Spacetime models

Tamás Matolcsi

2013

Contents

I Preface

The theory of special relativity was developed because experimental results concerning light phenomena (electromagnetic waves) contradicted the naive concepts of time and space. Though special relativity has become a well-working tool in modern physics, its usual formulation can give rise to misunderstandings: over and over again research papers appear describing 'paradoxes' in relativity theory (see Subsections 16.1 and 16.2).

The basic concept of special relativity – homogeneous and isotropic light propagation in every inertial frame (see Section 18) – and several of its consequences contradict our 'common sense', that is why it is continued to come under attacks. Attempts to deny the theory of relativity are supported by the fact that the usual formulations – both in the old, nonrelativistic and in the new, the relativistic case – are based

– on **intuitive notions** i.e. notions which are not precisely defined, which are supposed 'obviously known' for everybody and therefore do not require a deep explanation

- and on **tacit assumptions** i.e. seemingly suitable 'natural properties' of the intuitive notions

which, however, can contradict each other, leading to apparent paradoxes.

If we want – and we want – avoid misunderstandings and paradoxes then we have to eliminate the intuitive notions and tacit assumptions; we have to built up theories in which everything is perfectly defined.

We succeed in doing so taking into account that any physical theory is a mathematical model of reality. We can construct diverse models of the same physical object depending on our aim: for what we use the model, what kinds of experience are to be reflected in the model. An example: if our knowledge of water is to be used for shipping then water is described as a continuous medium which has viscosity etc.; on the other hand, if water is considered from chemical point of view then it is described as a set of H_2O molecules which have affinity etc. It is most important that a model is a mathematical object so we must give mathematically exact definitions and propositions.

All these hold for the models of spacetime, too. Not only for relativity theory but for every other theory, thus for the old, nonrelativistic theory, too.

Since spacetime appear in every physical theory, it is most important that spacetime shall get an exact mathematical framework; an eventual misunderstanding concerning spacetime can cause mistake in other theories, to.

This book aims to make clear the basic principles and notions of spacetime models, to deduce its mathematical formulae. Intuitive notions will be made 'sharp', tacit assumptions will be made 'clear'; attention will be called to the points that are confused in usual treatments and cause paradoxes.

A former book[1] expounds both the nonrelativistic spacetime model and the special relativistic spacetime model, without explaining how to arrive to them.

[1] T. Matolcsi: *Spacetime without reference frames*, 1993, Akdémiai Kiadó, Budapest

II Introduction

Now we examine the usual intuitive notions and tacit assumptions concerning space and time.

1. The first notion is a reference frame which is defined as a Cartesian coordinate system fixed to a body; for example, the axes of a coordinate system are the three edges of the wall meeting in a corner of a room. The intuitive notions – 'everybody knows what they are' – of this definition are

– the **space** in which the coordinate axes are fixed,

– the **straight line**,

– the **right angle**.

Tacit assumptions are the following:

– the body in question is **rigid** (how could we fix a Cartesian coordinate system to a room having rubber walls and shaken by an earthquake?),

– **there are straight lines in space**, more precisely the necessarily finite straight line segments experienced by us (such as the edges of the wall) can be extended to infinity,

– the straight lines **obey the rules of Euclidean geometry**, i.e. two straight lines can meet in one point only, it is meaningful that two straight lines are parallel, the sum of the angles of a triangle is the straight angle, a straight line segment has a length, the Pythagorean rule is valid, etc.

2. Inertial frames play a distinguished role; they are usually defined by Newton's first law: "there are reference frames – called inertial frames – in which a body left alone (free of forces) stays at rest or moves uniformly along a straight line".

Such a definition assumes the intuitive notion of

– **uniform motion**: "a body moves uniformly if it covers equal distances over equal times",

which involves the tacit assumption that

– **time passes uniformly**

and the intuitive notion

– **the time interval between two events occurred in two different space points**.

First of all let us observe that

– the usual definition of a reference frame – Cartesian coordinate system fixed to a rigid body –implies space only (by the intuitive notion of straight lines),

– whereas the definition of inertial frame involves time, too (by the intuitive notion of uniform velocity).

Thus, it is misleading to say that an inertial frame is a special reference frame. It is worth mentioning, too, that in the cited definition of inertial frames a coordinate system is not referred to, only straight lines in space appear. Thus inertial frame needs space and time but not coordinate systems.

Later on, giving an exact definition, we use the notion of reference frame in a sense different from the quoted intuitive notion which we will call a **spatial coordinate system** instead.

3. More serious questions arise in connection with the intuitive notion of uniform motion; not clarifying them one can get in trouble in the theory of relativity.

According to the usual definition "the motion of a body is uniform if it covers equal distances over equal times". In this definition

– the intuitive notion of **the time necessary for covering a distance**

appears („it need not be clarified, everyone knows what it is"). To see what is in question, let us formulate it precisely in the case of an everyday situation. Let us consider a foot race: how to measure the time elapsed from the start and the arrival of a racer?

We have three simple methods.

The first (usual) method requires a stop-watch which is set off and stopped by electric signals coming from the start and from the goal, when the sprinter starts and arrives, respectively.

The second method eliminates electric signals and requires two identical stop-watches. The two watches, placed closely together, are set off simultaneously. Then the two watches are transported to the start and to the goal, respectively. When the sprinter starts, the corresponding watch is stopped; and when the sprinter meets the goal, the corresponding watch is stopped. The desired duration is obtained as the difference of the values measured by the two watches.

The third method is an improvement of the previous two ones. It requires two stop-watches, one at the start, the other at the goal; a control centre, having equal distances from the start and from the goal sends simultaneously radio-signals which set off the corresponding watches. Then the procedure coincides with that of the second method.

According to the usual (superficial) view all the three methods – and every other one – give the same result. Let us investigate whether it is true or not.

In the first method the watch sets off and stops when the corresponding electric signal arrives at it, so the measured duration depends on the distance the watch is placed from the start and the goal, because the travel time of the electric signals is included in the measurement. This travel time is negligible in everyday situations, but for faster motions (near to the velocity of light) it can be significant.

The second method seems better than the earlier one because the watches stop exactly when the sprinter starts and arrives, respectively. The question arises, however, whether the time-keeping of the watches is influenced by their transport. It is not excluded that they go faster or slower during the transport. If so, the measured duration depends on the transport. In everyday situations we do not experience a change in the time-keeping of watches due to a transport, which means that such a change – if exists – is negligible. It is not excluded, however, that in extreme situations – e.g. in case of a transport with shaking very much – the change can be significant. (In fact, experiments with particle accelerators show such a significant change.)

The third method seems the best because “it is commonly known that light (electromagnetic wave) propagates everywhere in every direction with the same speed”, so the two watches are set off ‘in the same moment’ and no problem occurs in connection with their stops. But how we know that the speed of light has the mentioned property? To know this property we should measure the speed of light i.e. the time during which a light signal covers a given distance. None of the three methods is applicable and even we have no method for that. The astonished reader is asked to keep it in mind; later the exact meaning of the homogeneous and isotropic light propagation will be explained.

Recapitulating: ‘the time necessary for covering a given distance’ i.e. **the intuitive notion of velocity is rather questionable.**

Lastly, for better understanding, we call attention that ‘the time necessary for covering a given distance’ is a special case of the intuitive notion of

– **the time passed between two events that occurred in two different space points;**

one of the events is the start and the other is the arrival of the sprinter. In general, however, it is not necessary that two events be connected with the same material object; the problem is the same when we are interested in the time passed between the referee's first whistle and the first goal in a football match.

4. The usual treatments of spacetime which are based on spatial coordinate systems moving relative to each other run as follows.

Let us consider a Cartesian coordinate system (X, Y, Z) and another one (X', Y', Z') which moves with uniform velocity v in such a way that the corresponding axes of the systems are parallel, the axes X and X' coincide, as it is shown in Figure 1.

Let us remark immediately that such a treatment, at the very beginning, is built up on two intuitive notions:

– the **velocity**,

– the **parallelism of two straight lines moving with respect to each other**.

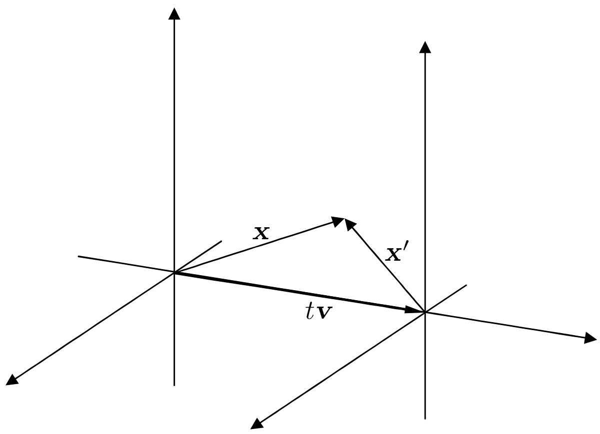

Figure 1 Galilean transformation rule

The latter is 'evident', is not? We can see from the figure that a time period t after the meeting of the origins, the vector $\mathbf{x}$ of the first coordinate system is observed by the other system as the vector $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$, in coordinates:

$$x' = x - vt, \qquad y' = y, \qquad z' = z.$$

This is the well known Galilean transformation rule.

Namely, the Galilean transformation rule is established on the basis of the arrows of Figure 1. It is rather peculiar that in the usual treatments of relativity theory, the figure of moving coordinate systems with parallel axes appear, too, but those arrows are not drawn at all, because they do not explain the Lorentz transformation rule

$$x' = \frac{1}{\sqrt{1-(v/c)^2}}(x - vt), \qquad y' = y, \qquad z' = z.$$

Thus, what the arrows of the figure show 'trivially' is not true in the relativistic case. Can we still attribute a 'trivial' meaning to the parallelism of moving straight lines?

5. Let us examine the parallelism of moving straight lines which "is so trivial that one need not clarify it" because looking at the figure one imagines that a little time earlier when the origins of the spatial coordinate systems coincided, then also the corresponding axes coincided, too. This idea is based on our everyday experience when we we see that an edge of a moving lorry coincides in a moment with an edge of a house.

Let us formulate this experience. There is a straight line in our space and its image in our eyes coincides in a moment with the image of *something* that moves; in other words, light arrives in the same moment at a single point (our eyes) from every point of our straight line and from every point of the moving object.

The following arguments need not be consider exact because they use the intuitive notion of light speed; however, they show very well in the usual framework the problem of parallelism of moving objects.

Let us take two orthogonal planes in our space, their intersection is the straight line L. A camera is in one of the planes, in the other a line is moving with velocity v. The lens of the camera is l distant from the line, i.e from a point A of L and it is $\sqrt{z^2+l^2}$ distant from another point B of L. If c is the light speed, then the travel times of light from the two points till the lens are l/c and $\sqrt{z^2+l^2}/c$, respectively. Thus, the light from B starts $\frac{1}{c}(\sqrt{z^2+l^2}-l)$ earlier than the light from A. The points A' and B' of the moving line send light to the lent when they meet A and B, respectively. The light starts earlier from B' than from A', therefore when A' meets A, B' has already left B; their distance is just $\frac{v}{c}(\sqrt{z^2+l^2}-l)$.

Consequently, the moving line is described by

$$\left(\frac{v}{c}(\sqrt{z^2+l^2}-l), z\right);$$

this is not a straight line. Of course, in everyday situation, it is practically straight because the first component is negligible in comparison with the second one. Indeed, if $l = 10$ m, $0 < z < 3$ m, $v = 30$ m/s, then – since $c = 3 \cdot 10^8$m/s – $\frac{v}{c}(\sqrt{z^2 + l^2} - l) < 4,4 \cdot 10^{-8}$m.

Our eyes (cameras) mislead us.

The attentive reader can notice that the reasoning above uses the time passed between two events that occurred in two different space points (the travel time of light from a point of the axis till the lens) and the velocity of light whose precise meaning was questioned previously, so he/she can have doubts about it. We reassure him/her: we have not questioned that these notions can be defined precisely.

It is sure, however, that the question of parallelism of lines moving with respect to each other remains open. Moreover, besides this notion of parallelism, the usual transformation rules contain even the assumption that the scales are the same on the lines, the open question can be formulated more sharply: **what does it mean that two space vectors that are moving with respect to each other are equal?**

So it is sure as well that without answering this question we cannot interpret the usual treatments.

III Construction of spacetime models

1 Heuristic of spacetime

In this section we gather our fundamental experiences that any spacetime model must be based on. We emphasize that **only physical facts will be taken into account for the construction of models**, artificial facts such as coordinate systems will be avoided.

1.1 Space

Space is one of our simplest and fundamental notion.

What is space? We are sitting in a room: a corner of the room, a spot on the carpet are points of our space, the table is a part of our space. Looking out the window, we see trees, chimneys, hills that form other parts of our space. A car travelling on the road is a phenomenon to us, not a part of our space.

On the other hand, the seats, the dashboard, etc. constitute the space for someone sitting in the car. Looking out he sees houses, trees, chimneys running; they are not parts of his/her space.

It is obvious then that the space for us in the room and the space for the one sitting in the car are different. We can assert that space is constituted by some material objects – let us call them **observers** – and **the spaces of different observes are different**, in other words, **space is a relative notion**. i.e **there is no absolute space**.

The spaces of different observers can have very different properties: e.g. the space of our room differs very much from the space of a room with rubber walls which is shaken by a earthquake.

The properties listed below are valid for the space of 'good observers'; such are the **inertial observers** determined by the physical fact that **their space**

points are material points free of any action and non moving with respect to each other .

In the sequel observer will mean an inertial observer.

Our simplest experience regarding the space of an observer lies in the following:

(**S1**) **There are straight lines in the space** and a **vector** i.e. an oriented straight line segment can be put between two space points; the vectors obeys the well known rules.

(**S2**) **Space is three dimensional** i.e. there are three essentially different directions – right-left, ahead-back, up-down – from which any other directions can be 'composed'.

Less simple experience (decay of K-mesons) demonstrates:

(**S3**) **Space is oriented** i.e. 'right-handed' order (right-ahead-up) and the and 'left-handed' order (left-back-down) are not equivalent.

We can measure distances and angles in our space. Distances and angles have the well-known Euclidean properties (e.g. the shortest way between two points is the straight line, the sum of the angles of a triangle is 180^o etc); thus

(**S4**) **Space has a Euclidean structure**.

Let us mention that the listed properties concern 'human size' experience . The question arises whether they are valid beyond human size, for instance, whether the vector between two stars of distant galaxies or the vector between two atoms of a crystal are meaningful. We do not answer such questions because we construct models extrapolating our human size experience; keep in mind that a model is not the reality itself but a human picture of reality.

1.2 Time

1.2.1 Time period, time point

Time, too, is one of our simplest and fundamental notion.

Processes indicate that time progresses: we are breathing, someone is speaking, a clock is ticking, the sun proceeds in the sky. **Time, too, is constituted by material objects**.

Both in everyday language and in the usual terminology of physics, time means

– time period: 'long time ago' 'a lot of time passed'

– time point: 'what is the time?' 'in the same time'.

Moreover, the same device (clock) *measures time periods* and *shows time points.* That is why the two distinct notions, time period and time points, are frequently confused, though it is extremely important to distinguish between them.

Superficially – according to the everyday experience – one could think that the time period between two time points is similar to the vector between two space points. However, this is not the case. Namely, both space points and the vectors between them are 'tangible' facts but this does not hold for time points and time periods, as we shall see.

1.2.2 Proper times

Time is passing: we grow older. Not only living organisms: our tools, furniture, house, the rocks in the mountain grow older, too (nowadays there are excellent methods for determining the age of a stone).

We experience the flight of time through time periods. I feel the time period between my breakfast and my dinner (by hunger) and you feel the time period between your breakfast and your dinner. I cannot feel the time period between your breakfast and your dinner. In the same way, an elementary particle feels the time period between its birth and death.

We can accept as a fundamental fact that time progresses for every material point. Visualizing this, we imagine that a minute quartz crystal is tied to the material point and the oscillations (ticks) of that crystal measure the time passed for the material point. Such a minute device measuring the progress of time will be called a **chronometer**. We intentionally do not say 'watch' or 'clock' in order to avoid the possibility that the everyday properties of our usual device mislead our thinking: a chronometer has no dial, it does not show what the time is ('it is two o'clock' makes no sense in this respect), it measures only the time passed – the number of ticks – between two arbitrary occurrences of the material point (e.g. between 'my having breakfast' and 'my having dinner').

We have the following fundamental experience regarding proper time:

(**T1**) **Proper time is one dimensional** because only forward (future) and backwards (past) exist.

(**T2**) **Proper time is oriented** because future and past cannot be interchanged.

It is emphasized, all these are valid for individual proper times. We have not stated anything about the relation of individual time passing.

It may happen – why not? – that different amounts of time passes for two material points between two of their meetings. For instance, let us consider me and my home as two material points. I leave my home, I fly to another country for holiday, then I return; it is not evident that in this period my chronometer has ticked just as many as the chronometer left at home. According to our everyday experience, the same proper times elapse for two material points between their two meetings, or better to say, the difference of elapsed proper times does not exceed the practical error of measurements. But this experience refers to 'mild' circumstances. Let us consider two chronometers resting together and then one of them being left there and the other one being revolved again and again very fast in a centrifugal machine; is it certain that the two chronometers have measured the same number of ticks between their meetings?

1.2.3 Simultaneity (synchronization)

A proper time period – the number of the ticks of a chronometer – is a physical fact. Such a proper time period is meaningful only between two events of a material point. The time period between 'my having breakfast' and 'your having dinner' cannot be measured by a chronometer (unless we are together).

The chronometer of a material point (a quartz crystal) measures the time passed (the number of ticks) between two arbitrary occurrences of the material point. Our usual **notion of time period concerns a physical fact**.

On the contrary, our usual **notion of a time point is only a human convention**.

A chronometer does not show what the time is; does it make sense at all when ('at what time') an occurrence takes place?

Now we investigate how the time point (instant) of an occurrence can be defined. If it can be defined somehow, then it makes sense that two occurrences (of two different material points) are **simultaneous** (they happen at the same instant). Let us put the question in this way: in what conditions are two occurrences simultaneous? How can we state e.g. that two explosions in different places happened at the same instant?

Our everyday experience regarding human size phenomena indicates that simultaneity has a sense independently of other facts. What is happening **now** in the street? Looking out the window, we see it. We can have doubts, however, as to this being true, since what we see does not happen 'now' but happened a little 'earlier' because light requires time – though ever so small – for arriving at us from the street. What is happening **now** in a town 240 kilometers far from us? The answer is not so immediate as previously. Say a firework explodes both in Budapest and in Vienna; how can we determine whether the explosions are simultaneous or not?

Proper time progresses (a chronometer ticks) in each space point of an observer but these proper times are not related to each other. A quartz crystal oscillates in Budapest, another one oscillates in Vienna, measuring how time progresses; they do not give information about – does it make sense at all? – when a tick in Budapest is simultaneous to a tick in Vienna. In other words, they do not say e.g. when midnight in the two towns is.

How do we define simultaneity between different places of the Earth? The ancient method uses the position of the Sun and the stars. Let us imagine the Sun and the rotating Earth in the huge system of fixed stars. In the following description, for historical reasons, we choose Greenwich as the 'centre' of synchronization. In a given day, knowing where the Earth is on its orbit around the Sun, we define midnight in Greenwich by prescribed positions of the stars with respect to Greenwich. Then the instant in Budapest corresponding to Greenwich-midnight is defined also by a certain prescribed position of the stars with respect to Budapest. Similarly, midday in Greenwich and in Budapest is defined by prescribed positions of the Sun with respect to the towns. (Note that the phrase 'prescribed position' refers to different skylines in the two towns; e.g. if midday in Greenwich is defined as the instant when the Sun is on the highest point of its orbit, then Greenwich-midday in Budapest is defined by a position of the Sun having past its summit by a given angle.) Of course, to determine simultaneity in this way requires highly precise measurement of directions, angles. The inaccuracy of measurement can result in an error of some seconds which is practically negligible.

Nowadays we define simultaneity in another way, too, with the aid of radio signals. This is an extrapolation of our everyday experience of simultaneity based on seeing. A radio signal (a pip) starts at midnight in Greenwich towards Budapest; knowing the distance of the two towns and the speed of light, it is known how much time the signal travels, and the time point at the arrival of the signal will be set correspondingly.

The movement of the Sun, the position of stars – i.e. the rotation of the Earth – made it evident how to determine simultaneity on the surface of the Earth; this notion of simultaneity, together with that based on seeing, was developed and became a part of our everyday life long ago. That is why it seems that simultaneity is an absolute law of nature. This fiction of absolute simultaneity was ruled out by the theory of relativity. All the paradoxes formulated against the theory of relativity involve implicitly absolute simultaneity or some of its consequences.

It is worth emphasizing: according to the everyday thinking, absolute simultaneity does exist and both the procedure using celestial bodies and the one using radio signals 'measure' the existing simultaneity, much the same way as we measure (by some procedure) the existing distance between two towns. In fact, however, simultaneity in itself does not exist, the mentioned procedures **define** simultaneity, perhaps different ones. Indeed, the simultaneity determined by the position of stars differs from that obtained by radio signals as it will be shown (see Paragraph 15.2); the difference, though being a practically negligible split of a second, is theoretically important: **simultaneity is a human convention, not a physical reality**.

An observer, in general, can define simultaneity for its space points by an arbitrary procedure; **synchronization** means establishing simultaneity in a continuous manner (we will come back to the precise definition).

Keep in mind that the same observer can, in general, establish many different synchronizations.

An observer can realize the time of a synchronization in such a way that a device, called **synchrometer** is put in each space point showing continuously the instants of the synchronization. Up to now we have illustrated simultaneity concerning towns (space points of an observer far from each other) but the same will hold for places of a building (space points of an observer near to each other), too. For instance, a synchrometer in every point of a school shows the instant of break opening, end of lessons etc.

We have introduced two different devices, the chronometer measuring only proper time periods of material points, and the synchrometer indicating only time points of a synchronization. In everyday usage, corresponding to the dial of a clock, a time point is represented by a number ('two o'clock'). A time point, however, in its original sense has nothing to do with numbers which is emphasized by us when speaking about instants (midnight, midday, end of lessons) that can be described without numbers.

Synchrometers and chronometers are different and independent devices: time points of a synchronization are different from and independent of time periods of a world line. It cannot be excluded that different times pass between two synchronization instants in two space points of an observer (between yesterday midnight and today midnight the chronometer in Budapest might not tick the same number of times as the chronometer in Vienna). If this is not experienced – at least within a practically negligible error –, then a chronometer and a synchrometer can be united in a single device, in a clock which represents time points by the corresponding time periods passed from a reference instant (midnight) (see Subsection 3.2.3).

1.3 Motions

1.3.1 Paths of motions

We experience that material objects moves with respect to us, i.e. they change their position in our space. In general, an observer can observe motions.

Motion is relative. Let us consider a comet far away in the universe. It does not move in itself, it exists. But it moves with respect to the Earth, moves with respect to the Mars, moves with respect to the Sun, it moves differently with respect to the different celestial bodies. A bird flies: it moves differently with respect to the street and with respect to a car travelling on the street. The comet exists, the bird exists; their existence (history) is perceived by different observers as different motions.

The path of a pointlike material point in the space of an observer is the collection of space points that meet the material point.

Our first simple experience is:

(**M1**) **Every straight line can be the path of a motion.**

1.3.2 Faster-slower

A swallow flies faster with respect to the Earth then a sparrow (this need not be true with respect to a car: let the two birds fly over the road in the same direction as the car travels; then the birds fly backwards with respect to the car, the sparrow faster than the swallow). How to determine the speed of a motion? This was outline in item 3 of the Introduction. According to 1.2.3, to give a quantity for the characterization of the speed – i.e. to establish velocity –, the simultaneity of different space points, i.e. a synchronization is necessary. Since synchronization is not a physical reality, the notion of velocity cannot be used in the construction of a spacetime model.

Recall, however, that the notion earlier-later is meaningful for every world line, in particular, in every space point of an observer. Let us consider a foot race. Two racers leave the start together (their leaving is the same occurrence of the start) but they arrive at the goal separately (their arrivals are different occurrences of the goal). The racer who arrives earlier is faster, who arrives later is slower. Thus, we know which of the racers is faster without knowing what their time results are (regardless of how it is measured). In other words, we know which of them is faster without knowing how fast they are (what their speeds are).

More precisely, we can declare the following important statement: although the speed of a motion is meaningful only with a synchronization, **it makes sense**

without synchronization which one is faster of two motions having the same path relative to an observer, and meeting at one point.

Regarding to the comparison of motions, we know the following:

(**M2**) **For every motion there is another, faster motion on the same path.**

(**M3**) **For every motion, all the smaller motions are possible on the same path.**

The content of the first statement is evident. The second one needs an explanation. Let us consider the motion of a racer1 on a path between a start and a goal. For an arbitrary time period T there is another racer2 which starts together with racer1 and arrives at the goal T time after racer1.

At last we emphasize that faster-slower on the same path is independent of synchronizations but can depend on observers: recall the swallow and the sparrow.

1.3.3 Uniform motion on a straight line

One of the most important law of physics is Newton's first law. We have seen in the Preface the problems connected with it. Its usual formulation requires a synchronization, and even a special one. Namely, it may happen that the same motion 'covers equal distances over equal times' with a synchronization, but it 'covers different distances over equal times' with another synchronization.

The right formulation of Newton's first law would be the following: "There are observers and synchronizations such that a body left alone (free of forces) stays at rest or moves uniformly along a straight line".

This cannot be used for our purposes because of synchronization. Let us leave synchronization for proper time in such a way that the body and the observer cooperate.

Let a body free of forces measure its proper time; let an arbitrary time period be prescribed; let the body mark the space points of the inertial observer that it meets at the end of every time period and let the observer measure the distance between the marked space points.

Then we can state: "An inertial body in the space on an inertial observer moves on a straight line (or is at rest) in such a way that the distances covered during equal proper time periods are equal."

This shows that

(**U**) **There is a uniform relation between inertial proper time and inertial distances.**

1.4 Equivalence of inertial observers

In general, it is stated as a fundamental principle that inertial observers are physically equivalent. The precise meaning of this principle is hardly conceivable. It seems, a basic condition for explaining this principle is a 'natural' relation between the spaces of the inertial observers which was mentioned in item 5 of the Introduction. Therefore we state:

(**E**) **It can be physically established what it means that a vector in the space of an inertial observer equals the space vector of another inertial observer.**

2 Flat spacetime models

In this chapter we treat a special type of spacetime models, using more mathematical tools. Who wants to deal with the non-relativistic and relativistic spacetime models only, can look over this chapter superficially without a deep understanding of the mathematical formulae.

2.1 Measure lines

First of all, we need to clarify the notion of physical dimensions (units of measurements), because speaking of time and space we cannot avoid measurements regarding time and space. In usual treatments time periods as well as spatial distances are given by real numbers, although they are clearly not just numbers. 3 hours or 3 minutes are time periods but 3 in itself is not. The numbers usually mean the corresponding multiple of a unit, but a unit of measurement remains an intuitive notion. Furthermore, the units of measurement, in general, are human conventions (not given by nature), thus their use implies an undesirable arbitrariness.

If we want to avoid – and we do – the complications coming from the use of intuitive notions, then we have to leave the set of real numbers and construct the precise mathematical model of physical dimensions (units of measurements). Of course, the precise mathematical model is based on the common properties that we are used to (e.g. all the values of an observable are unique multiples of an arbitrarily given non-zero value).

In general, the following can be said. Let A be the set of the magnitudes of an observable. Taking an arbitrary element a of A and a non-negative real number α, we can establish an element α times a of A, denoted by αa. This

multiplication by non-negative numbers, has the following properties for all $a \in \mathrm{A}$:

(i) $0a$ is the same element, called the **zero** of A and is denoted by 0 as well,

(ii) $1a = a$,

(iii) $\beta(\alpha a) = (\beta\alpha)a$,

(iv) if $a \neq 0$ then $\alpha a \neq \beta a$ for $\alpha \neq \beta$,

(v) if $a \neq 0$, then for all $b\boldsymbol{n}\mathrm{A}$ there is a β such that $b = \beta a$.

These properties allow us to define an **addition** on A: choosing an arbitrary nonzero element a, if $b = \beta a$ and $c = \gamma a$, then $b + c := (\beta + \gamma)a$. It is easy to show that this definition does not depend of a.

Let us introduce the notation $-\mathrm{A}$ as the set pairs $(-1, c)$ for $a \in \mathrm{A}$ and let us use the notation $-a := (-1, a)$. Lastly, let $\boldsymbol{A}$ the uniokn of $-\mathrm{A}$ and A. Then we can give a multiplication by real numbers and an addition on $\boldsymbol{A}$ that are trivial extensions of the operations given on A. For instance, for all $a \in \mathrm{A}$,

$$\begin{aligned} \alpha a : &= -|\alpha| a \quad \text{if} \quad \alpha < 0, \\ \alpha(-a) : &= -\alpha a \quad \text{if} \quad \alpha > 0, \\ \alpha(-a) : &= |\alpha| a \quad \text{if} \quad \alpha < 0. \end{aligned}$$

These operations turn $\boldsymbol{A}$ into a one dimensional vector space, whose two 'halves' have different importance: the original 'half' A contains the physically meaningful elements, the other 'half' $-\mathrm{A}$ is the result of a suitable mathematical construction. We express this fact mathematically in such a way that $\boldsymbol{A}$ is oriented by the elements of A (for further details on orientation can be found in tha mathematical supplement).

The preceding construction works e.g. for distance, mass, force magnitudes, etc. In some cases — e.g. for electric charge — we are given originally a one-dimensional real vector space of observable values which is oriented by us (we decide which of the electric charges is positive).

Thus we accept that an oriented one-dimensional real vector space, called a **measure line**, models the magnitudes of an observable. Choosing a unit of measurement means that we pick up a positive element of the measure line in question; doing so, we can represent the elements of the measure line by real numbers. For instance, if e is a positive element of $\boldsymbol{A}$, then every element a of $\boldsymbol{A}$ is a unique real multiple of e which is denoted by $\frac{a}{e}$.

In practice some units of measurement are deduced from other ones by multiplication and division; for instance, if kg, m and s are units of mass, distance and time period, respectively, then $\frac{\mathrm{kg\ m}}{\mathrm{s}^2}$ is the unit of force.

The question arises at once: how can we give a mathematically exact meaning to such a symbol? According to what has been said, kg, m and s are elements of one-dimensional vector spaces; how can we take their product and quotient? To give an answer let us list the rules associated usually with these operations; for instance,

$$(\alpha\mathrm{m})(\beta\mathrm{kg}) = (\alpha\beta)(\mathrm{m\,kg}), \qquad \frac{\alpha\mathrm{m}}{\beta\mathrm{s}} = \frac{\alpha}{\beta}\frac{\mathrm{m}}{\mathrm{s}}$$

for all $\alpha, \beta \in \mathbb{R}^+$. Extending these rules to negative numbers, too, we see that the usual multiplication is a bilinear map on the measure lines and the usual division is a linear-quotient map, with the additional property that the product and quotient of non-zero elements are not zero.

Consequently – on the base of the mathematical supplement –, we can state that the product and quotient of elements of measure lines are to be defined by their tensor product and tensor quotient, respectively. For the time being, the reader need not know too much about tensorial operations; it suffices bear in mind that the operations in question are well defined and have the usual properties. If $\mathbb{D}$, $\mathbb{I}$ and $\mathbf{G}$ denote the measure line of distance, time period and mass, respectively, and if $\mathrm{m} \in \mathbb{D}$, $\mathrm{s} \in \mathbb{I}$, $\mathrm{kg} \in \mathbf{G}$, then $\mathrm{m\,kg}$ and $\frac{\mathrm{m}}{\mathrm{s}}$ are well defined elements of $\mathbb{D} \otimes \mathbb{G}$ and $\frac{\mathbb{D}}{\mathbb{I}}$, respectively.

We have given the precise mathematical model of physical dimensions (units of measurements) and their operations, in such a way that (omitting the symbol of tensorial multiplication for elements of one-dimensional vector spaces, according to our convention), we obtain the common formulae regarding units of measurements.

2.2 Spacetime

We characterize events in such a way where and when they occurred. 'Where' refers to space point, 'when' refers to time point. A space point is physical reality, a time point – the result of a synchronization – is not. But synchronizations are possible and some of them can be related to proper times according to property (**U**).

Different observers and synchronizations associate different space points and time ppoints to the same event. Space points and time points of an event are special 'pictures' of an **absolute** entity.

In this way we arrive at the notion of **spacetime**. Roughly speaking, a spacetime point is a fusion of a space point and a time point: we can conceive a spacetime point as 'here and now' or 'there and then' keeping in mind that 'here and now' exists as an absolute unity not determining 'here' and 'now' separately.

The spacetime points can be illustrated by pointlike physical phenomena: two tiny balls collide, a lamp flashes, a firework explodes. That is why spacetime points are usually called 'events' which, unfortunately, has led to some misunderstandings. Namely, the word event has a definite meaning in probability theory applied in several areas of physics. The collision is an event of the balls, the flash is an event of the lamp, the explosion is an event of the firework: all these are not events of spacetime. They illustrate a spacetime point, the same spacetime point can be illustrated by two different events: the collison of two balls as well as the flash of a lamp can happen 'here and now'.

That is why we avoid the terminology 'event' and we call a spacetime point **occurrence** when relating a physical phenomena to it; also, instead of spacetime point we frequently say **world point** .

The properties (**S1**)-(**S3**) in 1.1 suggest that the space of an inertial observer is a three dimensional oriented affine space, (**T1**) and (**T2**) in 1.2.2 and (**U**) in 1.3.3 suggest that the proper time of an inertial material points is an oriented one dimensional affine space; at last, we can state according to property (**U**) that these affine structures are not indepenent.

Therefore we accept:

> The basic notion of a flat spacetime model is **spacetime, a four dimensional oriented affine space.**

Spacetime will be illustrated by the plane of the page. Thus, occurrences (world points) will be represented by the points of the page. Though such a representation is useful, we have to be cautious: we must not attribute the usual properties of the plane to spacetime. For instance, two points in the plane have a distance but two world points do not.

The affine spacetime will be denoted by $\mathtt{M}$, the underlying vector space will be denoted by $\mathbf{M}$.

It is important that the world points (elements of $\mathtt{M}$) are different from the **world vectors** (elements of $\mathbf{M}$). This difference gets lost in the usual treatments based on coordinates: both world points and world vectors are represented by quartets of numbers.

According to our general agreement, spacetime is illustrated in the plane of the page. The vector between two world points is represented by an arrow between the points (see Figure 2.2)

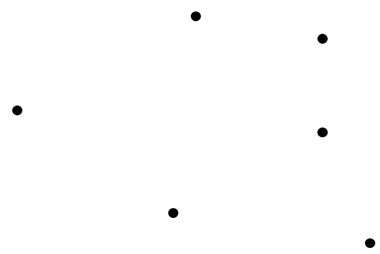

Figure 2.1 Occurrences

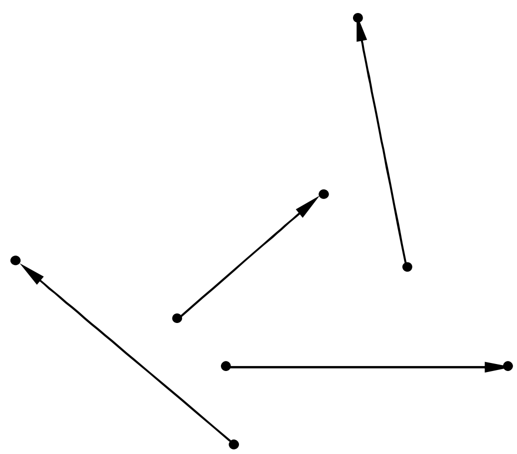

Figure 2.2 Affine structure of spacetime

Since we have no other choice, spacetime vectors, too, will be illustrated in the plane of the page; then vectors are represented by arrows starting from the zero (see Figure 2.3. In order to avoid misunderstandings, we always indicate in the figures whether `M` or **M** is illustrated.

Though such a representation of vectors is useful, we must be careful not to be deceived: in general, the length of a spacetime vector and the angle between two vectors are meaningful.

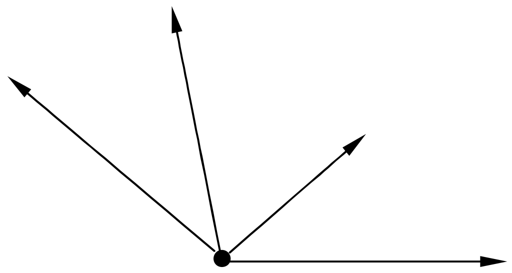

Figure 2.3 Spacetime vectors

2.3 Futurelike vectors

2.3.1 World lines

The life (history) of a 'tiny' material object, a (classical) pointlike body, is a continuous sequence of occurrences: this will be a **connected one-dimensional submanifold**, briefly a **curve** in spacetime.

The sequence of occurrences of a material point have a definite order: it makes sense, which of two occurrences of a material point is earlier or later than the other (e.g. if I am considered such a body, then 'my having breakfast' is earlier than 'my having dinner'). The relation earlier-later in the life (history) of a material point means an **orientation** on the corresponding curve in spacetime (as concerns the orientation of a curve see the mathematical supplement).

Therefore we accept that the life (history) of a classical material point is an **oriented curve** in spacetime, called **world line** in spacetime.

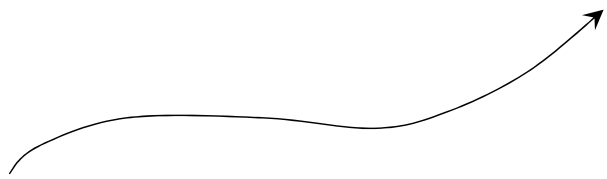

Figure 2.4 World line, directed from earlier to later

The affine structure allows us to formulate our experience that there is no distinguished part of spacetime: what can occur 'here and now', can occur 'there and later' (spacetime is 'homogeneous'). Thus we accept that every translated of a world line is a world line, too: if C is a world line, then $C + \boldsymbol{a}$ is world line for all $\mathbf{a} \in \mathbf{M}$.

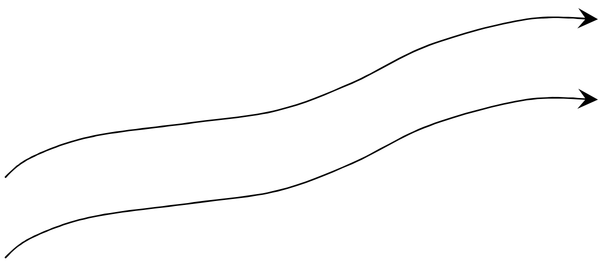

Figure 2.5 Translated world lines

Some oriented curves in spacetime are world lines, some others are not. If a curve with a given orientation is a world line, then with the opposite orientation it is not. Moreover, there are curves that cannot be world lines with any orientation. This can be illustrated as follows. Let us consider a string of lamps for decoration. The lamps can flash 'one after the other' or 'at the same time'. In the first case the sequence of flashes correspond to a world line because they can be obtained by a single lamp moving along the string and blinking. In the second case a single lamp cannot produce the flashes.

In what follows we establish which curves can be world lines in spacetime.

2.3.2 Inertial world lines

One of the arguments for the affine structure of spacetime was our experience of uniform motions on straight lines which refer to inertial material points. That is why we accept that **in flat spacetime, the world line of an inertial material point is an oriented straight line.**

An oriented straight line is determined uniquely by an arbitrary point of its and a direction vector. If an oriented straight line is a world line, then, according to the previous statement, every straight line parallel to that and oriented by the same direction vector is a world line, too.

Accordingly, the set of possible inertial world lines can be given by the possible direction vectors. Since the orientation of a world line expresses earlier-later, a direction vector goes from an earlier point to a later one. Such a direction vector is called **futurelike**. straight line is a world line, then its parallel transport is a world line, too. Thus, in order to determine the set of inertial world lines, we need to specify the set of futurelike vectors.

We know that the orientation of a straight line means a 'half line' of its direction vectors; therefore, the positive multiples of futurelike vectors must be futurelike (in other words, the set of futurelike vectors is a cone whose apex is the zero vector) and the negative multiples of futurelike vectors must not be futurelike.

Furthermore, properties (**M1**)-(**M3**) in 1.3 imply that the **set of of futurelike vectors must be convex and open** (which will be proved in 2.8.4).

The structure of a flat spacetime model must contain the set of **futurelike vectors**, a subset $\mathtt{T}^{\rightarrow}$ of $\mathbf{M}$, which is open and convex cone whose apex is at zero.

Thus, if $\boldsymbol{x}, \boldsymbol{y}$ are in $\mathtt{T}^{\rightarrow}$ and $\alpha, \beta \geq 0$ are real numbers, $\alpha + \beta \neq 0$, then $\alpha\boldsymbol{x} + \beta\boldsymbol{y}$ is in $\mathrm{T}^{\rightarrow}$, too.

Note that the zero vector (the apex of $\mathtt{T}^{\rightarrow}$) does not belong to $\mathtt{T}^{\rightarrow}$.

$\mathtt{T}^{\rightarrow}$, being convex, is a connected set.

The word 'cone' may lead someone to imagine $\mathrm{T}^{\rightarrow}$ as the body known from elementary school, but $\mathrm{T}^{\rightarrow}$, being an 'infinite' cone, can have many different forms.

Figure 2.6 shows three essentially different forms in three dimensions.

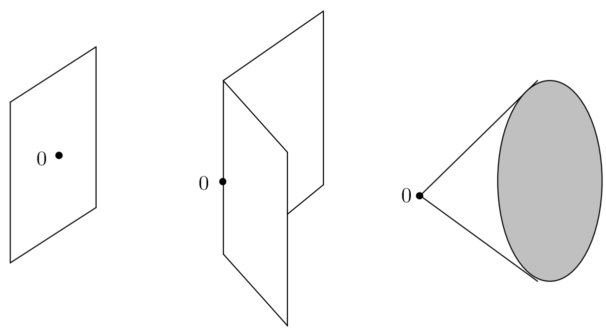

Figure 2.6 Futurelike cones

In our usual illustration in the plane of the page, the first and the third possibilities take the form shown in Figure 2.7 As for the second possibility depicted in Figure 2.6, different ‘natural’ two-dimensional illustrations can be drawn, according to the point of view.

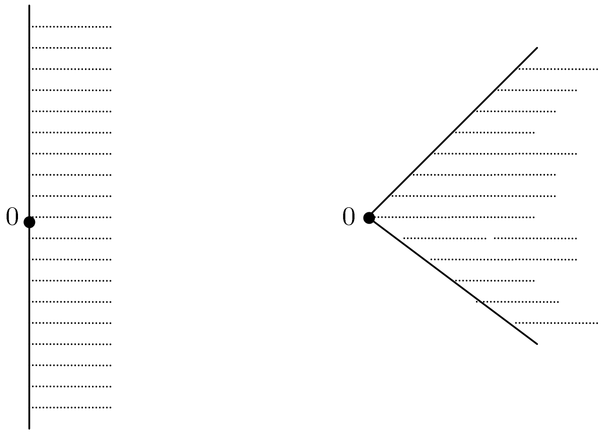

Figure 2.7 Futurelike vectors

A vector opposite to a futurelike vector is called **pastlike**, their set, $\mathrm{T}^{leftarrow} := -\mathtt{T}^{\to}$ is disjoint from $\mathtt{T}^{\to}$. A **timelike** vector is either futurelike or pastlike.

If x and y are spacetime points and $y - x \in \mathtt{T}^{\to}$, then y is called **futurelike or pastlike with respect to** x if $y - x \in \mathrm{T}^{\to}$ or $y - x \in \mathrm{T}^{\leftarrow}$, respectively.

Thus $x + \mathrm{T}^{\to} = \{x + \mathbf{x} \mid \mathbf{x} \in \mathrm{T}^{\to}\}$ is the set of world points futurelike with respect to the world point x.

The occurrences x and y are called **timelike separated** if $x-y$ is a timelike vector.

2.3.3 Properties of world lines

Since ‘in small scale’ every curve is approximately straight, it is reasonable to accept:

A **world line** is a **curve** whose **all tangents are timelike.**

This means that if p is a parametrization (see the mathematical supplement) of a world line, then $\dot{p}(a)$ is timelike for all parameter values a. The set of timelike vectors is the disjoint union of the set of futurelike vectors and the set of pastlike vectors; the derivative of a the parametrization is continuous, thus all $\dot{p}(a)$ are futurelike or pastlike. The orientation of a world line will be given by a parametrization – called **progressive** – such that $\dot{p}(a)$ is futurelike.

We can demonstrate a fundamental property of world lines which results from the set of futurelike vectors being open.

The vector between arbitrary two different occurrences of a world line is timelike.

In formula: if x and y are different point of the world line C, then $y - x$ is timelike. Equivalently: if x is an arbitrary point of a world line C, then any other point of C is futurelike or pastlike with respect to x, in other words,

$$\mathrm{C} \setminus \{x\} \subset x + T.$$

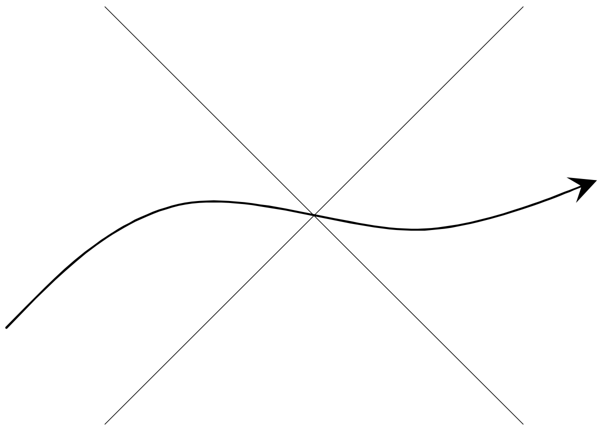

Figure 2.8 World line

Let p be a progressive parametrization of C and $x = p(0)$. Since p is continously differentiable and $\dot{p}(0)$ is futurelike, and $\mathtt{T}^{\rightarrow}$ is open, there is a parameter value $b^+ > 0$ such that $p(a) - x$ is futurelike for all $0 < a < b^+$ and there is parameter value $b^- < 0$ such that $x - p(c)$ is futurelike for all $< b^- < c < 0$. Let s^+ be the supremum of the nonvoid set $\{b^+ > 0 \mid p(a) - x \in \mathtt{T}^{\rightarrow}$ for all $0 < a < b\}$. It will be demonstrated that s^+ is not in the domain of p.

This is trivial if $s^+ = \infty$. Let us suppose that $s^+ < \infty$ is in the domain of p. Then, repeating the previous arguments, we can find a number $d^- > s^+$ such that $p(s^+) - p(c) \in \mathtt{T}^{\rightarrow}$ for all $d^- < c < s^+$. The definition of s^+ implies that $p(c) - p(0) \in \mathtt{T}^{\rightarrow}$ for such a c; thus, $p(s^+) - p(0) = \big(p(s^+) - p(c)\big) + \big(\big(p(c) - p(0)\big) \in \mathtt{T}^{\rightarrow}$. On the other hand, there is a number $d^+ > s^+$ such that $p(a) - p(s^+) \in \mathtt{T}^{\rightarrow}$ for all $s^+ < a < d^+$; thus, $p(a) - p(0) = \big(p(a) - p(s^+)\big) + \big(\big(p(s^+) - p(0)\big) \in \mathtt{T}^{\rightarrow}$. As a consequence, $p(a) - p(0) \in \mathtt{T}^{\rightarrow}$ for all $0 < a < d^+$ which contradict the definition of s^+, so s^+ cannot be in the domain of p.

Consequently, $p(a) - x$ is in $\mathtt{T}^{\rightarrow}$ for all $0 < a$ in the domain of p.

We can argue similarly for the pastlike part of C.

Therefore, the orientation of a world line can be defined in the following way, too:

An occurrence y of a world line is **later (earlier)** than an occurrence x of the same world line if y is futurelike (pastlike) with respect to x.

Accordingly, the orientation of world lines will not be marked in figures any more; the proper time on world lines – corresponding to the representation of futurelike vectors in Figure 2.7 – always progresses 'from left to right'.

Observe that the assertion implies the following: if x and y are world points and

– $y - x$ is timelike, then there is world line – e.g. a straight line (an inertial world line) – passing through them,

– $y - x$ is not timelike, then there is no world line passing through them.

2.4 Progress of time

2.4.1 Inertial times

First of all, we have to give the model of time periods:

A flat spacetime model must contain the **measure line of time periods**, denoted by $\mathbb{I}$.

According to (**U**) 1.3.3, we accept that time passes 'uniformly' on inertial world lines. This is formulated in the model in such a way the the time period between two occurrences of an inertial world line depends only on the vector between them, and, moreover, the time corresponding to the double, triple etc. of a futurelike vector $\mathbf{x}$ is the double, triple etc. of the time corresponding to $\mathbf{x}$.

Thus, we have to assign a positive element $\boldsymbol{P}(\mathbf{x})$ of $\mathbb{I}$ to every futurelike vector $\mathbf{x}$ in such a way that

$$\boldsymbol{P}(\alpha\mathbf{x}) = \alpha\boldsymbol{P}(\mathbf{x}) \qquad \text{(III1)}$$

for all $\alpha > 0$ which has the physical meaning that if x and y are world points, y is futurelike with respect to x, then $\boldsymbol{P}(y - x)$ is the proper time passed on the inertial world line between the two occurrences (inertial time elapsed between them).

Formula III1 is referred to as $\boldsymbol{P}$ is **positive homogeneous**. It is an evident requirement that $\boldsymbol{P}$ be continuous: time progresses 'approximately in the same way' on two inertial world lines whose directions are 'close to each other'; moreover, we assume that $\boldsymbol{P}$ is smooth (sufficiently many times differentiable) Suming up:

In a flat spacetime model there must be given the **inertial time progress**, a positive homogeneous and smooth function $\boldsymbol{P} : \mathrm{T}^{\to} \to \mathbb{I}^{+}$.

2.4.2 Proper time progress on a world line

We conceive that the time passed on a small piece of a world line equals approximately the inertial time corresponding to the straight line segment approximating the small piece in question. Approximating the world line by a broken line consisting of small straight line segments, we expect to get an approximation for the time progress on the world line.

Expressing this in a formula: if p is a progressive parametrization of the world line C and $a_1 < a_2 < \cdots < a_{n+1}$ are parameter values, then the time passed on the world line between the occurrences $x := p(a_1)$ and $y := p(a_{n+1})$, approximately equals

$$\sum_{k=1}^{n} \boldsymbol{P}(p(a_{k+1}) - p(a_k)) \approx \sum_{k=1}^{n} \boldsymbol{P}(\dot{p}(a_k))(a_{k+1} - a_k)$$

where we used the properties of $\boldsymbol{P}$ (positive homogeneous and continuous) and the fundamental equality of differential calculus, namely $p(a_{k+1}) - p(a_k) = \dot{p}(a_k)(a_{k+1} - a_k) + \mathrm{ordo}(a_{k+1} - a_k)$.

Note that the left-hand side of the above expression makes sense because, according to our earlier result, $p(a_{k+1}) - p(a_k)$ is futurelike for all k.

The right-hand side of the above expression looks like an approximation of an integral, that is why we accept: the proper time passed on the world line C between the occurrences x and y is

$$\boldsymbol{t}_{\mathrm{C}}(x, y) := \int_{p^{-1}(x)}^{p^{-1}(y)} \boldsymbol{P}(\dot{p}(a))da, \tag{III2}$$

where p is an arbitrary progressive parametrization of the world line.

Note that x and y are arbitrary occurrences of C. Depending on whether x is earlier or later than y, $\boldsymbol{t}_{\mathrm{C}}(x, y)$ is positive or negative.

The time progress on a world line, although defined by a progressive parametrization, is independent of parametrization.

Let q be another progressive parametrization such that both x and y are in the range of q. We know that $S := p^{-1} \circ q : \mathbb{R} \to \mathbb{R}$ is continuously differentiable and $S' > 0$. Therefore $q = p \circ S$ and $\dot{q}(b) = \dot{p}(, S(b))S'(b)$, so

$$\int_{q^{-1}(y)}^{q^{-1}(x)} \boldsymbol{P}(\dot{q}(b))db = \int_{q^{-1}(y)}^{q^{-1}(x)} \boldsymbol{P}(\dot{p}(S(b)))S'(b)db = \int_{p^{-1}(y)}^{p^{-1}(x)} \boldsymbol{P}(\dot{p}(a))da;$$

the first equality is based on the positive homogeneity of $\boldsymbol{P}$, while the second is the consequence of the well known formula of integration by substitution.

It is worth mentioning that our definition allows the possibility that different times pass on different world lines between two world points.

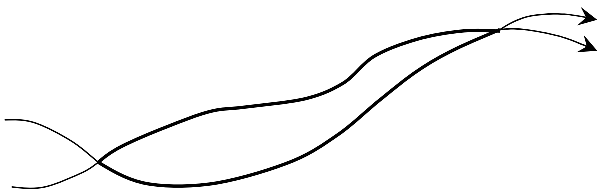

Figure 2.9 Different proper time periods

2.4.3 Absolute velocities

The futurelike vector are direction vectors of inertial world lines. The positive multiple of such a direction vector is a direction vector, too. Time progress makes it possible to introduce some 'unit' direction vecto,rs. Such vectors are obtained by dividing futurelike vector by the corresponding time period. The futurelike vectors are element of $\mathbf{M}$, the time p.eriods are elements of $\mathbb{I}$, thus the quotient will be in $\frac{\mathbf{M}}{\mathbb{I}}$. All the following formulae are understandable without that exact meaning; of this tensorial quotient can be found in the mathematical supplement. Namely, the usual rules are valid, e.g. $\frac{\boldsymbol{x}}{\boldsymbol{s}} + \frac{\boldsymbol{y}}{\boldsymbol{t}} = \frac{\boldsymbol{tx}+\boldsymbol{sy}}{\boldsymbol{st}}$ for $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{I}$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}$.

Let us introduce the notation

$$\mathtt{V}(1) := \left\{ \frac{\boldsymbol{x}}{\boldsymbol{P}(\boldsymbol{x})} \;\middle|\; \boldsymbol{x} \in \mathtt{T}^{\to} \right\} \subset \frac{\mathbf{M}}{\mathbb{I}}.$$

The elements of $\mathtt{V}(1)$ will be called **absolute velocities** because their physical interpretation is 'spacetime path covered over proper time period'.

An element of $\frac{\mathbf{M}}{\mathbb{I}}$ is called futurelike if it is the quotient of a futurelike vector by a positive time period. Then the above definition can be made more palpable, extending the definition of $\boldsymbol{P}$ to the futurelike element of $\frac{\mathbf{M}}{\mathbb{I}}$ by the property of positive homogeneity: if $\mathbf{x} \in \mathtt{T}^{\rightarrow}$ and $\boldsymbol{s} \in \mathbb{I}^{+}$, then let $\boldsymbol{P}\left(\frac{\mathbf{x}}{\boldsymbol{s}}\right) := \frac{\boldsymbol{P}(\mathbf{x})}{\boldsymbol{s}}$. Accordingly,

$$\mathtt{V}(1) := \left\{ \boldsymbol{u} \in \frac{\mathbf{M}}{\mathbb{I}} \,\middle|\, \boldsymbol{u} \text{ is futurelike}, \quad \boldsymbol{P}(\boldsymbol{u}) = 1 \right\}.$$

Thus, if $\boldsymbol{u} \in \mathtt{V}(1)$, then $\boldsymbol{t}\boldsymbol{u} \in \mathtt{T}^{\rightarrow}$ fora all $0 < \boldsymbol{t} \in \mathbb{I}$, and if $\mathbf{x} \in \mathtt{T}^{\rightarrow}$, then $\frac{\mathbf{x}}{\boldsymbol{P}(\mathbf{x})} \in \mathtt{V}(1)$.

Note that because of the positive homogeneity of $\boldsymbol{P}$, **if the absolute velocities $\boldsymbol{u}$ and $\boldsymbol{u}'$ are parallel, then $\boldsymbol{u} = \boldsymbol{u}'$.**

Since the word velocity inevitably reminds us of our everyday notion of relative velocity, it is worth keeping in mind that

– there is **no zero** absolute velocity,

– an absolute velocity has **no magnitude**,

– it makes **no sense** that an absolute velocity is **bigger or smaller** than another one,

– the **angle** between two absolute velocities is **not meaningful**.

2.5 Observers

2.5.1 Physical meaning of an observer

Let us mention in advance that the notions of observer, reference frame, coordinate system are used in the literature with several intuitive meanings, and it sometimes happens that in the same paper different terminologies refer to the same object, or the same terminology refers to different objects. Thus, it is highly important to clarify these notions.

One frequently considers a single material point (world line) as an observer. However, this is not satisfactory. A single material point cannot observe anything. A set of material points is necessary for making observations (experiments). For instance, if the cloud-chamber was one material point, then an ionization path could not come into being. Motion, relative velocity, etc., all these notions can only be defined by a cluster of material points.

An **observer** in a physical sense is a set of many material points; such observers are the room and the car treated in Subsection 1.1. It is important that these objects are extended and not pointlike.

A material point of an observer (a corner of the room, a button on the dashboard of the car) is conceived as a **space point of the observer**.

The history of a material point is a world line in spacetime. Thus a point of an observer is a world line, too.

We have to get accustomed to the following fact: **a space point perceived by us (a corner of the room) is a world line in spacetime**. Thinking about it a while, we do not find this so peculiar: the corner of the room existed yesterday, exists today and will exist (hopefully) tomorrow: the corner of the room is the history of the corresponding material point.

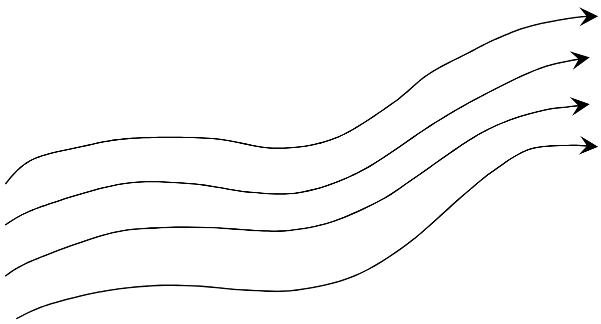

Figure 2.10 Space points of an observer

2.5.2 General observers

From now on tangents of world lines are considered to be elements of $\mathtt{V}(1)$.

In order that such an observer be suitable for natural purposes, its space points must be 'side by side', i.e. the corresponding world lines must fill 'continuously' a domain of spacetime. This can be best formulated with the aid of absolute velocities. Namely, attaching the tangent vector to all points of all world lines representing the space points of the observer, we get an absolute velocity field, i.e. a function assigning an absolute velocity to world points. It is well-known from the theory of differential equations that such a sufficiently smooth vector field determines uniquely the world lines whose tangents are the prescribed absolute velocities.

That is why we find it convenient to accept: an **observer** is an infinitely differentiable map $\boldsymbol{U} : \mathtt{M} \rightarrowtail \mathtt{V}(1)$ defined on a connected open subset.

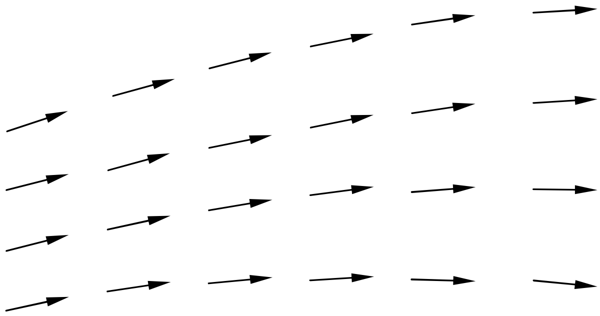

Figure 2.11 Observer

The **space points of the observer** are the maximal integral curves of the differential equation

$$(x : \mathbb{I} \to \mathtt{M})? \quad \dot{x} = \boldsymbol{U}(x).$$

Let us repeat: a space point of an observer is a subset of spacetime, a world line. The collection of these world lines is the space of the observer. **The space of the observer is not a subset of spacetime**. It is a set whose elements are subsets of spacetime.

2.5.3 Inertial observers

Heuristically, an inertial observer consists of inertial material points being 'at rest relative to each other'. This can be expressed by the fact that all space points of an inertial observer are inertial world lines having the same absolute velocity. Thus we accept:

An **observer is inertial** if the prescribed velocity in every world point is the same, i.e. if the velocity field is a constant map.

From now on we refer to an inertial observer by its constant velocity value, i.e. an inertial observer $\boldsymbol{u}$ means the observer $\boldsymbol{U}(x) = \boldsymbol{u}$ $(x \in \mathtt{M})$.

The **space points of the inertial observer $\boldsymbol{u}$** are parallel straight lines directed by $\boldsymbol{u}$. The world line directed by $\boldsymbol{u}$ and passing through the world point x is $x + \mathbb{I}\boldsymbol{u}$, where

$$\mathbb{I}\boldsymbol{u} := \{\boldsymbol{t}\boldsymbol{u} \mid \boldsymbol{t} \in \mathbb{I}\}$$

is a one-dimensional linear subspace in $\mathbf{M}$ which will appear frequently in the sequel.

We have excellent mathematical tools for handling the space of an inertial observer. The reader is asked to study the mathematical supplement (Subsection 19.2) treating a less common part of linear algebra which, however, is not difficult at all.

The space of the inertial observer $\boldsymbol{u}$, the collection of the straight lines directed by $\boldsymbol{u}$, is

$$\mathtt{E}_{\boldsymbol{u}} := \mathtt{M}/\mathbb{I}\boldsymbol{u}$$

which, endowed with the subtraction

$$(x + \mathbb{I}\boldsymbol{u}) - (y + \mathbb{I}\boldsymbol{u}) := (x - y) + \mathbb{I}\boldsymbol{u} \tag{III3}$$

is a three-dimensional affine space over the vector space $\mathbf{M}/\mathbb{I}\boldsymbol{u}$.

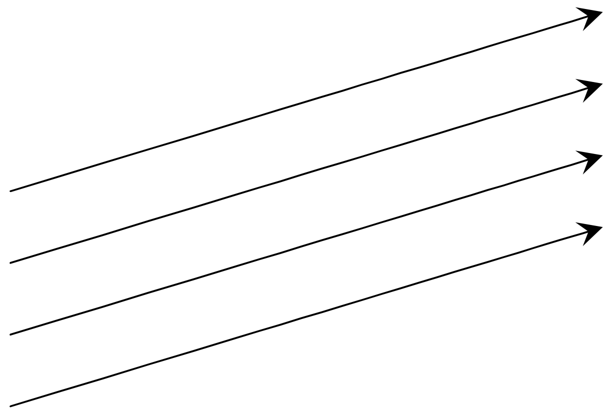

Figure 2.12 Space points of an inertial observer

2.5.4 The space vectors of an inertial observer

The model perfectly reflects the experimental fact that the space of an inertial observer is a three-dimensional affine space. We see the space determined by our room, the points of this space, a vector determined by such two points, e.g. the vector from a corner of the room to a spot on the carpet. We accepted a collection of world lines as a model of our space. Perhaps it was a little curious that a point of our space is a curve in spacetime but, hopefully, this was sufficiently explained in Subsection 2.5 and we have already got accustomed to it, so $\mathtt{E}_{\boldsymbol{u}} = \mathtt{M}/\mathbb{I}\boldsymbol{u}$ as the space of an inertial observer is fairly natural. It is also natural mathematically that $\mathtt{M}/\mathbb{I}\boldsymbol{u}$ is an affine space over the vector space $\mathbf{M}/\mathbb{I}\boldsymbol{u}$. On the other hand the **space vectors of the observer** of $\boldsymbol{u}$, the elements of $\mathbf{M}/\mathbb{I}\boldsymbol{u}$ cannot be related to our experimental vectors in a fairly natural way, e.g. we cannot represent them by arrows in the plane of the page.

We have perfect mathematical models for both space points and space vectors of an inertial observer; for the time being, we have to accept that the model of space points is apparent, the model of space vectors is not. In special cases – both in the nonrelativistic spacetime model nd in the relativistic one – we can improve the situation, the space vectors, too, can be represented in a straightforward way.

Let us recall that the experimental fact that our time is directed, so future and past are essentially different, is reflected in the model by the existence of futurelike and pastlike vectors. Further, the experimental fact that both time and space are oriented, is reflected in the model that spacetime is oriented. Thus the model contains spacetime orientation and time orientation; from them now we can deduce the orientation of the space of inertial observers.

Namely, the vector space $\mathbf{M}/\mathbb{I}\boldsymbol{u}$ (the set of space vectors of the inertial observer $\boldsymbol{u}$) can be endowed naturally with an orientation in such a way that a basis $\boldsymbol{x}_1 + \mathbb{I}\boldsymbol{u}$, $\boldsymbol{x}_2 + \mathbb{I}\boldsymbol{u}$, $\boldsymbol{x}_3 + \mathbb{I}\boldsymbol{u}$ is defined to be **positively oriented** if $\boldsymbol{t}\boldsymbol{u}, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ is a positively oriented basis of $\mathbf{M}$, where $\boldsymbol{t}$ is an arbitrary positive element of $\mathbb{I}$.

It is not difficult to see that the above definition is well posed, i.e. if $\boldsymbol{t}'\boldsymbol{u}, \boldsymbol{x}_1', \boldsymbol{x}_2', \boldsymbol{x}_3'$ is another positively oriented basis of $\mathbf{M}$ where $\boldsymbol{t}'$ is a positive element of $\mathbb{I}$, then $\boldsymbol{x}_1' + \mathbb{I}\boldsymbol{u}$, $\boldsymbol{x}_2' + \mathbb{I}\boldsymbol{u}$, $\boldsymbol{x}_3' + \mathbb{I}\boldsymbol{u}$ and $\boldsymbol{x}_1 + \mathbb{I}\boldsymbol{u}$, $\boldsymbol{x}_2 + \mathbb{I}\boldsymbol{u}$, $\boldsymbol{x}_3 + \mathbb{I}\boldsymbol{u}$ are equally oriented in $\mathbf{M}/\mathbb{I}\boldsymbol{u}$.

Introducing $x_0 := \boldsymbol{t}\boldsymbol{u}$, $x_0' := \boldsymbol{t}'\boldsymbol{u}$, we have that $x_i' = \sum_{k=0}^{3} A_{ki}x_k$ determines the 4×4 matrix $\{A_{ki} \mid i, k = 0, 1, 2, 3\}$ of changing from the 'unprimed' basis to the 'primed' basis of $\mathbf{M}$. Then the 3×3 matrix of changing the corresponding bases of $\mathbf{M}/\mathbb{I}\boldsymbol{u}$ is $\{A_{ki} \mid i, k = 1, 2, 3\}$. Since $A_{00} = \frac{\boldsymbol{t}'}{\boldsymbol{t}}$ and $A_{k0} = 0$ if $k = 1, 2, 3$, the determinant of the 4×4 matrix equals $\frac{\boldsymbol{t}'}{\boldsymbol{t}} > 0$ times the determinant of the 3×3 matrix, so the two determinants have the same sign.

2.6 Euclidean structures

To describe distances of space points first we have to introduce the mathematical model of distances:

A flat spacetime model must contain the **measure line of distances**, denoted by $\mathbb{D}$.

According to (**S4**) in 1.1 the space of an inertial observer is Euclidean. This means that for all $\boldsymbol{u} \in \mathtt{V}(1)$, there must be given a **Euclidean form** $\overline{\boldsymbol{d}}_{\boldsymbol{u}}$: $(\mathbf{M}/\mathbb{I}\boldsymbol{u}) \times (\mathbf{M}/\mathbb{I}\boldsymbol{u}) \to \mathbb{D} \otimes \mathbb{D}$, i.e. is a symmetric, bilinear, positive definite map which gives the length of vectors and the angle between space vectors in the usual way (see the mathematical supplement).

As we have mentioned, the space vectors of the inertial observer $\boldsymbol{u}$ cannot be illustrated in a fairly natural way, consequently, the Euclidean form on them is also a little complicated. Fortunately, we can replace it with a simpler, equivalent object.

The formula

$$\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y}) := \overline{\boldsymbol{d}}_{\boldsymbol{u}}(\boldsymbol{x} + \mathbb{I}\boldsymbol{u}, \boldsymbol{y} + \mathbb{I}\boldsymbol{u}) \tag{III4}$$

defines a symmetric, bilinear map $\mathbf{M} \times \mathbf{M} \to \mathbb{D} \otimes \mathbb{D}$ which is positive semidefinite and has $\mathbb{I}\boldsymbol{u}$ as its kernel, i.e.

– $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{x}) \geq \mathbf{0}$ and equality occurs if and only if $\boldsymbol{x}$ is parallel to $\boldsymbol{u}$,

– $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$ holds for all vectors $\boldsymbol{y}$ in $\mathbf{M}$ if and only if $\boldsymbol{x}$ is parallel to $\boldsymbol{u}$.

Conversely, if we give a symmetric, bilinear map

$$\boldsymbol{d}_{\boldsymbol{u}} : \mathbf{M} \times \mathbf{M} \to \mathbb{D} \otimes \mathbb{D}$$

which is positive semidefinite and has $\mathbb{I}\boldsymbol{u}$ as its kernel, then equality (III4) in the reverse sense (taking =: instead of :=) defines a well posed Euclidean form $\overline{\boldsymbol{d}}_{\boldsymbol{u}}$ on $\mathbf{M}/\mathbb{I}\boldsymbol{u}$. In other words:

$\overline{\boldsymbol{d}}_{\boldsymbol{u}}$ and $\boldsymbol{d}_{\boldsymbol{u}}$ determine each other uniquely.

Let the symmetric, bilinear map $\overline{\boldsymbol{d}}_{\boldsymbol{u}}$ be positive definite, i.e. $\overline{\boldsymbol{d}}_{\boldsymbol{u}}(\boldsymbol{x} + \mathbb{I}\boldsymbol{u}, \boldsymbol{x} + \mathbb{I}\boldsymbol{u}) \geq \mathbf{0}$ and equality holds if and only if $\boldsymbol{x} + \mathbb{I}\boldsymbol{u}$ is the zero vector of $\mathbf{M}/\mathbb{I}\boldsymbol{u}$, in other words if $\boldsymbol{x} + \mathbb{I}\boldsymbol{u} = \mathbb{I}\boldsymbol{u}$. Then $\boldsymbol{d}_{\boldsymbol{u}}$ defined by (III4) is symmetric, bilinear (because of the linearity of $\boldsymbol{x} \mapsto \boldsymbol{x} + \mathbb{I}\boldsymbol{u}$) and positive semidefinite. For all vectors $\boldsymbol{y}$ and $\boldsymbol{t} \in \mathbb{I}$ we have $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{t}\boldsymbol{u}, \boldsymbol{y}) = \overline{\boldsymbol{d}}_{\boldsymbol{u}}(\mathbb{I}\boldsymbol{u}, \boldsymbol{y} + \mathbb{I}\boldsymbol{u}) = \mathbf{0}$. Conversely, if $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$ for all vectors $\boldsymbol{y}$, then, in particular $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{x}) = \mathbf{0}$, i.e. $\overline{\boldsymbol{d}}_{\boldsymbol{u}}(\boldsymbol{x} + \mathbb{I}\boldsymbol{u}, \boldsymbol{x} + \mathbb{I}\boldsymbol{u}) = \mathbf{0}$, so $\boldsymbol{x} + \mathbb{I}\boldsymbol{u} = \mathbb{I}\boldsymbol{u}$, which means that $\boldsymbol{x}$ is parallel to $\boldsymbol{u}$. Consequently, the kernel of $\boldsymbol{d}_{\boldsymbol{u}}$ is $\mathbb{I}\boldsymbol{u}$.

Let the symmetric, bilinear map $\boldsymbol{d}_{\boldsymbol{u}}$ be positive semidefinite, having $\mathbb{I}\boldsymbol{u}$ as its kernel. First of all we have to prove that $\overline{\boldsymbol{d}}_{\boldsymbol{u}}$ is well defined by formula (III4), i.e. if $\boldsymbol{x}' + \mathbb{I}\boldsymbol{u} = \boldsymbol{x} + \mathbb{I}\boldsymbol{u}$ and $\boldsymbol{y}' + \mathbb{I}\boldsymbol{u} = \boldsymbol{y} + \mathbb{I}\boldsymbol{u}$, then $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}', \boldsymbol{y}') = \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y})$. This is true because $\boldsymbol{d}_{\boldsymbol{u}}$ is bilinear and has $\mathbb{I}\boldsymbol{u}$ as its kernel and there exist $\boldsymbol{t}, \boldsymbol{s} \in \mathbb{I}$ such that $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{t}\boldsymbol{u}$ and $\boldsymbol{y}' = \boldsymbol{y} + \boldsymbol{s}\boldsymbol{u}$. Then it is trivial that $\overline{\boldsymbol{d}}_{\boldsymbol{u}}$ is symmetric, bilinear and positive semidefinite. If $\overline{\boldsymbol{d}}_{\boldsymbol{u}}(\boldsymbol{x} + \mathbb{I}\boldsymbol{u}, \boldsymbol{x} + \mathbb{I}\boldsymbol{u}) = \mathbf{0}$, then $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{x}) = \mathbf{0}$ implying that $\boldsymbol{x}$ is parallel to $\boldsymbol{u}$, so $\boldsymbol{x} + \mathbb{I}\boldsymbol{u}$ equals $\mathbb{I}\boldsymbol{u}$ which is the zero vector of $\mathbf{M}/\mathbb{I}\boldsymbol{u}$; thus $\overline{\boldsymbol{d}}_{\boldsymbol{u}}$ is positive definite.

Thus, instead of $\overline{\boldsymbol{d}}_{\boldsymbol{u}}$, we shall consider $\boldsymbol{d}_{\boldsymbol{u}}$ and we require that it be smooth in $\boldsymbol{u}$ which means that the mapping $\boldsymbol{u} \mapsto \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y})$ should be smooth for all $\boldsymbol{x}, \boldsymbol{y} \in$

$\mathbf{M}$; this property expresses that distances in the space of inertial observers 'close to each other' are 'approximately the same'.

Summing up:

The structure of a flat spacetime model must contain for all $\boldsymbol{u} \in \mathtt{V}(1)$, a symmetric, bilinear, positive semidefinite map $\boldsymbol{d_u} : \mathbf{M} \times \mathbf{M} \to \mathbb{D} \otimes \mathbb{D}$ whose kernel is $\mathbb{I}\boldsymbol{u}$, and $\boldsymbol{d_u}$ depends smoothly on $\boldsymbol{u}$.

Let us repeat: the distance between the spacepoints $x + \mathbb{I}\boldsymbol{u}$ and $y + \mathbb{I}\boldsymbol{u}$ of the inertial observer $\boldsymbol{u}$ is $\sqrt{\boldsymbol{d_u}(x-y, x-y)}$; in other words, the distance of the $\boldsymbol{u}$-space points q and p (which are straight lines in $\mathtt{M}$ directed by $\boldsymbol{u}$) equals $\sqrt{\boldsymbol{d_u}(x-y, x-y)}$ where $x \in q$ and $y \in p$ are arbitrary.

Having defined the Euclidean structures in the space of inertial observers, summing 'small steps' (integrating), we can define distances in the space of an arbitrary observer. This can be done similarly (but in a more complicated way) as we defined proper time progress on arbitrary world lines from proper time progress on inertial world lines; details will not be given because we shall not use this notion.

2.7 Boosts

According to the principle (**E**) in 1.4, we have to give a relation expressing what it means that a vector in the space of an inertial observer equals the space vector of another inertial observer.

Such an equality, of course, must involve the equality of the lengths of the corresponding vectors and the equality of angles between vectors. In other words, for all absolute velocities $\boldsymbol{u}$ and $\boldsymbol{u}'$ we have to give a linear bijection from $\overline{\boldsymbol{B}}_{\boldsymbol{u}'\boldsymbol{u}} : \mathbf{M}/\mathbb{I}\boldsymbol{u} \to \mathbf{M}/\mathbb{I}\boldsymbol{u}'$ which preserves orientation and Euclidean structre.

To make our task easier, we require a linear bijection $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} : \mathbf{M} \to \mathbf{M}$ such that

$$\overline{\boldsymbol{B}}_{\boldsymbol{u}'\boldsymbol{u}}(\mathbf{x} + \mathbb{I}\boldsymbol{u}) = (\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \mathbf{x}) + \mathbb{I}\boldsymbol{u}'. \tag{III5}$$

Note that such a $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}$ exists but not uniquely; our requirement aims uniqueness.

In a flat spacetime model, for every absolute velocities $\boldsymbol{u}$ and $\boldsymbol{u}'$ there must be an orientation preserving linear bijection $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} : \mathbf{M} \to \mathbf{M}$, called the **boost from $\boldsymbol{u}$ to $\boldsymbol{u}'$** such that

$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{u} = \boldsymbol{u}',$$

$$\boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \mathbf{x}, \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{y}) = \boldsymbol{d}_{\boldsymbol{u}}(\mathbf{x}, \boldsymbol{y})$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}$; moreover $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}$ depends smoothly on $\boldsymbol{u}$ and $\boldsymbol{u}'$. The boost from $\boldsymbol{u}'$ to $\boldsymbol{u}$ is its inverse,

$$\boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'} = \boldsymbol{B}^{-1}_{\boldsymbol{u}'\boldsymbol{u}}.$$

The first condition above implies the existence of an orientation preserving linear bijection $\overline{\boldsymbol{B}}_{\boldsymbol{u}'\boldsymbol{u}} : \mathbf{M}/\mathbb{I}\boldsymbol{u} \to \mathbf{M}/\mathbb{I}\boldsymbol{u}'$ for which (III5) is satisfied. $(\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{x}) + \mathbb{I}\boldsymbol{u}'$ is the vector in the space of $\boldsymbol{u}'$ which is 'physically equal' to the vector $\mathbf{x} + \mathbb{I}\boldsymbol{u}$ in the space of $\boldsymbol{u}$.vektorral.

2.8 Again about the futurelike vectors

We have to prove that $\mathtt{T}^{\to}$ is convex and open. Knowing the space of inertial observers, we can do it. Unfortunately, here we must deal with the complicated space vectors.

2.8.1 Paths of motions

Let us take an inertial observer $\boldsymbol{u}$. A material point with world line C, in general, moves with respect to the observer; the **path of the motion** consists of the space points met by the material point:

$$\mathrm{C} + \mathbb{I}\boldsymbol{u} = \{x + \mathbb{I}\boldsymbol{u} \mid x \in \mathrm{C}\};$$

this is a curve (possibly a single point) in $\mathtt{E}_{\boldsymbol{u}}$.

In the sequel we consider only an inertial material point. Then there is a world point x_0 and an absolute velocity $\boldsymbol{u}'$ such that $\mathrm{C} = \{x_0 + \boldsymbol{s}\boldsymbol{u}' \mid \boldsymbol{s} \in \mathbb{I}\}$. The corresponding path in the $\boldsymbol{u}$-space is

$$\{x_0 + \boldsymbol{s}\boldsymbol{u}' + \mathbb{I}\boldsymbol{u} \mid \boldsymbol{s} \in \mathbb{I}\}.$$

This is a point if $\mathbf{u}' = \boldsymbol{u}$. If $\mathbf{u}' \neq \boldsymbol{u}$, then according to (III3), the $\boldsymbol{u}$-space vector between the points of the path corresponding to $\boldsymbol{s}+\boldsymbol{h}$ and $\boldsymbol{s}$ equals

$$(x_0 + (\boldsymbol{s}+\boldsymbol{h})\boldsymbol{u}' + \mathbb{I}\boldsymbol{u}) - (x_0 + \boldsymbol{s}\boldsymbol{u}' + \mathbb{I}\boldsymbol{u}) = \boldsymbol{h}\boldsymbol{u}' + \mathbb{I}\boldsymbol{u} = \boldsymbol{h}(\boldsymbol{u}' + \mathbb{R}\boldsymbol{u}).$$

It can be seen that the $\boldsymbol{u}$-space vector between arbitrary two points of the path is a multiple of the vector $\mathbf{u}' + \mathbb{R}\boldsymbol{u}$:

The path in the $\mathbf{u}$*-space of an inertial material point with absolute velocity* $\mathbf{u}' \neq \mathbf{u}$ *is a straight line with direction vector* $\mathbf{u}' + \mathbb{R}\mathbf{u}$*.*

2.8.2 Parallel paths

Let us consider two inertial world lines with absolute velocities $\boldsymbol{u}'$ and $\boldsymbol{u}''$, respectively. According to the previous result, their path in the $\boldsymbol{u}$-space are parallel if and only if $\mathbf{u}' + \mathbb{R}\boldsymbol{u}$ and $\boldsymbol{u}'' + \mathbb{R}\boldsymbol{u}$ are multiples of each other i.e. if there are real numbers α' and α'' in such a way that $\alpha'(\mathbf{u}' + \mathbb{R}\boldsymbol{u}) = \alpha''(\mathbf{u}'' + \mathbb{R}\boldsymbol{u})$; this is equivalent to the fact that there is a real number α in such a way that $\alpha'\mathbf{u}' - \alpha''\boldsymbol{u}'' = \alpha\mathbf{u}$.

The paths in the $\mathbf{u}$*-space of the inertial material points with absolute velocities* $\mathbf{u}'$ *and* $\mathbf{u}''$*, respectively, are parallel if and only if* $\mathbf{u}$*,* $\mathbf{u}'$ *and* $\mathbf{u}''$ *are in plane (in other words, are linearly dependent).*

It is worth emphasizing that it can happen that the paths of two material points are parallel in the space of an observer but are not parallel in the space of an other observer.

2.8.3 Faster-slower

Let us consider an inertial observer $\boldsymbol{u}$ and two inertial material points with absolute velocities $\boldsymbol{u}^-$ and $\boldsymbol{u}^+$, respectively.

Figure 2.13 shows that after the meeting of the two material points the one with absolute velocity $\boldsymbol{u}^-$ reaches another $\boldsymbol{u}$-space point later than the material point with absolute velocity $\boldsymbol{u}^+$.

Thus, we say that $\boldsymbol{u}^-$ is **slower** than $\boldsymbol{u}^+$ **with respect to** $\boldsymbol{u}$ if for all positive time period $\boldsymbol{t}^+$ there are positive time periods $\boldsymbol{t}, \boldsymbol{t}^-$ in such a way that

$$\boldsymbol{t}^+\boldsymbol{u}^+ + \boldsymbol{t}\boldsymbol{u} = \boldsymbol{t}^-\boldsymbol{u}^-.$$

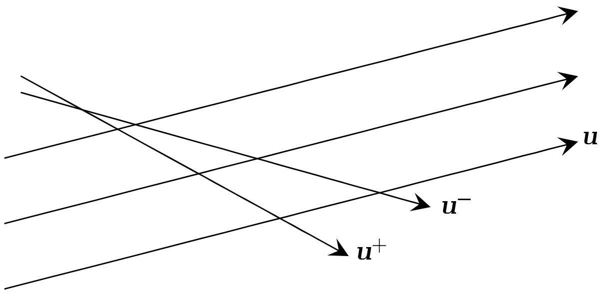

Figure 2.13 According to $\boldsymbol{u}$, $\boldsymbol{u}^-$ is slower than $\boldsymbol{u}^+$

Accordingly, $\boldsymbol{u}^+$ is **faster** than $\boldsymbol{u}^-$ **with respect to** $\boldsymbol{u}$ if for all positive time periods $\boldsymbol{t}^-$ there are positive time periods $\boldsymbol{t}, \boldsymbol{t}^+$ in such a way that

$$\boldsymbol{t}^-\boldsymbol{u}^- - \boldsymbol{t}\boldsymbol{u} = \boldsymbol{t}^+\boldsymbol{u}^+.$$

It is worth emphasizing: the notation + and − refers to faster-slower only with respect to the observer $\boldsymbol{u}$; the situation can be of other kind. Namely, if two inertial material points – called 1 and 2 – meet in a world point then there is an inertial observer

– for which the material point 1 is faster than the material point 2,

– for which the material point 2 is faster than the material point 1,

– which cannot be decide which of 1 and 2 is faster (because the paths are different in the observer space).

2.8.4 Convex and open set

According to (**M1**) in 1.3 a motion can be realized in every direction in $\boldsymbol{u}$-space which means that there are linearly independent vectors $\boldsymbol{h}_1$, $\boldsymbol{h}_2$ and $\boldsymbol{h}_3$ in $\mathtt{T}^{\to}$ in such a way that $\boldsymbol{h}_1 + \mathbb{I}\boldsymbol{u}$, $\boldsymbol{h}_2 + \mathbb{I}\boldsymbol{u}$ and $\boldsymbol{h}_3 + \mathbb{I}\boldsymbol{u}$ form a basis for the $\boldsymbol{u}$-space vectors. Then for arbitrary $\boldsymbol{t} \in \mathbb{I}^+$ the futurelike vectors $\boldsymbol{t}\boldsymbol{u}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3$ form a basis in $\mathbf{M}$. Consequently, every vector $\mathbf{x}$ of $\mathbf{M}$ can be represented as a sum $\mathbf{x} = \boldsymbol{k}_1 - \boldsymbol{k}_2$ where both $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ are in $\mathtt{T}^{\to}$. Indeed, if $\mathbf{x} \neq 0$ is not timelike then $\boldsymbol{k}_1$ is the sum of the basis elements with non-negative coefficients and $\boldsymbol{k}_2$ is the sum with negative coefficients. If $\boldsymbol{x}$ is futurelike then $\boldsymbol{k}_1 := 2\boldsymbol{x}$, $\boldsymbol{k}_2 := \boldsymbol{x}$ and if $\boldsymbol{x}$ is pastlike then $\boldsymbol{k}_1 := -\boldsymbol{x}$, $\boldsymbol{k}_2 := 2\boldsymbol{x}$.

According to (**M3**) in 1.3.2 and to the previous relation faster-slower for every absolute velocity $\boldsymbol{u}$ and futurelike vector $\boldsymbol{h}^+$ $\left(\boldsymbol{u}^+ := \frac{\boldsymbol{h}^+}{\boldsymbol{P}(\boldsymbol{h}^+)}\right)$ and for arbitrary positive element $\boldsymbol{t}^+$ of $\mathbb{I}$ there is a futurelike vector $\boldsymbol{h}^-$ $\left(\boldsymbol{u}^- := \frac{\boldsymbol{h}^-}{\boldsymbol{P}(\boldsymbol{h}^-)}\right)$ in such a way that $\boldsymbol{h}^+ + \boldsymbol{tu} = \boldsymbol{h}^-$.

Since every futurelike vector is of the form $\boldsymbol{tu}$ for some $\boldsymbol{t}$ and $\boldsymbol{u}$, we have the sum of two elements of$\mathtt{T}^{\rightarrow}$ is in $\mathtt{T}^{\rightarrow}$, too. Since the positive multiples of the elements of $\mathtt{T}^{\rightarrow}$ are in $\mathtt{T}^{\rightarrow}$, the set of futurelioke vectors is **convex cone with apex at zero** i.e. $\alpha\boldsymbol{h} + \beta\boldsymbol{k} \in \mathtt{T}^{\rightarrow}$ for all $\boldsymbol{h}, \boldsymbol{k} \in \mathtt{T}^{\rightarrow}$ and $\alpha, \beta > 0$.

According to (**M2**) in 1.3.2 and the previous relation faster-slower for every absolute velocity $\boldsymbol{u}$ and futurelike vector $\boldsymbol{h}^-$ $\left(\boldsymbol{u}^- := \frac{\boldsymbol{h}^-}{\boldsymbol{P}(\boldsymbol{h}^-)}\right)$ there are a positive element $\boldsymbol{t}^+$ of $\mathbb{I}$ and a futurelike vector $\boldsymbol{h}^+$ $\left(\boldsymbol{u}^+ := \frac{\boldsymbol{h}^+}{\boldsymbol{P}(\boldsymbol{h}^+)}\right)$ in such a way that $\boldsymbol{h}^- - \boldsymbol{tu} = \boldsymbol{h}^+$.

Since every futurelike vector is of the form $\boldsymbol{tu}$ for some $\boldsymbol{t}$ and $\boldsymbol{u}$, we have that for all $\boldsymbol{h}, \boldsymbol{k} \in \mathtt{T}^{\rightarrow}$ there is a $\beta > 0$ such that $\boldsymbol{h} - \beta\boldsymbol{k} \in \mathtt{T}^{\rightarrow}$. In view of the previous result, if $\eta < \beta$ then $\boldsymbol{h} - \eta\boldsymbol{k} = \boldsymbol{h} - \beta\boldsymbol{k} + (\beta - \eta)\boldsymbol{k}$ is in $\mathtt{T}^{\rightarrow}$. Therefore we can state that if $\boldsymbol{h}$is an arbitrary futurelike vector then for all $\boldsymbol{k} \in \mathtt{T}^{\rightarrow}$ there is number $\beta_{\boldsymbol{k}} > 0$ such that $\boldsymbol{h} - \eta\boldsymbol{k}$ is in $\mathtt{T}^{\rightarrow}$, too, for all $0 < \eta < \beta_{\boldsymbol{k}}$.

Let us consider an arbitrary non-zero vector $0 \neq \boldsymbol{x} = \boldsymbol{k}_1 - \boldsymbol{k}_2$ where $\boldsymbol{k}_1, \boldsymbol{k}_2 \in \mathtt{T}^{\rightarrow}$. Let $\boldsymbol{h}$ be an arbitrary futurelike vector. Let $\boldsymbol{k}_2$ be the previously obtained $\boldsymbol{k}$. Then $\boldsymbol{h} - \eta\boldsymbol{k}_2 \in \mathtt{T}^{\rightarrow}$ for all 'sufficiently small' positive η (depending on $\boldsymbol{k}_2$) $\mathtt{T}^{\rightarrow}$ is a convex cone, thus $\boldsymbol{h} + \eta\boldsymbol{k}_1 \in \mathtt{T}^{\rightarrow}$. At last, the half of the sum of the vectors in questions, i.e. $\boldsymbol{h} + \frac{\eta}{2}\boldsymbol{x}$ is futurelike for all 'sufficiently small' positive η (depending on $\boldsymbol{x}$). This means that $\mathtt{T}^{\rightarrow}$ contains a neighbourhood of its every element $\boldsymbol{h}$: $\mathtt{T}^{\rightarrow}$ is **open**.

2.9 Mathematical structure of flat spacetime models

2.9.1 Exact definition

Now we are in a position to compose an exact mathematical definition:

A **flat spacetime model** is $(\mathtt{M}, \mathbb{I}, \mathbb{D}, \mathtt{T}^{\to}, \boldsymbol{P}, \boldsymbol{d}, \boldsymbol{B})$ where

- $\mathtt{M}$ is **spacetime**, a four-dimensional oriented affine space over the vector space $\mathbf{M}$,
- $\mathbb{I}$ is the measure line of **time periods**,
- $\mathbb{D}$ is the measure line of **distances**,
- $\mathtt{T}^{\to}$ is the set of **futurelike vectors**, an open convex cone in $\mathbf{M}$ having the zero as its apex,
- $\boldsymbol{P} : \mathtt{T}^{\to} \to \mathbb{I}^{+}$ is the **proper time progress**, a smooth, positive homogeneous map,

and the latter two objects determine

$$\mathtt{V}(1) := \left\{ \boldsymbol{u} \in \frac{\mathbf{M}}{\mathbb{I}} \,\middle|\, \boldsymbol{u} \text{ is futurelike}, \quad \boldsymbol{P}(\boldsymbol{u}) = 1 \right\},$$

the set of absolute velocities,

- $\boldsymbol{d}$ is the collection of **Euclidean structures**, in other words, the expression for lengths and angles, which assings to every absolute velocity $\boldsymbol{u}$ a symmetric, bilinear, positive semidefinite map $\boldsymbol{d}_{\boldsymbol{u}} : \mathbf{M} \times \mathbf{M} \to \mathbb{D} \otimes \mathbb{D}$ whose kernel is $\mathbb{I}\boldsymbol{u}$, and this assignment is smooth with respect to $\boldsymbol{u}$,
- $\boldsymbol{B}$ is the collection of **boosts**, which assigns to every absolute velocity an orientation presesrving linear bijection $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} : \mathbf{M} \to \mathbf{M}$ such that
$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{u} = \boldsymbol{u}',$$
$$\boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{x}, \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{y}) = \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y})$$
for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}$ and
$$\boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'} = \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}^{-1},$$
and this assignment is smooth with respect to $\boldsymbol{u}$ and $\boldsymbol{u}'$.

In such a spacetime model

– the set of pastlike vectors and the set of timelike vectors are $\mathrm{T}^{\leftarrow} := -\mathtt{T}^{\to}$ and $\mathrm{T} := \mathrm{T}^{\leftarrow} \cup \mathtt{T}^{\to}$, respectively,

– a world line is a curve whose tangents are timelike,

– the proper time passed between the occurrences x and y on a world line C is

$$\boldsymbol{t}_{\mathrm{C}}(x,y) := \int_{p^{-1}(x)}^{p^{-1}(y)} \boldsymbol{P}(\dot{p}(a))da$$

where p is an arbitrary progressive parametrization of the world line,

– an inertial world line is a straight line, the time passed between the occurrences x and y of an inertial world line equals $\boldsymbol{P}(y-x)$,

– an observer is an infinitely differentiable velocity field $\boldsymbol{U} : \mathtt{M} \rightarrowtail \mathtt{V}(1)$ defined on a connected open subset,

– a space point of the observer $\boldsymbol{U}$ (briefly: a $\boldsymbol{U}$-space point) is a maximal integral curve of the vector field $\boldsymbol{U}$,

– an inertial observer is a constant velocity field, its space points are parallel straight lines; the space of the inertial observer $\boldsymbol{u}$ is $\mathtt{E}_{\boldsymbol{u}} := \mathtt{M}/\mathbb{I}\boldsymbol{u}$, a three-dimensional affine space over the vector space $\mathbf{M}/\mathbb{I}\boldsymbol{u}$,

– the inner product of the space vectors $\boldsymbol{x} + \mathbb{I}\boldsymbol{u}$ and $\boldsymbol{y} + \mathbb{I}\boldsymbol{u}$ of the inertial observer $\boldsymbol{u}$ is $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y})$; thus the distance between the space points $x + \mathbb{I}\boldsymbol{u}$ and $y + \mathbb{I}\boldsymbol{u}$ of the observer is $\sqrt{\boldsymbol{d}_{\boldsymbol{u}}(x-y, x-y)}$.

2.9.2 Isomorphisms

We can construct various spacetime models according to our experience. It may happen that two spacetime models, expressed in different mathematical formulae, have the same physical content. We expect that two models have the same physical content if and only if their mathematical structures are isomorphic as formulated below.

Let us consider the models $(\mathtt{M}, \mathbb{I}, \mathbb{D}, \mathtt{T}^{\rightarrow}, \boldsymbol{P}, \boldsymbol{d}, \boldsymbol{B})$ and $(\hat{\mathtt{M}}, \hat{\mathbb{I}}, \widehat{\mathbb{D}}, \hat{\mathtt{T}^{\rightarrow}}, \hat{\boldsymbol{P}}, \hat{\boldsymbol{d}}, \hat{\boldsymbol{B}})$. As concerns their mathematical structures, $\mathtt{M}$ and $\hat{\mathtt{M}}$, $\hat{\mathbb{I}}$ and $\hat{\mathbb{I}}$, as well as $\mathbb{D}$ and $\hat{\mathbb{D}}$ are isomorphic, i.e. there are (continuum many)

(i) an orientation preserving affine bijection $L : \mathtt{M} \to \hat{\mathtt{M}}$ (over the linear bijection $\boldsymbol{L} : \mathbf{M} \to \hat{\mathbf{M}}$),

(ii) an orientation preserving linear bijection $\boldsymbol{F} : \mathbb{I} \to \hat{\mathbb{I}}$,

(iii) an orientation preserving linear bijection $\boldsymbol{Z} : \mathbb{D} \to \hat{\mathbb{D}}$.

In other words, the first three components of every spacetime model are essentially the same; the difference appears in the last three components.

The two spacetime models are called **isomorphic** if there are L, $\boldsymbol{F}$ and $\boldsymbol{Z}$ which transform $\mathtt{T}^{\rightarrow}$ into $\widehat{\mathtt{T}^{\rightarrow}}$, $\boldsymbol{P}$ into $\hat{\boldsymbol{P}}$, $\boldsymbol{d}$ into $\hat{\boldsymbol{d}}$ and $\boldsymbol{B}$ into $\hat{\boldsymbol{B}}$ 'conveniently'.

The linear bijections regarding the measure lines can be given simply and transparently by choosing 'units' $\mathrm{s} \in \mathbb{I}$ and $\mathrm{m} \in \mathbb{D}$ as well as $\hat{\mathrm{s}} \in \hat{\mathbb{I}}$ and $\hat{\mathrm{m}} \in \hat{\mathbb{D}}$ and then

$$\boldsymbol{F}(\boldsymbol{t}) = \hat{\mathrm{s}}\frac{\boldsymbol{t}}{\mathrm{s}}, \qquad (\boldsymbol{t} \in \mathbb{I}),$$

$$\boldsymbol{Z}(\boldsymbol{h}) = \hat{\mathrm{m}}\frac{\boldsymbol{h}}{\mathrm{m}}, \qquad (\boldsymbol{h} \in \mathbb{D}).$$

Then we require that

(I) $\boldsymbol{L}[\mathtt{T}^{\rightarrow}] = \widehat{\mathtt{T}^{\rightarrow}}$,

(II) $\frac{\hat{\boldsymbol{P}}(\boldsymbol{L}\cdot\boldsymbol{x})}{\hat{\mathrm{s}}} = \frac{\boldsymbol{P}(\boldsymbol{x})}{\mathrm{s}}$ for all $\boldsymbol{x} \in \mathtt{T}^{\rightarrow}$.

Then $\boldsymbol{L}\cdot\boldsymbol{u} := \frac{\boldsymbol{L}(\mathrm{s}\boldsymbol{u})}{\hat{\mathrm{s}}}$ defines an affine bijection from $\mathtt{V}(1)$ to $\hat{\mathtt{V}(1)}$ which allows us to require

(III)

$$\frac{\hat{\boldsymbol{d}}_{\boldsymbol{L}\cdot\boldsymbol{u}}(\boldsymbol{L}\cdot\boldsymbol{x}, \boldsymbol{L}\cdot\boldsymbol{y})}{\hat{\mathrm{m}}} = \frac{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y})}{\mathrm{m}}$$

for all $\boldsymbol{u} \in \mathtt{V}(1)$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}$,

(IV)

$$\hat{\boldsymbol{B}}_{\boldsymbol{L}\cdot\boldsymbol{u}', \boldsymbol{L}\cdot\boldsymbol{u}}(\boldsymbol{L}\cdot\boldsymbol{x}) = \boldsymbol{L}\cdot\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}(\boldsymbol{x})$$

for all $\boldsymbol{u}', \boldsymbol{u} \in \mathtt{V}(1)$ and $\boldsymbol{x} \in \mathbf{M}$.

The triplet $(L, \boldsymbol{F}, \boldsymbol{Z})$ is called an **isomorphism** between the two spacetime models.

Note that two flat spacetime models cannot be isomorphic if their sets of futurelike vectors are of different types drawn in Figure 2.6 because a linear bijection cannot map the different cones onto each other.

2.9.3 Symmetries

An isomorphism of a spacetime model to itself is called an **automorphism**. A **symmetry** of the spacetime model is an automorphism if the involved linear maps $\mathbb{I} \to \mathbb{I}$ and $\mathbb{D} \to \mathbb{D}$ are the identity. Treating a symmetry, we omit these identities.

Accordingly, a symmetry of the spacetime model $(\mathtt{M}, \mathbb{I}, \mathbb{D}, \mathtt{T}^{\rightarrow}, \boldsymbol{P}, \boldsymbol{d}, \boldsymbol{B})$ is an orientation preserving affine bijection $\mathtt{M} \to \mathtt{M}$ and the underlying linear bijection $\boldsymbol{L} : \mathbf{M} \to \mathbf{M}$ – called **vectorial symmetry** – has the

(I) $\boldsymbol{L}[\mathtt{T}^{\rightarrow}] = \mathtt{T}^{\rightarrow}$

(II) $\boldsymbol{P}(\boldsymbol{L}\cdot\boldsymbol{x}) = \boldsymbol{P}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathtt{T}^{\rightarrow}$,

(III) $\boldsymbol{d}_{\boldsymbol{L}\cdot\boldsymbol{u}}(\boldsymbol{L}\cdot\boldsymbol{x}, \boldsymbol{L}\cdot\boldsymbol{y}) = \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{y})$ for all $\boldsymbol{u} \in \mathtt{V}(1)$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}$,

(IV) $\boldsymbol{B}_{\boldsymbol{L}\boldsymbol{u}',\boldsymbol{L}\boldsymbol{u}}(\boldsymbol{L}\cdot\mathbf{x}) = \boldsymbol{L}\cdot\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}(\boldsymbol{x},\boldsymbol{y})$ for all $\boldsymbol{u}', \boldsymbol{u} \in \mathtt{V}(1)$ and $\mathbf{x} \in \mathbf{M}$.

The symmetries have the following physical interpretations. We have an intuition when two phenomena, two procedures etc. can be considered the same from a physical point of view. This is expressed in the spacetime model that two objects are **physically equivalent** by definition if there is a symmetry that maps the objects into each other.

For instance, two world line C and C′ – i.e. the histories of two material points – physically equivalent if there is symmetry L such that $\mathrm{C}' = L[\mathrm{C}]$. In particular, the **translation** by an arbitrary vector $\boldsymbol{a}$, i.e. the map $x \mapsto x + \boldsymbol{a}$ is a symmetry, since the underlying linear map is the identity of $\mathbf{M}$. This gives an exact meaning to the homogeneity of spacetime mentioned in 2.3.2.

Further, the global (i.e. everywhere defined) observers $\boldsymbol{U}$ and $\boldsymbol{U}'$ are physically equivalent if there is spacetime symmetry L (over the vectorial symmetry $\boldsymbol{L}$) such that $\boldsymbol{U}'(x) = \boldsymbol{L}\cdot\boldsymbol{U}(L^{-1}(x))$ for all world points x.

Since we conceive that inertial observers are physically equivalent, according to what has been said in 1.4, we expect that **the boosts be vectorial symmetries**.

3 Other notions

In the sequel we consider a given spacetime model $(\mathtt{M}, \mathbb{I}, \mathbb{D}, \mathtt{T}^{\rightarrow}, \boldsymbol{P}, \boldsymbol{d}, \boldsymbol{B})$.

3.1 Synchronization

3.1.1 Simultaneity

We treated the problem of time points, the problem of synchronization in the Introduction as well as in 1.2.3. We can give a clear and exact meaning to these notions in the framework of a spacetime model.

An observer, in general, can define simultaneity for its space points by a 'natural' but an arbitrary procedure; **synchronization** means establishing simultaneity in a continuous manner.

Keep in mind that the same observer can, in general, establish many different synchronizations.

We accept as a fundamental property of synchronizations that **different occurrences of any world line** (different ticks of a chronometer) **cannot be simultaneous**.

As a matter of course, a synchronization – establishing simultaneity continuously – determines **time points (instants)**: simultaneous occurrences happen in the same instant.

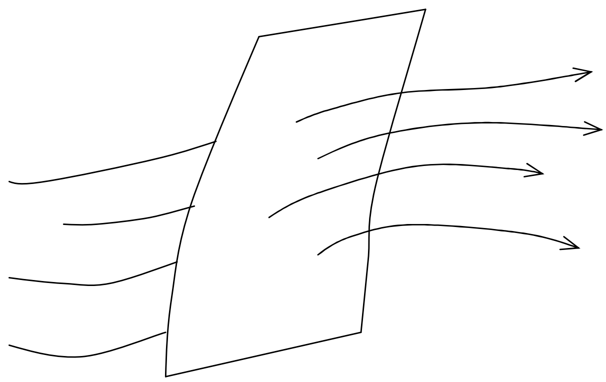

Figure 3.1 Synchronization instants

Now a very important idea follows. Take the space of an observer according to Figure 2.10, and let us imagine it being a little thicker, i.e. let us imagine space points (world lines) in front of and behind the plane of the page, too; then the world lines corresponding to the space points of the observer form a sheaf.

Let us mark the occurrences in each space point that are simultaneous according to a synchronization(say, for the sake of clarity, let us mark midnight in each space point). These simultaneous occurrences form a cross-section of the sheaf as it is seen in Figure 3.1. Such a cross-section is a three-dimensional submanifold in spacetime, called **world surface** (now, when sketching the scheme of the model we do not bother about its exact mathematical meaning).

This motivates us that it is convenient to consider an instant of a synchronization as the set of the corresponding simultaneous world points. For instance, midnight as an instant is the set of midnights in Greenwich, Budapest, Vienna, Prague etc.

We have to get accustomed to the following fact: **an instant (e.g. midnight) established by a synchronization is a world surface in spacetime**. Thinking about it a while, we do not find this so peculiar: midnight exists here, there, everywhere.

A **synchronization instant** is a world surface. The **time of a synchronization** is the collection of its instants.

Returning to the illustration in the plane of the page, a world surface will be drawn by a curve as it is seen in Figure 3.2

(time points)

(space points)

Figure 3.2 Space points of an observer and time points of a synchronization

Let us repeat for fixing in mind:

An instant of a synchronization is a subset of spacetime: a three-dimensional submanifold, a world surface.

The collection of these world surfaces is the time of the synchronization. Thus **the synchronization time is not a subset of spacetime**; it is a set whose elements are subsets of spacetime. The synchronization time becomes a one-dimensional manifold in a suitable sense (we do not yet bother about its exact mathematical meaning now); this is not a submanifold – because is not a subset – of spacetime.

Last but not least we emphasize: a synchronization is important only from a practical point of view, it is a human convention not a physical reality, that is why **a synchronization is not be a part of the structure of spacetime**; it is defined by the structure of the model.

3.1.2 Uniform synchronization

The mathematical formulation of a synchronization, in general, is somewhat complicated. Exceptions are the uniform synchronizations in which the time points are parallel three-dimensional hyperplanes.

Now we examine the uniform synchronizations form a purely mathematical point of view, not concerning the physical procedures which could establish such a synchronization.

Let $\mathbf{E}_s$ be a three dimensional linear subspace of $\mathbf{M}$. According to the **uniform synchronizationdetermined by $\mathbf{E}_s$** the world points x and y are

simultaneous if $y - x \in \mathbf{E}_s$.

The collection of world points simultaneous with the world point x according to the synchronization by $\mathbf{E}_s$ is $x + \mathbf{E}_s$. This means that a collection of simultaneous world point – a time point of the synchronization– is a hyperplane directed by $\mathbf{E}_s$. The synchronization time is the set of such time points which equals $\mathtt{M}/\mathbf{E}_s$.

According to the basic properties of synchronizations, no different occurrences of a world line can be simultaneous. Let us consider an inertial world line with absolute velocity $\boldsymbol{u}$. Its two different occurrences are not simultaneous with respect to $\mathbf{E}_s$ if and only if $\mathbb{I}\boldsymbol{u}$ and $\mathbf{E}_s$ meet in a sigle point (in the zero), in other words, if and only if $\boldsymbol{u}$ and $\mathbf{E}_s$ are transversal.

This means that $\mathbf{E}_s$ must be transversal to every futurelike vector. This is possible because $\mathtt{T}^{\rightarrow}$ is convex and open.

3.2 Reference frames

3.2.1 Splitting of spacetime

An observer and a synchronization together is called a **reference frame**.

A reference frame assigns to a world point the corresponding synchronization instant (world surface) and observer space point (world line) that contain the world point in question. We say that the reference frame **splits** spacetime into time and space.

This splitting has the physical meaning that the reference frame characterizes occurrences by 'when' and 'where' they happen. In other words, splitting of spacetime in the model corresponds to the fact that an observer, having introduced a synchronization, perceives spacetime as time and space separately.

Note that the splitting due to different reference frames can be totally different, i.e. different reference frames can assign totally different time points and space points to the same world point.

3.2.2 Description of motions

We describe the motion of a body by giving when it is where; thus, the description requires a synchronization (when: which time point) and an observer (where: which space point) i.e. a reference frame.

Let us consider a world line C representing the existence of a material point. The corresponding motion, with respect to an observer and a synchronization is a function $t \mapsto q(t)$ where t is a time point and $q(t)$ is a space point. This

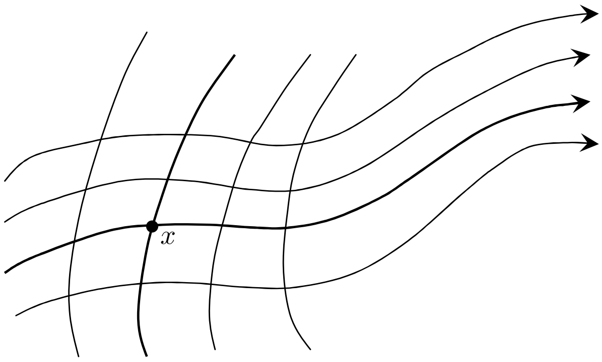

Figure 3.3 Splitting of spacetime to time and space

function is obtained as follows. The occurrence of the material point at the time point t (a world surface) of the synchronization is $t \cap \mathrm{C}$; this world point determines the space point $q(t)$ (a world line) of the observer for which $t \cap \mathrm{C} \in q(t)$ holds.

3.2.3 Inertial frames

A reference frame consisting of an inertial observer and a uniform synchronizationis called an **inertial frame**.

Thus an inertial frame is a pair $(\boldsymbol{u}, \mathbf{E}_{\mathrm{s}})$ where $\boldsymbol{u}$ is an inertial observer (in fact, an absolute velocity) and $\mathbf{E}_{\mathrm{s}}$ is a uniform synchronization (in fact, a three-dimensional linear subspace of $\mathbf{M}$ transverse to all absolute velocities).

Inertial frames have the following important property:

The same proper time elapses in every space point of an inertial observer between two time point of a uniform synchronization.

Let us consider two space points of the inertial observer $\boldsymbol{u}$, $x + \mathbb{I}\boldsymbol{u}$- and $y + \mathbb{I}\boldsymbol{u}$, supposing that x and y are simultaneous according to the uniform synchronization$\mathbf{E}_{\mathrm{s}}$ i.e $y - x \in \mathbf{E}_{\mathrm{s}}$. Then $x + \boldsymbol{t}\boldsymbol{u}$ and $y + \boldsymbol{s}\boldsymbol{u}$ are simultaneous i.e. $(y + \boldsymbol{s}\boldsymbol{u}) - (x + \boldsymbol{t}\boldsymbol{u}) \in \mathbf{E}_{\mathrm{s}}$ if and only if $\boldsymbol{s} = \boldsymbol{t}$.

Then it follows:

In an inertial frame, an inertial world line results in a uniform motion on straight line.

The proof is somewhat complicated, later, in actual spacetime models we get this result in a simpler way.

Let $\boldsymbol{u}$ and $\mathbf{E}_s$ be the observer and the synchronization, respectively. Let C be an inertial world line with absolute velocity $\boldsymbol{u}'$.

For two synchronization time points t and s put $x := t \cap \mathrm{C}$ and $y := s \cap \mathrm{C}$. The material point having the world line C meets the $\boldsymbol{u}$-space points $x + \mathbb{I}\boldsymbol{u}$ at t and $y + \mathbb{I}\boldsymbol{u}$ at s.

In these space points of the observer the same proper time $\boldsymbol{t}$ elapses. Since there is a $\boldsymbol{t}' \in \mathbb{I}$ such that $y - x = \boldsymbol{t}'\boldsymbol{u}'$, $\boldsymbol{t}$ is determined by the relation $(x + \boldsymbol{t}\boldsymbol{u}) - y = (x - y) + \boldsymbol{t}\boldsymbol{u} = \boldsymbol{t}\boldsymbol{u} - \boldsymbol{t}'\boldsymbol{u}' \in \mathbf{E}_s$. $\boldsymbol{u}'$ and $\mathbf{E}_s$ are transversal, thus $\boldsymbol{t}\boldsymbol{u}$ can be given uniquely as a sum of a vector in $\mathbb{I}\boldsymbol{u}'$ and a vector in $\mathbf{E}_s$. This means that the relation in question gives $\boldsymbol{t}'$ as a function of $\boldsymbol{t}$; let us denote this function by $\boldsymbol{t}'(\boldsymbol{t})$. It is evident that the double, the triple etc. of $\boldsymbol{t}$ results in the double, tiple etc. of $\boldsymbol{t}'$ i.e. the function is linear. Thus, there is a number α such that $\boldsymbol{t}'(\boldsymbol{t}) = \alpha\boldsymbol{t}$.

The $\boldsymbol{u}$-space vector between the space points in question is $(y + \mathbb{I}\boldsymbol{u}) - (x + \mathbb{I}\boldsymbol{u}) = (y - x) + \mathbb{I}\boldsymbol{u} = \boldsymbol{t}'\boldsymbol{u}' + \mathbb{I}\boldsymbol{u} = \alpha\boldsymbol{t}\boldsymbol{u}' + \mathbb{I}\boldsymbol{u}$.

We have obtained that the motion of the inertial material point is uniform on a straight line: the material point changes its $\boldsymbol{u}$-space position by the vector $(\alpha\boldsymbol{u}' + \mathbb{R}\boldsymbol{u})\boldsymbol{t}$ within the $\mathbf{E}_s$-time period $\boldsymbol{t}$.

3.3 Coordinate systems

It is worth repeating: the space of an observer and the proper time progressing in each space point of the observer are physical reality but a synchronization– though determined by a physical procedure in practice – is a human convention.

Thus, physical reality and human convention are mixed in a reference frame.

A further convention leads us to a **coordinate system** in which time points are represented by numbers and space points are represented by triplets of numbers, thus world points are represented by quartets of numbers.

In the simplest case such quartets are obtained as follows (considering an inertial reference frame).

One chooses an ‘origin’ in the observer space (Greenwich; the corner of our room) and an ‘origin’ in the synchronization time (midnight). Then one chooses a time unit (second: a given number of oscillations of a quartz crystal) and a distance unit (meter: the length of a straight chain of a given number of molecules in a quartz crystal). Furthermore, one chooses three orthogonal straight lines (axes) in the observer space passing through the origin.

A synchronization instant is represented by the real number ξ^0 if ξ^0 second time passed in Greenwich from midnight till the instant in question.

An observer space point is represented by the triplet (ξ^1, ξ^2, ξ^3) if the distances of the space point from the coordinate planes determined by the axes are ξ^1 meters, ξ^2 meters and ξ^3 meters.

Note that many different coordinate systems can be given to the same reference frame; different reference frames cannot define the same coordinate system.

Different coordinate systems assign totally different quartets to the same world point.

All this will be formulated later, in actual spacetime models, in a mathematically precise way.

We see, how many arbitrary and occasional data are involved in a coordinate system, consequently, how many unnecessary complications can arise in the usual treatment of spacetime with coordinates.

IV Absolute time

4 Basic notions and assumptions

In this section we take a flat spacetime model and, using the notions and notations of the previous chapter, we examine the special properties of $\mathtt{T}^{\rightarrow}$, $\boldsymbol{P}$, $\boldsymbol{d}$ and $\boldsymbol{B}$ that arise from the assumption of absolute time progress.

As a result, we obtain the so called nonrelativistic spacetime model which corresponds to the everyday thinking. Previously we mentioned that our simple, everyday experience indicates that the chronometer left at home and the one that accompanied me for my holidays have ticked the same number of times between their two meetings. The classical idea of spacetime – tacitly – is based on the assumption that time progresses in the same way for everybody: here are two (perfect) chronometers together, then they are separated and if they meet again, then one has ticked just as many as the other during their separation, independently of what they have undergone.

Now we construct a flat spacetime model that reflects our (superficial) experience:

(**I1**) **time progress is absolute** i.e. if two world line meet in two occurrences then the same the same proper time passes on both world lines between those two occurrences.

moreover assumes another of our seemingly natural (but superficial) experience:

(**I2**) In the space of an inertial observer, on every path **arbitrarily fast motion can be realized**.

This property is stronger than (**M2**) in 1.3.2. Let us consider a path between two space points of an inertial observer, called start and goal. Let $\boldsymbol{s}$ be the proper time period of a body from the start till the goal. Then for arbitrary time period $\boldsymbol{t} < \boldsymbol{s}$ there is another motion starting together with the previous one and arrives at the goal $\boldsymbol{t}$ time earlier.

5 Construction of the model

Who is interested only in the nonrelativistic spacetime model and its physical application and does not want to know the way that leads to it can skip this section without any harm in understanding the model and its applications.

5.1 Absolute time progress

Property (**I1**) can be formalized as follows: if x and y are world points such that $y - x$ is futurelike, then for arbitrary world lines C_1 and C_2 such that $x, y \in \mathrm{C}_1 \cap \mathrm{C}_2$, $\boldsymbol{t}_{\mathrm{C}_1}(x,y) = \boldsymbol{t}_{\mathrm{C}_2}(x,y)$ holds.

In particular, if C_1 is the straight line (inertial world line) passing through x and y, and C_2 consists of two straight line segments, one of which goes from x to z and the other goes from z to y, then we obtain that $\boldsymbol{P}(y-x) = \boldsymbol{P}(y-z) + \boldsymbol{P}(z-x)$. Since both $y-z$ and $z-x$ are futurelike, and $y-x = (y-z)+(z-x)$, we infer – taking into account that $\boldsymbol{P}$ is positive homogeneous – that

$$\boldsymbol{P}(\alpha\mathbf{x} + \beta\boldsymbol{y}) = \alpha\boldsymbol{P}(\mathbf{x}) + \beta\boldsymbol{P}(\boldsymbol{y})$$

for all positive real numbers α, β and futurelike vectors $\boldsymbol{x}, \boldsymbol{y}$. Then it follows:

There is a unique linear map $\boldsymbol{\tau} : \mathbf{M} \to \mathbb{I}$ whose restriction onto $\mathtt{T}^{\to}$ equals $\boldsymbol{P}$.

The statement means that $\boldsymbol{P}$ can be uniquely extended to a liner map on $\mathbf{M}$. According to 2.8.4 every element of of $\mathbf{M}$ is of the form $\boldsymbol{y} - \mathbf{x}$where $\mathbf{x}$ and $\boldsymbol{y}$ are in $\mathtt{T}^{\to}$. Let us define the linear map $\boldsymbol{\tau}$ by $\boldsymbol{\tau}(\boldsymbol{y} - \mathbf{x}) := \boldsymbol{P}(\boldsymbol{y}) - \boldsymbol{P}(\mathbf{x})$. Of course, it must be proved that this definition is correct i.e. if $\boldsymbol{y} - \mathbf{x} = \boldsymbol{y}' - \mathbf{x}'$ then $\boldsymbol{P}(\boldsymbol{y}) - \boldsymbol{P}(\mathbf{x}) = \boldsymbol{P}(\boldsymbol{y}') - \boldsymbol{P}(\mathbf{x}')$. This holds because then $\boldsymbol{y} + \mathbf{x}' = \boldsymbol{y}' + \mathbf{x}$ and here both sides are elements of $\mathtt{T}^{\to}$, so they are in the domain of $\boldsymbol{P}$ and applying its above special property of we get $\boldsymbol{P}(\boldsymbol{y}) + \boldsymbol{P}(\mathbf{x}') = \boldsymbol{P}(\boldsymbol{y}') + \boldsymbol{P}(\mathbf{x})$ which involves the necessary equality. The $\boldsymbol{\tau}$ is evidently linear by the positive homogeneity of $\boldsymbol{P}$.

In the sequel we write $\boldsymbol{\tau}$ instead of $\boldsymbol{P}$. Note that according to the general formula then for an absolute velocity $\boldsymbol{u}$ we have $\boldsymbol{\tau} \cdot \boldsymbol{u} = 1$.

For the sake of precision, we mention that here we have deviated from the original definition of world lines; namely, a subset consisting of two straight line segments is not a curve in our sense because it does not have a continuously differentiable parametrization. We could define a curve as a subset having a piecewise continuously differentiable parametrization and then there would be no trouble with the arguments above, but this more general notion is not necessary elsewhere. Therefore, we have decided to leave out the exact definition of this more general term, but here we have admitted the simple and clear notion of a curve consisting of two straight line segments.

Absolute time progress involves absolute simultaneity. This can be formulated as follows. Let us take an arbitrary world point z; then $x \in z + \mathtt{T}^{\rightarrow}$ and $y \in z + \mathtt{T}^{\rightarrow}$ are simultaneous 'in a natural way' according to z if the same times pass from z to x and to y i.e. $\boldsymbol{\tau} \cdot (x - z) = \boldsymbol{\tau} \cdot (y - z)$. Thus, $\boldsymbol{\tau} \cdot (y - x) = \boldsymbol{\tau} \cdot ((y - z) + (z - x)) = 0$; we see that z drops from the simultaneity of x and y, so we can state that x and y are **absolute simultaneous** if $\boldsymbol{\tau} \cdot (y - x) = 0$.

Let us introduce the notation

$$\mathbf{E} := \{\boldsymbol{q} \in \mathbf{M} \mid \boldsymbol{\tau} \cdot \boldsymbol{q} = 0\}.$$

This is the kernel of the linear map $\boldsymbol{\tau}$; since the range of $\boldsymbol{\tau}$ is one-dimensional, $\mathbf{E}$ is a three-dimensional linear subspace of $\mathbf{M}$.

Thus the occurrences x and y are absolute simultaneous if and only if $y - x \in \mathbf{E}$.

5.2 Futurelike vectors

Property (**I2**) can be now formalize as follows. Let s take the space points $x + \mathbb{I}\boldsymbol{u}$ and $z + \mathbb{I}\boldsymbol{u}$ of the inertial observer $\boldsymbol{u}$ – a start and a goal – and let a motion start at the occurrence and let it arrive at the occurrence x; then the proper time period od its motions equals $\boldsymbol{s} := \boldsymbol{\tau} \cdot (z - x)$. If $0 < \boldsymbol{t} < \boldsymbol{s}$ then there is another motion starting at x and arrives at $y := z - \boldsymbol{t}\boldsymbol{u}$. Consequently, $y - x$ is futurelike. The proper time period of this motion is $\boldsymbol{\tau} \cdot (y - x) = \boldsymbol{\tau} \cdot (z - \boldsymbol{t}\boldsymbol{u} - x) = \boldsymbol{\tau} \cdot (z - x) - \boldsymbol{t} = \boldsymbol{s} - \boldsymbol{t}$. Since $\boldsymbol{t}$ can be arbitrarily close to $\boldsymbol{s}$, this means that y and x cannot be on a world line only if $\boldsymbol{\tau} \cdot (y - x) = 0$ i.e. y and x are not absolutely simultaneous.

Thus, regarding vectors, we have that if $\boldsymbol{\tau} \cdot \mathbf{x} \neq 0$ then $\mathbf{x}$ or $-\mathbf{x}$ is futurelike. As a consequence, the set of futurelike vectors is a 'half space' in $\mathbf{M}$ having $\mathbf{E}$ as its boundary:

$$\mathtt{T}^{\rightarrow} = \{\mathbf{x} \in \mathbf{M} \mid \boldsymbol{\tau} \cdot \mathbf{x} > 0\}. \qquad \text{(IV1)}$$

Such a $\mathtt{T}^{\rightarrow}$ is illustrated by the first picture in Figure 2.6.

According to the definition of absolute velocities (see Paragraph 2.4.3) and the absolute time progress, now $\boldsymbol{u}$ is an absolute velocity if and only if $\boldsymbol{\tau} \cdot \boldsymbol{u} = 1$. Therefore,

$$\mathtt{V}(1) = \left\{ \boldsymbol{u} \in \frac{\mathbf{M}}{\mathbb{I}} \middle| \boldsymbol{\tau} \cdot \boldsymbol{u} = 1 \right\} \qquad \text{(IV2)}$$

which is an affine hyperplane over $\frac{\mathbf{E}}{\mathbb{I}}$.

5.3 Boosts

We have to define what it means that a vector (a straight line) in an observer space is equal (parallel) to a vector (straight line) in the space of an other observer, i.e. we must give the **boost** from $\boldsymbol{u}$ top $\boldsymbol{u}'$ (see 2.7).

Such a relation can be established in a natural way with the aid of absolute simultaneity. Let us consider a vector: rod with tip in the space of an inertial observer. The corresponding vector in the space of an other inertial observe is obtained by an **instantaneous print** i.e. the space points are marked that meet the points of the rod at the given instant.

Let q and p be the end points of the rod in the space of the observer $\boldsymbol{u}$. Let x and y are simultaneous occurrences of q and p, respectively, i.e. $\boldsymbol{\tau} \cdot (y - x) = 0$. The space points in question are straight lines in spacetime, $q = x + \mathbb{I}\boldsymbol{u}$ and $p = y + \mathbb{I}\boldsymbol{u}$. Their print in the $\boldsymbol{u}'$-space at the instant determined by x and y are $q' = x + \mathbb{I}\boldsymbol{u}'$ and $p' = y + \mathbb{I}\boldsymbol{u}'$.

Thus, the print in the $\boldsymbol{u}'$-space is determined by the same $y - x \in \mathbf{E}$ which determines the vector in the $\boldsymbol{u}$-space. Then it is evident that a boost maps the elements of $\mathbf{E}$ into themselves:

The boost form $\boldsymbol{u}$ to $\boldsymbol{u}'$ is determined by

$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{u} = \boldsymbol{u}', \qquad \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{q} = \boldsymbol{q} \qquad (\boldsymbol{q} \in \mathbf{E}).$$

It is trivial then that the condition $\boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'} = \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}^{-1}$ is fulfilled.

5.4 The absolute Euclidean structure

For every absolute velocity $\boldsymbol{u}$ there must be given a $\boldsymbol{d}_{\boldsymbol{u}} : \mathbf{M} \times \mathbf{M} \to \mathbb{D} \otimes \mathbb{D}$ bilinear, symmetric, positive semidefinite map whose kernel is $\mathbb{I}\boldsymbol{u}$. Since $\mathbf{E}$ is transversal to $\mathbb{I}\boldsymbol{u}$-ra, the restriction of $\boldsymbol{d}_{\boldsymbol{u}}$ to $\mathbf{E}$ is positive definite.

According to the condition imposed on boosts,

$$\boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}\cdot\boldsymbol{x},\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}\cdot\boldsymbol{y})=\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x},\boldsymbol{y})$$

must be satisfied for all vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. Then, because the restriction of boosts to $\mathbf{E}$ is the identity, it follows that the restriction of $\boldsymbol{d}_{\boldsymbol{u}'}$ and $\boldsymbol{d}_{\boldsymbol{u}}$ to $\mathbf{E}$ are equal; this involves:

> There is a uniquely determined
>
> $$\boldsymbol{b}:\mathbf{E}\times\mathbf{E}\to\mathbb{D}\otimes\mathbb{D} \tag{IV3}$$
>
> bilinear, symmetric, positive definite map, the **absolute Euclidean structure** such that the restriction of $\boldsymbol{d}_{\boldsymbol{u}}$ to $\mathbf{E}$ equals $\boldsymbol{b}$ for all $\boldsymbol{u}$.

For arbitrary spacetime vector $\boldsymbol{x}$ we have that $\boldsymbol{x}-\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{x})$ is in $\mathbf{E}$ (and similar relation holds for $\boldsymbol{y}$) and the kernel of $\boldsymbol{d}_{\boldsymbol{u}}$ is $\mathbb{I}\boldsymbol{u}$, thus $\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x}-\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{x}),\boldsymbol{y}-\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{y}))$; as a final result we can state that

$$\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{b}\big(\boldsymbol{x}-\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{x}),\boldsymbol{y}-\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{y})\big)\qquad(\boldsymbol{u}\in\mathtt{V}(1),\ \boldsymbol{x},\boldsymbol{y}\in\mathbf{M}). \tag{IV4}$$

6 Nonrelativistic spacetime model

The spacetime model introduced previously is called **nonrelativistic**; we have got it, besides the general assumptions, with the ones formulated in (**I1**) and (**I2**),

The nonrelativistic spatime model reflects our simple ideas about time and space and this spacetime model – not in an exact setting – form the background of classical mechanics, so it is deeply imbedded in physical considerations.

In this section, not using the previous one, i.e not referring to how we have arrived at the spacetime model in question, we define it and examine its properties. As a consequence, some formulae of the previous section appear again as a novelty.

6.1 Basic properties of the model

6.1.1 New notation of the model

Instead of the general notation $(\mathtt{M},\mathbb{I},\mathbb{D},\mathrm{T}^{\to},\boldsymbol{P},\boldsymbol{d},\boldsymbol{B})$, a nonrelativistic spacetime model will be referred to by the symbol $(\mathtt{M},\mathbb{I},\mathbb{D},\boldsymbol{\tau},\boldsymbol{b})$ where

– $\mathtt{M}$ is **spacetime**, a four-dimensional oriented affine space (over the vector space $\mathbf{M}$),
– $\mathbb{I}$ is the measure line of **time durations**,
– $\mathbb{D}$ is the measure line of **distances**,
– $\boldsymbol{\tau} : \mathbf{M} \to \mathbb{I}$ is a linear surjection, the **absolute time progress**, whose kernel is $\mathbf{E} := \{\boldsymbol{q} \in \mathbf{M} \mid \boldsymbol{\tau} \cdot \boldsymbol{q} = 0\}$, a three-dimensional liner subspace of $\mathbf{M}$,
– $\boldsymbol{b} : \mathbf{E} \times \mathbf{E} \to \mathbb{D} \otimes \mathbb{D}$ is a symmetric, positive definite bilinear map, the **absolute Euclidean form**
which give
– the set of futurelike vectors,

$$\mathrm{T}^{\rightarrow} := \{\boldsymbol{x} \in \mathbf{M} \mid \boldsymbol{\tau} \cdot \boldsymbol{x} > 0\},$$

which, being a half space, is indeed an open convex cone
– the inertial time progress $\boldsymbol{P}$ as the restriction of $\boldsymbol{\tau}$ onto $\mathrm{T}^{\rightarrow}$,
thus, the set of absolute velocities is

$$\mathtt{V}(1) := \left\{ \boldsymbol{u} \in \frac{\mathbf{M}}{\mathbb{I}} \;\middle|\; \boldsymbol{\tau} \cdot \boldsymbol{u} = 1 \right\},$$

– the Euclidean structure of inertial observers as

$$\boldsymbol{d_u}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{b}\big(\boldsymbol{x} - \boldsymbol{u}(\boldsymbol{\tau} \cdot \boldsymbol{x}), \boldsymbol{y} - \boldsymbol{u}(\boldsymbol{\tau} \cdot \boldsymbol{y})\big) \qquad (\boldsymbol{u} \in \mathtt{V}(1),\ \boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}). \quad \text{(IV5)}$$

which depends on $\boldsymbol{u}$ smoothly and is positive semidefinite having the kernel $\mathbb{I}\boldsymbol{u}$,
– the boosts by the formula

$$\boldsymbol{B}_{\boldsymbol{u'u}} = \mathbf{1} + (\boldsymbol{u}' - \boldsymbol{u}) \otimes \boldsymbol{\tau} \qquad (\boldsymbol{u}', \boldsymbol{u} \in \mathtt{V}(1))$$

which depends on $\boldsymbol{u}$ and $\boldsymbol{u}'$ and satisfies the required conditions
(i) $\boldsymbol{B}_{\boldsymbol{u'u}} \cdot \boldsymbol{u} = \boldsymbol{u}'$ (it is evident),
(ii) $\boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{B}_{\boldsymbol{u'u}} \cdot \boldsymbol{x}, \boldsymbol{B}_{\boldsymbol{u'u}} \cdot \boldsymbol{y}) = \boldsymbol{d_u}(\boldsymbol{x}, \boldsymbol{y})$, because $\boldsymbol{\tau} \cdot \boldsymbol{B}_{\boldsymbol{u'u}} = \boldsymbol{\tau}$ and so $\boldsymbol{B}_{\boldsymbol{u'u}} \cdot \boldsymbol{x} - \boldsymbol{u}'\boldsymbol{\tau} \cdot (\boldsymbol{B}_{\boldsymbol{u'u}} \cdot \boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{u}(\boldsymbol{\tau} \cdot \boldsymbol{x})$; as a consequence both the left hand side and the right hand side equal $\boldsymbol{b}(\boldsymbol{x} - \boldsymbol{u}(\boldsymbol{\tau} \cdot \boldsymbol{x}), \boldsymbol{x} - \boldsymbol{u}(\boldsymbol{\tau} \cdot \boldsymbol{x}))$,
(iii) $\boldsymbol{B}_{\boldsymbol{uu'}} = \boldsymbol{B}^{-1}_{\boldsymbol{u'u}}$ (it is evident).
Further, the sets of pastlike and timelike vectors (see 2.3.2) are

$$\mathrm{T}^{\leftarrow} = -\mathtt{T}^{\rightarrow}, \qquad \mathrm{T} = \mathrm{T}^{\leftarrow} \cup \mathtt{T}^{\rightarrow},$$

respectively.

That is why we call the elements of **E absolute spacelike vectors**. which is a three-dimensional linear subspace in **M**.

It is worthy to note that

$$\boldsymbol{d_u}(\boldsymbol{q},\boldsymbol{p}) = \boldsymbol{b}(\boldsymbol{q},\boldsymbol{p})$$

for all absolute velocities $\boldsymbol{u}$ and absolute spacelike vectors $\boldsymbol{q},\boldsymbol{p}$.

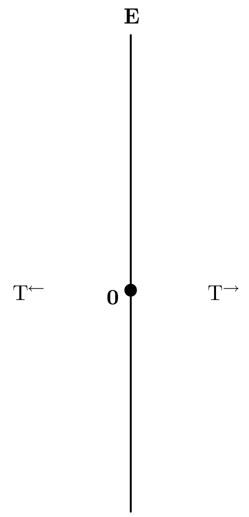

Figure 6.1 Spacetime vectors

The set of spacetime vectors is illustrated in the plane of the page according to Figure 6.1.

The set of absolute velocities is a three-dimensional affine subspace over $\frac{\mathbf{E}}{\mathbb{I}}$. Note that

– there is no zero absolute velocity,
– the magnitude of an absolute velocity makes no sense,
– the angle between two absolute velocities makes no sense;

let us be cautious: the absolute velocity $\boldsymbol{u}_2$ in Figure 6.2 is not longer than $\boldsymbol{u}_1$, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ have no angle between them, $\boldsymbol{u}_1$ is not orthogonal to $\frac{\mathbf{E}}{\mathbb{I}}$.

The same time durations on straight lines with different absolute velocities are represented by intervals of different lengths. The bigger the angle with the

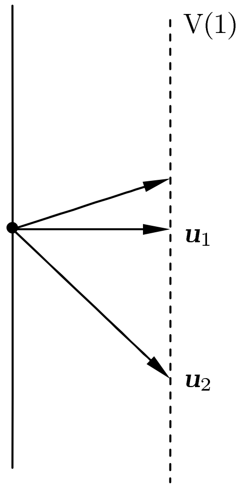

Figure 6.2 Absolute velocities

horizontal line, the longer the interval representing the same duration. The same length corresponds to the same duration for two absolute velocities that have the same angle with the horizontal line. That is why we illustrate two absolute velocities in this way when treating relations concerning them. Lastly, we emphasize again that length and angle are properties of the illustration, not of the absolute velocities.

6.1.2 Duals

The set of **absolute spacelike vectors**, $\mathbf{E}$, is a distinguished three-dimensional linear subspace of $\mathbf{M}$. Its dual, $\mathbf{E}^*$, the set of linear maps $\mathbf{E} \to \mathbb{R}$ is a three-dimensional vector space, too; the Euclidean form $\boldsymbol{b}$ establishes the identification

$$\frac{\mathbf{E}}{\mathbb{D} \otimes \mathbb{D}} \equiv \mathbf{E}^*, \qquad \frac{\boldsymbol{q}}{\boldsymbol{m}^2} \equiv \frac{\boldsymbol{b}(\boldsymbol{q}, \cdot)}{\boldsymbol{m}^2}.$$

We can conceive this in another way, too: the elements of $\mathbf{E}$ can be identified with linear maps $\mathbf{E} \to \mathbb{D} \otimes \mathbb{D}$: $\boldsymbol{q} \equiv \boldsymbol{b}(\boldsymbol{q}, \cdot)$. Thus, according to our convention regarding linear maps, we we write a dot product instead of $\boldsymbol{b}$:

$$\boldsymbol{q} \cdot \boldsymbol{p} := \boldsymbol{b}(\boldsymbol{q}, \boldsymbol{p}).$$

Note the important fact that, on the contrary, there is **no natural correspondence between M and M***.

E is a 'distinguished' linear subspace of **M**, its elements are called **absolute spacelike vectors**.

The transposes of the linear surjection $\mathbf{M} \to \mathbb{I}$ is $\boldsymbol{\tau}^* : \mathbb{I}^* \to \mathbf{M}^*$, a linear injection. Its range, $\boldsymbol{\tau}^*[\mathbb{I}^*]$, is a distinguished one-dimensional linear subspace of $\mathbf{M}^*$ whose elements are called **absolute timelike covectors**. In a simple form: $\boldsymbol{k}$ is an absolute timelike covector if and only if there is (uniquely determined) $\boldsymbol{e} \in \mathbb{I}^*$ such that $\boldsymbol{k} = \boldsymbol{e}\boldsymbol{\tau}$.

6.1.3 Proper times

The history of a material point is a world line in spacetime. If the world points x and y is such that $y - x \in \mathtt{T}^{\to}$ then the inertial time passed between them equals $\boldsymbol{\tau} \cdot (y - x)$. According to 2.4.2, the time passed between x and y on an arbitrary world line C, with a progressive parametrization p, is

$$\boldsymbol{t}_{\mathrm{C}}(x, y) = \int_{p^{-1}(x)}^{p^{-1}(y)} \boldsymbol{\tau} \cdot \dot{p}(a) da = \boldsymbol{\tau} \cdot \int_{p^{-1}(x)}^{p^{-1}(y)} \dot{p}(a) da = \\ = \boldsymbol{\tau} \cdot (y - x).$$

Our result is the **absolute time progress**:

The same proper time passes on arbitrary world line between two timelike separated occurrences.

6.1.4 Absolute time points

Absolute time progress gives the heuristic of absolute time points. An observer sends chronometers started at once from a space point (say Greenwich to all other space points (Paris, Prague, Budapest, etc.); the same number of ticks of these chronometers indicates simultaneous instants at the different space pons.

This is formulated in the model as follows. The occurrences x and y are simultaneous if $y - x \in \mathbf{E}$ i.e. $\boldsymbol{\tau} \cdot (y - x) = 0$; in other words, the set of occurrences, simultaneous with x is $x + \mathbf{E}$; Thus, $\mathbf{E}$ establishes this synchronization. Since the set of futurelike vectors is a half space having $\mathbf{E}$ as its boundary, $\mathbf{E}$ is the only three-dimensional subspace transverse to all futurelike vectors. Thus, in the nonrelativistic spacetime model only one synchronization is possible, called **absolute synchronization**, the one determined by $\mathbf{E}$.

The absolute instants are hyperplanes directed by $\mathbf{E}$, their collection is **absolute time**, $\mathtt{I} := \mathtt{M}/\mathbf{E}$.

Let us introduce the map

$$\tau : \mathtt{M} \to \mathtt{I}, \qquad x \mapsto x + \mathbf{E}, \tag{IV6}$$

called **time evaluation**. Time evaluation assigns to every occurrence the corresponding absolute time point[1].

The difference of two absolute instants t and s is defined as the absolute time duration between any of their occurrences:

$$t - s = \boldsymbol{\tau} \cdot (y - x), \qquad (y \in t, x \in s). \tag{IV7}$$

In other words,

$$(y + \mathbf{E}_{\boldsymbol{u}}) - (x + \mathbf{E}_{\boldsymbol{u}}) = \tau_{\boldsymbol{u}}(y) - \tau_{\boldsymbol{u}}(x) = \boldsymbol{\tau} \cdot (y - x). \tag{IV8}$$

The subtraction above turns $\mathtt{I}_{\boldsymbol{u}}$ into a one-dimensional affine space over $\mathbb{I}$ and $\tau_{\boldsymbol{u}}$ becomes an affine map over the linear map $\boldsymbol{\tau}_{\boldsymbol{u}} = -\boldsymbol{u}\cdot$.

If the occurrences x and y have the common property that $\tau(y) - \tau(x) = \boldsymbol{\tau} \cdot (y - x) > 0$, i.e. y is absolute futurelike with respect to x, then we write $\tau(y) > \tau(x)$ expressing that y is absolute later than x.

6.2 The arithmetic spacetime model

We can construct a nonrelativistic spacetime model with the aid of real numbers. In this **arithmetic nonrelativistic spacetime model**

– $\mathtt{M} = \mathbb{R}^4$ endowed with the standard orientation (then $\mathbf{M} = \mathbb{R}^4$, too),

– $\mathbb{I} = \mathbb{R}$ endowed with the standard orientation,

– $\mathbb{D} = \mathbb{R}$ endowed with the standard orientation,

– $\boldsymbol{\tau}(\xi^0, \xi^1, \xi^2, \xi^3) = \tau_i \xi^i = \xi^0$, i.e. $\boldsymbol{\tau}$ is a covector, in components $\boldsymbol{\tau} = (1, 0, 0, 0)$.

Now we remind that corresponding to the usual formulae, $(\mathbb{R}^4)^* \equiv \mathbb{R}^4$; since both $\mathbf{M}$ and $\mathbf{M}^*$ are $\mathbb{R}^4$ here, but there is no natural correspondence between

[1]In the book T. Matolcsi: *Spacetime without Reference Frames* (Budapest, 1993, Akadémiai Kiadó) the nonrelativistic spacetime model is $(\mathtt{M}, \mathtt{I}, \tau, \mathbb{D}, \boldsymbol{b})$; the present notation fits better the general treatment of flat spacetime models

the elements of $\mathbf{M}$ and $\mathbf{M}^*$, the components of vectors are described with superscripts and the components of covectors are described with subscripts; if $\boldsymbol{x} = (\xi^0, \xi^1, \xi^2, \xi^3)$ and $\boldsymbol{k} = (\kappa_0, \kappa_1, \kappa_2, \kappa_3)$, then $\boldsymbol{k} \cdot \boldsymbol{x} = \kappa_0\xi^0 + \kappa_1\xi^1 + \kappa_2\xi^2 + \kappa_3\xi^3 = \boldsymbol{k}_i\xi^i$ where the Einstein summation rule is applied.

As a consequence

$$\mathbf{E} = \{(\xi^0, \xi^1, \xi^2, \xi^3) \mid \xi^0 = 0\} = \{0\} \times \mathbb{R}^3 \equiv \mathbb{R}^3,$$

– $\boldsymbol{b}$ is the usual inner product on $\mathbb{R}^3$.

Furthermore, here

$$\mathrm{T}^{\to} = \{(\xi^0, \xi^1, \xi^2, \xi^3) \mid \xi^0 > 0\} = \mathbb{R}^+ \times \mathbb{R}^3,$$

$$\mathtt{V}(1) = \{(\xi^0, \xi^1, \xi^2, \xi^3) \mid \xi^0 = 1\} = \{1\} \times \mathbb{R}^3.$$

The matrix of the boosts is

$$(\boldsymbol{B}_{\nu'\nu})_i^k = \delta_i^k + (\nu' - \nu)^k \tau_i.$$

The world points $(\xi^0, \xi^1, \xi^2, \xi^3)$ and $(\eta^0, \eta^1, \eta^2, \eta^3)$ are simultaneous if and only if $\xi^0 = \eta^0$.

Consequently, the absolute instants (the hyperplanes directed by $\{0\} \times \mathbb{R}^3$) can be identified with real numbers in a natural way: the hyperplane corresponding to the real number t is $\{t\} \times \mathbb{R}^3$. Thus we can take the identification $\mathtt{I} \equiv \mathbb{R}$ and then the time evaluation τ coincides with $\boldsymbol{\tau}$.

Let us repeat: in the arithmetic spacetime model

– the spacetime $\mathtt{M}$ and the underlying vector space $\mathbf{M}$ are the same set,

– the absolute time $\mathtt{I}$ and the measure line $\mathbb{I}$ of time periods are the same set,

– the time evaluation τ and the time progress $\boldsymbol{\tau}$ are the same map.

– $\mathbb{I}$ and $\mathbb{D}$ are the same set, the real line, consequently every measure line is $\mathbb{R}$; therefore $\frac{\mathbf{M}}{\mathbb{I}} = \mathbf{M} = \mathbb{R}^4$, $\frac{\mathbf{E}}{\mathbb{D}\otimes\mathbb{D}} = \mathbf{E} = \mathbb{R}^3$, etc.

6.3 Isomorphisms

Isomorphism of models is an important notion expressing whether two spacetime models of different forms have the same physical content or not (see 2.9.2).

6.3.1 Isomorphism in the new notation

We can assert that the nonrelativistic spacetime model $(\hat{\mathtt{M}}, \hat{\mathbb{I}}, \hat{\mathbb{D}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{b}})$ is isomorphic to the model $(\mathtt{M}, \mathbb{I}, \mathbb{D}, \boldsymbol{\tau}, \boldsymbol{b})$ if and only if there are

(i) an orientation preserving affine bijection $L : \mathtt{M} \to \hat{\mathtt{M}}$ (over the linear bijection $\boldsymbol{L} : \mathbf{M} \to \hat{\mathbf{M}}$),

(ii) an orientation preserving linear bijection $\boldsymbol{B} : \mathbb{I} \to \hat{\mathbb{I}}$,

(iii) an orientation preserving linear bijection $\boldsymbol{Z} : \mathbb{D} \to \hat{\mathbb{D}}$

which send $\mathtt{T}^{\to}$ into $\widehat{\mathtt{T}^{\to}}$, $\boldsymbol{P}$ into $\hat{\boldsymbol{P}}$, $\boldsymbol{d}$ into $\hat{\boldsymbol{d}}$ and $\boldsymbol{B}$ into $\hat{\boldsymbol{B}}$ in a 'convenient way'.

$\boldsymbol{F}$ and $\boldsymbol{Z}$ are given by the 'units' s and ŝ, m and m̂, respectively, according to 2.9.2.

Since $\boldsymbol{\tau}$ determines both $\mathtt{T}^{\to}$ and $\boldsymbol{P}$, the first two general requirements can be fused:

(I)-(II)

$$\frac{\hat{\boldsymbol{\tau}}(\boldsymbol{L} \cdot \boldsymbol{x})}{\hat{\mathrm{s}}} = \frac{\boldsymbol{\tau}(\boldsymbol{x})}{\mathrm{s}}$$

for all $\mathbf{x} \in \mathbf{M}$.

The equality above implies that $\boldsymbol{L}$ maps $\mathbf{E} := \mathrm{Ker}\boldsymbol{\tau}$ onto $\hat{\mathbf{E}} := \mathrm{Ker}\hat{\boldsymbol{\tau}}$, thus the requirement imposed on the Euclidean structures is simplified as follows:

(III)

$$\frac{\hat{\boldsymbol{b}}(\boldsymbol{L} \cdot \boldsymbol{q}, \boldsymbol{L} \cdot \boldsymbol{p})}{\hat{\mathrm{m}}^2} = \frac{\boldsymbol{b}(\boldsymbol{q}, \boldsymbol{p})}{\mathrm{m}^2}$$

for all $\boldsymbol{q}, \boldsymbol{p} \in \mathbf{E}$.

The last requirement,

(IV)

$$\left(\hat{\mathbf{1}} + (\boldsymbol{L} \cdot \boldsymbol{u}' - \boldsymbol{L} \cdot \boldsymbol{u}) \otimes \hat{\boldsymbol{\tau}}\right) \cdot \boldsymbol{L} = \boldsymbol{L} \cdot (\mathbf{1} + (\boldsymbol{u}' - \boldsymbol{u}) \otimes \boldsymbol{\tau})$$

follows automatically from the previous ones.

It is worth noting that if the two spacetime models are isomorphic then there is an affine bijection $F : \mathtt{I} \to \hat{\mathtt{I}}$ over $\boldsymbol{F}$ such that $\hat{\tau} \circ L = F \circ \tau$ i.e. $F(x + \mathbf{E}) = L(x) + \hat{\mathbf{E}}$ for all occurrences x. Indeed, the previous equality is just the definition of F; we have to show only that the definition is correct: if $x + \mathbf{E} = y + \mathbf{E}$ then $L(x) + \hat{\mathbf{E}} = L(y) + \hat{\mathbf{E}}$ i.e. if $x - y \in \mathbf{E}$ then $L(x) - L(y) = \boldsymbol{L}(x - y) \in \hat{\mathbf{E}}$ which is satisfied.

6.3.2 Nonrelativistic spacetime models are isomorphic

We can easily demonstrate:

An arbitrary nonrelativistic spacetime model is isomorphic to the arithmetic one; this implies that arbitrary two nonrelativistic spacetime models are isomorphic to each other.

Considering an arbitrary nonrelativistic spacetime model $(\mathtt{M}, \mathbb{I}, \mathbb{D}, \boldsymbol{\tau}, \boldsymbol{b})$, let us choose

– a time unit $\mathrm{s} \in \mathbb{I}^+$,
– a distance unit $\mathrm{m} \in \mathbb{D}^+$,
– an 'origin' $o \in \mathtt{M}$.
– a futurelike vector $\boldsymbol{e}_0$ for which $\boldsymbol{\tau} \cdot \boldsymbol{e}_0 = \mathrm{s}$ holds,
– a positively oriented orthogonal basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, normed to m, in $\mathbf{E}$ (which means that $\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ is a positively oriented basis in $\mathbf{M}$ and $\boldsymbol{b}(\boldsymbol{e}_i, \boldsymbol{e}_k) = \mathrm{m}^2 \delta_{ik}$),

and let us put

$$L : \mathtt{M} \to \mathbb{R}^4, \qquad x \mapsto \{ \text{ coordinates of } x - o \text{ in the basis } \boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3 \},$$

$$\boldsymbol{B} : \mathbb{I} \to \mathbb{R}, \qquad \boldsymbol{t} \mapsto \frac{\boldsymbol{t}}{\mathrm{s}},$$

$$\boldsymbol{Z} : \mathbb{D} \to \mathbb{R}, \qquad \boldsymbol{h} \mapsto \frac{\boldsymbol{h}}{\mathrm{m}}.$$

It is simple that in such a way we have defined an isomorphism from the spacetime model in question to the arithmetic spacetime model. Indeed, if $\boldsymbol{x} = \sum_{i=0}^{3} \xi^i \boldsymbol{e}_i$, then $\boldsymbol{\tau} \cdot \boldsymbol{x} = \xi^0 \boldsymbol{e}_0$; if $\boldsymbol{q} = \sum_{i=1}^{3} \xi^i \boldsymbol{e}_i$ and $\boldsymbol{p} = \sum_{i=1}^{3} \eta^i \boldsymbol{e}_i$ are elements of $\mathbf{E}$, then $\boldsymbol{b}(\boldsymbol{q}, \boldsymbol{p}) = \left(\sum_{i=1}^{3} \xi^i \eta^i \right) \mathrm{m}^2$.

Our result on isomorphisms says that all nonrelativistic spacetime models have the same physical content. Therefore, we have only philosophical and practical reasons to make any specific choice. The arithmetic model is suitable for practical purposes, i.e. for solving actual problems. For theoretical considerations, however, a special model, like the arithmetic one, is less appropriate because such a model can have additional properties

– that have nothing to do with the structure of the spacetime model,
– that confuse some essential features of the spacetime model.

We emphasize again: a treatise in coordinates is not defective in itself, because all nonrelativistic spacetime models have the same physical content. Nevertheless, it is preferred not to use the arithmetic model for general considerations because its special properties can easily mislead us:

– spacetime points and spacetime vectors are given as elements of the same set,

– time instants and time periods are confused, time evaluation and time progress seem to be the same,

– a world point appears as a couple of a time point and a space point

– all measure lines are the real line,

etc. Using coordinates, if we do not want to make a mistake, we have to check permanently the physical validity of our ideas and formulae. This is very tiresome and, even if we do so, something can easily escape our attention. In Subsection 7.9 we find a convincing example how coordinates can mislead us.

6.3.3 Galilei and Noether transformations

The **Galilei transformations** in the spacetime model $(\mathtt{M}, \mathbb{I}, \mathbb{D}, \boldsymbol{\tau}, \boldsymbol{b})$ are linear bijections $\boldsymbol{L} : \mathbf{M} \to \mathbf{M}$ for which

(I) $\boldsymbol{\tau} \cdot (\boldsymbol{L} \cdot \boldsymbol{x}) = \pm \boldsymbol{\tau} \cdot \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbf{M}$,

(II) $\boldsymbol{b}(\boldsymbol{L} \cdot \boldsymbol{q}, \boldsymbol{L} \cdot \boldsymbol{p}) = \boldsymbol{b}(\boldsymbol{q}, \boldsymbol{p})$ for all $\boldsymbol{q}, \boldsymbol{p} \in \mathbf{E}$.

The composition of Galilei transformations is a Galilei transformation i.e. their collection is a group.

The orientation preserving Galilei transformations for which the positive sign appears in (I) are called **proper Galilei transformations**. According to the general definition in 2.9.3 and according to 6.3.1 they are the **vectorial symmetries** of the nonrelativistic spacetime model.

The proper Galilei transformations whose restriction to $\mathbf{E}$ is the identity are called **special Galilei transformations**. If $\boldsymbol{L}$ is a special Galilei transformation then $\boldsymbol{L} \cdot \boldsymbol{u}' - \boldsymbol{L} \cdot \boldsymbol{u} = \boldsymbol{L} \cdot (\boldsymbol{u}' - \boldsymbol{u}) = \boldsymbol{u}' - \boldsymbol{u}$ for all absolute velocities $\boldsymbol{u}$ and $\boldsymbol{u}'$, thus $\boldsymbol{L} \cdot \boldsymbol{u}' - \boldsymbol{u}' = \boldsymbol{L} \cdot \boldsymbol{u} - \boldsymbol{u} =: \mathbf{v}_{\boldsymbol{L}}$ is independent of the absolute velocities. As a consequence, $\boldsymbol{L} \cdot \boldsymbol{x} = \boldsymbol{x} + (\boldsymbol{\tau} \cdot \boldsymbol{x})\mathbf{v}_{\boldsymbol{L}}$ for all vectors $\boldsymbol{x}$ i.e. for every special Galilei transformation $\boldsymbol{L}$ there is a unique $\mathbf{v}_{\boldsymbol{L}} \in \frac{\mathbf{E}}{\mathbb{I}}$ such that

$$\boldsymbol{L} = \mathbf{1} + \mathbf{v}_{\boldsymbol{L}} \otimes \boldsymbol{\tau}. \tag{IV9}$$

The special Galilei transformations form a group and $\mathbf{v}_{\boldsymbol{LK}} = \mathbf{v}_{\boldsymbol{L}} + \mathbf{v}_{\boldsymbol{K}}$.

The boosts are special Galilei transformations, as expected.

The **Noether transformations** are the affine bijections $L : \mathtt{M} \to \mathtt{M}$ over the Galilei transformations. They form a group, too. The **proper Noether transformations** are the ones over the proper Galilei transformations. These are the **symmetries** of the nonrelativistic spacetime model.

A detailed treatment of the properties of the Galilei group and the Noether group can be found in the book

T. Matolcsi: *Spacetime without Reference Frames* (Budapest, 1993, Akadémiai Kiadó)

6.4 Space and space vectors of an inertial observer

6.4.1 Representation of the space vectors

Let us recall that the space of the inertial observer $\boldsymbol{u}$ is the set of stright lines in $\mathtt{M}$ directed by $\boldsymbol{u}$ (see 2.5.3),

$$\mathtt{E}_{\boldsymbol{u}} := \mathtt{M}/\mathbb{I}\boldsymbol{u},$$

and the set of $\boldsymbol{u}$-space vectors is the set $\mathbf{M}/\mathbb{I}\boldsymbol{u}$ of straight lines in $\mathbf{M}$ directed by $\boldsymbol{u}$.

The $\boldsymbol{u}$-space points are of the form $x+\mathbb{I}\boldsymbol{u}$ and the $\boldsymbol{u}$-space vectors are of the form $\boldsymbol{x}+\mathbb{I}\boldsymbol{u}$. $\mathtt{E}_{\boldsymbol{u}}$ is affine space over $\mathbf{M}/\mathbb{I}\boldsymbol{u}$ by the subtraction

$$(x+\mathbb{I}\boldsymbol{u}) - (y+\mathbb{I}\boldsymbol{u}) = (x-y)+\mathbb{I}\boldsymbol{u}.$$

The space vectors of an inertial observer cannot be fairly illustrated. The three-dimensional subspace $\mathbf{E}$ of absolute spacelike vectors is transverse to all absolute velocities, and no other subspace has this property, so the space vectors of any inertial observer can be identified with the elements of $\mathbf{E}$ by the linear bijection

$$\mathbf{E} \to \mathbf{M}/\mathbb{I}\boldsymbol{u}, \quad \boldsymbol{q} \mapsto \boldsymbol{q}+\mathbb{I}\boldsymbol{u}. \tag{IV10}$$

Let us introduce the notation

$$.\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{x} := \boldsymbol{x} - \boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{x}) \tag{IV11}$$

It is evident that

$$\boldsymbol{\sigma}_{\boldsymbol{u}} = \mathbf{1} - \boldsymbol{u}\otimes\boldsymbol{\tau} : \mathbf{M} \to \mathbf{E}$$

is a linear surjection and its is easy to see that

$$\mathbf{M}/\mathbb{I}\boldsymbol{u} \to \mathbf{E}, \quad \boldsymbol{x}+\mathbb{I}\boldsymbol{u} \mapsto \boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{x}$$

is the inverse of the linear bijection IV10.

Thus, the identification in question is

$$\mathbf{E} \equiv \mathbf{M}/\mathbb{I}\boldsymbol{u}, \qquad \boldsymbol{q} \equiv \boldsymbol{q}+\mathbb{I}\boldsymbol{u}$$

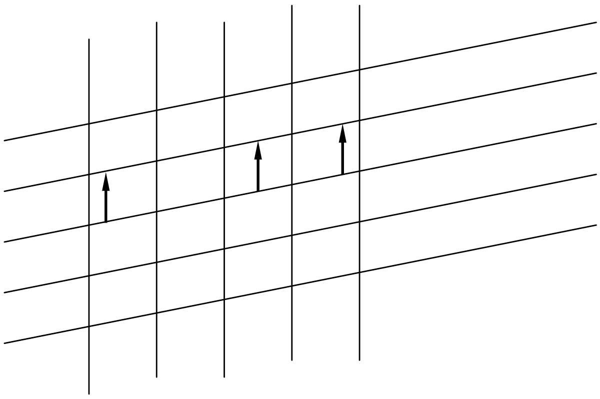

Figure 6.3 Space vectors of an inertial observer

or

$$\mathbf{M}/\mathbb{I}\boldsymbol{u} \equiv \mathbf{E}, \qquad \boldsymbol{x} + \mathbb{I}\boldsymbol{u} \equiv \boldsymbol{\sigma_u} \cdot \mathbf{x}.$$

E can be endowed with an orientation in a natural way. Let $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ be an ordered basis of **E** positively oriented if $(\boldsymbol{tu}, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ is a positively oriented basis of **M** for some (hence for arbitrary) positive element $\boldsymbol{t}$ of $\mathbb{I}$. It is easy to see that the definition is good, i.e. if $(\boldsymbol{tu}, \boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3')$ is a positively oriented basis in **M**, then $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ and $(\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3')$ are equally oriented in **E**. Further, it is evident, too, that this orientation corresponds to the one given in Paragraph 2.5.4.

The set **E** of absolute spacelike vectors is a three-dimensional Euclidean vector space, the related knowledge can be found in the mathematical supplement.

Let us introduce the notation

$$\sigma_{\boldsymbol{u}}(x) := x + \mathbb{I}\boldsymbol{u};$$

$\sigma_{\boldsymbol{u}}(x)$ is the $\boldsymbol{u}$-space point containing the occurrence x. According to the above identification, the subtraction in the $\boldsymbol{u}$-space becomes

$$\sigma_{\boldsymbol{u}}(x) - \sigma_{\boldsymbol{u}}(y) = (x + \mathbb{I}\boldsymbol{u}) - (y + \mathbb{I}\boldsymbol{u}) = (x - y) + \mathbb{I}\boldsymbol{u} \equiv \boldsymbol{\sigma_u} \cdot (x - y). \qquad \text{(IV12)}$$

In another way, if q and p are points in $\mathtt{E}_{\boldsymbol{u}}$, then

$$q - p := \boldsymbol{\sigma_u} \cdot (x - y) \qquad (x \in q,\ y \in p),$$

which equivalent to

$$q - p = x - y \qquad (x \in q,\ y \in p,\ x - y \in \mathbf{E}).$$

Note that (IV12) says that $\sigma_{\boldsymbol{u}} : \mathtt{M} \to \mathtt{E}_{\boldsymbol{u}}$, $x \mapsto x + \mathbb{I}\boldsymbol{u}$ is an affine map over the linear map $\boldsymbol{\sigma_u} : \mathbf{M} \to \mathbf{E}$.

On the base of the above identification, we can assert:

The spaces of different inertial observers are **different** three-dimensional affine spaces over the **same** vector space.

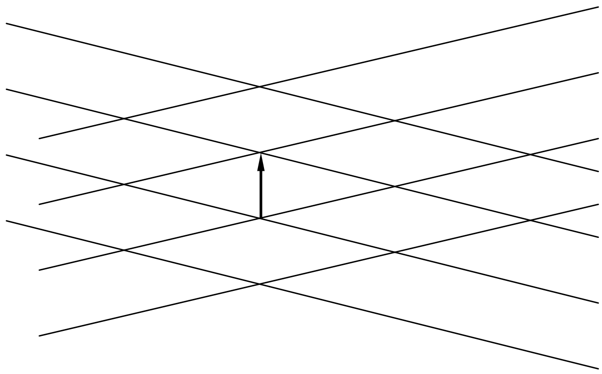

Figure 6.4 Different spaces, same space vectors

6.4.2 Properties of the boosts

The formula $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} = \mathbf{1} + (\boldsymbol{u}' - \boldsymbol{u}) \otimes \boldsymbol{\tau}$ of the boost from $\boldsymbol{u}$ to $\boldsymbol{u}'$ yields immediately that

$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{q} = \boldsymbol{q} \qquad (\boldsymbol{q} \in \mathbf{E}),$$

i.e. the restriction of the boosts onto $\mathbf{E}$ is the identity. By the way, this fact and the required property $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{u} = \boldsymbol{u}'$ determine uniquely the boost from $\boldsymbol{u}$ to $\boldsymbol{u}'$.

The equality of vectors in the spaces of different inertial observers can be realized physically in such a way that one of the observers makes an **instantaneous print** from a vector of the other inertial observer i.e. this observer marks its own space points that meet at an instant the space points corresponding to the vector in question (see Figure 6.4).

This makes evident that the restriction of boosts to $\mathbf{E}$ is the identity.

A simple consequence of the definition that the boosts are **transitive**:

$$\boldsymbol{B}_{\mathbf{u}''\mathbf{u}'}\boldsymbol{B}_{\mathbf{u}'\mathbf{u}} = \boldsymbol{B}_{\mathbf{u}''\mathbf{u}}$$

which is the same that

$$\boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\boldsymbol{B}_{\mathbf{u}''\mathbf{u}'}\boldsymbol{B}_{\mathbf{u}'\mathbf{u}} = \mathbf{1}$$

for all absolute velocities $\boldsymbol{u}$, $\boldsymbol{u}'$ and $\boldsymbol{u}''$.

The boosts are special Galilei transformations i.e. vectorial symmetries, as expected. Conversely, every special Galilei transformation is a boost: the transformation (IV9) is the boost from an arbitrary $\boldsymbol{u}$ to $\boldsymbol{u} + \mathbf{v}_{\boldsymbol{L}}$.

6.5 Relative velocities

Let us take an inertial observer and a straight line directed by $\boldsymbol{u}'$ which is the world line of an inertial material point. If $\boldsymbol{u}' \neq \boldsymbol{u}$, then the observer perceives that the material point is moving. The velocity of the material point relative to the observer is determined as follows.

Let x and y world points on the world line of the material point i.e. $y - x = \boldsymbol{s}\boldsymbol{u}'$ for some $\boldsymbol{s} \in \mathbb{I}$. The material point meets the $\boldsymbol{u}$-space points $x + \mathbb{I}\boldsymbol{u}$ and $y + \mathbb{I}\boldsymbol{u}$ at the instants $x + \mathbf{E}$ and $y + \mathbf{E}$, respectively. The $\boldsymbol{u}$-space vector between these $\boldsymbol{u}$-space points is $(y + \mathbb{I}\boldsymbol{u}) - (x + \mathbb{I}\boldsymbol{u}) = \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (y - x) = (\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{u}')\boldsymbol{s}$; the time period between the two meetings is $(y + \mathbf{E}) - (x + \mathbf{E}) = \boldsymbol{\tau} \cdot (y - x) = \boldsymbol{s}$. Consequently, the average relative velocity between the two $\boldsymbol{u}$-space points equals

$$\frac{(y + \mathbb{I}\boldsymbol{u}) - (x + \mathbb{I}\boldsymbol{u})}{(y + \mathbf{E}) - (x + \mathbf{E})} = \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{u}' = \boldsymbol{u}' - \boldsymbol{u}.$$

The result is independent of the choice of the instants: an inertial material point moves uniformly with respect to an inertial observer. The **relative velocity** of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$ is the difference of the absolute velocities:

$$\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} := \boldsymbol{u}' - \boldsymbol{u} \in \frac{\mathbf{E}}{\mathbb{I}}.$$

The relative velocity of an arbitrary absolute velocity with respect to another arbitrary absolute velocity lies in the same three-dimensional Euclidean space. As a consequence, in contrast to absolute velocities,

– there is a zero relative velocity,
– a relative velocity has a magnitude,

– the angle between two relative velocities makes sense.

The **reciprocity** of relative velocities,

$$\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'} = -\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}},$$

and their **transitivity**,

$$\mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}'} + \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} = \mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}}$$

are well seen from the definition.

All these facts correspond to our everyday experience.

6.6 Splitting of vectors and transformation rules

6.6.1 Splitting

Since $\mathbb{I}\boldsymbol{u}$ and $\mathbf{E}$ are transverse subspaces, every world vector can be given uniquely as a sum of vectors in those subspaces. Using the notation of (IV11), we get that sum in he form $\boldsymbol{x} = \boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{x}) + \boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{x}$ for all space vectors $\boldsymbol{x}$. We say that the $\boldsymbol{u}$ inertial observer **splits** the world vector $\boldsymbol{x}$ into the **timelike component** $\boldsymbol{\tau}\cdot\boldsymbol{x}$ and the **$\boldsymbol{u}$-spacelike component** $\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\mathbf{x}$. The linear bijection

$$\boldsymbol{h}_{\boldsymbol{u}} := (\tau, \boldsymbol{\sigma}_{\boldsymbol{u}}) : \mathbf{M} \to \mathbb{I}\times\mathbf{E}, \qquad \boldsymbol{x} \mapsto (\boldsymbol{\tau}\cdot\boldsymbol{x}, \boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{x}) \tag{IV13}$$

is called the **splitting of spacetime vectors** by $\boldsymbol{u}$.

Note that

$$\boldsymbol{h}_{\boldsymbol{u}}^{-1}(\boldsymbol{t},\boldsymbol{q}) = \boldsymbol{t}\boldsymbol{u} + \boldsymbol{q} \qquad (\boldsymbol{t},\boldsymbol{q}) \in \mathbb{I}\times\mathbf{E}.$$

Of course, the tensor products and quotients of $\mathbf{M}$ by measure lines such as $\frac{\mathbf{M}}{\mathbb{I}}$ or $\frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}}$ are split corresponding to the above formula, because the multiplication and division by element of measure lines can be interchanged with linear maps. For instance, the timelike component of an absolute velocity $\boldsymbol{u}'$ is $\boldsymbol{\tau}\cdot\boldsymbol{u}' = 1$, its $\boldsymbol{u}$-spacelike component is

$$\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{u}' = \boldsymbol{u}' - \boldsymbol{u}. \tag{IV14}$$

Thus, according to 6.5, the relative velocity of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$ is just the $\boldsymbol{u}$-spacelike component of $\boldsymbol{u}'$. To sum up:

$$\boldsymbol{h}_{\boldsymbol{u}}\cdot\boldsymbol{u}' = (1, \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}).$$

The splitting of vectors determines the splitting of covectors by the formula

$$\boldsymbol{r}_{\boldsymbol{u}} := \left(\boldsymbol{h}_{\boldsymbol{u}}^{-1}\right)^* : \mathbf{M}^* \to (\mathbb{I}\times\mathbf{E})^* = \mathbb{I}^*\times\mathbf{E}^*.$$

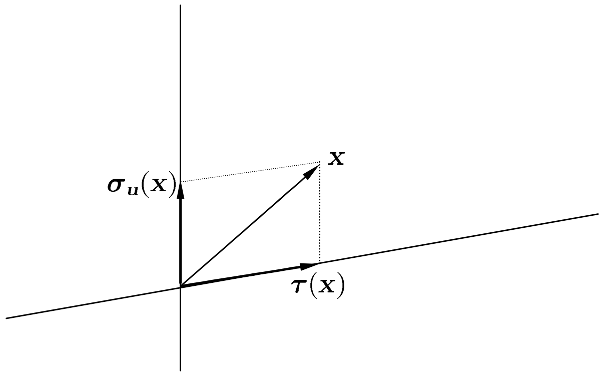

Figure 6.5 Splitting of vectors

If $\boldsymbol{k}$ is a covector and $(\boldsymbol{t}, \boldsymbol{q}) \in \mathbb{I} \times \mathbf{E}$ then $(\boldsymbol{r_u} \cdot \boldsymbol{k}) \cdot (\boldsymbol{t}, \boldsymbol{q}) = \big(\boldsymbol{k} \cdot \boldsymbol{h_u}^{-1}\big)(\boldsymbol{t}, \boldsymbol{q}) = (\boldsymbol{k} \cdot \boldsymbol{u})\boldsymbol{t} + \boldsymbol{k} \cdot \boldsymbol{q}$.

This splitting can be written in a convenient form by introducing the embedding $\boldsymbol{i} : \mathbf{E} \to \mathbf{M}$ i.e.

$$\boldsymbol{i} \cdot \boldsymbol{q} = \boldsymbol{q} \qquad (\boldsymbol{q} \in \mathbf{E},\ \boldsymbol{i} \cdot \boldsymbol{q} \in \mathbf{M})).$$

Its transpose is

$$\boldsymbol{i}^* : \mathbf{M}^* \to \mathbf{E}^*, \quad \boldsymbol{k} \mapsto \boldsymbol{k} \cdot \boldsymbol{i} = \boldsymbol{k}|_{\mathbf{E}},$$

where, as usual, $|_{\mathbf{E}}$ denotes the restriction onto $\mathbf{E}$.

As it is known (see the mathematical supplement), we can interchange the roles of $\boldsymbol{k}$ and $\boldsymbol{u}$ in the duality; but, contrary to the formulas in the supplement, we find convenient to write $\boldsymbol{u}^*$ instead of $\boldsymbol{u}$ in this role, for the better understanding. Thus, the **splitting of spacetime covectors** by the observer $\boldsymbol{u}$ is the liner bijection

$$\boldsymbol{k} \mapsto (\boldsymbol{u}^* \cdot \boldsymbol{k}, \boldsymbol{i}^* \cdot \boldsymbol{k}) = (\boldsymbol{k} \cdot \boldsymbol{u}, \boldsymbol{k} \cdot \boldsymbol{i})$$

$\boldsymbol{u}^* \cdot \boldsymbol{k} = \boldsymbol{k} \cdot \boldsymbol{u}$ and $\boldsymbol{i}^* \cdot \boldsymbol{k} = \boldsymbol{k} \cdot \boldsymbol{i}$ are the ***u*-timelike component** and the **spacelike component**, respectively, of the covector $\boldsymbol{k}$.

Concisely:

$$\boldsymbol{r_u} = (\boldsymbol{u}^*, \boldsymbol{i}^*).$$

The inverse of the splitting is $\boldsymbol{r}_{\boldsymbol{u}}^{-1} = \boldsymbol{h}_{\boldsymbol{u}}^* : \mathbb{I}^* \times \mathbf{E}^*$ which, according to

$$\Big(\boldsymbol{h}_{\boldsymbol{u}}^*(\boldsymbol{e},\boldsymbol{p})\Big)\cdot\boldsymbol{x} = (\boldsymbol{e},\boldsymbol{p})\cdot(\boldsymbol{h}_{\boldsymbol{u}}\cdot\boldsymbol{x}) = (\boldsymbol{e},\boldsymbol{p})\cdot\Big(\boldsymbol{\tau}\cdot\boldsymbol{x},\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{x}\Big) = \\ = (\boldsymbol{e}\boldsymbol{\tau} + \boldsymbol{p}\cdot\boldsymbol{\sigma}_{\boldsymbol{u}})\cdot\boldsymbol{x}$$

can be written in the form

$$\boldsymbol{r}_{\boldsymbol{u}}^{-1}(\boldsymbol{e},\boldsymbol{p}) = \boldsymbol{e}\boldsymbol{\tau} + \boldsymbol{p}\cdot\boldsymbol{\sigma}_{\boldsymbol{u}} \quad ((\boldsymbol{e},\boldsymbol{p})\in\mathbb{I}^*\times\mathbf{E}^*).$$

Note that **the timelike component of vectors are absolute**, i.e. independent of observers, whereas **the spacelike component of covectors are absolute**.

The spacelike component of a vector $\boldsymbol{x}$ depends on the observer unless the vector is absolute spacelike: the timelike component of $\boldsymbol{q}\in\mathbf{E}$ is zero ($\boldsymbol{\tau}\cdot\boldsymbol{q}=0$) and its $\boldsymbol{u}$-spacelike component is $\boldsymbol{q}$ itself for all $\boldsymbol{u}$ ($\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{q}=\boldsymbol{q}$).

The timelike component of a covector $\boldsymbol{k}$ depends on the observer unless the covector is absolute timelike, i.e. when it is in $\boldsymbol{\tau}^*[\mathbb{I}^*]$: the $\boldsymbol{u}$-timelike component of $\boldsymbol{k}=e\boldsymbol{\tau}$ is e for all $\boldsymbol{u}$ and its spacelike component is zero.

At present the splitting of spacetime vectors and covectors appeared as mathematical formulae. Later we shall see that they have a fundamental physical meaning: instead of the vectors themselves, an observer ‘perceives’ only their split components. For instance, it will be clear that the $\boldsymbol{u}$-spacelike component of $\boldsymbol{u}$ is just the relative velocity of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$.

6.6.2 Transformation rules

Different inertial observers split world vectors differently. More precisely, the timelike component is the same but the spacelike component depends on the observer. To see the difference in splitting, we have to compare them somehow. We perform the comparison as follows. Let $(\boldsymbol{t},\boldsymbol{q})$ and $(\boldsymbol{t}',\boldsymbol{q}')$ the split form of the same vector due to $\boldsymbol{u}$ and $\boldsymbol{u}'$. Then

$$\begin{aligned}(\boldsymbol{t}',\boldsymbol{q}') =& \boldsymbol{h}_{\boldsymbol{u}'}\big(\boldsymbol{h}_{\boldsymbol{u}}^{-1}(\boldsymbol{t},\boldsymbol{q})\big) = \boldsymbol{h}_{\boldsymbol{u}'}(\boldsymbol{t}\boldsymbol{u}+\boldsymbol{q}) = \\ =& (\boldsymbol{t},\boldsymbol{q}-\boldsymbol{t}(\boldsymbol{u}'-\boldsymbol{u})).\end{aligned}$$

Using the relative velocity $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} := \boldsymbol{u}'-\boldsymbol{u}$, we can write the above formula in the well known form of Galilean transformation rule:

$$\boldsymbol{t}'=\boldsymbol{t}, \qquad \boldsymbol{q}'=\boldsymbol{q}-\boldsymbol{t}\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}.$$

That is why, for the sake of simplicity, with the notation $\boldsymbol{v} := \boldsymbol{v}_{\boldsymbol{u}'\boldsymbol{u}}$, the linear map

$$\boldsymbol{h}_{\boldsymbol{u}'\boldsymbol{u}} := \boldsymbol{h}_{\boldsymbol{u}'} \cdot \boldsymbol{h}_{\boldsymbol{u}}^{-1} = \begin{pmatrix} 1 & 0 \\ -\boldsymbol{v} & \mathbf{1} \end{pmatrix} : (\mathbb{I} \times \mathbf{E}) \to (\mathbb{I} \times \mathbf{E})$$

is called the **Galilean transformation rule** .

In fact, this is not the usual form, because that one refes to coordinates, so $\mathbb{R}$ and $\mathbb{R}^3$ appear instead of $\mathbb{I}$ and $\mathbf{E}$, respectively, and then 'the corresponding axis of the spatial coordinate systems are parallel'.

We easily obtain the covectorial transformation rule

$$\boldsymbol{r}_{\boldsymbol{u}'\boldsymbol{u}} := \boldsymbol{r}_{\boldsymbol{u}'}\boldsymbol{r}_{\boldsymbol{u}}^{-1} = (\boldsymbol{h}_{\boldsymbol{u}}'^{-1})^* \boldsymbol{h}_{\boldsymbol{u}}^* = (\boldsymbol{h}_{\boldsymbol{u}'\boldsymbol{u}}^{-1})^* = \begin{pmatrix} 1 & \boldsymbol{v} \\ 0 & \mathbf{1} \end{pmatrix} : (\mathbb{I}^* \times \mathbf{E}^*) \to (\mathbb{I}^* \times \mathbf{E}^*).$$

Thus, if $(\boldsymbol{e}, \boldsymbol{p})$ and $(\boldsymbol{e}', \boldsymbol{p}')$ are the split forms of te same covector due to $\boldsymbol{u}$ and $\boldsymbol{u}'$, then

$$\boldsymbol{e}' = \boldsymbol{e} + \boldsymbol{p} \cdot \boldsymbol{v}, \qquad \boldsymbol{p}' = \boldsymbol{p}.$$

We see that the covectorial transformation rule is different from the vectorial transformation rule.

6.7 Tensorial splitting and transformation rules

6.7.1 Splitting

In a number of physical theories – e.g. in electromagnetism – not only vectors and covectors but various tensors appear, too. The mathematical supplement helps the reader to be familiar with tensors.

The inertial observer$\boldsymbol{u}$ splits the various tensors, i.e. the elements of $\mathbf{M} \otimes \mathbf{M}$, $\mathbf{M} \otimes \mathbf{M}^*$, $\mathbf{M}^* \otimes \mathbf{M}$ and $\mathbf{M}^* \otimes \mathbf{M}^*$. These tenors can be considered as liner maps $\mathbf{M}^* \to \mathbf{M}$, $\mathbf{M} \to \mathbf{M}$, $\mathbf{M}^* \to \mathbf{M}^*$ and $\mathbf{M} \to \mathbf{M}^*$, respectively.

The split form of $\boldsymbol{G} \in \mathbf{M} \otimes \mathbf{M}$ i.e. the linear map $\boldsymbol{G} : \mathbf{M}^* \to \mathbf{M}$ is the tensor

$$\boldsymbol{h}_{\boldsymbol{u}} \cdot \boldsymbol{G} \cdot \boldsymbol{h}_{\boldsymbol{u}}^* : (\mathbb{I} \times \mathbf{E})^* \to (\mathbb{I} \times \mathbf{E}).$$

Since $\boldsymbol{h}_{\boldsymbol{u}} = (\boldsymbol{\tau}, \boldsymbol{\sigma}_{\boldsymbol{u}})$ and $\boldsymbol{h}_{\boldsymbol{u}}^* = (\boldsymbol{\tau}^*, \boldsymbol{\sigma}_{\boldsymbol{u}}^*)$, moreover $(\mathbb{I} \times \mathbf{E})^* = \mathbb{I}^* \times \mathbf{E}^*$, the split tensor can be written in a matrix form,

$$\boldsymbol{h}_{\boldsymbol{u}} \cdot \boldsymbol{G} \cdot \boldsymbol{h}_{\boldsymbol{u}}^* = = \begin{pmatrix} \boldsymbol{\tau} \cdot \boldsymbol{G} \cdot \boldsymbol{\tau}^* & \boldsymbol{\tau} \cdot \boldsymbol{G} \cdot \boldsymbol{\sigma}_{\boldsymbol{u}}^* \\ \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{G} \cdot \boldsymbol{\tau}^* & \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{G} \cdot \boldsymbol{\sigma}_{\boldsymbol{u}}^* \end{pmatrix},$$

whose components, explicitly, are

$$\boldsymbol{\tau}\cdot\boldsymbol{G}\cdot\boldsymbol{\sigma}_{\boldsymbol{u}}^{*}=\boldsymbol{\tau}\cdot\boldsymbol{G}-\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{G}\cdot\boldsymbol{\tau}^{*}),\quad \boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{G}\cdot\boldsymbol{\tau}^{*}=\boldsymbol{G}\cdot\boldsymbol{\tau}^{*}-\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{G}\cdot\boldsymbol{\tau}^{*}),$$

$$\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{G}\cdot\boldsymbol{\sigma}_{\boldsymbol{u}}^{*}=\boldsymbol{G}-\boldsymbol{u}\otimes(\boldsymbol{\tau}\cdot\boldsymbol{G})-(\boldsymbol{G}\cdot\boldsymbol{\tau}^{*})\otimes\boldsymbol{u}+\boldsymbol{u}\otimes\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{G}\cdot\boldsymbol{\tau}^{*}).$$

We obtain similarly, using the formula $\boldsymbol{r}_{\boldsymbol{u}}^{*}=(\boldsymbol{u},\boldsymbol{i})$, the $\boldsymbol{u}$-split form of $\boldsymbol{L}\in\mathbf{M}\otimes\mathbf{M}^{*}$:

$$\boldsymbol{h}_{\boldsymbol{u}}\cdot\boldsymbol{L}\cdot\boldsymbol{r}_{\boldsymbol{u}}^{*}=\begin{pmatrix}\boldsymbol{\tau}\cdot\boldsymbol{L}\cdot\boldsymbol{u} & \boldsymbol{\tau}\cdot\boldsymbol{L}\cdot\boldsymbol{i}\\ \boldsymbol{L}\cdot\boldsymbol{u}-\boldsymbol{u}(\boldsymbol{\tau}\cdot\boldsymbol{L}\cdot\boldsymbol{u}) & \boldsymbol{L}\cdot\boldsymbol{i}-\boldsymbol{u}\otimes(\boldsymbol{\tau}\cdot\boldsymbol{L}\cdot\boldsymbol{i})\end{pmatrix}. \tag{IV15}$$

The $\boldsymbol{u}$-split form of $\boldsymbol{P}\in\mathbf{M}^{*}\otimes\mathbf{M}$ is

$$\boldsymbol{r}_{\boldsymbol{u}}\cdot\boldsymbol{P}\cdot\boldsymbol{h}_{\boldsymbol{u}}^{*}=\begin{pmatrix}\boldsymbol{u}\cdot\boldsymbol{P}\cdot\boldsymbol{\tau}^{*} & \boldsymbol{u}\cdot\boldsymbol{P}\cdot\boldsymbol{\sigma}_{\boldsymbol{u}}^{*}\\ \boldsymbol{i}^{*}\boldsymbol{P}\cdot\boldsymbol{\tau}^{*} & \boldsymbol{i}^{*}\boldsymbol{P}\cdot\boldsymbol{\sigma}_{\boldsymbol{u}}^{*}\end{pmatrix}.$$

The $\boldsymbol{u}$-split form of $\boldsymbol{F}\in\mathbf{M}^{*}\otimes\mathbf{M}^{*}$ is

$$\boldsymbol{r}_{\boldsymbol{u}}\cdot\boldsymbol{F}\cdot\boldsymbol{r}_{\boldsymbol{u}}^{*}=\begin{pmatrix}\boldsymbol{u}\cdot\boldsymbol{F}\cdot\boldsymbol{u} & \boldsymbol{u}\cdot\boldsymbol{F}\cdot\boldsymbol{i}\\ \boldsymbol{i}^{*}\cdot\boldsymbol{F}\cdot\boldsymbol{u} & \boldsymbol{i}^{*}\cdot\boldsymbol{F}\cdot\boldsymbol{i}\end{pmatrix}.$$

The antisymmetric elements of $\mathbf{M}\otimes\mathbf{M}$ and $\mathbf{M}^{*}\otimes\mathbf{M}^{*}$ are particularly interesting.

If $\boldsymbol{G}$ is an antisymmetric tensor i.e. $\boldsymbol{G}=-\boldsymbol{G}^{*}$ then $\boldsymbol{\tau}\cdot\boldsymbol{G}=-\boldsymbol{G}\cdot\boldsymbol{\tau}^{*}$ and $\boldsymbol{\tau}\cdot\boldsymbol{G}\cdot\boldsymbol{\tau}^{*}=0$, that is why it has the $\boldsymbol{u}$-split form

$$\begin{pmatrix}0 & \boldsymbol{\tau}\cdot\boldsymbol{G}\\ -\boldsymbol{\tau}\cdot\boldsymbol{G} & \boldsymbol{G}-\boldsymbol{u}\wedge(\boldsymbol{\tau}\cdot\boldsymbol{G})\end{pmatrix};$$

here the two 'lower' components determine the other ones, therefore we refer to the split form as

$$(\!(-\boldsymbol{\tau}\cdot\boldsymbol{G},\boldsymbol{G}-\boldsymbol{u}\wedge(\boldsymbol{\tau}\cdot\boldsymbol{G}))\!)\in(\mathbf{E}\otimes\mathbb{I})\times(\mathbf{E}\wedge\mathbf{E}).$$

The first one is called the **timelike component** of $\boldsymbol{G}$ – which is independent of $\boldsymbol{u}$ –, the second one is called the**$\boldsymbol{u}$-spacelike component**.

It is a simple fact that if the $\boldsymbol{u}$-split form of $\boldsymbol{G}$ is $(\!(\boldsymbol{D},\boldsymbol{H}_{\boldsymbol{u}})\!)$ then

$$\boldsymbol{G}=\boldsymbol{H}_{\boldsymbol{u}}-\boldsymbol{u}\wedge\boldsymbol{D}.$$

Similarly, if $\boldsymbol{F}$ is an antisymmetric cotensor i.e. $\boldsymbol{F}=-\boldsymbol{F}^{*}$ then $\boldsymbol{u}^{*}\cdot\boldsymbol{F}=-\boldsymbol{F}\cdot\boldsymbol{u}$ and $\boldsymbol{u}^{*}\cdot\boldsymbol{F}\cdot\boldsymbol{u}=0$, that is why it has the $\boldsymbol{u}$-split form

$$\begin{pmatrix}0 & -\boldsymbol{i}^{*}\cdot\boldsymbol{F}\cdot\boldsymbol{u}\\ \boldsymbol{i}^{*}\cdot\boldsymbol{F}\cdot\boldsymbol{u} & \boldsymbol{i}^{*}\cdot\boldsymbol{F}\cdot\boldsymbol{i}\end{pmatrix}.$$

As previously, only the 'lower' components will be considered for referring to the split form:

$$(\!(\boldsymbol{i}^* \cdot \boldsymbol{F} \cdot \boldsymbol{u}, \boldsymbol{i}^* \cdot \boldsymbol{F} \cdot \boldsymbol{i})\!) \in (\mathbf{E}^* \otimes \mathbb{I}^*) \times (\mathbf{E}^* \wedge \mathbf{E}^*).$$

The first one is called the **$\boldsymbol{u}$-timelike component** of $\boldsymbol{F}$, the second one is called the**spacelike component** which is independent of $\boldsymbol{u}$..

Using the simple facts $\boldsymbol{\sigma_u} \cdot \boldsymbol{i} = \boldsymbol{i}$ and $\boldsymbol{\tau} \cdot \boldsymbol{i} = 0$, we find that if the $\boldsymbol{u}$-split form of $\boldsymbol{F}$ is $(\!(\boldsymbol{E_u}, \boldsymbol{B})\!)$ then

$$\boldsymbol{F} = \boldsymbol{\sigma}_{\boldsymbol{u}}^* \cdot \boldsymbol{B} \cdot \boldsymbol{\sigma_u} - \boldsymbol{\tau}^* \wedge \boldsymbol{\sigma}_{\boldsymbol{u}}^* \cdot \boldsymbol{E_u}$$

where $\boldsymbol{\tau}^* : \mathbf{M}^* \to \mathbb{I}^*$ is considered, according to the usual identification, an element of $\frac{\mathbf{M}}{\mathbb{I}}$.

6.7.2 Transformation rules

Comparing the tensorial splitting due to different observers, we get the tensorial transformation rules. We treat only the formulae concerning antisymmetric tensors and cotensors.

If $((\boldsymbol{D}, \boldsymbol{H}))$ and $((\boldsymbol{D}', \boldsymbol{H}'))$ are the $\boldsymbol{u}$-split form and the $\boldsymbol{u}'$-split form, respectively, of an antisymmetric tensor, then $((\boldsymbol{D}', \boldsymbol{H}')) = \boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{u}} \cdot ((\boldsymbol{D}, \boldsymbol{H})) \cdot \boldsymbol{H}^*_{\boldsymbol{u}'\boldsymbol{u}}$; in matrix form and with the notation $\mathbf{v} = \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$

$$\begin{pmatrix} 1 & 0 \\ -\mathbf{v} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{0} & -\boldsymbol{D} \\ \boldsymbol{D} & \boldsymbol{H} \end{pmatrix} \begin{pmatrix} 1 & -\mathbf{v} \\ 0 & \mathbf{1} \end{pmatrix} = \begin{pmatrix} 0 & -\boldsymbol{D} \\ \boldsymbol{D} & \mathbf{v} \wedge \boldsymbol{D} + \boldsymbol{H} \end{pmatrix},$$

thus

$$\boldsymbol{D}' = \boldsymbol{D}, \qquad \boldsymbol{H}' = \mathbf{v} \wedge \boldsymbol{D} + \boldsymbol{H}.$$

If $((\boldsymbol{E}, \boldsymbol{B}))$ and $((\boldsymbol{E}', \boldsymbol{B}'))$ are the $\boldsymbol{u}$-split form and the $\boldsymbol{u}'$-split form, respectively, of the an antisymmetric cotensor, then in matrix form

$$\begin{pmatrix} 1 & \mathbf{v} \\ 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{0} & -\boldsymbol{E} \\ \boldsymbol{E} & \boldsymbol{B} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \mathbf{v} & \mathbf{1} \end{pmatrix} = \begin{pmatrix} 0 & -(\boldsymbol{E} + \boldsymbol{B} \cdot \mathbf{v}) \\ \boldsymbol{E} + \boldsymbol{B} \cdot \mathbf{v} & \boldsymbol{B} \end{pmatrix},$$

thus

$$\boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{B} \cdot \mathbf{v}, \qquad \boldsymbol{B}' = \boldsymbol{B}.$$

6.8 Splitting of spacetime and transformation rules

6.8.1 Splitting

According to subsection 3.2.3, a uniform synchronization and an inertial observer together form an inertial frame. Since the resent model admits only a single synchronization, an inertial observer determines uniquely an inertial frame. That is why now we can omit referring the the synchronization and we say **inertial observer instead of inertial frame**.

Thus, an inertial observer $\boldsymbol{u}$ characterizes the occurrences by giving when and where they happen: it **splits** spacetime into time and space; to a world point x it assigns the corresponding absolute time point $\tau(x) = x + \mathbf{E}$ and the corresponding $\boldsymbol{u}$-space point $\sigma_{\boldsymbol{u}}(x) = x + \mathbb{I}\boldsymbol{u}$. The splitting

$$h_{\boldsymbol{u}} : \mathtt{M} \to \mathtt{I} \times \mathtt{E}_{\boldsymbol{u}}, \qquad x \mapsto \big(\tau(x), \sigma_{\boldsymbol{u}}(x)\big). \tag{IV16}$$

is affine bijection over the vectorial splitting $\boldsymbol{h}_{\boldsymbol{u}}$ as it is well seen from equalities (IV8), (IV12) and (IV13). The inverse of this splitting – which gives the occurrence corresponding to a time point and a $\boldsymbol{u}$-space point – is

$$h_{\boldsymbol{u}}^{-1} : \mathtt{I} \times \mathtt{E}_{\boldsymbol{u}} \to \mathtt{M}, \qquad (t, q) \mapsto t \cap q.$$

Instead of affine spaces it is often more suitable to deal with the underlying vector spaces; therefore an inertial observer – corresponding to the everyday usage when timepoints are represented by time intervals that passed from a given time point and space points are represented by vectors from an origin –, choosing an 'initial' time point t_o and a $\boldsymbol{u}$-'origin' q_o, vectorizes time and $\boldsymbol{u}$-space by the assignment

$$\mathtt{I} \times \mathtt{E}_{\boldsymbol{u}} \to \mathbb{I} \times \mathbf{E}, \qquad (t, q) \mapsto (t - t_o, q - q_o).$$

Choosing a t_o and a q_o is equivalent to choosing a 'spacetime origin' o: $o = t_o \cap q_o$, $t_o = \tau(o) = o + \mathbf{E}$, $q_o = \sigma_{\boldsymbol{u}}(o) = o + \mathbb{I}\boldsymbol{u}$. Then it is a simple fact that the $\boldsymbol{u}$-splitting of spacetime followed by the vectorization of time and $\boldsymbol{u}$-space gives the **vectorized splitting of spacetime** by o and $\boldsymbol{u}$:

$$h_{\boldsymbol{u},o} : \mathtt{M} \to \mathbb{I} \times \mathbf{E}, \qquad x \mapsto \boldsymbol{h}_{\boldsymbol{u}} \cdot (x - o) = \big(\boldsymbol{\tau} \cdot (x - o), \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (x - o)\big). \tag{IV17}$$

It has the inverse

$$\mathbb{I} \times \mathbf{E} \to \mathtt{M}, \qquad (\boldsymbol{t}, \boldsymbol{q}) \mapsto o + \boldsymbol{t}\boldsymbol{u} + \boldsymbol{q}.$$

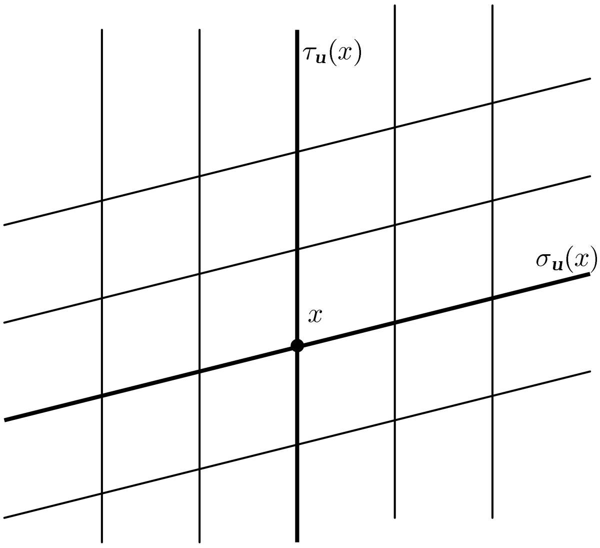

Figure 6.6 Splitting of spacetime

6.8.2 Transformation rules

The transformation rule between spacetime splittings would be $h_{\mathbf{u}'} \circ h_{\mathbf{u}}^{-1} : I \times E_{\boldsymbol{u}} \to \mathtt{I} \times E_{\mathbf{u}'}$. This is not suitable for comparison because the domain and the range are different sets. That is why we compare always vectorial splittings where the range and domain is the same $\mathbb{I} \times \mathbf{E}$.

Let $(\boldsymbol{t}, \boldsymbol{q})$ and $(\boldsymbol{t}', \boldsymbol{q}')$ the vectorized split form of the same world point, due to $(\boldsymbol{u}, o)$ and $(\mathbf{u}', o')$, respectively. Then

$$\begin{aligned}(\boldsymbol{t}', \boldsymbol{q}') =& h_{\mathbf{u}',o'}\big(h^{-1}_{\mathbf{u},o}(\boldsymbol{t}, \boldsymbol{q})\big) = \\ =& h_{\mathbf{u}',o'}(o + \boldsymbol{t}\boldsymbol{u} + \boldsymbol{q}) = \big(\boldsymbol{\tau} \cdot (\boldsymbol{t}\boldsymbol{u} + \boldsymbol{q} + o - o'),\ \boldsymbol{\sigma}_{\mathbf{u}'} \cdot (o - o' + \boldsymbol{t}\boldsymbol{u} + \boldsymbol{q})\big) = \\ =& \big(\boldsymbol{t} + \boldsymbol{\tau} \cdot (o - o'), \boldsymbol{q} - \boldsymbol{t}\mathbf{v}_{\mathbf{u}'\mathbf{u}} + \boldsymbol{\sigma}_{\mathbf{u}'} \cdot (o - o')\big),\end{aligned}$$

that is, with notations $\boldsymbol{t}_o := \boldsymbol{\tau} \cdot (o - o')$ and $\boldsymbol{q}_o = \boldsymbol{\sigma}_{\mathbf{u}'} \cdot (o - o')$

$$\boldsymbol{t}' = \boldsymbol{t} + \boldsymbol{t}_o, \qquad \boldsymbol{q}' = \boldsymbol{q} - \boldsymbol{t}\mathbf{v}_{\mathbf{u}'\boldsymbol{u}} + \boldsymbol{q}_o$$

which is the well known **inhomogeneous Galilean transformation rule**.

This equals the vectorial Galilean transformation rule if the observers choose the same spacetime origin ($o' = o$).

Note that

– the inhomogeneous Galilean transformation rule is an affine map, it serves for comparing the splitting of spacetime,

– the Galilean transformation rule is a linear map, it serves for comparing the splitting of spacetime vectors.

6.9 Transformations and transformation rules

Let us consider $\boldsymbol{u}$-split form $\boldsymbol{h_u} \cdot \boldsymbol{L} \cdot \boldsymbol{h_u}^{-1}$ (see IV15) of the special Galilei transformation $\boldsymbol{L} = \mathbf{1} + \boldsymbol{v} \otimes \boldsymbol{\tau}$ (see (IV9)). $\boldsymbol{L} \cdot \boldsymbol{u} = \boldsymbol{u} + \boldsymbol{v}$ and $\boldsymbol{L} \cdot \boldsymbol{i}$ is the identity of $\mathbf{E}$, thus $\boldsymbol{\tau} \cdot \boldsymbol{L} \cdot \boldsymbol{u} = 1$ and $\boldsymbol{\tau} \cdot \boldsymbol{L} \cdot \boldsymbol{i} = 0$; consequently, in a matrix form (since $\boldsymbol{L} \in \mathbf{M} \otimes \mathbf{M}^*$) we have

$$\begin{pmatrix} 1 & 0 \\ \boldsymbol{v} & \mathbf{1} \end{pmatrix}. \tag{IV18}$$

Introducing the notation $\boldsymbol{u}' := \boldsymbol{u} + \boldsymbol{v}$, we obtain $\boldsymbol{v} = \boldsymbol{v}_{\boldsymbol{u}'\boldsymbol{u}} = -\boldsymbol{v}_{\boldsymbol{u}\boldsymbol{u}'}$. Thus, the result is just the Galilean transformation rule from $\boldsymbol{u}'$ to $\boldsymbol{u}$ (see 6.6.2). Since $\boldsymbol{L} = \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}$, we can state that

$$\boldsymbol{h}_{\boldsymbol{u}\boldsymbol{u}'} = \boldsymbol{h_u} \cdot \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{h_u}^{-1}$$

i.e. the vectorial transformation rule from $\boldsymbol{u}'$ to $\boldsymbol{u}$ is the $\boldsymbol{u}$-split form of the boost from $\boldsymbol{u}'$ to $\boldsymbol{u}$.

The vectorized splitting of spacetime converts the special Noether transformations into the inhomogeneus Galilean transzformations.

Though there is some relation between special (Noether) Galilei transformations and (inhomogeneous) Galilean transformation rules, they differ essentially both from a mathematical and a physical point of view. The **transformations** are maps $\mathtt{M} \to \mathtt{M}$ and $\mathbf{M} \to \mathbf{M}$, repectively, that reflects the structure of spacetime (symmetries), whereas the **transformation rules** are maps $\mathbb{I} \times \mathbf{E} \to \mathbb{I} \times \mathbf{E}$ that compare different splittings.

In usual treatments based on coordinates the transformations (spacetime symmetries) and the transformation rules are confused because all of them are maps in $\mathbb{R} \times \mathbb{R}^3$. This often causes conceptual errors, e.g. when one says that a symmetry of spacetime is that the inertial frames are physically equivalent and one explains the symmetries of a physical system by showing that a quantity transforms in a convenient way when changing reference frames[2].

[2] An excellent example for such a confusion can be found in L.D.Landau-E.M.Lifshitz: *Mechanics*, (Addison-Wesley, 1960). In deducing the Lagrange function of a free material

We can find hints in the literature to that two different objects appear in the same form, when one distinguishes between active transformations (which correspond to spacetime symmetries) and passive transformations (which correspond to transformation rules).

6.10 Coordinatizations

Time intervals are usually characterized by numbers giving them as a multiple of a time unit $\mathrm{s} \in \mathbb{I}^+$ (second); in formula, time intervals are coordinatized by

$$\mathbb{I} \to \mathbb{R}, \qquad \boldsymbol{t} \mapsto \frac{\boldsymbol{t}}{\mathrm{s}}.$$

Lenghts (distances) are usually characterized by numbers giving them as a multiple of a length unit $\mathrm{m} \in \mathbb{D}^+$ (meter); in formula, lengths are coordinatized by

$$\mathbb{D} \to \mathbb{R}, \qquad \boldsymbol{d} \mapsto \frac{\boldsymbol{d}}{\mathrm{m}}.$$

The space vectors are usually characterized by triplets of numbers in such a way that a laength unit m is chosen, further coordinate axes determined by three 'right handed' orthogonal vectors of length m. In formula, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ is a positively oriented orthogonal basis in $\mathbf{E}$, every basis element having the length m, and we take the coordinates of vectors corresponding to that basis:

$$\mathbf{E} \to \mathbb{R}^3, \qquad \boldsymbol{q} \mapsto \left(\frac{\boldsymbol{e}_1 \cdot \boldsymbol{q}}{\mathrm{m}^2}, \frac{\boldsymbol{e}_2 \cdot \boldsymbol{q}}{\mathrm{m}^2}, \frac{\boldsymbol{e}_3 \cdot \boldsymbol{q}}{\mathrm{m}^2} \right).$$

The inertial observer $\boldsymbol{u}$ coordinatizes the world vectors in such a way that
– it splits $\mathbf{M}$ into $\mathbb{I} \times \mathbf{E}$ according to 6.6.1,
– it coordinatizes $\mathbb{I}$ by s,
– it coordinatizes $\mathbf{E}$ by $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$.

In other words, with the notation $\boldsymbol{e}_0 := \mathrm{s}\boldsymbol{u}$, the coordinatization of world vectors is the linear bijection

$$\mathbf{M} \to \mathbb{R}^4, \qquad \boldsymbol{x} \mapsto \{\text{coordinates of } \boldsymbol{x} \text{ in the basis } \boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}.$$

point, that book says on one hand that according to the homogeneity of time and space the properties of a closed physical system do not change when the system is translated parallel to itself (thus, spacetime translations are considered – correctly – as symmetries of the physical system), and on the other hand that no essential change enters the description if one changes to another coordinate system moving uniformly (thus, a transformation rule is considered – incorrectly – as a symmetry of the physical system; correctly it should be said that the properties of a closed physical system do not change when the system is put in uniform motion parallel to itself)

Thus, if $(\xi^0, \xi^1, \xi^2, \xi^3)$ are the coordinates of the vector $\boldsymbol{x}$, then $\boldsymbol{x} = \sum_{i=0}^{3} \xi^i \boldsymbol{e}_i$ and it is easy to see that

$$\xi^0 = \frac{\boldsymbol{\tau} \cdot \mathbf{x}}{\mathrm{s}}, \quad \xi^i = \frac{\boldsymbol{e}_i \cdot (\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathbf{x})}{\mathrm{m}^2}, \quad (i = 1, 2, 3).$$

Note that the indices of a vector are usually written as superscripts.

To coordinatize spacetime, the inertial observer vectorizes spacetime by a chosen spacetime origin o and then applies the previous procedure. The result is

$$\mathtt{M} \to \mathbb{R}^4, \qquad x \mapsto \{\text{coordinates of } x - o \text{ in the basis } \boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}.$$

The coordinatization can be described in a physically more transparent way as follows. The observer

– splits $\mathtt{M}$ to time and $\boldsymbol{u}$-space according to IV17,

– vectorizes time and $\boldsymbol{u}$-space by choosing a 'time origin' t_o and a '$\boldsymbol{u}$-space origin'q_o,

which is equivalent to the vectorized splitting of spacetime with the aid of the 'world origin' $o := t_o \cap q_o$ (see IV17), and then

– coordinatizes $\mathbb{I}$ and $\mathbf{E}$ as previously given.

According to these formulae, an **inertial coordinate system** is $(o, \mathrm{s}, \mathrm{m}, \boldsymbol{u}, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$, where o is a world point, s is a time unit, m is a distance unit, $\boldsymbol{u}$ is an inertial observer, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ is a positively oriented orthogonal basis, normed to m, in $\mathbf{E}$.

Returning to Paragraph 6.3.2, we can state that the coordinatization of spacetime by an inertial coordinate system is just an isomorphism to the arithmetic spacetime model. We see very well, how many arbitrary objects are hidden in the arithmetic spacetime model: a spacetime origin, a time unit, a lenght unit, an inertial observerand a spacelike basis.

The coordinate system represents covectors and various tensors by coordinates, too. For instance, the coordinates of the covector $\boldsymbol{k}$ are

$$(\chi_i := \boldsymbol{k} \cdot \boldsymbol{e}_i \mid i = 0, 1, 2, 3).$$

Note that the indices of a covector are usually written as subscripts. This usage serves for distinguishing the coordinates of covectors from those of vectors.

The coordinates of a tensor $\boldsymbol{G} \in \mathbf{M} \otimes \mathbf{M}$ are $(i, k = 1, 2, 3)$:

$$G^{00} := \frac{\boldsymbol{\tau} \cdot \boldsymbol{G} \cdot \boldsymbol{\tau}^*}{\mathrm{s}^2}, \quad G^{0k} = \frac{\boldsymbol{\tau} \cdot \boldsymbol{G} \cdot \boldsymbol{\sigma}_{\boldsymbol{u}}^* \cdot \boldsymbol{e}_k}{\mathrm{sm}}, \quad G^{k0} = \frac{\boldsymbol{e}_k \cdot \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{G} \cdot \boldsymbol{\tau}^*}{\mathrm{sm}},$$

$$G^{ik} = \frac{\boldsymbol{e}_i \cdot \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{G} \cdot \boldsymbol{\sigma}^*_{\boldsymbol{u}} \cdot \boldsymbol{e}_k}{\mathrm{m}^2}$$

The coordinates of a cotensor $\boldsymbol{F} \in \mathbf{M}^* \otimes \mathbf{M}^*$ are $(i, k = 0, 1, 2, 3)$:

$$F_{ik} = \boldsymbol{e}_i \cdot \boldsymbol{F} \cdot \boldsymbol{e}_k.$$

6.11 Derivatives

Let us consider a differentiable function $f : \mathtt{M} \to \mathbb{R}$. It derivative at a point x is a linear map $\mathrm{D}f(x) : \mathbf{M} \to \mathbb{R}$ i.e. it is an element of $\mathbf{M}^*$ (see the mathematical supplement). This covector is split by the inertial observer into the $\boldsymbol{u}$-timelike component

$$(\mathrm{D}f(x)) \cdot \boldsymbol{u} =: \mathrm{D}_{\boldsymbol{u}} f(x)$$

and

$$(\mathrm{D}f(x))|_{\mathbf{E}} =: \nabla f(x)$$

spacelike component which have a direct meaning as follows.

Let us restrict f onto the straight line passing through x and directed by $\boldsymbol{u}$, i.e. let us consider the function $\mathbb{I} \to \mathbb{R}$, $\boldsymbol{t} \mapsto f(x + \boldsymbol{t}\boldsymbol{u})$. The derivative at zero of this function – according to the rule of differentiation of composite functions – is $\mathrm{D}f(x) \cdot \boldsymbol{u}$.

Let us restrict f onto the hyperplane passing through x and directed by $\boldsymbol{u}$, i.e. let us consider the function $\mathbf{E} \to \mathbb{R}$, $\boldsymbol{q} \mapsto f(x + \boldsymbol{q})$. The derivative at zero of this function – according to the rule of differentiation of composite functions – is $(\mathrm{D}f(x))|_{\mathbf{E}}$.

That is why $\mathrm{D}_{\boldsymbol{u}} f$ and ∇f are called the **$\boldsymbol{u}$-timelike derivative** and the **spacelike derivative** of f, respectively.

If spacetime is coordinatized in the usual way (see Subsection 6.10), the function f is given in the form $\mathbb{R}^4 \to \mathbb{R}$, $(\xi^0, \xi^1, \xi^2, \xi^3) \mapsto \hat{f}(\xi^0, \xi^1, \xi^2, \xi^3) := f(o + \sum_{k=0}^{3} \xi^k \boldsymbol{e}_k)$. Then $\partial_k \hat{f}(\xi^0, \xi^1, \xi^2, \xi^3) = \mathrm{D}f(o + \sum_{k=0}^{3} \xi^k \boldsymbol{e}_k) \cdot \boldsymbol{e}_k$, i.e. the partial derivatives are the coordinates of $\mathrm{D}f$. For the sake of simplicity, admitting a little abuse of notation, let us omit the 'hat'; the we can write

$$\mathrm{D}f \quad \text{in coordinates is} \quad \partial_k f \quad (k = 0, 1, 2, 3).$$

It is evident that the zeroth partial derivative is the coordinatized form of the $\boldsymbol{u}$-timelike derivative, the other three partial derivatives constitute the coordinatized form of the spacelike derivative.

In general, the differentiation D can be conceived as a symbolic covector whose $\boldsymbol{u}$-split form is $(\mathrm{D}_{\boldsymbol{u}}, \nabla) := (\boldsymbol{u} \cdot \mathrm{D}, \boldsymbol{i}^* \cdot \mathrm{D})$.

For instance, the values of the derivative $\mathrm{D}\mathcal{J}$ of a vector field $\mathcal{J} : \mathtt{M} \to \mathbf{M}$ are elements of $\mathbf{M} \otimes \mathbf{M}^*$. As said in the mathematical supplement, it is suitable to use the transpose of the derivative which will be denoted by $\mathrm{D} \otimes \mathcal{J}$; it has values in $\mathbf{M}^* \otimes \mathbf{M}$. Its $\boldsymbol{u}$-split form is

$$\begin{pmatrix} \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{\tau} \cdot \mathcal{J}) & \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathcal{J}) \\ \nabla(\boldsymbol{\tau} \cdot \mathcal{J}) & \nabla \otimes (\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathcal{J}) \end{pmatrix},$$

in coordinates $\partial_i J^k$ $(i, k = 0, 1, 2, 3)$.

Taking the trace of $(\mathrm{D} \otimes \mathcal{J})(x)$ (see the mathematical supplement), we define the **divergence** of $\mathcal{J}$:

$$(\mathrm{D} \cdot \mathcal{J})(x) := \mathrm{Tr}(\mathrm{D} \otimes \mathcal{J}(x)),$$

in a split form

$$\mathrm{D} \cdot \mathcal{J} = \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{\tau} \cdot \mathcal{J}) + \nabla \cdot (\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathcal{J}),$$

which reads in coordinates as $\sum_{k=0}^{3} \partial_k J^k$.

For a better survey, let us introduce the symbol $\sim$ which will refer to what a split form and a coordinatized form has the derivative in question. Cooordinates always run through the values $0, 1, 2, 3$ and a summation is has to be effected for equal subscripts and superscripts (Einstein summation rule). Thus, if

$$\mathcal{J} \quad \sim \quad (\rho, \boldsymbol{j}_{\boldsymbol{u}}) \quad \sim \quad J^k,$$

then

$$\mathrm{D} \cdot \mathcal{J} \quad \sim \quad \mathrm{D}_{\boldsymbol{u}}\rho + \nabla \cdot \boldsymbol{j}_{\boldsymbol{u}} \quad \sim \quad \partial_k J^k.$$

The (transpose of) the derivative $\mathrm{D} \otimes \boldsymbol{K}$ of a covector field $\boldsymbol{K} : \mathtt{M} \to \mathbf{M}^*$ has values in $\mathbf{M}^* \otimes \mathbf{M}^*$; its split form is

$$\begin{pmatrix} \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{u} \cdot \boldsymbol{K}) & \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{i}^* \cdot \boldsymbol{K}) \\ \nabla(\boldsymbol{u} \cdot \boldsymbol{K}) & \nabla \otimes (\boldsymbol{i}^* \cdot \boldsymbol{K}) \end{pmatrix},$$

in coordinates $\partial_i K_k$ $(i, k = 0, 1, 2, 3)$.

Taking the antisymmetric part of $(\mathrm{D} \otimes \boldsymbol{K})(x)$, we define the **exterior derivative** of $\boldsymbol{K}$:

$$\mathrm{D} \wedge \boldsymbol{K} := \mathrm{D} \otimes \boldsymbol{K} - (\mathrm{D} \otimes \boldsymbol{K})^*,$$

in a split form

$$\big(\big(\nabla(\boldsymbol{u} \cdot \boldsymbol{K}) - \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{i}^* \cdot \boldsymbol{K}), \nabla \wedge (\boldsymbol{i}^* \cdot \boldsymbol{K})\big)\big),$$

which reads in coordinates as $\partial_k K_i - \partial_i K_k$.

With the previous survey: if

$$\boldsymbol{K} \quad \sim \quad (-V_{\boldsymbol{u}}, \boldsymbol{A}) \quad \sim \quad K_k,$$

then

$$\mathrm{D} \wedge \boldsymbol{K} \quad \sim \quad ((-\nabla V_{\boldsymbol{u}} - \mathrm{D}_{\boldsymbol{u}} \cdot \boldsymbol{A}, \nabla \wedge \boldsymbol{A})) \quad \sim \quad \partial_i K_k - \partial_k K_i.$$

Note that

– **a vector field has divergence and has not exterior derivative,**

– **a covector field has exterior derivative and has not divergence.**

Similarly, an antisymmetric tensor field $\boldsymbol{G} : \mathtt{M} \to \mathbf{M} \wedge \mathbf{M}$ has divergence $\mathrm{D} \cdot \boldsymbol{G}$ which takes values in $\mathbf{M}$; if

$$\boldsymbol{G} \quad \sim \quad ((\boldsymbol{D}, \boldsymbol{H}_{\boldsymbol{u}})) \quad \sim \quad G^{ik},$$

then

$$\mathrm{D} \cdot \boldsymbol{G} \quad \sim \quad (\nabla \cdot \boldsymbol{D}, -\mathrm{D}_{\boldsymbol{u}} \boldsymbol{D} + \nabla \cdot \boldsymbol{H}_{\boldsymbol{u}}) \quad \sim \quad \partial_i G^{ik}. \qquad \text{(IV19)}$$

An antisymmetric cotensor field $\boldsymbol{F} : \mathtt{M} \to \mathbf{M}^* \wedge \mathbf{M}^*$ has exterior derivative $\mathrm{D} \wedge \boldsymbol{F}$ which takes values in $\mathbf{M}^* \wedge \mathbf{M}^* \wedge \mathbf{M}^*$; if

$$\boldsymbol{F} \quad \sim \quad ((\boldsymbol{E}_{\boldsymbol{u}}, \boldsymbol{B})) \quad \sim \quad F_{ik},$$

then

$$\mathrm{D} \wedge \boldsymbol{F} \quad \sim \quad (((\nabla \wedge \boldsymbol{E} + \mathrm{D}_{\boldsymbol{u}} \boldsymbol{B}, \nabla \wedge \boldsymbol{B}))) \quad \sim \quad \partial_j F_{ik} + \partial_k F_{ji} + \partial_i F_{kj} \quad \text{(IV20)}$$

where the triple of brackets means that the two objects between them determine the whole antisymmetric tensor.

7 Fundamentals of point mechanics in the spacetime model

In this section we treat in the nonrelativistic spacetime model the notions and relations of the simplest physical theory. Classical mechanics is a well elaborated theory in the framework of coordinates, thus it offers us a good possibility to deepen our knowledge on spacetime.

7.1 World line functions

The history of a material point is a world line in spacetime. Such a world line – a curve – can be parameterized by absolute time in a natural way.

A world line C is parameterized in such a way that the world point $t \cap \mathrm{C}$ is assigned to the instant t (recall that t is a three dimensional affine hyperplane in spacetime). So we get the notion of the **world line function**: a (sufficiently many times) differentiable function which assigns to every instant t a world point whose absolute time point is t:

$$r : \mathtt{I} \to \mathtt{M}, \qquad \tau(r(t)) = t.$$

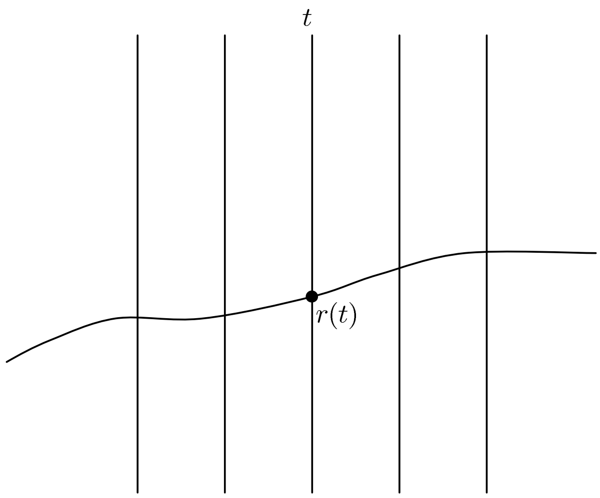

Figure 7.1 World line functiton

It is evident that a world line function and its range – a world line – determine each other uniquely.

The the derivative of such a function, $\dot{r}(t) \in \frac{\mathbf{M}}{\mathbb{I}}$ has the property

$$\boldsymbol{\tau} \cdot \dot{r}(t) = 1$$

i.e. $\dot{r}(t)$ is an absolute velocity. Then it is immediate that the absolute acceleration $\ddot{r}(t) \in \frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}}$ has the property

$$\boldsymbol{\tau} \cdot \ddot{r}(t) = 0,$$

i.e. it is absolute spacelike, an element of

$$\frac{\mathbf{E}}{\mathbb{I}\otimes\mathbb{I}}.$$

It is quite trivial that every element of $\frac{\mathbf{E}}{\mathbb{I}\otimes\mathbb{I}}$ can be the acceleration of a world line function (e.g. $\boldsymbol{a}$ is the acceleration of $t \mapsto o + (t-\tau(o))\boldsymbol{u} + \frac{(t-\tau(o))^2\boldsymbol{a}}{2}$ for all world points o), thus, the set of **absolute accelerations** is a three-dimensional Euclidean vector space. In contrast to absolute velocities,

- three is a zero absolute acceleration,
- an absolute acceleration has magnitude,
- the angle between two absolute accelerations makes sense.

7.2 Motions

An observer, in general, can describe motions only by choosing a synchronization because it must be given *when* the body in question is *where*. Motion is meaningful only with respect to a reference frame. In the nonrelativistic spacetime model a single synchronization exists, the absolute one; consequently, an observer determines a unique reference frame. Therefore in the following we say inertial observer, too, instead of inertial reference frame.

7.2.1 Relative velocities

An inertial observer $\boldsymbol{u}$ – i.e. an inertial frame because of the absolue synchronization – perceives the history of a material point (a world line) as a motion and describes it by assigning to an absolute instant t (a hyperplane directed by $\mathbf{E}$) the $\boldsymbol{u}$-space point (straight line directed by $\boldsymbol{u}$) which meets the material point (the world line) at t. Therefore, the $\boldsymbol{u}$-motion corresponding to the world line function r is

$$\mathtt{I} \to \mathtt{E}_{\boldsymbol{u}}, \quad t \mapsto r_{\boldsymbol{u}}(t) := \sigma_{\boldsymbol{u}}(r(t)).$$

The velocity of the material point relative to the observer is the time derivative of the $\boldsymbol{u}$-motion; using the formula (IV12), we get

$$\dot{r}_{\boldsymbol{u}}(t) = \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \dot{r}(t) = \dot{r}(t) - \boldsymbol{u}. \qquad \text{(IV21)}$$

This is the generalization of our result concerning inertial world lines: the **relative velocity** of the absolute velocity $\boldsymbol{u}'$ with respect $\boldsymbol{u}$ equals

$$\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} := \boldsymbol{u}' - \boldsymbol{u} \in \frac{\mathbf{E}}{\mathbb{I}}.$$

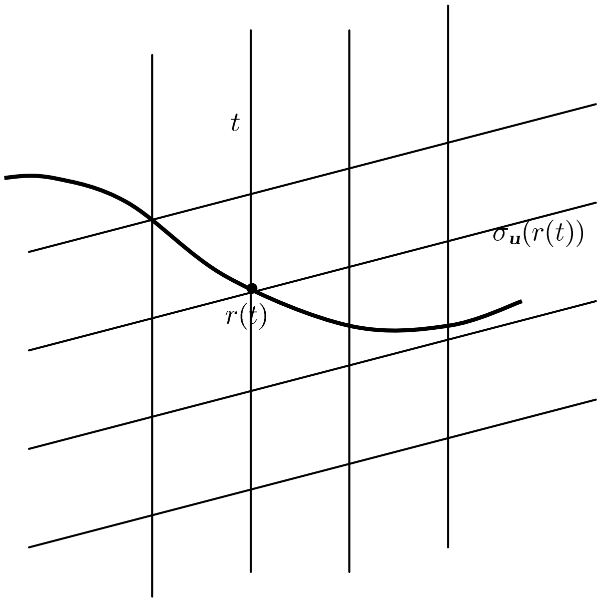

Figure 7.2 Description of motion

7.2.2 Relative accelerations

It is immediate fromIV21 that the **relative acceleration equals the absolute acceleration**:

$$\ddot{r}_{\boldsymbol{u}}(t) = \ddot{r}(t) \in \frac{\mathbf{E}}{\mathbb{I} \otimes \mathbb{I}}.$$

7.3 Absolute Newtonian equation

7.3.1 Measure line of masses

The unit of mass, kg in everyday practice is independent of the time unit s and the distance unit m. To treat the fundamentals of mechanics in the nonrelativistic spacetime model, we have to introduce the measure line of masses which seems an object outside of the model. Quantum mechanics, however, found out that the Planck constant $\hbar$, a qunatity distinguished by Nature, establishes a relation among the units of mass, time and distance:

$$\hbar = (1,05\dots)10^{-34}\frac{\mathrm{m}^2\mathrm{kg}}{\mathrm{s}}.$$

Thus, we can choose $\frac{\mathbb{I}}{\mathbb{D}\otimes\mathbb{D}}$ for the measure lie of masses in such a way that

$$\mathrm{kg} := (9,48\dots)10^{33}\frac{\mathrm{s}}{\mathrm{m}^2};$$

then the Planck constant becomes the real number 1.

It is evident that this choice makes easier the theoretical exposition because we not deal with a new measure line and so the formulae become simpler.

Of course, for practical application the everyday unit kg is more advantageous.

7.3.2 Absolute forces

We conceive that the absolute Newtonian equation has the form ,mass×absolute acceleration = absolute force' where the absolute force can depend on world points and absolute velocities.

Since the values of ,mass×absolute acceleration' are in the vector space $\frac{\mathbb{I}}{\mathbb{D}\otimes\mathbb{D}} \otimes \frac{\mathbf{E}}{\mathbb{I}\otimes\mathbb{I}} = \frac{\mathbf{E}}{\mathbb{D}\otimes\mathbb{D}\otimes\mathbb{I}}$, an absolute force is described by a function

$$\boldsymbol{f} : \mathtt{M} \times \mathtt{V}(1) \to \frac{\mathbf{E}}{\mathbb{D}\otimes\mathbb{D}\otimes\mathbb{I}} \equiv \frac{\mathbf{E}^*}{\mathbb{I}}.$$

Thus, the possible world line functions of a material point with mass m under the action of the absolute force $\boldsymbol{f}$ is determined by the **absolute Newtonian equation**

$$(x : \mathtt{I} \to \mathtt{M})? \qquad m\ddot{x} = \boldsymbol{f}(x, \dot{x}) \tag{IV22}$$

which is a second order differential equation.

To have a unique solution of that differential equation, the initial spacetime position and absolute velocity of the mass point must be given. That is why, if $t \mapsto r(t)$ is a solution, it is suitable to consider the pair $(r, \dot{r})$ the **process** of the mass point: its value at an arbitrary instant determines the whole function.

The **evolution space** of a mass point is the set in which the processes take values: $\mathtt{M} \times \mathtt{V}(1)$.

In what follows, we accept the notation (widely used in physics) that

– the elements of the evolution space are written in the form $(x, \dot{x})$ (note that then $\dot{x}$ is independent of x, it can denote an arbitrary absolute velocity),

– an arbitrary ('abstract') process – i.e. a time function – is denoted by $(x, \dot{x})$ as well,

– an actual process is denoted by $(r, \dot{r})$.

As it is known, the forces having a **potential** play a peculiar role. A potential is twice differentiable function

$$\boldsymbol{K} : \mathtt{M} \to \mathbf{M}^*.$$

The **field strength** corresponding to the potential is the exterior derivative of the potential:

$$\boldsymbol{F} := \mathrm{D} \wedge \boldsymbol{K},$$

and the corresponding force equals

$$\boldsymbol{f}(x, \dot{x}) := \boldsymbol{i}^* \cdot \boldsymbol{F}(x) \cdot \dot{x}.$$

Such a definition of potential and field strength will be explained in Paragraph 7.5.2.

7.4 Momenta

Let us consider a material point with mass m having the absolute velocity $\dot{x}$.

We accept the definition

'absolute momentum = mass×absolute velocity' $(m\dot{x})$

Since

'time derivative of absolute momentum = mass×absolute acceleration', $((m\dot{x})^{\cdot} = m\ddot{x})$,

the absolute Newtonian equation can be conceived in two ways:

'mass×absolute acceleration = absolute force',

'time derivative of absolute momentum = absolute force'.

Let us consider an inertial observer $\boldsymbol{u}$. According to the usual notions,

'$\boldsymbol{u}$-relative momentum = mass× $\boldsymbol{u}$-relative velocity' $(m\mathbf{v}_{\dot{x}\boldsymbol{u}})$,

for which we have, too, that

'$\boldsymbol{u}$-relative momentum = $\boldsymbol{u}$-spacelike component of absolute momentum' $(\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (m\dot{x}))$.

Further,

'time derivative of $\boldsymbol{u}$-relative momentum = mass× $\boldsymbol{u}$-relative acceleration',

thus the relative Newtonian equation, too, can be conceived in two ways:

'mass times $\boldsymbol{u}$-relative acceleration = $\boldsymbol{u}$-relative force',

'time derivative of $\boldsymbol{u}$-relative momentum = $\boldsymbol{u}$-relative force'.

This involved sequence of statements will be significant in comparing non-relativistic and relativistic dynamics.

7.5 Relative Newtonian equation

7.5.1 Definition

An inertial observer $\boldsymbol{u}$ perceives the world line function r as a motion, described by the function $r_{\boldsymbol{u}} : \mathtt{I} \to \mathtt{E}_{\boldsymbol{u}}$, $t \mapsto \sigma_{\boldsymbol{u}}(r(t))$ (see Paragraph 7.2.1. This motion satisfies a relative Newtonian equation which, according to the previous paragraph, can be written either in the form 'mass×$\boldsymbol{u}$-relative acceleration = $\boldsymbol{u}$-relative force' or in the form 'time derivative of $\boldsymbol{u}$-relative momentum = $\boldsymbol{u}$-relative force'.

The $\boldsymbol{u}$-relative force can depend on time points, $\boldsymbol{u}$-space points and $\boldsymbol{u}$-relative velocity; thus, the $\boldsymbol{u}$-relative Newtonian equation has the form

$$(q : \mathtt{I} \to \mathtt{E}_{\boldsymbol{u}})? \qquad m\ddot{q} = \boldsymbol{f}_{\boldsymbol{u}}(t, q, \dot{q}).$$

7.5.2 Relative forces

Since relative acceleration equals the absolute one, the left hand side of the relative Newtonian equation equals the left hand side of the absolute Newtonian equation, the relative force is **essentially equals** the absolute force; the only difference is that the variables of the absolute force have to be expressed by the relative variables. The world point determined by the time instant t (a hyperplane directed by $\mathbf{E}$) and the $\boldsymbol{u}$-space point q (a straight line directed by $\boldsymbol{u}$) is $t \cap q$; the absolute velocity corresponding to the $\boldsymbol{u}$-relative velocity is $\boldsymbol{u} + \dot{q}$, therefore

$$\boldsymbol{f}_{\boldsymbol{u}}(t, q, \dot{q}) := \boldsymbol{f}(t \cap q, \boldsymbol{u} + \dot{q}).$$

Note that the absolute force can be easily recovered from the relative one:

$$\boldsymbol{f}(x, \dot{x}) = \boldsymbol{f}_{\boldsymbol{u}}(\tau(x), \sigma_{\boldsymbol{u}}(x), \mathbf{v}_{\dot{x}\boldsymbol{u}}). \tag{IV23}$$

Let us examine the form of a relative force corresponding to an absolute force having the potential $\boldsymbol{K}$.

Recall the formulae of Subsection 6.11 for $\boldsymbol{K}$.

Let

$$\boldsymbol{K} \quad \text{have the } \boldsymbol{u}\text{-split form} \quad (-V_{\boldsymbol{u}}, \boldsymbol{A});$$

then

$$\boldsymbol{F} := \mathrm{D} \wedge \boldsymbol{K} \quad \text{have the } \boldsymbol{u}\text{-split form} \quad ((-\nabla V_{\boldsymbol{u}} - \mathrm{D}_{\boldsymbol{u}}\boldsymbol{A}, \nabla \wedge \boldsymbol{A})) =: ((\boldsymbol{E}_{\boldsymbol{u}}, \boldsymbol{B})).$$

The force, by definition, is

$$\boldsymbol{i}^* \cdot \boldsymbol{F}(x) \cdot \dot{x} = \boldsymbol{i}^* \cdot \boldsymbol{F}(x) \cdot \boldsymbol{u} + \boldsymbol{i}^* \cdot \boldsymbol{F}(x) \cdot (\dot{x} - \boldsymbol{u}).$$

The first member on the right hand side is just the $\boldsymbol{u}$-timelike compoent of $\boldsymbol{F}$, $(\dot{x} - \boldsymbol{u})$ in the second member is absolutely spacelike, so it can be substituted by $\boldsymbol{i} \cdot (\dot{x} - \boldsymbol{u})$; thus, there we find the spacelike component of $\boldsymbol{F}$ and the $\boldsymbol{u}$-relative velocity.

As a consequence,

$$\boldsymbol{i}^* \cdot \boldsymbol{F}(x) \cdot \dot{x} \quad \text{has the } \boldsymbol{u}\text{-split form} \quad \boldsymbol{E}_{\boldsymbol{u}} + \boldsymbol{B} \cdot \mathbf{v}_{\dot{x}\boldsymbol{u}}.$$

We recognize: in the electromagnetic case $V_{\boldsymbol{u}}$ is the scalar potential, $\boldsymbol{A}$ is the vector potential, $\boldsymbol{E}_{\boldsymbol{u}}$ is the electric force and $\boldsymbol{B} \cdot \mathbf{v}_{\dot{x}\boldsymbol{u}}$ is the magnetic Lorentz force.

Of course, the above formulae are valid not only for electromagnetism but for gravitation and elasticity where $\boldsymbol{i}^* \cdot \boldsymbol{K} = \boldsymbol{A} = 0$, so $\boldsymbol{B} = 0$. According to Paragraph 19.4, in such a case the potential is absolute timelike, i.e. there is an **absolute scalar potential** $V : \mathtt{M} \to \mathbb{I}^*$ such that

$$\boldsymbol{K} = V\boldsymbol{\tau}.$$

Then the force does not depend on velocity.

7.6 Some special absolute forces

In this subsection we examine the form of the most often treated usual forces in spacetime.

7.6.1 The simplest cases

a) Force **independent of velocity**: there is a function $\boldsymbol{h} : \mathtt{M} \to \frac{\mathbf{E}^*}{\mathbb{I}}$ and

$$\boldsymbol{f}(x, \dot{x}) = \boldsymbol{h}(x).$$

The corresponding $\boldsymbol{u}$-relative force is

$$\boldsymbol{f}_{\mathbf{u}}(t, q, \dot{q}) = \boldsymbol{h}(t \cap q).$$

b) Force **dependig only on time**: there is a function $\boldsymbol{h} : \mathtt{I} \to \frac{\mathbf{E}^*}{\mathbb{I}}$ and

$$\boldsymbol{f}(x, \dot{x}) = \boldsymbol{h}(\tau(x)).$$

The corresponding $\boldsymbol{u}$-relative force is

$$\boldsymbol{f}_{\mathbf{u}}(t, q, \dot{q}) = \boldsymbol{h}(t).$$

c) **Constant** force: there is a $\boldsymbol{h} \in \frac{\mathbf{E}^*}{\mathbb{I}}$ and

$$\boldsymbol{f}(x, \dot{x}) = \boldsymbol{h}$$

The gravitational force near the Earth surface is modelled by such a force.

This force has an absolute scalar potential,

$$V(x) = -\boldsymbol{h} \cdot (\boldsymbol{\sigma}_{\mathbf{u}} \cdot (x - o))$$

for arbitrary $o \in M$ and $\mathbf{u} \in \mathtt{V}(1)$; the potential itself is

$$\boldsymbol{K}(x) = -\boldsymbol{h} \cdot (\boldsymbol{\sigma}_{\mathbf{u}} \cdot (x - o))\boldsymbol{\tau}. \tag{IV24}$$

d) In usual treatments of mechanics we often meet static forces i.e. forces independent of time and velocity, in other words, depending only on 'space'. Note the important fact that **there is no static absolute force**, except the constant forces. It may happen that the absolute force is constant on the straight lines directed by an absolute velocity $\boldsymbol{u}_c$ i.e.

$$\boldsymbol{f}(x, \dot{x}) = \boldsymbol{f}(x + \boldsymbol{t}\boldsymbol{u}_c)$$

holds for all $\boldsymbol{t} \in \mathbb{I}$. Then we say that the absolute force is **$\boldsymbol{u}_c$-static**. Equivalently, there is a function $\boldsymbol{h} : \mathbf{E} \to \frac{\mathbf{E}^*}{\mathbb{I}}$ ansd an $o \in M$ such that

$$\boldsymbol{f}(x, \dot{x}) = \boldsymbol{h}(\boldsymbol{\sigma}_{\mathbf{u}_c} \cdot (x - o)).$$

The corresponding $\boldsymbol{u}$-relative force is

$$\boldsymbol{f}_{\mathbf{u}}(t,q,\dot{q}) = \boldsymbol{h}(q - q_o - (t - t_o)\mathbf{v}_{\mathbf{u}_c\boldsymbol{u}})$$

where $q_o := \sigma_{\boldsymbol{u}}(o)$ and $t_o := \tau(o)$. To see that this formula is correct, we can apply equality (IV23) t is replaced by $\tau(x)$, q is replaced by $\sigma_{\boldsymbol{u}}(x)$ and it is used that $\sigma_{\boldsymbol{u}}(x) - \sigma_{\boldsymbol{u}}(o) = (x - o) - \boldsymbol{\tau} \cdot (x - o)\boldsymbol{u}$.

It is evident then this is time independent if and only $\boldsymbol{u} = \boldsymbol{u}_c$:

$$\boldsymbol{f}_{\mathbf{u}_c}(t,q,\dot{q}) = \boldsymbol{h}(q - q_o).$$

7.6.2 Central forces

Roughly speaking, we say that a force is central if it depends only on the space vector from a centre. Such forces ares the gravitational force due to a mass point and the elastic force. The centre means a material point, i.e. a world line in spacetime. The force in a world point is determined by the vector between the world point and the simultaneous occurrence of the world line (Figure 7.3). Thus, a central force is given by a world line function r_c and a map $a : \mathbb{D} \to \frac{\mathbb{R}}{\mathbb{D}\otimes\mathbb{D}\otimes\mathbb{I}}$ in such a way

$$\boldsymbol{f}(x,\dot{x}) := a\big(|x - r_c(\tau(x))|\big)\big(x - r_c(\tau(x))\big).$$

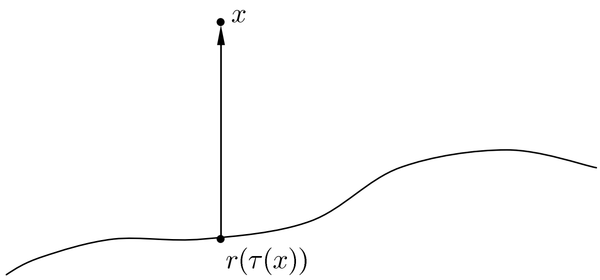

Figure 7.3 Central force

The corresponding $\boldsymbol{u}$-relative force is

$$\boldsymbol{f}_{\mathbf{u}}(t,q,\dot{q}) = a\big(|q - q_c(t)|\big)\big(q - q_c(t)\big)$$

where $q_c(t) := \sigma_{\boldsymbol{u}}(r_c(t))$ is the position of the centre in $\boldsymbol{u}$-space.

A special case: the centre is inertial with absolute velocity $\boldsymbol{u}_c$. Then with an arbitrary occurrence o of the world line

$$\boldsymbol{f}(x,\dot{x}) = a\big(|\boldsymbol{\sigma}_{\mathbf{u}_c}\cdot(x-o)|\big)\big(\boldsymbol{\sigma}_{\mathbf{u}_c}\cdot(x-o)\big).$$

This force is $\boldsymbol{u}_c$-static and has an absolute scalar potential

$$V(x) := -b(|\boldsymbol{\sigma}_{\mathbf{u}_c}\cdot(x-o)|)$$

where b is a primitive function of $\xi \mapsto a(\xi)\xi$. The potential itself is

$$\boldsymbol{K}(x) := -b(|\boldsymbol{\sigma}_{\mathbf{u}_c}\cdot(x-o)|)\boldsymbol{\tau}. \tag{IV25}$$

This force has the $\boldsymbol{u}_c$-relative form

$$\mathbf{f}_{\mathbf{u}_c}(t,q,\dot{q}) = a(|q-q_o|)(q-q_o)$$

where $q_o := \sigma_{\mathbf{u}_c}(o)$ is the (resting) position of the centre in $\boldsymbol{u}_c$-space.

In particular:

a) **Gravitational** force of a mass point:

$$\boldsymbol{f}(x,\dot{x}) = -\frac{\gamma m_c m}{|x-r_c(\tau(x))|^3}\big(x-r_c(\tau(x)\big)$$

where γ is the gravitational constant, and m_c is the mass of the centre and m the mass of the material upon which the centre acts.

b) **Elastic** force:

$$\boldsymbol{f}(x,\dot{x}) = -k(x-r_c(\tau(x))$$

where k is a constant.

7.7 Kinetic energy and power

The $\boldsymbol{u}$**-kinetic energy** of a material point with mass m and absolute velocity $\dot{x}$ is, by definition,

$$\frac{m|\mathbf{v}_{\dot{x}\mathbf{u}}|^2}{2} \in \frac{\mathbb{R}}{\mathbb{I}} \equiv \mathbb{I}^*.$$

This quantity evidently depends on the inertial observer $\boldsymbol{u}$. Since its value is in $\mathbb{I}^*$, we could think that it is the $\boldsymbol{u}$-timelike component of a covector but this is not true: the above expression is quadratic in $\boldsymbol{u}$ (not linear), so **there is no covector whose $\boldsymbol{u}$-timelike component is the $\boldsymbol{u}$-kinetic energy**.

Multiplying the absolute Newtonian equation (IV22) by the relative velocity $\mathbf{v}_{\dot{x}\boldsymbol{u}} = \dot{x} - \boldsymbol{u}$ and taking into account that $\ddot{x} = (\dot{x} - \boldsymbol{u})^{\cdot}$, we get

$$\left(\frac{m|\mathbf{v}_{\dot{x}\boldsymbol{u}}|^2}{2}\right)^{\cdot} = \boldsymbol{f}(x, \dot{x}) \cdot \mathbf{v}_{\dot{x}\boldsymbol{u}}.$$

The right hand side is the time derivative of the kinetic energy; the quantity on the left hand is called the **$\boldsymbol{u}$-relative power** of the force $\boldsymbol{f}$. .

Now it is worthy to make a little digression for a better understanding of the relativistic case. The absolute force has values in $\frac{\mathbf{E}^*}{\mathbb{I}}$; that is why we can think that it is the spacelike component of a covector. We find quickly a solution: let $\boldsymbol{f}(x, \dot{x})$ be the spacelike component of the covector whose $\dot{x}$-timelike component is zero. This condition determines uniquely the function

$$\hat{\boldsymbol{f}} : \mathtt{M} \times \mathtt{V}(1) \to \frac{\mathbf{M}^*}{\mathbb{I}}, \qquad \hat{\boldsymbol{f}}(x, \dot{x}) := \boldsymbol{f}(x, \dot{x})(\mathbf{1} - \dot{x} \otimes \boldsymbol{\tau})$$

for which

$$\hat{\boldsymbol{f}}(x, \dot{x}) \cdot \dot{x} = 0, \qquad \boldsymbol{i}^* \cdot \hat{\boldsymbol{f}}(x, \dot{x}) = \boldsymbol{f}(x, \dot{x})$$

holds. Then

$$\boldsymbol{f}(x, \dot{x}) \cdot \mathbf{v}_{\dot{x}\boldsymbol{u}} = \hat{\boldsymbol{f}}(x, \dot{x}) \cdot (\dot{x} - \boldsymbol{u}) = -\hat{\boldsymbol{f}}(x, \dot{x}) \cdot \boldsymbol{u},$$

i.e. there is a unique absolute covector function such that its negative $\boldsymbol{u}$-timelike component is the $\boldsymbol{u}$-power and its spacelike component is the force.

Warning: $\hat{\boldsymbol{f}}$ is never independent of $\dot{x}$, it depends even if $\boldsymbol{f}$ is independent.

7.8 Conservation laws

7.8.1 Action-reaction

Let us consider two material points – briefly particles – which without touching exert a force one to other (action at a distance).

According to Newton's law of action-reaction the interaction between the particles is **instantaneous** and the forces are **opposite** to each other.

In formula: let $\boldsymbol{f}_{12}(x_1, \dot{x}_1, x_2, \dot{x}_2)$ be the force acting on the 'first' particle being in the world point x_1 and having absolute velocity $\dot{x}_1$, due to the 'second' particle being in the world point x_2 and having absolute velocity $\dot{x}_2$ and let $\boldsymbol{f}_{12}(x_1, \dot{x}_1, x_2, \dot{x}_2)$ be the other force. Then

$$(\tau(x_1) = \tau(x_2)), \qquad \boldsymbol{f}_{21}(x_2, \dot{x}_2, x_1, \dot{x}_1) = -\boldsymbol{f}_{12}(x_1, \dot{x}_1, x_2, \dot{x}_2)$$

must hold.

If the particles have masses m_1 and m_2, then the Newtonian equations for them are

$$\big((x_1, x_2) : I \to M \times M\big)? \qquad m_1\ddot{x}_1 = \boldsymbol{f}_{12}(x_1, \dot{x}_1, x_2, \dot{x}_2),$$
$$m_2\ddot{x}_2 = \boldsymbol{f}_{21}(x_2, \dot{x}_2, x_1, \dot{x}_1).$$

Then the law of action-reaction yields immediately that

$$(m_1\dot{x}_1 + m_2\dot{x}_2)^{\cdot} = 0,$$

i.e. the **total momentum is coserved**: it does not change in the course of the interaction.

Let us suppose that the force has an absolute scalar potential which depends only on the distance of the particles, i.e. there is continuously differentiable function $V : \mathbb{D} \to \mathbb{I}^*$ such that

$$\boldsymbol{f}_{12}(x_1, \dot{x}_1, x_2, \dot{x}_2) = -\frac{\partial V(|x_1 - x_2|)}{\partial x_1} = -V'(|x_1 - x_2|)\frac{x_1 - x_2}{|x_1 - x_2|},$$

where the prime denotes differentiation with respect to the variable of V.

Multiplying the first Newtonian equation by $(\dot{x}_1 - \boldsymbol{u})$ and the second one by $(\dot{x}_2 - \boldsymbol{u})$ and taking their sum we get that the time derivative of the total $\boldsymbol{u}$-kinetic energy stands on the left hand side and the time derivative of $-V(|x_1 - x_2|)$ stands on the right hand side. Considering V as a **potential energy**, we can state that for all $\boldsymbol{u}$ the **$\boldsymbol{u}$-mechanical energy** is conserved

$$\frac{m_1|\mathbf{v}_{\dot{x}_1\boldsymbol{u}}|^2}{2} + \frac{m_2|\mathbf{v}_{\dot{x}_2\boldsymbol{u}}|^2}{2} + V(|x_1 - x_2|)$$

is conserved: it does not change in the course of the interaction.

7.8.2 Collisions

We accept as a fundamental physical fact that absolute momentum is conserved in processes, too, that cannot be described by forces and Newtonian equations, e.g. in collisions.

Let two particles meet and join together (inelastic collision). Let the particles have masses m_1 and m_2 and absolute velocities $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, respectively. Let the arising new particle have mass m_3 and absolute velocity $\boldsymbol{u}_3$. The the conservation of total momentum gives

$$m_1\boldsymbol{u}_1 + m_2\boldsymbol{u}_2 = m_3\boldsymbol{u}_3.$$

The timelike component of this equality is just the **conservation of mass**:

$$m_1 + m_2 = m_3.$$

Of course, the $\boldsymbol{u}$-spacelike component is the **conservation of $\boldsymbol{u}$-relative momentum** for arbitrary inertial observer $\boldsymbol{u}$:

$$m_1 \mathbf{v}_{\boldsymbol{u}_1 \boldsymbol{u}} + m_2 \mathbf{v}_{\boldsymbol{u}_2 \boldsymbol{u}} = m_3 \mathbf{v}_{\boldsymbol{u}_3 \boldsymbol{u}}.$$

The total kinetic energy of the particles, relative to the observer with respect to the arising particle rests, before collisison is

$$\frac{m_1|\boldsymbol{u}_1 - \boldsymbol{u}_3|^2}{2} + \frac{m_2|\boldsymbol{u}_2 - \boldsymbol{u}_3|^2}{2} \qquad \text{(IV26)}$$

and after collision is 0. Thus, if $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are not equal to $\boldsymbol{u}_3$ (this is a proper collision), then the $\boldsymbol{u}_3$-kinetic energy vanishes in the collision.

It is experienced, however, that the arising particle has a property which can be explained by this lost of energy, e.g. chemical bond, higher temperature. That is why we introduce the notion of **internal energy** and we conceive that the kinetic energy before collision is transformed into internal energy in the collision, i.e. the total energy is conserved.

More precisely, accepting the definition

'$\boldsymbol{u}$-relative energy:=$\boldsymbol{u}$-kinetic energy + internal energy'

we state the **conservation of $\boldsymbol{u}$-relative energy** for all inertial observers $\boldsymbol{u}$. Of course, we can must show that the definition is right: if it is true for some – say, for $\boldsymbol{u}_3$ above – then it is true for an arbitrary $\boldsymbol{u}$. Simply: the difference between the $\boldsymbol{u}$-kinetic energy before collision and the $\boldsymbol{u}$-kinetic energy after collision equals the similar difference regarding $\boldsymbol{u}_3$.

This is true because $|\boldsymbol{u}_1 - \boldsymbol{u}|^2 = |\boldsymbol{u}_1 - \boldsymbol{u}_3|^2 + 2(\boldsymbol{u}_1 - \boldsymbol{u}_3)\cdot(\boldsymbol{u}_3 - \boldsymbol{u}) + |\boldsymbol{u}_3 - \boldsymbol{u}|^2$ and the same equality holds for the subscript 2, too; furthermore, using $m_1(\boldsymbol{u}_1 - \boldsymbol{u}_3) + m_2(\boldsymbol{u}_2 - \boldsymbol{u}_3) = m_1\boldsymbol{u}_1 + m_2\boldsymbol{u}_2 - (m_1 + m_2)\boldsymbol{u}_3 = 0$ we get

$$\frac{m_1|\boldsymbol{u}_1 - \boldsymbol{u}|^2}{2} + \frac{m_2|\boldsymbol{u}_2 - \boldsymbol{u}|^2}{2} - \frac{(m_1+m_2)|\boldsymbol{u}_3 - \boldsymbol{u}|^2}{2} = \\ = \frac{m_1|\boldsymbol{u}_1 - \boldsymbol{u}_3|^2}{2} + \frac{m_2|\boldsymbol{u}_2 - \boldsymbol{u}_3|^2}{2}.$$

7.9 The rocket equation

We ca find in some books treating mechanics that the forms of the relative Newtonian equation given in Subsection 7.4 (the absolute Newtonian equation is not known there) are not equivalent if the mass varies in time; then the second one is accepted: the time derivative of the relative moment equals the relative force. However, the force 'does not know' whether the mass is constant or not (consider for instance that electromagnetic force acts on a body whose electric charge is constant but it looses mass). Thus, in the case of varying mass the right hand side of the relative Newtonian equation remains the same but the left hand would change, resulting in the equation $(m\dot{q})^{\cdot} = \dot{m}\dot{q} + m\ddot{q} = \mathbf{f_u}(t, q, \dot{q})$ which seems meaningful.

The trouble becomes evident if we look for the corresponding absolute equation. The left hand side, with the aid of the explicit splitting, has the form $(m\boldsymbol{\sigma_u} \cdot \dot{x})^{\cdot} = \boldsymbol{\sigma_u} \cdot (\dot{m}\dot{x} + m\ddot{x})$, therefore the absolute Newtonian equation would read $(m\dot{x})^{\cdot} = \dot{m}\dot{x} + m\ddot{x} = \boldsymbol{f}(x, \dot{x})$. Applying $\boldsymbol{\tau}$ to this equation we get (since $\boldsymbol{f}$ and $m\ddot{x}$ have spacelike values) that $\dot{m} = 0$: the mass does not change. This means that the "the time derivative of the relative moment equals the relative force" is not the correct form of the relative Newtonian equation if mass changes.

This is an excellent example how the relative objects – and even coordinates – can mislead one.

It is interesting, that the correct relative equation, in general, appears in the same books, as well. This is the rocket equation whose absolute form is the following.

The mass of rocket is given as a function of time, $m : \mathtt{I} \to \frac{\mathbb{I}}{\mathbb{D}\otimes\mathbb{D}}$, and the relative velocity of the mass outflow with respect to the rocket, as a function of time, $\mathbf{v} : \mathtt{I} \to \frac{\mathbf{E}}{\mathbb{I}}$, and then the **absolute rocket equation** is

$$(x : \mathtt{I} \mapsto \mathtt{M})? \qquad m\ddot{x} - \dot{m}\mathbf{v} = \boldsymbol{f}(x, \dot{x}).$$

Of course, this is valid not only for rockets but for every body whose mass changes, and the change can be mass increase, too (e.g. a raindrop gains weight passing through a cloud).

The $\boldsymbol{u}$-relative form of the rocket equation is

$$m\ddot{q} - \dot{m}\mathbf{v} = \mathbf{f_u}(t, q, \dot{q}).$$

Let us remark that the relative velocity $\mathbf{v}$ has the following meaning: $\mathbf{v}(t)$ is the relative velocity of the mass outflow with respect to the inertial observer in which the rocket is at rest at the instant t i.e. the absolute velocity of the observer is $\dot{r}(t)$. Thus, the absolute velocity of the mass outflow equals $\dot{r}(t) + \mathbf{v}(t)$.

8 Fundamentals of electromagnetism in spacetime

Electromagnetism is somewhat more complicated than mechanics but it has a well elaborated theory in the usual framework relative to observers. Just because of its more complicated nature it offers further possibilities to deepen our knowledge on the nonrelativistic spacetime model. Moreover, electromagnetism plaid a fundamental role in composing the relativity theory; our treatment aids us to understand better that role.

8.1 Maxwell equations

8.1.1 Relative Maxwell equations

The usual Maxwell equations are composed by quantities related to an inertial observer, thus, they are relative equations. Now we examine them for obtaining the absolute Maxwell equations.

The existence of the elementary electric charge makes possible to measure electric charges by real numbers i.e. we can choose the real line as the measure line of electric charges; though this does not fit the everyday practice, makes simpler the formulae in the theory.

Accordingly, the charge density ρ which is $\frac{\text{charge}}{\text{volume}}$ has the physical dimension $\frac{\mathbb{R}}{\mathbb{D}^{\otimes 3}}$, The physical dimension of the electric current density $\boldsymbol{j}$ is $\frac{\mathbb{R}}{\mathbb{D}^{\otimes 2}\otimes\mathbb{I}}$.

The usual Maxwell equations, of course, in relative quantities read as follows:

$$\mathrm{div}\boldsymbol{D} = \rho, \tag{IV27}$$

$$-\partial_0\boldsymbol{D} + \mathrm{curl}\boldsymbol{H} = \boldsymbol{j}, \tag{IV28}$$

$$\mathrm{curl}\boldsymbol{E} + \partial_0\boldsymbol{B} = 0, \tag{IV29}$$

$$\mathrm{div}\boldsymbol{B} = 0, \tag{IV30}$$

where ∂_0 denotes partial derivative with respect to time. These equations fix the physical dimension of the electric displacement $\boldsymbol{D}$ and the magnetic displacement $\boldsymbol{H}$. The physical dimension of the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ will be given by the formula of force acting on charges.

Furthermore it is important that in this usual treatment the values of the magnetic quantities $\boldsymbol{H}$ and $\boldsymbol{B}$ are axial vectors; these are, in fact, antisymmetric tensors (see the mathematical supplement). Let us change the quantities accordingly; then, using the 'vector' ∇ of spacelike differentiation we shall write

$$\nabla \cdot \boldsymbol{H} \text{ instead of curl}\boldsymbol{H} \ (\nabla \times \boldsymbol{H}),$$

$$\nabla \wedge \boldsymbol{B} \text{ instead of div}\boldsymbol{B} \ (\nabla \cdot \boldsymbol{B}).$$

Since the electric quantities are vectors,

$$\nabla \cdot \boldsymbol{D} \text{ will stand instead of div}\boldsymbol{D},$$

$$\nabla \wedge \boldsymbol{E} \text{ will stand instead of curl}\boldsymbol{E}.$$

Without doubt, the antisymmetric tensors instead of axial vectors make more complicated some formulae but they are correct from a theoretical point of view.

8.1.2 Absolute Maxwell equations

The previous electromagnetic quantities are relative to an inertial observer.

Now we look for the corresponding absolute forms.

Let us begin with the quantities connected with charges. The history of a pointlike charge is a world line. The history of many charges is a collection of world lines. The collection of world lines 'closely together' is similar to an observer, we can describe it by a futurelike vector field: the integral curves of the vector field are the world lines of the charges. There is a difference, however: the values of the futurelike vector will not be absolute velocities, they 'length' will reflect how close the charges are with respect to each other, i.e. the density of charges in spacetime. Thus, the history of charges in spacetime will be described by a futurelike vector field $\mathcal{J}$ in such a way that its 'length' $\boldsymbol{\tau} \cdot \mathcal{J}$ gives the density. Consequently, we have **the absolute electric flow density**

$$\mathcal{J} : \mathtt{M} \to \frac{\mathbf{M}}{\mathbb{D}^{\otimes 3} \otimes \mathbb{I}},$$

whose $\boldsymbol{u}$-split form is

$$\rho := \boldsymbol{\tau} \cdot \mathcal{J} : \mathtt{M} \to \frac{\mathbb{R}}{\mathbb{D}^{\otimes 3}}, \qquad \boldsymbol{j_u} = \boldsymbol{\sigma_u} \cdot \mathcal{J} : \mathtt{M} \to \frac{\mathbf{E}}{\mathbb{D}^{\otimes 3} \otimes \mathbb{I}}$$

the charge density and the electric $\boldsymbol{u}$-flow density, respectively. The charge density has the following meaning: an instant t is a hyperplane over the Euclidean vector space $\mathbf{E}$ and integrating the restriction of ρ onto t with respect to the canonical measure (see 21) for a subset we get the quantity of charges in that subset in the instant t. Similarly, integrating $\boldsymbol{j_u}$ to a two dimensional surface in t, we get the quantity of charges passing through that surface in the $\boldsymbol{u}$-space in the instant t.

The usual formula of Lorenz force read as follows: according to an observer that the force acting on a charge moving with velocity $\boldsymbol{v}$ in an electric filed $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$, considered an axial vector equals $\boldsymbol{E} + \boldsymbol{B} \cdot \boldsymbol{v}$. Formulae in 7.5.2 show that the electric field and the magnetic field relative to an observer $\boldsymbol{u}$ are split components of a cotensor field $\boldsymbol{F} : \mathtt{M} \to \mathbf{M}^* \wedge \mathbf{M}^*$, the **electromagnetic field**:

$$\boldsymbol{E_u} := \boldsymbol{i}^* \cdot \boldsymbol{F} \cdot \boldsymbol{u}, \qquad \boldsymbol{B} := \boldsymbol{i}^* \cdot \boldsymbol{F} \cdot \boldsymbol{i}$$

where we denoted explicitly that the electric field depends on the observer.

Further, we easily obtain that the other electric and magnetic quantities, too, are $\boldsymbol{u}$-split form of an antisymmetric tensor field $\boldsymbol{G} : \mathtt{M} \to \frac{\mathbf{M} \wedge \mathbf{M}}{\mathbb{D}^{\otimes 3} \otimes \mathbb{I}}$, the **electromagnetic displacement**:

$$\boldsymbol{D} := \boldsymbol{G} \cdot \boldsymbol{\tau}^*, \qquad \boldsymbol{H_u} := \boldsymbol{G} - (\boldsymbol{G} \cdot \boldsymbol{\tau}^*) \wedge \boldsymbol{u}.$$

The usual Maxwell equations, considered $\boldsymbol{u}$-split forms of some absolute equations, are rewritten according to the previous facts; then partial time derivative is replaced by $\boldsymbol{u}$-timelike derivative:

$$\nabla \cdot \boldsymbol{D} = \rho, \tag{IV31}$$
$$-\mathrm{D}_{\boldsymbol{u}} \boldsymbol{D} + \nabla \cdot \boldsymbol{H_u} = \boldsymbol{j_u}, \tag{IV32}$$
$$\nabla \wedge \boldsymbol{E_u} + \mathrm{D}_{\boldsymbol{u}} \boldsymbol{B} = 0, \tag{IV33}$$
$$\nabla \wedge \boldsymbol{B} = 0. \tag{IV34}$$

Then it is evident from Subsection 6.11 that the **absolute Maxwell equations** are

$$\mathrm{D} \cdot \boldsymbol{G} = \mathcal{J}, \qquad \mathrm{D} \wedge \boldsymbol{F} = 0. \tag{IV35}$$

According to the first equation, the quantity $\boldsymbol{G}$ gives how the charges produce electromagnetic action.

The electromagnetic field $\boldsymbol{F}$ gives the electromagnetic force acting on a charge; namely, the force acting at the world point x on the pointlike charge

e with absolute velocity $\dot{x}$ equals $e\boldsymbol{i}^* \cdot \boldsymbol{F}(x) \cdot \dot{x}$ (this expression appeared in Paragraph 7.3.2 where the charge was omitted for the sake of simplicity).

Note the important fact that $\boldsymbol{G}$ **is a tensor**, whereas $\boldsymbol{F}$ **is a cotensor**.

8.2 Constitutive relations

8.2.1 General formulae

Of course, $\boldsymbol{G}$ and $\boldsymbol{F}$ must be related somehow. It is known that the same charge-current density in different media produces different electromagnetic field. Usually one considers the equations $\boldsymbol{D} = \epsilon \boldsymbol{E}$ and $\boldsymbol{H} = \frac{1}{\mu}\boldsymbol{B}$ with some coefficients ϵ called permittivity and μ called permeability.

Accordingly, in general, we have to give a **constitutive relation** between the electromagnetic quantities,

$$\boldsymbol{G} = \Gamma(\boldsymbol{F})$$

which reflects how a material medium in spacetime influences the electromagnetic phenomena.

Then we get the **absolute Maxwell equations with a constitutive relation** in which only the electromagnetic field appear:

$$\mathrm{D} \cdot \Gamma(\boldsymbol{F}) = \mathcal{J}, \qquad \mathrm{D} \wedge \boldsymbol{F} = 0.$$

An inertial observer splits the constitutive relation, too, obtaining the result

$$\boldsymbol{D} = \eta_{\boldsymbol{u}}(\boldsymbol{E}_{\boldsymbol{u}}, \boldsymbol{B}), \qquad \boldsymbol{H}_{\boldsymbol{u}} = \gamma_{\boldsymbol{u}}(\boldsymbol{E}_{\boldsymbol{u}}, \boldsymbol{B}). \tag{IV36}$$

8.2.2 A special case

Let us suppose an infinite medium consisting of inertial material points which rest with respect to each other. As a matter of fact, such a medium is an inertial observer; let $\boldsymbol{u}_{\mathrm{o}}$ be its absolute velocity. Then the usual $\boldsymbol{u}_{\mathrm{o}}$-relative constitutive relation has the form

$$\boldsymbol{D} = \frac{\epsilon}{c}\boldsymbol{E}_{\boldsymbol{u}_{\mathrm{o}}}, \qquad \boldsymbol{H}_{\boldsymbol{u}_{\mathrm{o}}} = \frac{c}{\mu}\boldsymbol{B}, \tag{IV37}$$

where c is positive element of $\frac{\mathbb{D}}{\mathbb{I}}$.

From that we can construct the absolute constitutive relation:

$$\boldsymbol{G} = \boldsymbol{H}_{\boldsymbol{u}_{\mathrm{o}}} + \boldsymbol{D} \wedge \boldsymbol{u}_{\mathrm{o}} = \frac{c}{\mu}\boldsymbol{i}^* \cdot \boldsymbol{F} \cdot \boldsymbol{i} + \frac{\epsilon}{c}(\boldsymbol{i}^* \cdot \boldsymbol{F} \cdot \boldsymbol{u}_{\mathrm{o}}) \wedge \boldsymbol{u}_{\mathrm{o}}.$$

It is trivial that for any $\boldsymbol{u} \neq \boldsymbol{u}_o$ the $\boldsymbol{u}$-relative constitutive relation differs from the $\boldsymbol{u}_o$-relative one, $\boldsymbol{D} \neq \frac{\epsilon}{c}\boldsymbol{E}_{\boldsymbol{u}}$ and $\boldsymbol{H}_{\boldsymbol{u}} \neq \frac{c}{\mu}\boldsymbol{B}$. This is straightforward from a physical point of view: the observer $\boldsymbol{u}_o$ is distinguished by the really existing medium.

8.3 What is the trouble?

Let us emphasize: the relative form (IV31) of Maxwell equations are the same for all inertial observers.

This is important because it is often stated in the introduction of the theory of relativity that the Maxwell equations 'do not transform in a correct way' in the nonrelativistic case, i.e. they have different forms for different observers.

Our result is exact; is not true the incorrect transformation?

It is true in a convenient formulation.

Let us consider the vacuum: there is no medium. What an absolute constitutive relation can be given? In usual treatments it is stated that the permittivity and permeability of vacuum is 1, therefore one writes

$$\boldsymbol{D} = \frac{1}{c}\boldsymbol{E}, \qquad \boldsymbol{H} = c\boldsymbol{B}. \tag{IV38}$$

Relative quantities appear in these equalities but the symbol referring to an observer is missing; therefore the question 'for what observer is that true?' does not arise at all in the usual treatments. Since there is no medium, there would be one acceptable answer: for all. This is impossible, however; if we write the missing symbol – i.e. we take $\boldsymbol{E}_{\boldsymbol{u}}$ instead of $\boldsymbol{E}$ and $\boldsymbol{H}_{\boldsymbol{u}}$ instead of $\boldsymbol{H}$ – then we see that one of the sides of the equalities do depend on $\boldsymbol{u}$ and the other side does not.

In order to have a relative constitutive relation of form in (IV38), we have to point out an inertial observer. In this way we obtain the ether, a fictitious medium which is neutral from the electromagnetic point of view i.e. having permittivity and permeability 1.

Resuming: there is no absolute constitutive relation corresponding to the vacuum; in other words, it is the usual vacuum constitutive relation (IV38) that 'does not transform correctly' i.e. cannot hold for all observers. Of course, then the relative Maxwell equations with constitutive relations corresponding to the vacuum do not transform well.

Once again: let us suppose that the constitutive relation

$$\boldsymbol{D} = \frac{\epsilon}{c}\boldsymbol{E}_{\boldsymbol{u}_o}, \qquad \boldsymbol{H}_{\boldsymbol{u}_o} = \frac{c}{\mu}\boldsymbol{B}$$

holds for some inertial observer $\boldsymbol{u}_o$. Then the $\boldsymbol{u}_o$-relative Maxwell equations with constitutive relation, too, have the usual form. For an $\boldsymbol{u} \neq \boldsymbol{u}_o$, however, neither the $\boldsymbol{u}$-relative constitutive relation nor the $\boldsymbol{u}$-relative Maxwell equations with constitutive relation have the usual form.

These are understandable and all right when an observable $\boldsymbol{u}_o$ distingushed by a real medium is in question.

In the case of vacuum – when there is no real medium – all those (of course with $\epsilon = \mu = 1$) are not acceptable. In order to make manageable the situation, one had to find out a fictitious medium, the ether.

9 Noninertial observers

The nonrelativistic spacetime model admits an excellent treatment of noninertial observers, rigid observers, in particular uniformly accelerated observers, uniformly rotating observers, and important notions related to them such as inertial forces, centrifugal forces, Coriolis forces. All these can be found in the book

T. Matolcsi: *Spacetime without Reference Frames* (Budapest, 1993, Akadémiai Kiadó)

V Absolute light propagation

10 Basic notions and assumptios

In this section we take a flat spacetime model and, using the notions and notations of Chapter 2, we examine the special properties of $\mathtt{T}^{\rightarrow}$, $\boldsymbol{P}$, $\boldsymbol{d}$ ans $\boldsymbol{B}$ that arise from the assumption of absolute light propagation.

As a result, we obtain the so called (special) relativistic spacetime model which is strange for the everyday thinking. Who does not want to know the way that leads to it can skip this section without any harm in understanding the model and its applications.

10.1 Light signals

Up to now we have considered histories of material points only, leading us to the notion of world lines in the model. According to our experience, a 'pointlike light packet in vacuo', let us call it a **light signal**, is similar to a material point in some respect: it moves in our space; the path of a free (unhindered) light signal is a straight line. In other respect, a light signal is different from a material point: it cannot be at rest in the space of any observer.

The history of a light signal, too, is modelled by a curve in spacetime. In flat spacetime the history of a free light signal is a straight line (or a curve consisting of straight line segments, if the light signal reflects on mirrors). The direction vector of such a straight line cannot be futurelike (if it were, then the light signal would be at rest in the space of an inertial observer).

As a consequence, the direction vectors of light signals in the nonrelativistic spacetime model would be in $\mathbf{E}$, so the light signals would be simultaneously in every point of its path with respect to any observer; as a consequence, a light signal reflected on a mirror would arrive back at its source in the very instant of its start. This contradicts our experience. Thus, light signals cannot be treated in the nonrelataivistic spacetime model:

Absolute simultaneity and the right description of light signals exclude each other.

10.2 Heuristics of light propagation

10.2.1 Motion of light signals

The fundamental facts about free light signals with respect to an arbitrary inertial observer are the following:

(**L1**) **The path of a free light signal** in the observer space is a (segment of a) **straight line**,

(**L2**) **Every straight line** in the space of the observer **can be the path** of a light signal,.

(**L3**) **A light signal is faster than any material point on the same path**.

(**L4**) **The motion of any light signal can be arbitrarily approximated by the motion of a material point**.

The first two statements do not require a comment. It is worth noting that the third statement makes sense without a synchronization.

The fourth statement makes the following sense (without a synchronization, too). Let us take two points in our space: a start and a goal. A light signal and a material point are launched together at the start: the light signal reaches the goal before the material point does. In the goal the time difference between the arrivals can be measured. The statement in question says that for any given time interval a material point can be started in such a way that it arrives at the goal after the light signal with a delay equalling the prescribed time interval.

Though the third and fourth properties are simple, together they have a vital importance. First of all, they imply:

Light signals on the same path are equally fast.

Indeed, if one of them was faster than an other, then a material point whose motion approximates sufficiently the first light signal, would be faster than the second light signal.

Then the above statement has the consequence:
The history of a light signal is independent of its source.

This can be illustrated by the following example: there are a lamp on the rail-track and a lamp on the moving train; both lamps flash when they meet: the two light signals will propagate together both in the space of the rail (Earth) and in the space of the train and even in the space of an arbitrary observer (the colour i.e. the frequency of the light signals emitted from identical lamps will be different for the different observers).

Thus, we can state:

Light propagation in spacetime is absolute .

10.2.2 Homogeneous and isotropic light propagation

The homogeneous and isotropic light propagation is the starting point in usual treatments of relativity theory[1] which means that light signals in the space of an arbitrary inertial observer, starting anywhere and propagating in any direction, have the same speed. This assertion, however, refers to the speed of motions along different paths, which makes sense only when a synchronization is given (see Paragraph 1.3.2).

Since absolute synchronization and light signals exclude each other, one would have to specify, which synchronization the homogeneous and isotropic light propagation holds for (because if it holds for a synchronization, then it does not hold for another one). The usual treatments mention a synchronization only after the assumption of homogeneous and isotropic light propagation. This gives rise to the false image that the synchronization is a consequence of the homogeneous and isotropic light propagation.

It is clear then that the homogeneous and isotropic light propagation in this formulation, without specifying a synchronization, cannot be accepted as a fundamental fact.

By the way, the homogeneous and isotropic light propagation is astonishing from the point of view of everyday thinking: if you move relative to me and something moves relative to me in the same direction, then that thing 'must' move slower relative to you than to me.

Nevertheless, the homogenous and isotropic light propagation seems to be an experimental fact. Is this a contradiction? No, it is not. Experiments concern a fact that do not refer to a synchronization. This should be a well-known fact

[1] "the velocity of propagation of light in vacuo must be the same constant value $c = 3 \cdot 10^8 m/s$ in every system of inertia."(M. C. Møller: *The Theory of Relativity*, Clarendon Press, 1972; p.30)

for a long time[2] but, unfortunately, it is not generally known.

Namely, experiments regarding light speed concern **two-way light propagation**: a light signal starts in a space point (source) of an inertial observer, reflects in an other space point (mirror) and returns to the source. Knowing the distance between the source and the mirror, and measuring at the source the time elapsed between the start and the arrival, we can calculate the **two-way speed of light**[3]. It is evident, that this procedure uses only the time elapsed in one space point and does not require a synchronization.

Varying the place of the source and the mirror, we can measure the two-way light speed in all directions and over all distances. Similarly, when the signal returns to the source after more reflections or even running round in a path, the **the round-way propagation** of light. We find the experimental fact:

We emphasize: the two-way speed does not say anything about the one-way speed, which is meaningful only if a synchronization is specified.

Experiments in which light signals return to the source after two or more reflections – and even running round in a path –indicate also:

(**L5**) **The round-way light speed** for any inertial observer **is the same over all paths** (homogeneous and isotropic round-way light propagation).

This round-way light speed is the natural constant $c := (2,99793\dots)10^8$m/s.

Keep in mind that the round-way speed does not imply anything on the one-way speed which requires a synchronization.

10.2.3 Measuring distances by time intervals

The homogeneous and isotropic light propagation gives us a possibility to to measure the distance between two points in the space of an inertial observer by the half of the time interval which passes in one of the space points between the start and the return of a light signal reflecting in the other space point. Though this method is not used in the everyday practice, it is common in astronomy.

The formulae of the spacetime model will be simpler with this method of measuring distances. Consequently, in the relativistic spacetime model the measure line $\mathbb{D}$ of distances will coincide with the measure line $\mathbb{I}$ of time intervals in such a way that

$$\mathrm{m} := (3,336\dots)10^{-9}\mathrm{s} >$$

[2]H. Reichenbach: *The Philosophy of Space and Time*, Dover, 1957

[3]Both Fizeau and Foucault measured the light speed in century XIX essentially in this manner.

according to this choice the round-way speed of light will be the real number 1.

Of course, then – taking the Planck constant to be 1, too – the measure line of mass will be $\frac{\mathbb{I}}{\mathbb{D}\otimes\mathbb{D}} = \frac{\mathbb{I}}{\mathbb{I}\otimes\mathbb{I}} = \frac{\mathbb{R}}{\mathbb{I}} = \mathbb{I}^*$ and

$$\mathrm{kg} := (8,55\dots)10^{50}\frac{1}{\mathrm{s}}.$$

11 Construction of the model

Who is interested only in the relativistic spacetime model and its physical application and does not want to know the way that leads to it can skip this section without any harm in understanding the model and its applications.

11.1 Lightlike vectors

According to properties (**L1**) and (**L5**) the light signals in the space of an arbitrary inertial observer move to and fro uniformly there and back uniformly on a straight line; this and the affine structure of spacetime suggest naturally the the history of a free light signal in spacetime is a straight line, similarly to the history of inertial material points.

Recall that the straight line of the history of an inertial material point is taken to be oriented, the orientation is given by the earlier-later of proper time. Though we cannot state that some proper time is passing for a light signal (there isno experience in this respect), the earlier-later relation of material points results in an orientation on the straight lines of light signals: the reflected signal returns after the start.

An oriented straight line describing the history of a light signal will be called a **light line**.

As said in connection with world lines, according to the affine structure of spacetime, arbitrary translation of a light line is a light line as well. Therefore, as in the case of inertial world lines, it suffices to give the possible direction vectors which will be called **future-lightlike vectors**.

In the space of the inertial observer $\boldsymbol{u}$ every straight line is obtained as the path of an inertial material point, i.e. every $\boldsymbol{u}$-space vector is of the form $\boldsymbol{h}+\mathbb{I}\boldsymbol{u}$ where $\boldsymbol{h}$ is futurelike. Since the direction of the path in the $\boldsymbol{u}$-space of a light signal having the future-lightlike vector $\boldsymbol{a}$ is $\boldsymbol{a}+\mathbb{I}\boldsymbol{u}$, we can state on the base of (**L2**) that for every futurelike vector $\boldsymbol{h}$ there is future-lightlike vector $\boldsymbol{a}$ such that $\boldsymbol{a}+\mathbb{I}\boldsymbol{u} = \boldsymbol{h}+\mathbb{I}\boldsymbol{u}$. Conversely, for every future-lightlike vector $\boldsymbol{a}$ there is a futurelike vector $\boldsymbol{h}$ such that $\boldsymbol{a} + \mathbb{I}\boldsymbol{u} = \boldsymbol{h} + \mathbb{I}\boldsymbol{u}$. In other words:, if $\boldsymbol{u}$ as an arbitrary absolute velocity then

(a) *for all absolute velocities* $\mathbf{u}$ *and for all futurelike vectors* $\mathbf{h}$ *there is a future-lightlike vector* $\mathbf{a}$ *and a* $\boldsymbol{t} \in \mathbb{I}$ *in such a way that* $\mathbf{a} + \boldsymbol{t}\mathbf{u} = \boldsymbol{h}$,

(b) *for all absolute velocities* $\mathbf{u}$ *and for all future-lightlike vectors* $\mathbf{a}$ *there is a futurelike vector* $\boldsymbol{h}$ *and a* $\boldsymbol{t} \in \mathbb{I}$ *in such a way that* $\mathbf{a} + \boldsymbol{t}\mathbf{u} = \boldsymbol{h}$.

According to property (**L3**), a future-lightlike vector cannot be futurelike. Moreover, as we found in 2.8.4 ($\boldsymbol{h} = \boldsymbol{t}^{-}\boldsymbol{u}^{-}$), the previous time intervals $\boldsymbol{t}$ must be positive.

Property (**L4**) can be formalized as follows: let $\boldsymbol{a}$ be the future-lightlike vector between the arrival occurrence y in a $\boldsymbol{u}$-space point and the start occurrence x of the light signal. For all $\boldsymbol{t}$ time intervals there is a material point starting together with the light and has the arrival occurrence $y + \boldsymbol{t}\boldsymbol{u}$. Therefore, we can state more than (b):

(c) *if* $\mathbf{a}$ *is a future-lightlike vector then* $\mathbf{a} + \boldsymbol{t}\mathbf{u}$ *is futurelike i.e. is in* $\mathtt{T}^{\rightarrow}$ *for all absolute velocities* $\mathbf{u}$ *and for all* $\boldsymbol{t} \in \mathbb{I}^{+}$.

Taking the limit $\boldsymbol{t} \to 0$, we find that $\boldsymbol{a}$ is the element of the boudary of $\mathtt{T}^{\rightarrow}$.

Lett $\boldsymbol{b} \neq 0$ is an element in the boundary of $\mathtt{T}^{\rightarrow}$. Then every open subset containg $\boldsymbol{b}$ – so $\boldsymbol{b} + \mathtt{T}^{\rightarrow}$ as well – intersect $\mathtt{T}^{\rightarrow}$. This means that for all $\boldsymbol{u}$ there is a $\boldsymbol{s} > 0$ such that $\boldsymbol{b} + \boldsymbol{s}\boldsymbol{u} =: \boldsymbol{h} \in \mathtt{T}^{\rightarrow}$. According to the statement (a) there is a future-lightlike vector $\boldsymbol{a}$ and a $\boldsymbol{t} > 0$ such that $\boldsymbol{a} + \boldsymbol{t}\boldsymbol{u} = \boldsymbol{b} + \boldsymbol{s}\boldsymbol{u}$. If $\boldsymbol{t}$ were bigger than $\boldsymbol{s}$ than $\boldsymbol{a} + (\boldsymbol{t} - \boldsymbol{s})\boldsymbol{u} = \boldsymbol{b}$ would hold which is impossible because the left hand side is in $\mathtt{T}^{\rightarrow}$ according to (c). Similarly, the case $\boldsymbol{s} > \boldsymbol{t}$ is impossible, too. Thus, $\boldsymbol{b} = \boldsymbol{a}$ i.e. every nonzero element in the boundary of $\mathtt{T}^{\rightarrow}$ is future-lightlike.

Summing up: the collection of the **future-lightlike vectors** is the boundary of of the set of futurelike vectors, except the zero:

$$\mathtt{L}^{\rightarrow} := \partial\mathtt{T}^{\rightarrow} \setminus \{0\}.$$

Correspondingly, $\mathtt{L}^{\leftarrow} := -\mathtt{L}^{\rightarrow}$ is the collection of **past-lightlike vectors** and $\mathtt{L} := \mathtt{L}^{\leftarrow} \cup \mathtt{L}^{\rightarrow}$ is the collection of **lightlike vectors**.

It is evident that $\mathtt{L}^{\rightarrow}$ is a cone with apex at zero i.e. if $\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$ and $\alpha > 0$ is a real number then $\alpha\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$.

The previous statements can be reformulated in such a way that we consider a futurelike vector $\boldsymbol{k}$ instead of $\mathbf{u}$, according to the sense.

(a') *For all* $\mathbf{k} \in \mathtt{T}^{\rightarrow}$ *and* $\boldsymbol{h} \in \mathtt{T}^{\rightarrow}$ *there are a* $\mathbf{a} \in \mathtt{L}^{\rightarrow}$ *and a positive number* $\alpha > 0$ *such that* $\mathbf{a} + \alpha\mathbf{k} = \boldsymbol{h}$,

(b') *for all* $\mathbf{k} \in \mathtt{T}^{\rightarrow}$ *and* $\mathbf{a} \in \mathtt{L}^{\rightarrow}$ *there are an* $\boldsymbol{h} \in \mathtt{T}^{\rightarrow}$ *and a positive number* $\alpha > 0$ *such that* $\mathbf{a} + \alpha\mathbf{k} = \boldsymbol{h}$, (c') *if* $\boldsymbol{h} \in \mathtt{T}^{\rightarrow}$ *and* $\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$ *then* $\boldsymbol{h} + \boldsymbol{a} \in \mathtt{T}^{\rightarrow}$*.*

Further, we have:

(d) *If* $\boldsymbol{h} \in \mathtt{T}^{\rightarrow}$ *and* $\mathbf{x}$ *is a vector non parallel to* $\boldsymbol{h}$ *then there is a nonzero real number* α *such that* $\boldsymbol{h} + \alpha\mathbf{x} \in \mathtt{L}^{\rightarrow}$*.*

Since $\mathtt{T}^{\rightarrow}$ is open, there is a λ such that $\boldsymbol{h} + \lambda\boldsymbol{x} \in \mathtt{T}^{\rightarrow}$. This cannot hold for every λ because $\mathtt{T}^{\rightarrow}$ does not contain a whole straight line. Thus, the supremum or infimum of such λ-s (maybe both) must be finite; let it be α. $\boldsymbol{h} + \alpha\boldsymbol{x}$ is in the boundary of $\mathtt{T}^{\rightarrow}$ and cannot be zero.

(e) *If* $\boldsymbol{h} \in \mathtt{T}^{\rightarrow}$ *and* $\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$ *then there is a real number* $\beta > 0$ *such that* $\beta\boldsymbol{h} - \boldsymbol{a} \in \mathtt{L}^{\rightarrow}$*.*

Let us apply (d) to $\boldsymbol{h}$ and $\boldsymbol{a}$: there is a number α such that $\boldsymbol{h} + \alpha\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$. Then α is necessarily negative because if it were positive then $\alpha\boldsymbol{a}$ would be in $\mathtt{L}^{\rightarrow}$ which is excluded by (c). Multiplying by $\beta := \frac{1}{|\alpha|}$, we get the desired result.

11.2 Homogeneous and isotropic light propagation

11.2.1 Round-way propagation in the model

Let us take an inertial observer $\boldsymbol{u}$. A light signal launched from a $\boldsymbol{u}$-space point – source – and reflected in another $\boldsymbol{u}$-space point – mirror – returns to the mirror. Let $\boldsymbol{a}$ and $\boldsymbol{a}'$ be the future-lightlike vector of the light signal from the source to the mirror and from the mirror the source, respectively.

The distance between the source and the mirror is (see 2.6) $\sqrt{\boldsymbol{d_u}(\boldsymbol{a},\boldsymbol{a})} = \sqrt{\boldsymbol{d_u}(\boldsymbol{a}',\boldsymbol{a}')}$. Thus, according to the homogeneous and isotropic two-way light propagation, the time elapsed in the source between the departure and the return of the light signal equals

$$\frac{\sqrt{\boldsymbol{d_u}(\boldsymbol{a},\boldsymbol{a})} + \sqrt{\boldsymbol{d_u}(\boldsymbol{a}',\boldsymbol{a}')}}{\mathrm{c}}$$

where c is the round-way speed of light. According to our agreement, we use measure lines such that c = 1, thus the round-way speed will not appear in the following formulae.

Since $\boldsymbol{t}\boldsymbol{u} = \boldsymbol{a} + \boldsymbol{a}'$ (see Figure 11.1), we have that the homogeneous and isotropic light propagation is characterized in our model as follows:

If $\boldsymbol{u} \in \mathtt{V}(1)$ and $\boldsymbol{a}, \boldsymbol{a}'$ are elements of $\mathtt{L}^{\rightarrow}$ such that for some $\boldsymbol{t} \in \mathbb{I}$

$$\boldsymbol{t}\boldsymbol{u} = \boldsymbol{a} + \boldsymbol{a}', \quad \text{then} \quad \boldsymbol{t} = \sqrt{\boldsymbol{d_u}(\boldsymbol{a},\boldsymbol{a})} + \sqrt{\boldsymbol{d_u}(\boldsymbol{a}',\boldsymbol{a}')}. \qquad \text{(V1)}$$

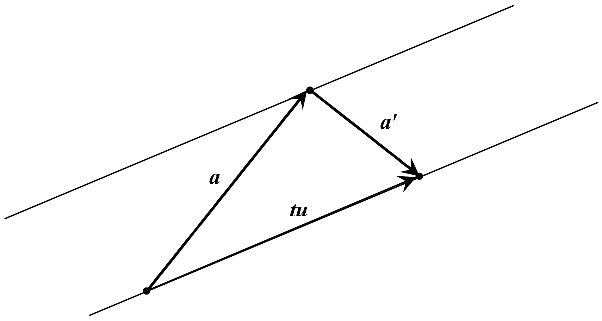

Figure 11.1 Two-way light propagation

In a similar way, considering a light signal reflected from more mirrors, we can formulate the homogeneous and isotropic round-way propagation of light in the model as follows:

> If $\boldsymbol{u} \in \mathtt{V}(1)$, $n \geq 2$ is a natural number and $\boldsymbol{a}_1, \dots \boldsymbol{a}_n$ are elements of $\mathrm{L}^{\rightarrow}$ such that for some $\boldsymbol{t} \in \mathbb{I}$
>
> $$\boldsymbol{tu} = \sum_{k=1}^{n} \boldsymbol{a}_k, \quad \text{then} \quad \boldsymbol{t} = \sum_{k=1}^{n} \sqrt{\boldsymbol{d_u}(\boldsymbol{a}_k, \boldsymbol{a}_k)}. \tag{V2}$$

Then we have

$$\sum_{k=1}^{n} \left(\boldsymbol{a}_k - \sqrt{\boldsymbol{d_u}(\boldsymbol{a}_k, \boldsymbol{a}_k)}\boldsymbol{u} \right) = 0. \tag{V3}$$

All these impose a relation between $\boldsymbol{d}$ and $\mathtt{L}^{\rightarrow}$; in other words, if we want – and we do want – to put the homogeneous and isotropic round-way light propagation in the model, then $\boldsymbol{d}$ cannot be given independently of $\mathtt{L}^{\rightarrow}$.

11.2.2 Standard space vectors of inertial observer

Formula (V3) holds only for $n \geq 2$; n cannot be 1 because a lightlike vector cannot be equal a timelike vector.

Nevertheless, this formula suggests investigating the set of vectors given by a similar formula for $n = 1$, more closely, the set

$$\mathbf{E}_{\boldsymbol{u}} := \{\mathbf{a} - \sqrt{\boldsymbol{d_u}(\mathbf{a}, \mathbf{a})}\boldsymbol{u}/\mathrm{c} \mid \mathbf{a} \in \mathrm{L}^{\rightarrow}\} \cup \{0\}. \tag{V4}$$

We have the following important fact:

$\mathbf{E}_{\boldsymbol{u}}$ *is a three-dimensional linear subspace which does not containt either timelike or lightlike vectors; in particular, it is transversal to* $\mathbb{I}\boldsymbol{u}$.

Because of the positive homogeneity of $\mathtt{L}^{\rightarrow}$, $\mathbf{E}_{\boldsymbol{u}}$ is positive homogeneous, too.

Fort all $\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$ and $\boldsymbol{s} \in \mathbb{I}^{+}$ putting $\boldsymbol{h} := \boldsymbol{su}$ the statement (e) in 11.1 results that there is a number $\beta > 0$ such that $\beta\boldsymbol{su} - \boldsymbol{a} \in \mathtt{L}^{\rightarrow} =: \boldsymbol{a}'$. Reformulated: for all $\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$ there are a positive $\boldsymbol{t} \in \mathbb{I}$ and an $\boldsymbol{a}' \in \mathtt{L}^{\rightarrow}$ such that $\boldsymbol{tu} = \boldsymbol{a} + \boldsymbol{a}'$. According the formula (V3) for $n = 2$ we have $\sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}', \boldsymbol{a}')}\boldsymbol{u} - \boldsymbol{a}' = \boldsymbol{a} - \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{a})}\boldsymbol{u}$. Thus, if $\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}}$ then there is a $\boldsymbol{q}' \in \mathbf{E}_{\boldsymbol{u}}$ such that $\boldsymbol{q}' = -\boldsymbol{q}$ i.e. $-\boldsymbol{q}$ is in $\mathbf{E}_{\boldsymbol{u}}$ as well.

As a consequence, $\mathbf{E}_{\boldsymbol{u}}$ is homogeneous, i.e. if $\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}}$, then $\alpha\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}}$ for all real numbers α.

Let $\boldsymbol{a}_1, \boldsymbol{a}_2 \in \mathtt{L}^{\rightarrow}$; then we know that $\boldsymbol{a}_1 + \boldsymbol{a}_2 \in \mathtt{L}^{\rightarrow} \cup \mathtt{T}^{\rightarrow}$.

If $\boldsymbol{a}_1 + \boldsymbol{a}_2 \in \mathtt{L}^{\rightarrow}$ then the previous reasoning yields that there are a positive $\boldsymbol{t} \in \mathbb{I}$ and an $\boldsymbol{a}_3 \in \mathtt{L}^{\rightarrow}$ such that $\boldsymbol{tu} = (\boldsymbol{a}_1 + \boldsymbol{a}_2) + \boldsymbol{a}_3$.

If $\boldsymbol{a}_1 + \boldsymbol{a}_2 \in \mathtt{T}^{\rightarrow}$ then taking it for the role $\boldsymbol{k}$ in the statement (b') and taking $\boldsymbol{h} := \boldsymbol{su}$, we get that there are an $\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$ and a number $\alpha > 0$ such that $\boldsymbol{a} + \alpha(\boldsymbol{a}_1 + \boldsymbol{a}_2) = \boldsymbol{su}$. Dividing by α and introducing $\boldsymbol{a}_3 := \boldsymbol{a}/\alpha$, $\boldsymbol{t} := \boldsymbol{s}/\alpha$, we obtain the previous equality.

Thus, in both cases, formula (V3) implies $\boldsymbol{a}_1 - \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}_1, \boldsymbol{a}_1)}\boldsymbol{u} + \boldsymbol{a}_2 - \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}_2, \boldsymbol{a}_2)}\boldsymbol{u} = \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}', \boldsymbol{a}')}\boldsymbol{u} - \boldsymbol{a}_3$.

As a consequence, $\mathbf{E}_{\boldsymbol{u}}$ is additive, i.e. if $\boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbf{E}_{\boldsymbol{u}}$, then $\boldsymbol{q}_1 + \boldsymbol{q}_2 \in \mathbf{E}_{\boldsymbol{u}}$.

Then, after all, $\mathbf{E}_{\boldsymbol{u}}$ is a linear subspace.

Now let us consider the restriction of the canonical surjection $\mathbf{M} \to \mathbf{M}/\mathbb{I}\boldsymbol{u}$ onto $\mathbf{E}_{\boldsymbol{u}}$, i.e. the linear map $\mathbf{E}_{\boldsymbol{u}} \to \mathbf{M}/\mathbb{I}\boldsymbol{u}$, $\boldsymbol{q} \mapsto \boldsymbol{q} + \mathbb{I}\boldsymbol{u}$. For a non zero element of $\mathbf{E}_{\boldsymbol{u}}$ we have $\boldsymbol{a} - \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{a})}\boldsymbol{u} + \mathbb{I}\boldsymbol{u} = \boldsymbol{a} + \mathbb{I}\boldsymbol{u} \neq \mathbb{I}\boldsymbol{u}$, thus the linear map in question is injective. It is surjective as well because we find a basis in $\mathbf{M}$ whose image lies in $\mathbf{E}_{\boldsymbol{u}}$. According to (a) in 11.1, for all $\boldsymbol{h} \in \mathtt{T}^{\rightarrow}$ there is an $\boldsymbol{a} \in \mathtt{L}^{\rightarrow}$ such that $\boldsymbol{h} + \mathbb{I}\boldsymbol{u} = \boldsymbol{a} + \mathbb{I}\boldsymbol{u}$ i.e. the elements of the form $\boldsymbol{h} + \mathbb{I}\boldsymbol{u}$ are in $\mathbf{E}_{\boldsymbol{u}}$; since $\mathtt{T}^{\rightarrow}$ is an opern set, there is a basis of $\mathbf{M}$ consisting of elements in $\mathtt{T}^{\rightarrow}$.

As a consequence, $\mathbf{E}_{\boldsymbol{u}}$ is three-dimensional.

Since $\mathbf{E}_{\boldsymbol{u}}$ is a linear subspace, it suffices to show that it does not contain either futurelike or future lightlike vectors.

If $\boldsymbol{x} \in \mathtt{T}^{\rightarrow} \cup \mathtt{L}^{\rightarrow}$ would be equal to a vector $\boldsymbol{a} - \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{a})}\boldsymbol{u}/c$ in $\mathbf{E}_{\boldsymbol{u}}$, then $\boldsymbol{a}$ would be equal to $\boldsymbol{x} + \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{a})}\boldsymbol{u}/c$ which is impossible according to the statement (c') in Paragraph 11.1.

The elements of $\mathbf{E}_{\boldsymbol{u}}$ will be called the **standard space vectors** of the inertial observer $\boldsymbol{u}$ because the $\boldsymbol{u}$-space vectors i.e. the elements of $\mathbf{M}/\mathbb{I}\boldsymbol{u}$ can be identified in a natual way with elements of $\mathbf{E}_{\boldsymbol{u}}$; later (see 12.4.1) we deal with $\mathbf{E}_{\boldsymbol{u}}$ in more details.

For further convenience, let us note that according to (V3)

$$\mathbf{E}_{\boldsymbol{u}} = \left\{ \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}', \boldsymbol{a}')}\boldsymbol{u} - \boldsymbol{a}' \mid \boldsymbol{a}' \in \mathtt{L}^{\rightarrow} \right\} \cup \{0\} \tag{V5}$$

holds as well, moreover, if

$$\boldsymbol{q} = \boldsymbol{a} - \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{a})}\boldsymbol{u} = \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}', \boldsymbol{a}')}\boldsymbol{u} - \boldsymbol{a}'$$

then

$$\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}, \boldsymbol{a}) = \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a}', \boldsymbol{a}') = \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{q}, \boldsymbol{q}), \tag{V6}$$

so

$$\mathbf{E}_{\boldsymbol{u}} = \left\{ \boldsymbol{q} \in \mathbf{M} \;\middle|\; \boldsymbol{q} + \sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{q},\boldsymbol{q})}\,\boldsymbol{u} \in \mathtt{L}^{\rightarrow} \right\} \cup \{0\}. \qquad \text{(V7)}$$

Since $\mathbf{E}_{\boldsymbol{u}}$ and $\mathbb{I}\boldsymbol{u}$ are complementary subspaces, there is a unique liner map $\boldsymbol{\tau}_{\boldsymbol{u}} : \mathbf{M} \to \mathbb{I}$ such that

$$\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u} = 1$$

and

$$\mathbf{E}_{\boldsymbol{u}} = \{ \boldsymbol{q} \in \mathbf{M} \mid \boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{q} = 0 \}. \qquad \text{(V8)}$$

According to equality (V4)

$$\sqrt{\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a},\boldsymbol{a})} = \boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{a} \qquad \text{(V9)}$$

for all future-lightlike vectors $\boldsymbol{a}$; as a consequence,

$$\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{a},\boldsymbol{a}) - (\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{a})^2 = 0 \qquad (\boldsymbol{a} \in \mathtt{L}^{\rightarrow}, \boldsymbol{u} \in \mathtt{V}(1)). \qquad \text{(V10)}$$

11.2.3 Different observers, different standard representations

If $\mathbf{u}$ *and* $\mathbf{u}'$ *are different absolute velocities, then the linear subspaces* $\mathbf{E}_{\boldsymbol{u}}$ *and* $\mathbf{E}_{\boldsymbol{u}'}$ *of* $\mathbf{M}$ *are different.*

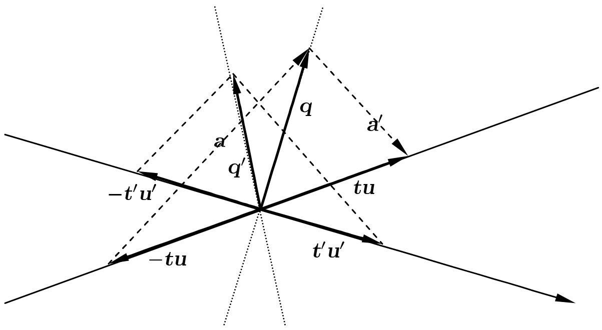

Figure 11.2 $\mathbf{E}_{\boldsymbol{u}}$ and $\mathbf{E}_{\boldsymbol{u}'}$ are different

Let $\boldsymbol{q}$ a multiple of $\boldsymbol{u}' - (\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}')\boldsymbol{u}$; $\boldsymbol{q}$ is in $\mathbf{E}_{\boldsymbol{u}}$ because $\boldsymbol{\tau}_{\boldsymbol{u}}$ maps it in zero. Thus, according to (V4), (V5) and (V6) there are future-lightlike vectors $\boldsymbol{a}$ and $\boldsymbol{a}'$ and $\boldsymbol{t} \in \mathbb{I}$ such that

$$\boldsymbol{q} = \boldsymbol{a} - \boldsymbol{t}\boldsymbol{u} = \boldsymbol{t}\boldsymbol{u} - \boldsymbol{a}'. \qquad \text{(V11)}$$

$\boldsymbol{a}$ and $\boldsymbol{a}'$ are in the plane spanned by $\boldsymbol{u}$ and $\boldsymbol{u}'$ and they cannot be parallel because a lightlike vector cannot be equal a timelike vector. Therefore the plane in question is spanned by $\boldsymbol{a}$ and $\boldsymbol{a}'$ as well.

Let us suppose that $\boldsymbol{q}$ is an element of $\mathbf{E}_{\boldsymbol{u}'}$, too. Then – with straightforward notations – we find that

$$\boldsymbol{q} = \alpha\boldsymbol{a} - \boldsymbol{t}'\boldsymbol{u}' = \boldsymbol{t}'\boldsymbol{u}' - \alpha'\boldsymbol{a}'. \tag{V12}$$

(V11) and (V12) imply $2\boldsymbol{q} = \boldsymbol{a} - \boldsymbol{a}' = \alpha\boldsymbol{a} - \alpha'\boldsymbol{a}'$ i.e. $(1-\alpha)\boldsymbol{a} = (1-\alpha')\boldsymbol{a}'$. Since $\boldsymbol{a}$ and $\boldsymbol{a}'$ are not parallel, this is possible only with $\alpha = \alpha' = 1$. As a consequence, $\boldsymbol{t}'\boldsymbol{u}' = \boldsymbol{t}\boldsymbol{u}$ holds which gives $\boldsymbol{u}' = \boldsymbol{u}$, a contradiction.

$\mathbf{E}_{\boldsymbol{u}}$ and $\mathbf{E}_{\boldsymbol{u}'}$ are different three-dimensional linear subspaces in the four dimensional vector space $\mathbf{M}$, so their intersection $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$ is a two dimensional linear subspace.

Let us note that the previous proof gives the results:

- the multiple of $\boldsymbol{u}' - (\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}')\boldsymbol{u}$ are elements of $\mathbf{E}_{\boldsymbol{u}}$ and are not in $\mathbf{E}_{\boldsymbol{u}'}$,
- the multiple of $\boldsymbol{u} - (\boldsymbol{\tau}'_{\boldsymbol{u}} \cdot \boldsymbol{u})\boldsymbol{u}'$ are elements of $\mathbf{E}_{\boldsymbol{u}'}$ and are not in $\mathbf{E}_{\boldsymbol{u}}$.

Furthermore, since neither $\mathbf{E}_{\boldsymbol{u}}$ nor $\mathbf{E}_{\boldsymbol{u}'}$ contains futurelike vectors,

$$\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}' > 0, \quad \text{and} \quad \boldsymbol{\tau}'_{\boldsymbol{u}} \cdot \boldsymbol{u} > 0$$

must hold.

11.3 Boosts

We have to define what it means that a vector (straight line) in the space of an inertial observer is equal (parallel) to a vector (straight line) in the space of another inertial observeri.e. we have to give the **boost** from $\boldsymbol{u}$ to $\boldsymbol{u}'$ for all absolute velocities $\boldsymbol{u}$ and $\boldsymbol{u}'$ (see 2.7).

The absolute light propagation gives a natural method for this. Later we will describe the corresponding physical procedure (see 12.5.6. Now we derive the boosts in a straightforward way from the mathematical relations obtained previously.

First, it is natural that the boost from $\boldsymbol{u}$ to $\boldsymbol{u}'$ maps the elements of $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$ onto themselves. Second, it is natural as well that the elements $\boldsymbol{u}' - (\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}')\boldsymbol{u}$ and $\boldsymbol{u} - (\boldsymbol{\tau}'_{\boldsymbol{u}} \cdot \boldsymbol{u})\boldsymbol{u}'$ which are not in $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$ and which play similar role, be connected by the boost. Accordingly, the boost is defined by

$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{u} = \boldsymbol{u}', \qquad \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{q} = \boldsymbol{q} \qquad (\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}),$$

$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \big(\boldsymbol{u}' - (\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}')\boldsymbol{u}\big) = -\big(\boldsymbol{u} - (\boldsymbol{\tau}'_{\boldsymbol{u}} \cdot \boldsymbol{u})\boldsymbol{u}'\big);$$

the negative sign appears because in such a manner will be the boost orientation preserving.

It is evident that the necessary condition $\boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'} = \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}^{-1}$ holds.

11.4 The absolute Lorentz form

11.4.1 Construction of the Lorentz form

For all $\boldsymbol{u}$ there must be given a $\boldsymbol{d_u} : \mathbf{M} \times \mathbf{M} \to \mathbb{I} \otimes \mathbb{I}$ positive semidefinite symmetric bilinear map whose kernel is $\mathbb{I}\boldsymbol{u}$. Since $\mathbf{E}_{\boldsymbol{u}}$ is transversal to $\mathbb{I}\boldsymbol{u}$-ra, the restriction of $\boldsymbol{d_u}$ to $\mathbf{E}_{\boldsymbol{u}}$ is positive definite. This restriction gives the Euclidean structure of the $\boldsymbol{u}$-space because $\mathbf{E}_{\boldsymbol{u}}$ represents the $\boldsymbol{u}$-space.

The boost form $\boldsymbol{u}$ to $\boldsymbol{u}'$ must satisfy

$$\boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{x}, \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{y}) = \boldsymbol{d_u}(\boldsymbol{x}, \boldsymbol{y})$$

for all spacetime vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. This equality and the property of the boost imply that the restrictions of $\boldsymbol{d}_{\boldsymbol{u}'}$ and $\boldsymbol{d_u}$ to $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$ must be equal:

$$\boldsymbol{d_u}(\boldsymbol{q}, \boldsymbol{q}) = \boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{q}, \boldsymbol{q}) \qquad \text{if} \qquad \boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}; \tag{V13}$$

this quantity, for the sake of simplicity, will be denoted by $|\boldsymbol{q}|^2$.

If $\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$ then both $\boldsymbol{q} + |\boldsymbol{q}|\boldsymbol{u}'$ and $-\boldsymbol{q} + |\boldsymbol{q}|\boldsymbol{u}'$ are future-lightlike vectors, thus, by (V10), $\boldsymbol{d_u}(\pm\boldsymbol{q} + |\boldsymbol{q}|\boldsymbol{u}', \pm\boldsymbol{q} + |\boldsymbol{q}|\boldsymbol{u}') = \big(\boldsymbol{\tau_u} \cdot (\boldsymbol{q} \pm |\boldsymbol{q}|\boldsymbol{u}')\big)^2$ i.e.

$$|\boldsymbol{q}|^2 \pm 2|\boldsymbol{q}|\boldsymbol{d_u}(\boldsymbol{q}, \boldsymbol{u}') + |\boldsymbol{q}|^2\boldsymbol{d_u}(\boldsymbol{u}', \boldsymbol{u}') = |\boldsymbol{q}|^2(\boldsymbol{\tau_u} \cdot \boldsymbol{u}')^2,$$

which implies that

$$\boldsymbol{d_u}(\boldsymbol{q}, \boldsymbol{u}') = 0 \qquad (\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}), \tag{V14}$$

and

$$(\boldsymbol{\tau_u} \cdot \boldsymbol{u}')^2 - \boldsymbol{d_u}(\boldsymbol{u}', \boldsymbol{u}') = 1. \tag{V15}$$

Changing the role of $\boldsymbol{u}$ and $\boldsymbol{u}'$ we get similar relations:

$$\boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{q}, \boldsymbol{u}) = 0 \qquad (\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}),$$

$$(\boldsymbol{\tau}_{\boldsymbol{u}'} \cdot \boldsymbol{u})^2 - \boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{u}, \boldsymbol{u}) = 1.$$

Then the properties of the boosts yield

$$\boldsymbol{d}_{\boldsymbol{u}'}\big(\boldsymbol{u} - (\boldsymbol{\tau}_{\boldsymbol{u}'} \cdot \boldsymbol{u})\boldsymbol{u}', \boldsymbol{u} - (\boldsymbol{\tau}_{\boldsymbol{u}'} \cdot \boldsymbol{u})\boldsymbol{u}'\big) = \boldsymbol{d_u}\big(\boldsymbol{u}' - (\boldsymbol{\tau_u} \cdot \boldsymbol{u}')\boldsymbol{u}, \boldsymbol{u}' - (\boldsymbol{\tau_u} \cdot \boldsymbol{u}')\boldsymbol{u}\big),$$

which results in

$$\boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{u},\boldsymbol{u}) = \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{u}',\boldsymbol{u}').$$

This and the previous formulae imply

$$\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}' = \boldsymbol{\tau}_{\boldsymbol{u}'} \cdot \boldsymbol{u}. \tag{V16}$$

An arbitrary spacetime vector $\boldsymbol{x}$ can be decomposed in the form

$$\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{t}\boldsymbol{u} + \boldsymbol{t}'\boldsymbol{u}'$$

where $\boldsymbol{q}$ is in $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$, $\boldsymbol{t}$ and $\boldsymbol{t}'$ are in $\mathbb{I}$. According to (V14) and (V15) it follows then that

$$\boldsymbol{d}_{\boldsymbol{u}}(\mathbf{x},\mathbf{x}) = |\boldsymbol{q}|^2 + (\boldsymbol{t}')^2(\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}')^2$$

Further,

$$\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \mathbf{x} = \boldsymbol{t} + \boldsymbol{t}'\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}',$$

so

$$\boldsymbol{d}_{\boldsymbol{u}}(\mathbf{x},\mathbf{x}) - (\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \mathbf{x})^2 = |\boldsymbol{q}|^2 - \boldsymbol{t}^2 - (\boldsymbol{t}')^2 - 2\boldsymbol{t}\boldsymbol{t}'\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{u}'.$$

Similar result is obtained for $\boldsymbol{u}'$ instead of $\boldsymbol{u}'$; therefore, by (V16) we have

$$\boldsymbol{d}_{\boldsymbol{u}}(\mathbf{x},\mathbf{x}) - (\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \mathbf{x})^2 = \boldsymbol{d}_{\boldsymbol{u}'}(\mathbf{x},\mathbf{x}) - (\boldsymbol{\tau}_{\boldsymbol{u}'} \cdot \mathbf{x})^2 \tag{V17}$$

for all absolute velocities $\boldsymbol{u}$ and $\boldsymbol{u}'$ and spacetime vectors $\boldsymbol{x}$. Applying this equality for $\boldsymbol{y}$ and $\boldsymbol{x}+\boldsymbol{y}$, we arrive at

$$\boldsymbol{d}_{\boldsymbol{u}}(\mathbf{x},\boldsymbol{y}) - (\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \mathbf{x})(\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{y}) = \boldsymbol{d}_{\boldsymbol{u}'}(\mathbf{x},\boldsymbol{y}) - (\boldsymbol{\tau}_{\boldsymbol{u}'} \cdot \mathbf{x})(\boldsymbol{\tau}_{\boldsymbol{u}'} \cdot \mathbf{x}).$$

This means that there is symmetric bilinear map $\boldsymbol{g} : \mathbf{M} \times \mathbf{M} \to \mathbb{I} \otimes \mathbb{I}$ such that for all $\boldsymbol{u} \in \mathtt{V}(1)$

$$\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}) := -(\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \mathbf{x})(\boldsymbol{\tau}_{\boldsymbol{u}} \cdot \boldsymbol{y}) + \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{x},\boldsymbol{y}). \tag{V18}$$

This $\boldsymbol{g}$ has negative values on $\mathbb{I}\boldsymbol{u} \times \mathbb{I}\boldsymbol{u}$ and positive definit on $\mathbf{E}_{\boldsymbol{u}} \times \mathbf{E}_{\boldsymbol{u}}$, ths $\boldsymbol{g}$ is a Lorentz form of type $1-3$ (see the Minkowski spaces in mathematical supplement).

For the sake of simplicity, later on we write a dot product instead of $\boldsymbol{g}$ (unless we want emphasis $\boldsymbol{g}$ for some reason):

$$\boldsymbol{x} \cdot \boldsymbol{y} := \boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}). \tag{V19}$$

11.4.2 Futurelike vectors

According to the Lorentz form (V18) and the equality (V10), $\boldsymbol{x}\cdot\boldsymbol{x} = 0$ for all future-lightlike vectors $\mathbf{x}$ and, of course, for all lightlike vectors, too. Further, $\boldsymbol{u}\cdot\boldsymbol{u} = -(\boldsymbol{\tau}\cdot\boldsymbol{u})^2 = -1$ for all absolute velocities $\boldsymbol{u}$, thus $\boldsymbol{x}\cdot\boldsymbol{x} < 0$ for all futurelike vectors $\boldsymbol{x}$ and, of course, for timelike, too. Since $\mathtt{L}^{\rightarrow}$ is the boundary of $\mathtt{T}^{\rightarrow}$, except the zero, the properties of the a Lorentz-forms imply that the equalities and inequalities above characterize the lightlike and and timelike vectors, respectively. The set of futurelike vectors in the set of timelike vectors means an **arrow orientation** of the Lorentz form. Summing up:

$$\mathtt{T}^{\rightarrow} = \{\boldsymbol{x}\in\mathbf{M} \mid \boldsymbol{x}\cdot\boldsymbol{x} < 0,\ \boldsymbol{x} \text{ has positive arrow}\}. \qquad \text{(V20)}$$

Consequently,

$$\mathtt{L}^{\rightarrow} = \{\boldsymbol{x}\in\mathbf{M} \mid \boldsymbol{x}\cdot\boldsymbol{x} = 0,\ \boldsymbol{x}\neq\mathbf{0},\ \boldsymbol{x} \text{ has positive arrow}\} \qquad \text{(V21)}$$

where having positive arrow means that $\boldsymbol{x}$ is a boundary point of $\mathrm{T}^{\rightarrow}$. Lastly, we have

$$\mathtt{V}(1) = \left\{\boldsymbol{u}\in\frac{\mathbf{M}}{\mathbb{I}}\middle|\boldsymbol{u}\cdot\boldsymbol{u} = -1,\ \boldsymbol{u} \text{ has positive arrow}\right\}. \qquad \text{(V22)}$$

11.4.3 Inertial proper time progress

We know about the function $\boldsymbol{P}$ of proper time progress that if $\boldsymbol{x}$ is futurelike, then $\frac{\boldsymbol{x}}{\boldsymbol{P}(\boldsymbol{x})}$ is an absolute velocity, thus $\frac{\boldsymbol{x}}{\boldsymbol{P}(\boldsymbol{x})}\cdot\frac{\boldsymbol{x}}{\boldsymbol{P}(\boldsymbol{x})} = -1$. As a consequence, time progress can be expressed by the Lorentz form as follows:

$$\boldsymbol{P}(\boldsymbol{x}) = \sqrt{-\boldsymbol{x}\cdot\boldsymbol{x}} \qquad (\boldsymbol{x}\in\mathtt{T}^{\rightarrow}). \qquad \text{(V23)}$$

11.4.4 Euclidean structures

Since the vector $\boldsymbol{q}$ is an element of $\mathbf{E}_{\boldsymbol{u}}$ if and only if $\boldsymbol{\tau}_{\boldsymbol{u}}\cdot\boldsymbol{q} = 0$, we see from the definition of the Lorentz form that $\boldsymbol{q}\cdot\boldsymbol{q} = \boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{q},\boldsymbol{q})$ for all $\boldsymbol{u}$ and $\boldsymbol{q}\in\mathbf{E}_{\boldsymbol{u}}$. Therefore (V7) (recall that c = 1) and (V21) yield that $\boldsymbol{q}$ is in $\mathbf{E}_{\boldsymbol{u}}$ if and only if $(\boldsymbol{q}+\sqrt{\boldsymbol{q}\cdot\boldsymbol{q}}\ \boldsymbol{u})\cdot(\boldsymbol{q}+\sqrt{\boldsymbol{q}\cdot\boldsymbol{q}}\ \boldsymbol{u}) = 0$, from which we infer that

$$\mathbf{E}_{\boldsymbol{u}} = \{\boldsymbol{q}\in\mathbf{M} \mid \boldsymbol{u}\cdot\boldsymbol{q} = 0\}.$$

Thus for all vectors $\boldsymbol{x}$, the equality $\boldsymbol{\tau_u}\cdot\mathbf{x}=0$ is equivalent to that $\boldsymbol{u}\cdot\boldsymbol{x}=0$; moreover, $\boldsymbol{\tau_u}\cdot\boldsymbol{u}=1=-\boldsymbol{u}\cdot\boldsymbol{u}$, so

$$\boldsymbol{\tau_u}\cdot\mathbf{x}=-\boldsymbol{u}\cdot\mathbf{x}\qquad(\boldsymbol{x}\in\mathbf{M})\tag{V24}$$

for all vectors $\boldsymbol{x}$. At last, we have got:

$$\boldsymbol{d_u}(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{x}\cdot\boldsymbol{y}+(\boldsymbol{u}\cdot\boldsymbol{x})(\boldsymbol{u}\cdot\boldsymbol{y})=\boldsymbol{g}\big(\boldsymbol{x}-\boldsymbol{u}(\boldsymbol{\tau_u}\cdot\boldsymbol{x}),\boldsymbol{y}-\boldsymbol{u}(\boldsymbol{\tau_u}\cdot\boldsymbol{y})\big)\tag{V25}$$
$$(\boldsymbol{u}\in\mathtt{V}(1)\quad\text{and}\quad\boldsymbol{x},\boldsymbol{y}\in\mathbf{M}).$$

where we have written $\boldsymbol{g}$ for emphasizing the similarity to formula (IV5).

12 The relativistic spacetime model

The special relativistic spacetime model introduced in the previous section is constructed on the following assumptions:

Light propagation is absolute in spacetime and has the properties listed in Subsection 10.2.

The description of light signals is not compatible with absolute simultaneity; the properties of the nonrelativistic spacetime model that are related to absolute simultaneity, are not valid here. Thus, we have to get rid of a number of familiar ideas.

In this section, not using the previous one, i.e not referring to how we have arrived at the spacetime in question, we define it and examine its properties. As a consequence, some formulae of the previous section appear again as a novelty.

We mention that this spacetime model, corresponding to the usual terminology, would be called special relativistic spacetime model. We deal only with flat spacetime models, so general relativistic spacetime models are excluded from this book, therefore we have omitted the adjective ‘special’ from our terminology.

12.1 Basic properties of the model

12.1.1 New notation

Instead of the general notation $(\mathtt{M},\mathbb{I},\mathbb{D},\mathrm{T}^{\to},\boldsymbol{P},\boldsymbol{d},\boldsymbol{B})$, a relativistic spacetime model will be referred to by the symbol $(\mathtt{M},\mathbb{I},\boldsymbol{g})$ where

– $\mathtt{M}$ is **spacetime**, a four-dimensional oriented affine space (over the vector space $\mathbf{M}$),

– $\mathbb{I}$ is the measure line of **time durations and distances**,

– $\boldsymbol{g} : \mathbf{M} \times \mathbf{M} \to \mathbb{I} \otimes \mathbb{I}$, $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{x} \cdot \boldsymbol{y}$ is an arrow oriented Lorentz form which, for the sake of simplicity, will be written as a dot-product, i.e. , $\boldsymbol{x} \cdot \boldsymbol{y} := \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$; we shall use the notions (pseudolength) and results (reversed Cauchy inequality) of Minkowskian vector spaces treated in the mathematical supplement; then

– the set of futurelike vectors is

$$\mathtt{T}^{\to} := \{\boldsymbol{x} \in \mathbf{M} \mid \boldsymbol{x} \cdot \boldsymbol{x} < 0,\ \boldsymbol{x} \text{ has positive arrow}\}$$

which is indeed an open convex cone with apex at zero,

– the inertial time progress is $\boldsymbol{P}(\boldsymbol{x}) := \sqrt{-\mathbf{x} \cdot \mathbf{x}} = |\boldsymbol{x}|$ ($\boldsymbol{x} \in \mathtt{T}^{\to}$) which is indeed smooth and positive homogeneous,
consequently, the set of absolute velocites is

$$\mathtt{V}(1) := \left\{ \boldsymbol{u} \in \frac{\mathbf{M}}{\mathbb{I}} \middle| \boldsymbol{u} \cdot \boldsymbol{u} = -1,\ \boldsymbol{u} \text{ has positive arrow} \right\}.$$

– $\mathbb{D} = \mathbb{I}$ and the Euclidean structure of inertial observers is given by

$$\boldsymbol{d_u}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{g}\big(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{x}), \boldsymbol{y} + \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{y})\big) = \boldsymbol{x} \cdot \boldsymbol{y} + (\boldsymbol{u} \cdot \boldsymbol{x})(\boldsymbol{u} \cdot \boldsymbol{y}) \qquad \text{(V26)}$$

for all $\boldsymbol{u} \in \mathtt{V}(1)$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}$) which depends smoothly on $\boldsymbol{u}$ and is positive semidefinite having the kernel $\mathbb{I}\boldsymbol{u}$,

– the boosts are given by the formula

$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} = \mathbf{1} + \frac{(\boldsymbol{u}' + \boldsymbol{u}) \otimes (\boldsymbol{u}' + \boldsymbol{u})}{1 - \boldsymbol{u}' \cdot \boldsymbol{u}} - 2\boldsymbol{u}' \otimes \boldsymbol{u} \qquad (\boldsymbol{u}', \boldsymbol{u} \in \mathtt{V}(1)) \qquad \text{(V27)}$$

which depends smoothly on $\boldsymbol{u}$ and $\boldsymbol{u}'$ and satisfies the requirements (the symbol of tensorial products is that e.g. $\boldsymbol{u}' \otimes \boldsymbol{u}$ is the linear map assigning $\boldsymbol{u}'(\boldsymbol{u} \cdot \boldsymbol{x})$ to the vector $\boldsymbol{x}$):

(i) $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{u} = \boldsymbol{u}'$ is evident,

(ii) $\boldsymbol{d}_{\boldsymbol{u}'}(\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \mathbf{x}, \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{y}) = \boldsymbol{d_u}(\boldsymbol{x}, \boldsymbol{y})$ because $\boldsymbol{u}' \cdot \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} = \boldsymbol{u}$ and a simple but somewhat lengthy calculation yields that

$$(\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{x}) \cdot (\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y} \qquad (\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}), \qquad \text{(V28)}$$

so $\boldsymbol{x} \cdot \boldsymbol{y} + (\boldsymbol{u}' \cdot \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \mathbf{x})(\boldsymbol{u}' \cdot \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y} + (\boldsymbol{u} \cdot \boldsymbol{x})(\boldsymbol{u} \cdot \boldsymbol{y})$,

(iii) $\boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'} = \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}^{-1}$ is obtained quite simply.

Observe the important fact that the time progress and the Euclidean structures in the nonrelativistic case are given by two different mathematical objects – $\boldsymbol{\tau}$ and $\boldsymbol{b}$ – whereas in the relativistic case they are obtained from a single one, $\boldsymbol{g}$.

Further notions:

$$\mathrm{T}^{\leftarrow} := -\mathrm{T}^{\rightarrow}, \qquad \mathrm{T} := \mathrm{T} \cup \mathrm{T}^{\rightarrow} = \{\boldsymbol{x} \in \mathbf{M} \mid \boldsymbol{x} \cdot \boldsymbol{x} < 0\}$$

are the set of **pastlike** and **timelike** vectors, respectively,

$$\mathrm{L}^{\rightarrow} := \{\boldsymbol{x} \in \mathbf{M} \mid \boldsymbol{x} \cdot \boldsymbol{x} = 0,\ \boldsymbol{x} \text{ has positive arrow}\}$$

is the set of **future-lightlike** vectors,

$$\mathrm{L}^{\leftarrow} := -\mathrm{L}^{\rightarrow}, \qquad \mathrm{L} := \mathrm{L}^{\leftarrow} \cup \mathrm{L}^{\rightarrow} = \{\boldsymbol{x} \in \mathbf{M} \mid \boldsymbol{x} \cdot \boldsymbol{x} = 0,\ \boldsymbol{x} \neq \mathbf{0}\}$$

are the set of **past-lightlike** and **lightlike** vectors, respectively,

$$\mathrm{S} := \{\boldsymbol{x} \in \mathbf{M} \mid \boldsymbol{x} \cdot \boldsymbol{x} > 0\} \cup \{0\}$$

is the set of **spacelike** vectorsa

This classification of spacetime vectors is represented in Figure 12.1 which corresponds to the arithmetic Lorentz form in two dimensions (see 12.2), without coordinate axes. The figure is somewhat misleading because the set of timelike vectors and the set of spacelike vectors are shown in a similar form; we get a better image by rotating the figure around a 'horizontal axis'.

The set of absolute velocities is three-dimensional submanifold in $\frac{\mathbf{M}}{\mathbb{I}}$, shown by Figure 12.2. Note that

– there is no zero absolute velocity,
– the magnitude of an absolute velocity makes no sense,
– the angle between absolute velocities makes no sense;

let us be cautious: the absolute velocity $\boldsymbol{u}_2$ in Figure 12.2 is not longer than $\boldsymbol{u}_1$, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ have no angle between them, $\boldsymbol{u}_1$ is not the central line of the cone $\mathrm{T}^{\rightarrow}$.

The same time durations on straight lines with different absolute velocities are represented by intervals of different lengths. The bigger the angle with the horizontal line, the longer the interval representing the same duration. The same length corresponds to the same duration for two absolute velocities that have the same angle with the horizontal line. That is why we illustrate two

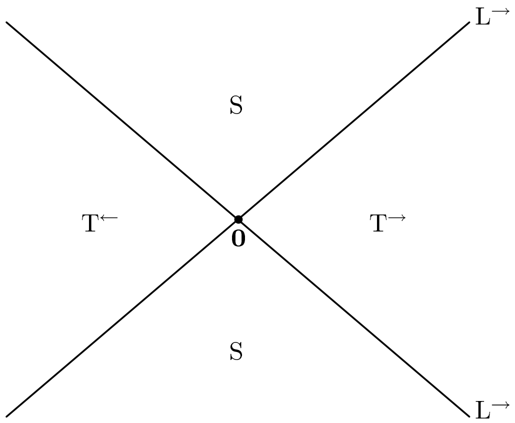

Figure 12.1 Spacetime vectors

absolute velocities in this way when treating relations concerning them. Lastly, we emphasize again that length and angle are properties of the illustration, not of the absolute velocities.

According to the reversed Cauchy inequality we have the important relation:

$$-\boldsymbol{u} \cdot \boldsymbol{u}' \geq 1 \qquad (\boldsymbol{u}, \boldsymbol{u}' \in V(1))$$

where equality holds if and only if $\boldsymbol{u} = \boldsymbol{u}'$.

For $\boldsymbol{u} \in \mathtt{V}(1)$, the set

$$\mathbf{E}_{\boldsymbol{u}} := \{\boldsymbol{q} \in \mathbf{M} \mid \boldsymbol{u} \cdot \boldsymbol{q} = 0\}$$

is a three dimensional spacelike linear subspace in $\mathbf{M}$, its elements are called **$\boldsymbol{u}$-spacelike vectors**; the restriction of the Lorentz form to $\mathbf{E}_{\boldsymbol{u}}$ is a **Euclidean structure** (i.e. positive definite). If $\boldsymbol{q}, \boldsymbol{p} \in \mathbf{E}_{\boldsymbol{u}}$ then

$$\boldsymbol{d}_{\boldsymbol{u}}(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{g}(\boldsymbol{q}, \boldsymbol{p}).$$

According to the rules of our illustration, $\mathbf{E}_{\boldsymbol{u}}$ is represented in our figures by a line which bends to the light cone with the same angle as $\boldsymbol{u}$ (see Figure 12.4). It is emphasized that those angles have no physical meaning, they are only tools or illustration.

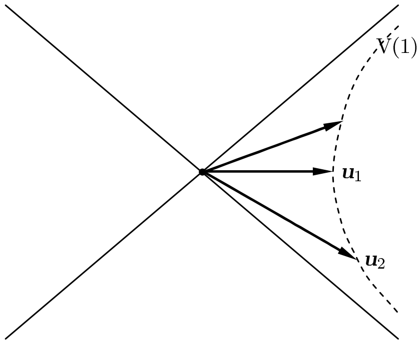

Figure 12.2 Absolute velocities

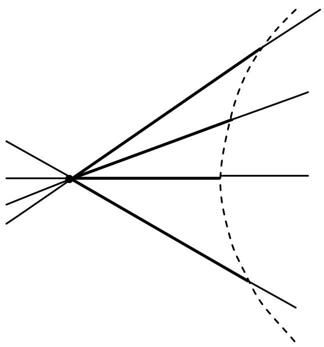

Figure 12.3 Same time intervals

12.1.2 Duals

The dual of $\mathbf{M}$, $\mathbf{M}^*$, the set of linear maps (covectors) $\mathbf{M} \to \mathbb{R}$ is a four-dimensional vector space, too; the Lorentz form $\boldsymbol{g}$ establishes the identification

$$\frac{\mathbf{M}}{\mathbb{I} \otimes \mathbb{I}} \equiv \mathbf{E}^*, \qquad \frac{\boldsymbol{x}}{\boldsymbol{s}^2} \equiv \frac{\boldsymbol{g}(\mathbf{x}, \cdot)}{\boldsymbol{s}^2}.$$

We can conceive this in another way, too: the elements of $\mathbf{M}$ can be identified with linear maps $\mathbf{M} \to \mathbb{I} \otimes \mathbb{I}$: $\mathbf{x} \equiv \boldsymbol{g}(\mathbf{x}, \cdot)$. This reflected in our convention, too, that a dot-product is written instead of $\boldsymbol{g}$.

From the usual point of view, the identification above means that coordinates of vectors are quartets of numbers having a 'superscript', coordinates of

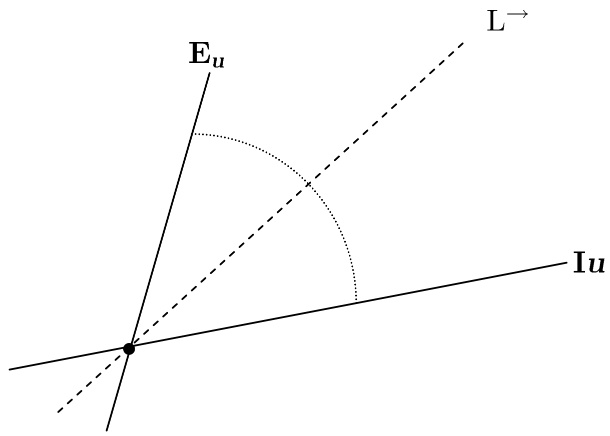

Figure 12.4 $\boldsymbol{u}$-spacelike vectors

covectors are quartets of numbers having a 'subscript' and there is a relation between the quantities with superscript and the quantities with subscript.

12.1.3 Proper times

The history of a material point in spacetime is a world line. The proper time passed on the world line C between the world points x and y, according to 2.4.2, now becomes

$$\boldsymbol{t}_{\mathrm{C}}(x,y) = \int_{p^{-1}(x)}^{p^{-1}(y)} |\dot{p}(a)|da,$$

where p is a progressive parametrization of C and $|\ |$ denotes the pseudolength i.e. $|\dot{p}(a)| = \sqrt{-\dot{p}(a)\cdot\dot{p}(a)}$.

Let y be futurelike with respect to x (i.e. $y-x\in \mathtt{T}^{\to}$). Let us approximate the value of the integral by taking a world point $z\in \mathrm{C}$ between x and y and by considering the proper times passed on the straight lines form x to z and from z to y. The result is $|(z-x)|+|y-z|$; according to the reversed Cauchy inequality, this is less then $|y-x|$, the inertial time from x to y. (except that the three points lie on a straight line). Taking further approximations by broken lines, the results become smaller and smaller.

At last, we get:

If $\boldsymbol{t}(x,y)$ denotes the inertial time between the world points x and y $(y-x\in \mathtt{T}^{\to})$ and C is not an inertial world line containing x and y then

$$\boldsymbol{t}_{\mathrm{C}}(x,y)<\boldsymbol{t}(x,y).$$

12.2 The arithmetic spacetime model

We can construct a relativistic spacetime model with the aid of real numbers. In this **arithmetic relativistic spacetime model**

– $\mathtt{M}=\mathbb{R}^4$ endowed with the standard orientation (then $\mathbf{M}=\mathbb{R}^4$, too),
– $\mathbb{I}=\mathbb{R}$ endowed with the standard orientation,
– $\boldsymbol{g}$ is the usual Lorentz form on $\mathbb{R}^4$, i.e.

$$\boldsymbol{g}((\xi^0,\xi^1,\xi^2,\xi^3),(\eta^0,\eta^1,\eta^2,\eta^3))=-\xi^0\eta^0+\sum_{i=1}^{3}\xi^i\eta^i.$$

The identification $\mathbf{M}\equiv\mathbf{M}^*$ induced by the Lorentz form $\mathbf{M}\equiv\mathbf{M}^*$ is the map

$$(\xi^0,\xi^1,\xi^2,\xi^3)\mapsto(-\xi^0,\xi^1,\xi^2,\xi^3)=:(\xi_0,\xi_1,\xi_2,\xi_3)$$

Therefore, the Lorentz product is usually written in the form

$$\boldsymbol{g}(\xi,\eta)=\xi_i\eta^i$$

where the Einstein summation is used.

Consequently,

$$\mathtt{T}^{\to}=\{(\xi^0,\xi^1,\xi^2,\xi^3)\mid \xi_i\xi^i<0,\ \xi^0>0\}.$$

Further, here

$$\mathtt{V}(1)=\{(\nu^0,\nu^1,\nu^2,\nu^3)\mid \nu_i\nu^i=-1,\ \nu^0>0\}.$$

The matrix of the boost is

$$(\boldsymbol{B}_{\nu'\nu})^k_i=\delta^k_i+\frac{(\nu'+\nu)^k(\nu'+\nu)_i}{1-(\nu')_m\nu^m}-2(\nu')^k\nu_i.$$

Let us repeat: in the arithmetic spacetime model

– the spacetime $\mathtt{M}$ and the underlying vector space $\mathbf{M}$ are the same set,
– $\mathbb{I}$ is the real line, consequently every measure line is $\mathbb{R}$; therefore $\frac{\mathbf{M}}{\mathbb{I}}=\mathbf{M}=\frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}}=\mathbb{R}^4$, etc.

12.3 Isomorphisms

Isomorphism of models is an important notion expressing whether two spacetime models of different forms have the same physical content or not.

12.3.1 Isomorphism in the new notation

We can assert that the relativistic spacetime model $(\hat{\mathtt{M}}, \hat{\mathbb{I}}, \hat{\boldsymbol{g}})$ is isomorphic to the relativistic spacetime model $(\mathtt{M}, \mathbb{I}, \boldsymbol{g})$ if and only if there are

(i) an orientation preserving affine bijection $L : \mathtt{M} \to \hat{\mathtt{M}}$ (over the linear bijection $\boldsymbol{L} : \mathbf{M} \to \hat{\mathbf{M}}$),

(ii) an orientation preserving linear bijection $\boldsymbol{F} : \mathbb{I} \to \hat{\mathbb{I}}$

which send $\mathtt{T}^{\to}$ into $\widehat{\mathtt{T}^{\to}}$, $\boldsymbol{P}$ into $\hat{\boldsymbol{P}}$, $\boldsymbol{d}$ into $\hat{\boldsymbol{d}}$ and $\boldsymbol{B}$ into $\hat{\boldsymbol{B}}$ in a 'convenient way'.

$\boldsymbol{F}$ is given by the 'units' s and ŝ according to 2.9.2.

Since $\boldsymbol{g}$ determines both $\mathtt{T}^{\to}$, $\boldsymbol{P}$ and $\boldsymbol{d}$, the first three general requirements can be fused:

(I)-(II)-(III)

$$\frac{\hat{\boldsymbol{g}}(\boldsymbol{L}\cdot\boldsymbol{x}, \boldsymbol{L}\cdot\boldsymbol{y})}{\hat{\mathrm{s}}^2} = \frac{\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})}{\mathrm{s}^2}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}$.

The last requirement,

(IV)

$$\left(\hat{\mathbf{1}} + \frac{(\boldsymbol{L}\cdot\boldsymbol{u}' + \boldsymbol{L}\cdot\boldsymbol{u})\otimes(\boldsymbol{L}\cdot\boldsymbol{u}' + \boldsymbol{L}\cdot\boldsymbol{u})}{1 - \boldsymbol{L}\cdot\boldsymbol{u}'\cdot\boldsymbol{L}\cdot\boldsymbol{u}} - 2\boldsymbol{L}\cdot\boldsymbol{u}'\otimes\boldsymbol{L}\cdot\boldsymbol{u}\right)\cdot\boldsymbol{L} =$$
$$= \boldsymbol{L}\cdot\left(\mathbf{1} + \frac{(\boldsymbol{u}'-\boldsymbol{u})\otimes(\boldsymbol{u}'-\boldsymbol{u})}{1-\boldsymbol{u}'\cdot\boldsymbol{u}} - 2\boldsymbol{u}'\otimes\boldsymbol{u}\right)$$

follows automatically from the previous ones (because e.g. the definition of the tensor product implies $\boldsymbol{L}\cdot(\boldsymbol{u}'\otimes\boldsymbol{u}) = (\boldsymbol{L}\cdot\boldsymbol{u}')\otimes\boldsymbol{u}$ and the previous equality gives $(\boldsymbol{L}\boldsymbol{u}'\otimes\boldsymbol{L}\boldsymbol{u})\cdot\boldsymbol{L} = \boldsymbol{L}\boldsymbol{u}'\otimes\boldsymbol{u}$).

12.3.2 Relativistic spacetime models are isomorphic

We can easily demonstrate:

> An arbitrary relativistic spacetime model is isomorphic to the arithmetic one; this implies that arbitrary two relativistic spacetime models are isomorphic to each other.

Considering an arbitrary relativistic spacetime model $(\mathtt{M}, \mathbb{I}, \boldsymbol{g})$, let us choose

– a time unit $\mathrm{s} \in \mathbb{I}^+$,

– an 'origin' $o \in \mathtt{M}$.

– a positively oriented $\boldsymbol{g}$-orthogonal basis $\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, normed to s, in $\mathbf{M}$, such that e_0 is futurelike,

and let us put

$$L : \mathtt{M} \to \mathbb{R}^4, \qquad x \mapsto \{\text{ coordinates of } x - o \text{ in the basis } \boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\},$$

$$\boldsymbol{B} : \mathbb{I} \to \mathbb{R}, \qquad \boldsymbol{t} \mapsto \frac{\boldsymbol{t}}{\mathrm{s}}.$$

It is simple that in such a way we have defined an isomorphism from the spacetime model in question to the arithmetic spacetime model.

Our result on isomophisms says that all relativistic spacetime models have the same physical content. Therefore, we have only philosophical and practical reasons to make any specific choice. The arithmetic model is suitable for practical purposes, i.e. for solving actual problems. For theoretical considerations, however, a special model, like the arithmetic one, is less appropriate because such a model can have additional properties

– that have nothing to do with the structure of the spacetime model,

– that confuse some essential features of the spacetime model,

We emphasize again: a treatise in coordinates is not defective in itself, because all relativistic spacetime models have the same physical content. Nevertheless, it is preferred not to use the arithmetic model for general considerations because its special properties can easily mislead us:

– spacetime points and spacetime vectors are given as elements of the same set,

– all measure lines are the real line,

etc. Using coordinates, if we do not want to make a mistake, we have to check permanently the physical validity of our ideas and formulae. This is very tiresome and, even if we do so, something can easily escape our attention.

12.3.3 Lorentz and Poincar transformations

The **Lorentz transformations** in the spacetime model $(\mathtt{M}, \mathbb{I}, \boldsymbol{g})$ are linear bijections $\boldsymbol{L} : \mathbf{M} \to \mathbf{M}$ for which

$$\boldsymbol{g}(\boldsymbol{L} \cdot \boldsymbol{x}, \boldsymbol{L} \cdot \boldsymbol{y}) = \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$$

holds for all $\boldsymbol{q}, \boldsymbol{p} \in \mathbf{E}$.

The composition of Lorentz transformations is a Lorentz transformation i.e. their collection is a group.

The orientation and arrow orientation preserving Lorentz transformations are called **proper Lorentz transformations**. According to the general definition in 2.9.3 and according to 6.3.1 they are the **vectorial symmetries** of the relativistic spacetime model.

In contrast to the nonrelativistic case, there are no special Lorentz transformations because now there is no distinguished three-dimensional linear subspace in $\mathbf{M}$.

The boosts are Lorentz transformations, as expected.

The **Poincaré transformations** are the affine bijections $L : \mathtt{M} \to \mathtt{M}$ over the Lorentz transformations. They form a group, too. The **proper Poincaré transformations** are the ones over the proper Lorentz transformations. These are the **symmetries** of the relativistic spacetime model.

A detailed treatment of the properties of the Lorentz group and the Poincaré group can be found in the book

T. Matolcsi: *Spacetime without Reference Frames* (Budapest, 1993, Akadémiai Kiadó)

12.4 Space and space vectors of an inertial observer

12.4.1 Standard representation of space vectors

Let us recall that the space of the inertial observer $\boldsymbol{u}$ is the set of straight lines in $\mathtt{M}$ directed by $\boldsymbol{u}$ (see 2.5.3),

$$\mathtt{E}_{\boldsymbol{u}} := \mathtt{M}/\mathbb{I}\boldsymbol{u},$$

and the set of $\boldsymbol{u}$-space vectors is $\mathbf{M}/\mathbb{I}\boldsymbol{u}$, the set of straight lines in $\mathbf{M}$ directed by $\boldsymbol{u}$.

The $\boldsymbol{u}$-space points are of the form $x + \mathbb{I}\boldsymbol{u}$ and the $\boldsymbol{u}$-space vectors are of the form $\mathbf{x} + \mathbb{I}\boldsymbol{u}$. $\mathtt{E}_{\boldsymbol{u}}$ is affine space over $\mathbf{M}/\mathbb{I}\boldsymbol{u}$ by the subtraction

$$(x + \mathbb{I}\boldsymbol{u}) - (y + \mathbb{I}\boldsymbol{u}) = (x - y) + \mathbb{I}\boldsymbol{u}.$$

The space vectors of an inertial observer cannot be fairly illustrated.

In the nonrelativistic spacetime model the space vectors of an arbitrary inertial observer are represented by the absolute spacelike vectors in a natural way. Now the situation is more complicated because there are more three-dimensional linear subspaces that are transversal to all absolute velocities. The Lorentz form, however, selects the three-dimensional linear subspace

$$\mathbf{E}_{\boldsymbol{u}} = \{\boldsymbol{q} \in \mathbf{M} \mid \boldsymbol{u} \cdot \boldsymbol{q} = 0\}$$

for every inertial observer $\boldsymbol{u}$.

Since $\mathbb{I}\boldsymbol{u}$ and $\mathbf{E}_{\boldsymbol{u}}$ are transversal, the $\boldsymbol{u}$-space vectors can be **identified** in a natural – standard – way by the elements of $\mathbf{E}_{\boldsymbol{u}}$ by the linear bijection

$$\mathbf{E}_{\boldsymbol{u}} \to \mathbf{M}/\mathbb{I}\boldsymbol{u}, \qquad \boldsymbol{q} \mapsto \boldsymbol{q} + \mathbb{I}\boldsymbol{u}. \tag{V29}$$

Let us introduce the notation

$$\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{x} := \boldsymbol{x} + \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{x}). \tag{V30}$$

It is evident that

$$\boldsymbol{\sigma}_{\boldsymbol{u}} = \mathbf{1} + \boldsymbol{u} \otimes \boldsymbol{u} : \mathbf{M} \to \mathbf{E}_{\boldsymbol{u}}$$

is a linear surjection and its is easy to see that

$$\mathbf{M}/\mathbb{I}\boldsymbol{u} \to \mathbf{E}_{\boldsymbol{u}}, \quad \boldsymbol{x} + \mathbb{I}\boldsymbol{u} \mapsto \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{x}$$

is the inverse of the linear bijection V29.

Thus, the identification in question can be made transparent by

$$\mathbf{E}_{\boldsymbol{u}} \equiv \mathbf{M}/\mathbb{I}\boldsymbol{u}, \qquad \boldsymbol{q} \equiv \boldsymbol{q} + \mathbb{I}\boldsymbol{u}$$

or, equivalently,

$$\mathbf{M}/\mathbb{I}\boldsymbol{u} \equiv \mathbf{E}_{\boldsymbol{u}}, \qquad \boldsymbol{x} + \mathbb{I}\boldsymbol{u} \equiv \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{x}.$$

Using this identification, later on we shall consider the elements of $\mathbf{E}_{\boldsymbol{u}}$ the space vectors of the observer, calling them **$\boldsymbol{u}$-spacelike vectors**.

The Euclidean structure defined in (V26) is given by the restriction of the Lorentz to $\mathbf{E}_{\boldsymbol{u}}$.

Thus, $\mathbf{E}_{\boldsymbol{u}}$ is a three-dimensional Euclidean vector space which can be endowed with an orientation in a natural way. Let $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ be an ordered basis of $\mathbf{E}$ positively oriented if $(\boldsymbol{tu}, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ is a positively oriented basis of $\mathbf{M}$ for some (hence for arbitrary) positive element $\boldsymbol{t}$ of $\mathbb{I}$. Further, it has the same properties as $\boldsymbol{e}$ in the nonrelativistic case (length of vectors, angles between them, axial vectors etc.)

Let us introduce the notation

$$\sigma_{\boldsymbol{u}}(x) := x + \mathbb{I}\boldsymbol{u};$$

$\sigma_{\boldsymbol{u}}(x)$ is the $\boldsymbol{u}$-space point containing the occurrence x. According to the above identification, the subtraction in the $\boldsymbol{u}$-space becomes

$$\sigma_{\boldsymbol{u}}(x) - \sigma_{\boldsymbol{u}}(y) = (x + \mathbb{I}\boldsymbol{u}) - (y + \mathbb{I}\boldsymbol{u}) = (x - y) + \mathbb{I}\boldsymbol{u} \equiv \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (x - y). \qquad \text{(V31)}$$

In another way, if q and p are points in $\mathtt{E}_{\boldsymbol{u}}$, then

$$q - p := \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (x - y) \qquad (x \in q,\ y \in p),$$

which equivalent to

$$q - p = x - y \qquad (x \in q,\ y \in p,\ x - y \in \mathbf{E}_{\boldsymbol{u}}).$$

Note that (V31) says that $\sigma_{\boldsymbol{u}} : \mathtt{M} \to \mathtt{E}_{\boldsymbol{u}}$, $x \mapsto x + \mathbb{I}\boldsymbol{u}$ is an affine map over the linear map $\boldsymbol{\sigma}_{\boldsymbol{u}} : \mathbf{M} \to \mathbf{E}_{\boldsymbol{u}}$.

As said, according to the rules of our illustration (which comes from the two-dimensional arithmetic Lorentz form), $\mathbf{E}_{\boldsymbol{u}}$ is represented in our figures by a line which bends to the light cone with the same angle as $\boldsymbol{u}$ (see Figure 12.4). It is emphasized that those angles have no physical meaning, they are only tools or illustration.

12.4.2 Different observers, different standard representations

It is easy to see that if $\boldsymbol{u} \neq \boldsymbol{u}'$, then $\mathbf{E}_{\boldsymbol{u}} \neq \mathbf{E}_{\boldsymbol{u}'}$. This follows from the properties of the Lorentz form, but can be demonstrated directly:

$$\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} := \frac{\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{u}'}{-\boldsymbol{u} \cdot \boldsymbol{u}'} = \frac{\boldsymbol{u}'}{-\boldsymbol{u} \cdot \boldsymbol{u}'} - \boldsymbol{u} \qquad \text{(V32)}$$

is an element of $\frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}}$ but is not in $\frac{\mathbf{E}_{\boldsymbol{u}'}}{\mathbb{I}}$. Moreover, it is orthogonal to $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$; indeed, if $\boldsymbol{q}$ is in the intersection then $\boldsymbol{u} \cdot \boldsymbol{q} = \boldsymbol{u}' \cdot \boldsymbol{q} = 0$. Similarly,

$$\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'} := \frac{\boldsymbol{\sigma}_{\boldsymbol{u}'} \cdot \boldsymbol{u}}{-\boldsymbol{u}' \cdot \boldsymbol{u}} = \frac{\boldsymbol{u}}{-\boldsymbol{u}' \cdot \boldsymbol{u}} - \boldsymbol{u}' \in \frac{\mathbf{E}_{\boldsymbol{u}'}}{\mathbb{I}} \qquad \text{(V33)}$$

is orthogonal to the intersection.

Therefore, we can state:

The spaces of different inertial observers are **different** three-dimensional affine spaces over the **different** vector spaces.

Since $\mathbf{E}_{\boldsymbol{u}}$ and $\mathbf{E}_{\boldsymbol{u}'}$ are different three-dimensional linear subspaces in the four-dimensional vector space $\mathbf{M}$, their intersection $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$ is a two-dimensional linear subspace.

It is worth keeping in mind that both $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$ and $\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}$ are orthogonal to the intersection:

$$\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} \perp \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'} \perp \mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}.$$

Moreover,

$$|\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}|^2 = |\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}|^2 = 1 - \frac{1}{(\boldsymbol{u}\cdot\boldsymbol{u}')^2} < 1 \tag{V34}$$

holds which implies

$$-\boldsymbol{u}\cdot\boldsymbol{u}' = \frac{1}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}|^2}} = \frac{1}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}|^2}}. \tag{V35}$$

The vectors 'emerging out' from the intersection have a clear physical meaning. Namely, according to 2.8.1, the path of an inertial material point with absolute velocity $\boldsymbol{u}'$ in the space of the inertial observer $\boldsymbol{u}$ is a straight line whose direction vector (divided by $\mathbb{I}$) is $\boldsymbol{u}' + \mathbb{R}\boldsymbol{u}$. According to the identification above, $\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{u}'$ is the corresponding space vector.

It is important to remark that in the usual treatments the space vectors of every observers is represented by the same $\mathbb{R}^3$ which can mislead one.

12.4.3 Properties of boosts

With the aid of the previously treated facts, we can give palpable properties of the boosts described by (V27). First, the boost from $\boldsymbol{u}$ to $\boldsymbol{u}'$ is the identity on the intersection of the $\boldsymbol{u}$-space and $\boldsymbol{u}'$-space,

$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{q} = \boldsymbol{q} \qquad (\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}).$$

Second,it maps the $\boldsymbol{u}$-space vector orthogonal to the intersection into the negative of the $\boldsymbol{u}'$-space vector, orthogonal to the intersection:

$$\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} = -\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}$$

i.e. the direction of the motion of $\boldsymbol{u}'$ in the $\boldsymbol{u}$-space is mapped into the negative of the direction of motion of $\boldsymbol{u}$ in the $\boldsymbol{u}'$-space (so it realizes our idea that "if you move in some direction relative to me then I move in the opposite direction relative to you").

These properties and $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{u} = \boldsymbol{u}'$ determine uniquely the boost from $\boldsymbol{u}$ to $\boldsymbol{u}'$.

The boosts are proper Lorentz transformations, i.e. vectorial symmetries as expected. Contrary to the nonrelataivistic case, however, they do not form a subgroup, in other words, they are **not transitive**:

$$\boldsymbol{B}_{\boldsymbol{u}''\boldsymbol{u}'}\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \neq \boldsymbol{B}_{\boldsymbol{u}''\boldsymbol{u}} \tag{V36}$$

for the absolute velocities $\boldsymbol{u}$,$\boldsymbol{u}'$ and $\boldsymbol{u}''$, in general; *equality occurs if and only if the three absolute velocities are coplanar.*

If they are coplanar, say $\boldsymbol{u}'' = \alpha\boldsymbol{u} + \alpha'\boldsymbol{u}'$ thena direct calculation shows the equality.

Let us suppose that the equality holds and $\boldsymbol{u}$, $\boldsymbol{u}'$, $\boldsymbol{u}''$ are different. Applying both sides to a vector $\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}''}$ we get $\boldsymbol{q} + (\boldsymbol{u}' \cdot \boldsymbol{q})f(\boldsymbol{u}, \boldsymbol{u}', \boldsymbol{u}'') = \boldsymbol{q}$ where $f(\boldsymbol{u}, \boldsymbol{u}', \boldsymbol{u}'')$ is a linear combination of the absolute velocities. This can hold only if

1. $f(\boldsymbol{u}, \boldsymbol{u}', \boldsymbol{u}'') = 0$ from which is evident the absolute velocities are coplanar;

2. $\boldsymbol{u}' \cdot \boldsymbol{q} = 0$ when we can argue: $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}''}$, $\boldsymbol{u}$ and $\boldsymbol{u}''$ span $\mathbf{M}$, so $\boldsymbol{u}'$ can be given in the form of a corresponding linear combination; since $\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}''}$ is arbitrary, the equality implies that $\boldsymbol{u}'$ is in the two-dimensional subspace spanned by $\boldsymbol{u}$ and $\boldsymbol{u}''$.

We can reformulate our result:

$$\boldsymbol{R}_{\boldsymbol{u}(\boldsymbol{u}'\boldsymbol{u}'')} := \boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}''}\boldsymbol{B}_{\boldsymbol{u}''\boldsymbol{u}'}\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}}$$

is the identity if and only if $\boldsymbol{u}$, $\boldsymbol{u}'$ and $\boldsymbol{u}''$ are coplanar.

$\boldsymbol{R}_{\boldsymbol{u}(\boldsymbol{u}'\boldsymbol{u}'')}$ is a vectorial symmetry (orientation and arrow orientation preserving Lorentz transformation) which maps $\boldsymbol{u}$ into $\boldsymbol{u}$; therefore, its restriction to $\mathbf{E}_{\boldsymbol{u}}$ is an orientation preserving Euclidean transformation i.e. a rotation, called the **Thomas-rotation** in $\boldsymbol{u}$, determined by $\boldsymbol{u}'$ and $\boldsymbol{u}''$.

Let us see the importance of the non-transitivity of boosts. A boost establishes a natural correspondence between different observer spaces which makes sense to the parallelism of straight lines in different spaces (in particular, to the usual tacit assumption that the coordinate axes of different moving coordinate systems are parallel).

Let me, you and him sit in different space ships. If my vector boosted to you equals a a vector of yours (my straight line is parallel to one of yours) then your vector boosted to me equals my vector (your straight line is parallel to mine). On the other hand, his vectors obtained from a vector of mine

a) by boosting from me to you and then from you to him,

b) by boosting from me to him,

are not equal, in general.

A recent paradox of relativity theory is based on the fact that one takes it for granted that parallelism between straight lines in different observer spaces is transitive (see Subsection 16.1).

It will treated later (12.5.6) how the boosts are put in practice.

12.5 Light signals

12.5.1 Light lines

We defined lightlike vectors in 12.1.1 but they have not yet appeared.

Up to now we have considered the history of material points, resulting the world lines in the model. As our experience show, a 'point like light packet' in vacuo, called **light signal** behaves similarly to a material point in some respect.

The history of a light signal will be described as a curve in spacetime, called **light line**; a curve whose every tangent vector is lightlike. It is convenient to take the elements, called **light directions**, of

$$\mathtt{V}^{\to} := \left\{ \boldsymbol{w} \in \frac{\mathbf{M}}{\mathbb{I}} \;\middle|\; \mathbb{I}^{+}\boldsymbol{w} \subset \mathtt{L}^{\to} \right\}$$

as tangent vectors (because the absolute velocities considered to be tangents of world lines are the elements of $\frac{\mathbf{M}}{\mathbb{I}}$ as well). In a simple setting, $\boldsymbol{w} \in \frac{\mathbf{M}}{\mathbb{I}}$ is a light direction if and only if it is future-lightlike. Unlike to absolute velocities, if $\boldsymbol{w}$ is a light direction then – because of the zero the pseudolength – $\alpha\boldsymbol{w}$ is a pseudolength, too, for all positive number α.

The Lorentz product of future-lightlike vectors and futurelike vectors is negative (see the mathematical supplement). Thus,

$$-\boldsymbol{u} \cdot \boldsymbol{w} > 0$$

for all absolute velocities $\boldsymbol{u}$ and light directions $\boldsymbol{w}$.

12.5.2 Absolute light propagation

The history of a free light signal is described by straight line in spacetime, similarly to the history if an inertial material point.

A fundamental property of light signals is that their history is independent of their source; this absolute light propagation is shown by the following experimental facts (which were basic tools in the construction of the relativistic spacetime model): in the space of an inertial observer

1) *the path of a free light signal is a straight line,*

2) *every straight line can be the path of a light signal,*

3) *a light signal is faster than any material point on the same path,*

4) *the motion of any light sinal can be arbitrarily approximated by the motion of a material point.*

Now we show that these assertions are true in the spacetime model

Let $\boldsymbol{u}$ be an inertial observer and $\boldsymbol{w}$ a light direction.

To prove 1) we can argue as in 2.8.1, taking $\boldsymbol{w}$ instead of $\boldsymbol{u}'$: the path in the $\boldsymbol{u}$-space of a light signal with direction w is the straight line directed by $\boldsymbol{w} + \mathbb{R}\boldsymbol{u}$. According to the standard representation of $\boldsymbol{u}$-space vectors, this vector corresponds to $\boldsymbol{\sigma_u} \cdot \boldsymbol{w}$.

To prove 2) we have to show that for all $0 \neq \boldsymbol{v} \in \frac{\mathbf{E_u}}{\mathbb{I}}$ there is a light direction $\boldsymbol{w}$ such that $\boldsymbol{\sigma_u} \cdot \boldsymbol{w} = \boldsymbol{v}$. This is simple: $\boldsymbol{w} := |\boldsymbol{v}|\boldsymbol{u} + \boldsymbol{v}$.

To prove 3) we let us observe that if the light signal with direction w and the material point with absolute velocity $\boldsymbol{u}'$ move on the same path in $\boldsymbol{u}$-space then there is a number $\beta' > 0$ such that $\boldsymbol{\sigma_u} \cdot \boldsymbol{w} = \beta'\boldsymbol{\sigma_u} \cdot \boldsymbol{u}'$. Rearranging the equality

$$\boldsymbol{w} + (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{u} = \beta'(\boldsymbol{u}' + (\boldsymbol{u} \cdot \boldsymbol{u}')\boldsymbol{u})$$

we obtain

$$\boldsymbol{w} = \alpha\boldsymbol{u}' + (\alpha\boldsymbol{u} \cdot \boldsymbol{u}' - \boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{u} =: \beta'\boldsymbol{u}' - \beta\boldsymbol{u}$$

According to 2.8.2 and 2.8.3, the motion of the light signal is faster than that of the material point if β' and β are positive. It is known that $\beta' > 0$; if β were negative then the right hand side would be futurelike (see the mathematical supplement) and the left hand side would be future-lightlike which is impossible.

Assertion 4) can be formulated as follows: given an absolute velocity $\boldsymbol{u}$ (inertial observer) and a light direction $\boldsymbol{w}$ (light signal), then for all numbers $\beta > 0$ there is an absolute velocity $\boldsymbol{u}'$ (inertial material point) and number $\beta' > 0$ such that $\boldsymbol{w} = \beta'\boldsymbol{u}' - \beta\boldsymbol{u}$. This is obvious because the sum of a futurelike vector and a future-lightlike vector is futurelike (see the mathematical supplement); thus, $\beta\boldsymbol{u} + \boldsymbol{w}$ is futurelike and so $\boldsymbol{u}' := \frac{\beta\boldsymbol{u}+\boldsymbol{w}}{|\beta\boldsymbol{u}+\boldsymbol{w}|}$.

The proved facts imply:

The history of two light signals is the same if they arise in the same world point and they move on the same path in the space of an arbitrary inertial observer.

This is an important property, called absolute light propagation which we prove independently of the previous demonstrations.

Let the light signals with $\boldsymbol{w} \in \mathtt{V}^{\to}$ and $\boldsymbol{w}' \in \mathtt{V}^{\to}$ move in the same direction in the $\boldsymbol{u}$-space. Since instead of w' we can take its arbitrary positive multiples, we can choose it in such a way that $\boldsymbol{\sigma_u} \cdot \boldsymbol{w} = \boldsymbol{\sigma_u} \cdot \boldsymbol{w}'$ be satisfied. This means that there is a positive number α such that $\boldsymbol{w}' = \boldsymbol{w} + \alpha\boldsymbol{u}$. Taking the the Lorentz square of both sides we get $0 = 2\alpha(\boldsymbol{u} \cdot \boldsymbol{w}) - \alpha^2$. Further, the Lorentz product of both sides by $\boldsymbol{u}$ results in $(\boldsymbol{u} \cdot \boldsymbol{w}') = \boldsymbol{u} \cdot \boldsymbol{w} - \alpha$. The square of this equality together with the previous one yields $\boldsymbol{u} \cdot \boldsymbol{w}' = \boldsymbol{u} \cdot \boldsymbol{w}$ implying $\alpha = 0$ i.e. $\boldsymbol{w} = \boldsymbol{w}'$.

Two world lines are on Figure 12.5 symbolizing two lamps. When they meet both emit a light signal. The light signals propagate together in spacetime.

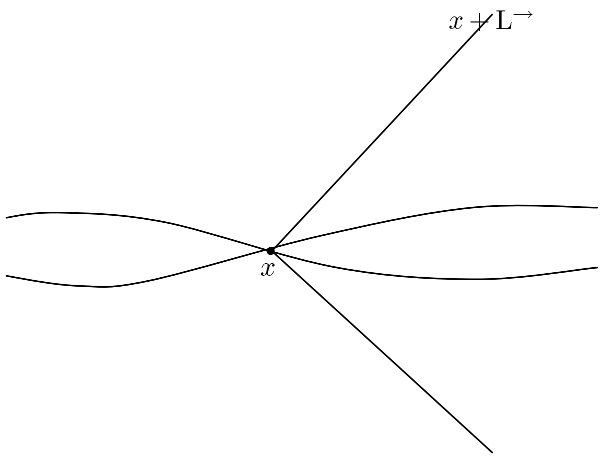

Figure 12.5 Absolute light propagation

12.5.3 Round-way speed of light

A most important experimental fact regarding light signals (which was a basic tool in the construction of the relativistic spacetime model) is the following:

5) *In the space of every inertial observer the round-way speed of light is homogenous and isotropic.*

In more details this can be expounded as follows. Let a light signal, starting form a space point of an inertial observer, return to the start after running on an arbitrary path; the ratio of the length of the path to the time interval between the start and the arrival is independent of the position of the source and the path.

We demonstrate that this is true in the spacetime model; for the sake of simplicity, we consider only two-way speed i.e. the path is straight line segment to and fro between a source and a mirror.

Let us consider an inertial observer $\boldsymbol{u}$. Let$\boldsymbol{a}$ and $\boldsymbol{a}'$ be the future-lightlike vector of the light signal from the source to the mirror and from the mirror to the source, respectively (see Figure 11.1).

If $2\boldsymbol{t}$ is the time interval in the source between the start and the arrival then $\boldsymbol{a} + \boldsymbol{a}' = 2\boldsymbol{t}\boldsymbol{u}$, so $2\boldsymbol{t} = -\boldsymbol{u} \cdot (\boldsymbol{a} + \boldsymbol{a}')$.

The distance between the source and the mirror, according to (V31), is $\boldsymbol{d} := |\boldsymbol{\sigma_u} \cdot \boldsymbol{a}| = |\boldsymbol{\sigma_u} \cdot \boldsymbol{a}'|$. Since $|\boldsymbol{\sigma_u} \cdot \boldsymbol{a}|^2 = |\boldsymbol{a} + (\boldsymbol{u} \cdot \boldsymbol{a})\boldsymbol{u}|^2 = (\boldsymbol{u} \cdot \boldsymbol{a})^2$ and a similar relation holds for $\boldsymbol{a}'$, too, $2\boldsymbol{d} = -\boldsymbol{u} \cdot (\boldsymbol{a} + \boldsymbol{a}') = \boldsymbol{t}$; thus, the two-way speed of light is $2\frac{\boldsymbol{d}}{2\boldsymbol{t}} = 1$.

The result is independent of the observer, the positions of the source and the mirror.

12.5.4 Propagation of light signals

Let us consider two inertial observers, $\boldsymbol{u}$ and $\boldsymbol{u}'$, and a light signal with direction $\boldsymbol{w}$ in spacetime. The directions of motion of $\boldsymbol{u}'$ and w in the $\boldsymbol{u}$-space are $\boldsymbol{\sigma_u} \cdot \boldsymbol{u}'$ and $\boldsymbol{\sigma_u} \cdot \boldsymbol{w}$, respectively; it is worth taking some multiples of them, namely the vectors $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} := \frac{\boldsymbol{u}'}{-\boldsymbol{u}\cdot\boldsymbol{u}'} - \boldsymbol{u}$ and $\mathbf{v}_{\boldsymbol{wu}} := \frac{\boldsymbol{w}}{-\boldsymbol{u}\cdot\boldsymbol{w}} - \boldsymbol{u}$ (later we see that these are the standard relative velocities). It is simple fact that $|\mathbf{v}_{\boldsymbol{wu}}| = 1$. Similar formulae give the directions of motion in the $\boldsymbol{u}'$-space.

Then we can state:

If the light signal moves in $\mathbf{u}$*-space in the same direction as* $\mathbf{u}'$ *moves then the same light signal moves in* $\mathbf{u}'$*-space in the opposite direction as* $\mathbf{u}$ *moves i.e.*

– *if* $\mathbf{v}_{\boldsymbol{wu}} = \frac{\mathbf{v}_{u'u}}{|\mathbf{v}_{u'u}|}$ *then* $\mathbf{v}_{\boldsymbol{wu}'} = -\frac{\mathbf{v}_{uu'}}{|\mathbf{v}_{uu'}|}$.

The proof runs as follows. With the notation

$$v := |\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}| = |\mathbf{v}_{\boldsymbol{uu}'}| \tag{V37}$$

we have

$$-\boldsymbol{u} \cdot \boldsymbol{u}' = \frac{1}{\sqrt{1 - v^2}},$$

so

$$\frac{\boldsymbol{w}}{-\boldsymbol{u} \cdot \boldsymbol{w}} - \boldsymbol{u} = \frac{1}{v}(\sqrt{1 - v^2}\boldsymbol{u}' - \boldsymbol{u})$$

from which we get

$$\frac{\boldsymbol{w}}{-\boldsymbol{u}\cdot\boldsymbol{w}} = \frac{\sqrt{1-v}}{v}(\sqrt{1+v}\boldsymbol{u}' - \sqrt{1-v}\boldsymbol{u}). \tag{V38}$$

The Lorentz product by $\boldsymbol{w}$ yields

$$-\boldsymbol{u}'\cdot\boldsymbol{w}\sqrt{1+v} = -\boldsymbol{u}\cdot\boldsymbol{w}\sqrt{1-v}.$$

Then $\frac{\boldsymbol{w}}{-\boldsymbol{u}'\cdot\boldsymbol{w}} = \frac{\sqrt{1+v}\boldsymbol{w}}{-\boldsymbol{u}\cdot\boldsymbol{w}\sqrt{1-v}}$ which, finally, results in

$$\frac{\boldsymbol{w}}{-\boldsymbol{u}'\cdot\boldsymbol{w}} - \boldsymbol{u}' = -\frac{1}{v}(\sqrt{1-v^2}\boldsymbol{u} - \boldsymbol{u}').$$

12.5.5 Delay of light signals

The inertial observer $\boldsymbol{u}$ sends two light signals from the same source towards $\boldsymbol{u}'$ i.e. parallel to the direction of motion of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$. Let $\boldsymbol{t}$ be the time interval between the starts of the light signals. Depending on whether the direction of light signals is $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$ or $-\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$, thy arrive at a $\boldsymbol{u}'$-space point with time delay $\boldsymbol{t}'_+$ and $\boldsymbol{t}'_-$.

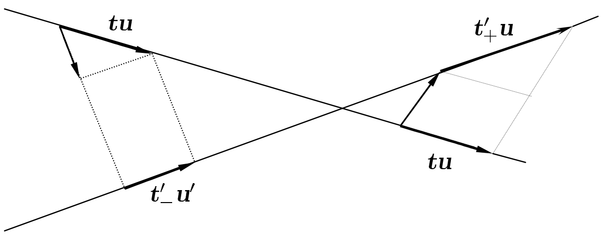

Figure 12.6 Delay of light signals

Then as Figure 12.6 shows, the light directions in spacetime are $\boldsymbol{t}'_+\boldsymbol{u}' - \boldsymbol{tu}$ and $\boldsymbol{tu} - \boldsymbol{t}'_-\boldsymbol{u}'$, respectively. Of course, we can take the light directions $\boldsymbol{w}_\pm$ in such a way that $\frac{\boldsymbol{w}_+}{-\boldsymbol{u}\cdot\boldsymbol{w}_+} = \boldsymbol{t}'_+\boldsymbol{u}' - \boldsymbol{tu}$ and $\frac{\boldsymbol{w}_-}{-\boldsymbol{u}\cdot\boldsymbol{w}_-} = \boldsymbol{tu} - \boldsymbol{t}'_-\boldsymbol{u}'$ be satisfied.

Equality (V38) gives immediately that

$$\boldsymbol{t}'_+ = \sqrt{\frac{1+v}{1-v}}\ \boldsymbol{t}, \qquad \boldsymbol{t}'_- = \sqrt{\frac{1-v}{1+v}}\ \boldsymbol{t}.$$

12.5.6 Physical method of establishing boosts

The boosts between inertial observers are defined in a 'mathematically natural' way (see 12.4.3). Now a physical procedure is described how a space vector of an observer can be identified with a space vector of another observer, with the aid of absolute light propagation.

To get an intuition, let us consider two (sufficiently large) spaceships as two observers. I am sitting in one of them ($\boldsymbol{u}$), you are sitting in the other ($\boldsymbol{u}'$). I see that you are coming towards me (or you are moving away from me) and you have a similar experience regarding me.

Recall (see 12.5.4) that if I send a light signal towards you (in the direction in which you are moving with respect to me) then the light signal arrives at you and you see that it comes from me (from a direction which is opposite to my movement with respect to you)

This simple fact gives the idea of making a correspondence (boost) between our space vectors.

I take a plane in my space which is orthogonal to the direction of your motion. Any vector in my space can be given as a sum of two components, one of them lying in that plane and the other being parallel to the direction of your motion; thus it suffices to give how these components are boosted.

Taking a vector in that plane, I send towards you a green light signal from the base point and a red light signal from the end point of the vector.The light signals hit a plane in your space which is orthogonal to the direction of my motion. The vector determined by these green and red points in your plane will correspond to the vector in my plane.

Then taking a vector in my space whose direction equals the direction of your motion, I send two yellow light signals at once from the end point of the vector; one of them towards you, and the other towards the base point of the vector where it reflects; then the two light signals hit your plane at a single point with a time delay from which you can calculate the length of my vector; the vector in your space with that length and opposite to my motion will correspond to my vector. More closely, let my vector have the direction of $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$ and length $\boldsymbol{t}$; then the second light signal reaches the end point again after a time interval $2\boldsymbol{t}$. Thus, if you measure a time delay $2\boldsymbol{t}'_+$ between the arrivals of the light signals then your corresponding vector will have the direction of $-\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}$ and length $\boldsymbol{t}'_+\sqrt{\frac{1-v}{1+v}}$.

A similar procedure is effectuated by blue light signals in the case of a vector whose direction is opposite to the direction of your motion (the colour of the light signals indicates the direction).

It remains the question, how to measure v; the answer will be given in 12.6.5.

12.6 Standard inertial frames

12.6.1 Standard synchronization

In everyday practice we establish a synchronization by light signals (radio signals) on the (unbased!) conviction that light propagates in the space of an inertial observer homogeneously and isotropically at the speed $c := 2,99793... \cdot 10^8$m/s.

According to this conviction, a one-way light signal covers a distance d in time d/c; thus if the light signal starts in the instant t from a 'centre' of the observer, then it arrives at the instant $t + d/c$ at a place having distance d from the centre. An example: a pip is launched in Greenwich at midnight; in Budapest, 1800 km far from Greenwich, when the pip arrives, the clock is set to $3 \cdot 10^{-5}$ seconds after midnight. Similarly, the clocks in Vienna, Prague etc. are set according to the pips arriving from Greenwich.

This method means in the spacetime model that the observer establishes – if it can! – a synchronization in such a way that the one-way light speed be equal to the round-way speed. Recalling what has been said in 12.5.3, the synchronization in question means that the the reflection of the light signal at the mirror be simultaneous with the occurrence of the source which is in the middle between the start and the return of the light signal.

Using the notations in 12.5.3, $\boldsymbol{q} := \boldsymbol{a} - \boldsymbol{t}\boldsymbol{u} = \boldsymbol{t}\boldsymbol{u} - \boldsymbol{a}'$ is the vector between the simultaneous occurrences of the source and the mirror. Thus both $\boldsymbol{q} + \boldsymbol{t}\boldsymbol{u}$ and $\boldsymbol{t}\boldsymbol{u} - \boldsymbol{q}$ are future-lightlike, so their pseudolength is zero, $\boldsymbol{q} \cdot \boldsymbol{q} \pm 2\boldsymbol{t}\boldsymbol{u} \cdot \boldsymbol{q} - \boldsymbol{t}^2 = 0$ which implies $\boldsymbol{u} \cdot \boldsymbol{q} = 0$ i.e. $\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}}$.

Our results shows that this synchronization is independent of the source ('centre'). Thus, according to such a synchronization established by the inertial observer $\boldsymbol{u}$, the world point x and y are simultaneous if and only if $y - x \in \mathbf{E}_{\boldsymbol{u}}$. In other words, the collection of world points simultaneous with x in such a way is $x + \mathbf{E}_{\boldsymbol{u}}$.

This uniform synchronization is called the **standard synchronization** of the observer.

We call attention to that – since $\mathbf{E}_{\boldsymbol{u}} \neq \mathbf{E}_{\boldsymbol{u}'}$ if $\boldsymbol{u} \neq \boldsymbol{u}'$ – **different standard synchronizations correspond to different inertial observers.**

According to 3.2.3, a uniform synchronization and an inertial observer together form an inertial frame. The inertial observer $\boldsymbol{u}$ and its standard synchronization together is called the **$\boldsymbol{u}$-standard inertial frame**.

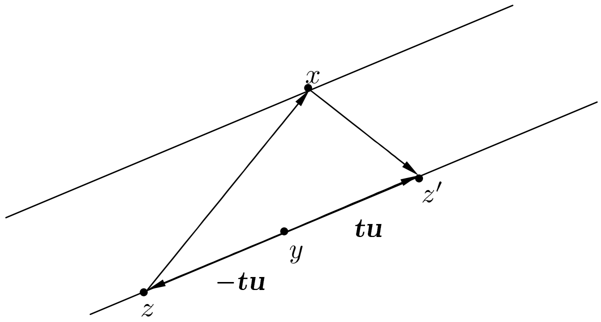

Figure 12.7 Standard synchronization

The method of such synchronizations shows well what has been emphasized earlier: a synchronization is not a physical fact, it is an artificial (human) construction. We are not obliged to use the standard synchronization. We could synchronize in such a way that the light speed from Greenwich to Budapest be the half of the light speed from Budapest to Greenwich. The standard synchronization distinguished by its simple and nice feature

Keep in mind that fundamental physical laws must be formulated without synchronizations.

It is worth to put the following questions. Here we considered the Earth an inertial observer though we know that it is not. Can we apply on the Earth in practice the treated light speed synchronization and if yes, does it not lead to some problem? Further, there is another, ancient method of synchronization-which is established by the position of stars (or the Sun): midnight (midday) is in Greenwich and in Budapest if the stars have (te Sun has) a prescribed position. Are the two synchronizations the same? Later we give an answer.

12.6.2 Standard time points

The $\boldsymbol{u}$-standard instants are the hyperplanes directed by $\mathbf{E}_{\boldsymbol{u}}$, their collection is the **$\boldsymbol{u}$-standard time**, $\mathtt{I}_{\boldsymbol{u}} := \mathtt{M}/\mathbf{E}_{\boldsymbol{u}}$. In order to make a comparison with the nonrelativistic case, we introduce the **$\boldsymbol{u}$-time evaluation**

$$\tau_{\boldsymbol{u}} : \mathtt{M} \to \mathtt{I}_{\boldsymbol{u}}, \qquad x \mapsto x + \mathbf{E}_{\boldsymbol{u}}.$$

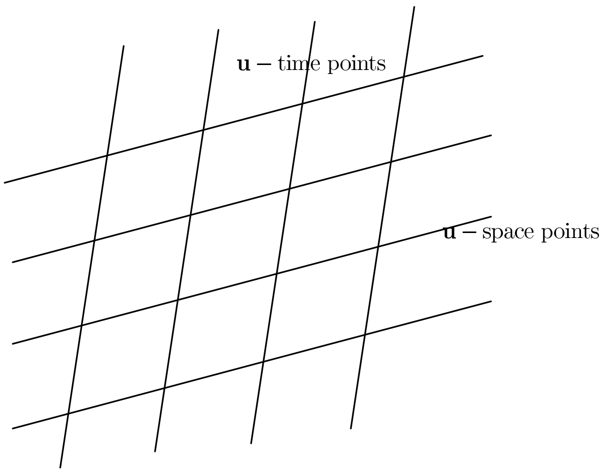

Figure 12.8 $\boldsymbol{u}$-time points

The difference of two $\boldsymbol{u}$-instants t and s is defined as the proper time duration in an arbitrary $\boldsymbol{u}$-space point between them: $t - s := |y - x| = \boldsymbol{u} \cdot (y - x)$ if $y \in t, x \in s$ and $y - x$ is parallel to $\boldsymbol{u}$.

If $y \in t$ and $x \in s$ are arbitrary then there are $x' \in s$ and $\boldsymbol{y}' \in t$ such that $y - x'$ is parallel to $\boldsymbol{u}$-val and, of course, $x - x' \in \mathbf{E}_{\boldsymbol{u}}$. Therefore we have

$$t - s = -\boldsymbol{u} \cdot (y - x), \qquad (y \in t, x \in s). \tag{V39}$$

In other words,

$$(y + \mathbf{E}_{\boldsymbol{u}}) - (x + \mathbf{E}_{\boldsymbol{u}}) = \tau_{\boldsymbol{u}}(y) - \tau_{\boldsymbol{u}}(x) := -\boldsymbol{u} \cdot (y - x) \tag{V40}$$

where, again for a comparison with the nonrelativistic case, we introduced the notation $\boldsymbol{\tau}_{\boldsymbol{u}}$ (which appeared in the construction of the model).

The subtraction above turns $\mathtt{I}_{\boldsymbol{u}}$ into a one-dimensional affine space over $\mathbb{I}$ and $\tau_{\boldsymbol{u}}$ becomes an affine map over the linear map $\boldsymbol{\tau}_{\boldsymbol{u}} = -\boldsymbol{u}\cdot$.

12.6.3 Standard relative velocities

Recall that the nonrelativistic spacetime model admits a single synchronization(the absolute one), thus, there we need not bother about synchronizations, we can say observer instead of reference frame, and we can say relative velocity without specifying a synchronization. Let us be cautious not to follow this practice when dealing with the relativistic spacetime model, there are a lot of misunderstandings and, of course, not founded paradoxes which aroused from the fact that in usual treatments observers and reference frames are not distinguished and one speaks about relative velocity without specifying a synchronization.

Let us consider a standard inertial reference frame $\boldsymbol{u}$ and the world line of an inertial material point i.e. straight line in spacetime, directed by an absolute velocity $\boldsymbol{u}'$. If $\boldsymbol{u}' \neq \boldsymbol{u}$ then the material point moves with respect to the observer; the relative velocity of the material point with respect to the reference frame is defined as follows.

Let x and y be world points on the world line i.e. $y - x = \boldsymbol{s}\boldsymbol{u}'$ for some $\boldsymbol{s} \in I$. The material point meets the $\boldsymbol{u}$-space points $x + \mathbb{I}\boldsymbol{u}$ and $y + \mathbb{I}\boldsymbol{u}$ at the $\boldsymbol{u}$-instants $x + \mathbf{E}_{\boldsymbol{u}}$ and $y + \mathbf{E}_{\boldsymbol{u}}$, respectively. The $\boldsymbol{u}$-space vector between these $\boldsymbol{u}$-space points is $(y+\mathbb{I}\boldsymbol{u}) - (x+\mathbb{I}\boldsymbol{u}) = \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (y-x) = (\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{u}')\boldsymbol{s}$; the time period between the two meetings is $(y + \mathbf{E}) - (x + \mathbf{E}) = -\boldsymbol{u} \cdot (y - x) = \boldsymbol{s}(-\boldsymbol{u} \cdot \boldsymbol{u}')$. Consequently, the average $\boldsymbol{u}$-relative velocity between the two $\boldsymbol{u}$-space points equals

$$\frac{(y+\mathbb{I}\boldsymbol{u}) - (x+\mathbb{I}\boldsymbol{u})}{(y+\mathbf{E}_{\boldsymbol{u}}) - (x+\mathbf{E}_{\boldsymbol{u}})} = \frac{\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{u}'}{-\boldsymbol{u} \cdot \boldsymbol{u}'} = \frac{\boldsymbol{u}'}{-\boldsymbol{u} \cdot \boldsymbol{u}'} - \boldsymbol{u}.$$

The result is independent of the choice of the $\boldsymbol{u}$-instants: an inertial material point moves uniformly with respect to a standard inertial reference frame. The **standard relative velocity** of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$ is

$$\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} := \frac{\boldsymbol{u}'}{-\boldsymbol{u} \cdot \boldsymbol{u}'} - \boldsymbol{u} \in \frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}}.$$

Note, that contrary to the relativistic case, this is not just the difference of the two absolute velocities.

The standard relative velocities with respect to an inertial reference frame form a Euclidean space. As a consequence, in contrast to absolute velocities,

– there is a zero $\boldsymbol{u}$-relative velocity,
– a $\boldsymbol{u}$- relative velocity has a magnitude,
– the angle between two $\boldsymbol{u}$-relative velocities makes sense.

Two relative velocities corresponding two different inertial frames are in different three-dimensional Euclidean vector spaces, so the angle between them is not necessarily meaningful.

Interchanging the role of the absolute velocities, we have

$$\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'} = \frac{\boldsymbol{u}}{-\boldsymbol{u}'\cdot\boldsymbol{u}} - \boldsymbol{u}'.$$

As a consequence of the reversed Cauchy inequality, if $\boldsymbol{u} \neq \boldsymbol{u}'$ then $-\boldsymbol{u}\cdot\boldsymbol{u}' > 1$ which implies

$$\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'} \neq -\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} \quad \text{if} \quad \boldsymbol{u} \neq \boldsymbol{u}'.$$

The **reciprocity of relative velocities**, contrary to the nonrelativistic case, **does not hold**. Unfortunately, the reciprocity is taken for granted in usual treatments which results in paradoxes (see 16.1).

It is true, however, that the boosts map the corresponding relative velocities into the opposit of each other. Recall the formulae from paragraph 11.2.3 and 12.4.3 :

– $\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}$ and $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$ are in different three-dimensional Euclidean vector spaces, $\frac{\mathbf{E}_{\boldsymbol{u}'}}{\mathbb{I}}$ and $\frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}}$, respectively,
– both $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$ and $\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}$ are orthogonal to $\mathbf{E}_{\boldsymbol{u}} \cap \mathbf{E}_{\boldsymbol{u}'}$,
– $\boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} = -\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}$,
– $v := |\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}|^2 = |\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}|^2 = 1 - \frac{1}{(\boldsymbol{u}\cdot\boldsymbol{u}')^2} < 1$, so

$$-\boldsymbol{u}\cdot\boldsymbol{u}' = \frac{1}{\sqrt{1-v^2}}.$$

The last quantity appears many times in formulae, its usual name is the **relativistic factor**.

Lastly we remind again that relative velocity here concerns the standard synchronization. Other synchronization gives other relative velocity (see 17.4).

Similarly, the free light signal with light direction (light line directed by) $\boldsymbol{w}$ has the standard $\boldsymbol{u}$-relative velocity

$$\mathbf{v}_{\boldsymbol{w}\boldsymbol{u}} := \frac{\boldsymbol{w}}{-\boldsymbol{u}\cdot\boldsymbol{w}} - \boldsymbol{u}$$

for which $|\mathbf{v}_{\boldsymbol{w}\boldsymbol{u}}| = 1$ holds i.e. the one-way speed of light is the same in every direction in a standard inertial frame, as it must be, since this was a condition for defining the standard synchronization.

12.6.4 Addition of relative velocities

It is highly important that, contrary to the nonrelativistic case, that the **transitivity of relative velocities does not hold** i.e. except the trivial case

$\boldsymbol{u}'' = \boldsymbol{u}'$ and $\boldsymbol{u}' = \boldsymbol{u}$,

$$\mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}} \neq \mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}'} + \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}.$$

It is a simple fact that such an equality cannot hold, in general, because $\mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}'}$ in the three-dimensional vector space $\mathbf{E}_{\boldsymbol{u}'}/\mathbb{I}$ and the two others in the different three-dimensional vector space $\mathbf{E}_{\boldsymbol{u}}/\mathbb{I}$.

The addition of relative velocities is given by the formula

$$\mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}} = \frac{\beta}{\gamma}\boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'} \cdot \mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}'} + \frac{\alpha(\beta+\gamma)}{\gamma(1+\alpha)}\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} \tag{V41}$$

where

$$\alpha := -\boldsymbol{u}' \cdot \boldsymbol{u} = \frac{1}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}|^2}}, \qquad \beta := -\boldsymbol{u}'' \cdot \boldsymbol{u}' = \frac{1}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}'}|^2}},$$

$$\gamma := -\boldsymbol{u}'' \cdot \boldsymbol{u} = \alpha\beta(1 + \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'} \cdot \mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}'}).$$

Theoretically the proof is simple: substituting the actual expressions of the quantities to both sides, we get an equality; practically, however, it requries rather complicated manipulations; we omit the details.

Lastly, we mention an interesting fact. We know that the boost from $\boldsymbol{u}''$ to $\boldsymbol{u}'$ followed by the boost from $\boldsymbol{u}'$ to $\boldsymbol{u}''$ equals the boost from $\boldsymbol{u}$ to $\boldsymbol{u}''$ if and only if (see (V36))

– $\boldsymbol{u}$, $\boldsymbol{u}'$ and $\boldsymbol{u}''$ are coplanar.

It is simple to show that this is equivalent to that

– $\mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}}$ is parallel to $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$ and $\mathbf{v}_{\boldsymbol{u}''\boldsymbol{u}'}$ is parallel to $\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}$.

12.6.5 How to measure a standard relative velocity

In usual practice, using radar apparatus, one measures relative velocity on the base that the one-way velocity of light is the same everywhere in every direction. Two light signals are emitted from a radar apparatus (being in a space point of the Earth considered an inertial observer) at the standard instant t and $t + \boldsymbol{s}$, respectively. The light signals are reflected on a car and return after the time intervals $2\boldsymbol{t}_1$ and $\boldsymbol{t}_2$, respectively. Then the distances of the car from the apparatus at the standard instants $t + \boldsymbol{t}_1$ and $t + \boldsymbol{s} + \boldsymbol{t}_2$ are $c\boldsymbol{t}_1$ and $c\boldsymbol{t}_2$, respectively. Thus, the magnitude of the relative velocity is

$$v := \left|\frac{c\boldsymbol{t}_2 - c\boldsymbol{t}_1}{(t+\boldsymbol{s}+\boldsymbol{t}_2) - (t+\boldsymbol{t}_1)}\right| = \frac{\boldsymbol{t}_1 - \boldsymbol{t}_2}{\boldsymbol{s} - (\boldsymbol{t}_1 - \boldsymbol{t}_2)}c.$$

Here only time intervals, measured in a single space point (apparatus) appear.

12.7 Standard vectorial splitting and transformation rules

12.7.1 Splitting

Since $\mathbb{I}\boldsymbol{u}$ and $\mathbf{E}_{\boldsymbol{u}}$) are transverse linear subspaces and span $\mathbf{M}$, every spacetime vector can be given uniquely as a sum of vectors in these subspaces. Using the notation (V30), for $\boldsymbol{x} \in \mathbf{M}$ we have $\boldsymbol{x} = -\boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{x}) + \boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{x}$; we say that $\boldsymbol{u}$ **splits** $\boldsymbol{x}$ into the $\boldsymbol{u}$-timelike component $-\boldsymbol{u}\cdot\boldsymbol{x}$ and the $\boldsymbol{u}$**-spacelike component** $\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{x}$. The linear bijection

$$\boldsymbol{h}_{\boldsymbol{u}} := \mathbf{M} \to \mathbb{I} \times \mathbf{E}, \qquad \boldsymbol{x} \mapsto (-\boldsymbol{u}\cdot\boldsymbol{x}, \boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{x}) \tag{V42}$$

is the **standard splitting of spacetime vectors according to** $\boldsymbol{u}$.

For a clear comparison with the nonrelativistic case let us take the notation

$$\boldsymbol{\tau}_{\boldsymbol{u}} : \mathbf{M} \to \mathbb{I}, \qquad \boldsymbol{x} \mapsto -\boldsymbol{u}\cdot\boldsymbol{x};$$

then $\mathbf{E}_{\boldsymbol{u}}$ is the kernel of the linear map $\boldsymbol{\tau}_{\boldsymbol{u}}$ and the splitting has the form

$$\boldsymbol{h}_{\boldsymbol{u}}\cdot\boldsymbol{x} = (\boldsymbol{\tau}_{\boldsymbol{u}}\cdot\boldsymbol{x}, \boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\boldsymbol{x}).$$

Note that

$$\boldsymbol{h}_{\boldsymbol{u}}^{-1}(\boldsymbol{t},\boldsymbol{q}) = \boldsymbol{t}\boldsymbol{u} + \boldsymbol{q} \qquad (\boldsymbol{t},\boldsymbol{q}) \in \mathbb{I}\times\mathbf{E}_{\boldsymbol{u}}.$$

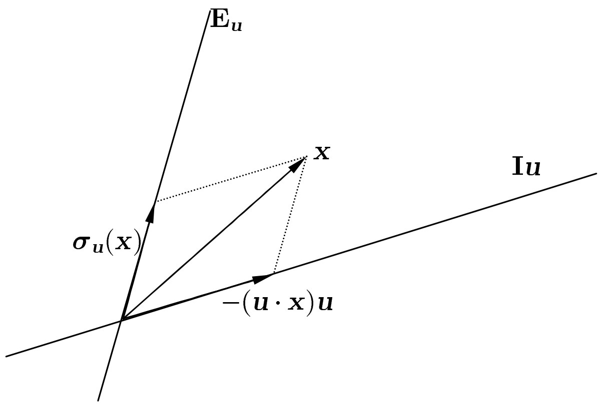

Figure 12.9 Splitting of vectors

Of course, the tensor products and quotients of $\mathbf{M}$ by measure lines such as $\frac{\mathbf{M}}{\mathbb{I}}$ or $\frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}}$ are split corresponding to the above formula, because the multiplication and devision by element of measure lines can be interchanged with linear maps. For instance, the $\boldsymbol{u}$-timelike component of an absolute velocity $\boldsymbol{u}'$ is $-\boldsymbol{u}\cdot\boldsymbol{u}' =\geq 1$, its $\boldsymbol{u}$-spacelike component is

$$\boldsymbol{\sigma_u}\cdot\boldsymbol{u}' = \boldsymbol{u}' + (\boldsymbol{u}\cdot\boldsymbol{u}')\boldsymbol{u}. \tag{V43}$$

Thus, according to the formulae of 12.6.3,

– the $\boldsymbol{u}$-timelike component of $\boldsymbol{u}'$ is the relativistic factor between $\boldsymbol{u}'$ and $\boldsymbol{u}$,

– the $\boldsymbol{u}$-spacelike component of $\boldsymbol{u}'$ is the standard relative velocoty of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$, multiplied by the relativistic factor.

As a consequence,

$$\boldsymbol{h_u}\cdot\boldsymbol{u}' = \frac{1}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}|^2}}(1,\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}).$$

The splitting of vectors determines the splitting of covectors by the formula

$$\boldsymbol{r_u} : \left(\boldsymbol{h}_{\boldsymbol{u}}^{-1}\right)^* : \mathbf{M}^* \to \mathbb{I}^* \times \mathbf{E}_{\boldsymbol{u}}^*.$$

For a covector $\boldsymbol{k}$ and $(\boldsymbol{t},\boldsymbol{q}) \in \mathbb{I}\times\mathbf{E}_{\boldsymbol{u}}$ we have $\big(\boldsymbol{r_u}\cdot\boldsymbol{k})\cdot(\boldsymbol{t},\boldsymbol{q})\big) = \boldsymbol{k}\cdot\boldsymbol{h}_{\boldsymbol{u}}^{-1}(\boldsymbol{t},\boldsymbol{q})\big) = (\boldsymbol{k}\cdot\boldsymbol{u})\boldsymbol{t} + \boldsymbol{k}\cdot\boldsymbol{q}$.

Now we can handle the splitting of covectors in a simpler way than in the nonrelativistic case because of the identification $\mathbf{M}^* \equiv \frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}}$. If the covector is considered a vector (of cotype $\mathbb{I}\otimes\mathbb{I}$) then the dot product means Lorentz product. Thus, we see that

– the **$\boldsymbol{u}$-timelike component** $\boldsymbol{k}\dot{\boldsymbol{u}} = \boldsymbol{u}\cdot\boldsymbol{k}$ in the covectorial splitting is the negative of the $\boldsymbol{u}$-timelike component in the vectorial splitting,

– the **$\boldsymbol{u}$-spacelike component** $\boldsymbol{k}|_{\mathbf{E}_{\boldsymbol{u}}} = \boldsymbol{k}\cdot\boldsymbol{\sigma_u} = \boldsymbol{\sigma_u}\cdot\boldsymbol{k}$ in the covectorial splitting equals the $\boldsymbol{u}$-timelike component in the vectorial splitting.

Concisely:

$$\boldsymbol{r_u} = (\boldsymbol{u}\cdot, \boldsymbol{\sigma_u}).$$

Consequently,

$$\boldsymbol{r}_{\boldsymbol{u}}^{-1}(\boldsymbol{e},\boldsymbol{p}) = -\boldsymbol{e}\boldsymbol{u} + \boldsymbol{p} \left((\boldsymbol{e},\boldsymbol{p}) \in \mathbb{I}^*\times\mathbf{E}_{\boldsymbol{u}}^* \equiv \frac{\mathbb{R}\times\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}\otimes\mathbb{I}}\right).$$

At present the splitting of spactime vectors and covectors appeared as mathematical formulae. Later we shall see that it has a fundamental physical meaning: instead of the vectors themselves, a standard reference frame 'perceives' only their split components. For instance, it will be clear that the $\boldsymbol{u}$-spacelike component of $\boldsymbol{u}'$ divided by its $\boldsymbol{u}$-timelike component is just the relative velocity of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$.

12.7.2 Transformation rules

Different inertial observers split spacetime vectors differently. The comparison of different splittings is not so simple as in the nonrelativistic case because the split forms are elements of different vector spaces. More closely, the split forms $(\boldsymbol{t}, \boldsymbol{q})$ and $(\boldsymbol{t}', \boldsymbol{q}')$ of the same vector according to $\boldsymbol{u}$ and $\boldsymbol{u}'$ are elements of $\mathbb{I}\times\mathbf{E}_{\boldsymbol{u}}$ and $\mathbb{I}\times\mathbf{E}_{\boldsymbol{u}'}$, respectively. Thus, $\boldsymbol{h}_{\boldsymbol{u}'}\cdot\boldsymbol{h}_{\boldsymbol{u}}^{-1} : \mathbb{I}\times\mathbf{E}_{\boldsymbol{u}} \to \mathbb{I}\times\mathbf{E}_{\boldsymbol{u}'}$ is not suitable for the comparison.

Now we can utilize the boost which makes a natural correspondence between $\mathbf{E}_{\boldsymbol{u}'}$ and $\mathbf{E}_{\boldsymbol{u}}$. Namely, the **vectorial transformation rule** from the $\boldsymbol{u}$-splitting to the $\boldsymbol{u}'$-splitting is defined to be

$$\boldsymbol{h}_{\boldsymbol{u}'\boldsymbol{u}} := (1, \boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'})\left(\boldsymbol{h}'_{\boldsymbol{u}}\cdot\boldsymbol{h}_{\boldsymbol{u}}^{-1}\right) \tag{V44}$$

where $(1, \boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'})$ means that the first component (an element of $\mathbb{I}$) is multiplied by 1 and the $\boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'}$ is applied to the second component (an element of $\mathbf{E}_{\boldsymbol{u}'}$).

After some manipulations, with the standard relative velocity $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} := \frac{\boldsymbol{u}'}{-\boldsymbol{u}\cdot\boldsymbol{u}'} - \boldsymbol{u}$ and the notations $v := |\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}|$,

$$\kappa(v) := \frac{1}{\sqrt{1-|v|^2}},$$

$$\boldsymbol{J}(\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}) := \frac{1}{\kappa(v)}\left(\mathrm{id}_{\mathbf{E}_{\boldsymbol{u}}} + \frac{\kappa(v)^2}{\kappa(v)+1}\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}\otimes\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}\right)$$

we obtain the transformation rule (a linear map $\mathbb{I}\times\mathbf{E}_{\boldsymbol{u}} \to \mathbb{I}\times\mathbf{E}_{\boldsymbol{u}}$) in a matrix form as follows:

$$\boldsymbol{h}_{\boldsymbol{u}'\boldsymbol{u}} = \kappa(v)\begin{pmatrix} 1 & -\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} \\ -\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} & \boldsymbol{J}(\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}) \end{pmatrix}.$$

This is called the **Lorentz transformation rule**.

Thus, if $(\boldsymbol{t}, \boldsymbol{q})$ is the split form of a vector according to $\boldsymbol{u}$ then its split form $(\boldsymbol{t}', \boldsymbol{q}')$ according to $\boldsymbol{u}'$ **boosted** to $\boldsymbol{u}$ is

$$\boldsymbol{t}' = \kappa(v)(\boldsymbol{t} - \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}\cdot\boldsymbol{q}), \qquad \boldsymbol{q}' = \kappa(v)(\boldsymbol{J}(\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}})\cdot\boldsymbol{q} - \boldsymbol{t}\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}).$$

We get a simpler form if the element $\boldsymbol{q}$ of $\mathbf{E}_{\boldsymbol{u}}$ is given as a sum of two components, $\boldsymbol{q}_{\|}$ being parallel to $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$ and $\boldsymbol{q}_{\perp}$ being perpendicular (orthogonal) to $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$. Then $(\boldsymbol{t},\boldsymbol{q}) = (\boldsymbol{t},\boldsymbol{q}_{\|}) + (\boldsymbol{0},\boldsymbol{q}_{\perp})$ and

$$\boldsymbol{q}'_{\perp} = \boldsymbol{q}_{\perp},$$

$$\boldsymbol{t}' = \frac{1}{\sqrt{1-v^2}}(\boldsymbol{t} - \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{q}_{\|}), \qquad \boldsymbol{q}'_{\|} = \frac{1}{\sqrt{1-v^2}}(\boldsymbol{q}_{\|} - \boldsymbol{t}\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}).$$

The formulae above are similar to the usual Lorentz transformation rule but are not the same because those concern coordinates, so they concern $\mathbb{R}$ instead of $\mathbb{I}$ and $\mathbb{R}^3$ instead of $\mathbf{E}_{\boldsymbol{u}}$ when "the corresponding axes of the moving coordinate systems are parallel" which is considered a fact not to be explained. **The boost between different observer spaces is hidden hidden in the usual Lorentz transformation rule**.

Because of the Lorentz identification of covectors and vectors the splitting of covectors is essentially (apart from minus sign) the same as the splitting of vectors, the covector transformation rule will be similar to, too. More closely,

$$\boldsymbol{r}_{\boldsymbol{u}'\boldsymbol{u}} := \left((\boldsymbol{h}_{\boldsymbol{u}',\boldsymbol{u}})^{-1}\right)^* = \kappa(v)\begin{pmatrix} 1 & \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} \\ \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} & \boldsymbol{J}(\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}) \end{pmatrix}.$$

Thus, if $(\boldsymbol{e},\boldsymbol{p})$ is the split form of a covector according to $\boldsymbol{u}$ then its split form $(\boldsymbol{e}',\boldsymbol{p}')$ according to $\boldsymbol{u}'$ **boosted** to $\boldsymbol{u}$ is

$$\boldsymbol{p}'_{\perp} = \boldsymbol{p}_{\perp},$$

$$\boldsymbol{e}' = \frac{1}{\sqrt{1-v^2}}(\boldsymbol{e} + \mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{p}_{\|}), \qquad \boldsymbol{p}'_{\|} = \frac{1}{\sqrt{1-v^2}}(\boldsymbol{p}_{\|} + \boldsymbol{e}\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}).$$

12.8 Standard tensorial splitting and transformation rules

12.8.1 Splitting

In a number of physical theories – e.g. in electromagnetism – not only vectors and covectors but various tensors appear, too. The mathematical supplement helps the reader to be familiar with tensors.

The inertial observer $\boldsymbol{u}$ splits the various tensors, i.e. the elements of $\mathbf{M}\otimes\mathbf{M}$, $\mathbf{M}\otimes\mathbf{M}^*$, $\mathbf{M}^*\otimes\mathbf{M}$ and $\mathbf{M}^*\otimes\mathbf{M}^*$. These tenors can be considered as liner maps $\mathbf{M}^*\to\mathbf{M}$, $\mathbf{M}\to\mathbf{M}$, $\mathbf{M}^*\to\mathbf{M}^*$ and $\mathbf{M}\to\mathbf{M}^*$, respectively.

The formulae are obtained as in the nonrelativistic case but in a simpler form. Namely, because of the identification of covectors and vectors, it suffices to consider only e.g. tensors, the split form of the other ones will quite the same apart from some negative sign. As concerns the notations used in the nonrelativistic case, we can stat:

– $-\boldsymbol{u}\cdot$ and $-\cdot\boldsymbol{u}$ are to be taken instead of $\boldsymbol{\tau}$ and $\boldsymbol{\tau}^*$, respectively,

– $\boldsymbol{u}\cdot$ is to be taken instead of $\boldsymbol{u}^*$,

– $\boldsymbol{i}$ is to substituted by the canonical embedding $\boldsymbol{i_u} : \mathbf{E_u} \to \mathbf{M}$ for which $\boldsymbol{i_u} = \boldsymbol{\sigma}^*_{\boldsymbol{u}}$ holds, so $\boldsymbol{i}^*$ is to be substituted by $\boldsymbol{\sigma_u}$.

The standard $\boldsymbol{u}$-split form of $\boldsymbol{G} \in \mathbf{M} \otimes \mathbf{M}$ i.e. $\boldsymbol{G} : \mathbf{M}^* \to \mathbf{M}$ is the tensor

$$\boldsymbol{h_u} \cdot \boldsymbol{G} \cdot \boldsymbol{h}^*_{\boldsymbol{u}} : (\mathbb{I} \times \mathbf{E_u})^* \to (\mathbb{I} \times \mathbf{E_u}).$$

Since $\boldsymbol{h_u} = (\boldsymbol{\tau_u}, \boldsymbol{\sigma_u})$ and $\boldsymbol{h}^*_{\boldsymbol{u}} = (\boldsymbol{\tau}^*_{\boldsymbol{u}}, \boldsymbol{\sigma}^*_{\boldsymbol{u}})$, moreover $(\mathbb{I} \times \mathbf{E})^* = \mathbb{I}^* \times \mathbf{E}^*$, the split tensor can be written in a matrix form,

$$\boldsymbol{h_u} \cdot \boldsymbol{G} \cdot \boldsymbol{h}^*_{\boldsymbol{u}} = \ = \begin{pmatrix} \boldsymbol{u} \cdot \boldsymbol{G} \cdot \boldsymbol{u} & -\boldsymbol{u} \cdot \boldsymbol{G} \cdot \boldsymbol{\sigma}^*_{\boldsymbol{u}} \\ -\boldsymbol{\sigma_u} \cdot \boldsymbol{G} \cdot \boldsymbol{u} & \boldsymbol{\sigma_u} \cdot \boldsymbol{G} \cdot \boldsymbol{\sigma}^*_{\boldsymbol{u}} \end{pmatrix},$$

whose component, detailed more explicitly, are

$$-\boldsymbol{u} \cdot \boldsymbol{G} \cdot \boldsymbol{\sigma}^*_{\boldsymbol{u}} = -\boldsymbol{u} \cdot \boldsymbol{G} - \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{G} \cdot \boldsymbol{u}), \quad -\boldsymbol{\sigma_u} \cdot \boldsymbol{G} \cdot \boldsymbol{u} = -\boldsymbol{G} \cdot \boldsymbol{u} - \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{G} \cdot \boldsymbol{u}),$$

$$\boldsymbol{\sigma_u} \cdot \boldsymbol{G} \cdot \boldsymbol{\sigma}^*_{\boldsymbol{u}} = \boldsymbol{G} + \boldsymbol{u} \otimes (\boldsymbol{u} \cdot \boldsymbol{G}) + (\boldsymbol{G} \cdot \boldsymbol{u}) \otimes \boldsymbol{u} + \boldsymbol{u} \otimes \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{G} \cdot \boldsymbol{u}).$$

As a consequence, the $\boldsymbol{u}$-split form of $\boldsymbol{F} \in \mathbf{M}^* \otimes \mathbf{M}^*$ is

$$\begin{pmatrix} \boldsymbol{u} \cdot \boldsymbol{F} \cdot \boldsymbol{u} & \boldsymbol{u} \cdot \boldsymbol{F} + \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{F} \cdot \boldsymbol{u}) \\ \boldsymbol{F} \cdot \boldsymbol{u} + \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{F} \cdot \boldsymbol{u}) & \boldsymbol{F} + \boldsymbol{u} \otimes (\boldsymbol{u} \cdot \boldsymbol{F}) + (\boldsymbol{F} \cdot \boldsymbol{u}) \otimes \boldsymbol{u} + \boldsymbol{u} \otimes \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{F} \cdot \boldsymbol{u}) \end{pmatrix}.$$

Antisymmetric elements of $\mathbf{M}\otimes\mathbf{M}$ and $\mathbf{M}^*\otimes\mathbf{M}^*$ are particularly interesting.

If $\boldsymbol{G}$ is an antisymmetric tensor i.e. $\boldsymbol{G} = -\boldsymbol{G}^*$ then $\boldsymbol{u} \cdot \boldsymbol{G}\boldsymbol{u} = 0$, that is why it has the $\boldsymbol{u}$-split form

$$\begin{pmatrix} 0 & \boldsymbol{G} \cdot \boldsymbol{u} \\ -\boldsymbol{G} \cdot \boldsymbol{u} & \boldsymbol{G} + (\boldsymbol{G} \cdot \boldsymbol{u}) \wedge \boldsymbol{u} \end{pmatrix};$$

here the two 'lower' components determine the other ones, therefore we refer to the split form as

$$(\!(-\boldsymbol{G} \cdot \boldsymbol{u}, \boldsymbol{G} + (\boldsymbol{G} \cdot \boldsymbol{u}) \wedge \boldsymbol{u})\!) \in (\mathbf{E_u} \otimes \mathbb{I}) \times (\mathbf{E_u} \wedge \mathbf{E_u}).$$

They are called **$\boldsymbol{u}$-timelike component** and the **$\boldsymbol{u}$-spacelike component** of $\boldsymbol{G}$, respectively.

It is a simple fact that if $\boldsymbol{G}$ has the $\boldsymbol{u}$-split form $(\!(\boldsymbol{D_u}, \boldsymbol{H_u})\!)$ then

$$\boldsymbol{G} = \boldsymbol{H_u} - \boldsymbol{u} \wedge \boldsymbol{D_u}.$$

Similarly, the $\boldsymbol{u}$-split form of the antisymmetric cotensor $\boldsymbol{F}$ is

$$(\!(\boldsymbol{F} \cdot \boldsymbol{u}, \boldsymbol{F} + (\boldsymbol{F} \cdot \boldsymbol{u}) \wedge \boldsymbol{u})\!),$$

and if $\boldsymbol{F}$ has the $\boldsymbol{u}$-split form $(\!(\boldsymbol{E_u}, \boldsymbol{B_u})\!)$ then

$$\boldsymbol{F} = \boldsymbol{B_u} + \boldsymbol{u} \wedge \boldsymbol{E_u}.$$

12.8.2 Transformation rules

Comparing the tensorial splittings due to different observers, we get the tensorial transformation rules. We treat only the formulae concerning antisymmetric tensors and cotensors.

If $((\boldsymbol{D}, \boldsymbol{H}))$ and $((\boldsymbol{D}', \boldsymbol{H}'))$ are the $\boldsymbol{u}$-split form and the $\boldsymbol{u}'$-split form **boosted** to $\boldsymbol{u}$, respectively, of an antisymmetric tensor, then with the notation $\boldsymbol{v} := \boldsymbol{v_{u'u}}$

$$(\!(\boldsymbol{D}', \boldsymbol{H}')\!) = \boldsymbol{h_{u'u}} \cdot (\!(\boldsymbol{D}, \boldsymbol{H})\!) \cdot \boldsymbol{h}^*_{\boldsymbol{u'u}} = \\ = \kappa(v)^2 \begin{pmatrix} 1 & -\boldsymbol{v} \\ -\boldsymbol{v} & \mathbf{J}(\boldsymbol{v}) \end{pmatrix} \begin{pmatrix} \mathbf{0} & -\boldsymbol{D} \\ \boldsymbol{D} & \boldsymbol{H} \end{pmatrix} \begin{pmatrix} 1 & -\boldsymbol{v} \\ -\boldsymbol{v} & \mathbf{J}(\boldsymbol{v}) \end{pmatrix}.$$

Here, too, the decomposition

$$\boldsymbol{D} = \boldsymbol{D}_\perp + \boldsymbol{D}_\parallel, \qquad \boldsymbol{H} = \boldsymbol{H}_\perp + \boldsymbol{H}_\parallel$$

gives a more tractable form where $\boldsymbol{D}_\perp$ and $\boldsymbol{D}_\parallel$ aare the vectors perpendicular and orthogonal to $\boldsymbol{v}$, respectively, $\boldsymbol{H}_\perp$ and $\boldsymbol{H}_\parallel$ are antisymmetric $\boldsymbol{u}$-tensors whose kernel is perpendicular and parallel to $\boldsymbol{v}$, respectively. Then

$$\boldsymbol{D}'_\parallel = \boldsymbol{D}_\parallel, \qquad \boldsymbol{H}'_\parallel = \boldsymbol{H}_\parallel,$$

$$\boldsymbol{D}'_\perp = \frac{1}{\sqrt{1-v^2}}(\boldsymbol{D}_\perp + \boldsymbol{H}_\perp \cdot \boldsymbol{v}) \qquad \boldsymbol{H}'_\perp = \frac{1}{\sqrt{1-v^2}}(\boldsymbol{v} \wedge \boldsymbol{D}_\perp + \boldsymbol{H}_\perp).$$

Apart from tw minus signs, similar transformation rule holds for antisymmetric cotensors, too. More closely, if $((\boldsymbol{E}, \boldsymbol{B}))$ and $((\boldsymbol{E}', \boldsymbol{B}'))$ are the $\boldsymbol{u}$-split form and the $\boldsymbol{u}'$-split form **boosted** to $\boldsymbol{u}$, respectively, of an antisymmetric cotensor then

$$\boldsymbol{E}'_{\|} = \boldsymbol{E}_{\|}, \qquad \boldsymbol{B}'_{\|} = \boldsymbol{B}_{\|},$$

$$\boldsymbol{E}'_{\perp} = \frac{1}{\sqrt{1-v^2}}(\boldsymbol{E}_{\perp} - \boldsymbol{B}_{\perp}\cdot\boldsymbol{v}) \qquad \boldsymbol{B}'_{\perp} = \frac{1}{\sqrt{1-v^2}}(-\boldsymbol{v}\wedge\boldsymbol{E}_{\perp} + \boldsymbol{B}_{\perp}).$$

12.9 Standard splitting of spacetime and transformatin rules

12.9.1 Splitting

A standard reference frame $\boldsymbol{u}$ characterizes the occurrences by giving when and where they happen: it **splits** spacetime into time and space; to a world point x it assigns the corresponding $\boldsymbol{u}$-time point $\tau(x) = x + \mathbf{E}_{\boldsymbol{u}}$ and the corresponding $\boldsymbol{u}$-space point $\sigma_{\boldsymbol{u}}(x) = x + \mathbb{I}\boldsymbol{u}$. The splitting

$$h_{\boldsymbol{u}} : \mathtt{M} \to \mathtt{I}_{\boldsymbol{u}} \times \mathtt{E}_{\boldsymbol{u}}, \qquad x \mapsto \big(\tau(x), \sigma_{\boldsymbol{u}}(x)\big). \tag{V45}$$

is affine bijection over the vectorial splitting $\boldsymbol{h}_{\boldsymbol{u}}$ as it is well seen from equalities (V40), (V31) and (V42). The inverse of this splitting – which gives the occurrence corresponding to a $\boldsymbol{u}$-time point and a $\boldsymbol{u}$-space point – is

$$h_{\boldsymbol{u}}^{-1} : \mathtt{I} \times \mathtt{E}_{\boldsymbol{u}} \to \mathtt{M}, \qquad (t, q) \mapsto t \cap q.$$

Instead of affine spaces it is often more suitable to deal with the underlying vector spaces; therefore a standard reference frame – corresponding to the everyday usage when time points are represented by time intervals that passed from a given time points and space point are represented by vectors from an origin –, choosing an 'initial' $\boldsymbol{u}$-time point t_o and a $\boldsymbol{u}$-'origin' q_o, vectorizes time and $\boldsymbol{u}$-space by the assignment

$$\mathtt{I} \times \mathtt{E}_{\boldsymbol{u}} \to \mathbb{I} \times \mathbf{E}, \qquad (t, q) \mapsto (t - t_o, q - q_o).$$

Choosing a t_o and a q_o is equivalent to choosing a 'spacetime origin' o: $o = t_o \cap q_o$, $t_o = \tau(o) = o + \mathbf{E}$, $q_o = \sigma_{\boldsymbol{u}}(o) = o + \mathbb{I}\boldsymbol{u}$. Then it is a simple fact that the $\boldsymbol{u}$-splitting of spacetime followed by the vectorization of $\boldsymbol{u}$-time and $\boldsymbol{u}$-space gives the **vectorized splitting of spacetime** by o and $\boldsymbol{u}$:

$$h_{\boldsymbol{u},o} : \mathtt{M} \to \mathbb{I} \times \mathbf{E}_{\boldsymbol{u}}, \qquad x \mapsto \boldsymbol{h}_{\boldsymbol{u}} \cdot (x - o) = \big(-\boldsymbol{u}\cdot(x-o), \boldsymbol{\sigma}_{\boldsymbol{u}}\cdot(x-o)\big). \tag{V46}$$

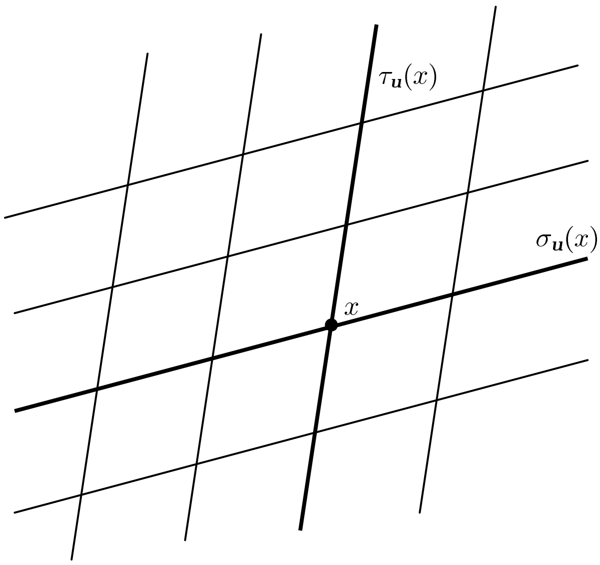

Figure 12.10 Splitting of spacetime

It has the inverse

$$\mathtt{I} \times \mathbf{E}_{\boldsymbol{u}} \to \mathtt{M}, \qquad (\boldsymbol{t}, \boldsymbol{q}) \mapsto o + \boldsymbol{t}\boldsymbol{u} + \boldsymbol{q}.$$

12.9.2 Transformation rules

The transformation rule between spacetime splittings would be $h_{\boldsymbol{u}'} \circ h_{\boldsymbol{u}}^{-1} : \mathtt{I}_{\boldsymbol{u}} \times E_{\boldsymbol{u}} \to \mathtt{I}_{\boldsymbol{u}} \times E_{\boldsymbol{u}'}$. We face the problem again that the domain and the range of this map are different. To overcome this problem, we consider vectorized splittings combined with boosts, getting a transformation rule $\mathbb{I} \times \mathbf{E}_{\boldsymbol{u}} \to \mathbb{I} \times \mathbf{E}_{\boldsymbol{u}}$.

Let $(\boldsymbol{t}, \boldsymbol{q})$ and $(\boldsymbol{t}', \boldsymbol{q}')$ the vectorized split form of the same world point, due to $(\boldsymbol{u}, o)$ and $(\boldsymbol{u}', o')$, respectively, the latter one boosted to $\boldsymbol{u}$.

$$(\boldsymbol{t}', \boldsymbol{q}') = (1, \boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'}) h_{\boldsymbol{u},o'}\big(h_{\boldsymbol{u},o}^{-1}(\boldsymbol{t}, \boldsymbol{q})\big).$$

Omitting a trivial steps coming from the previous formulae and putting $\boldsymbol{t}_o := -\boldsymbol{u}' \cdot (o - o')$ and $\boldsymbol{q}_o = \boldsymbol{\sigma}'_{\boldsymbol{u}} \cdot (o - o')$ we obtain

$$\boldsymbol{t}' = \kappa(\mathbf{v})(\boldsymbol{t} - \mathbf{v} \cdot \boldsymbol{q}) + \boldsymbol{t}_o, \qquad \boldsymbol{q}' = \kappa(\mathbf{v})(\boldsymbol{J}(\mathbf{v}) \cdot \boldsymbol{q} - \boldsymbol{t}\mathbf{v}) + \boldsymbol{q}_o,$$

which is the well known **inhomogeneous Lorentz transformation rule**.

This equals the vectorial Lorentz transformation rule if the observers choose the same spacetime origin ($o' = o$).

Note that

– the inhomogeneous Lorentz transformation rule is an affine map, it serves for comparing the splitting of spacetime,

– the Lorentz transformation rule is a linear map, it serves for comparing the splitting of spacetime vectors.

12.10 Transformations and transformation rules

The Lorentz transformations are linear map $\mathbf{M} \to \mathbf{M}$, in other words,, are elements of $\mathbf{M} \otimes \mathbf{M}^*$. The formula of tensorial splitting (see 12.8.1) shows that the $\boldsymbol{u}$-split form of a Lorentz transformation is more complicated than the that of a Galilei transformation (see (IV18)) which is connected with the fact that here there is no three dimensional linear subspace invariant for all Lorentz transformations.

We derive from (V44) that

$$\boldsymbol{h}_{\boldsymbol{u}\boldsymbol{u}'} = \boldsymbol{h}_{\boldsymbol{u}} \cdot \boldsymbol{B}_{\boldsymbol{u}'\boldsymbol{u}} \cdot \boldsymbol{h}_{\boldsymbol{u}}^{-1}$$

i.e. the vectorial transformation rule from $\boldsymbol{u}'$ to $\boldsymbol{u}$ is the $\boldsymbol{u}$-split form of the boost from $\boldsymbol{u}'$ to $\boldsymbol{u}$.

The vectorized splitting of spacetime converts the Poincaré transformation over boosts into the inhomogeneous Lorentz transformations.

Though there is some relation between special (Poincaré) Lorentz transformations and (inhomogeneous) Lorentz transformation rules, they differ essentially both from a mathematical and a physical point of view. The **transformations** are maps $\mathtt{M} \to \mathtt{M}$ and $\mathbf{M} \to \mathbf{M}$, respectively, that reflects the structure of spacetime (symmetries), whereas the **transformation rules** are maps $\mathbb{I} \times \mathbf{E} \to \mathbb{I} \times \mathbf{E}$ that compare different splittings.

In usual treatments based on coordinates the transformations (spacetime symmetries) and the transformation rules are confused because all of them are maps in $\mathbb{R} \times \mathbb{R}^3$. This often causes conceptual errors as we pointed out in the nonrealtivistic case (see 6.9).

We can find hints in the literature to that two different objects appear in the same form, when one distinguishes between active transformations (which correspond to spacetime symmetries) and passive transformations (which correspond to transformation rules).

12.11 Standard coordinatizations

Time intervals are usually characterized by numbers giving them as a multiple of a time unit $\mathrm{s} \in \mathbb{I}^+$ (second); in formula, time intervals are coordinatized by

$$\mathbb{I} \to \mathbb{R}, \qquad \boldsymbol{t} \mapsto \frac{\boldsymbol{t}}{\mathrm{s}}.$$

Now the lengths (distances), too, are measured by time intervals (the distance between two spacepoint of an inertial observer is the half of the time interval of light signal running to and fro).

The space vectors of the inertial observer $\boldsymbol{u}$ are usually characterized by triplets of numbers in such a way that a time unit s is chosen, further coordinate axes determined by three 'right handed' orthogonal vectors of length s. In formula, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ is a positively oriented orthogonal basis in $\mathbf{E}_{\boldsymbol{u}}$, every basis element having the length s, and we take the coordinates of vectors corresponding to that basis:

$$\mathbf{E}_{\boldsymbol{u}} \to \mathbb{R}^3, \qquad \boldsymbol{q} \mapsto \left(\frac{\boldsymbol{e}_1 \cdot \boldsymbol{q}}{\mathrm{s}^2}, \frac{\boldsymbol{e}_2 \cdot \boldsymbol{q}}{\mathrm{s}^2}, \frac{\boldsymbol{e}_3 \cdot \boldsymbol{q}}{\mathrm{s}^2} \right).$$

The standard inertial frame $\boldsymbol{u}$ coordinatizes the world vectors in such a way that

- it splits $\mathbf{M}$ into $\mathbb{I} \times \mathbf{E}_{\boldsymbol{u}}$ according to V42,
- it coordinatizes $\mathbb{I}$ by s,
- it coordinatizes $\mathbf{E}_{\boldsymbol{u}}$ by $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$.

In other words, with the notation $\boldsymbol{e}_0 := \mathrm{s}\boldsymbol{u}$, the coordinatization of world vectors is the linear bijection

$$\mathbf{M} \to \mathbb{R}^4, \qquad \boldsymbol{x} \mapsto \{\text{coordinates of } \boldsymbol{x} \text{ in the basis } \boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}.$$

Thus, if $(\xi^0, \xi^1, \xi^2, \xi^3)$ are the coordinates of the vector $\boldsymbol{x}$, then $\boldsymbol{x} = \sum_{i=0}^{3} \xi^i \boldsymbol{e}_i$ and it is easy to see that

$$\xi^0 = -\frac{\boldsymbol{e}_0 \cdot \boldsymbol{x}}{\mathrm{s}^2}, \quad \xi^i = \frac{\boldsymbol{e}_i \cdot (\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{x})}{\mathrm{s}^2}, \quad (i = 1, 2, 3).$$

Note that the indices of a vector are usually written as superscripts.

To coordinatize spacetime, the standard inertial frame vectorizes spacetime by a chosen spacetime origin o and then applies the previous procedure. The result is

$$\mathtt{M} \to \mathbb{R}^4, \qquad x \mapsto \{\text{coordinates of } x - o \text{ in the basis } \boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}.$$

The coordinatization can be described in a physically more transparent way as follows. The standard inertial frame

– splits M to $\boldsymbol{u}$-time and $\boldsymbol{u}$-space according to V45,

– vectorizes time and $\boldsymbol{u}$-space by choosing a 'time origin' t_o and a '$\boldsymbol{u}$-space origin'q_o,

which is equivalent to the vectorized splitting of spacetime with the aid of the 'world origin' $o := t_o \cap q_o$ (see V46), and then

– coordinatizes $\mathbb{I}$ and $\mathbf{E}_{\boldsymbol{u}}$ as previously given.

According to these formulae, a **standard inertial coordinate system** is $(o, \mathrm{s}, \boldsymbol{u}, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$, where o is a world point, s is a time unit, $\boldsymbol{u}$ is an inertial observer, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ is a positively oriented orthogonal basis, normed to s, in $\mathbf{E}_{\boldsymbol{u}}$.

Returning to Paragraph 12.3.2, we can state that the coordinatization of spacetime by an inertial coordinate system is just an isomorphism to the arithmetic spacetime model. We see very well, how many arbitrary objects are hidden in the arithmetic spacetime model: a spacetime origin, a time unit, an inertial observer, the standard synchronization of the inertial observer and a spacelike basis.

The coordinate system represents covectors and various tensors by coordinates, too. For instance, the coordinates of the covector $\boldsymbol{k}$ are

$$(\chi_i := \boldsymbol{k} \cdot \boldsymbol{e}_i \mid i = 0, 1, 2, 3).$$

The indices of a covector are usually written as subscripts, for distingushing them from the coordinates of vectors.

Now, becasue of the $\mathbf{M}^* \equiv \frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}}$ Lorentz identification, coordinates of covectors are obtained essentially in the same way as coordinates of vectors: the difference is a minus sign in the zeroth (timelike) coordinate. Thus, one can consider the vectorial coordinates $(x^i : i = 0, 1, 2, 3)$ of a vector $\mathbf{x}$ (which are the coordinates of $\mathbf{x}$ itself) and the covectorial coordinates $(x_i : i = 0, 1, 2, 3)$ of $\mathbf{x}$ (which are the coordinates of $\frac{\mathbf{x}}{\mathrm{s}^2}$ as covector); then $\xi_0 = -\xi^0$, $\xi_k = \xi^k$ $(k = 1, 2, 3)$.

The coordinates of a tensor $\boldsymbol{G} \in \mathbf{M} \otimes \mathbf{M}$ are $(i, k = 1, 2, 3)$:

$$G^{00} := \frac{\boldsymbol{e}_0 \cdot \boldsymbol{G} \cdot \boldsymbol{e}_0}{\mathrm{s}^2}, \quad G^{0k} = -\frac{\boldsymbol{e}_0 \cdot \boldsymbol{G} \cdot \boldsymbol{e}_k}{\mathrm{s}^2}, \quad G^{k0} = -\frac{\boldsymbol{e}_k \cdot \boldsymbol{G} \cdot \boldsymbol{e}_0}{\mathrm{s}^2},$$

$$G^{ik} = \frac{\boldsymbol{e}_i \cdot \boldsymbol{G} \cdot \boldsymbol{e}_k}{\mathrm{s}^2}.$$

The coordinates of a cotensor $\boldsymbol{F} \in \mathbf{M}^* \otimes \mathbf{M}^*$ are $(i, k = 0, 1, 2, 3)$:

$$F_{ik} = \boldsymbol{e}_i \cdot \boldsymbol{F} \cdot \boldsymbol{e}_k.$$

12.12 Comparison of lengths and time intervals

12.12.1 Prints

Though boosts are already defined by light signals, it is worth examining instantaneous prints that gave the boosts in the nonrelativistic case.

Let the standard inertial frame $\boldsymbol{u}$ make an instantaneous print of the space vectors of an inertial observer $\boldsymbol{u}'$. The print of the vector $\boldsymbol{q}' \in \mathbf{E}_{\boldsymbol{u}'}$ is the vector $\boldsymbol{q} \in \mathbf{E}_{\boldsymbol{u}}$ for which $\boldsymbol{q}' - \boldsymbol{q}$ is parallel to $\boldsymbol{u}'$ (see Figure 12.11); this means that $\boldsymbol{q}$ is the projection of $\boldsymbol{q}'$ onto $\mathbf{E}_{\boldsymbol{u}}$ along $\boldsymbol{u}'$. This projection is obtained by the linear map

$$\mathbf{1} + \frac{\boldsymbol{u}' \otimes \boldsymbol{u}}{-\boldsymbol{u}' \cdot \boldsymbol{u}} \tag{V47}$$

because it maps $\boldsymbol{u}'$ to zero and the elements of $\mathbf{E}_{\boldsymbol{u}}$ to themselves; thus,

$$\boldsymbol{q} = \left(\mathbf{1} + \frac{\boldsymbol{u}' \otimes \boldsymbol{u}}{-\boldsymbol{u}' \cdot \boldsymbol{u}}\right)\boldsymbol{q}'. \tag{V48}$$

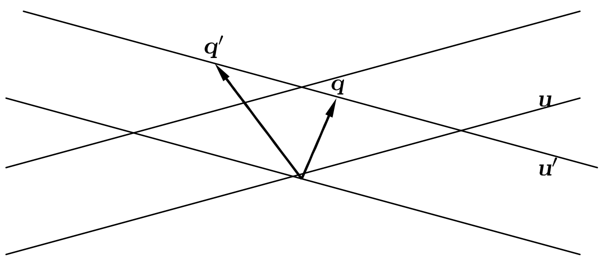

Figure 12.11 Instantaneous print

$\boldsymbol{q}$ and $\boldsymbol{q}'$ have different lengths, so the instantaneous print is not suitable for establishing a boost.

12.12.2 Lorentz contraction

The equality $\boldsymbol{u}' \cdot \boldsymbol{q}' = 0$ and (V48) result in

$$\boldsymbol{q} = \boldsymbol{q}' + \boldsymbol{u}'\frac{\boldsymbol{u} \cdot \boldsymbol{q}'}{-\boldsymbol{u}' \cdot \boldsymbol{u}} = \boldsymbol{q}' + \boldsymbol{u}'\left(\frac{\boldsymbol{u}}{-\boldsymbol{u}' \cdot \boldsymbol{u}} - \boldsymbol{u}'\right) \cdot \boldsymbol{q}' =$$
$$= \boldsymbol{q}' + \boldsymbol{u}'(\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'} \cdot \boldsymbol{q}'),$$

from which it follows that

$$|\boldsymbol{q}|^2 = |\boldsymbol{q}'|^2 - (\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'} \cdot \boldsymbol{q}')^2.$$

This is the famous **Lorentz contraction**: if the vector is perpendicular to the relative velocity then the length of the vector and that of its print are equal. Otherwise the print is shorter; it is the shortest if the vector is parallel to the relative velocity: then its length is $\sqrt{1 - |\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}|^2}$ times the length of the vector.

The Lorentz contraction is often declared so that a moving rod is contracted. This is incorrect. We emphasize: **the Lorentz contraction is not a physical fact** but an illusion connected with a synchronizationwhich is not a physical reality but a human convention. Using a synchronizationdifferent from the standard one, we get a different result (see 17.5).

We call attention to that **a physical fact must not be explained by the Lorentz contraction**. The physical relations which are demonstrated by the Lorentz contraction in the usual literature can be accepted only if a demonstration without synchronizationcan be found.

12.12.3 The tunnel paradox

The famous tunnel paradox originated from the Lorentz contraction.

Namely, the roles of $\boldsymbol{u}$ and $\boldsymbol{u}'$ in the previous formula can be interchanged: then the vectors of $\boldsymbol{u}$ will be contracted.

Let us consider a train moving on a straight rail with uniform velocity. The train approaches a tunnel. The proper length of the train equals the proper length of the tunnel.

The train moves with respect to the tunnel, so the tunnel 'perceives' that the train is shorter, so the tunnel will have inside the entire train during a time period.

The tunnel moves with respect to the train, so the train 'perceives' that the tunnel is shorter, so the train never will be entirely in the tunnel.

We can remove the paradox by an exact formulation. Let us consider the occurrence that the front of the train meets the exit of the tunnel. According to the standard synchronization of the tunnel, at the corresponding instant, the end of the train has already passed the entrance. According to the standard synchronization of the train, at the corresponding instant, the end of the train has not yet passed the the entrance (see Figure 12.12).

Thus, both the tunnel and the train have right but, of course, concerning two different synchronizations.

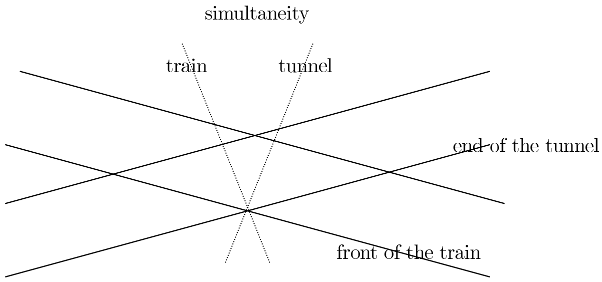

Figure 12.12 Train and tunnel

Someone who is not convinced by the argument above says: let the tunnel block the train when it is inside, shutting the exit with a strong gate; then the tunnel will have right not the train.

Let us examine this situation. Of course, the front of the train will collide the exit gate which forces the train to stop. How does the train stop? Let us suppose that all the wagons stop simultaneously according to the synchronization of the tunnel. Then the train, according to its own synchronization, 'feels' that when the engine stops then the first wagon is still in motion, when the first wagon stops the the second one is still in motion, etc: the train will squeeze (see Figure 12.13.

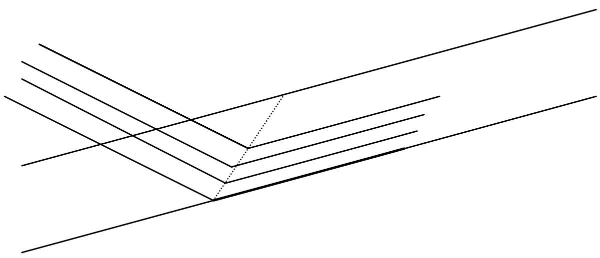

Figure 12.13 Stopped train

Indeed, the braked train is inside in the tunnel. This procedure, however, terminates the reciprocity between the tunnel and the train because the train is not inertial during the braking which causes a real change in the length of the train.

12.12.4 Time dilation

An inertial observer wants to measure how a moving chronometer works i.e. the frequency of its ticks.

The chronometer moves with respect to the observer, so different ticks occur in different space points of the observer. As a consequence, the observer cannot compare the frequency of the moving chronometer to the frequency of one of its own chronometer (resting in some space point of the observer). To fulfil the task, at least two chronometer of the observer is necessary, and a synchronizationof the chronometers.

Let us consider a standard inertial frame $\boldsymbol{u}$ measuring the frequency of ticks of a chronometer with absolute velocity $\boldsymbol{u}'$.

The synchronization time period $\boldsymbol{t}$ corresponding to the proper time period $\boldsymbol{t}'$ of the chronometer moving with respect to the observer is the $\boldsymbol{u}$-timelike component of the vector $\boldsymbol{t}'\boldsymbol{u}'$ (see Figure 12.14:

$$\boldsymbol{t} := \boldsymbol{t}'(-\boldsymbol{u}\cdot\boldsymbol{u}') = \frac{\boldsymbol{t}'}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}|^2}}.$$

This is the famous **time dilation**: the measured time period is longer than the proper time period.

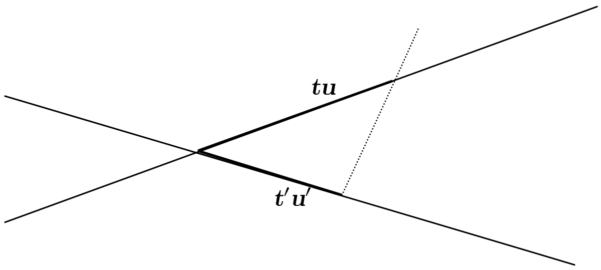

Figure 12.14 Time dilation

The time dilation is often declared so that a moving clock (chronometer) works more slowly than a resting one. This is incorrect. We emphasize: **the time dilation is not a physical fact** but an illusion connected with a synchronizationwhich is not a physical reality but a human convention. Using a synchronizationdifferent from the standard one, we get a different result (see 17.6).

We call attention to that **a physical fact must not be explained by the time dilation**. The physical relations which are demonstrated by the time dilation in the usual literature can be accepted only if a demonstration without synchronizationcan be found.

12.12.5 The twin paradox

The famous twin paradox originated from the time dilation.

Namely, the roles of $\boldsymbol{u}$ and $\boldsymbol{u}'$ in the previous formula can be interchanged: then the chronometers of $\boldsymbol{u}$ will work more slowly.

Let us consider two twins, Alice and Bob, who are separated when being born, continue to live in two different inertial spaceships.

According to Alice, Bob is moving, so she finds that when she is twenty five years old then Bob is only twenty years old: Bob is younger than Alice.

According to Bob, Alice is moving, so he finds that when he is twenty five years old then Alice is only twenty years old: Alice is younger than Bob.

It is easy to remove the paradox: the "when ... then ..." concerns two different synchronizations in the two cases.

Someone who is not convinced by the argument above says: let us see which of the twins will be younger if they meet.

But they never meet if both remain inertial.

If they meet in such a way that Alice (Bob) goes to Bob (Alice) i.e. Bob (Alice) remains inertial then Alice (Bob) will be younger because we know that the inertial time period between two occurrences is the longest. Of course, it may happen that neither of them remains inertial and just in such a way that they will be equally old at their meeting.

Someone can ask: how can Alice (Bob) have information about the age of Bob (Alice), since they are far from each other. The method is the following. Alice (Bob), being ten years old, sends a radio message to Bob (Alice): how old are you are? Bob (Alice) answers immediately: I am twenty. The answer returns to Alice (Bob) when she (he) is forty years old. Then she (he) calculates that thirty years passed between the start of her (his) message and the arrival of his (her) message, so it is 'evident' that both messages travelled fifteen years, thus, when I was twenty five when he (she) was twenty.

This calculation of simultaneity corresponds to the standard synchronization.

12.13 Derivatives

Let us consider a differentiable function $f : \mathtt{M} \to \mathbb{R}$. It derivative at a point x is a linear map $\mathrm{D}f(x) : \mathbf{M} \to \mathbb{R}$ i.e. it is an element of $\mathbf{M}^*$ (see the mathematical supplement). This covector is split by the inertial observer into the $\boldsymbol{u}$-timelike component

$$(\mathrm{D}f(x)) \cdot \boldsymbol{u} =: \mathrm{D}_{\boldsymbol{u}} f(x)$$

and into the $\boldsymbol{u}$-spacelike component

$$(\mathrm{D}f(x))|_{\mathbf{E}_{\boldsymbol{u}}} =: \nabla_{\boldsymbol{u}} f(x)$$

which have a direct meaning as follows.

Let us restrict f onto the straight line passing through x and directed by $\boldsymbol{u}$, i.e. let us consider the function $\mathbb{I} \to \mathbb{R}$, $\boldsymbol{t} \mapsto f(x + \boldsymbol{t}\boldsymbol{u})$. The derivative at zero of this function – according to the rule of differentiation of composite functions – is $\mathrm{D}f(x) \cdot \boldsymbol{u}$.

Let us restrict f onto the hyperplane passing through x and directed by $\boldsymbol{u}$, i.e. let us consider the function $\mathbf{E}_{\boldsymbol{u}} \to \mathbb{R}$, $\boldsymbol{q} \mapsto f(x + \boldsymbol{q})$. The derivative at zero of this function – according to the rule of differentiation of composite functions – is $(\mathrm{D}f(x))|_{\mathbf{E}_{\boldsymbol{u}}}$.

That is why $\mathrm{D}_{\boldsymbol{u}} f$ and $\nabla_{\boldsymbol{u}} f$ are called the **$\boldsymbol{u}$-timelike derivative** and the **$\boldsymbol{u}$-spacelike derivative** of f, respectively.

If spacetime is coordinatized in the usual way (see Subsection 12.11), the function f is given in the form $\mathbb{R}^4 \to \mathbb{R}$, $(\xi^0, \xi^1, \xi^2, \xi^3) \mapsto \hat{f}(\xi^0, \xi^1, \xi^2, \xi^3) := f(o + \sum_{k=0}^{3} \xi^k \boldsymbol{e}_k)$. Then $\partial_k \hat{f}(\xi^0, \xi^1, \xi^2, \xi^3) = \mathrm{D}f(o + \sum_{k=0}^{3} \xi^k \boldsymbol{e}_k) \cdot \boldsymbol{e}_k$, i.e. the partial derivatives are the coordinates of $\mathrm{D}f$. For the sake of simplicity, admitting a little abuse of notation, let us omit the 'hat'; the we can write

$$\mathrm{D}f \quad \text{in coordinates is} \quad \partial_k f \quad (k = 0, 1, 2, 3).$$

It is evident that the zeroth partial derivative is the coordinatized form of the $\boldsymbol{u}$-timelike derivative, the other three partial derivatives constitute the coordinatized form of the $\boldsymbol{u}$-spacelike derivative.

In general, the differentiation D can be conceived as a symbolic covector whose $\boldsymbol{u}$-split form is $(\mathrm{D}_{\boldsymbol{u}}, \nabla_{\boldsymbol{u}}) := (\boldsymbol{u} \cdot \mathrm{D}, \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathrm{D})$.

For instance, the values of the derivative $\mathrm{D}\mathcal{J}$ of a vector field $\mathcal{J} : \mathtt{M} \to \mathbf{M}$ are elements of $\mathbf{M} \otimes \mathbf{M}^*$. As said in the mathematical supplement, it is suitable to use the transpose of the derivative which will be denoted by $\mathrm{D} \otimes \mathcal{J}$; it has values in $\mathbf{M}^* \otimes \mathbf{M}$. Its $\boldsymbol{u}$-split form is

$$\begin{pmatrix} \mathrm{D}_{\boldsymbol{u}}(-\boldsymbol{u} \cdot \mathcal{J}) & \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathcal{J}) \\ \nabla_{\boldsymbol{u}}(-\boldsymbol{u} \cdot \mathcal{J}) & \nabla_{\boldsymbol{u}} \otimes (\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathcal{J}) \end{pmatrix},$$

in coordinates $\partial_i J^k$ $(i, k = 0, 1, 2, 3)$.

Taking the trace of $(\mathrm{D} \otimes \mathcal{J})(x)$ (see the mathematical supplement), we define the **divergence** of $\mathcal{J}$:

$$(\mathrm{D} \cdot \mathcal{J})(x) := \mathrm{Tr}(\mathrm{D} \otimes \mathcal{J}(x)),$$

in a split form

$$\mathrm{D} \cdot \mathcal{J} = \mathrm{D}_{\boldsymbol{u}}(-\boldsymbol{u} \cdot \mathcal{J}) + \nabla_{\boldsymbol{u}} \cdot (\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathcal{J}),$$

which reads in coordinates as $\sum_{k=0}^{3} \partial_k J^k$.

For a better survey, let us introduce the symbol $\sim$ which will refer to what a split form and a coordinatized form has the derivative in question. Cooordinates always run through the values $0, 1, 2, 3$ and a summation is has to be effected for equal subscripts and superscripts (Einstein summation rule). Thus, if

$$\mathcal{J} \quad \sim \quad (\rho_{\boldsymbol{u}}, \boldsymbol{j}_{\boldsymbol{u}}) \quad \sim \quad J^k,$$

then

$$\mathrm{D} \cdot \mathcal{J} \quad \sim \quad \mathrm{D}_{\boldsymbol{u}} \rho_{\boldsymbol{u}} + \nabla_{\boldsymbol{u}} \cdot \boldsymbol{j}_{\boldsymbol{u}} \quad \sim \quad \partial_k J^k.$$

The (transpose of) the derivative $\mathrm{D} \otimes \boldsymbol{K}$ of a covector field $\boldsymbol{K} : \mathtt{M} \to \mathbf{M}^*$ has values in $\mathbf{M}^* \otimes \mathbf{M}^*$; its split form is

$$\begin{pmatrix} \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{u} \cdot \boldsymbol{K}) & \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{i}_{\boldsymbol{u}}^* \cdot \boldsymbol{K}) \\ \nabla_{\boldsymbol{u}}(\boldsymbol{u} \cdot \boldsymbol{K}) & \nabla_{\boldsymbol{u}} \otimes (\boldsymbol{i}_{\boldsymbol{u}}^* \cdot \boldsymbol{K}) \end{pmatrix},$$

in coordinates $\partial_i K_k$ $(i, k = 0, 1, 2, 3)$.

Taking the antisymmetric part of $(\mathrm{D} \otimes \boldsymbol{K})(x)$, we define the **exterior derivative** of $\boldsymbol{K}$:

$$\mathrm{D} \wedge \boldsymbol{K} := \mathrm{D} \otimes \boldsymbol{K} - (\mathrm{D} \otimes \boldsymbol{K})^*,$$

in a split form

$$\big((\nabla_{\boldsymbol{u}}(\boldsymbol{u} \cdot \boldsymbol{K}) - \mathrm{D}_{\boldsymbol{u}}(\boldsymbol{i}_{\boldsymbol{u}}^* \cdot \boldsymbol{K}), \nabla_{\boldsymbol{u}} \wedge (\boldsymbol{i}_{\boldsymbol{u}}^* \cdot \boldsymbol{K})) \big),$$

which reads in coordinates as $\partial_k K_i - \partial_i K_k$.

With the previous survey: if

$$\boldsymbol{K} \quad \sim \quad (-V_{\boldsymbol{u}}, \boldsymbol{A}_{\boldsymbol{u}}) \quad \sim \quad K_k,$$

then

$$\mathrm{D} \wedge \boldsymbol{K} \quad \sim \quad ((-\nabla_{\boldsymbol{u}} V_{\boldsymbol{u}} - \mathrm{D}_{\boldsymbol{u}} \cdot \boldsymbol{A}_{\boldsymbol{u}}, \nabla_{\boldsymbol{u}} \wedge \boldsymbol{A}_{\boldsymbol{u}})) \quad \sim \quad \partial_i K_k - \partial_k K_i.$$

Note that, originally,
– a vector field has divergence and has not exterior derivative,
– a covector field has exterior derivative and has not divergence.

The vector fields and covector fields, however, can be identified via the identification $\mathbf{M}^* \equiv \frac{\mathbf{M}}{\mathbb{I}\otimes}$, so **in applications we can take the exterior derivative of a vector field and the divergence of a covector field.**

Similarly, an antisymmetric tensor field $\boldsymbol{G} : \mathtt{M} \to \mathbf{M} \wedge \mathbf{M}$ has **divergence** $\mathrm{D} \cdot \boldsymbol{G}$ which takes values in $\mathbf{M}$; if

$$\boldsymbol{G} \quad \sim \quad ((\boldsymbol{D}_{\boldsymbol{u}}, \boldsymbol{H}_{\boldsymbol{u}})) \quad \sim \quad G^{ik},$$

then

$$\mathrm{D} \cdot \boldsymbol{G} \quad \sim \quad (\nabla_{\boldsymbol{u}} \cdot \boldsymbol{D}_{\boldsymbol{u}}, -\mathrm{D}_{\boldsymbol{u}} \boldsymbol{D}_{\boldsymbol{u}} + \nabla_{\boldsymbol{u}} \cdot \boldsymbol{H}_{\boldsymbol{u}}) \quad \sim \quad \partial_i G^{ik}. \qquad \text{(V49)}$$

An antisymmetric cotensor field $\boldsymbol{F} : \mathtt{M} \to \mathbf{M}^* \wedge \mathbf{M}^*$ has the **exterior derivative** $\mathrm{D} \wedge \boldsymbol{F}$ which takes values in $\mathbf{M}^* \wedge \mathbf{M}^* \wedge \mathbf{M}^*$; if

$$\boldsymbol{F} \quad \sim \quad ((\boldsymbol{E}_{\boldsymbol{u}}, \boldsymbol{B}_{\boldsymbol{u}})) \quad \sim \quad F_{ik},$$

then

$$\mathrm{D} \wedge \boldsymbol{F} \quad \sim \quad (((\nabla_{\boldsymbol{u}} \wedge \boldsymbol{E}_{\boldsymbol{u}} + \mathrm{D}_{\boldsymbol{u}} \boldsymbol{B}_{\boldsymbol{u}}, \nabla_{\boldsymbol{u}} \wedge \boldsymbol{B}_{\boldsymbol{u}}))) \quad \sim \quad \partial_j F_{ik} + \partial_k F_{ji} + \partial_i F_{kj} \qquad \text{(V50)}$$

where the triple of brackets means that the two objects between them determine the whole antisymmetric tensor.

As a consequence of the identification of tensores and cotensors, we can take the exterior derivative of a tensor field and the divergence of a cotensor field.

13 Fundamentals of point mechanics in the spacetime model

In this section we treat in the special relativistic spacetime model the notions and relations of the simplest physical theory. Classical mechanics is a well elaborated theory in the framework of coordinates, thus it offers us a good possibility to deepen our knowledge on spacetime; moreover, it shows the essential differences between the nonrelativistic and relativistic theory.

13.1 World line functions

In the nonrelativistic spacetime model a world lime can be parameterized by absolute time in a natural way. Here such a possibility does not exist, nevertheless we can give natural parameterizations by the proper time of the world line.

Let us fix a point x_0 of a world line C; assigning to every point x of C the proper time $\boldsymbol{t}_{\mathrm{C}}(x_0, x)$ passed from from x_0 to x (see 12.1.3), we get a proper time function of C.

The function $\mathrm{C} \to \mathbb{I}$, $x \mapsto \boldsymbol{t}_{\mathrm{C}}(x_0, x)$ *is injective, its inverse, denoted by* r, *is a continuously differentiable function from* $\mathbb{I}$ *into* $\mathtt{M}$, *and for every progressive parametrization* p *of* C

$$\dot{r}(r^{-1}(x)) = \frac{\dot{p}(p^{-1}(x))}{|\dot{p}(p^{-1}(x))|} \tag{V51}$$

where $|\ |$ denotes the pseudolength.

Let p be a progressive parametrization $p(0) = x_0$. Then the function

$$\mathbb{R} \to \mathbb{I}, \qquad a \mapsto X(a) := \int_0^a |\dot{p}(b)| db$$

gives the proper time function in the form $X \circ p^{-1}$. According to a well known result of integral calculus, X is continuously differentiable and $\dot{X}(a) = |\dot{p}(a)| > 0$ for all a. This means that X is strictly monotone increasing (thus, it is injective), its inverse, too, is continuously differentiable and $(X^{-1})^{\cdot} = \frac{1}{\dot{X} \circ X^{-1}}$. As a consequence, the proper time function is injective and its inverse $r := p \circ X^{-1}$ is injective and continuously differentiable as well. All these imply that equality (V51) holds.

According to the original definition, a parametrization of a curve is established by real numbers but, of course, the elements of a measure line, too, can be used for a parameterization. In this sense we can say that r is a progressive parametrization of C which is called the **proper time parametrization**. $r(\boldsymbol{s})$ is the point of the world line till which $\boldsymbol{t}$ proper time passed from the point x_0.

(V51) implies $|\dot{r}(\boldsymbol{s})| = 1$ i.e. the values of $\dot{r}$ are absolute velocities.

A **world line function** is a sufficiently many time differentiable function $r : \mathbb{I} \to \mathtt{M}$ whose derivative has values in $\mathtt{V}(1)$.

Evidently, the range of a world line function is world line and every world line can be given as the range of (a lot of) world line functions. The world line functions whose range is the same differ only only in a translation of their domain: if $\mathrm{Ran} r_1 = \mathrm{Ran} r_2$ then there is $\boldsymbol{s}_0 \in \mathbb{I}$ such that $r_2(\boldsymbol{s}) = r_1(\boldsymbol{s} + \boldsymbol{s}_0)$.

Since world line function r has the property $\dot{r}(\boldsymbol{s}) \cdot \dot{r}(\boldsymbol{s}) = -1$ we get immediately that

$$\dot{r}(\boldsymbol{s}) \cdot \ddot{r}(\boldsymbol{s}) = 0$$

holds for the absolute acceleration $\ddot{r}(\boldsymbol{s})$ i.e.

$$\ddot{r}(\boldsymbol{s}) \in \frac{\mathbf{E}_{\dot{r}(\boldsymbol{s})}}{\mathbb{I} \otimes \mathbb{I}};$$

The absolute acceleration has spacelike values, an instantaneous value is Lorentz orthogonal to the corresponding absolute velocity; thus the collection **absolute accelerations belonging to the absolute velocity** $\boldsymbol{u}$ is the three-dimensional Euclidean space

$$\frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I} \otimes \mathbb{I}}$$

In contrast to absolute velocities,
– three is a zero absolute acceleration,
– an absolute acceleration has magnitude,
– the angle between two absolute accelerations belonging to the same absolute velocity makes sense (but need not make sense for absolute accelerations belonging to different absolute velocity).

13.2 Motions

An observer, in general, can describe motions only by choosing a synchronization because it must be given *when* the body in question is *where*. Motion is meaningful only with respect to a reference frame. In the relativistic spacetime model more synchronizations can be given. In the sequel we consider standard inertial frames.

13.2.1 Relative velocities

A standard inertial frame $\boldsymbol{u}$ perceives the history of a material point (a world line) as a motion and describes it by assigning to a $\boldsymbol{u}$-instant t (a hyperplane directed by $\mathbf{E}_{\boldsymbol{u}}$) the $\boldsymbol{u}$-space point (straight line directed by $\boldsymbol{u}$) which meets the material point (the world line) at t. Therefore, the $\boldsymbol{u}$-motion corresponding to the world line function r is

$$\mathtt{I}_{\boldsymbol{u}} \to \mathtt{E}_{\boldsymbol{u}}, \quad t \mapsto r_{\boldsymbol{u}}(t) := \sigma_{\boldsymbol{u}}(t \cap \mathrm{Ran} r).$$

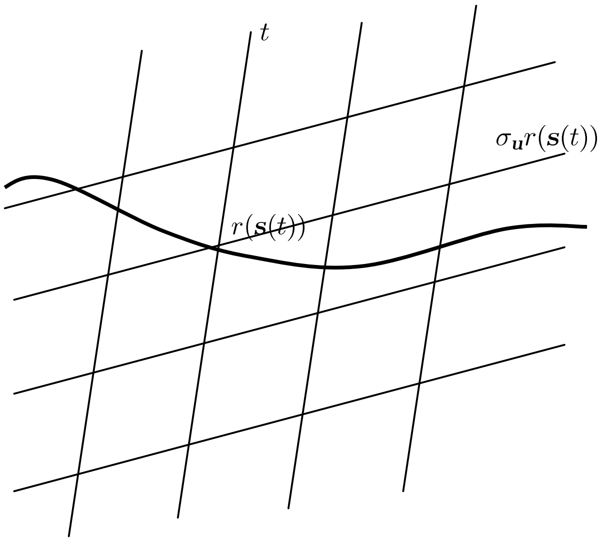

Figure 13.1 Description of motion

This function can be well handled if we give the relation between the proper time of the world line (the variable of r) and the synchronization time (the variable of $r_{\boldsymbol{u}}$).

This relation is described by the function which assigns to the proper time value $\boldsymbol{s}$ the $\boldsymbol{u}$-time instant (hyperplane directed by $\mathbf{E}_{\boldsymbol{u}}$) containing $r(\boldsymbol{s})$:

$$t(\boldsymbol{s}) := r(\boldsymbol{s}) + \mathbf{E}_{\boldsymbol{u}} = \tau_{\boldsymbol{u}}(r(\boldsymbol{s})).$$

Subtraction (V40) yields

$$\frac{\mathrm{d}t(\boldsymbol{s})}{\mathrm{d}\boldsymbol{s}} = \boldsymbol{\tau}_{\boldsymbol{u}} \cdot \dot{r}(\boldsymbol{s}) = -\boldsymbol{u} \cdot \dot{r}(\boldsymbol{s}) \geq 1;$$

therefore, the function is strictly monotone increasing (injective); for its inverse, denoted by $\boldsymbol{s}(t)$,

$$\frac{\mathrm{d}\boldsymbol{s}(t)}{\mathrm{d}t} = \frac{1}{-\boldsymbol{u} \cdot \dot{r}(\boldsymbol{s}(t))} \qquad \text{(V52)}$$

holds. As a consequence,

$$t \cap \mathrm{Ran} r = r(\boldsymbol{s}(t)).$$

The velocity of the material point relative to standard inertial frame is the time derivative of the $\boldsymbol{u}$-motion:

$$\frac{\mathrm{d}r_{\boldsymbol{u}}(t)}{\mathrm{d}t} = \frac{\mathrm{d}\sigma_{\boldsymbol{u}}(r(\boldsymbol{s}(t))}{\mathrm{d}t} = \frac{\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \dot{r}(\boldsymbol{s}(t))}{-\boldsymbol{u} \cdot \dot{r}(\boldsymbol{s}(t))} = \left(\frac{\dot{r}}{-\boldsymbol{u} \cdot \dot{r}} - \boldsymbol{u}\right)(\boldsymbol{s}(t)).$$

This is the generalization of our result concerning inertial world lines: the **relative velocity** of the absolute velocity $\boldsymbol{u}'$ with respect $\boldsymbol{u}$ equals

$$\mathbf{v}_{\mathbf{u}'\boldsymbol{u}} := \boldsymbol{u}' - \boldsymbol{u} \in \frac{\mathbf{E}}{\mathbb{I}}.$$

13.2.2 Relative accelerations

In the sequel we often met functions depending on the frame time via the proper time. For the sake of avoiding involved formulae, we refer to that by the symbol $\bullet$. If f is a function of proper time then $f\bullet$ depends of frame time: $(f\bullet)(t) = f(\boldsymbol{s}(t))$. The derivation with respect to proper time is denoted by a dot, the derivation with respect to the frame time is denoted by an inverted comma. Thus, on the base of equality (V52) we have

$$(f\bullet)' = \left(\frac{\dot{f}}{-\boldsymbol{u} \cdot \dot{r}}\right)\bullet .$$

Using this notation, the $\boldsymbol{u}$-relative velocity corresponding to the world line function r is

$$r'_{\boldsymbol{u}} = \left(\frac{\dot{r}}{-\boldsymbol{u} \cdot \dot{r}} - \boldsymbol{u}\right)\bullet = \mathbf{v}_{\dot{r}\boldsymbol{u}}\bullet .$$

Further, we recall equality (V35) in a form, convenient for the present case

$$(-\boldsymbol{u} \cdot \dot{r}) = \frac{1}{\sqrt{1 - |\mathbf{v}_{\dot{r}\boldsymbol{u}}|^2}}. \tag{V53}$$

Accordingly, the $\boldsymbol{u}$-relative acceleration is

$$r''_{\boldsymbol{u}} = \left(\left(\ddot{r} + \frac{\dot{r}(\boldsymbol{u} \cdot \ddot{r})}{-\boldsymbol{u} \cdot \dot{r}}\right)\frac{1}{(\boldsymbol{u} \cdot \dot{r})^2}\right)\bullet .$$

Let us add $(\boldsymbol{u}\cdot\ddot{r})\boldsymbol{u}$ to $\ddot{r}$ in the inner parenthesis and let us subtract it from the second member; because of $\dot{r}\cdot\ddot{r}=0$ we get $\boldsymbol{u}\cdot\ddot{r}=-\mathbf{v}_{\dot{r}\boldsymbol{u}}\cdot\ddot{r}$, therefore

$$r''_{\boldsymbol{u}} = \big((1-\mathbf{v}_{\dot{r}\boldsymbol{u}}^2)\,(\mathbf{1}-\mathbf{v}_{\dot{r}\boldsymbol{u}}\otimes\mathbf{v}_{\dot{r}\boldsymbol{u}})\,\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\ddot{r}\big)\bullet$$

It is not hard to see that this gives

$$(\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\ddot{r})\bullet = \frac{1}{1-|r'_{\boldsymbol{u}}|^2}\left(\mathbf{1}+\frac{r'_{\boldsymbol{u}}\otimes r'_{\boldsymbol{u}}}{1-|r'_{\boldsymbol{u}}|^2}\right)r''_{\boldsymbol{u}}.$$

We see that, contrary to the nonrelativistic case, the relative acceleration is far from being equal the absolute acceleration. This not so surprising since the $\boldsymbol{u}$-relative acceleration has all its values in $\frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}\otimes\mathbb{I}}$ while the absolute acceleration at the proper time instant $\boldsymbol{s}$is in $\frac{\mathbf{E}_{\dot{r}(s)}}{\mathbb{I}\otimes\mathbb{I}}$.

13.3 Absolute Newtonian equation

13.3.1 Measure line of mass

As in the nonrelativistic case, we can choose $\frac{\mathbb{I}}{\mathbb{D}\otimes\mathbb{D}}$ for the measure line of mass; now, with the identification $\mathbb{D}=\mathbb{I}$, $2,99\ldots 10^8\text{m} := \text{s}$ the measure line of mass becomes $\frac{\mathbb{R}}{\mathbb{I}}=\mathbb{I}^*$; then

$$\text{kg} = 8,47\ldots 10^{50}\frac{1}{\text{s}}.$$

This choice is unusual in practical applications but it makes easier the theoretical exposition.

13.3.2 Absolute forces

We conceive that the absolute Newtonian equation has the form ‚mass×absolute acceleration = absolute force' where the absolute force can depend on world points and absolute velocities.

Since the values of ‚mass×absolute acceleration' are in the vector space $\frac{\mathbb{R}}{\mathbb{I}}\otimes\frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}}=\frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}\otimes\mathbb{I}}$, an absolute force is described by a function

$$\boldsymbol{f}:\mathtt{M}\times\mathtt{V}(1)\to\frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}\otimes\mathbb{I}}\equiv\frac{\mathbf{M}^*}{\mathbb{I}}.$$

Thus, the possible world line functions of a material point with mass m under the action of the absolute force $\boldsymbol{f}$ is determined by the **absolute Newtonian equation**

$$(x:\mathtt{I}\to\mathtt{M})? \qquad m\ddot{x}=\boldsymbol{f}(x,\dot{x}) \tag{V54}$$

which is a second order differential equation.

To have a unique solution of that differential equation, the initial spacetime position and absolute velocity of the mass point must be given. That is why, if $\boldsymbol{s} \mapsto r(\boldsymbol{s})$ is a solution, it is suitable to consider the pair $(r, \dot{r})$ the **process** of the mass point: its value at an arbitrary instant determines the whole function.

The **evolution space** of a mass point is the set in which the processes take values: $\mathtt{M} \times \mathtt{V}(1)$.

In what follows, we accept the notation (widely used in physics) that

– the elements of the evolution space are written in the form $(x, \dot{x})$ (note that then $\dot{x}$ is independent of x, it can denote an arbitrary absolute velocity),

– an arbitrary ('abstract') process – i.e. a time funcion – is denoted by $(x, \dot{x})$ as well,

– an actual process is denoted by $(r, \dot{r})$.

Call attention to the fact that a nonrelativistic force maps in the three-dimensional vector space $\frac{\mathbf{E}}{\mathbb{D}\otimes\mathbb{D}\otimes\mathbb{I}} \equiv \frac{\mathbf{E}^*}{\mathbb{I}}$, its values are absolute spacelike. A relativistic force maps in the four-dimensional vector space $\frac{\mathbf{M}}{\mathbb{I}\otimes\mathbb{I}\otimes\mathbb{I}} \equiv \frac{\mathbf{M}^*}{\mathbb{I}}$, its values, however, are 'three-dimensional' and absolute spacelike. Namely, the absolute acceleration is Lorentz-orthogonal to the absolute velocity, so the force must satisfy the equality

$$\boldsymbol{f}(x, \dot{x}) \cdot \dot{x} = 0$$

i.e. its values at the absolute velocity $\dot{x}$ are $\dot{x}$-spacelike, that is in $\frac{\mathbf{E}_{\dot{x}}}{\mathbb{I}\otimes\mathbb{I}\otimes\mathbb{I}}$.

As it is known, the forces having a **potential** play a peculiar role. A potential is twice differentiable function

$$\boldsymbol{K} : \mathtt{M} \to \mathbf{M}^*.$$

The **field strength** corresponding to the potential is the exterior derivative of the potential:

$$\boldsymbol{F} := \mathrm{D} \wedge \boldsymbol{K},$$

and the corresponding force equals

$$\boldsymbol{f}(x, \dot{x}) := \boldsymbol{i}^* \cdot \boldsymbol{F}(x) \cdot \dot{x}.$$

Such a definition of potential and field strength will be explained in Paragraph 7.5.2.

13.4 Momenta

The reader is asked for recalling subsection 7.4 in order to better understand the followings.

Let us consider a material point with mass m having the absolute velocity $\dot{x}$.

We accept the definition

'absolute momentum = mass×absolute velocity' $(m\dot{x})$

Since

'proper time derivative of absolute momentum = mass×absolute acceleration',

$((m\dot{x})^{\cdot} = m\ddot{x})$,

the absolute Newtonian equation can be conceived in two ways:

'mass×absolute acceleration = absolute force',

'proper time derivative of absolute momentum = absolute force'.

Let us consider a standard inertial frame $\boldsymbol{u}$. We find (omitting the notation • for the sake of simplicity and without the danger of confusion)

'mass× $\boldsymbol{u}$-relative velocity' $(m\mathbf{v}_{\dot{x}\boldsymbol{u}})$

and

'$\boldsymbol{u}$-spacelike component of absolute momentum' $\left(\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (m\dot{x}) = \frac{m\mathbf{v}_{\dot{x}\boldsymbol{u}}}{\sqrt{1-|\mathbf{v}_{\dot{x}\boldsymbol{u}}|^2}}\right)$,

are not equal, contrary to the nonrelativistic case. It is a question, which of them is accepted for being the $\boldsymbol{u}$-relative momentum. Theory does not answer the question. Experimental facts suggest the decision:

'$\boldsymbol{u}$-relative momentum := $\boldsymbol{u}$-spacelike component of absolute momentum'.

Further,

'mass× $\boldsymbol{u}$-relative acceleration' $\left(m\left(\left(\ddot{x} + \frac{\dot{x}(\boldsymbol{u}\cdot\ddot{x})}{-\boldsymbol{u}\cdot\dot{x}}\right)\frac{1}{(\boldsymbol{u}\cdot\dot{x})^2}\right)\right)$

and

'$\boldsymbol{u}$-time derivative of $\boldsymbol{u}$-relative momentum' $\left(\left(\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (m\dot{x})\right)' = m\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \ddot{x}\frac{1}{-\boldsymbol{u}\cdot\dot{x}}\right)$

are not equal, contrary to the nonrelativistic case. It is a question, which of them is accepted for being on the left-hand side of the relative Newtonian equation. Theory does not answer the question. Experimental facts suggest the decision that the $\boldsymbol{u}$-relative Newtonian equation reads:

'$\boldsymbol{u}$-time derivative of $\boldsymbol{u}$-relative momentum = $\boldsymbol{u}$-relative force'.

13.5 Relative Newtonian equation

13.5.1 Definition

A standard inertial frame $\boldsymbol{u}$ the existence of material point with world line function r perceives as a motion described by the function $r_{\boldsymbol{u}} : \mathtt{I}_{\boldsymbol{u}} \to \mathtt{E}_{\boldsymbol{u}}$, $t \mapsto \sigma_{\boldsymbol{u}}(r(\boldsymbol{s}(t)))$ where $t \mapsto \boldsymbol{s}(t)$ is the proper time of the material point as a function of the frame time (see 13.2.1). This motion satisfies a relative Newtonian equation which, according to what has been said previously, is accepted in the form 'time derivative of relative momentum = relative force'.

The $\boldsymbol{u}$-relative force can depend on $\boldsymbol{u}$-time points, $\boldsymbol{u}$-space points and $\boldsymbol{u}$-relative velocity; the $\boldsymbol{u}$-relative momentum can be given by $\boldsymbol{u}$-relative velocity,

$$m(\dot{x} + \mathbf{u}(\mathbf{u} \cdot \dot{x})) = m(-\mathbf{u} \cdot \dot{x}) \left(\frac{\dot{x}}{-\mathbf{u} \cdot \dot{x}} - \mathbf{u} \right) = \frac{m\mathbf{v}_{\dot{x}\mathbf{u}}}{\sqrt{1 - |\mathbf{v}_{\dot{x}\mathbf{u}}|^2}},$$

thus,the **$\boldsymbol{u}$-relative Newtonian equation** has the form

$$(q : \mathtt{I}_{\boldsymbol{u}} \to \mathtt{E}_{\boldsymbol{u}})? \qquad \left(\frac{mq'}{\sqrt{1 - |q'|^2}} \right)' = \boldsymbol{f}_{\boldsymbol{u}}(t, q, q').$$

Executing the differentiation, we get for the left-hand side:

$$\frac{m}{\sqrt{1 - |q'|^2}} \left(\mathbf{1} + \frac{q' \otimes q'}{1 - |q'|^2} \right) q''.$$

This shows the important fact that, in contrast with the nonrelativistic case, the **relative acceleration is not parallel to relative force**.

13.5.2 Relative forces

Contrary to the nonrelativstic case, the relative force differs from the absolute one not only in substituting the absolute variables by the relative ones.

Taking into account the time derivative of the relative momentum

$$(\boldsymbol{\sigma_u}\cdot(m\dot{x})\bullet)' = \left(\boldsymbol{\sigma_u}\cdot(m\ddot{x})\frac{1}{-\boldsymbol{u}\cdot\dot{x}}\right)\bullet$$

and the absolute Newtonian equation, we get $\left(\boldsymbol{\sigma_u}\cdot\boldsymbol{f}(x,\dot{x})\frac{1}{-\boldsymbol{u}\cdot\dot{x}}\right)\bullet$ for the $\boldsymbol{u}$-relative force where, of course, the absolute variables are to be substituted by the relative ones; finally:

$$\boldsymbol{f_u}(t,q,q') = \boldsymbol{\sigma_u}\cdot\boldsymbol{f}\left(t\cap q, \frac{\boldsymbol{u}+q'}{\sqrt{1-|q'|^2}}\right)\sqrt{1-|q'|^2}$$

where $(t,q,q') \in I_{\boldsymbol{u}} \times \mathrm{E}_u \times \frac{\mathbf{E}_u}{\mathbb{I}}$.

Let us examine the form of a relative force corresponding to an absolute force having the potential $\boldsymbol{K}$.

Recall the formulae of Subsection 12.13 for $\boldsymbol{K}$.

Let

$$\boldsymbol{K} \quad \text{have the } \boldsymbol{u}\text{-split form} \quad (-V_{\boldsymbol{u}}, \boldsymbol{A_u});$$

then

$$\boldsymbol{F} := \mathrm{D}\wedge\boldsymbol{K} \quad \text{have the } \boldsymbol{u}\text{-split form} \quad ((-\nabla_{\boldsymbol{u}}V_{\boldsymbol{u}} - \mathrm{D}_{\boldsymbol{u}}\boldsymbol{A_u}, \nabla_{\boldsymbol{u}}\wedge\boldsymbol{A_u})) =: ((\boldsymbol{E_u}, \boldsymbol{B_u})).$$

Further,

$$\boldsymbol{\sigma_u}\cdot\boldsymbol{F}(x)\cdot\dot{x}\frac{1}{-\boldsymbol{u}\cdot\dot{x}} = \boldsymbol{\sigma_u}\cdot\boldsymbol{F}(x)\cdot\boldsymbol{u} + \boldsymbol{\sigma_u}\cdot\boldsymbol{F}(x)\cdot\left(\frac{\dot{x}}{-\boldsymbol{u}\cdot\dot{x}} - \boldsymbol{u}\right);$$

the first member is just the $\boldsymbol{u}$-time component of $\boldsymbol{F}$, in the second member $\mathbf{v}_{\dot{x}\boldsymbol{u}} = \boldsymbol{\sigma_u}\cdot\mathbf{v}_{\dot{x}\boldsymbol{u}}$ stands beside $\boldsymbol{F}(x)$, so there the $\boldsymbol{u}$-spacelike component of $\boldsymbol{F}$ and the $\boldsymbol{u}$-relative velocity appear, thus

$$\boldsymbol{\sigma_u}\cdot\boldsymbol{F}(x)\cdot\dot{x}\frac{1}{-\boldsymbol{u}\cdot\dot{x}} \quad \text{in the } \boldsymbol{u}\text{-split form is} \quad \boldsymbol{E_u} + \boldsymbol{B_u}\cdot\mathbf{v}_{\dot{x}\boldsymbol{u}}.$$

We recognize: in the electromagnetic case $V_{\boldsymbol{u}}$ is the scalar potential, $\boldsymbol{A_u}$ is the vector potential, $\boldsymbol{E_u}$ is the electric force and $\boldsymbol{B_u}\cdot\mathbf{v}_{\dot{x}\boldsymbol{u}}$ is the magnetic Lorentz force. Of course, the above formulae are valid not only for electromagnetism but for other forces, too, having a potential. In contrast to the nonrelativistic case, however, here the possibilities are less as we shall see.

Finally we show how the absolute force can be restored from the relative one. For the sake of simplicity, we omit denoting variables. The $\boldsymbol{u}$-relative force from the absolute one is obtained in the form

$$\boldsymbol{f_u} = \frac{\mathbf{1} + \boldsymbol{u} \otimes \boldsymbol{u}}{-\boldsymbol{u} \cdot \dot{x}} \cdot \boldsymbol{f}$$

Since

$$\left(\mathbf{1} + \frac{\boldsymbol{u} \otimes \dot{x}}{-\boldsymbol{u} \cdot \dot{x}}\right)(\mathbf{1} + \boldsymbol{u} \otimes \boldsymbol{u}) = \mathbf{1} + \frac{\boldsymbol{u} \otimes \dot{x}}{-\boldsymbol{u} \cdot \dot{x}},$$

and $\dot{x} \cdot \boldsymbol{f} = 0$, we get

$$\boldsymbol{f} = (-\boldsymbol{u} \cdot \dot{x}) \left(\mathbf{1} + \frac{\boldsymbol{u} \otimes \dot{x}}{-\boldsymbol{u} \cdot \dot{x}}\right) \boldsymbol{f_u} = (-\boldsymbol{u} \cdot \dot{x})\boldsymbol{f_u} + \boldsymbol{u}(\dot{x} \cdot f_{\boldsymbol{u}}) = (\boldsymbol{u} \wedge \boldsymbol{f_u}) \cdot \dot{x}. \quad \text{(V55)}$$

Thus, we have the explicit result, denoting the variables, too:

$$\boldsymbol{f}(x, \dot{x}) = \big(\boldsymbol{u} \wedge \boldsymbol{f_u}(\tau_{\boldsymbol{u}}(x), \sigma_{\boldsymbol{u}}(x), \mathbf{v}_{\dot{x}\boldsymbol{u}})\big) \cdot \dot{x}.$$

13.5.3 The role of mass

In the nonrelativistic case the same quantity appears in three different roles:

1) m is the multiplier of absolute acceleration in the absolute Newtonian equation,

2) m is the multiplier of relative velocity in the formula of relative momentum,

3) m is the multiplier of relative acceleration in the relative Newtonian equation.

In the relativistic case, however, three different quantities appear in three different roles.

1) m is the multiplier of absolute acceleration in the absolute Newtonian equation,

2) $\frac{m}{\sqrt{1-|q'|^2}}$ is the multiplier of relative velocity in the formula of relative momentum,

3) $\frac{m}{\sqrt{1-|q'|^2}}\left(\mathbf{1} + \frac{q' \otimes q'}{1-|q'|^2}\right)$ is the 'multiplier' of relative acceleration in the relative Newtonian equation.

The quantities above are usually called

1) 'rest mass',

2) 'motion mass',

3) an amalgamation of 'longitudinal mass' and 'transversal mass'.

In general, the origin of these names is not pointed out clearly which can lead to conceptual troubles.

The best is to call **mass** only the **first quantity** above.

The **second quantity** has another (by the way, usual) name: **relative energy** (see subsection 13.7).

The third quantity is quite unnecessary to name at all.

13.6 Some special absolute forces

Most of the well tractable and important forces in the nonrelativistic case have no relativistic counterpart. The reason is that the absolute force must be Lorentz orthogonal to the actual absolute velocity.

13.6.1 The simplest cases

a) There is no velocity independent absolute force, in particular, there is no constant force.

It may happen, however, that a $\boldsymbol{u}$-relative force is e.g. constant. Then the absolute force can be restored by formula (V55).

If the $\boldsymbol{u}$-relative force is the constant $\boldsymbol{h} \in \frac{\mathbf{E}_{\boldsymbol{u}}^*}{\mathbb{I}}$ then the absolute force is

$$\boldsymbol{f}(x, \dot{x}) = (\boldsymbol{u} \wedge \boldsymbol{h}) \cdot \dot{x}.$$

This force has a potential,

$$\boldsymbol{K}(x) = -\big(\boldsymbol{h} \cdot (x - o)\big)\boldsymbol{u}$$

for arbitrary $o \in \mathtt{M}$. Taking into account that $-\boldsymbol{u} = \boldsymbol{\tau}_{\boldsymbol{u}}$ and now $\boldsymbol{h} \cdot (x - o) = \boldsymbol{h} \cdot (\boldsymbol{\sigma}_{\boldsymbol{u}}(x - o))$, we find that the present formula agrees the nonrelativistic (IV24).

b) The $\mathbf{u}_c$-static absolute force can be defined as nonrelativistically:

$$\boldsymbol{f}(x, \dot{x}) = \boldsymbol{f}(x + \boldsymbol{t}\mathbf{u}_c, \dot{x})$$

holds for all $\boldsymbol{t} \in \mathbb{I}$. Equivalently, there is a $\boldsymbol{h} : \mathbf{E}_{\boldsymbol{u}} \times \mathtt{V}(1) \to \frac{\mathbf{M}^*}{\mathbb{I}}$ function and $o \in \mathtt{M}$ such that

$$\boldsymbol{f}(x, \dot{x}) = \boldsymbol{h}(\boldsymbol{\sigma}_{\mathbf{u}_c} \cdot (x - o), \dot{x}).$$

13.6.2 Central forces

In general, there are no central forces because there is no world point of the centre absolutely simultaneous whit a world point x.

Exceptions are the inertial centres when we can consider standard simultaneity determined by the centre. Let $\boldsymbol{u}_c$ be the absolute velocity of the centre. Then we accept that the $\boldsymbol{u}_c$-relative force has the form known in the nonrelativistic case:.

$$\boldsymbol{f}_{\boldsymbol{u}_c}(t,q,q') = a(|q-q_o|)(q-q_o)$$

where q_o is the position of the centre in the space of the inertial observer $\boldsymbol{u}_c$ and $a : \mathbb{I} \to \frac{\mathbb{R}}{\mathbb{I}\otimes\mathbb{I}\otimes\mathbb{I}}$ is a function.

Then, if o is an arbitrary occurrence of q_o i.e. $q_o = o+\mathbb{I}\boldsymbol{u}_c$ – since $(x+\mathbb{I}\boldsymbol{u}_c)-(o+\mathbb{I}\boldsymbol{u}_c) = \boldsymbol{\sigma}_{\boldsymbol{u}_c}\cdot(x-o)$ and $\boldsymbol{\sigma}_{\boldsymbol{u}_c}\cdot(x-o)\wedge\boldsymbol{u}_c = (x-o)\wedge\boldsymbol{u}_c$ – the absolute force is

$$\boldsymbol{f}(x,\dot{x}) = a(|\boldsymbol{\sigma}_{\boldsymbol{u}_c}\cdot(x-o)|)\big(\boldsymbol{u}_c\wedge(x-o)\big)\dot{x}.$$

This force has a potential,

$$\boldsymbol{K}(x) = -b(|\boldsymbol{\sigma}_{\boldsymbol{u}_c}\cdot(x-o)|)\boldsymbol{u}_c$$

where b is a primitive function of $\mathrm{s}\mapsto a(\mathrm{s})\mathrm{s}$. Taking into account that $-\boldsymbol{u}_c = \boldsymbol{\tau}_{\boldsymbol{u}_c}$ this formula, too, agrees the nonrelativistic (IV25).

13.7 Kinetic energy and power

We had to decide which possibility is to be chosen regarding the relative Newtonian equation. Now we make a similar decision: corresponding to the nonrelativistic case, we accept that the **relative power is the product of relative force and relative velocity.**

For the sake of brevity, in the sequel the variables of forces (both absolute and relative ones) are omitted from the notations.

Thus, the $\boldsymbol{u}$-relative power of of the absolute force f is

$$(\boldsymbol{f}_{\boldsymbol{u}}\cdot\mathbf{v}_{\dot{x}\boldsymbol{u}})\bullet = \left(\frac{\boldsymbol{f}+\boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{f})}{-\boldsymbol{u}\cdot\dot{x}}\cdot\left(\frac{\dot{x}}{-\boldsymbol{u}\cdot\dot{x}}-\boldsymbol{u}\right)\right)\bullet = \left(\frac{-\boldsymbol{u}\cdot\boldsymbol{f}}{-\boldsymbol{u}\cdot\dot{x}}\right)\bullet \qquad \text{(V56)}$$

which is just the negative of the $\boldsymbol{u}$-timelike component of the force considered a covector, multiplied by the reciprocal of the relativistic factor. The reader is asked to what has been said in 7.7.

Further, we have

$$\left(\frac{-\boldsymbol{u}\cdot\boldsymbol{f}}{-\boldsymbol{u}\cdot\dot{x}}\right)\bullet=\left(\frac{-\boldsymbol{u}\cdot m\ddot{x}}{-\boldsymbol{u}\cdot\dot{x}}\right)\bullet=\big(m(-\boldsymbol{u}\cdot\dot{x})\bullet\big)'=\left(\frac{m}{\sqrt{1-|\mathbf{v}_{\dot{x}\boldsymbol{u}}|^2}}\bullet\right)'.$$

On the base of this formula, one usually calls relative energy the quantity in the paranthesis of the last (or last but one) right hand because its time derivative is the power. This is, however, not correct because adding an arbitrary constant to it, the Further, in the nonrelativistic case the time derivative of kinetic energy equals the product of relative force and relative velocity. Thus, we have to find a quantity in the parenthesis which is zero for zero relative velocity.

That is why we accept that

$$m(-\boldsymbol{u}\cdot\dot{x})-m=\frac{m}{\sqrt{1-|\mathbf{v}_{\dot{x}\boldsymbol{u}}|^2}}-m$$

is the $\boldsymbol{u}$**-kinetic energy**.

This choice is supported be the fact, too, that for relative velocities whose magnitude is much less than 1 (which is the light speed) we have – from the power series of the square root – that the kinetic energy approximately equals the nonrelativistic kinetic energy $m|\mathbf{v}_{\dot{x}\boldsymbol{u}}|^2/2$.

Nevertheless,

$$m(-\boldsymbol{u}\cdot\dot{x})=\frac{m}{\sqrt{1-|\mathbf{v}_{\dot{x}\boldsymbol{u}}|^2}}$$

can be accepted as the $\boldsymbol{u}$**-relative energy** but this does not follow from the previous considerations; this will be supported by considering processes in which material points join and separate or emit light particles (photons) treated later.

13.8 Conservation laws

13.8.1 There is no action-reaction

The interaction of two material points – briefly particles – which do not contact cannot be described by forces: contrary to the nonrelativistic case, there are no absolute instants, so instantaneous action does not exist. In other words, there are no actions at an instant.

The possibility of interaction occurs only if the particles contact. Of course, it may happen that a particle 'radiates' tiny objects which 'hit' the other particle, and vice versa. Then the interaction is realized by the contact of the particles and the tiny objects.

13.8.2 Collisions

We accept as a fundamental physical fact that absolute momentum is conserved in collisions.

Let two particles meet and join together (inelastic collision). Let the particles have masses m_1 and m_2 and absolute velocities $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, respectively. Let the arising new particle have mss m_3 and absolute velocity $\boldsymbol{u}_3$. The the conservation of total momentum gives

$$m_1\boldsymbol{u}_1 + m_2\boldsymbol{u}_2 = m_3\boldsymbol{u}_3.$$

The $\boldsymbol{u}$-spacelike component of this equality gives the **conservation of $\boldsymbol{u}$-relative moment**:

$$m_1\big(\boldsymbol{u}_1 + \boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{u}_1)\big) + m_2\big(\boldsymbol{u}_2 + \boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{u}_2)\big) = m_3\big(\boldsymbol{u}_3 + \boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{u}_3)\big)$$

which, expressed by relative velocities is

$$\frac{m_1\mathbf{v}_{\boldsymbol{u}_1\boldsymbol{u}}}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}_1\boldsymbol{u}}|^2}} + \frac{m_2\mathbf{v}_{\boldsymbol{u}_2\boldsymbol{u}}}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}_2\boldsymbol{u}}|^2}} = \frac{m_3\mathbf{v}_{\boldsymbol{u}_3\boldsymbol{u}}}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}_3\boldsymbol{u}}|^2}}.$$

Of course, the $\boldsymbol{u}$-timelike component gives

$$m_1(-\boldsymbol{u}\cdot\boldsymbol{u}_1) + m_2(-\boldsymbol{u}\cdot\boldsymbol{u}_2) = m_3(-\boldsymbol{u}\cdot\boldsymbol{u}_3)$$

which, supported by later arguments, is the **conservation of $\boldsymbol{u}$-relative energy**. Expressed by relative velocities,

$$\frac{m_1}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}_1\boldsymbol{u}}|^2}} + \frac{m_2}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}_2\boldsymbol{u}}|^2}} = \frac{m_3}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}_3\boldsymbol{u}}|^2}}.$$

Let us take $\boldsymbol{u}_3$ for the role of $\boldsymbol{u}$ (let us consider the standard inertial frame in which the new particle is at rest). Then

$$\frac{m_1}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}_1\boldsymbol{u}_3}|^2}} + \frac{m_2}{\sqrt{1-|\mathbf{v}_{\boldsymbol{u}_2\boldsymbol{u}_3}|^2}} = m_3;$$

the multipliers on the left-hand side are greater than 1 – if $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ does not equal $\boldsymbol{u}_3$ (this is the case of a real collision) –, consequently,

$$m_1 + m_2 < m_3.$$

Mass is not conserved when two particles join in a collision.

Experience suggested us to introduce in the nonrelativistic case the notion of inner energy which admitted us to formulate energy conservation: the inner energy of the particle after collision is greater then the sum of the inner energies of the particles before collision, the difference is just the difference between the kinetic energies after and before collision, respectively.

Now the difference between the $\boldsymbol{u}$-kinetic energies aftrer and before collision is

$$\big(m_1(-\boldsymbol{u}\cdot\boldsymbol{u}_1)-m_1\big)+\big(m_2(-\boldsymbol{u}\cdot\boldsymbol{u}_2)-m_2\big)-\big(m_3(-\boldsymbol{u}\cdot\boldsymbol{u}_1)-m_3\big)=m_3-(m_1+m_2).$$

Thus, the the difference between the $\boldsymbol{u}$-kinetic energies is just the difference between the masses after and before collision. In analogy with the nonrelativistic case, we can say that here mass takes the role of inner energy, therefore we accept:

'$\boldsymbol{u}$-relative energy:= $\boldsymbol{u}$-kinetic energy + mass =
= $\boldsymbol{u}$-timelike component of absolute momentum.'

13.8.3 Particles and photons

The previous idea regarding energy will be more convincing if we consider light emission and absorption. It is a simple empirical fact that a body emitting or absorbing light will be colder or warmer, respectively.

The light emission and absorption as well as reflection can be well treated as a collision of particles and photons. The latter ones are considered 'tiny' objects that have future-lightlike momentum.

Note that the absolute momentum $\boldsymbol{p}$ of a particle with mass m is a futurelike vector for which $\boldsymbol{p}\cdot\boldsymbol{p}=-m^2$ holds.

If $\boldsymbol{k}$ is the absolute momentum of a photon then $\boldsymbol{k}\cdot\boldsymbol{k}=0$ which is conceived that a photon has no mass.

Let a particle absorb a photon; the balance of momentum is

$$\mathrm{m}_1\boldsymbol{u}_1+\boldsymbol{k}=\mathrm{m}_2\boldsymbol{u}_2$$

Lorentz multiplying by $-\boldsymbol{u}_2$ we arrive at

$$m_1(-\boldsymbol{u}_2\cdot\boldsymbol{u}_1)-\boldsymbol{u}_2\cdot\boldsymbol{k}=m_2.$$

Since $-\boldsymbol{u}_2\cdot\boldsymbol{u}_1>1$ and $-\boldsymbol{u}_2\cdot\boldsymbol{k}>0$, we see that $m_2>m_1$. We have got that light absorption which means inner energy increase in the nonrelativistic case results in mass increase in the relativistic case. The increase of mass is

$$m_2 - m_1 = \big(m_1(-\boldsymbol{u}_2 \cdot \boldsymbol{u}_1) - m_1\big) - \boldsymbol{u}_2 \cdot \boldsymbol{k}.$$

The $\boldsymbol{u}_2$-kinetic energy before absorption is $m_1(-\boldsymbol{u}_2 \cdot \boldsymbol{u}_1) - m_1$, after absorption is zero. Thus, according to the nonrelativistic reasoning we would say that the kinetic energy before collision and the energy of the photon is transformed into inner energy i.e. the sum of the energy of the particle and the photon before collision equals the inner energy increase of the particle. In the right-hand side of the above equality we find the total energy before collision while the left-hand side is the mass increase. This support our idea that in the relativistic case mass takes the role of inner energy.

More interesting is the process in which a particle is annihilated by radiating two photons. It has the balance of momentum

$$m\boldsymbol{u} = \boldsymbol{k}_1 + \boldsymbol{k}_2$$

from which we get

$$m = -\boldsymbol{u} \cdot \boldsymbol{k}_1 - \boldsymbol{u} \cdot \boldsymbol{k}_2.$$

The particle is annihilated, mass disappears and photons appear. These photons can be absorbed by another particle which causes mass increase (inner energy increase from a nonrelativistic point of view) of that particle.

13.8.4 Equivalence of mass and energy?

Let us examine the famous result of Einstein which is formulated as the equivalence of mass and energy. Summarizing our results, we can say that

in the nonrelativistic case we have

- relative energy= kinetic energy + inner energy,
- kinetic energy → increase of inner energy,
- relative momentum= mass × relative velocity,

whereas in the relativistic case we have

- relative energy = kinetic energy + mass,
- kinetic energy → increase of mass,
- relative momentum= relative energy × relative velocity.

In the first and second relations mass takes the role of inner energy. In the third relation relative energy takes the role of mass. Even if we called the change of roles, we would run into trouble because neither of the two statements would be true: 1) mass is equivalent to inner energy, 2) mass is equivalent to relative energy.

Since mass is an absolute quantity (independent of reference frames) both in the nonrelativistic and the relativistic cases and inner energy is absolute, perhaps the best is to say that the relativistic mass unites the notions of nonrelativistic mass and inner energy (but neither this statement reflects exactly the situation).

13.9 The rocket equation

The relativistic rocket equation, as the nonrelativistic one, is based on the conservation of absolute momentum. The mass of rocket is given as a function of the proper time of the rocket, $m : \mathtt{I} \to \mathbb{I}^*$, and the relative velocity of the mass outflow with respect to the rocket, as a function of proper time, $\boldsymbol{v} : \mathtt{I} \to \frac{\mathbf{M}}{\mathbb{I}}$ in such a way that if r is the world line function of the rocket then $\boldsymbol{v}(\boldsymbol{s})$ in $\frac{\mathbf{E}_{\dot{r}(s)}}{\mathbb{I}}$ i.e. $\dot{r}(\boldsymbol{s}) \cdot \boldsymbol{v}(\boldsymbol{s}) = 0$.

Let us write the balance of absolute momentum for the proper time values $\boldsymbol{s}$ and $\boldsymbol{s} + \boldsymbol{h}$; contrary to the nonrelativistic case, since mass is not conserve, we cannot state that the mass outflow between the two proper time points is $m(\boldsymbol{s}) - m(\boldsymbol{s}+\boldsymbol{h})$. For the time being, we can say only that it is of the form $\mu(\boldsymbol{s})\boldsymbol{h} +$ ordo$(\boldsymbol{h})$ and has the absolute velocity $\frac{\dot{r}(\boldsymbol{s}) + \boldsymbol{v}(\boldsymbol{s})}{\sqrt{1 - |\boldsymbol{v}(\boldsymbol{s})|^2}} + \mathrm{ordo}(\boldsymbol{h})$. Consequently,

$$m(\boldsymbol{s})\dot{r}(\boldsymbol{s}) = m(\boldsymbol{s}+\boldsymbol{h})\dot{r}(\boldsymbol{s}+\boldsymbol{h}) + \big(\mu(\boldsymbol{s})\boldsymbol{h} + \mathrm{ordo}(\boldsymbol{h})\big)\left(\frac{\dot{r}(\boldsymbol{s}) + \boldsymbol{v}(\boldsymbol{s})}{\sqrt{1 - |\boldsymbol{v}(\boldsymbol{s})|^2}} + \mathrm{ordo}(\boldsymbol{h})\right).$$

Adding to and subtracting from the right-hand side $m(\boldsymbol{s})\dot{r}(\boldsymbol{s}+\boldsymbol{h})$ and rearranging the terms, then dividing by $\boldsymbol{h}$ and letting it tend to zero, we get

$$m(\boldsymbol{s})\ddot{r}(\boldsymbol{s}) + \dot{m}(\boldsymbol{s})\dot{r}(\boldsymbol{s}) + \mu(\boldsymbol{s})\frac{\dot{r}(\boldsymbol{s}) + \boldsymbol{v}(\boldsymbol{s})}{\sqrt{1 - |\boldsymbol{v}(\boldsymbol{s})|^2}} = 0.$$

Lorentz multiplying by $\dot{r}(\boldsymbol{s})$ we obtain

$$\mu = -\dot{m}\sqrt{1 - |\boldsymbol{v}|^2}.$$

Finally, taking into account the force acting on the rocket as well, we have the **rocket equation**

$$(x : \mathbb{I} \mapsto \mathtt{M})? \qquad m\ddot{x} - \dot{m}\boldsymbol{v} = \boldsymbol{f}(x, \dot{x})$$

which is formally the same as the nonrelativistic one.

There is a trouble, however. Namely, this would determine the world line of the rocket but the function $\boldsymbol{v}$ can be given if the world line function of the rocket is known because $\boldsymbol{v}(\boldsymbol{s}) \cdot \dot{r}(\boldsymbol{s}) = 0$ must hold. We can avoid this trouble if we imagine the rocket in the space of the inertial observer $\boldsymbol{u} := r(0)$, we give the relative velocity of the outflow in that space, i.e. a function $\hat{\boldsymbol{v}} : \mathbb{I} \to \frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}}$ which will be boosted to the actual absolute velocity, $\boldsymbol{v}(\boldsymbol{s}) := \boldsymbol{B}_{\dot{r}(\boldsymbol{s}),\boldsymbol{u}} \cdot \hat{\boldsymbol{v}}(\boldsymbol{s})$. Then the equation

$$(x : \mathbb{I} \mapsto \mathtt{M})? \qquad m\ddot{x} - \dot{m}\boldsymbol{B}_{\dot{x},\boldsymbol{u}} \cdot \hat{\boldsymbol{v}} = \boldsymbol{f}(x, \dot{x})$$

is well defined.

14 Fundamental properties of electromagnetism in spacetime

The equations of electromagnetism – Maxwell equations – plaid a fundamental role in constucting relativity theory. Light is an electromagnetic phenomenon. We based the relativistic spacetime model for the properties of light propagation. This section will show how the relativistic spacetime model corrects the troubles of nonrelativistic electromagnetism.

14.1 Maxwell equations

The experimentally verified usual relative Maxwell equations (IV27) hold relativistically, too, if we consider them concerning a standard inertial frame $\boldsymbol{u}$; then the formulae of 12.8.1 and 12.13 give

$$\begin{aligned}
\nabla_{\boldsymbol{u}} \cdot \boldsymbol{D}_{\boldsymbol{u}} &= \rho_{\boldsymbol{u}}, \\
-\mathrm{D}_{\boldsymbol{u}}\boldsymbol{D}_{\boldsymbol{u}} + \nabla_{\boldsymbol{u}} \cdot \boldsymbol{H}_{\boldsymbol{u}} &= \boldsymbol{j}_{\boldsymbol{u}}, \\
\nabla_{\boldsymbol{u}} \wedge \boldsymbol{E}_{\boldsymbol{u}} + \mathrm{D}_{\boldsymbol{u}}\boldsymbol{B}_{\boldsymbol{u}} &= 0, \\
\nabla_{\boldsymbol{u}} \wedge \boldsymbol{B}_{\boldsymbol{u}} &= 0,
\end{aligned}$$

where

$$\rho_{\boldsymbol{u}} := -\boldsymbol{u} \cdot \mathcal{J}, \qquad \boldsymbol{j}_{\boldsymbol{u}} = \boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \mathcal{J},$$

$$-\boldsymbol{D}_{\boldsymbol{u}} := -\boldsymbol{G} \cdot \boldsymbol{u}, \qquad \boldsymbol{H}_{\boldsymbol{u}} := \boldsymbol{G} - \boldsymbol{u} \wedge (\boldsymbol{G} \cdot \boldsymbol{u}),$$

$$\boldsymbol{E}_{\boldsymbol{u}} := \boldsymbol{F} \cdot \boldsymbol{u}, \qquad \boldsymbol{B}_{\boldsymbol{u}} := \boldsymbol{F} - \boldsymbol{u} \wedge (\boldsymbol{F} \cdot \boldsymbol{u}),$$

where $\mathcal{J}$, $\boldsymbol{G}$ and $\boldsymbol{F}$ have the same physical meaning as in the nonrelativistic and their mathematical expression is the same as well if we substitute $\mathbb{I}$ for $\mathbb{D}$.

Of course, the absolute Maxwell equations, too, have the same form:

$$\mathrm{D} \cdot \boldsymbol{G} = \mathcal{J}, \qquad \mathrm{D} \wedge \boldsymbol{F} = 0. \tag{V57}$$

14.2 Vacuum constituive relations

As in the nonrelativistic case, we have to give a **constitutive relation**

$$\boldsymbol{G} = \Gamma(\boldsymbol{F})$$

which reflects how a material medium in spacetime influences the electromagnetic phenomena.

The constitutive relation due to a real medium can be formulated here in the same way. There is an essential difference for vacuum because now we have a convenient constitutive relation **without introducing the ether, a fictitious medium**.

Namely, since now $\mathbb{D} = \mathbb{I}$, the electromagnetic displacement $\boldsymbol{G}$ has values in $\frac{\mathbf{M} \wedge \mathbf{M}}{\mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I}}$ which can be identified with $\mathbf{M}^* \wedge \mathbf{M}^*$, so $\boldsymbol{G}$ and $\boldsymbol{F}$ become quantities of the same kind.

Therefore the vacuum constitutive relation can be formulated as

$$\boldsymbol{G} = \boldsymbol{F}$$

and the absolute Maxwell equations in vacuum reads

$$\mathrm{D} \cdot \boldsymbol{F} = \mathcal{J}, \qquad \mathrm{D} \wedge \boldsymbol{F} = 0.$$

15 Noninertial observers

Noninertial observers can also be well treated in the special relativistic spacetime model but in a more complicated way than in the nonrelativistic case. The reason is the absence of absolute simultaneity. Now we treat only some problems. Further information such as a detailed treatment of uniformly accelerated observers and uniformly rotating observers can be found in the book

T. Matolcsi: *Spacetime without Reference Frames* (Budapest, 1993, Akadémiai Kiadó)

15.1 Nearly standard local synchronizations

What would be the standard synchronizationof a noninertial observer?

Of course, a noninertial observer can send light signals form a 'centre' and makes them reflect; the mid time between the start and the return of the light signal is defined to be simultaneous with the reflection. Such a synchronization, if exists at all, may have the following unpleseant properties:

– different times pass in different space points between simultaneous occurrences,

– different 'centres' establish different synchronizations.

A noninertial observer, choosing an arbitrary point of its space, can establish a synchronization in a neighbourhood of that space point in such a way that light speed in that space point is the same in every direction. This **nearly standard local synchronization** in the given space point results in isotropic one-way light propagation in that space point.

In the model, the instants of nearly standard local synchronizationin the space point q of an observe $\boldsymbol{U}$ are defined as follows: let the elements (world points) of $x + \mathbf{E}_{\boldsymbol{U}(x)}$ be simultaneous with the occurrence x of the $\boldsymbol{U}$-space point (world line). As Figure 15.1 shows different such hyperplanes can meet, therefore such a definition can be good only in a neighbourhood of q; it can be proved that such a neighbourhood exists.

Let us proceed in this way applying 'tiny steps', from Greenwich to London, from London to Dover, from Dover to Paris etc. In the model, applying the limit with infinitesimal steps we obtain a world surface whose tangent space at every point x is Lorentz orthogonal (Ux). to the absolute velocity. Such a synchronization can be called the synchronization of the observer.

In this way, however, we have not eliminated the problems. First of all, **in general, no standard synchronization exists to a noninertial observer**. And even if it exists, the problem remains that different times pass in different space points between simultaneous occurrences.

15.2 Synchronizations of a uniformly rotating observer

Recall that we mentioned two synchronizations on the Earth: the one is established by light signals (from Greenwich to Budapest etc.), the other one is established by the position of stars (Sun) (see Paragraph 12.6.1).

The stars form an inertial observer. The axis of the Earth is considered to be at rest with respect to this inertial observer. The synchronization corresponding to the stars is the standard synchronizationof an inertial observer.

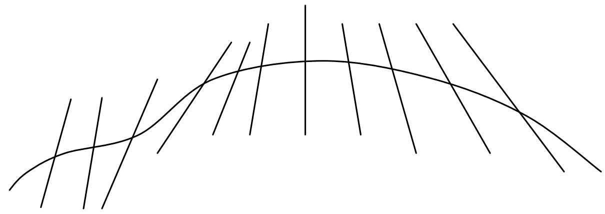

Figure 15.1 Nearly standard local synchronization

The synchronization with light signal is defined in such a way that one-way light speed in the 'centre' is the same in every direction. Thus, this synchronization is the nearly standard local synchronizationin a point of the Earth surface.

These synchronizations are theoretically different but practically same. Namely, the synchronization corresponding to the stars gives that the light speed on the Equator towards east and west are

$$c_+ = \frac{1}{1 - 1,6 \cdot 10^{-6}}, \qquad c_- = \frac{1}{1 + 1,6 \cdot 10^{-6}},$$

respectively.

16 Two recent paradoxes

16.1 Velocity addition paradox

The paradox can be described as follows[4]. Let us consider three standard inertial frames: me, you and him for the sake of easy formulation. Your velocity $\boldsymbol{v}$ relative to me and his velocity $\boldsymbol{w}$ relative to you determine his velocity $\boldsymbol{v} \oplus \boldsymbol{w}$ relative to me by the formula

$$\boldsymbol{v} \oplus \boldsymbol{w} = \frac{\alpha\beta}{\gamma}\left(\boldsymbol{v} + \boldsymbol{w} + \frac{\alpha}{1+\alpha}\boldsymbol{v} \times (\boldsymbol{v} \times \boldsymbol{w})\right) = \frac{\alpha(\beta+\gamma)}{\gamma(1+\alpha)}\boldsymbol{v} + \frac{\beta}{\gamma}\boldsymbol{w} \qquad \text{(V58)}$$

where

$$\alpha := \frac{1}{\sqrt{1 - |\boldsymbol{v}|^2}}, \qquad \beta := \frac{1}{\sqrt{1 - |\boldsymbol{w}|^2}}, \qquad \gamma := \alpha\beta(1 + \boldsymbol{v} \cdot \boldsymbol{w}).$$

[4]C.I.Mocanu, *Foundations of Physics Letters* 5(1992)73

Similarly, your velocity $\hat{\boldsymbol{w}}$ relative to him and my velocity $\hat{\boldsymbol{v}}$ relative to you determine my velocity $\hat{\boldsymbol{w}}\oplus\hat{\boldsymbol{v}}$ relative to him by the same formula. We 'evidently' have $\hat{\boldsymbol{w}} = -\boldsymbol{w}$, $\hat{\boldsymbol{v}} = -\boldsymbol{v}$ and $\hat{\boldsymbol{w}} \oplus \hat{\boldsymbol{v}} = -\boldsymbol{v} \oplus \boldsymbol{w}$; however, the actual formula for the addition $\oplus$ shows that, in general,

$$(-\boldsymbol{w}) \oplus (-\boldsymbol{v}) \neq -(\boldsymbol{v} \oplus \boldsymbol{w}), \qquad \text{or, equivalently,} \qquad \boldsymbol{w} \oplus \boldsymbol{v} \neq \boldsymbol{v} \oplus \boldsymbol{w}. \quad \text{(V59)}$$

The paradox arose in the usual formalism using coordinates where, instead of vectors in the spaces of different observers, one considers tacitly the corresponding vectors boosted in the space of the observer hidden in the coordinates (the 'rest frame'), which implies the incorrect tacit assumption that making boosts is a transitive relation. In fact, the relative velocities $\boldsymbol{v}$ and $\boldsymbol{w}$ as well as $\hat{\boldsymbol{w}}$ and $\hat{\boldsymbol{v}}$ are considered to be elements of $\mathbb{R}^3$, their sum and vectorial product appear in the formulae of $\boldsymbol{v} \oplus \boldsymbol{w}$ and $\hat{\boldsymbol{w}} \oplus \hat{\boldsymbol{v}}$, yielding elements of $\mathbb{R}^3$.

Now, let us see the explanation of the paradox returning to our notations[5]. Let $\boldsymbol{u}$, $\boldsymbol{u}'$ and $\boldsymbol{u}''$ denote me, you and him, respectively. Then $\boldsymbol{v}_{\boldsymbol{u}'\boldsymbol{u}}$ plays the role of $\boldsymbol{v}$ and $\boldsymbol{v}_{\boldsymbol{u}''\boldsymbol{u}'}$ plays the role of $\boldsymbol{w}$. However, $\boldsymbol{v}_{\boldsymbol{u}'\boldsymbol{u}}$ and $\boldsymbol{v}_{\boldsymbol{u}''\boldsymbol{u}'}$ are in the different three-dimensional vector spaces $\frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}}$ and $\frac{\mathbf{E}_{\boldsymbol{u}'}}{\mathbb{I}}$, their linear combination does not lie either in $\frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}}$ or in $\frac{\mathbf{E}_{\boldsymbol{u}'}}{\mathbb{I}}$ and their vectorial product is not meaningful.

The above velocity addition formula is meaningful and holds true only if the second relative velocity ($\boldsymbol{w}$) is boosted into the space of the observer ($\boldsymbol{u}$) to which the first velocity ($\boldsymbol{v}$) and the resulting one ($\boldsymbol{v} \oplus \boldsymbol{w}$) are related.

Thus we have to take $\boldsymbol{v} = \boldsymbol{v}_{\boldsymbol{u}'\boldsymbol{u}}$ and $\boldsymbol{w} = \boldsymbol{B}_{\boldsymbol{u}\boldsymbol{u}'}\boldsymbol{v}_{\boldsymbol{u}''\boldsymbol{u}'}$ and then

$$\boldsymbol{v} \oplus \boldsymbol{w} = \boldsymbol{v}_{\boldsymbol{u}''\boldsymbol{u}}. \tag{V60}$$

Regarding the other addition in the paradox involving $\hat{\boldsymbol{w}}$ and $\hat{\boldsymbol{v}}$, we must be careful: since the velocity of $\boldsymbol{u}$ relative to $\boldsymbol{u}''$ is calculated by the addition formula from the velocity of $\boldsymbol{u}'$ relative to $\boldsymbol{u}''$ and from the velocity of $\boldsymbol{u}$ relative to $\boldsymbol{u}'$, this last relative velocity must be boosted into the $\boldsymbol{u}''$-space, so

$$\hat{\boldsymbol{w}} := \boldsymbol{v}_{\boldsymbol{u}'\boldsymbol{u}''}, \qquad \hat{\boldsymbol{v}} := \boldsymbol{B}_{\boldsymbol{u}''\boldsymbol{u}'}\boldsymbol{v}_{\boldsymbol{u}\boldsymbol{u}'} \tag{V61}$$

and then

$$\hat{\boldsymbol{w}} \oplus \hat{\boldsymbol{v}} = \boldsymbol{v}_{\boldsymbol{u}\boldsymbol{u}''}. \tag{V62}$$

Note that $\hat{\boldsymbol{w}}$ and $-\boldsymbol{w}$ as well as $\hat{\boldsymbol{v}}$ and $-\boldsymbol{v}$ are in different spaces: the first ones in the $\boldsymbol{u}''$-space, the second ones in the $\boldsymbol{u}$-space. The 'evidence' used in

[5] T.Matolcsi-A.Goher, *Studies in History and Philosophy of Modern Physics* 32(2001)83

the formulation of the paradox that $\hat{\boldsymbol{w}}$ equals $-\boldsymbol{w}$ and $\hat{\boldsymbol{v}}$ equals $-\boldsymbol{v}$ would mean correctly that $\hat{\boldsymbol{w}}$ and $\hat{\boldsymbol{v}}$, boosted to the $\boldsymbol{u}$-space result in $-\boldsymbol{w}$ and $-\boldsymbol{v}$, respectively.

But this is not true:

$$\begin{aligned} \boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\hat{\mathbf{w}} &= \boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\mathbf{v}_{\mathbf{u}'\mathbf{u}''} = -\boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\boldsymbol{B}_{\mathbf{u}''\mathbf{u}'}\mathbf{v}_{\mathbf{u}''\mathbf{u}'} \\ &= -\boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\boldsymbol{B}_{\mathbf{u}''\mathbf{u}'}\boldsymbol{B}_{\mathbf{u}'\mathbf{u}}\mathbf{w} = -\mathbf{R}_{\boldsymbol{u}(\mathbf{u}',\mathbf{u}'')}\boldsymbol{w}, \end{aligned} \tag{V63}$$

$$\begin{aligned} \boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\hat{\mathbf{v}} &= \boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\boldsymbol{B}_{\mathbf{u}''\mathbf{u}'}\mathbf{v}_{\mathbf{u}\mathbf{u}'} = \\ &= -\boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\boldsymbol{B}_{\mathbf{u}''\mathbf{u}'}\boldsymbol{B}_{\mathbf{u}'\mathbf{u}}\mathbf{v}_{\mathbf{u}'\mathbf{u}} = -\mathbf{R}_{\boldsymbol{u}(\mathbf{u}',\mathbf{u}'')}\boldsymbol{v}. \end{aligned}$$

Then it is not surprising that $\hat{\boldsymbol{w}} \oplus \hat{\boldsymbol{v}}$, boosted to the $\boldsymbol{u}$-space is not equal to $\boldsymbol{v} \oplus \boldsymbol{w}$ either. All this is the consequence of the non-transitivity of the Lorentz boosts.

It is worth explaining the origin of the paradox from another point of view. The relative velocity $\mathbf{v}_{\mathbf{u}'\mathbf{u}''}$, boosted to $\boldsymbol{u}''$, will be opposite to the relative velocity $\mathbf{v}_{\mathbf{u}''\mathbf{u}'}$, and the relative velocity $\mathbf{v}_{\mathbf{u}''\mathbf{u}'}$, boosted to $\boldsymbol{u}'$, will be opposite to the relative velocity $\mathbf{v}_{\mathbf{u}'\mathbf{u}''}$. On the contrary, those relative velocities, boosted to $\boldsymbol{u}$, will not be opposite to each other: $\boldsymbol{B}_{\mathbf{u}\mathbf{u}'}\mathbf{v}_{\mathbf{u}''\mathbf{u}'} \neq -\boldsymbol{B}_{\mathbf{u}\mathbf{u}''}\mathbf{v}_{\mathbf{u}'\mathbf{u}''}$ (unless the three absolute velocities are coplanar which is equivalent to that the relative velocities $\mathbf{v}_{\mathbf{u}'\boldsymbol{u}}$ and $\mathbf{v}_{\mathbf{u}''\boldsymbol{u}}$ are collinear).

16.2 Light propagation paradox

Consider a light source on the rim of a rotating disk (platform). Light signals are sent around the rim (e.g. with the aid of mirrors) both forwards and backwards according to the direction of the rotation. We can measure the proper time periods of the source between the start and the arival of a light signals. Then knowing the distance covered by the light signals, (the circumference of the circle), we can determine the **round way** light speeds c_+ (forwards) and c_- (backwards). It can be shown [6] that

$$c_+ = \frac{1}{1+\omega d}, \qquad c_- = \frac{1}{1-\omega d}. \tag{V64}$$

where ω is the angular velocity of the rotation and d is the distance of the source from the centre of rotation.

The paradox appeared as follows[7].

[6] T. Matolcsi (1998) *Foundations of Physics* **27** 1865
[7] F.Selleri *Foundations of Physics Letters* 10(1997)73

It was stated that $\frac{1-\omega d}{1+\omega d}$ "does not give the ratio of global light velocities for a full trip around the platform in the two opposite directions, but the local ratio as well: isotropy of space ensures that the velocities of light are the same in all points of the rim and therefore the average value coincides with the local ones".

Then it is argued: consider uniformly rotating observers whose angular speed ω is smaller and smaller, and take their small pieces whose distance d from the centre is larger and larger such that $\beta := \omega d$ is constant. Then the ratio of light speeds in the opposite directions is the same $\frac{1-\beta}{1+\beta} \neq 1$ for all such pieces. However, these pieces become more and more similar to pieces of a limit inertial observer moving with speed β with respect to the centre. Thus the ratio will differ from unity for that limit inertial observer, in contradiction to special relativity, which asserts that light speed is the same in all directions with respect to inertial observers. Thus, accepting special relativity, we have a discontinuity which is not confirmed by experiments.

This conclusion is erroneous[8]. Namely, according to special relativity, **light speed is the same in all directions with respect to an inertial observer if and only if the standard synchronization of the observer is used**. And this error indicates the origin of the paradox as well. Namely, one speaks about light speed without specifying synchronization and confuses round-way (in fact circle-way) speed with one-way speed. Formula (V64) concerns the **circle-way speed** of light which does not tell anything about 'local', i.e. one-way speed. The circle-way speeds are meaningful without synchronization, but one-way speed makes sense only if a synchronization is given (and depends on the synchronization); so **any assertion regarding one-way (local) light speeds – e.g. a formula for their ratio – would be meaningful if a synchronization had been specified**.

The forward and backward one-way speeds are equal to the forward and backward circle way speeds, respectively, if the synchronization of the centre is used. The limit inertial observer moves with speed β with respect to the centre, and instead of its standard synchronization it uses the synchronization of the centre, therefore, it is not surprising that light speed is not the same in its space directions.

If the nearly standard local synchronization is applied at the light source, then in the limit we will get the standard synchronization of the limiting inertial observer.

[8]T.Matolcsi, *Foundations of Physics* 27(1998)1685

17 Non-standard formulae

17.1 Synchronization

In usual treatments of special relativity, the coordinates always refer – mostly in an implicit way – to standard synchronizations, therefore the importance of synchronizations gets lost, e.g. in connection with relative velocities. One speaks about relative velocity without specifying a synchronization and the basic axiom is formulated in such a (false) way that one-way light propagation is homogeneous and isotropic with respect to every inertial frame A recent paradox of relativity comes from an implicit application of a non-standard synchronization which results, of course, in a non-isotropic one-way light propagation which is in contrast with the usual (false) basic axiom (see Subsection 16.2).

Thus it is worth investigating a little the **non-standard synchronizations**; we shall get very instructive results.

A uniform synchronization is given by a three-dimensional linear subspace which is transverse to all absolute velocities; we suppose now that such a linear subspace does not contain lightlike vectors, i.e. it is spacelike. Then there is an absolute velocity $\boldsymbol{u}_\mathrm{s}$ such that the uniform synchronization in question is established by

$$\mathbf{E}_\mathrm{s} := \mathbf{E}_{\boldsymbol{u}_\mathrm{s}} = \{\mathbf{x} \in \mathbf{M} \mid \boldsymbol{u}_\mathrm{s} \cdot \mathbf{x} = 0\}.$$

The synchronization itself is denoted by $\mathbf{E}_\mathrm{s}$; the instants of this synchronization are hyperplanes parallel with $\mathbf{E}_\mathrm{s}$, their collection is denoted by $\mathtt{I}_\mathrm{s}$.

Taking an inertial observer $\boldsymbol{u}$, let us consider the **inertial reference frame** $(\boldsymbol{u}, \mathbf{E}_\mathrm{s})$.

The inertial observer $\boldsymbol{u}$ measures the time interval between two $\mathbf{E}_\mathrm{s}$-instants by the proper time passing in an arbitrary $\boldsymbol{u}$-space point. Thus, the time interval $\boldsymbol{t}$ between the $\mathbf{E}_\mathrm{s}$-instants t and s is determined as follows: if $y \in t$ and $x \in s$ then $(x - y - \boldsymbol{t}\boldsymbol{u}) \cdot \boldsymbol{u}_\mathrm{s} = 0$ which implies $\boldsymbol{t} = \frac{-\boldsymbol{u}_\mathrm{s} \cdot (x-y)}{uu \cdot \boldsymbol{u}_\mathrm{s}}$. Therefore, introducing the notation

$$\boldsymbol{\tau}_{\boldsymbol{u}\mathbf{s}} := \frac{-\boldsymbol{u}_\mathrm{s}}{-\boldsymbol{u} \cdot \boldsymbol{u}_\mathrm{s}}, \tag{V65}$$

we have that $\mathtt{I}_\mathrm{s}$ is an affine space overt $\mathbb{I}$ with the subtraction

$$(x + \mathbf{E}_\mathrm{s}) -_{\boldsymbol{u}} (y + \mathbf{E}_\mathrm{s}) := \boldsymbol{\tau}_{\boldsymbol{u}\mathbf{s}} \cdot (x - y). \tag{V66}$$

Now the inertial observer $\boldsymbol{u}$ represents its space vectors by the elements of $\mathbf{E}_\mathrm{s}$ in such a way that the vector $\boldsymbol{q}_\mathrm{s}$ between the $\boldsymbol{u}$-space points (straight lines

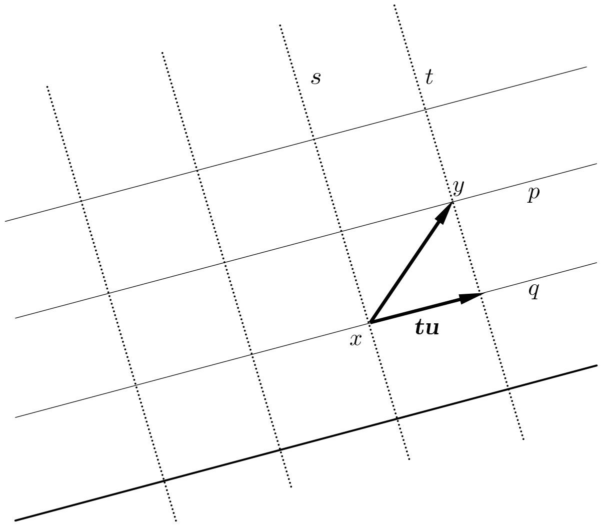

Figure 17.1 Non-standard synchronization

directed by $\boldsymbol{u}$) q and p is determined as follows: if $x \in q$ and $y \in p$ then $(x - y - \boldsymbol{q}_\mathrm{s})$ is parallel to $\boldsymbol{u}$.

It is simple to show that with the notation

$$\boldsymbol{\sigma}_{\mathbf{us}} := \mathbf{1} + \frac{\boldsymbol{u} \otimes \boldsymbol{u}_\mathrm{s}}{-\boldsymbol{u} \cdot \boldsymbol{u}_\mathrm{s}} \tag{V67}$$

the vector in question is

$$\boldsymbol{q}_\mathrm{s} = \boldsymbol{\sigma}_{\mathbf{us}} \cdot (x - y),$$

i.e. the observer space $\mathtt{E}_{\boldsymbol{u}}$ is an affine space over $\mathbf{E}_\mathrm{s}$ with the subtraction

$$(x + \mathbb{I}\boldsymbol{u}) -_{\mathbf{E}_\mathrm{s}} (y + \mathbb{I}\boldsymbol{u}) := \boldsymbol{\sigma}_{\mathbf{us}} \cdot (x - y). \tag{V68}$$

Of course, the Euclidean structure of the observer does not depend on the synchronization. Thus, the distance between the $\boldsymbol{u}$-space points $x + \mathbb{I}\boldsymbol{u}$ and $y + \mathbb{I}\boldsymbol{u}$ is $|\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot (x - y)|$. Since $\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{\sigma}_{\mathbf{us}} = \boldsymbol{\sigma}_{\boldsymbol{u}}$, if the vector between the $\boldsymbol{u}$-space points is represented by $\boldsymbol{q}_\mathrm{s} = \boldsymbol{\sigma}_{\mathbf{us}} \cdot (x - y) \in \mathbf{E}_\mathrm{s}$ then

$$|\boldsymbol{q}_\mathrm{s}|^2_{\boldsymbol{u}} := |\boldsymbol{\sigma}_{\boldsymbol{u}} \cdot \boldsymbol{q}_s|^2 = |\boldsymbol{q}_\mathrm{s}|^2 + (\boldsymbol{u} \cdot \boldsymbol{q}_\mathrm{s})^2. \tag{V69}$$

Keep this formula in mind, not to commit an error: $|\boldsymbol{q}_\mathrm{s}|_{\boldsymbol{u}} \neq |\boldsymbol{q}_\mathrm{s}|$ if $\boldsymbol{u}_\mathrm{s} \neq \boldsymbol{u}$.

17.2 Splitting

The inertial frame $(\boldsymbol{u}, \mathbf{E}_{\rm s})$ decomposes the spacetime vectors in the sum of of two components, one is parallel to $\boldsymbol{u}$ and another one is in $\mathbf{E}_{\rm s}$. This decomposition of the vector $\boldsymbol{x}$ has the form

$$\boldsymbol{x} = \frac{-\boldsymbol{u}_{\rm s}\cdot\boldsymbol{x}}{-\boldsymbol{u}\cdot\boldsymbol{u}_{\rm s}}\boldsymbol{u} + \left(\boldsymbol{x} + \frac{\boldsymbol{u}_{\rm s}\cdot\boldsymbol{x}}{-\boldsymbol{u}\cdot\boldsymbol{u}_{\rm s}}\boldsymbol{u}\right) = (\boldsymbol{\tau}_{\boldsymbol{u}\mathrm{s}}\cdot\boldsymbol{x})\boldsymbol{u} + \boldsymbol{\sigma}_{\boldsymbol{u}\mathrm{s}}\cdot\boldsymbol{x}.$$

Only the coefficient of $\boldsymbol{u}$ is interesting in the first component, therefore we define the **splitting of spacetime vectors** according to $(\boldsymbol{u}, \mathbf{E}_{\rm s})$:

$$\boldsymbol{h}_{\boldsymbol{u}\mathrm{s}} := (\boldsymbol{\tau}_{\boldsymbol{u}\mathrm{s}}, \boldsymbol{\sigma}_{\boldsymbol{u}\mathrm{s}}) : \mathbf{M} \to \mathbb{I} \times \mathbf{E}_{\rm s}.$$

Accordingly, a covector $\boldsymbol{k}$ has the split form

$$(\boldsymbol{k}\cdot\boldsymbol{u},\ \boldsymbol{k}\cdot\boldsymbol{\sigma}_{\boldsymbol{u}_{\rm s}}) \in \mathbb{I}^* \times \mathbf{E}_{\rm s}^*$$

where $\boldsymbol{\sigma}_{\boldsymbol{u}_{\rm s}}$ is defined corresponding to (V30).

The covector $\boldsymbol{k}$ can be considered as a vector. We see that its vectorial split form is highly different from its covectorial split form, the relation between them is not so simple as in the case of standard synchronizations.

Let $(\mathbf{e}_\bullet, \mathbf{p}_\bullet)$ and $(\mathbf{e}^\bullet, \mathbf{p}^\bullet)$ denote the covectorial and the vectorial split form of a covector, respectively. A little manipulation with the above formulae gives us that

$$\mathbf{e}^\bullet = -\sqrt{1-v^2}\left(\sqrt{1-v^2}\ \mathbf{e}_\bullet - \mathbf{p}_\bullet\cdot\mathbf{v}\right), \qquad \mathbf{p}^\bullet = \mathbf{p}_\bullet + \left(\sqrt{1-v^2}\ \mathbf{e}_\bullet - \mathbf{p}_\bullet\cdot\mathbf{v}\right)\mathbf{v}$$

where $\boldsymbol{v} := \frac{\boldsymbol{u}}{-\boldsymbol{u}_{\rm s}\cdot\boldsymbol{u}} - \boldsymbol{u}_{\rm s}$ (the standard relative velocity of $\boldsymbol{u}$ with respect to $\boldsymbol{u}_{\rm s}$) and $v := |\boldsymbol{v}|$.

In coordinates corresponding to such a splitting this means that the covectorial coordinates $(\kappa_0, \kappa_1, \kappa_2, \kappa_3)$ and the vectorial coordinates $(\kappa^0, \kappa^1, \kappa^2, \kappa^3)$ of a covector are related in the following way:

$$\kappa^0 = -\sqrt{1-v^2}\left(\sqrt{1-v^2}\kappa_0 - \kappa_k v^k\right), \qquad \kappa^i = \kappa_i + \left(\sqrt{1-v^2}\kappa_0 - \kappa_k v^k\right)v^i$$

$$(i = 1, 2, 3).$$

17.3 Transformation rules

We want to compare the splitting of spacetime vectors due to the inertial frame $(\boldsymbol{u}, \mathbf{E}_{\rm s})$ with that due to $(\boldsymbol{u}', \mathbf{E}_{\rm s'})$. Of course, as in the case of standard frames, we have to make a boost from $\mathbf{E}_{\rm s'}$ into $\mathbf{E}_{\rm s}$. This boost is realized as follows:

In general, we get a rather complicated transformation rule.

Let us consider the special case when $\mathbf{E}_{\rm s'} = \mathbf{E}_{\rm s}$ and $\boldsymbol{u}' = \boldsymbol{u}_{\rm s}$ i.e. the inertial frames are $(\boldsymbol{u}, \mathbf{E}_{\boldsymbol{u}'})$ and $(\boldsymbol{u}', \mathbf{E}_{\boldsymbol{u}'})$ (the 'primed' frame is standard, and the 'unprimed' frame uses the synchronization of the 'primed' one). Then we need not apply a boost.

Let $(\boldsymbol{t}, \boldsymbol{q}) \in \mathbb{I} \times \mathbf{E}_{\boldsymbol{u}'}$ the $(\boldsymbol{u}, \mathbf{E}_{\boldsymbol{u}'})$-split form of a vector and $(\boldsymbol{t}', \boldsymbol{q}') \in \mathbb{I} \times \mathbf{E}_{\boldsymbol{u}'}$ the $(\boldsymbol{u}', \mathbf{E}_{\boldsymbol{u}'})$-split form – i.e the standard one – of the same vector. The the vector itself is $\boldsymbol{u}\boldsymbol{t} + \boldsymbol{q}$, so

$$\boldsymbol{t}' = \frac{-\boldsymbol{u}}{uu \cdot \boldsymbol{u}'} \cdot (\boldsymbol{u}\boldsymbol{t} + \boldsymbol{q}) = \frac{\boldsymbol{t}}{-\boldsymbol{u}' \cdot \boldsymbol{u}} = \boldsymbol{t}\sqrt{1 - v^2},$$

$$\boldsymbol{q}' = \left(\mathbf{1} + \frac{\boldsymbol{u}' \otimes \boldsymbol{u}}{-\boldsymbol{u}' \cdot \boldsymbol{u}}\right) \cdot (\boldsymbol{u}\boldsymbol{t} + \boldsymbol{q}) = (\boldsymbol{u} - (\boldsymbol{u}' \cdot \boldsymbol{u})\boldsymbol{u}')\boldsymbol{t} + \boldsymbol{q} = -\boldsymbol{v}\boldsymbol{t} + \boldsymbol{q}$$

where $\boldsymbol{v} := \frac{\boldsymbol{u}'}{-\boldsymbol{u}' \cdot \boldsymbol{u}} - \boldsymbol{u}$ (the standard relative velocity of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$) and $v := |\boldsymbol{v}|$.

As a means of denying the theory of relativity, such and more complicated transformation rules are suggested in the literature instead of the Lorentz transformation rule[9]. Of course, there everything is formulated in coordinates and it is quite impossible to be aware that relativity theory is in question in some (which?) coordinates.

As we see, such transformation rules have their place in the relativistic spacetime model, so they cannot deny it.

17.4 Relative velocities

The motion of the inertial material point with absolute velocity $\boldsymbol{u}'$ with respect to the inertial frame $(\boldsymbol{u}, \mathbf{E}_{\rm s})$ is uniform i.e. has constant relative velocity.

Let x and y be two occurrences of the material point i.e. $y - x = \boldsymbol{s}\boldsymbol{u}'$ for some $\boldsymbol{s} \in I$. At the $\mathbf{E}_{\rm s}$-intants $x + \mathbf{E}_{\rm s}$ and $y + \mathbf{E}_{\rm s}$, the material points meets the $\boldsymbol{u}$-space points $x + \mathbb{I}\boldsymbol{u}$ and $y + \mathbb{I}\boldsymbol{u}$, respectively; the $\boldsymbol{u}$-space vector betwen them is $(y + \mathbb{I}\boldsymbol{u}) -_{\boldsymbol{u}_{\rm s}} (x + \mathbb{I}\boldsymbol{u}) = \boldsymbol{\sigma}_{\mathbf{us}} \cdot (y - x) = \boldsymbol{s}(\boldsymbol{\sigma}_{\mathbf{us}} \cdot \boldsymbol{u}')$. The $\mathbf{E}_{\rm s}$ duration between

[9]S.Marinov, *General Relativity and Gravitation*, **12**(1980)57; F. Goy - F. Selleri,*Foundations of Physics Letters* **10**(1997)17

the two meetings is $(y+\mathbf{E}_s) -_{\boldsymbol{u}} (x+\mathbf{E}_s) = \boldsymbol{\tau}_{\boldsymbol{u}\mathrm{s}}\cdot(y-x) = \boldsymbol{s}(\boldsymbol{\tau}_{\boldsymbol{u}\mathrm{s}}\cdot\boldsymbol{u}')$. Therefore, the relative velocity is

$$\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u},\boldsymbol{u}_s} := \frac{(y+\mathbb{I}\boldsymbol{u}) -_{\boldsymbol{u}_s} (x+\mathbb{I}\boldsymbol{u})}{(y+\mathbf{E}_s) -_{\boldsymbol{u}} (x+\mathbf{E}_s)} = \frac{\boldsymbol{\sigma}_{\boldsymbol{u}\mathrm{s}}\cdot\boldsymbol{u}'}{\boldsymbol{\tau}_{\boldsymbol{u}\mathrm{s}}\cdot\boldsymbol{u}'} = \frac{(-\boldsymbol{u}_s\cdot\boldsymbol{u})\boldsymbol{u}'}{-\boldsymbol{u}_s\cdot\boldsymbol{u}'} - \boldsymbol{u},$$

for which

$$|\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u},\boldsymbol{u}_s}|_{\boldsymbol{u}} = |\boldsymbol{\sigma}_{\boldsymbol{u}}\cdot\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u},\boldsymbol{u}_s}| = |\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}|\frac{(-\boldsymbol{u}'\cdot\boldsymbol{u})(-\boldsymbol{u}_s\cdot\boldsymbol{u})}{-\boldsymbol{u}_s\cdot\boldsymbol{u}'}$$

holds where $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}$ the standard relative velocity of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$.

Similar formulae hold for light signals, i.e. for a light direction $\boldsymbol{w}$ instead of the absolute velocity $\boldsymbol{u}'$. We can see better the physical content of such a formula if we write

$$\boldsymbol{u}_s = \frac{\boldsymbol{u}+v_s\boldsymbol{n}_s}{\sqrt{1-v_s^2}}, \qquad \boldsymbol{w} = \boldsymbol{u}+\boldsymbol{n}_{\boldsymbol{w}},$$

where $\boldsymbol{n}_s, \boldsymbol{n}_{\boldsymbol{w}} \in \frac{\mathbf{E}_{\boldsymbol{u}}}{\mathbb{I}}$ are unit vectors. $\boldsymbol{n}_s$ is a direction in $\boldsymbol{u}$-space, characterizing the synchronization, more closely, $v_s := |\mathbf{v}_{\boldsymbol{u}_s,\boldsymbol{u}}|$, $\boldsymbol{n}_s = \frac{\mathbf{v}_{\boldsymbol{u}_s,\boldsymbol{u}}}{v_s}$. $\boldsymbol{n}_{\boldsymbol{w}}$ is the direction of the light signal in $\boldsymbol{u}$-space. Then

$$|\mathbf{v}_{\boldsymbol{w}\boldsymbol{u},\boldsymbol{u}_s}|_{\boldsymbol{u}} = \frac{1-v_s\boldsymbol{n}_s\cdot\boldsymbol{n}_{\boldsymbol{w}}}{1-v_s^2}.$$

The light speed in the $\boldsymbol{u}$-space depends on the direction of light propagation (except, of course, the trivial case $v_s = 0$ which corresponds to the standard synchronization). The minimal light speed is $\frac{1}{1+v_s}$ – when $\boldsymbol{n}_{\boldsymbol{w}} = \boldsymbol{n}_s$ – and the maximal light speed is $\frac{1}{1-v_s}$ – when $\boldsymbol{n}_{\boldsymbol{w}} = -\boldsymbol{n}_s$.

17.5 Comparison of lengths

Let us take the $\mathbf{E}_s$-instantaneous print of the $\boldsymbol{u}'$-space vectors in the $\boldsymbol{u}$-space.

The print of the vector $\boldsymbol{q}' \in \mathbf{E}_{\boldsymbol{u}'}$ will be the vector $\boldsymbol{q}_s \in \mathbf{E}_s$ for which $\boldsymbol{q}'-\boldsymbol{q}_s$ is parallel to $\boldsymbol{u}'$; this means that $\boldsymbol{q}_s$ is the projection of $\boldsymbol{q}'$ onto $\mathbf{E}_s$ along $\boldsymbol{u}'$. Applying V47 according to the sense, we get

$$\boldsymbol{q}_s = \left(\mathbf{1}+\frac{\boldsymbol{u}'\otimes\boldsymbol{u}_s}{-\boldsymbol{u}'\cdot\boldsymbol{u}_s}\right)\boldsymbol{q}'.,$$

Then V69 gives

$$|\boldsymbol{q}_s|_{\boldsymbol{u}}^2 = \left|(\mathbf{1}+\boldsymbol{u}\otimes\boldsymbol{u})\left(\mathbf{1}+\frac{\boldsymbol{u}'\otimes\boldsymbol{u}_s}{-\boldsymbol{u}'\cdot\boldsymbol{u}_s}\right)\boldsymbol{q}'\right|^2.$$

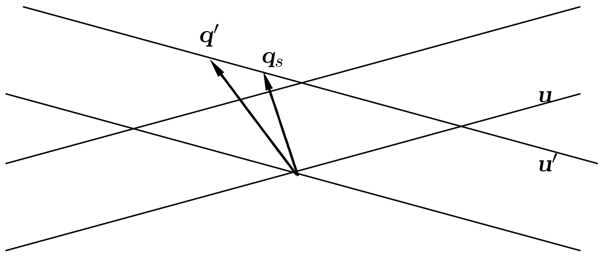

Figure 17.2 Non-standard print

This is rather complicated in general. Let us consider the special cases when the standard relative velocity of $\boldsymbol{u}'$ with respect to $\boldsymbol{u}$ is parallel to $\boldsymbol{n}_\mathrm{s}$ i.e. $\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}} = v\boldsymbol{n}_\mathrm{s}$; equivalently, $\boldsymbol{u}' = \frac{\boldsymbol{u}+v\boldsymbol{n}_\mathrm{s}}{\sqrt{1-v^2}}$ ($v = |\mathbf{v}_{\boldsymbol{u}'\boldsymbol{u}}| = |\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}|$). Then

$$(\mathbf{1} + \boldsymbol{u}\otimes\boldsymbol{u})\left(\mathbf{1} + \frac{\boldsymbol{u}'\otimes\boldsymbol{u}_\mathrm{s}}{-\boldsymbol{u}'\cdot\boldsymbol{u}_\mathrm{s}}\right) = \mathbf{1} + \boldsymbol{u}\otimes\boldsymbol{u} + \frac{v\boldsymbol{n}_\mathrm{s}\otimes\boldsymbol{u}_\mathrm{s}}{1-v_\mathrm{s}v}$$

and $\boldsymbol{u}'\cdot\boldsymbol{q}' = 0$ imply $\boldsymbol{n}_\mathrm{s}\cdot\boldsymbol{q}' = \frac{-\boldsymbol{u}\cdot\boldsymbol{q}'}{v}$, so $\boldsymbol{u}_\mathrm{s}\cdot\boldsymbol{q}' = \left(1-\frac{v_\mathrm{s}}{v}\right)\boldsymbol{u}\cdot\boldsymbol{q}'$; further, $\boldsymbol{u}\cdot\boldsymbol{q}' = \frac{\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}\cdot\boldsymbol{q}'}{\sqrt{1-v^2}}$. Consequently,

$$|\boldsymbol{q}'|^2 - (\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}\cdot\boldsymbol{q}')^2\frac{2\beta-1-\beta^2v^2}{1-v^2}$$

where

$$\beta := \frac{(1-\frac{v_\mathrm{s}}{v})}{1-v_\mathrm{s}v}.$$

In particular, if $v_\mathrm{s} = 0$ (i.e. $\boldsymbol{u}_\mathrm{s} = \boldsymbol{u}$) then $\beta = 1$ and we have the well known Lorentz contraction corresponding to the standard synchronization.

Another special case is $v_\mathrm{s} = v$ (i.e.$\boldsymbol{u}_\mathrm{s} = \boldsymbol{u}'$) when $\beta = 0$, so

$$|\boldsymbol{q}'|^2 + \frac{(\mathbf{v}_{\boldsymbol{u}\boldsymbol{u}'}\cdot\boldsymbol{q}')^2}{1-v^2};$$

the print is longer than the vector.

At last, a third special case is $v_\mathrm{s} = \frac{1-\sqrt{1-v^2}}{v}$ when $2\beta - 1 - \beta^2v^2 = 0$; then the length of the print equals the length of the vector.

These actual formulae show well that the **Lorentz contraction is not a physical relity**: it is eventual, artificial illusion in a specific synchronization. Other synchronization can result in a length dilation.

17.6 Comparison of time periods

Let t and s be two $\mathbf{E}_\mathrm{s}$-instants. According to the observer $\boldsymbol{u}$, time passed between t and s is

$$\boldsymbol{t} = \boldsymbol{\tau}_{\mathbf{us}} \cdot (x-y) = \frac{-\boldsymbol{u}_\mathrm{s} \cdot (x-y)}{-\boldsymbol{u} \cdot \boldsymbol{u}_\mathrm{s}}$$

where $x \in s$, $y \in t$. Similarly, the time interval according to the observer $\boldsymbol{u}'$ is

$$\boldsymbol{t}' = \frac{-\boldsymbol{u}_\mathrm{s} \cdot (x-y)}{-\boldsymbol{u}' \cdot \boldsymbol{u}_\mathrm{s}}$$

which gives

$$\boldsymbol{t} = \frac{-\boldsymbol{u}' \cdot \boldsymbol{u}_\mathrm{s}}{-\boldsymbol{u} \cdot \boldsymbol{u}_\mathrm{s}} \boldsymbol{t}'.$$

Let us take again $\boldsymbol{u}' := \frac{\boldsymbol{u}+v\boldsymbol{n}_\mathrm{s}}{\sqrt{1-v^2}}$. Then

$$\boldsymbol{t} = \frac{1 - v_\mathrm{s} v}{\sqrt{1-v^2}} \boldsymbol{t}'.$$

In particular, if $v_\mathrm{s} = 0$ then we get the time dilation corresponding to the standard synchronization.

Another special case is $v_\mathrm{s} = v$ when

$$\boldsymbol{t} = \sqrt{1-v^2}\,\boldsymbol{t}';$$

the synchronizationtime interval is shorter than the proper time interval.

At last, a third special case is $v_\mathrm{s} = \frac{1-\sqrt{1-v^2}}{v}$; then the two time intervals are equal.

These actual formulae show well that the **Lorentz time dilation is not a physical reality**: it is an eventual, artificial illusion in a specific synchronization. Other synchronization can result in a time contraction.

18 Some remarks

The usual treatments of relativity in coordinates can mislead one in several aspects. For instance, it gets lost that

– the space vectors of different observers are different (contrary to the nonrelativistic case),

– the equality of vectors (parallelism of straight lines) in different observer spaces is not obvious; a consequence is the velocoty addition paradox.

Further, coordinates always refer to standard synchronizations, thus,

– it seems as if the standard synchronization were a necessity; consequently, some statements, e.g. Lorentz contraction and time dilation, seem to express real facts,

– the importance of synchronizations is neglected in the definition of relative velocity; a consequence is the light propagation paradox.

It is stated frequently (almost always) that special relativity is the theory of inertial observers, while general observers are objects of general relativity. This is erroneous as it was emphasized even fifty years ago [10], and as it is apparent from our treatment in which observers (and reference frames), instead of being fundamental intuitive building blocks of the theory, are notions derived in the framework of the theory. Noninertial observers, too, can be well treated in the special relativistic spacetime model.

What is true is the following: the treatment of noninertial reference frames in the special relativistic spacetime model requires the same higher mathematical tools as the treatment of general relativistic spacetime models.

General relativistic spacetime models are models of gravitational actions and the special relativistic spacetime model is the model of the lack of gravitation.

[10] J.L.Synge: *Relativity: The Special Theory*, North-Holland, 1955; *Relativity: The General Theory*, North-Holland,1964

VI Mathematical tools

The fundamental notions and notations of set theory – subset, intersection, union, Cartesian product, function, composition of functions etc. – are supposed to be known.

The fundamental notions of vector spaces – linear combination, basis, dimension, linear subspaces, linear maps, etc. – are supposed to be known.

The vector spaces in this book are real and finite dimensional.

It is known that all norms on a finite dimensional vector space are equivalent, thus the notions of analysis are meaningful without specifying an actual norm.

The fundamental notions of analysis – open sets, closed sets, closure of a set, convergence of a sequence, continuity of a function etc. – are supposed to be known.

We list other necessary mathematical notions and results without proof; further knowlegdge and detailed proofs can be found in Spacetime without Reference Frames.

19 Vector spaces

As mentioned, the fundamental notions of vector spaces – linear combination, basis, dimension, linear subspaces, linear maps, etc. – are supposed to be known.

The vector spaces in this book are **real and finite-dimensional**.

The action of a linear map will be denoted by a dot: $\boldsymbol{L}\cdot\boldsymbol{v}$ is the value of the linear map $\boldsymbol{L}$ at the vector $\boldsymbol{v}$. Similarly, the composition of linear maps will be denoted by a dot: $\boldsymbol{L}\cdot\boldsymbol{K}$ is the composition ('product') of the linear maps $\boldsymbol{L}$ and $\boldsymbol{K}$.

We often use the following ensemble operations: if $\boldsymbol{v}$ is a vector, α is a real number and A, B are subsets of a vector space, then

$$\boldsymbol{v} + \mathrm{A} := \{\boldsymbol{v}+\boldsymbol{x} \mid \boldsymbol{x}\in\mathrm{A}\}, \qquad \alpha\mathrm{A} := \{\alpha\boldsymbol{x} \mid \boldsymbol{x}\in\mathrm{A}\},$$

$$\mathrm{A} + \mathrm{B} := \{\boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{x} \in \mathrm{A}, \boldsymbol{y} \in \mathrm{B}\}.$$

19.1 Complementary subspaces

The linear subspaces $\mathbf{E}$ and $\mathbf{F}$ of the vector space $\mathbf{V}$ are **transversal** if $\mathbf{E} \cap \mathbf{F} = \mathbf{0}$.

The linear subspaces $\mathbf{E}$ and $\mathbf{F}$ of the vector space $\mathbf{V}$ are **complementary** if $\mathbf{E} \cap \mathbf{F} = \mathbf{0}$ and $\dim \mathbf{E} + \dim \mathbf{F} = \dim \mathbf{V}$. Then $\mathbf{E} + \mathbf{F} = \mathbf{V}$.

If $\mathbf{E}$ and $\mathbf{F}$ are complementary, then for all $\boldsymbol{v} \in \mathbf{V}$ there are unique elemets $\mathbf{v}_{\mathbf{E}}$ in $\mathbf{E}$ and $\mathbf{v}_{\mathbf{F}}$ in $\boldsymbol{F}$ such that $\boldsymbol{v} = \mathbf{v}_{\mathbf{E}} + \mathbf{v}_{\boldsymbol{F}}$.

Then the mapping $\boldsymbol{v} \mapsto \mathbf{v}_{\mathbf{E}}$ is linear and is called the **projection onto E along F**.

19.2 Factor spaces

Let $\mathbf{E}$ be a linear subspace of the vector space $\mathbf{V}$.

A subset of $\mathbf{V}$ is an **affine subspace directed by E** if it is of the form $\boldsymbol{v} + \mathbf{E}$ for some $\boldsymbol{v} \in \mathbf{V}$. Note that $\boldsymbol{v} + \mathbf{E} = \boldsymbol{u} + \mathbf{E}$ if and only if $\boldsymbol{v} - \boldsymbol{u} \in \mathbf{E}$.

An affine subspace directed by a one dimensional linear subspace is called a **straigth line** and an affine subspace directed by a linear subspace of co-dimension one is called a **hyperplane**. Two affine subspaces directed by the same linear subspace are called **parallel**.

The set of affine subspaces directed by $\mathbf{E}$ is denoted by $\mathbf{V}/\mathbf{E}$.

Endowed with the following well defined addition and multiplication by real numbers,

$$(\boldsymbol{v} + \mathbf{E}) + (\boldsymbol{u} + \mathbf{E}) := (\boldsymbol{v} + \boldsymbol{u}) + \mathbf{E}, \qquad \alpha(\boldsymbol{v} + \mathbf{E}) := (\alpha \boldsymbol{v}) + \mathbf{E},$$

$\mathbf{V}/\mathbf{E}$ becomes a vector space.

$\dim(\mathbf{V}/\mathbf{E}) = \dim \mathbf{V} - \dim \mathbf{E}$.

If $\boldsymbol{F}$ is a linear subspace complementary to $\mathbf{E}$, then every affine subspace directed by $\mathbf{E}$ contains exactly one element of $\boldsymbol{F}$. More closely, $\boldsymbol{F} \to \mathbf{V}/\mathbf{E}$, $\boldsymbol{v} \mapsto \boldsymbol{v} + \mathbf{E}$ is a linear bijection. This linear bijection allows us to represent the elements of $\mathbf{V}/\mathbf{E}$ by the elements of $\boldsymbol{F}$.

19.3 Orientation

Two ordered bases $(\mathbf{v}_1, \ldots, \mathbf{v}_N)$ and $(\mathbf{v}'_1, \ldots, \mathbf{v}'_N)$ of $\mathbf{V}$ are called **equally oriented** if the linear map defined by $\mathbf{v}_i \mapsto \mathbf{v}'_i \ (i = 1, \ldots, N)$ has positive determinant. An equivalence class of equally oriented bases is called an **orientation** of $\mathbf{V}$. $\mathbf{V}$ is **oriented** if an orientation of $\mathbf{V}$ is given; the bases in the chosen equivalence class are called **positively oriented,** the other ones are called **negatively ortiented**.

A linear bijection between oriented vector spaces is **orientation-preserving** (resp. **orientation-reversing**) if it sends positively oriented bases into positively (rsp. negatively) oriented ones.

Let $\mathbf{A}$ be a one-dimensional vector space.

Two bases $\boldsymbol{a}$ and $\boldsymbol{a}'$ of $\mathbf{A}$ are equally oriented if and only if $\boldsymbol{a}'$ is a positive multiple of $\boldsymbol{a}$. Thus the equally oriented bases form a 'half line'.

A non-zero element $\boldsymbol{a}$ of the oriented one-dimensional vector space $\mathbf{A}$ is called **positive,** in symbols $\mathbf{0} < \boldsymbol{a}$, if the corresponding basis is positively oriented.

Moreover, we write $\boldsymbol{a} \leq \boldsymbol{b}$ if $\mathbf{0} \leq \boldsymbol{b} - \boldsymbol{a}$. It is easily shown that in this way a total ordering is defined on $\mathbf{A}$ for which

– if $\boldsymbol{a} \leq \boldsymbol{b}$ and $\boldsymbol{c} \leq \mathbf{d}$, then $\boldsymbol{a} + \boldsymbol{c} \leq \boldsymbol{b} + \mathbf{d}$,

– if $\boldsymbol{a} \leq \boldsymbol{b}$ and α is a positive real number, then $\alpha\boldsymbol{a} \leq \alpha\boldsymbol{b}$.

We introduce the notations

$$\mathbf{A}^+ := \{\boldsymbol{a} \in \mathbf{A} \mid \mathbf{0} < \boldsymbol{a}\}, \qquad \mathbf{A}^+_{\mathbf{0}} := \mathbf{A}^+ \cup \{\mathbf{0}\}.$$

Furthermore, the absolute value of $\boldsymbol{a} \in \mathbf{A}$ is

$$|\boldsymbol{a}| := \begin{cases} \boldsymbol{a} & \text{if } \boldsymbol{a} \in \mathbf{A}^+ \\ \mathbf{0} & \text{if } \boldsymbol{a} = \mathbf{0} \\ -\boldsymbol{a} & \text{if } \boldsymbol{a} \notin \mathbf{A}^+. \end{cases}$$

If $\mathbf{A}$ is oriented, an orientation of $\mathbf{A} \otimes \mathbf{A}$ is induced by the orientation of $\mathbf{A}$ and then

$$\mathbf{A}^+_{\mathbf{0}} \to (\mathbf{A} \otimes \mathbf{A})^+_{\mathbf{0}}, \qquad \boldsymbol{a} \mapsto \boldsymbol{a}^2 := \boldsymbol{a} \otimes \boldsymbol{a}$$

is a bijection whose inverse is the square root mapping $\sqrt{\ } : (\mathbf{A} \otimes \mathbf{A})^+_{\mathbf{0}} \to \mathbf{A}^+_{\mathbf{0}}$.

19.4 Vectors, covectors

Let $\mathbf{V}$ be a vector space; its elements are called vectors. A linear map $\mathbf{V} \to \mathbb{R}$ is called a **covector**. The set of covectors is the **dual** of $\mathbf{V}$ and is denoted by $\mathbf{V}^*$. $\mathbf{V}^*$ is a vector space whose dimension equals the dimension of $\mathbf{V}$.

The dual of $\mathbf{V}^*$ can be identified with $\mathbf{V}$ (in other words, the co-covectors are vectors, $\mathbf{V}^{**} \equiv \mathbf{V}$) by considering the vector $\boldsymbol{v}$ as the linear map $\mathbf{V}^* \to \mathbb{R}$, $\boldsymbol{f} \mapsto \boldsymbol{f} \cdot \boldsymbol{v}$. Accordingly, we find it convenient to write $\boldsymbol{v} \cdot \boldsymbol{f} = \boldsymbol{f} \cdot \boldsymbol{v}$.

19.5 Transposes

The **transpose** of a linear map $\boldsymbol{L} : \mathbf{V} \to \mathbf{U}$ is the linear map

$$\boldsymbol{L}^* : \mathbf{U}^* \to \mathbf{V}^*, \qquad \boldsymbol{f} \mapsto \boldsymbol{f} \circ \boldsymbol{L},$$

i.e.

$$(\boldsymbol{L}^* \cdot \boldsymbol{f}) \cdot \boldsymbol{v} = \boldsymbol{f} \cdot (\boldsymbol{L} \cdot \boldsymbol{v})$$

or, with the above identification,

$$\boldsymbol{v} \cdot \boldsymbol{L}^* \cdot \boldsymbol{f} = \boldsymbol{f} \cdot \boldsymbol{L} \cdot \boldsymbol{v} \qquad (\boldsymbol{f} \in \mathbf{U}^*,\ \boldsymbol{v} \in \mathbf{V}).$$

If $\boldsymbol{L}, \boldsymbol{K} : \mathbf{V} \to \mathbf{U}$ are linear maps and α is a real number, then

$$(\boldsymbol{L} + \boldsymbol{K})^* = \boldsymbol{L}^* + \boldsymbol{K}^*, \qquad (\alpha \boldsymbol{L})^* = \alpha \boldsymbol{L}^*.$$

If $\boldsymbol{L} : \mathbf{V} \to \mathbf{U}$ and $\boldsymbol{K} : \mathbf{U} \to \mathbf{W}$ are linear maps then

$$(\boldsymbol{K} \cdot \boldsymbol{L})^* = \boldsymbol{L}^* \cdot \boldsymbol{K}^*.$$

Moreover – because of the identifications $\mathbf{V}^{**} \equiv \mathbf{V}$, $\mathbf{U}^{**} \equiv \mathbf{U}$ – we have

$$\boldsymbol{L}^{**} = \boldsymbol{L}.$$

Furthermore,
- $\boldsymbol{L}$ is injective if and only if $\boldsymbol{L}^*$ is surjective,
- $\boldsymbol{L}$ is surjective if and only if $\boldsymbol{L}^*$ is injective.

If $\boldsymbol{L}$ is bijective, then

$$\left(\boldsymbol{L}^{-1}\right)^* = (\boldsymbol{L}^*)^{-1}.$$

19.6 Cotensors, tensors

A bilinear map $\mathbf{V} \times \mathbf{V} \to \mathbb{R}$ is called a **cotensor**. The set of cotensors is is denoted by $\mathbf{V}^* \otimes \mathbf{V}^*$. $\mathbf{V}^* \otimes \mathbf{V}^*$ is a vector space whose dimension equals the square of the dimension of $\mathbf{V}$.

The subspace of antisymmetric cotensors is denoted by $\mathbf{V}^* \wedge \mathbf{V}^*$. The **antisymmetrization** of the cotensor $\boldsymbol{F}$ is the bilinear map $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}) - \boldsymbol{F}(\boldsymbol{y}, \boldsymbol{x})$.

A cotensor $\boldsymbol{F}$ can be considered as a linear map $\mathbf{V} \to \mathbf{V}^*$, $\boldsymbol{v} \mapsto \boldsymbol{F} \cdot \boldsymbol{v} := \boldsymbol{F}(\cdot, \boldsymbol{v})$.

For covectors $\boldsymbol{q}$ and $\boldsymbol{p}$ we define the cotensor $\boldsymbol{q} \otimes \boldsymbol{p}$ as the bilinear map $(\boldsymbol{u}, \boldsymbol{v}) \mapsto (\boldsymbol{q} \cdot \boldsymbol{u})(\boldsymbol{p} \cdot \boldsymbol{v})$. Its antisymmetrization is denoted by $\boldsymbol{q} \wedge \boldsymbol{p}$. Considered as a linear map, $\boldsymbol{q} \otimes \boldsymbol{p}$ becomes $\boldsymbol{v} \mapsto \boldsymbol{q}(\boldsymbol{p} \cdot \boldsymbol{v})$.

Interchanging the roles of $\mathbf{V}$ and $\mathbf{V}^*$ – because of the identification $\mathbf{V}^{**} = \mathbf{V}$ – , we get the **tensors** as the bilinear maps $\mathbf{V}^* \times \mathbf{V}^* \to \mathbb{R}$, whose set is denoted by $\mathbf{V} \otimes \mathbf{V}$.

Then we define antisymmetric tensors and antsymmetrization of tensors as previously. A tensor $\mathbf{T}$ can be considered as a linear map $\mathbf{V}^* \to \mathbf{V}$, $\boldsymbol{q} \mapsto \mathbf{T} \cdot \boldsymbol{q} := \mathbf{T}(\cdot, \boldsymbol{q})$.

Mixed tensors are bilinear maps $\mathbf{V} \times \mathbf{V}^* \to \mathbb{R}$ or $\mathbf{V}^* \times \mathbf{V} \to \mathbb{R}$ whose set is denoted by $\mathbf{V}^* \otimes \mathbf{V}$ and $\mathbf{V} \otimes \mathbf{V}^*$, respectively. Of course, there is no antisymmetric mixed tensor.

More generally, for a natural number n, the n-cotensors are n-linear maps $\mathbf{V}^n \to \mathbb{R}$. In this context, the 0-cotensors are defined to be the real numbers, the 1-cotensors are the covectors, the 2-cotensors are the cotensors defined previously. The n-tensors and mixed n-tensors are defined similarly.

19.7 Coordinates

An ordered basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_N$ of the vector space $\mathbf{V}$ establishes a **coordinatization** of $\mathbf{V}$, i.e. allows us to represent an arbitrary vector $\boldsymbol{v}$ by an N-tuple of real numbers, denoted by v^i where the index i runs from 1 to N, in such a way that $\boldsymbol{v} = \sum_{i=1}^{N} v^i \boldsymbol{e}_i$.

Then $\mathbf{V}^*$ is coordinatized as well; the covector $\boldsymbol{p}$ is represented by the N-tuple of real numbers $p_i := \boldsymbol{p} \cdot \boldsymbol{e}_i$ where the index i runs from 1 to N.

In this way we have that $\boldsymbol{p} \cdot \boldsymbol{v} = \sum_{i=1}^{N} p_i v^i$; to make the notations simpler, the Einstein summation rule will be accepted: the symbol $\sum$ is omitted and an automatic summation has to be carried out for equal lower and upper indices: $\boldsymbol{p} \cdot \boldsymbol{v} = p_i v^i$.

The basis allows us to represent linear maps by matrices; in particular,

– a linear map $\boldsymbol{L} : \mathbf{V} \to \mathbf{V}$ is represented by L^i_k in such a way that the coordinates of $\boldsymbol{L} \cdot \boldsymbol{v}$ are $L^i_k v^k$ (Einstein summation!),

– a linear map $\boldsymbol{F} : \mathbf{V} \to \mathbf{V}^*$ is represented by F_{ik} in such a way that the coordinates of $\boldsymbol{F} \cdot \boldsymbol{v}$ are $F_{ik} v^k$ (Einstein summation!).

Then bilinear maps are represented by matrices as well. In particular, the matrix cooresponding to the cotensor $\boldsymbol{F} : \mathbf{V} \times \mathbf{V} \to \mathbb{R}$ will be $F_{ik} := \boldsymbol{F}(\boldsymbol{e}_i, \boldsymbol{e}_k)$.

The matrix forms reflect that tensors, cotensors, mixed tensors can be identified with various linear maps.

20 Tensorial operations

20.1 Tensor products

As a generalization of the previous notions, taking two vector spaces $\mathbf{V}$ and $\mathbf{U}$, we find it convenient to use the symbol $\mathbf{U}^* \otimes \mathbf{V}^*$ for the set of bilinear maps $\mathbf{U} \times \mathbf{V} \to \mathbb{R}$, which is identified with the set of linear maps $\mathbf{V} \to \mathbf{U}^*$: the bilinear map $\boldsymbol{F} : \mathbf{U} \times \mathbf{V} \to \mathbb{R}$ is considered as a linear map $\mathbf{V} \to \mathbf{U}^*$, $\boldsymbol{v} \mapsto \boldsymbol{F} \cdot \boldsymbol{v} := \boldsymbol{F}(\cdot, \boldsymbol{v})$.

The tensor product $\boldsymbol{q} \otimes \boldsymbol{p}$ of $\boldsymbol{q} \in \mathbf{U}^*$ and $\boldsymbol{p} \in \mathbf{V}^*$ is the bilinear map

$$\mathbf{U} \times \mathbf{V} \to \mathbb{R}, \qquad (\boldsymbol{u}, \boldsymbol{v}) \mapsto (\boldsymbol{q} \cdot \boldsymbol{u})(\boldsymbol{p} \cdot \boldsymbol{v}),$$

or the linear map

$$\mathbf{V} \to \mathbf{U}^*, \qquad \boldsymbol{v} \mapsto \boldsymbol{q}(\boldsymbol{p} \cdot \boldsymbol{v}).$$

Interchanging the roles of vectors and covectors, we write $\mathbf{U} \otimes \mathbf{V}$ for the set of bilinear maps $\mathbf{U}^* \times \mathbf{V}^*$, which is identified with the set of linear maps $\mathbf{V}^* \to \mathbf{U}$, and we define $\boldsymbol{u} \otimes \boldsymbol{v}$ by a formula similar to the previous one.

Furthermore, we write $\mathbf{U} \otimes \mathbf{V}^*$ for the set of bilinear maps $\mathbf{U}^* \times \mathbf{V}$ ... etc.

If $\boldsymbol{u} \otimes \boldsymbol{v}$ is considered as a linear map $\mathbf{V}^* \to \mathbf{U}$, then its transpose is a linear map $\mathbf{U}^* \to \mathbf{V}$; one finds that $(\boldsymbol{u} \otimes \boldsymbol{v})* = \boldsymbol{v} \otimes \boldsymbol{u}$.

If one of the vector spaces is one-dimensional, then we identify a linear map with its transpose, i.e. we consider the tensor product commutative:

$$\mathbf{A} \otimes \mathbf{V} \equiv \mathbf{V} \otimes \mathbf{A}, \quad \boldsymbol{a} \otimes \boldsymbol{v} \equiv \boldsymbol{v} \otimes \boldsymbol{a} \qquad \text{if} \quad \dim \mathbf{A} = 1;$$

moreover, in this case we omit the symbol $\otimes$ between the vectors:

$$\boldsymbol{a}\boldsymbol{v} := \boldsymbol{a} \otimes \boldsymbol{v}.$$

If $\dim \mathbf{A} = 1$, then every element of $\mathbf{A} \otimes \mathbf{V}$ has the form $\boldsymbol{a}\boldsymbol{v}$.

Though, in general, $\mathbf{A} \otimes \mathbf{V} \neq \mathbf{V}$, it makes sense (if $\dim \mathbf{A} = 1$) that an element $\boldsymbol{z}$ of $\mathbf{A} \otimes \mathbf{V}$ is **parallel** to an element $\boldsymbol{v}$ of $\mathbf{V}$: if there is an $\boldsymbol{a} \in \mathbf{A}$ such that $\boldsymbol{z} = \boldsymbol{a}\boldsymbol{v}$.

20.2 Tensor quotients

Let $\mathbf{A}$ be a one-dimensional vector space.

If $\boldsymbol{a}$ is a non-zero element of $\mathbf{A}$, then for every $\boldsymbol{b} \in \mathbf{A}$ there is a unique real number, denoted by $\frac{\boldsymbol{b}}{\mathbf{a}}$ such that $\frac{\boldsymbol{b}}{\mathbf{a}}\boldsymbol{a} = \boldsymbol{b}$.

If $\boldsymbol{v}$ is a vector in $\mathbf{V}$ and $\boldsymbol{a} \neq 0$ in $\mathbf{A}$, then we define the linear map

$$\frac{\boldsymbol{v}}{\boldsymbol{a}} : \mathbf{A} \to \mathbf{V}, \qquad \boldsymbol{b} \mapsto \frac{\boldsymbol{b}}{\boldsymbol{a}}\boldsymbol{v},$$

called the **tensor quotient of $\boldsymbol{v}$ by $\boldsymbol{a}$**. Every linear map $\mathbf{A} \to \mathbf{V}$ is of this form, thus we can use the symbol $\frac{\mathbf{V}}{\mathbf{A}}$ for the set of such linear maps.

Since the linear maps $\mathbf{A} \to \mathbf{A}$ are the multiplications by real numbers, $\frac{\mathbf{A}}{\mathbf{A}} = \mathbb{R}$, and $\frac{\boldsymbol{b}}{\mathbf{a}}$ as a tensor quotient equals the real number introduced at the beginning of the subsection.

Though, in general, $\frac{\mathbf{V}}{\mathbf{A}} \neq \mathbf{V}$, it makes sense that an element $\boldsymbol{z}$ of $\frac{\mathbf{V}}{\mathbf{A}}$ is **parallel** to an element $\boldsymbol{v}$ of $\mathbf{V}$: if there is an $\mathbf{0} \neq \boldsymbol{a} \in \mathbf{A}$ such that $\boldsymbol{z} = \frac{\mathbf{v}}{\mathbf{a}}$.

20.3 Tensorial identifications

Let us consider $\mathbf{U} \otimes \mathbf{V}$. We easily find that $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \boldsymbol{u} \otimes \boldsymbol{v}$ is a bilinear map from $\mathbf{U} \times \mathbf{V}$ into $\mathbf{U} \otimes \mathbf{V}$ and

(i) every element of $\mathbf{U} \otimes \mathbf{V}$ is of the form $\sum\limits_{i=1}^{n} \boldsymbol{u}_i \otimes \boldsymbol{v}_i$ for some natural number n and vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in \mathbf{U}$, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathbf{V}$,

(ii) if $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are linearly independent and $\sum\limits_{i=1}^{n} \boldsymbol{u}_i \otimes \boldsymbol{v}_i = 0$, then $\boldsymbol{u}_i = 0$ for all $i = 1, \ldots, n$.

In particular, we have that $\boldsymbol{u} \otimes \boldsymbol{v} = 0$ if and only if $\boldsymbol{u} = 0$ or $\boldsymbol{v} = 0$.

The tensor product of $\mathbf{U}$ and $\mathbf{V}$, as it has been defined, can be considered either as a space of bilinear maps or as a space of linear maps, i.e. the tensor product is not unique. More generally, we consider the **abstract tensor product $\mathbf{U} \otimes \mathbf{V}$** which is an arbitrary vector space to which a bilinear map $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \boldsymbol{u} \otimes \boldsymbol{v}$ is given from $\mathbf{U} \times \mathbf{V}$, having the properties (i) and (ii) above.

Similarly, we easily find that $(\boldsymbol{v}, \boldsymbol{a}) \mapsto \frac{\mathbf{v}}{\mathbf{a}}$ is a linear-quotient map from $\mathbf{V} \times \mathbf{A} \setminus \{\mathbf{0}\}$ into $\frac{\mathbf{V}}{\mathbf{A}}$, i.e. it is linear in variable $\boldsymbol{v}$ and $\frac{\mathbf{v}}{\alpha\boldsymbol{a}} = \frac{1}{\alpha}\frac{\mathbf{v}}{\mathbf{a}}$ for all nonzero real numbers α. Moreover,

(i) $\frac{\mathbf{v}}{\mathbf{a}} = \mathbf{0}$ if and only if $\boldsymbol{v} = \mathbf{0}$,

(ii) every element of $\frac{\mathbf{V}}{\mathbf{A}}$ is of the form $\frac{\mathbf{v}}{\mathbf{a}}$.

Then we consider the **abstract tensor quotient** $\frac{\mathbf{V}}{\mathbf{A}}$ as an arbitrary vector space to which a linear-quotient map $(\boldsymbol{v}, \boldsymbol{a}) \mapsto \frac{\boldsymbol{v}}{\boldsymbol{a}}$ is given from $\mathbf{V} \times \mathbf{A} \setminus \{\mathbf{0}\}$ having the properties (i) and (ii) above.

We have the following identifications of various tensor products and quotients for vector spaces $\mathbf{V}, \mathbf{U}, \mathbf{A}$ and $\mathbf{B}$ with $\dim \mathbf{A} = \dim \mathbf{B} = 1$:

$$\frac{\mathbb{R}}{\mathbf{A}} \equiv \mathbf{A}^*, \qquad \frac{\alpha}{\boldsymbol{a}} \cdot \boldsymbol{b} \equiv \alpha \frac{\boldsymbol{b}}{\boldsymbol{a}};$$

$$\frac{\mathbf{V}}{\mathbf{A}} \equiv \mathbf{V} \otimes \mathbf{A}^*, \qquad \frac{\boldsymbol{v}}{\boldsymbol{a}} \equiv \boldsymbol{v} \otimes \frac{1}{\boldsymbol{a}};$$

$$\frac{\mathbf{V}^*}{\mathbf{A}^*} \equiv \left(\frac{\mathbf{V}}{\mathbf{A}}\right)^*, \qquad \frac{\boldsymbol{p}}{\mathbf{h}} \cdot \frac{\boldsymbol{v}}{\boldsymbol{a}} \equiv \frac{\boldsymbol{p} \cdot \boldsymbol{v}}{\mathbf{h}\boldsymbol{a}};$$

$$\frac{\left(\frac{\mathbf{V}}{\mathbf{A}}\right)}{\mathbf{B}} \equiv \frac{\mathbf{V}}{\mathbf{A} \otimes \mathbf{B}}, \qquad \frac{\left(\frac{\boldsymbol{v}}{\boldsymbol{a}}\right)}{\boldsymbol{b}} \equiv \frac{\boldsymbol{v}}{\boldsymbol{a}\boldsymbol{b}};$$

$$\frac{\mathbf{V}}{\mathbf{A}} \otimes \frac{\mathbf{U}}{\mathbf{B}} \equiv \frac{\mathbf{V} \otimes \mathbf{U}}{\mathbf{A} \otimes B} \equiv \frac{\mathbf{V}}{\mathbf{A} \otimes \mathbf{B}} \otimes \mathbf{U} \equiv \text{etc.}$$

$$\frac{\boldsymbol{v}}{\boldsymbol{a}} \otimes \frac{\boldsymbol{u}}{\boldsymbol{b}} \equiv \frac{\boldsymbol{v} \otimes \boldsymbol{u}}{\boldsymbol{a}\boldsymbol{b}} \equiv \frac{\boldsymbol{v}}{\boldsymbol{a}\boldsymbol{b}} \otimes \boldsymbol{u} \equiv \text{etc.}$$

Note that according to the last two identifcations, the rules of tensorial multiplication and division coincide with those well known for numbers.

As a combination of the preceding identifications, we have

$$\mathbf{A} \otimes \mathbf{A}^* = \mathbf{A}^* \otimes \mathbf{A} \equiv \mathbb{R}, \qquad \boldsymbol{a}\mathbf{h} = \mathbf{h}\boldsymbol{a} \equiv \mathbf{h} \cdot \boldsymbol{a}.$$

A linear map $\boldsymbol{J} : \mathbf{V} \to \mathbf{B}$ with $\dim(\mathbf{B}) = 1$ can be considered as a linear map $\frac{\mathbf{V}}{\mathbf{A}} \to \frac{\mathbf{B}}{\mathbf{A}}$ for all one dimensional vector spaces $\mathbf{A}$ in such a way that

$$\boldsymbol{J} \cdot \frac{\boldsymbol{v}}{\boldsymbol{a}} := \frac{\boldsymbol{J} \cdot \boldsymbol{v}}{\boldsymbol{a}}.$$

This corresponds to the identifications above. Namely, because of $\mathbf{A} \otimes \boldsymbol{A}^* \equiv \mathbb{R}$, we have

$$\mathbf{B} \otimes \mathbf{V}^* \equiv \frac{\mathbf{B} \otimes \mathbf{V}^*}{\mathbf{A} \otimes \mathbf{A}^*} \equiv \frac{\mathbf{B}}{\mathbf{A}} \otimes \left(\frac{\mathbf{V}}{\mathbf{A}}\right)^*.$$

20.4 Contractions

The assignments $\boldsymbol{v}\otimes\boldsymbol{p}\mapsto\boldsymbol{v}\cdot\boldsymbol{p}$ and $\boldsymbol{p}\otimes\boldsymbol{v}\mapsto\boldsymbol{p}\cdot\boldsymbol{v}$ can be extended to linear maps

$$\mathrm{Tr}:\mathbf{V}\otimes\mathbf{V}^*\to\mathbb{R}\qquad\text{and}\qquad\mathbf{V}^*\otimes\mathbf{V}\to\mathbb{R},$$

called **trace**.

Keep in mind that only mixed tensors have a trace; tensors and cotensors do not have.

More generally, if $\mathbf{U}$ is a vector space, then we can define the trace

$$\mathbf{U}\otimes\mathbf{V}\otimes\mathbf{V}^*\to\mathbf{U},\quad \boldsymbol{u}\otimes\boldsymbol{v}\otimes\boldsymbol{p}\mapsto(\boldsymbol{p}\cdot\boldsymbol{v})\boldsymbol{u}.$$

If $\boldsymbol{F}:\mathbf{Z}\to\mathbf{V}$ and $\boldsymbol{G}:\mathbf{V}\to\mathbf{U}$ are linear maps, then their composition is denoted by a dot product $\boldsymbol{G}\cdot\boldsymbol{F}$. In the language of tensors, this corresponds to the dot product

$$(\mathbf{U}\otimes\mathbf{V}^*)\times(\mathbf{V}\otimes\mathbf{Z})\to\mathbf{U}\otimes\mathbf{Z},$$

$$(\boldsymbol{u}\otimes\boldsymbol{p})\cdot(\boldsymbol{v}\otimes\boldsymbol{z}):=(\boldsymbol{p}\cdot\boldsymbol{v})\boldsymbol{u}\otimes\boldsymbol{z},$$

called **contraction**.

A similar formula is valid for tensor quotients, too.

21 Euclidean vector spaces

A Euclidean vector space is a triplet $(\mathbf{E},\mathbb{D},\boldsymbol{b})$ where

- $\mathbf{E}$ is a finite dimensional real vector space,
- $\mathbb{D}$ is an oriented one dimensional real vector space,
- $\boldsymbol{b}:\mathbf{E}\times\mathbf{E}\to\mathbb{D}\otimes\mathbb{D}$ is a symmetric positive definite bilinear map.

Recall that $\mathbb{D}\otimes\mathbb{D}$ is oriented, therefore it is totally ordered (see 19.3), thus positive definitness means that $\boldsymbol{b}(\boldsymbol{x},\boldsymbol{x})>0$ for all non-zero $\boldsymbol{x}$.

The Euclidean structure $\boldsymbol{b}$ allows us to make the identification

$$\frac{\mathbf{E}}{\mathbb{D}\otimes\mathbb{D}}\equiv\mathbf{E}^*$$

in such a way that $\frac{\boldsymbol{q}}{\mathbf{m}^2}$ is identified with the linear form (covector) $\boldsymbol{x}\mapsto\frac{\boldsymbol{b}(\boldsymbol{q},\boldsymbol{x})}{\mathbf{m}^2}$.

According to this identification and the dot product of contractions (see 20.4), we shall write

$$\boldsymbol{x}\cdot\boldsymbol{y}:=\boldsymbol{b}(\boldsymbol{x},\boldsymbol{y})\in\mathbb{D}\otimes\mathbb{D}\qquad(\boldsymbol{x},\boldsymbol{y}\in\mathbf{E}).$$

The above identificataion results in another one, namely

$$\frac{\mathbf{E}}{\mathbb{D}} \equiv \frac{\mathbf{E}}{\mathbb{D}}^* .$$

If $\mathbf{A}$ is a measure line, the elements of $\mathbf{A} \otimes \frac{\mathbf{E}}{\mathbb{D}}$ will be called **vectors of type A**. Thus a vector (an element of $\mathbf{E} = \mathbb{D} \otimes \frac{\mathbf{E}}{\mathbb{D}}$) is a vector of type $\mathbb{D}$.

According to our earlier convention, we can take the dot product of a vector $\boldsymbol{x}$ of type $\mathbf{A}$ and a vector $\boldsymbol{y}$ of type $\mathbf{B}$:

$$\boldsymbol{x} \cdot \boldsymbol{y} \in \mathbf{A} \otimes \mathbf{B}.$$

Moreover, we define the length of a vector $\boldsymbol{x}$ of type $\mathbf{A}$ by

$$|\boldsymbol{x}| := \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} \in \mathbf{A}_{\mathbf{0}}^{+}.$$

This length has the following fundamental properties
- $|\boldsymbol{x}| = \mathbf{0}$ if and only if $\boldsymbol{x} = \mathbf{0}$,
- $|\alpha \boldsymbol{x}| = |\alpha||\boldsymbol{x}|$,
- $|\boldsymbol{x} + \boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{y}|$

for all vectors $\boldsymbol{x}, \boldsymbol{y}$ of type $\mathbf{A}$ and real numbers α. The last relation is called the **triangle inequality** and is proved by the Cauchy–Schwartz inequality:

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq |\boldsymbol{x}| \; |\boldsymbol{y}|$$

and equality holds if and only if $\boldsymbol{x}$ and $\boldsymbol{y}$ are parallel.

This inequality allows us to define the **angle** formed by $\boldsymbol{x} \neq 0$ and $\boldsymbol{y} \neq 0$:

$$\arg(\boldsymbol{x}, \boldsymbol{y}) := \arccos \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{|\boldsymbol{x}| \; |\boldsymbol{y}|}.$$

If $\dim \mathbf{E} = 3$ and $\mathbf{E}$ is oriented, then we can define the cross product

$$\frac{\mathbf{E}}{\mathbb{D}} \times \frac{\mathbf{E}}{\mathbb{D}} \to \frac{\mathbf{E}}{\mathbb{D}} \qquad (\boldsymbol{x}, \boldsymbol{y}) \to \boldsymbol{x} \times \boldsymbol{y}$$

in such a way that
- $\boldsymbol{x} \times \boldsymbol{y}$ be orthogonal to both $\boldsymbol{x}$ and $\boldsymbol{y}$, i.e. $(\boldsymbol{x} \times \boldsymbol{y}) \cdot \boldsymbol{x} = (\boldsymbol{x} \times \boldsymbol{y}) \cdot \boldsymbol{y} = 0$,
- $|\boldsymbol{x} \times \boldsymbol{y}| = |\boldsymbol{x}||\boldsymbol{y}| \sin(\arg(\boldsymbol{x}, \boldsymbol{y}))$,
- $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x} \times \boldsymbol{y}$ form a positively oriented basis.

Of course, the cross product of a vector $\boldsymbol{x}$ of type $\mathbf{A}$ by a vector $\boldsymbol{y}$ of type $\mathbf{B}$ can be defined as a vector of type $\boldsymbol{A} \otimes \mathbf{B}$.

Recall that according to our identifications, $\frac{\mathbf{E}}{\mathbb{D}} \wedge \frac{\mathbf{E}}{\mathbb{D}}$ is the set of real valued antisymmetric bilinear forms on $\frac{\mathbf{E}}{\mathbb{D}}$ (antisymmetric tensors). The cross product establishes the identification

$$\frac{\mathbf{E}}{\mathbb{D}} \wedge \frac{\mathbf{E}}{\mathbb{D}} \equiv \frac{\mathbf{E}}{\mathbb{D}}, \qquad \boldsymbol{x} \wedge \boldsymbol{y} \equiv \boldsymbol{x} \times \boldsymbol{y}$$

which can be transferred to antisymmetric tensors of arbitrary type, e.g.

$$\mathbf{E}^* \wedge \mathbf{E}^* \equiv \frac{\mathbf{E}}{\mathbb{D} \otimes \mathbb{D}} \wedge \frac{\mathbf{E}}{\mathbb{D} \otimes \mathbb{D}} \equiv \frac{\mathbf{E}}{\mathbb{D}^{\otimes 3}} \equiv \frac{\mathbf{E}^*}{\mathbb{D}}.$$

22 Minkowskian vector spaces

A Minkowskian vector space is a triplet $(\mathbf{M}, \mathbb{I}, \boldsymbol{g})$, where

– $\mathbf{M}$ is a finite dimensional real vector space, whose dimension is $1 + N$ where N is a positive integer,

– $\mathbb{I}$ is an oriented one dimensional real vector space,

– $\boldsymbol{g} : \mathbf{M} \times \mathbf{M} \to \mathbb{I} \otimes \mathbb{I}$ is a Lorentzian form, i.e. a symmetric bilinear map of type (1,N) which means that if $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_N$ is a set of $\boldsymbol{g}$-orthogonal vectors, i.e. $\boldsymbol{g}(\boldsymbol{e}_i, \boldsymbol{e}_k) = 0$ if $i \neq k$, then (with conveniently chosen numbering) $\boldsymbol{g}(\boldsymbol{e}_0, \boldsymbol{e}_0) < 0$ and $\boldsymbol{g}(\boldsymbol{e}_i, \boldsymbol{e}_i) > 0$ for $i = 1, \ldots, N$.

The Lorentzian form $\boldsymbol{g}$ allows us to make the identification

$$\frac{\mathbf{M}}{\mathbb{I} \otimes \mathbb{I}} \equiv \mathbf{M}^*$$

in such a way that $\frac{\boldsymbol{y}}{\mathbf{s}^2}$ is identified with the linear form (covector) $\boldsymbol{x} \mapsto \frac{\boldsymbol{g}(\boldsymbol{y}, \boldsymbol{x})}{\mathbf{s}^2}$.

According to this identification and the dot product of contractions (see 20.4), we shall write

$$\boldsymbol{x} \cdot \boldsymbol{y} := \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{I} \otimes \mathbb{I} \qquad (\boldsymbol{x}, \boldsymbol{y} \in \mathbf{M}),$$

and similarly, if $\boldsymbol{k}, \boldsymbol{n} \in \frac{\mathbf{M}}{\mathbb{I}}$ then

$$\boldsymbol{n} \cdot \boldsymbol{x} \in \mathbb{I}, \qquad \boldsymbol{k} \cdot \boldsymbol{n} \in \mathbb{R}$$

, etc.

Since $\mathbb{I}$ is oriented, we can take the square root of non-negative elements of $\mathbb{I} \otimes \mathbb{I}$, so we define the **pseudo-length** of vectors:

$$|\boldsymbol{x}| := \sqrt{|\boldsymbol{x} \cdot \boldsymbol{x}|}.$$

We find that
- $|\mathbf{x}| = \mathbf{0}$ if $\mathbf{x} = \mathbf{0}$ but $|\mathbf{x}| = \mathbf{0}$ does not imply $\mathbf{x} = \mathbf{0}$,
- $|\alpha\mathbf{x}| = |\alpha||\mathbf{x}|$ for all $\alpha \in \mathbb{R}$,
- there is no definite relation between $|\mathbf{x} + \boldsymbol{y}|$ and $|\mathbf{x}| + |\boldsymbol{y}|$.

Recall that $\mathbb{I} \otimes \mathbb{I}$ is oriented, so the following notation is well defined:

$$\mathrm{T} := \{\boldsymbol{z} \in \mathbf{M}| \ \boldsymbol{z} \cdot \boldsymbol{z} < \mathbf{0}\}.$$

If $\boldsymbol{z}$ is in T, then $\{\mathbf{x} \in \mathbf{M} \mid \boldsymbol{z} \cdot \mathbf{x} = \mathbf{0}\}$ is a linear subspace and the restriction of $\boldsymbol{g}$ onto this linear subspace gives a Euclidean structure, i.e. the pseudo-length on this subspace becomes a length having the properties listed in the previous subsection.

The famous **reversed Cauchy inequality** reads as follows: if $\mathbf{x}, \boldsymbol{y} \in \mathrm{T}$ then

$$|\mathbf{x} \cdot \boldsymbol{y}| \geq |\mathbf{x}||\boldsymbol{y}| > \mathbf{0}$$

and equality holds if and only if $\mathbf{x}$ and $\boldsymbol{y}$ are parallel.

This implies the **reversed triangle inequality**: if $\mathbf{x}, \boldsymbol{y} \in \mathrm{T}$ and $\mathbf{x} \cdot \boldsymbol{y} < \mathbf{0}$, then

$$|\mathbf{x} + \boldsymbol{y}| \geq |\mathbf{x}| + |\boldsymbol{y}|$$

and equality holds if and only if $\mathbf{x}$ and $\boldsymbol{y}$ are parallel. The elements $\mathbf{x}$ and $\boldsymbol{y}$ of T are said to **have the same arrow** if $\mathbf{x} \cdot \boldsymbol{y} < \mathbf{0}$.

Having the same arrow is an equivalence relation on T and there are two equivalence classes, called **arrow classes**.

The arrow classes of T are open convex cones whose apex is the zero vector: if $\mathbf{x}$ and $\boldsymbol{y}$ have the same arrow then $\alpha\mathbf{x} + \beta\boldsymbol{y}$ is in their arrow class for all α, β positive real numbers.

We say that $\boldsymbol{g}$ is **arrow oriented** if we select one of the arrow classes of T.

23 Affine spaces

23.1 Fundamentals

An **affine space** is a triplet $(\mathtt{V}, \mathbf{V}, -)$ where
- $\mathtt{V}$ is a non-void set,
- $\mathbf{V}$ is a vector space,
- $-$ is a map from $\mathtt{V} \times \mathtt{V}$ into $\mathbf{V}$, denoted by

$$(x, y) \mapsto x - y,$$

having the properties

1) for every $o \in \mathtt{V}$ the map $O_o : \mathtt{V} \to \mathbf{V}$, $x \mapsto x - o$ is bijective,

2) $(x - y) + (y - z) + (z - x) = \mathbf{0}$ for all $x, y, z \in \mathtt{V}$.

O_o is often called the **vectorization** of $\mathtt{V}$ **with origin** o.

As usual, we shall denote an affine space by a single letter; we say that $\mathtt{V}$ is an affine space over the vector space $\mathbf{V}$ and we call the map $-$ **subtraction.**

In particular, a vector space $\mathbf{V}$, endowed with the vectorial subtraction, is an affine space over itself.

The **dimension** of an affine space $\mathtt{V}$ is, by definition, the dimension of the underlying vector space $\mathbf{V}$. $\mathtt{V}$ is **oriented** if $\mathbf{V}$ is oriented.

From properties 1) and 2) above we infer that for all $x, y \in \mathtt{V}$

– $x - y = \mathbf{0}$ if and only if $x = y$,

– $x - y = -(y - x)$;

moreover, for a natural number $n \geq 3$ and $x_1, x_2, \dots, x_n \in \mathtt{V}$,

$$(x_1 - x_2) + (x_2 - x_3) + \dots + (x_n - x_1) = \mathbf{0}.$$

Observe that the sign '$-$' on the right-hand side of the second item above denotes two different objects. Inside the parantheses it means the subtraction in the affine space, outside it means the subtraction in the underlying vector space. This ambiguity does not cause confusion if we are careful. We even find it convenient to increase the ambiguity a little furhter:

For given $y \in \mathtt{V}$, the inverse of the map O_y is denoted by

$$\mathbf{V} \to \mathtt{V}, \qquad \mathbf{x} \mapsto y + \mathbf{x}. \tag{*}$$

Hence, by definition,

$$y + (x - y) = x$$

and a simple reasoning shows that

$$(x + \mathbf{x}) + \boldsymbol{y} = x + (\mathbf{x} + \boldsymbol{y}).$$

Here the symbol '+' on the left-hand side stands twice for the operation $(*)$, on the right-hand side first it denotes this operation and then the addition of vectors.

Keep in mind the following:

– the sum and the difference of two vectors, and the multiple of a vector by a number are meaningful, they are vectors;

– the difference of two elements of an affine space is meaningful, it is a vector (sums and multiples make no sense);

– the sum of an affine space element and of a vector is meaningful, it is an element of the affine space.

According to these, we can apply all the usual rules of addition and subtraction paying always attention to that the operations be meaningful; for instance, the equality $(x - y) + (u - v) = (x - v) - (y - u)$ is correct but the equality $(x - y) + (u - v) = (x + u) - (y + v)$ is not, since the right-hand side makes no sense.

23.2 Factor spaces

Let $\mathbf{E}$ be a linear subspace of the vector space $\mathbf{V}$.

A subset of the affine space $\mathtt{V}$ is an **affine subspace directed by E** if it is of the form $v + \mathbf{E}$ for some $v \in \mathtt{V}$. Note that $v + \mathbf{E} = u + \mathbf{E}$ if and only if $v - u \in \mathbf{E}$.

An affine subspace directed by $\mathbf{E}$, endowed with the subtraction inherited from $\mathtt{V}$, is an affine space over $\mathbf{E}$.

Points of $\mathtt{V}$ are zero-dimensional affine subspaces. An affine subspace directed by a one-dimensional linear subspace is called a **straigth line**, and an affine subspace directed by a linear subspace of co-dimension one is called a **hyperplane**. Two affine subspaces directed by the same linear subspace are called **parallel**.

The set of affine subspaces directed by $\mathbf{E}$ is denoted by $\mathtt{V}/\mathbf{E}$.

Endowed with the following well defined subtraction

$$(v + \mathbf{E}) - (u + \mathbf{E}) := (v - u) + \mathbf{E},$$

$\mathtt{V}/\mathbf{E}$ becomes an affine space over $\mathbf{V}/\mathbf{E}$.

23.3 Affine maps

Let $\mathtt{U}$ and $\mathtt{V}$ be affine spaces over $\mathbf{U}$ and $\mathbf{V}$, respectively. A map $L : \mathtt{V} \to \mathtt{U}$ is called **affine** if there is a linear map $\boldsymbol{L} : \mathbf{V} \to \mathbf{U}$ such that

$$L(y) - L(x) = \boldsymbol{L} \cdot (y - x) \qquad (x, y, \in \mathtt{V}).$$

The linear map $\boldsymbol{L}$ is unique. We say that L is an affine map over $\boldsymbol{L}$. If L is a bijection, $\mathtt{U}$ and $\mathtt{V}$ are oriented, L is called **orientation-preserving** or **orientation-reversing** if $\boldsymbol{L}$ has that property.

The formula above is equivalent to

$$L(x + \mathbf{x}) = L(x) + \boldsymbol{L} \cdot \mathbf{x} \qquad (x \in \mathtt{V}, \mathbf{x} \in \mathbf{V}).$$

The affine map $L : \mathtt{V} \to \mathtt{U}$ is injective or surjective if and only if $\boldsymbol{L}$ is injective or surjective, respectively; if L is bijective then L^{-1} is an affine bijection over $\boldsymbol{L}^{-1}$;

For all u in the range of L, $\{v \in \mathtt{V} \mid L(v) = u\}$ is an affine subspace of $\mathtt{V}$, directed by the kernel of $\boldsymbol{L}$, i.e. by the linear subspace $\{\mathbf{v} \in \mathbf{V} \mid \boldsymbol{L} \cdot \boldsymbol{v} = \mathbf{0}\}$.

23.4 Differentiation

The reader is supposed to be familiar with the fundamental notions of analysis connected with metrics: open subsets, closed subsets, convergence, continuity, etc.

If $\mathtt{V}$ is an affine space over the (finite dimensional, real) vector space $\mathbf{V}$ and $\| \ \|$ is a norm on $\mathbf{V}$, then $\mathtt{V} \times \mathtt{V} \to \mathbb{R}, \qquad (x, y) \mapsto \|x - y\|$ is a metrics on $\mathtt{V}$.

Since all the norms on a finite-dimensional vector space are equivalent, i.e. they determine the same open subsets, closed subsets, convergent sequences, continuous functions etc. we can speak about open subsets, closed subsets, continuity etc. without giving an actual norm. Linear, bilinear, multilinear maps and affine maps are automatically continuous (we are in finite dimensions!).

As usual, if $\mathbf{U}$ and $\mathbf{V}$ are finite-dimensional vector spaces, ordo : $\mathbf{V} \rightarrowtail \mathbf{U}$ denotes a function such that

– it is defined in a neighbourhood of $\mathbf{0} \in \mathbf{V}$,

– $\lim\limits_{\boldsymbol{x} \to \mathbf{0}} \frac{\mathrm{ordo}(\boldsymbol{x})}{\|\boldsymbol{x}\|} = \mathbf{0}$ for some (hence for every) norm $\| \cdot \|$ on $\mathbf{V}$.

Let $\mathtt{U}$ and $\mathtt{V}$ be affine spaces. A map $F : \mathtt{V} \rightarrowtail \mathtt{U}$ is called **differentiable** at an interior point x of the domain of F if there is a linear map $\mathrm{D}F(x) : \mathbf{V} \to \mathbf{U}$ such that

$$F(y) - F(x) = \mathrm{D}F(x) \cdot (y - x) + \mathrm{ordo}(y - x).$$

$\mathrm{D}F(x)$ is the derivative of F at x.

F is **differentiable** if it is differentiable at every point of its domain (which is necessarily open in this case). F is **continuously differentiable** if it is differentiable and $\mathtt{V} \to \mathrm{Lin}(\mathbf{V}, \mathbf{U})$, $x \mapsto \mathrm{D}F(x)$ is continuous.

F is **twice differentiable** if it is differentiable and $\mathtt{V} \rightarrowtail \mathrm{Lin}(\mathbf{V}, \mathbf{U})$ $x \mapsto \mathrm{D}F(x)$ is differentiable.

Differentiability of higher order is defined similarly. An infinitely many times differentiable map is called **smooth.**

The derivative of F at x, the linear map $\mathrm{D}F(x)$, in the language of tensors, is an element of $\mathbf{U} \otimes \mathbf{V}^*$. Then the second derivative of F at x, denoted by $\mathrm{D}^2F(x)$ is an element of $\mathbf{U} \otimes \mathbf{V}^* \otimes \mathbf{V}^*$, and the n-th derivative is an element of $\mathbf{U} \otimes \mathbf{V}^{\otimes n}$.

Functions defined in $\mathtt{V}$ and having values in $\mathbf{V}$ (vector fields) or in $\mathbf{V}^*$ (covector fields) have a special interest.

The derivative of a vector field $\boldsymbol{X} : \mathtt{V} \rightarrowtail \mathbf{V}$ at x is a mixed tensor $\mathrm{D}\boldsymbol{X}(x) \in \mathbf{V} \otimes \mathbf{V}^*$, thus we can take its trace: $(\mathrm{D} \cdot \boldsymbol{X})(x) := \mathrm{Tr}(\mathrm{D}\boldsymbol{X}(x))$. Then the map $x \mapsto (\mathrm{D} \cdot \boldsymbol{X})(x)$ is called the **divergence** of $\boldsymbol{X}$. The divergence of a tensor field is defined similarly (see 20.4)

The derivative of a covector field $\boldsymbol{K} : \mathtt{V} \rightarrowtail \mathbf{V}^*$ at x is a cotensor $\mathrm{D}\boldsymbol{K}(x) \in \mathbf{V}^* \otimes \mathbf{V}^*$, thus we can take its antisymmetrization denoted by $-(\mathrm{D} \wedge \boldsymbol{K})(x)$. Then the map $x \mapsto (\mathrm{D} \wedge \boldsymbol{K})(x)$ is called the **curle** or the **antisymmetric derivative** of $\boldsymbol{K}$. The antsymmetric derivative of an antsymmetric cotensor field can be defined similarly.

The derivative of a function defined in a one dimensional affine space can be obtained in a more explicit form. Let $\mathtt{A}$ be a one-dimensional affine space over the vector space $\mathbf{A}$. According to the definition, the derivative of $f : \mathtt{A} \rightarrowtail \mathtt{V}$ at a, $\mathrm{D}f(a)$ is a linear map $\mathbf{A} \to \mathbf{V}$; in the language of tensors, it can be considered as an element of $\frac{\mathbf{V}}{\mathbf{A}}$. Then we use the notation $f'(a)$ or $\dot{f}(a)$ instead of $\mathrm{D}f(a)$ and we can show that

$$\dot{f}(a) = \lim_{b \to a} \frac{f(b) - f(a)}{b - a}.$$

23.5 Curves

Let $\mathtt{V}$ be an affine space of dimension greater than 1. A subset C in $\mathtt{V}$ is a **curve** or **line** if there is a mapping $p : \mathbb{R} \rightarrowtail \mathtt{V}$, called a **parametrization** of C, such that

- the domain of p is an open interval, the range of p equals C,
- p is continuously differentiable and $\dot{p}(\xi) \neq 0$ for all ξ in the domain of p,
- p is injective and p^{-1} is continuous.

If p and q are paramterizations of the curve C, then $p^{-1} \circ q : \mathbb{R} \rightarrowtail \mathbb{R}$ is continuously differentiable and for all $x \in \mathrm{C}$ $\dot{p}(p^{-1}(x))$ is parallel to $\dot{q}(q^{-1}(x))$.

For $x \in \mathrm{C}$, the vectors parallel to $\dot{p}(p^{-1}(x))$ are called **tangent vectors** of C at x and the **tangent space** of C at x is the one-dimensional vector space of all tangent vectors.

According to the previous equality, the tangent vectors, though defined by a parametrization, are independent of the parametrization.

Two parametrizations p and q of a curve are **equally oriented** if $\dot{p}(p^{-1}(x))$ and $\dot{q}(q^{-1}(x))$ are positive multiple of each other for all $x \in \mathrm{C}$. An equivalence class of equally oriented parametrizations is called an **orientation** of C. C is **oriented** if an orientation of C is given.

In particular, if C is a straight line i.e. a one-dimensional affine subspace of $\mathtt{V}$, then it has the form $x + \mathbf{C}$ where x is an arbitrary point of C and $\mathbf{C}$ is a one-dimensional linear subspace of $\mathrm{V}^{\rightarrow}$. The tangent space of the straight line C at every point is $\mathbf{C}$. An orientation of C is equivalent to an orientation of $\mathbf{C}$ which is a 'half line' (see Subsection19.3).

23.6 Submanifolds

Let $\mathtt{V}$ be an affine space of dimension $N \geq 2$ ad let M be a natural number $1 \leq M \leq N$.

A subset H of $\mathtt{V}$ is called an M-**dimensional simple submanifold** if there is a mapping $p : \mathbb{R}^M \rightarrowtail \mathtt{V}$, called a **parametrization** of H, such that

– the domain of p is open and connected, the range of p equals H,

– p is continuously differentiable and $\mathrm{D}p(\xi)$ is injective for all ξ in the domain of p,

– p is injective and p^{-1} is continuous.

Recall that $\mathrm{D}p(\xi)$ is a linear map $\mathbb{R}^M \to \mathbf{V}$.

A subset H of $\mathtt{V}$ is called an M-**dimensional submanifold** if every $x \in$ H has a neighbourhood $\mathrm{G}(x)$ in $\mathtt{V}$ such that $\mathrm{G}(x) \cap \mathrm{H}$ is an M-dimensional simple submanifold. A paramterization of such a subset of H is called a **local parametrization** of the submanifold.

The inverse of a local parametrization is called a **local coordinatization**.

If p and q are local parametrizations of the M-dimensional submanifold H whose ranges meet, then $p^{-1} \circ q : \mathbb{R}^M \rightarrowtail \mathbb{R}^M$ is continuously differentiable and if $x \in \mathrm{Ran}p \cap \mathrm{Ran}q$, then

$$\mathrm{Ran}\left(\mathrm{D}p(p^{-1}(x))\right) = \mathrm{Ran}\left(\mathrm{D}q(q^{-1}(x))\right).$$

For $x \in$ H, the elements in the range of $\mathrm{D}p(p^{-1}(x))$ are called **tangent vectors** of H at x, and

$$T_x(\mathrm{H}) := \mathrm{Ran}\left(\mathrm{D}p(p^{-1}(x))\right)$$

is the **tangent space** of H at x where p is an arbitrary local parametrization of H such that x is in the range of p.

The tangent space is an M dimensional linear subspace of $\mathbf{V}$ and, according to the previous equality, is independent of the local parametrization.

Note that curves are one-dimensional simple submanifolds. $(N-1)$-dimensional simple submanifolds are called **hypersurfaces**.