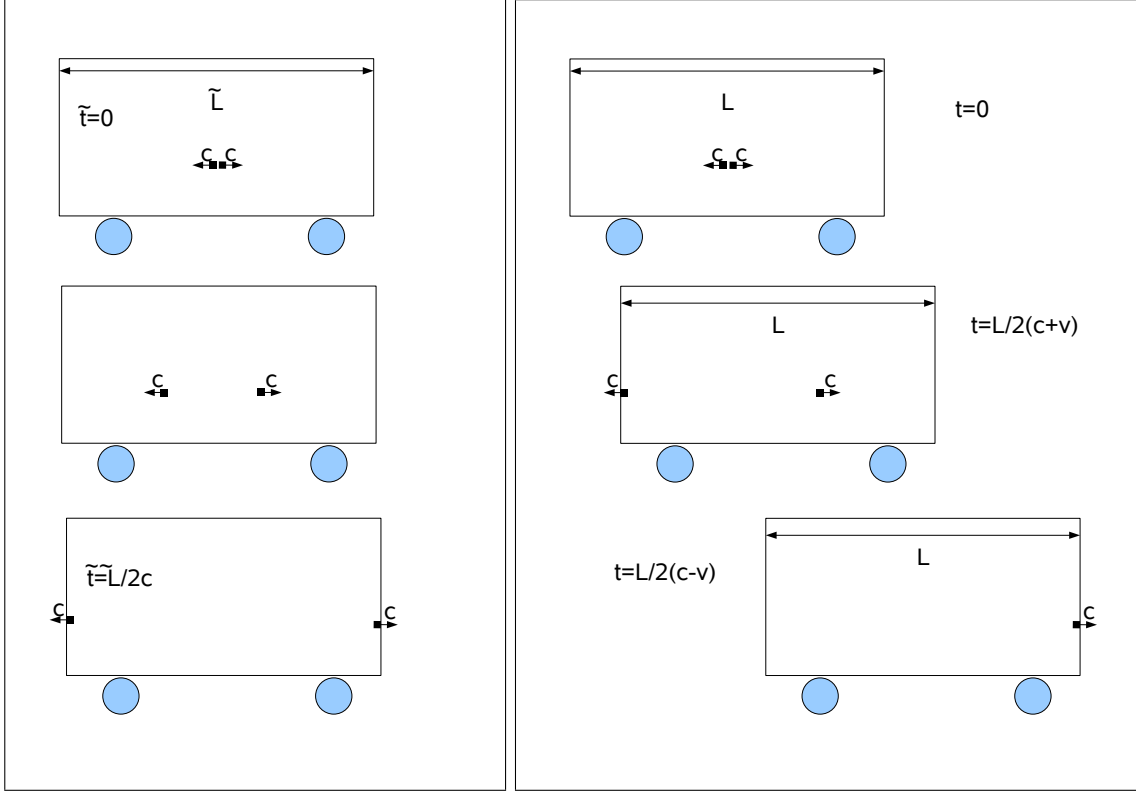


Figure 2.1: Thought experiment number 1. *Left panel:* Events as seen in the carriage frame. *Right panel:* Events as seen in the platform frame.

The crowd on the platform, however, see things differently. As both ends of the carriage move to the right the pulse sent to the left has to cover shorter distance and reaches its end earlier than the pulse sent to the right (right panel of figure 2.1). In fact, the required times are

$$t_{left} = \frac{L}{2} \left( \frac{1}{c+v} \right) \quad \text{and} \quad t_{right} = \frac{L}{2} \left( \frac{1}{c-v} \right),$$

where  $L$  is the carriage length as measured by the crowd (obtain these results.) As we shall see later  $L \neq \tilde{L}$ ). Thus, two events, which are simultaneous in the carriage frame are not simultaneous in the platform frame. *This implies that temporal order of events depends of the frame of reference and hence that the absolute time does no longer exists. Instead, each inertial frame must have its own time.* The next experiment supports this conclusion, showing that the same events may have different durations in different frames.

### 2.2.2 Experiment 2. Time dilation

This time a single pulse is fired from one side of the carriage perpendicular to the track, reflects of the other side, and returns back. In the carriage frame (left panel of figure 2.2) the pulse covers the distance  $2W$ , where  $W$  is the carriage width, and this takes time

$$\Delta\tilde{t} = \frac{2W}{c}.$$

In order to find  $\Delta t$ , the elapsed time as measured in the platform frame, we notice that we can write the distance covered by the pulse as  $c\Delta t$  and as  $2\sqrt{W^2 + (v\Delta t/2)^2}$  (right panel of figure 2.2). Thus,

$$(c\Delta t/2)^2 = W^2 + (v\Delta t/2)^2.$$

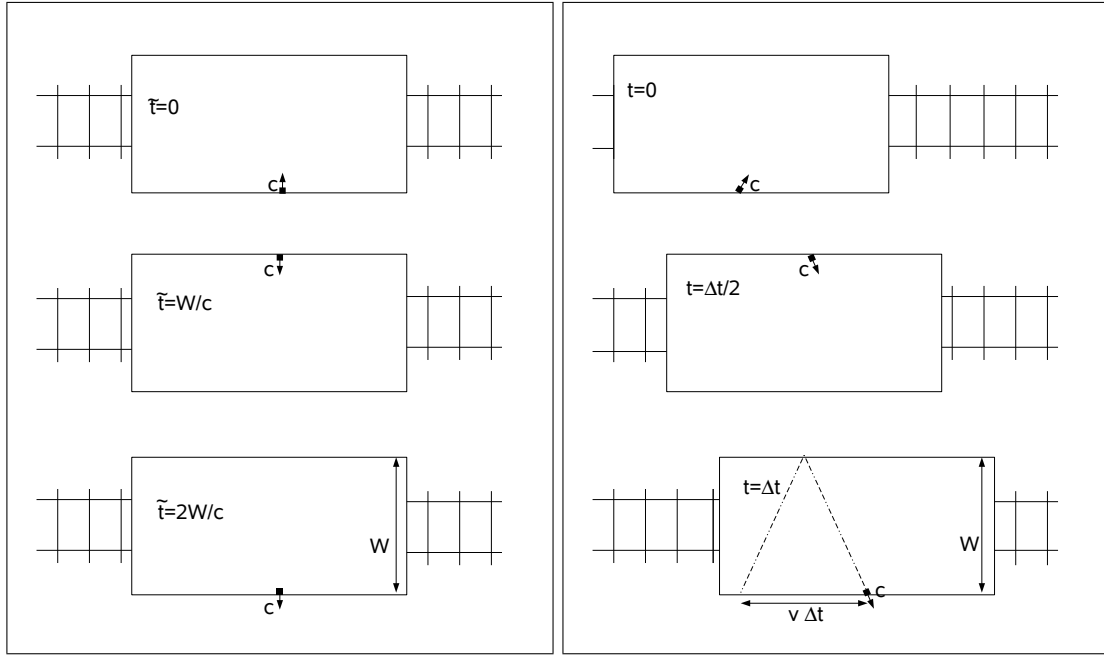


Figure 2.2: Thought experiment number 2. *Left panel:* Events as seen in the carriage frame. *Right panel:* Events as seen in the platform frame.

From this we find

$$\Delta t = \frac{2W}{c} \frac{1}{\sqrt{1-v^2/c^2}} = \Delta \tilde{t} \frac{1}{\sqrt{1-v^2/c^2}}.$$

Introducing the so-called Lorentz factor

$$\gamma = 1/\sqrt{1-v^2/c^2}. \quad (2.1)$$

we can write this result as

$$\Delta t = \gamma \Delta \tilde{t}. \quad (2.2)$$

This shows us that not only simultaneity is relative but also the duration of events. (Notice that we assumed here that the carriage width is the same in both inertial frames. We will come back to this assumption later.)

### Proper time and time of inertial frames.

The best way of measuring time is via some periodic process. Each standard clock is based on such a process. Its time is called the *proper time* of the clock. When instead of a clock we have some physical body it is also useful to introduce the proper time of this body – this can be defined as the time that would be measured by a standard clock moving with this body. Often, the proper time is denoted using the Greek letter  $\tau$ .

As to the *time of inertial frame*, it can be based on the time of a single standard clock co-moving with this frame. It is easy to see how to use such a clock for measuring time of events which occur at the same location as this clock. For remote events one has to have a system of communication between various locations of the frame. The best type of signal for such a system is light, as its speed is given before hand. The clock records the reception time of the light signal sent by the event and the time of the event is this time minus the distance to the event divided by the speed of light. However, an identical result will be obtained if a whole grid of closely spaced standard clocks is

build for the frame, each at rest in this frame, and synchronised with other clocks using light signals. When an event occurs one simply has to use the reading of the clock at the same location as the event.

When it comes to our thought experiment, then  $\Delta\tilde{t}$  could be measured by means of conductor's standard clock only – the pulse is fired from its location and then comes back to its location. Thus,  $\Delta\tilde{t}$  is the proper time of the conductor's clock. In contrast,  $\Delta t$  is the time of the inertial frame of the platform and not a proper time of any of its clocks. Indeed, two clocks, located at points A and B, are required to determine this time interval.

Given these definitions, we conclude that the result (2.2) implies that

$$\Delta t = \gamma \Delta\tau, \quad (2.3)$$

where  $\tau$  is the proper time of some standard clock or other physical body and  $t$  is the time of the inertial frame where this clock moves with the Lorentz factor  $\gamma$ .

Since  $\gamma > 1$  for any  $v \neq 0$  this shows us that  $\Delta t > \Delta\tau$ . What does this imply? One can say that according to the time system of a given inertial frame, any moving clock slows down. Since this result does not depend on the physical nature of the clock mechanism it inevitably implies that all physical processes within a moving body slow down compared to the processes of a similar body at rest, when checked against the time system of the inertial frame where these time measurements are made. This effect is called the *time dilation*.

### 2.2.3 Experiment 3. Length contraction

This time the light pulse is fired from one end of the carriage along the track, gets reflected off the other end and comes back. In the carriage frame the time of pulse journey in both directions is equal to  $\tilde{L}/c$  where  $\tilde{L}$  is the carriage length as measured in the carriage frame (left panel of figure 2.3). Thus, the total time of the pulse journey is

$$\Delta\tilde{t} = 2\tilde{L}/c.$$

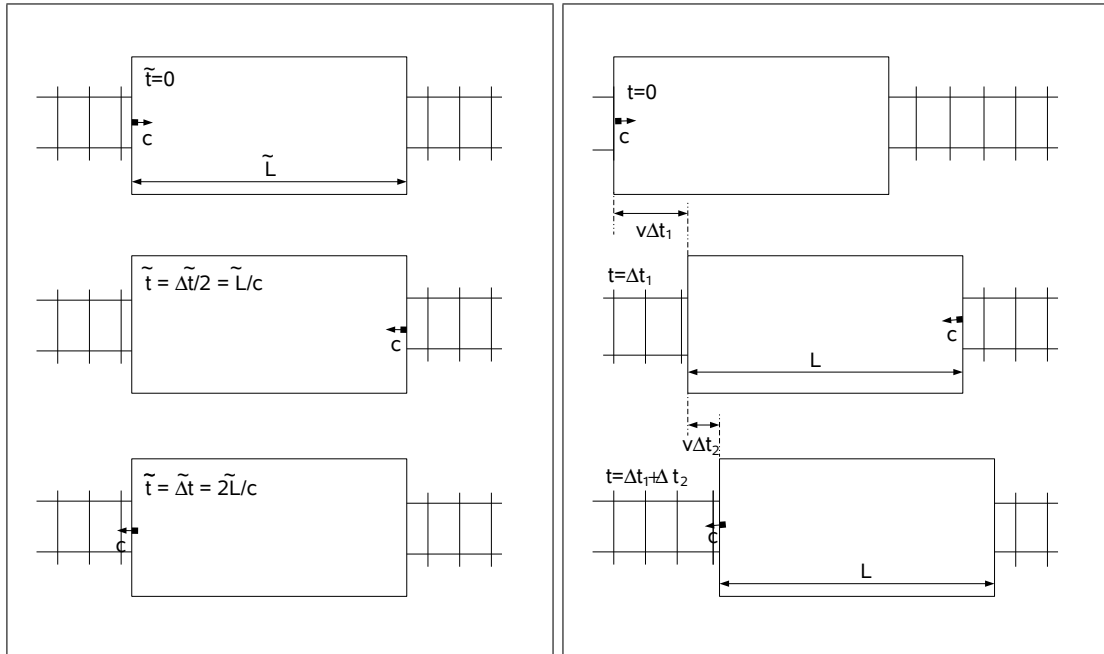


Figure 2.3: Thought experiment number 3. *Left panel:* Events as seen in the carriage frame. *Right panel:* Events as seen in the platform frame.



In the platform frame the first leg of the journey (before the reflection) takes some time  $\Delta t_1$  and the second leg takes  $\Delta t_2$  which is less than  $\Delta t_1$  because of the carriage motion. To find  $\Delta t_1$  we notice that the distance covered by the pulse during the first leg can be expressed as  $c\Delta t_1$  and also as  $L + v\Delta t_1$ , where  $L$  is the length of the carriage as measured in the platform frame (see the right panel of figure 2.3). Thus,

$$\Delta t_1 = L/(c - v).$$

To find  $\Delta t_2$  we notice that the distance covered by the pulse during the second leg can be expressed as  $c\Delta t_2$  and also as  $L - v\Delta t_2$ , (see the right panel of figure 2.3). Thus,

$$\Delta t_2 = L/(c + v).$$

The total time is

$$\Delta t = \Delta t_1 + \Delta t_2 = \frac{L}{c - v} + \frac{L}{c + v} = \frac{2L}{c} \gamma^2.$$

Since  $\Delta \tilde{t}$  is actually a proper time interval we can apply eq.2.3 and write

$$\Delta t = \gamma \Delta \tilde{t} = \gamma \frac{2\tilde{L}}{c}.$$

Combining the last two results we obtain

$$\frac{2L}{c} \gamma^2 = \gamma \frac{2\tilde{L}}{c}$$

which gives us

$$L = \tilde{L}/\gamma \tag{2.4}$$

Thus, the length of the carriage in the platform frame is different from that in the carriage frame. *This shows that if we accept that the speed of light is the same in all inertial frames then we have to get rid of the absolute space as well!*

Similarly to the definition of the proper time interval, the *proper length* of an object, which we will denote as  $L_0$ , is defined as the length measured in the frame where this object is at rest. In this experiment the proper length is  $\tilde{L}$ . Thus we can write

$$L = L_0/\gamma \tag{2.5}$$

In this equation the lengths are measured along the direction of relative motion of two inertial frames. Since,  $\gamma > 1$  we conclude that  $L$  is always shorter than  $L_0$ . Hence the name of this effect – *length contraction*.

Consider two identical bars. When they are rested one alongside the other they have exactly the same length. Set them in relative motion in such a way that they are aligned with the direction of motion (see figure 2.4). In the frame where one bar is at rest the other bar is shorter, and the other way around. At first this may seem contradictory. However, this is in full agreement with the Principle of Relativity. In both frames we observe the same phenomenon – the moving bar becomes shorter. Moreover, the relativity of simultaneity explains how this can be actually possible. In order to measure the length of the moving bar the observer should mark the positions of its ends simultaneously and then to measure the distance between the marks. This way the observer in the rest frame of the upper bar in Figure 2.4 finds that the length of the lower bar is  $L = L_0/\gamma$ . However, the observer in the rest frame of the lower bar finds that the positions are not marked simultaneously, but the position of the left end is marked before the position of the right end (Recall the thought experiment 1 in order to verify this conclusion.). As the result, the distance between the marks,  $L'$  is even smaller than  $L_0/\gamma$ , in fact the actual calculations give  $L' = L_0/\gamma^2 = L/\gamma$ , in agreement with results obtained in the frame of the upper bar.

Lengths measured perpendicular to the direction of relative motion of two inertial frames must be the same in both frames. To show this, consider two identical bars perpendicular to the direction of motion (see Fig. 2.5). In this case there is in no need to know the simultaneous positions of

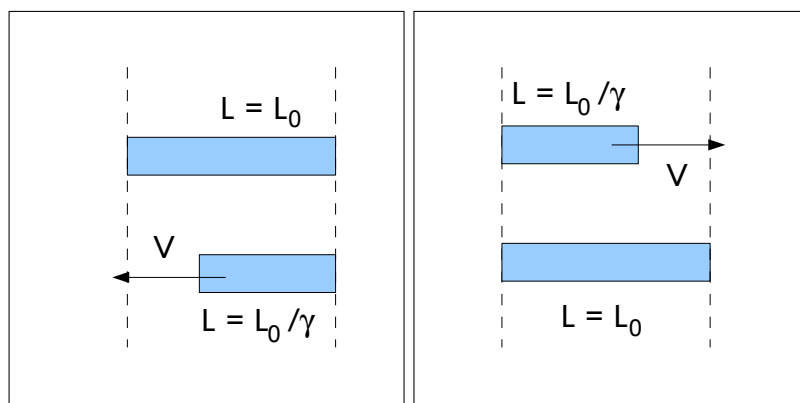


Figure 2.4: Two identical bars are aligned with the direction of their relative motion. In the rest frames of both bars the same phenomenon is observed – a moving bar is shorter.

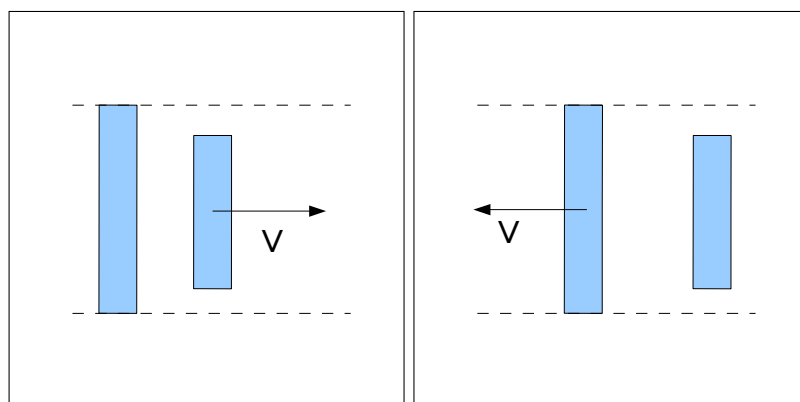


Figure 2.5: Two identical bars are aligned perpendicular to the direction of their relative motion. If the bars did not retain equal lengths then the equivalence of inertial frames would be broken. In one frame a moving bar gets shorter, whereas in the other it gets longer. This would contradict to the Principle of Relativity.

the ends as they do not move in the direction along which the length is measured and, thus, the relativity of simultaneity is no longer important. For example, one could use two strings stretched parallel to the  $x$  axis so that the ends of one of the bars slide along these strings (Fig. 2.5). By observing whether the other bar fits between these strings or not one can decide if it is longer or shorter in the absolute sense – it does not matter which inertial observer makes this observation, the result will be the same. Let us say that the right bar in Figure 2.5 is shorter. Then in the frame of the left bar moving bars contract, whereas in the frame of the right bar moving bars lengthen. This breaks the equivalence of inertial frames postulated in the Relativity Principle. Similarly, we show that the Relativity Principle does not allow the right bar to be longer than the left one. There is no conflict with this principle only if the bars have the same length.

## 2.2.4 Synchronization of clocks

Consider a set of standard clocks placed on the carriage deck along the track. To make sure that all these clocks can be used for consistent time measurements the carriage passengers should synchronise them. This can be done by selecting the clock in the middle to be a reference clock and then by making sure that all other clocks show the same time simultaneously with the reference one. One

way of doing this is by sending light signals at time  $t_0$  from the reference clock to all the others. When a carriage clock receives this signal it should show time  $t = t_0 + l/c$ , where  $l$  is the distance to the reference clock. Now consider set of standard clocks placed on the platform along the track. This set can also be synchronised using the above procedure (now the reference clock will be in the middle of the platform). Recalling the result of the thought experiment 1 we are forced to conclude that to the platform crowd the carriage clocks will appear desynchronised (see fig.2.6) and the other way around – to the carriage passengers the platform clocks will appear desynchronised (see fig.2.7). That is if the passengers inside the carriage standing next to each of the carriage clocks are asked

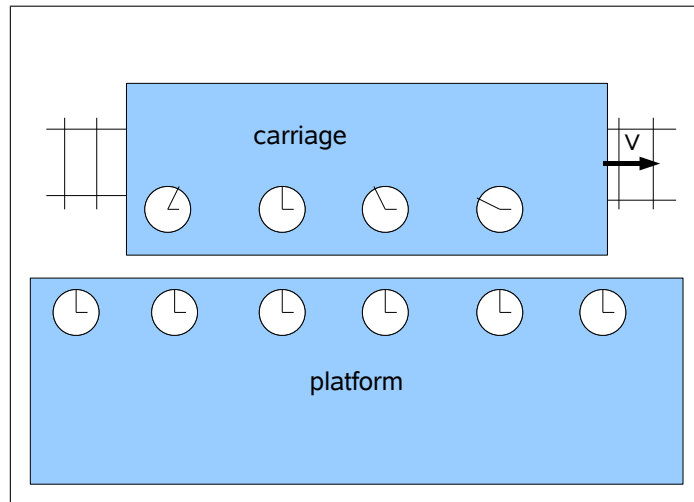


Figure 2.6: Clocks synchronised in the carriage frame appear desynchronised in the platform frame.

to report the time shown by the platform clock which is located right opposite to his clock when his clock shows time  $t_1$  their reports will all have different readings, and the other way around. So in general, a set of clocks synchronised in one inertial frame will appear as desynchronised in another inertial frame moving relative to the first one.

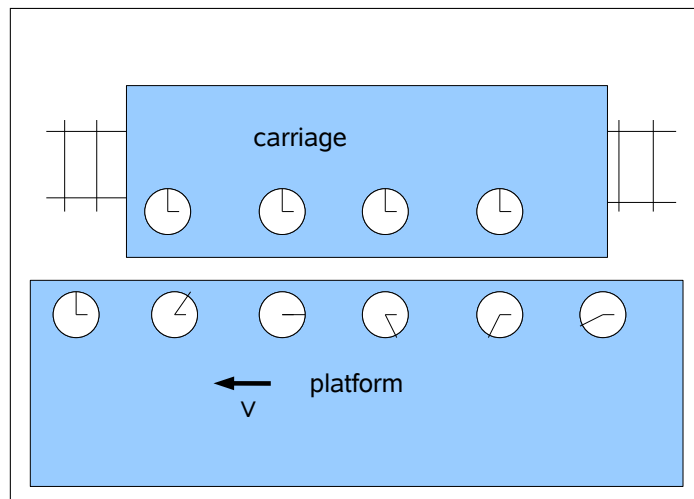


Figure 2.7: Clocks synchronised in the platform frame appear desynchronised in the carriage frame.

## 2.3 Lorentz transformation

In Special Relativity the transition from one inertial frame to another (in standard configuration) is no longer described by the Galilean transformation but by the Lorentz transformation. This transformation ensures that light propagates with the same speed in all inertial frames.

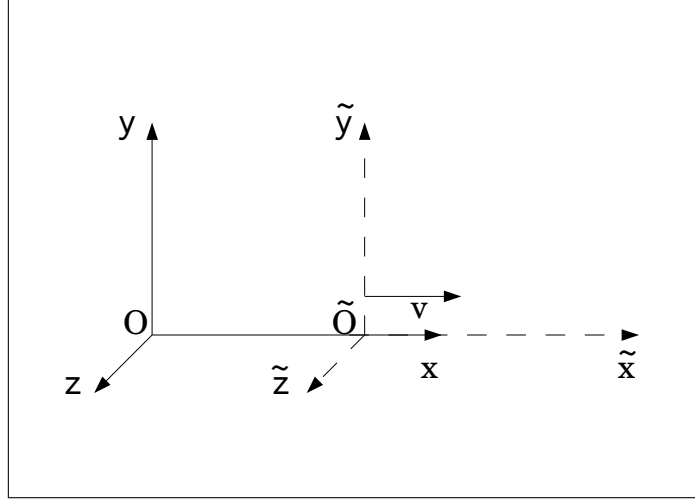


Figure 2.8: Two inertial frames in standard configuration

### 2.3.1 Derivation

The Galilean transformation

$$\begin{cases} t = \tilde{t} \\ x = \tilde{x} + v\tilde{t} \\ y = \tilde{y} \\ z = \tilde{z} \end{cases} \quad \begin{cases} \tilde{t} = t \\ \tilde{x} = x - vt \\ \tilde{y} = y \\ \tilde{z} = z \end{cases} \quad (2.6)$$

is inconsistent with the second postulate of Special Relativity. The new transformation should have the form

$$\begin{cases} t = f(\tilde{t}, \tilde{x}, v) \\ x = g(\tilde{t}, \tilde{x}, v) \\ y = \tilde{y} \\ z = \tilde{z} \end{cases} \quad \begin{cases} \tilde{t} = f(t, x, -v) \\ \tilde{x} = g(t, x, -v) \\ \tilde{y} = y \\ \tilde{z} = z \end{cases} \quad (2.7)$$

just because of the symmetry between the two frames. Indeed, the only difference between the frames  $S$  and  $\tilde{S}$  is the direction of relative motion: If  $\tilde{S}$  moves with speed  $v$  relative to  $S$  then  $S$  moves with speed  $-v$  relative to  $\tilde{S}$ . Hence comes the change in the sign of  $v$  in the equations of direct and inverse transformations (2.7).  $y$  and  $z$  coordinates are invariant because lengths normal to the direction of motion are unchanged (see Sec.2.2.3). Now we need to find functions  $f$  and  $g$ .

(2.6) is a linear transformation. Assume that (2.7) is linear as well (if our derivation fall through we will come back and try something less restrictive). Then

$$x = \gamma(v)\tilde{x} + \delta(v)\tilde{t} + \eta(v). \quad (2.8)$$

Clearly, we should have  $\tilde{x} = -v\tilde{t}$  for any  $\tilde{t}$  if  $x = 0$ . Thus,

$$(-v\gamma + \delta)\tilde{t} + \eta(v) = 0$$

for any  $\tilde{t}$ . This requires

$$\eta = 0 \quad \text{and} \quad \delta = v\gamma.$$

Thus,

$$x = \gamma(\tilde{x} + v\tilde{t}), \quad (2.9)$$

$$\tilde{x} = \gamma(x - vt). \quad (2.10)$$

In principle, the symmetry of direct and inverse transformation is preserved both if  $\gamma(-v) = \gamma(v)$  and  $\gamma(-v) = -\gamma(v)$ . However, it is clear that for  $\tilde{x} \rightarrow +\infty$  we should have  $x \rightarrow +\infty$  as well. This condition selects  $\gamma(-v) = \gamma(v) > 0$ .

Now to the main condition that allows us to fully determine the transformation. Suppose that a light signal is fired at time  $t = \tilde{t} = 0$  in the positive direction of the x axis. (Note that we can always ensure that the standard time-keeping clocks located at the origins of  $S$  and  $\tilde{S}$  show the same time when the origins coincide.) Eventually, the signal location will be  $x = ct$  and  $\tilde{x} = c\tilde{t}$ . Substitute these into eq.(2.10,2.9) to find

$$ct = \gamma c\tilde{t}(1 + v/c), \quad (2.11)$$

$$c\tilde{t} = \gamma ct(1 - v/c). \quad (2.12)$$

Now we substitute  $c\tilde{t}$  from eq.2.12 into eq.2.11

$$ct = \gamma^2 ct(1 - v/c)(1 + v/c)$$

and derive

$$\gamma = 1/\sqrt{(1 - v^2/c^2)}. \quad (2.13)$$

We immediately recognise the Lorentz factor.

In order to find function  $g$  of Eq.2.7 we simply substitute  $x$  from Eq.2.9 into Eq.2.10 and then express  $t$  as a function of  $\tilde{t}$  and  $\tilde{x}$ :

$$\tilde{x} = \gamma[\gamma\tilde{x} + v\gamma\tilde{t} - vt],$$

$$\gamma vt = \gamma^2 v\tilde{t} - \tilde{x}(1 - \gamma^2).$$

It is easy to show that

$$(1 - \gamma^2) = -v^2\gamma^2/c^2.$$

Thus,

$$\gamma vt = \gamma^2 v\tilde{t} + \frac{v^2\gamma^2}{c^2}\tilde{x}$$

and finally

$$t = \gamma\left(\tilde{t} + \frac{v}{c^2}\tilde{x}\right). \quad (2.14)$$

Summarising, the coordinate transformations that keep the speed of light unchanged are

$$\begin{cases} t = \gamma(\tilde{t} + (v/c^2)\tilde{x}) \\ x = \gamma(\tilde{x} + v\tilde{t}) \\ y = \tilde{y} \\ z = \tilde{z} \end{cases} \quad \begin{cases} \tilde{t} = \gamma(t - (v/c^2)x) \\ \tilde{x} = \gamma(x - vt) \\ \tilde{y} = y \\ \tilde{z} = z \end{cases} \quad (2.15)$$

They are due to Lorentz(1853-1928) and Larmor(1857-1942).

### 2.3.2 Newtonian limit

Consider the Lorentz transformations in the case of  $v \ll c$ . This is the realm of our everyday life. In fact even the fastest rockets fly with speeds much less than the speed of light. In this limit

$$\gamma = (1 - v^2/c^2)^{-1/2} \rightarrow 1$$

and the transformation law for the  $x$  coordinate reduces from

$$x = \gamma(\tilde{x} + v\tilde{t}) \quad \tilde{x} = \gamma(x - vt)$$

to the old good Galilean form

$$x = \tilde{x} + v\tilde{t} \quad \tilde{x} = x - vt.$$

This is why the Galilean transformation appears to work so well.

Similarly, we find that the time equation of the Lorentz transformation reduces to

$$t = \tilde{t} + (v/c^2)\tilde{x}$$

If  $T$  and  $L$  are the typical time and length scales of our everyday life, then  $\tilde{t} \sim T$  and  $(v/c^2)\tilde{x} \sim (v/c)(L/c)$ . Let us compare these two terms. If  $L = 1$  km then

$$(v/c)(L/c) \sim (v/c)10^{-5}\text{s} \ll 10^{-5}\text{s}.$$

This is a very short time indeed and much smaller than the timescale  $T$  we normally deal with. Hence, with very high accuracy we may put

$$t = \tilde{t}.$$

This is what has led to the concept absolute time of Newtonian physics.

Formally, the Newtonian limit can be reached by letting  $c \rightarrow +\infty$ , or just replacing  $c$  with  $+\infty$ .

## 2.4 Relativistic velocity "addition"

For  $|v| > c$  the Lorentz factor

$$\gamma = (1 - v^2/c^2)^{-1/2}$$

becomes imaginary and the equations of Lorentz transformation, as well as the time dilation and Lorentz contraction equations, become meaningless. This suggests that  $c$  is the **maximum possible speed in nature**. How can this possibly be the case? If in the frame  $\tilde{S}$  we have a body moving with speed  $\tilde{w} > c/2$  to the right and this frame moves relative to the frame  $S$  with speed  $v > c/2$ , then in the frame  $S$  this body should move with speed  $w > c/2 + c/2 = c$ . However, in this calculation we have used the velocity addition law of Newtonian mechanics,  $w = \tilde{w} + v$ , which is based on the Galilean transformation, not the Lorentz transformation! So what does the Lorentz transformation tell us in this regard?

### 2.4.1 One-dimensional velocity "addition"

Consider a particle moving in the frame  $\tilde{S}$  with speed  $\tilde{w}$  along the  $x$  axis. Then we can write

$$\frac{d\tilde{x}}{d\tilde{t}} = \tilde{w}, \quad \frac{d\tilde{y}}{d\tilde{t}} = \frac{d\tilde{z}}{d\tilde{t}} = 0,$$

and

$$\frac{dx}{dt} = w, \quad \frac{dy}{dt} = \frac{dz}{dt} = 0.$$

From the first two equations of the Lorentz transformation (2.15) one has

$$\begin{cases} dt = \gamma(d\tilde{t} + (v/c^2)d\tilde{x}) \\ dx = \gamma(d\tilde{x} + v d\tilde{t}) \end{cases}. \quad (2.16)$$

Thus,

$$w = \frac{dx}{dt} = \frac{d\tilde{x} + v d\tilde{t}}{d\tilde{t} + (v/c^2)d\tilde{x}} =$$

$$= \frac{d\tilde{x}/d\tilde{t} + v}{1 + (v/c^2)d\tilde{x}/d\tilde{t}} = \frac{\tilde{w} + v}{1 + (v\tilde{w}/c^2)}.$$

The result is the relativistic velocity "addition" law for one-dimensional motion

$$w = \frac{\tilde{w} + v}{1 + (v\tilde{w}/c^2)}. \quad (2.17)$$

Obviously, this differs from the Newtonian result, but reduces to it when  $v, \tilde{w} \ll 1$ .

Now let us consider the numbers  $v = \tilde{w} = c/2$ . In this case  $w \neq c$  but according to Eq.2.17

$$w = \frac{c/2 + c/2}{1 + 1/4} = \frac{4}{5}c < c!$$

Simple analysis of Eq.2.17 shows that for any  $-c < v, \tilde{w} < c$  we obtain  $-c < w < c$ . Thus indeed, the speed of light cannot be exceeded!

Now we can use eqs.(2.17,2.18) in order to verify, that the speed of light is indeed the same in both frames. Substitute  $\tilde{w} = c$  in eq.2.17

$$w = \frac{c + v}{1 + vc/c^2} = c \frac{c + v}{c + v} = c$$

for any  $v$ . Next substitute  $\tilde{w} = -c$

$$w = \frac{-c + v}{1 - vc/c^2} = c \frac{-c + v}{c - v} = -c$$

for any  $v$ . In fact, it is easy to show that for any  $-c < \tilde{w} < c$  and  $-c < v < c$  eq.2.17 gives  $-c < w < c$ . This result is consistent with our earlier conclusion that  $c$  must be the highest possible speed.

The inverse to Eq.2.17 law is the same as eq.2.17 up to the sign of  $v$ :

$$\tilde{w} = \frac{w - v}{1 - (vw/c^2)}. \quad (2.18)$$

The change of sign of  $v$  in this equation follows directly from the change of sign of  $v$  in the inverse Lorentz transformation (2.15). (Another way of deriving eq.2.18 is via finding  $\tilde{w}$  from eq.2.17)

### 2.4.2 Three-dimensional velocity "addition"

In Sec.2.4.1 we considered the motion only along the x axis. If instead we allow the particle to move in arbitrary directions then using the same method we can derive the following result

$$w^x = \frac{w^{\tilde{x}} + v}{1 + vw^{\tilde{x}}/c^2}, \quad w^y = \frac{w^{\tilde{y}}}{\gamma(1 + vw^{\tilde{x}}/c^2)}, \quad w^z = \frac{w^{\tilde{z}}}{\gamma(1 + vw^{\tilde{x}}/c^2)}. \quad (2.19)$$

for the direct transformation and

$$w^{\tilde{x}} = \frac{w^x - v}{1 - vw^x/c^2}, \quad w^{\tilde{y}} = \frac{w^y}{\gamma(1 - vw^x/c^2)}, \quad w^{\tilde{z}} = \frac{w^z}{\gamma(1 - vw^x/c^2)}. \quad (2.20)$$

for the inverse one. (Notice again that the only difference in the direct and the inverse equations is the sign of  $v$ .)

## 2.5 Aberration of light

A light signal propagates at the angle  $\tilde{\theta}$  to the x axis of frame  $\tilde{S}$ . What is the corresponding angle measured in frame  $S$ .

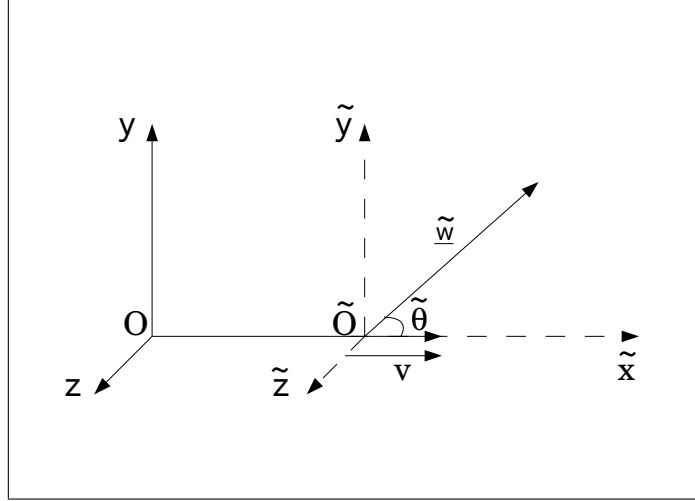


Figure 2.9:

We can always rotate both frames about the x-axis so that the signal is confined to the XOY plane as in figure 2.9. Then the velocity vector of the signal has the following components

$$w^{\tilde{x}} = c \cos \tilde{\theta}, \quad w^{\tilde{y}} = c \sin \tilde{\theta}, \quad w^{\tilde{z}} = 0,$$

in frame  $\tilde{S}$  and

$$w^x = c \cos \theta, \quad w^y = c \sin \theta, \quad w^z = 0,$$

in frame  $S$ . Substitute these into the first two equations in (2.19) and find that

$$\cos \theta = \frac{\cos \tilde{\theta} + \beta}{1 + \beta \cos \tilde{\theta}} \quad (2.21)$$

$$\sin \theta = \frac{\sin \tilde{\theta}}{\gamma(1 + \beta \cos \tilde{\theta})}, \quad (2.22)$$

where  $\beta = v/c$ . From these one also finds

$$\tan \theta = \frac{\sin \tilde{\theta}}{\gamma(\beta + \cos \tilde{\theta})}. \quad (2.23)$$

Let us analyse the results. Differentiate eq.2.21 with respect to  $\beta$

$$-\sin \theta \frac{d\theta}{d\beta} = \frac{\sin^2 \tilde{\theta}}{(1 + \beta \cos \tilde{\theta})^2}$$

and substitute  $\sin \theta$  from eq.2.22 to obtain

$$\frac{d\theta}{d\beta} = -\frac{\gamma \sin \tilde{\theta}}{1 + \beta \cos \tilde{\theta}}.$$

This derivative is negative or zero ( $\tilde{\theta} = 0, \pi$ ). Thus,  $\theta$  decreases with  $\beta$  and the direction of light signal is closer to the x direction in frame  $S$ . To visualise this effect imagine an isotropic source of light moving with speed close to the speed of light. Its light emission will be beamed in the direction of motion (see fig.2.10). In fact, eq.2.21 shows that in the limit  $v \rightarrow c$  ( $\beta \rightarrow 1$ ) we have  $\cos \theta \rightarrow 1$  and hence  $\theta \rightarrow 0$  (with exception of  $\tilde{\theta} = \pi$  which always gives  $\theta = \pi$ ).

In the Newtonian limit ( $c \rightarrow \infty$ ) Equations (2.21-2.23) reduce to

$$\theta = \tilde{\theta},$$

showing that there is no aberration of light effect.



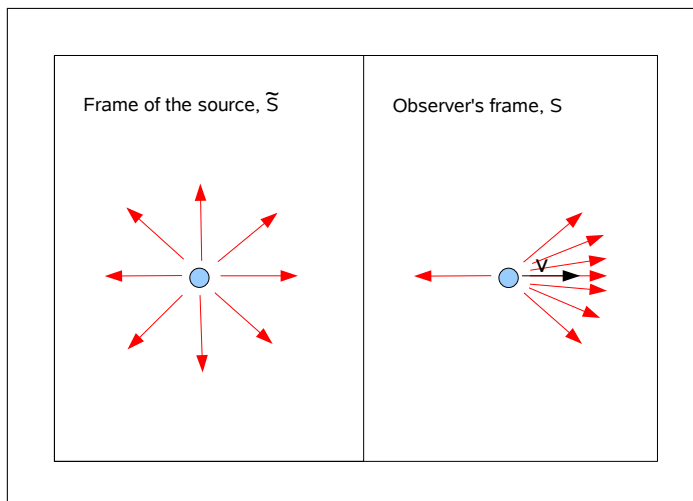


Figure 2.10: The effect of aberration of light on the radiation of moving sources. If in the rest frame of some light source its radiation is isotropic (left panel) then in the frame where the source is moving with relativistic speed its radiation is beamed in the direction of motion (right panel).

## 2.6 Doppler effect

The Doppler effect describes the variation of periodic light signal due to the motion of emitter relative to receiver.

### 2.6.1 Transverse Doppler effect

*Problem setup:* A source (emitter) of periodic light signal (e.g. monochromatic wave) moves perpendicular to the line of sight of some observer with speed  $v$ . Find the frequency  $\nu$  of the signal received by the observer if in the source frame the frequency of emitted light is  $\nu_0$ .

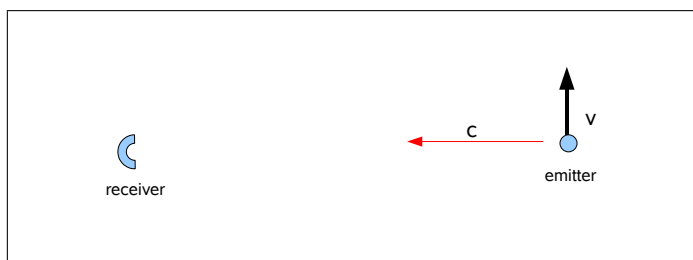


Figure 2.11:

*Solution:*  $T_0 = 1/\nu_0$  is the period of the emitter in its rest frame. Suppose, that in the frame of the emitter this signal is emitted during the time  $\Delta t_0 = NT_0$ , where  $N$  is the number of produced wave crests. By definition, both these times are the proper times of the emitter.

In the frame of receiver the emission process takes a longer time,  $\Delta t_e$ , given by the time dilation formula

$$\Delta t_e = \gamma \Delta t_0.$$

Assuming that the distance between the emitter and the receiver is so large that its change during the emission is negligible, the signal is received during the same time interval of the same length,  $\Delta t_r = \Delta t_e$ . Since the number of crests remains the same, the period of the received light is

$$T_r = \Delta t_r / N = \gamma T_0. \quad (2.24)$$

Thus, the frequency of the received signal is

$$\nu = \frac{1}{T_r} = \frac{1}{\gamma T_0} = \frac{\nu_0}{\gamma}. \quad (2.25)$$

Since  $\gamma > 1$  we have  $\nu < \nu_0$ . One can see that the transverse Doppler effect is entirely due to the time dilation.

### 2.6.2 Radial Doppler effect

*Problem setup:* Now we consider the case where the source is moving along the line of sight and so the distance between the emitter and the receiver is increasing during the process of emission. This has an additional effect on the frequency of the received signal.

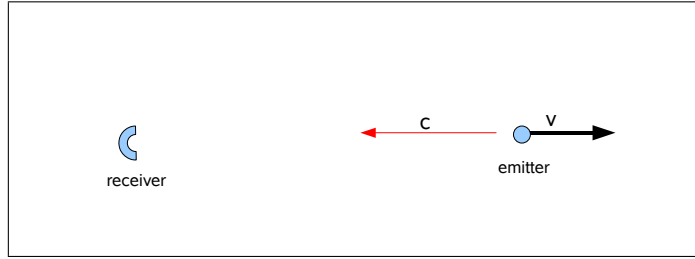


Figure 2.12:

*Solution:* The emission time as measured in the frame of the observer is still given by the time dilation formula

$$\Delta t_e = \gamma \Delta t_0.$$

During the time  $\Delta t_e$  the emitter moves away from the observer by the distance

$$\Delta l = \Delta t_e v.$$

This leads to the increase in the period of reception time by  $\Delta l/c$ . Thus, the total reception time in this case is

$$\Delta t_r = \Delta t_e + \frac{\Delta t_e v}{c} = \Delta t_0 \gamma (1 + \beta) = \Delta t_0 \sqrt{\frac{1 + \beta}{1 - \beta}}.$$

This yields the observed period of the received signal

$$T_r = \frac{\Delta t_r}{N} = \frac{\Delta t_0}{N} \sqrt{\frac{1 + \beta}{1 - \beta}} = T_0 \sqrt{\frac{1 + \beta}{1 - \beta}},$$

and the observed frequency

$$\nu = \nu_0 \sqrt{\frac{1 - \beta}{1 + \beta}}. \quad (2.26)$$

Notice that in this case  $\nu < \nu_0$  if  $v > 0$  (the emitter is moving away from the receiver) and  $\nu > \nu_0$  if  $v < 0$  (the emitter is approaching the receiver).

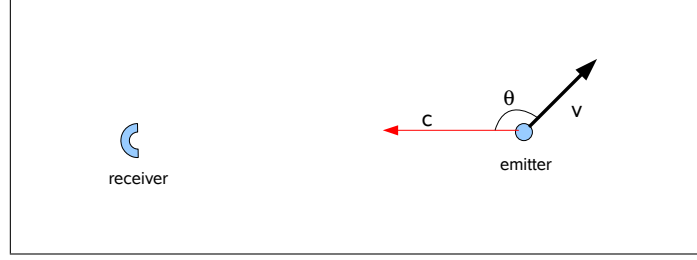


Figure 2.13:

### 2.6.3 General case

*Problem setup:* In this case, the source is moving at angle  $\theta$  with the line of sight. This includes the transverse ( $\theta = \pi/2$ ) and the radial cases ( $\theta = 0, \pi$ ).

*Solution:* The only difference compared to the radial case is in the rate of increase of the distance between the emitter and the receiver. Now

$$\Delta l = -\Delta t_e v \cos \theta,$$

where we assume that  $v > 0$  and that the emitter is far away from the receiver. In order to see this, consider the triangle with sides  $l_0$ ,  $l$ , and  $v\Delta t_e$ , where  $l_0$  is the distance between the emitter and the receiver at the beginning of emission,  $l$  is this distance at the end of emission and  $v\Delta t_e$  is the distance covered by the emitter during the time of emission. From this triangle,

$$l^2 = l_0^2 + (v\Delta t_e)^2 - 2l_0 v\Delta t_e \cos \theta,$$

or

$$l^2 = l_0^2 \left( 1 - 2 \cos \theta \left( \frac{v\Delta t_e}{l_0} \right) + \left( \frac{v\Delta t_e}{l_0} \right)^2 \right).$$

For a distant emitter  $v\Delta t_e/l_0 \ll 1$  and we have

$$l^2 \simeq l_0^2 \left( 1 - 2 \cos \theta \left( \frac{v\Delta t_e}{l_0} \right) \right).$$

Moreover,

$$l \simeq l_0 \left( 1 - 2 \cos \theta \left( \frac{v\Delta t_e}{l_0} \right) \right)^{1/2} \simeq l_0 \left( 1 - \cos \theta \left( \frac{v\Delta t_e}{l_0} \right) \right) = l_0 - \cos \theta v\Delta t_e.$$

Repeating the calculations of the radial case we then find that

$$T_r = T_0 \gamma (1 - \beta \cos \theta),$$

and

$$\nu = \frac{\nu_0}{\gamma (1 - \beta \cos \theta)}. \quad (2.27)$$

Now it is less clear whether  $\nu > \nu_0$  or otherwise, and further analysis is required.

In the Newtonian limit ( $c \rightarrow \infty$ ) one has  $\nu = \nu_0$ , so no effect is seen. In the case of  $\beta \ll 1$ , the Taylor expansion of Eq.2.27 gives

$$\nu = \nu_0 (1 + \beta \cos \theta + O(\beta^2)).$$



# Chapter 3

## Space-time

### 3.1 Minkowski diagrams

The difference between the Newtonian and relativistic visions of space and time is nicely illustrated with the help of Minkowski diagrams. The idea is to represent instantaneous localised events as points of a two dimensional graph. Surely, some information is lost because of we really need four dimensions to fully represent space time events. So what is actually shown is time and one of the spatial coordinates (usually  $x$ ) as measured in some inertial frame, and it is assumed that other coordinates of displayed events vanish,  $y = z = 0$ .

The left panel of Fig.3.1 is a Newtonian version of Minkowski diagram. Along the horizontal axis we show the  $x$  coordinate of some inertial frame (frame  $S$ ) and along the vertical axis its time multiplied by the speed of light ( $ct$  has the dimension of length, just like  $x$ ). Any two events, like A and B in this figure, that belong to a line parallel to the time axis, are simultaneous in this frame. The whole such line represents all real and potential events occurring at the same time  $t = t_0$  and shows the spatial separation between them. One may say that it represents the space at time  $t = t_0$ . Any particle is represented by a line on this diagram, called the *world line* of the particle. For example, the time axis is the world line of the frame's origin. The world line of a particle moving with constant speed along the  $x$  axis of the frame is also straight but now it is inclined to the time axis. The inclination angle varies between 0 and  $\pi/2$ , depending on the particle speed.

Consider another inertial frame (frame  $\tilde{S}$ ) which has the same orientation of axes and moves along the  $x$ -axis with constant speed  $v$ . (As always, we also assume that the origins of both frames coincide at time  $\tilde{t} = t = 0$ , where  $\tilde{t}$  is the time of the frame  $\tilde{S}$ ). The line  $\tilde{t}$  is the world line of the origin of the frame  $\tilde{S}$  – it can be considered as the time axis of this frame. The angle between the  $t$  and  $\tilde{t}$  axes follows from the equations of motion of the origin

$$x = \beta ct, \quad \text{where } \beta = v/c.$$

From this we have

$$\tan \theta = \frac{x}{ct} = \beta. \tag{3.1}$$

According to the Galilean transformation, the events A and B, shown in the left panel of Fig.3.1, are also simultaneous in the frame  $\tilde{S}$  and occur at exactly the same time,  $\tilde{t} = t_0$ . Thus, the space of this frame at time  $\tilde{t} = t_0$  is the same as the space of the frame  $S$  at time  $t = t_0$ . Both spaces are represented by the same line parallel to the  $x$  axis. For the same reason the  $x$  axis coincides with the  $\tilde{x}$  axis in the diagram - they show the space at  $t = \tilde{t} = 0$ .

The right panel of Fig.3.1 is a Minkowski diagram of special relativity for the same combination of two inertial frames. One obvious difference is that  $\tilde{x}$  axis does no longer coincide with the  $x$  axis. Indeed, the axis shows all events that occur at time  $\tilde{t} = 0$ . They are simultaneous in the “moving” frame ( $\tilde{S}$ ). However, they are not simultaneous in the “stationary” frame ( $S$ ) and thus cannot be represented by the  $x$  axis of the diagram. To find the equation of the  $\tilde{x}$  axis we should use the equation of Lorentz transformation

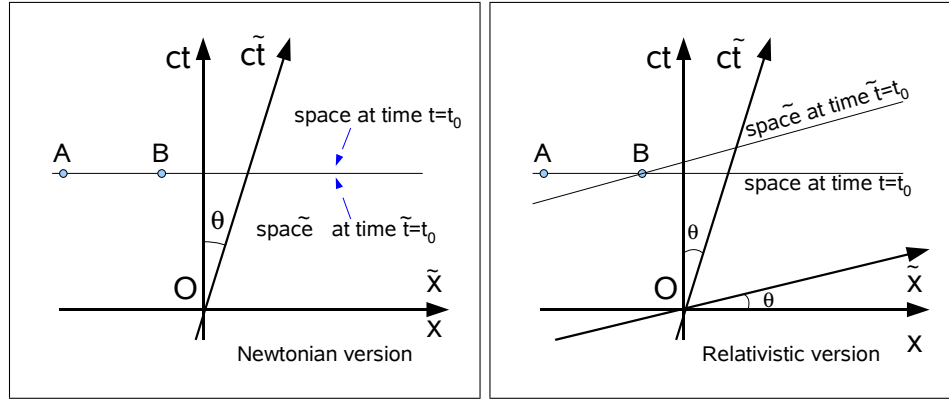


Figure 3.1: Newtonian (left panel) and Minkowskian (right panel) visions of space and time.

$$c\tilde{t} = \gamma(ct - \beta x).$$

Putting  $\tilde{t} = 0$  we find

$$ct = \beta x,$$

which is the equation of a straight line passing through the origin at the angle  $\theta = \arctan \beta$  (see Eq.3.1) with the  $x$  axis. This line shows the space of frame  $\tilde{S}$  at the time  $\tilde{t} = 0$ , the set of all actual and potential events at time  $\tilde{t} = 0$  and their spacial separation. Similarly, one deduces that the space of frame  $\tilde{S}$  at  $\tilde{t} \neq 0$  is represented by a line parallel to the  $\tilde{x}$  axis. It is obvious from this diagram that these two frames have different spaces and different times!

Since  $\beta < 1$  we have  $\theta < \pi/4$ . Thus, the time and the  $x$  axis of frame  $\tilde{S}$  never merge, with the time axis staying above and the  $x$  axis below the line  $ct = x$  which is the world line of a light signal sent from the origin at time  $t = 0$ .

## 3.2 Space-time

In Newtonian physics the physical space Euclidean. By definition, a space is Euclidean if one can introduce such coordinates  $x, y, z$  that its metric form becomes

$$dl^2 = dx^2 + dy^2 + dz^2 \quad (3.2)$$

at every point of the space. Such coordinates are called Cartesian coordinates. There are infinitely many different Cartesian coordinates. For example, one can obtain one system of Cartesian coordinates from another via rotation about the  $z$  axis by angle  $\theta$ :

$$\begin{aligned} \tilde{x} &= x \cos \theta + y \sin \theta, \\ \tilde{y} &= -x \sin \theta + y \cos \theta, \\ \tilde{z} &= z. \end{aligned}$$

This transformation keeps the metric form invariant;

$$dl^2 = d\tilde{x}^2 + d\tilde{y}^2 + d\tilde{z}^2. \quad (3.3)$$

In Newtonian Physics it is assumed that the lengths of objects and distances between them do not depend on the selection of inertial frame where the length measurements are carried out. The geometry of Physical Space is reassuringly the same for every one. In Special Relativity the lengths of objects and distances between them do depend on the inertial frame used for the length measurements. The metric form (3.2) is not invariant under the Lorentz transformation. The

geometry of the Physical Space is no longer reassuringly fixed. In mathematical terms the Physical Space can no longer be modelled as a metric space. This almost amounts to saying that the Physical Space is not quite a proper physical reality.

Einstein's teacher at Zurich Polytechnic, professor Minkowski (1864-1909), discovered that the following "generalised metric form"

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \quad (3.4)$$

is in fact invariant under the Lorentz transformations. Obviously  $dy^2 + dz^2 = d\tilde{y}^2 + d\tilde{z}^2$ , so we only need to show that

$$-c^2 dt^2 + dx^2 = -c^2 d\tilde{t}^2 + d\tilde{x}^2.$$

The calculations are most concise when we use the Lorentz transformation written in terms of hyperbolic functions

$$\begin{aligned} ct &= c\tilde{t} \cosh \theta + \tilde{x} \sinh \theta, \\ x &= \tilde{x} \cosh \theta + c\tilde{t} \sinh \theta. \end{aligned} \quad (3.5)$$

To see that the Lorentz transformation can indeed be written this way, we first write it as

$$\begin{aligned} ct &= \gamma(c\tilde{t} + \beta\tilde{x}) \\ x &= \gamma(\tilde{x} + \beta c\tilde{t}), \end{aligned} \quad (3.6)$$

and then introduce  $\theta$  via  $\sinh \theta = \gamma\beta$ , from which it follows that  $\gamma = \cosh \theta$ .

From the first two equations in (3.5) we have

$$\begin{aligned} cdt &= c\tilde{d}t \cosh \theta + d\tilde{x} \sinh \theta, \\ dx &= d\tilde{x} \cosh \theta + c\tilde{d}t \sinh \theta. \end{aligned}$$

Then

$$\begin{aligned} -c^2 dt^2 + dx^2 &= -(c\tilde{d}t \cosh \theta + d\tilde{x} \sinh \theta)^2 + (d\tilde{x} \cosh \theta + c\tilde{d}t \sinh \theta)^2 = \\ &= -c^2 \tilde{d}t^2 \cosh^2 \theta - 2cd\tilde{x}\tilde{d}t \cosh \theta \sinh \theta - d\tilde{x}^2 \sinh^2 \theta + \\ &+ c^2 \tilde{d}t^2 \sinh^2 \theta + 2cd\tilde{x}\tilde{d}t \cosh \theta \sinh \theta + d\tilde{x}^2 \cosh^2 \theta = \\ &= -c^2 \tilde{d}t^2 (\cosh^2 \theta - \sinh^2 \theta) + d\tilde{x}^2 (\cosh^2 \theta - \sinh^2 \theta) = \\ &= -c^2 \tilde{d}t^2 + d\tilde{x}^2. \end{aligned}$$

To make the last step we use the well known result  $\cosh^2 \theta - \sinh^2 \theta = 1$ . Thus, the quadratic form (3.4) is indeed Lorentz invariant and this fact suggests that space and time should be united into a single 4-dimensional metric space with metric form (3.4)!!! Like in the case of Minkowski diagrams points of this space are called *events*. This space is called *the space-time or the Minkowski space-time*.  $ds^2$  is the generalised distance between points of this space – it is called the *space-time interval*. Denoting

$$x^0 = ct, x^1 = x, x^2 = y, x^3 = z \quad (3.7)$$

we can rewrite eq.3.4 as

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2. \quad (3.8)$$

Given the similarity of this metric form with the metric form of Euclidean space in Cartesian coordinates the space-time of Special relativity is often called *pseudo-Euclidean* (where "pseudo" refers to the sign "-" in front of the first term) and the coordinates (3.7) are often called *pseudo-Cartesian*. Thus, there is a one-to-one correspondence between inertial frames with their time and Cartesian grid and systems of pseudo-Cartesian coordinates in space-time! The Lorentz transformation can now be considered as a transformation from one system of pseudo-Cartesian coordinates to another system of pseudo-Cartesian coordinates in space-time. Minkowski diagrams are simply maps of space-time onto a 2D Euclidean plane of a paper sheet. Notice that the usual Euclidean distances between the points on these diagrams do not reflect the generalised distances between the corresponding events in space-time.

In arbitrary curvilinear coordinates the space-time metric becomes

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \quad (3.9)$$

where the coefficients  $g_{\mu\nu}$  are quite arbitrary as well. They can be written as components of a  $4 \times 4$  matrix

$$\begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix}$$

Here and in the rest of the notes we will assume that Greek indexes run from 0 to 3 whereas Latin indexes run from 1 to 3. One can always impose the symmetry condition on the metric form

$$g_{\mu\nu} = g_{\nu\mu}, \quad (3.10)$$

and this is what we will always assume here. In pseudo-Cartesian coordinates

$$g_{00} = -1, \quad g_{11} = g_{22} = g_{33} = 1, \\ \text{and } g_{\mu\nu} = 0 \quad \text{if } \mu \neq \nu, \quad (3.11)$$

or in the matrix form

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.12)$$

### 3.3 Light cone

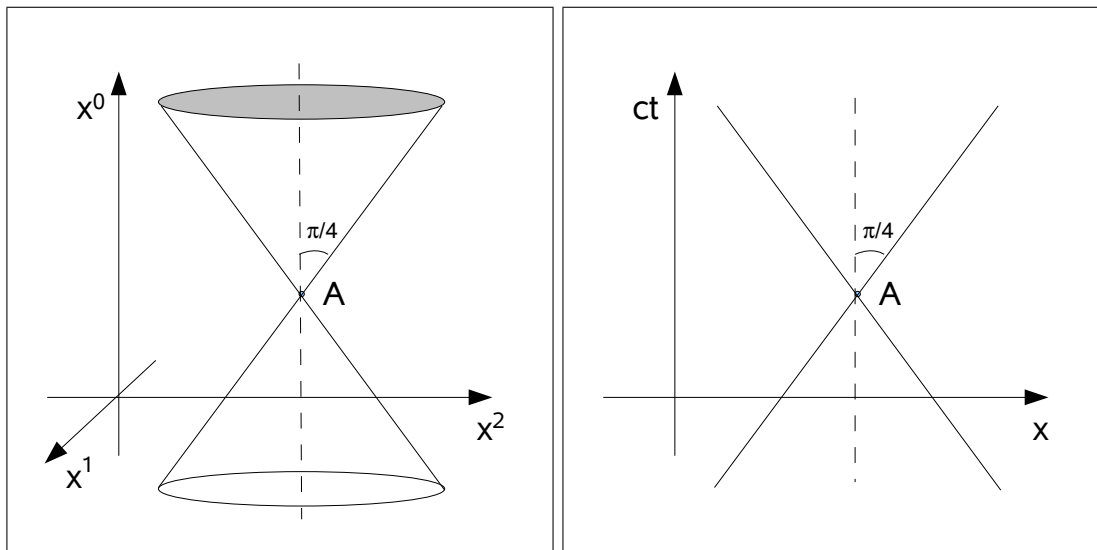


Figure 3.2: *Left panel:* The light cone of event A. *Right panel:* Appearance of the light cone in Minkowski diagram.

Select some event, say event A, in space-time. Let  $x_A^\nu$  to be the pseudo-Cartesian coordinates of this event. Consider the surface determined by

$$\Delta s^2 = 0$$



or

$$(\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2 = (\Delta x^0)^2, \quad (3.13)$$

where  $\Delta x^\nu = x^\nu - x_A^\nu$ . This surface is called the light cone of event A. It is made out of all events that are connected to A by world lines of light signals. It is easy to see that this is indeed a four-dimensional cone whose half-opening angle is  $\pi/4 = 45^\circ$ , whose symmetry axis is parallel to the  $x^0$  axis, and whose tip is located at A (see the left panel of fig.3.2).

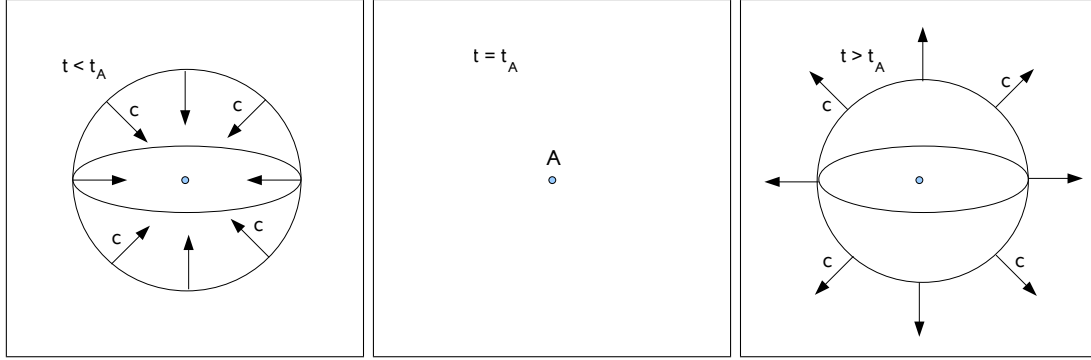


Figure 3.3: In 3-dimensional physical space of any inertial frame the lower half of the light cone of event A corresponds to a spherical light front converging of the location of event A (left panel,  $t < t_A$ ), and its upper half corresponds to a spherical light front diverging from the location of this event (right panel,  $t > t_A$ ).

In Minkowski diagrams the light cone of event A gives two straight lines passing through A at the angle of  $\pi/4$  to the  $x^0$  axis (time axis). They are the world lines of two light signals passing through A (see the right panel of fig.3.2).

In the physical space of any inertial frame the bottom half of the light cone corresponds to a spherical light front converging onto the spacial location of A

$$c(t - t_A) = -\sqrt{(x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2} \quad (3.14)$$

and its upper half corresponds to a spherical light front diverging from the spacial location of A

$$c(t - t_A) = +\sqrt{(x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2} \quad (3.15)$$

(see fig.3.3). Due to the length contraction effect shapes and sizes of physical objects as well as distances between them differ in different inertial frames. However, the shape and size of these light fronts remain the same, being described by eqs.(3.14,3.15) in *all inertial frames*.

Consider a particle travelling with constant velocity in some inertial frame. Then its world line is a straight line. Provided that it passes through A its equation is

$$(x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2 = v^2(t - t_A)^2 < c^2(t - t_A)^2.$$

Thus, such a world line is located inside the light cone of A (see fig.3.4).

### 3.4 Causal structure of space-time

Physical objects interact with each other by means of waves or particles they produce. Let us call these interaction agents as *signals*. If an event A is caused or triggered by some other event B then this is arranged via a signal that B sends to A and that reaches the location of A before A occurs. Since the signal speed must be less or equal than the speed of light its world line is inclined at an angle to the time axis that is less than or equal to  $\pi/4$ . Thus the event B must lie inside the lower half of the light cone of event A (see the left panel of fig.3.5). On the other hand, the world lines

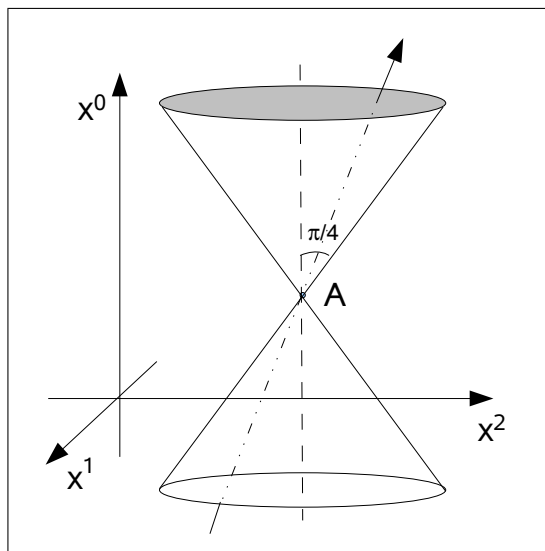


Figure 3.4: The world line of any physical object that passes through A lies inside the light cone of A.

of all possible signals sent by the event A fill the upper half of the light cone of A. Thus, the events that may be caused by A are located inside the upper half of the light cone of A (see the left panel of fig.3.5). As to the events that lie outside of the light cone of A they can neither cause nor be caused by A – they are causally unrelated to A.

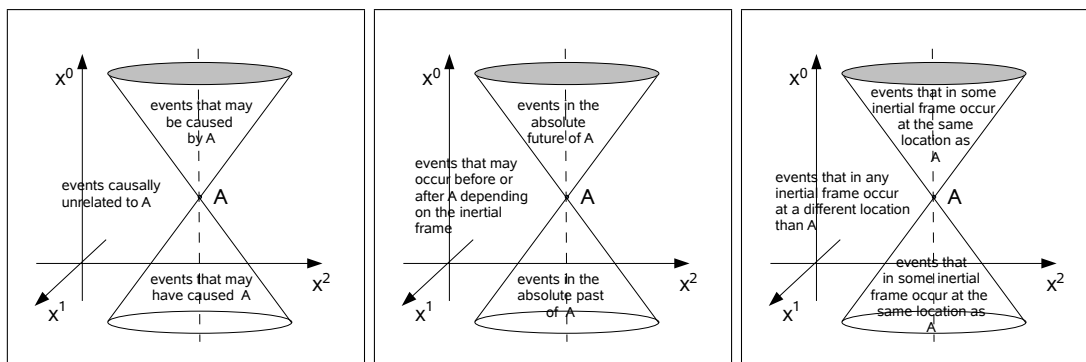


Figure 3.5: The light cone of event A, and the causal, spatial, and temporal connection of this event to all other events.

Consider an event B inside the light cone of A. Construct a straight line connecting A and B. Since this line is inside the light cone it is inclined at an angle less than  $\pi/4$  with the time axis and can be a world line of some physical object. For example this can be a grid point of some inertial frame. In this frame A and B have the same spatial location. For any event C located outside the light cone of A the line connecting A and C cannot be a world line of any physical object as this would imply motion with speed exceeding the speed of light. Thus, in any inertial frame A and C will have different spatial locations (see the right panel of fig.3.5).

Comparing this with the causal structure of space time we conclude that the events that can be causally connected are also the events that can have the same location in some inertial frame. On the contrary, the events that cannot be causally connected are spatially separated in all inertial frames.

The casual structure is also related with the temporal order of events. The events that can be caused by A are in the future of A ( $t > t_A$ ) in all inertial frames and the events that could cause A are in the past of A ( $t < t_A$ ) in all inertial frames (This is because  $\theta$  in the right panel of Figure 3.1 cannot exceed  $\pi/2$ ). Thus, the events that have unambiguous temporal location with respect to A are all inside the light cone of A. On the other, hand the events that cannot be causally connected to A (they are outside of the light cone of A) can precede A in some frames and follow it in other frames (See Figure 3.5).

### 3.5 Types of space-time intervals

For any event B that lies outside of the light cone of event A the space-time interval separating these two events is positive,

$$\Delta s_{AB}^2 = -(\Delta x_{AB}^0)^2 + \sum_{i=1}^3 (\Delta x_{AB}^i)^2 > 0, \quad (3.16)$$

where  $\Delta x_{AB}^\nu = x_B^\nu - x_A^\nu$ . Such intervals are called *space-like*. It is easy to show that there always exists an inertial frame where A and B are simultaneous and

$$\Delta s_{AB}^2 = \sum_{i=1}^3 (\Delta x_{AB}^i)^2.$$

In some inertial frames moving relative to this one A occurs before B, in others B occurs before A. Thus, the temporal order of A and B depends on the choice of inertial frame.

For any event B on the surface of the light cone of event A the space-time interval separating these two events is zero,

$$\Delta s_{AB}^2 = 0. \quad (3.17)$$

Such intervals are called *null*.

For any event B that lies inside the light cone of event A the space-time interval separating these two events is negative,

$$\Delta s_{AB}^2 = -(\Delta x_{AB}^0)^2 + \sum_{i=1}^3 (\Delta x_{AB}^i)^2 < 0. \quad (3.18)$$

Such intervals are called *time-like*. It is easy to show that there always exists an inertial frame where A and B are separated only in time but not in space and hence

$$\Delta s_{AB}^2 = -(\Delta x_{AB}^0)^2. \quad (3.19)$$

If B lies inside the upper half of the light cone of A then in any inertial frame A precedes B and otherwise if B lies inside the lower half of the light cone of A then in any inertial frame B precedes A. To see we use the Minkowski diagram (figure 3.6) that shows A, B and the axes of two inertial frames centred on the event A (this can always be done simply by resetting the zero time and changing the location of origin.). As we have already seen the angle  $\theta$  is always less than  $\pi/4$ . Since the lines of simultaneity of frame  $\tilde{S}$  also make angle  $\theta$  with the x axis the event A precedes the event B in this frame as well as in frame  $S$ . One can say that B belongs to *the absolute future of A*. Similarly, we see that any event C that lies inside the lower half of the light cone of A always precedes A and thus belongs to *the absolute past of A*.

These temporal properties are summarised in the middle panel of fig.3.5.

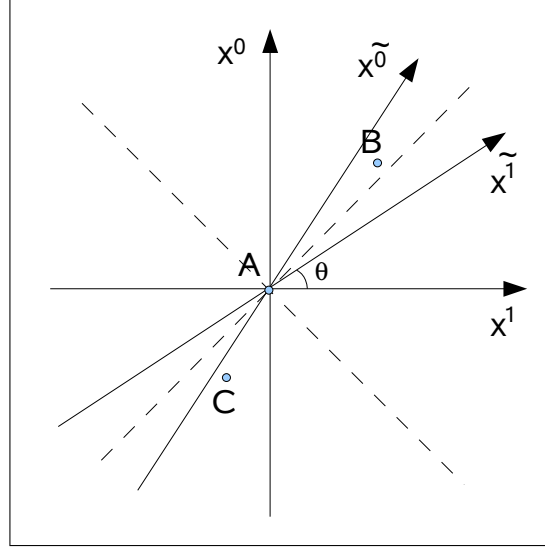


Figure 3.6:

## 3.6 Vectors

The visualisation/interpretation of vectors as straight arrows whose length describes the vector magnitude works fine for Euclidean space but is not fully suitable for curved manifolds (e.g. surfaces in Euclidean space) and space-time. To see the shortcomings of this interpretation in space-time consider the light cone of event A and a straight line on the surface of this cone (see fig.3.7). Take two points, B and C, on this line. They seem to define two different vectors  $\vec{AB}$  and  $\vec{AC}$ . However, the generalised distances (squared) between these points are equal to zero

$$\Delta s_{AB}^2 = \Delta s_{AC}^2 = 0$$

and so are the generalised lengths of the arrows. Thus, arrows  $\vec{AB}$  and  $\vec{AC}$  have the same lengths and the same directions! Moreover, within this interpretation it is impossible to address the difference between vectors that connect points with space-like separation and vectors that connect points with time-like separation. Below we give the modern generalised definition of geometric vectors due to Cartan, which is very robust and suits very well the Theory of Relativity, both Special and General.

### 3.6.1 Definition

Consider a point (we denote it as A) in space-time (in fact this theory applies to any other kind of space or manifold) and a curve (a one-dimensional continuous string of points) passing through this point. Let  $\lambda$  to be a parameter of this curve (coordinate of its points). This parameter defines a directional derivative  $d/d\lambda$  at point A. Indeed, if  $F$  is a function defined on this space (space-time, or manifold) then on this curve it is a function of  $\lambda$  and can be differentiated with respect to it:  $F = F(\lambda)$  and  $F' = dF/d\lambda$ .

Next introduce some coordinates in the space-time  $\{x^\nu\}$ . Now the curve can be described by the functions  $x^\mu = x^\mu(\lambda)$ . Moreover,  $F = F(x^\nu)$  and can be differentiated with respect to  $x^\nu$  as well. According to the chain rule

$$\frac{dF}{d\lambda} = \frac{dx^\nu}{d\lambda} \frac{\partial F}{\partial x^\nu},$$

where we applied the Einstein summation convention. Thus,

$$\frac{d}{d\lambda} = \frac{dx^\nu}{d\lambda} \frac{\partial}{\partial x^\nu}. \quad (3.20)$$

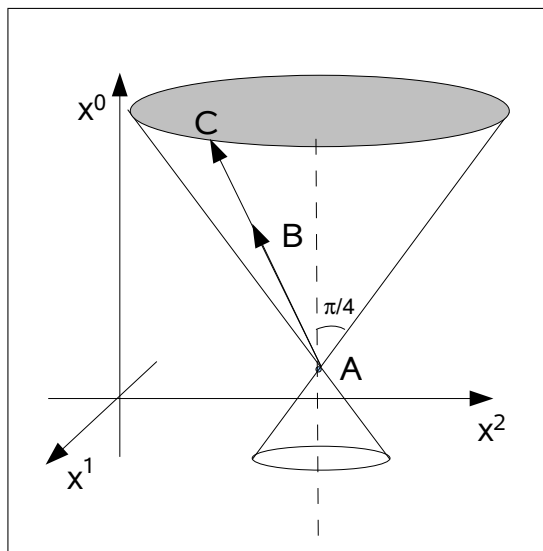


Figure 3.7:

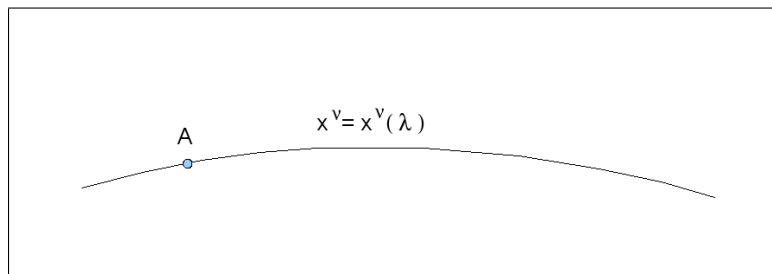


Figure 3.8:

It is easy to verify that all directional derivatives defined in this way at point A form a 4-dimensional *abstract vector space*, which we will denote as  $T_A$ . The set of operators

$$\left\{ \frac{\partial}{\partial x^\nu} \right\}, \quad \nu = 0, \dots, 3$$

is called the coordinate basis in  $T_A$ , induced by the coordinates  $\{x^\nu\}$ , and

$$\frac{dx^\nu}{d\lambda}, \quad \nu = 0, \dots, 3$$

are the components of the directional derivative  $d/d\lambda$  in this basis. These directional derivatives are identified with vectors in modern geometry. They are often called *Cartan's vectors*.  $d/d\lambda$  is also called the *tangent vector* to the curve with parameter  $\lambda$  at point A.

As it has been stressed this definition is very robust. It can be applied to the Euclidean space of Newtonian physics equally well. To see this consider the trajectory of a particle in Euclidean space and parametrise it using the Newtonian absolute time,  $t$ . The directional derivative  $d/dt$  is in fact the velocity vector of this particle,

$$\underline{v} = \frac{d}{dt}.$$

Indeed,

$$\frac{d}{dt} = \frac{dx^i}{dt} \frac{\partial}{\partial x^i}$$

and thus the components of  $\underline{v}$  in the coordinate basis are

$$v^i = \frac{dx^i}{dt},$$

which is the familiar result of Particle Kinematics.

### 3.6.2 Operations of addition and multiplication

The operations of addition and multiplication by real number for vectors in  $T_A$  are defined in terms of their components. Let

$$\vec{a} = a^\nu \frac{\partial}{\partial x^\nu}, \quad \vec{b} = b^\nu \frac{\partial}{\partial x^\nu}, \quad \vec{c} = c^\nu \frac{\partial}{\partial x^\nu}, \quad \alpha \in R^1.$$

Then we say that

$$\vec{c} = \vec{a} + \vec{b} \quad \text{iff} \quad c^\nu = a^\nu + b^\nu \quad (3.21)$$

and

$$\vec{c} = \alpha \vec{a} \quad \text{iff} \quad c^\nu = \alpha a^\nu. \quad (3.22)$$

These definitions ensure all the familiar properties of vector addition and multiplication (commutative law, associative law etc.) Notice that we use “arrow” to indicate 4-dimensional vectors of space-time, e.g.  $\vec{a}$ ,  $\vec{b}$ , whereas “underline” is reserved for 3-dimensional usual vectors of space, e.g.  $\underline{a}$ ,  $\underline{b}$ .

### 3.6.3 Coordinate transformation

Introduce new coordinates,  $\{x^{\tilde{\nu}}\}$ , in space-time. Obviously,

$$x^{\tilde{\nu}} = x^{\tilde{\nu}}(x^\mu) \quad \text{as well as} \quad x^\mu = x^\mu(x^{\tilde{\nu}})$$

The transformation from  $\{x^\nu\}$  to  $\{x^{\tilde{\nu}}\}$  delivers the transformation matrix

$$A_{\tilde{\mu}}^{\tilde{\nu}} = \frac{\partial x^{\tilde{\nu}}}{\partial x^{\tilde{\mu}}}. \quad (3.23)$$

We agree to consider the upper index in this expression as the row index and the lower one as the column index. The inverse transformation, from  $\{x^{\tilde{\nu}}\}$  to  $\{x^\nu\}$ , delivers another transformation matrix

$$A_\nu^{\tilde{\mu}} = \frac{\partial x^\mu}{\partial x^{\tilde{\nu}}}. \quad (3.24)$$

These two matrices are inverse to each other. Indeed their multiplication results in the unit matrix

$$A_{\tilde{\mu}}^{\tilde{\nu}} A_\beta^\mu = \frac{\partial x^{\tilde{\nu}}}{\partial x^{\tilde{\mu}}} \frac{\partial x^\mu}{\partial x^{\tilde{\beta}}} = \frac{\partial x^{\tilde{\nu}}}{\partial x^{\tilde{\beta}}} = \delta_{\tilde{\beta}}^{\tilde{\nu}}.$$

Here  $\delta_{\tilde{\beta}}^{\tilde{\nu}}$  is Kronecker’s delta symbol:

$$\delta_{\tilde{\beta}}^{\tilde{\nu}} = \begin{cases} 1 & \text{if } \tilde{\nu} = \tilde{\beta} \\ 0 & \text{if } \tilde{\nu} \neq \tilde{\beta} \end{cases}.$$

The new coordinates will introduce new coordinate basis at every point of space-time and in this new basis vectors will have different components. First we find the transformation of the coordinate basis of Cartan’s vectors. According to the chain rule

$$\frac{\partial f}{\partial x^\nu} = \frac{\partial f}{\partial x^{\tilde{\mu}}} \frac{\partial x^{\tilde{\mu}}}{\partial x^\nu}.$$

Thus,

$$\frac{\partial}{\partial x^\nu} = A_{\tilde{\nu}}^{\tilde{\mu}} \frac{\partial}{\partial x^{\tilde{\mu}}}. \quad (3.25)$$

Similarly one finds

$$\frac{\partial}{\partial x^{\tilde{\mu}}} = A_{\tilde{\mu}}^{\nu} \frac{\partial}{\partial x^{\nu}}. \quad (3.26)$$

Notice that in these expressions we automatically select the correct transformation matrix when we apply the Einstein summation rule. Now we can find the transformation law for the components of vectors. On one hand

$$\vec{u} = u^\nu \frac{\partial}{\partial x^\nu} = u^\nu A_{\tilde{\nu}}^{\tilde{\mu}} \frac{\partial}{\partial x^{\tilde{\mu}}}.$$

On the other hand

$$\vec{u} = u^{\tilde{\mu}} \frac{\partial}{\partial x^{\tilde{\mu}}}.$$

Comparing these two results we immediately obtain

$$u^{\tilde{\mu}} = A_{\tilde{\nu}}^{\tilde{\mu}} u^\nu. \quad (3.27)$$

Similarly we find

$$u^\nu = A_{\tilde{\mu}}^{\nu} u^{\tilde{\mu}}. \quad (3.28)$$

In these equations, we automatically select the correct transformation matrix when we apply the Einstein summation rule. In the old fashion textbooks on the Theory of Relativity, it is the transformation laws 3.27 and 3.28 that define vectors (or to be more precise the so-called contravariant vectors).

### 3.6.4 Infinitesimal displacement vectors

Consider a curve with parameter  $\lambda$  and its tangent vector  $d/d\lambda$  at the point with coordinates  $x^\nu$ . The infinitesimal increment of parameter due to infinitesimal displacement along the curve  $d\lambda$  is a scalar by which we can multiply the tangent vector and obtain the *infinitesimal displacement vector*

$$\vec{ds} = d\lambda \frac{d}{d\lambda}. \quad (3.29)$$

The components of this vector in the coordinate basis can be found via

$$ds^\nu = d\lambda \frac{dx^\nu}{d\lambda} = dx^\nu. \quad (3.30)$$

Thus, the components of  $\vec{ds}$  are the infinitesimal increments of coordinates  $dx^\nu$  corresponding to

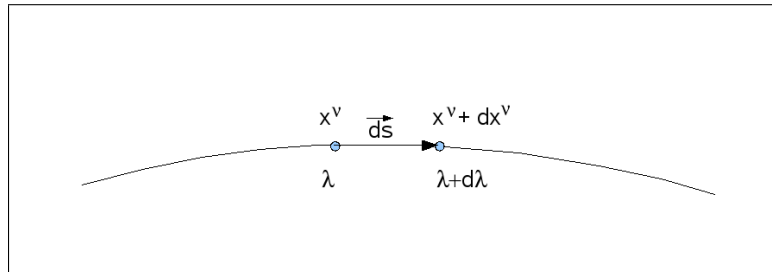


Figure 3.9:

the infinitesimal increment of the curve parameter  $\lambda$ ,

$$\vec{ds} = dx^\nu \frac{\partial}{\partial x^\nu}. \quad (3.31)$$

## 3.7 Tensors

### 3.7.1 Definition

In brief, tensors defined at point  $A$  are linear scalar operators acting on vectors from  $T_A$ . In other words, a tensor is a linear scalar function of, in general, several vector variables. The number of vectors the tensor needs to be “fed” with in order to produce a scalar is called its rank.

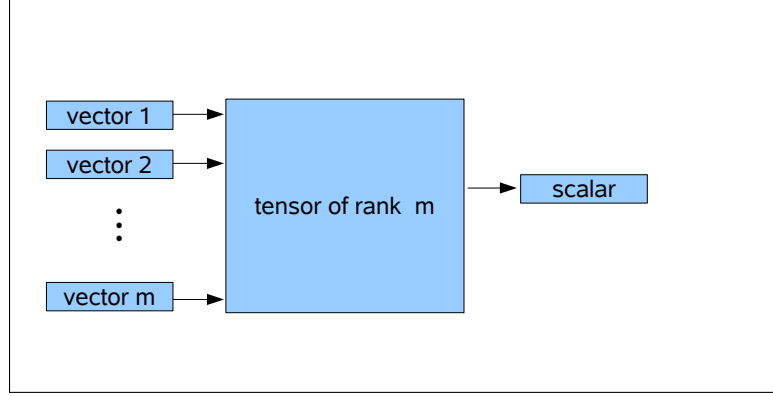


Figure 3.10:

Let, for example,  $T$  to be a second rank tensor defined at point  $A$ ,  $\vec{u}, \vec{v}, \vec{w} \in T_A$ , and  $\alpha, \beta \in R^1$ . Then according to the definition

$$T(\vec{u}, \vec{v}) \in R^1; \quad (3.32)$$

$$T(\alpha\vec{u} + \beta\vec{w}, \vec{v}) = \alpha T(\vec{u}, \vec{v}) + \beta T(\vec{w}, \vec{v}); \quad (3.33)$$

$$T(\vec{v}, \alpha\vec{u} + \beta\vec{w}) = \alpha T(\vec{v}, \vec{u}) + \beta T(\vec{v}, \vec{w}). \quad (3.34)$$

### 3.7.2 Components of tensors

Components of tensors are defined via their action on the basis vectors. For example

$$T_{\nu\mu} = T\left(\frac{\partial}{\partial x^\nu}, \frac{\partial}{\partial x^\mu}\right).$$

The action of tensors on vectors can be fully described in terms of their components. For example,

$$T(\vec{u}, \vec{v}) = T\left(u^\nu \frac{\partial}{\partial x^\nu}, v^\mu \frac{\partial}{\partial x^\mu}\right) = u^\nu v^\mu T\left(\frac{\partial}{\partial x^\nu}, \frac{\partial}{\partial x^\mu}\right) = u^\nu v^\mu T_{\nu\mu}.$$

Thus,

$$T(\vec{u}, \vec{v}) = u^\nu v^\mu T_{\nu\mu}. \quad (3.35)$$

This is often described as a *contraction* of  $T_{\nu\mu}$  with  $u^\nu$  (over index  $\nu$ ) followed by a *contraction* with  $v^\mu$  (over index  $\mu$ ).

### 3.7.3 Coordinate transformation

Introduce new coordinates  $\{x^{\bar{\mu}}\}$ . The new components  $T_{\bar{\nu}\bar{\mu}}$  of tensor  $T$  are

$$\begin{aligned} T_{\bar{\nu}\bar{\mu}} &= T\left(\frac{\partial}{\partial x^{\bar{\nu}}}, \frac{\partial}{\partial x^{\bar{\mu}}}\right) = T\left(A_{\bar{\nu}}^\alpha \frac{\partial}{\partial x^\alpha}, A_{\bar{\mu}}^\beta \frac{\partial}{\partial x^\beta}\right) = \\ &= A_{\bar{\nu}}^\alpha A_{\bar{\mu}}^\beta T\left(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\right) = A_{\bar{\nu}}^\alpha A_{\bar{\mu}}^\beta T_{\alpha\beta} \end{aligned}$$



Thus,

$$T_{\tilde{\nu}\tilde{\mu}} = A_{\tilde{\nu}}^{\alpha} A_{\tilde{\mu}}^{\beta} T_{\alpha\beta} \quad (3.36)$$

Note that the transformation matrix  $A$  appears once per each index of  $T$ . One of the indexes of  $A$  is always the summation index (also known as *dummy* or *contraction index*) – this is  $\alpha$  in the first transformation matrix in eq.3.36 and  $\beta$  in the second one. Whether this index is upper or lower one is determined by the Einstein summation convention – in general it must be the opposite case to that of the transformed tensor contracted with  $A$ . The other index (known as *free index* because it is not involved in the contraction) is the same both in case and letter as the corresponding index of  $T$  on the left-hand side – this is  $\tilde{\nu}$  in the first transformation matrix in eq.3.36 and  $\tilde{\mu}$  in the second one.

## 3.8 Metric tensor

### 3.8.1 Definition

Consider the infinitesimal displacement vector “connecting” points  $x^{\nu}$  and  $x^{\nu} + dx^{\nu}$ .

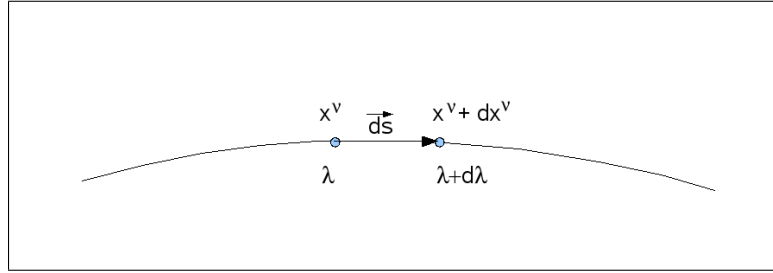


Figure 3.11:

A second rank tensor  $g$  which is symmetric,

$$g(\vec{u}, \vec{v}) = g(\vec{v}, \vec{u}), \quad (3.37)$$

and when acting on  $\vec{ds}$  produces the generalised distance between the points  $x^{\nu}$  and  $x^{\nu} + dx^{\nu}$

$$g(\vec{ds}, \vec{ds}) = ds^2 \quad (3.38)$$

is called the *metric tensor*. Since the components of  $\vec{ds}$  are  $dx^{\nu}$  we have

$$ds^2 = g_{\nu\mu} dx^{\nu} dx^{\mu}, \quad (3.39)$$

where  $g_{\nu\mu}$  are the components of the metric tensor. We immediately recognise this as the metric form.

The metric tensor is also used to define the *scalar product of vectors* and their magnitudes,

$$(\vec{u} \cdot \vec{v}) = g(\vec{u}, \vec{v}) \quad (3.40)$$

and

$$|\vec{u}|^2 = g(\vec{u}, \vec{u}). \quad (3.41)$$

In components these results read

$$(\vec{u} \cdot \vec{v}) = g_{\nu\mu} u^{\nu} v^{\mu} \quad (3.42)$$

and

$$|\vec{u}|^2 = g_{\nu\mu} u^{\nu} u^{\mu}. \quad (3.43)$$

### 3.8.2 Classification of space-time vectors and other results specific to space-time

The theory described above in Sections 3.2-3.5 is very general and does not use properties specific to space-time. Here we begin to apply this theory to space-time. Since the space-time of Special Relativity is four dimensional scalars, vectors and tensors of space-time are often called 4-scalars, 4-vectors and 4-tensors. In contrast the usual 3-dimensional scalars, vectors and tensors are called 3-scalars, 3-vectors and 3-tensors.

A space-time vector  $\vec{a}$  is called

- *time-like* if  $|\vec{a}|^2 < 0$ ,
- *space-like* if  $|\vec{a}|^2 > 0$
- *null* if  $|\vec{a}|^2 = 0$

To compute  $|\vec{a}|^2$  we have to use eq.3.43. The calculations are easiest in pseudo-Cartesian coordinates  $\{x^\nu\}$  as in these coordinates

$$g_{\nu\mu} = \begin{cases} -1 & \text{if } \nu = \mu = 0; \\ +1 & \text{if } \nu = \mu = 1, 2, 3; \\ 0 & \text{if } \nu \neq \mu. \end{cases}$$

and hence

$$|\vec{a}|^2 = g_{\nu\mu} a^\nu a^\mu = \sum_{\nu=0}^3 g_{\nu\nu} a^\nu a^\nu = -(a^0)^2 + (a^1)^2 + (a^2)^2 + (a^3)^2. \quad (3.44)$$

For example, consider space-time vector  $\vec{a} = (1, 1, 0, 0)$  whose components are those in the coordinate basis of some pseudo-Cartesian coordinates. Then

$$|\vec{a}|^2 = -1^2 + 1^2 + 0^2 + 0^2 = 0$$

and thus this vector is null.

In fact all sorts of calculations also easiest in pseudo-Cartesian coordinates. For example, the scalar product of two vectors is simply

$$(\vec{a} \cdot \vec{b}) = g_{\nu\mu} a^\nu b^\mu = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3. \quad (3.45)$$

What are the magnitudes of pseudo-Cartesian basis vectors and the angles between them? This is easy to answer as

$$\left| \frac{\partial}{\partial x^\nu} \right|^2 = g \left( \frac{\partial}{\partial x^\nu}, \frac{\partial}{\partial x^\nu} \right) = g_{\nu\nu}.$$

Thus  $\partial/\partial x^0$  is a unit time-like vector as

$$\left| \frac{\partial}{\partial x^0} \right|^2 = g_{00} = -1, \quad (3.46)$$

and  $\partial/\partial x^i$  ( $i = 1, 2, 3$ ) are unit space-like vector as

$$\left| \frac{\partial}{\partial x^i} \right|^2 = g_{ii} = 1. \quad (3.47)$$

Similarly we find that

$$\left( \frac{\partial}{\partial x^\nu} \cdot \frac{\partial}{\partial x^\mu} \right) = g_{\nu\mu} = 0 \quad \text{if } \nu \neq \mu,$$

and thus these unit vectors are orthogonal to each other.

Finally, we construct the transformation matrix of Lorentz transformations. In the case of pseudo-Cartesian coordinates the Lorentz transformation reads

$$\begin{cases} x^{\bar{0}} = \gamma(x^0 - \beta x^1) \\ x^{\bar{1}} = \gamma(x^1 - \beta x^0) \\ x^{\bar{2}} = x^2 \\ x^{\bar{3}} = x^3 \end{cases} . \quad (3.48)$$

From this we find that

$$\begin{cases} dx^{\bar{0}} = \gamma(dx^0 - \beta dx^1) \\ dx^{\bar{1}} = \gamma(dx^1 - \beta dx^0) \\ dx^{\bar{2}} = dx^2 \\ dx^{\bar{3}} = dx^3 \end{cases} \quad (3.49)$$

and thus

$$A_{\nu}^{\bar{\mu}} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \quad (3.50)$$

Thus, pseudo-Cartesian components of any 4-vector transform according to the same law as eq.3.48:

$$\begin{cases} a^{\bar{0}} = \gamma(a^0 - \beta a^1) \\ a^{\bar{1}} = \gamma(a^1 - \beta a^0) \\ a^{\bar{2}} = a^2 \\ a^{\bar{3}} = a^3 \end{cases} \quad (3.51)$$

In fact eq.3.49 is a particular example of this rule as  $dx^{\nu}$  are the components of infinitesimal displacement vector  $\vec{ds}$ .



## Chapter 4

# Relativistic particle mechanics

### 4.1 Tensor equations and the Principle of Relativity

The Einstein's Principle of Relativity requires all laws of Physics to be the same in all inertial frames and hence to be invariant under the Lorentz transformation. Since the laws of Newtonian physics are Galilean invariant they cannot be Lorentz invariant and have to be modified. How can this be done? Before, we do this we should try to answer the related question: "How to write Lorentz invariant equations?"

The great value of the concept of space-time and of the tensor calculus in the Theory of Relativity is that they suggest a straightforward way of constructing Lorentz invariant equations. Indeed, the Lorentz transformation, which in physical terms describes the transformation between two inertial frames in standard configuration, corresponds to a particular type of coordinate transformations in spacetime, namely a sub-class of transformations between two systems of pseudo-Cartesian coordinates. However, vector equations involving space-time vectors like

$$\vec{a} = m \vec{b}, \quad (4.1)$$

where  $m$  is a space-time scalar, are fully meaningful without any use of coordinates, and hence fully independent on the choice of coordinates in space-time. Once we introduce some coordinates we can write the representation of this equations in terms of components of vectors  $\vec{a}$  and  $\vec{b}$

$$a^\nu = m b^\nu \quad (\nu = 0, 1, 2, 3). \quad (4.2)$$

Clearly, the form of equation (4.2) is the same for any choice of coordinates in space-time and hence this equation is Lorentz invariant too (We also note that the left and the right hand sides of eq.4.2 transform in exactly the same way; see eq.3.27.). The same applies to all proper tensor equations in general, like

$$G^{\nu\mu} = a T^{\nu\mu}.$$

However, the detailed study of tensor operations is out of the scope of this course and we will deal only with vector equations.

Thus, we now know how to construct Lorentz invariant equations. But we also need to figure out how to relate such equations with the corresponding equations of Newtonian physics. The Newtonian equations do not comply with the Einstein's Relativity Principle but they are very accurate when the involved speeds are much lower compared to the speed of light. Thus, the relativistic equations must reduce to the Newtonian ones in the limit of low speeds. In order, to compare the relativistic and Newtonian equation we need to find out how the 4-vector equations can be reduced to the 3-vector and 3-scalar equations characteristic for Newtonian Physics.

Equations like eq.4.2 differ from the vector equations of classical physics in that they have not 3 but 4 components. The following simple analysis shows that any 4-vector equation corresponds to a pair of traditional Newtonian-like equations, one them being a 3-scalar equation and the other one a 3-vector equation. Indeed, let us select an arbitrary inertial frame. This frame introduces

hyper-surfaces (“hyper” because they are 3-dimensional) of simultaneity,  $t = t_0$ , in space-time. Each such hypersurface is the space of this inertial frame at a particular time. (Another inertial frame introduces different hypersurfaces of this sort and has different space and time.)

Now consider a transformation of spacial coordinates ( $x^i$ ,  $i = 1, 2, 3$ ) in the space of this frame. The transformation matrix of this transformation is

$$A_{\tilde{j}}^{\tilde{i}} = \frac{\partial x^{\tilde{i}}}{\partial x_j}, \quad (i, j = 1, 2, 3).$$

Quantities that remain invariants under such transformations are the usual spacial scalars (or 3-scalars). The components of usual spacial vectors (or 3-vectors) vary according to the vector transformation law

$$a^{\tilde{i}} = A_{\tilde{j}}^{\tilde{i}} a^j \quad (i = 1, 2, 3).$$

This transformation can also be viewed as a transformation of space-time coordinates but of a rather specific sort, namely

$$x^{\tilde{0}} = x^0, \quad x^{\tilde{i}} = x^{\tilde{i}}(x^j), \quad (i, j = 1, 2, 3).$$

Thus,

$$A_0^{\tilde{0}} = 1, \quad A_i^{\tilde{0}} = 0, \quad A_0^{\tilde{i}} = 0, \quad (4.3)$$

and so we have

$$A_{\tilde{\nu}}^{\tilde{\mu}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & A_1^{\tilde{1}} & A_2^{\tilde{1}} & A_3^{\tilde{1}} \\ 0 & A_1^{\tilde{2}} & A_2^{\tilde{2}} & A_3^{\tilde{2}} \\ 0 & A_1^{\tilde{3}} & A_2^{\tilde{3}} & A_3^{\tilde{3}} \end{pmatrix}. \quad (4.4)$$

Next we apply this transformation law to a space-time vector. Whereas in general

$$a^{\tilde{\mu}} = A_{\tilde{\nu}}^{\tilde{\mu}} a^{\nu}$$

in our case we have

$$a^{\tilde{0}} = a^0 \quad (4.5)$$

and

$$a^{\tilde{i}} = A_{\tilde{j}}^{\tilde{i}} a^j \quad (i = 1, 2, 3). \quad (4.6)$$

Thus, the  $a^0$  is invariant and hence behaves as a 3-scalar, whereas  $a^i$  behave as components of a 3-vector. This result is summarised in the following notation

$$\vec{a} = (a^0, \underline{a}). \quad (4.7)$$

Thus, any 4-vector equation splits into a 3-scalar equation and a 3-vector equation. For example, the 4-vector equation

$$\vec{a} = m \vec{b}$$

splits into the 3-scalar equation

$$a^0 = mb^0$$

and the 3-vector equation

$$\underline{a} = m\underline{b}.$$

This result shows that Newtonian physics missed important connections between various physical quantities and between various equations. These are uncovered in Special Relativity.

## 4.2 4-velocity and 4-momentum

Before we attempt to write relativistic versions of the most important laws of particle dynamics, we need to introduce relevant 4-scalars and 4-vectors. The 3-velocity vector of Newtonian physics is the tangent vector along the particle trajectory with the absolute time as a parameter

$$\underline{v} = \frac{d}{dt} \quad (v^i = dx^i/dt)$$

(see the end of Sec.3.2.1).

In the same manner we can introduce the 4-velocity vector of a particle. Let us see how this is done. First instead of a curve in space, which is the particle trajectory, we need a curve in space-time. The only obvious choice is the world-line of the particle. Second, we need a time-like parameter along this curve. However, we no longer have the absolute time and each inertial frame has its own time. The only natural choice which gives us a unique parameter is the proper time of the particle,  $\tau$ , which is the time as measured by a fiducial standard clock moving with this particle (the comoving clock). From this definition, it follows that  $\tau$  is a 4-scalar. Indeed, it is defined without any reference to a coordinate system in space-time. Thus, the 4-velocity of a particle can be defined as the directional derivative

$$\vec{u} = \frac{d}{d\tau} \quad (4.8)$$

along the world line of this particle.

Once a coordinate system is introduced for the space-time, we can find the components of  $\vec{u}$  in the corresponding coordinate basis,

$$u^\nu = \frac{dx^\nu}{d\tau}. \quad (4.9)$$

Since  $dx^\nu$  are also the components of the infinitesimal displacement 4-vector  $\vec{ds}$  along the world line, in the same basis, we can also write

$$\vec{u} = \frac{\vec{ds}}{d\tau}. \quad (4.10)$$

This can be understood as the 4-vector  $\vec{ds}$  divided by the 4-scalar  $d\tau$ , the corresponding increment of the proper time.

Consider an arbitrary inertial frame with arbitrary spatial coordinates  $x^i$  and the **normalised** time coordinate  $x^0 = ct$ , where  $t$  is the time of this inertial frame. Then all the basis 4-vectors  $\partial/\partial x^i$  are orthogonal to  $\partial/\partial x^0$  but in general they are not unit 4-vectors. Whereas the basis 4-vector

$$\frac{\partial}{\partial x^0} = \frac{1}{c} \frac{\partial}{\partial t}$$

is unit. In this basis,

$$u^0 = \frac{dx^0}{d\tau} = c \frac{dt}{d\tau}.$$

Since  $dt$ ,  $d\tau$ , and the particle Lorentz factor  $\Gamma$  are related via the time dilation formula

$$dt = \Gamma d\tau,$$

we can write

$$u^0 = \Gamma c.$$

As to the other components,

$$u^i = \frac{dx^i}{d\tau} = \frac{dt}{d\tau} \frac{dx^i}{dt} = \Gamma v^i,$$

where  $v^i = dx^i/dt$  are the components of the 3-velocity vector of this particle in this inertial frame in the coordinate basis of its space. These results are summarised using the following notation

$$\vec{u} = (\Gamma c, \Gamma \underline{v}). \quad (4.11)$$

Notice that  $\Gamma$  is a 3-scalar. (The most frequently used notation for the Lorentz factor is  $\gamma$ . However, in these notes we reserve  $\gamma$  specifically for the Lorentz factor of the relative motion of the two inertial frames involved in the Lorentz transformation and use  $\Gamma$  for the Lorentz factor associated with any other kind of motion.)

Given the above results, we can find the magnitude of the 4-velocity:

$$|\vec{u}|^2 = -\Gamma^2 c^2 + \Gamma^2 v^2 = -c^2 \Gamma^2 (1 - v^2/c^2) = -c^2.$$

Thus, 4-velocity is a time-like vector. Recall, that  $|\vec{u}|^2$  is a 4-scalar and in any inertial frame and for any selection of space-time coordinates, the calculations of this quantity will yield exactly the same result,  $-c^2$ . For example, in the particle rest frame (the inertial frame where this particle is instantaneously at rest)  $\Gamma = 1$ ,  $\underline{v} = \underline{0}$  and

$$\vec{u} = (c, 0, 0, 0) \quad (4.12)$$

From this we immediately find that,

$$|\vec{u}|^2 = -c^2. \quad (4.13)$$

This example demonstrates how easy the calculations can become when we make a good choice of coordinate system/inertial frame.

Another important particle parameter in Newtonian mechanics is its inertial mass, which is assumed to be the same in all inertial frames. However, we cannot be sure that this assumption, which makes the inertial mass a 4-scalar, will be consistent with the relativistic mechanics. (In fact, as we will see very soon, it is not.) Yet, we need a 4-scalar to be used in 4-tensor equations. To make sure, we can use the same trick as with the proper time. We introduce the particle *rest mass* (or *proper mass*),  $m_0$ , which is defined as the inertial mass measured in the particle rest frame. This defines  $m_0$  uniquely, yielding the same quantity no matter which inertial frame is actually used to describe the particle motion, or which coordinates we introduce in space-time for this purpose.

In Newtonian mechanics, the product of particle's inertial mass and its 3-velocity is another important 3-vector – the 3-vector of momentum,

$$\underline{p} = m\underline{v}.$$

This suggests to introduce the 4-vector of momentum (or the 4-momentum vector) via

$$\vec{P} = m_0 \vec{u}. \quad (4.14)$$

Using eq.4.13 we immediately obtain

$$|\vec{P}|^2 = -m_0^2 c^2 \quad (4.15)$$

and thus the 4-momentum vector is also time-like. Using eq.4.11 we find that

$$\vec{P} = (m_0 \Gamma c, m_0 \Gamma \underline{v}). \quad (4.16)$$

The relativistic generalisation of the second law of Newton is also quite straightforward

$$\frac{d\vec{P}}{d\tau} = \vec{F}, \quad (4.17)$$

where  $\vec{F}$  is a 4-vector of force or just 4-force. We will come back to this law later after we consider the relativistic laws of conservation of energy and momentum.



### 4.3 Energy-momentum conservation

Consider an isolated system of  $N$  particles. The term isolated means that these particles may interact with each other but not with their surrounding. Moreover, our analysis here is restricted to only short-range interaction between particles (collisions) whereas the long-range interactions and the corresponding potential energy are ignored. The Newtonian mechanics states that the total energy and the total momentum of such a system are conserved, that is they do not change in time:

$$\sum_{k=1}^N E_{(k)} = \text{const} \quad (4.18)$$

and

$$\sum_{k=1}^N \underline{p}_{(k)} = \text{const}, \quad (4.19)$$

where  $E_{(k)} = m_{(k)}v_{(k)}^2/2$  and  $\underline{p}_{(k)} = m_{(k)}\underline{v}_{(k)}$  are the energy and momentum of  $k$ -th particle respectively,  $m_{(k)}$  is its inertial mass and  $\underline{v}_{(k)}$  is its 3-velocity. Note that we use round brackets here to indicate that  $k$  is the particle number (or name). This helps to avoid potential confusion with tensor indexes.

We should start writing the relativistic laws of mechanics as 4-tensor equations and the equation

$$\sum_{k=1}^N \vec{P}_{(k)} = \text{const} \quad (4.20)$$

seems to be a good candidate for the 4-vector counterpart of momentum conservation. If this is a good choice then we should recover the Newtonian law of momentum conservation in the limit of slow speeds ( $v \ll c$ ). Let us split this equation into a 3-scalar and a 3-vector equations following the discussion in Sec.4.1. Using eq.4.16 we obtain

$$\sum_{k=1}^N P_{(k)}^0 = \text{const} \quad \text{or} \quad \sum_{k=1}^N m_{0,(k)}\Gamma_{(k)}c = \text{const} \quad (4.21)$$

and

$$\sum_{k=1}^N P_{(k)}^i = \text{const} \quad \text{or} \quad \sum_{k=1}^N m_{0,(k)}\Gamma_{(k)}\underline{v}_{(k)} = \text{const}. \quad (4.22)$$

Quick inspection shows that the last equation appears to be the relativistic version of 3-momentum conservation and suggests to define the relativistic 3-momentum of a particle as

$$\underline{p} = m_0\Gamma\underline{v}. \quad (4.23)$$

In fact, we can make the expression for  $\underline{p}$  to be identical to the Newtonian one

$$\underline{p} = m\underline{v} \quad (4.24)$$

if we define the inertial mass of moving particle as

$$m = m_0\Gamma. \quad (4.25)$$

This definition make the inertial mass to be different from the rest mass and depending of the particle speed. Only at small speeds ( $v \ll c$  and hence  $\Gamma \simeq 1$ ) we have  $m_{(k)} = m_{0,(k)}$ .

In order to understand the meaning of eq.4.21 we consider its Newtonian limit ( $v \ll c$ ). If we simply replace each  $\Gamma$  in eq.4.21 with unity then we obtain

$$\sum_{k=1}^N m_{0,(k)} = \text{const}.$$

Clearly this is the mass conservation law – the total mass of the system is conserved. However, we get deeper insight into the nature of eq.4.21 if we consider the Taylor expansion of  $\Gamma$

$$\Gamma = \left(1 - \left(\frac{v}{c}\right)^2\right)^{-1/2} = 1 + \frac{1}{2} \left(\frac{v}{c}\right)^2 + \frac{3}{4} \left(\frac{v}{c}\right)^4 + \dots$$

and keep not only the first but also the second term in this expansion. Then

$$m_0 \Gamma c \simeq m_0 c \left(1 + \frac{1}{2} \frac{v^2}{c^2}\right) = \frac{1}{c} \left(m_0 c^2 + \frac{1}{2} m_0 v^2\right) \quad (4.26)$$

and eq.4.21 reads

$$\sum_{k=1}^N \left(m_{0,(k)} c^2 + \frac{1}{2} m_{0,(k)} v^2\right) = \text{const} . \quad (4.27)$$

Since in Newtonian mechanics the mass conservation is a separate law then eq.4.27 reduces to the Newtonian energy conservation

$$\sum_{k=1}^N \frac{1}{2} m_{0,(k)} v^2 = \text{const} . \quad (4.28)$$

This implies that

- the 3-scalar equation (4.21) is the relativistic generalisation of the energy conservation law;
- the 4-vector equation (4.20) unites the energy and momentum conservation laws into a single *energy-momentum conservation law*;
- the 4-momentum vector describes not only the particle 3-momentum but also its energy (a 3-scalar). For this reason it is often called *the energy-momentum vector*.

What should we consider as the particle energy in the relativistic case? Whatever is the definition it should deliver a conserved quantity! Given the results (4.26,4.28) we seem to have only two options. The first one is

$$K = E - m_0 c^2. \quad (4.29)$$

As we have seen, in the Newtonian limit  $K$  reduces to the usual expression for the kinetic energy, which seems to make this a sensible choice. If, however, we use this option then the total energy of the system will be conserved only provided that the total rest mass of the system is conserved. In fact, the laboratory experiments with atoms and sub-atomic particles show that the total rest mass of particles is not always conserved in particle collisions

$$\sum_{k=1}^N m_{0,(k)} c^2 \neq \text{const}.$$

For this reason this option is unsatisfactory. The second option is

$$E = P^0 c = m_0 \Gamma c^2 = m c^2 \quad (4.30)$$

and eq.4.21 shows that in this case the total energy of an isolated system is always conserved

$$\sum_{k=1}^N E_{(k)} = \text{const}.$$

Thus,  $E = m c^2$ , is the only satisfactory definition of particle energy. In the case,  $v = 0$ , it reduces to

$$E_0 = m_0 c^2, \quad (4.31)$$

which is called the *rest mass-energy* of the particle. In general,  $E \neq E_0$  and the difference between them

$$K = E - E_0 \quad (4.32)$$

is called the particle *kinetic energy*.

The lack of rest mass-energy conservation in particle collisions means that the total kinetic energy of an isolated system may increase or decrease provided its total rest mass-energy decreases or increases respectively. In fact it is the reduction of the total rest mass-energy of nuclear fuel in nuclear reactors during the process of radioactive decay that is the source of heat in nuclear power stations.

The definitions (4.24,4.30) allow us to write

$$\vec{P} = (E/c, \underline{p}). \quad (4.33)$$

That is given some inertial frame the energy-momentum vector splits into the particle energy  $E$  (3-scalar) and momentum  $\underline{p}$  as measured in this frame. Thus, Special Relativity uncovers the deep connection between energy and momentum.

In Sec.4.2, eq.4.15, we derived an expression for the magnitude of 4-momentum vector of a particle with rest mass  $m_0$ . Using eq.4.33 we can find another expression for it. Utilising pseudo-Cartesian coordinates we find that

$$|\vec{P}|^2 = -\frac{E^2}{c^2} + (p^1)^2 + (p^2)^2 + (p^3)^2 = -\frac{E^2}{c^2} + p^2.$$

Combining this result with the one of eq.4.15 we find the following identity

$$E^2 = p^2 c^2 + m_0^2 c^4 \quad \text{or} \quad E^2 = E_0^2 + p^2 c^2, \quad (4.34)$$

where  $E_0 = m_0 c^2$  is the rest mass-energy.

## 4.4 Photons

In 1900 German physicist Max Planck concluded that electromagnetic energy is emitted not in arbitrary amounts but in *quanta of energy*

$$E = h\nu, \quad (4.35)$$

where  $\nu$  is the frequency of radiation and  $h$  is a constant, known as the Planck constant. Later, in 1905, Einstein concluded that the electromagnetic energy is not only emitted but also absorbed in quanta and proposed that it must actually propagate in quanta. Thus, in some respects light can be considered as a collection of particles, called *photons*.

It is natural to assume that the speed of a photon equals to the speed of light. This makes world lines of photons null and we cannot introduce 4-velocity of photons in the same fashion as we have done for massive particles. Indeed, the 4-velocity is defined as  $\vec{ds}/d\tau$  where  $\tau$  is the proper time of the particle. This time is measured by the standard clock moving with the same speed as the particle. However, no such clock can be forced to move with the speed of light. Indeed, if the rest mass of the clock is  $m_0$  and it is moving with the Lorentz factor  $\Gamma$ , then according to eq.4.30 the total energy of this clock is  $E = m_0 \Gamma c^2$ . Thus, as  $v \rightarrow c$  we have  $\Gamma \rightarrow \infty$  and hence  $E \rightarrow \infty$  and given the finite energy resources available to man the speed of light cannot be reached by the clock. This reasoning invites the following question: How can photons move with the speed of light and still have the finite energy given by eq.4.35? The only possibility is to assume that for a photon

$$m_0 = 0. \quad (4.36)$$

This makes sense as photons cannot be found at rest in any inertial frame – they always move with the speed of light.

Particles with zero rest mass are usually called massless (in contrast to “massive” particles which have  $m_0 \neq 0$ ). However, this does not mean that the inertial mass of such “massless” particles is also zero. In fact, combining Einstein’s  $E = mc^2$  with Planck’s  $E = h\nu$  we find

$$m = h\nu/c^2. \quad (4.37)$$

Using this result we can find the 3-momentum vector of a photon

$$\underline{p} = \frac{E}{c}\underline{n}, \quad (4.38)$$

where  $\underline{n}$  is the unit vector in the direction of motion. Indeed,

$$\underline{p} = m\underline{v} = \frac{E}{c^2}c\underline{n} = \frac{E}{c}\underline{n}.$$

This allows us to compile the 4-momentum vector of photon as

$$\vec{P} = \left(\frac{E}{c}, \frac{E}{c}\underline{n}\right) = \frac{E}{c}(1, \underline{n}) = \frac{h\nu}{c}(1, \underline{n}) \quad (4.39)$$

even if we cannot introduce 4-velocity for particles of this sort. From this equation we find that

$$|\vec{P}|^2 = 0 \quad (4.40)$$

## 4.5 Particle collisions

The simplest and yet very important class of problems in mechanics is the particle collisions. In such problems a number of particles interact with each other by means of some short-range forces which can be rather complex and the actual interaction can be difficult to describe. However, the conservation of energy and momentum allows to answer a number of important questions on the final outcome of collisions. Here we consider few simple examples.

### 4.5.1 Nuclear recoil

**Problem:** If not excited via collisions with other particles atomic nuclei reside in the ground states.

However, via collisions they can be moved to states with higher rest-mass energy, the excited states, and then spontaneously return to the ground state. The energy excess is then passed to another particle emitted by the nucleus during the transition. This particle can be a photon.

Denote the rest mass-energy of the ground state as  $E_g$  and the rest mass-energy of excited state  $E_g + \delta E$ . Find the energy of emitted photon in the frame where the excited nucleus is initially at rest.

**Solution:**

In the rest frame of the excited nucleus its 4-momentum is

$$\vec{Q}_0 = \left(\frac{E_g + \delta E}{c}, \underline{0}\right). \quad (4.41)$$

In the same frame its 4-momentum after the transition to the ground state is

$$\vec{Q}_1 = \left(\frac{E}{c}, \underline{p}\right), \quad (4.42)$$

where  $\underline{p}$  is its 3-momentum and  $E > E_g$  is its energy which includes the kinetic energy as well. In the same frame the 4-momentum of the emitted photon is

$$\vec{P} = \left(\frac{E_p}{c}, \frac{E_p}{c}\underline{n}\right), \quad (4.43)$$

where  $\underline{n}$  is the direction of motion of the photon. The energy-momentum conservation law requires

$$\vec{Q}_0 = \vec{Q}_1 + \vec{P}. \quad (4.44)$$

From this point there are two ways to proceed, one is more elegant than the other.

*The less elegant way:* Here we simply use the fact that the corresponding components of 4-vectors on the left and right hand sides of Eq.4.44 must be the same:

$$\begin{cases} E_g + \delta E = E + E_p, \\ \underline{0} = \underline{p} + \frac{E_p}{c} \underline{n} \end{cases} \quad (4.45)$$

From this we find

$$E = E_g + \delta E - E_p,$$

and

$$p^2 = \frac{E_p^2}{c^2}.$$

The next step is to utilize Eq.4.34, which in our case reads

$$E^2 = E_g^2 + p^2 c^2.$$

Substituting the above expressions for  $E$  and  $p$  into this equation we obtain

$$(E_g + \delta E - E_p)^2 = E_g^2 + \frac{E_p^2}{c^2} c^2.$$

From this we can find  $E_p$  as a function of  $E_g$  and  $\delta E$  only. Simple calculations yield

$$E_p = \delta E \frac{\delta E + 2E_g}{2\delta E + 2E_g}.$$

Notice that  $E_p < \delta E$ . This is because a fraction of  $\delta E$  has been transformed into the kinetic energy of the nucleus. When the photon is emitted the nucleus receives a kick in the opposite direction (nuclear recoil) so that the total momentum of the system “nucleus+photon” remains zero.

*The more elegant way:* One can rewrite Eq.4.44 as

$$\vec{Q}_0 - \vec{P} = \vec{Q}_1.$$

This may seem not much different but in fact this significantly simplifies calculations because when we square both sides of this equation the result will include only the quantities which are given and the quantity we need to find.

Let us show this:

$$|\vec{Q}_0 - \vec{P}|^2 = |\vec{Q}_1|^2 \quad (4.46)$$

means that

$$|\vec{Q}_0|^2 + |\vec{P}|^2 - 2\vec{Q}_0 \cdot \vec{P} = |\vec{Q}_1|^2. \quad (4.47)$$

According to the general results (4.15) and (4.40), for any particle the magnitude squared of its 4-momentum equals to

$$|\vec{P}|^2 = -E_0^2/c^2, \quad (4.48)$$

where  $E_0 = m_0 c^2$  is the rest mass-energy of the particle. Thus, we have

$$|\vec{Q}_0|^2 = -E_g^2/c^2, \quad |\vec{Q}_1|^2 = -(E_g + \delta E)^2/c^2, \quad \text{and} \quad |\vec{P}|^2 = 0.$$

Moreover,

$$\vec{Q}_0 \cdot \vec{P} = -\frac{(E_g + \delta E) E_p}{c}.$$

Substituting these into Eq.4.47 we find

$$-(E_g + \delta E)^2 + 2(E_g + \delta E)E_p = -E_g^2,$$

which includes only one unknown quantity,  $E_p$ . Solving this for  $E_p$  we find the same answer as before

$$E_p = \delta E \frac{\delta E + 2E_g}{2\delta E + 2E_g}. \quad (4.49)$$

### 4.5.2 Absorption of neutrons

**Problem:** A neutron of rest mass-energy  $E_n$  and kinetic energy  $2E_n$  is absorbed by a stationary  $B^{10}$  (Boron-10) nucleus of rest mass-energy  $E_{10} = 10E_n$  and becomes  $B^{11}$  (Boron-11) nucleus. Find the rest mass-energy  $E_{11}$  of  $B^{11}$  immediately after the absorption ( and hence before it moves from the excited state to the ground state).

**Solution:** Use the rest frame of  $B^{10}$  for the calculations. The 4-momentum of  $B^{10}$  and the neutron before the collision are

$$\vec{Q}_{10} = \left( \frac{10E_n}{c}, \underline{0} \right),$$

$$\vec{Q}_n = \left( \frac{3E_n}{c}, \underline{p} \right),$$

The 4-momentum of  $B^{11}$  after the collision denote as  $\vec{Q}_{11}$ . The energy-momentum conservation reads

$$\vec{Q}_{10} + \vec{Q}_n = \vec{Q}_{11}.$$

From this equation we have

$$|\vec{Q}_{10}|^2 + |\vec{Q}_n|^2 + 2\vec{Q}_{10} \cdot \vec{Q}_n = |\vec{Q}_{11}|^2. \quad (4.50)$$

Next we use the relation between the rest mass-energy of a particle and the magnitude of its 4-momentum vector:

$$|\vec{Q}_{10}|^2 = -\frac{(10E_n)^2}{c^2}, \quad |\vec{Q}_n|^2 = -\frac{(E_n)^2}{c^2}, \quad |\vec{Q}_{11}|^2 = -\frac{(E_{11})^2}{c^2},$$

and calculate

$$\vec{Q}_{10} \cdot \vec{Q}_n = -30 \frac{E_n^2}{c^2}.$$

Finally we substitute these results into Eq.4.50

$$-E_n^2 - 100E_n^2 - 60E_n^2 = -E_{11}^2$$

and obtain

$$E_{11} = \sqrt{161}E_n \approx 12.7E_n.$$

## 4.6 4-acceleration and 4-force

The 4-acceleration of a massive particle is defined as

$$\vec{a} = \frac{d\vec{u}}{d\tau}. \quad (4.51)$$

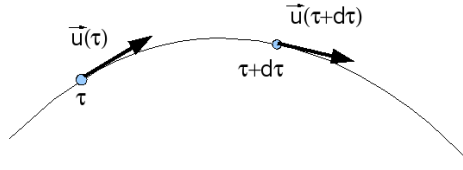


Figure 4.1:

(In this definition  $d\vec{u} = \vec{u}(\tau + d\tau) - \vec{u}(\tau)$  and thus we subtract vectors defined at two different points. We can do this only if we can identify two vectors defined at two different places. In Special Relativity we say that two such vectors are identical if they have the same components in the corresponding coordinate bases of the same system of pseudo-Cartesian coordinates. ) It is easy to see that

$$(\vec{a} \cdot \vec{u}) = 0. \quad (4.52)$$

Indeed,

$$(\vec{a} \cdot \vec{u}) = \vec{u} \cdot \frac{d}{d\tau} \vec{u} = \frac{1}{2} \frac{d}{d\tau} (\vec{u} \cdot \vec{u}) = -\frac{1}{2} \frac{dc^2}{d\tau}.$$

In terms of pseudo-Cartesian components eq.4.51 reads

$$\vec{a} = \frac{d}{d\tau} (\Gamma c, \Gamma \underline{v}) = \Gamma \frac{d}{dt} (\Gamma c, \Gamma \underline{v}) = \left( c\Gamma \frac{d\Gamma}{dt}, \Gamma \frac{d\Gamma}{dt} \underline{v} + \Gamma^2 \frac{d\underline{v}}{dt} \right).$$

Thus,

$$\vec{a} = \left( c\Gamma \frac{d\Gamma}{dt}, \Gamma \frac{d\Gamma}{dt} \underline{v} + \Gamma^2 \underline{a} \right), \quad (4.53)$$

where  $\underline{a} = d\underline{v}/dt$  is the usual 3-acceleration vector.

The 4-force vector is defined as

$$\vec{F} = \frac{d\vec{P}}{d\tau}. \quad (4.54)$$

Since for massive particle  $\vec{P} = m_0 \vec{u}$  we obtain that

$$\vec{F} = \frac{dm_0}{d\tau} \vec{u} + m_0 \vec{a} \quad (4.55)$$

and thus in general the 4-force is not parallel to the 4-acceleration. This fact gives rise to the following classification of forces in Special Relativity.

- A 4-force is called *pure* if

$$dm_0/d\tau = 0.$$

Eq.4.55 tells us that a pure 4-force is parallel to the 4-acceleration of particle

$$\vec{F} = m_0 \vec{a}. \quad (4.56)$$

Combining this result with eq.4.52 we also find that

$$\vec{F} \cdot \vec{a} = 0. \quad (4.57)$$

- A 4-force is called *heat-like* if

$$\vec{a} = 0.$$

Eq.4.55 tells us that a heat-like 4-force is parallel to the 4-velocity of particle

$$\vec{F} = \frac{dm_0}{d\tau} \vec{u}. \quad (4.58)$$

Given the splitting (4.33) of 4-momentum we can split the 4-force vector of space-time into a 3-scalar and a 3-vector as well. We obtain

$$\vec{F} = \frac{d}{d\tau}(E/c, \underline{p}) = \Gamma \frac{d}{dt}(E/c, \underline{p}) = \left( \frac{\Gamma}{c} \frac{dE}{dt}, \Gamma \frac{d\underline{p}}{dt} \right).$$

Using the same definition of 3-force as in Newtonian mechanics,

$$\underline{f} = \frac{d\underline{p}}{dt}, \quad (4.59)$$

we can write the above result as

$$\vec{F} = \left( \frac{\Gamma}{c} \frac{dE}{dt}, \Gamma \underline{f} \right). \quad (4.60)$$

This equation tells us that the 4-force determines not only the rate of change of particle momentum but also the change its energy. Since  $\vec{u} = (\Gamma c m \Gamma \underline{v})$  we have

$$(\vec{F} \cdot \vec{u}) = \gamma^2 \left( -\frac{dE}{dt} + \underline{f} \cdot \underline{v} \right).$$

On the other hand, using eqs.(4.52,4.55) we also obtain

$$(\vec{F} \cdot \vec{u}) = -c^2 \frac{dm_0}{d\tau} = -c^2 \gamma \frac{dm_0}{dt}.$$

Combining these two result we find

$$\frac{dE}{dt} = \frac{c^2}{\gamma} \frac{dm_0}{dt} + (\underline{f} \cdot \underline{v}). \quad (4.61)$$

Similar to Newtonian mechanics each kind of interaction should be described by additional laws specifying the 4-force. This is the subject for more advanced course on Special Relativity.