Séminaire Poincaré 1 (2005) 1 – 45 **Séminaire Poincaré**

The Geometry of Relativistic Spacetime: from Euclid's Geometry to Minkowski's Spacetime

Jacques BROS
Service de Physique Théorique, C.E. Saclay,
91191 Gif-sur-Yvette, France

"...the word *relativity-postulate* for the requirement of the invariance under the group G_c seems to me very feeble. Since the postulate comes to mean that only the four-dimensional world in space and time is given by phenomena, but that the projection in space and in time may still be undertaken with a certain degree of freedom, I prefer to call it the *postulate of the absolute world* (or briefly the world-postulate)."

H. MINKOWSKI

Cologne Conference, September 1908

Introduction and general survey

From a variety of viewpoints, the theory of relativity appears as one of the major conceptual events that have ever happened in the adventure of knowledge. It is therefore highly pertinent that the scientific community celebrates the “century commemoration” of the revelation of special relativity by two of the four fundamental papers that were published by Einstein in the year 1905. Since then, the historians of science have been able to accumulate a crop of information about the complex genesis and the multiple and intricate aspects of that extraordinary intellectual adventure. However, strangely enough an important pedagogical work still remains to be done, if one retains from that adventure one of its most striking aspects, namely the existence of a united geometrical representation of space and time, called *spacetime*, and the logical necessity of its introduction on the basis of the special properties of the velocity of light. In fact, we think it worthwhile and possible to communicate this geometrical representation not only to learned scientists, but also to any scientifically-curious and/or philosophically-minded student. Let us explain why we think that it is 1) worthwhile and 2) possible.

1) A wide communication of it is worthwhile, because we have here to deal with a genuine “jewel of human knowledge”, in which Physics, Mathematics (at a rather elementary level, see 2) below) and Philosophy are intimately related. Physics at first: one century after its discovery, one can say that in our present knowledge of the universe, the validity of this joint representation of space and time extends from the spacetime scales of microphysics to those of cosmology, which represents a scaling factor of more than 10^{40}. Then Philosophy and Mathematics: we have to deal with an overwhelming ”ontological fusion” of the categories of space and time, through a mental representation which belongs to the platonician world of geometrical concepts. Here is what can be felt as a real shock for the human mind ! With respect to our usual separate perceptions of space and time, the new geometrical conception of spacetime is as much revolutionary as was the idea of the sphericity of the earth and the computation of its circumference by Eratosthenes with respect to the primitive conception of a flat earth. In the latter case, it is only the development of long-distance travels that have made this idea more and more acceptable for the ”common sense” throughout the centuries. In the former case, only motions whose velocity is substantial compared with the velocity of light provide an evidence that the new spacetime framework gives a correct representation of the physical reality. This is indeed attested as well by the motions of particles which are the ultimate components of matter as by the motions of astronomical objects observed by telescopes. It is only the fact (basic in our social existence!) that all of us are ”slowly moving travellers with respect to one another” which comforts us every day in our feeling that the flow of physical time is the same for all of us and therefore perceived as *absolute* (our watches run at the same rythm!); but this viewpoint, which is encoded in the usual “Galilean kinematics” is only the low-velocity approximation of the physically relevant representation of spacetime. The basic character of the physical spacetime is that the lapse of time measured by an experimentalist between two successive events A and B depends on the particular motion which has been adopted by this experimentalist for proceeding from A to B. But this fact becomes conceivable to us if we compare it with the following one which is familiar to our perception: the distance which is measured by an experimentalist between two given points A and B of space depends on the particular path which has been adopted by this experimentalist for going from A to B. As a matter of fact, what may seem here as purely metaphoric turns out to be a deep structural analogy in geometrical terms.

2) A wide communication of it is possible, once one has realized that these purely geometrical aspects of relativity theory can actually be transmitted in the old Greek spirit of Euclid’s geometry. In fact, let us recall (if forgotten) that this so-called ”elementary geometry”, revived in a second golden age by the European geometers (from seventeenth to nineteenth centuries), was given to the pupils of secondary schools of the old Europe as the most secure guide for training the faculties of logics and rational thought ! Here we would like to make the point that (at the age of computers. . .) this framework might also be the most secure one for transmitting to everyone who is interested a simple, but sound idea of *what is the spacetime of relativity theory* ! The simplest the argument, the strongest the impact for the mind !

From the viewpoint of the historian of science, the adventure of relativistic theory can be seen as the unexpected, although unavoidable issue of the major crisis of nineteenth-century physics, in which the concept of a fixed reference medium in the universe, called the *ether*, was in open conflict with the recently discovered laws of electromagnetism. Among a lot of experimental as well as theoretical results, crucial experiments had been proposed and performed as soon as 1887: these were the famous Michelson and Morley experiments about the constancy of the velocity of light. Then almost twenty years of maturation were still necessary for the conceptual elaboration of the theory of special relativity to be performed. Although it was revealed to the scientific community in the year 1905 by Einstein's revolutionary paper entitled "On the electrodynamics of moving bodies" [E1], the theory made a basic use of formulae established previously by Lorentz; moreover its further formulation greatly benefitted from the group-theoretical analysis of Poincaré also delivered in 1905 [4], while it found its achievement in 1908 through Minkowski's illuminating geometrical work [3]. It is indeed the latter which has to be granted for introducing the appropriate new concept of *absolute spacetime*, a concept whose fate was to go far beyond the theory of special relativity, since it played an essential role in the further discovery and formulation of the theory of general relativity by Einstein in 1916.

It will be precisely our purpose to focus on the concept of spacetime and at first on its logical introduction, which may be presented in a spirit that parallels the axiomatization of Euclid's geometry, thanks to an appropriate axiom about the "universality" of the velocity of light. This spacetime, which can be regarded after Minkowski as an *absolute* framework for describing the kinematics of special relativity, is a representation space whose points are interpreted as the "physical events". Any motion which is physically possible between two given events A and B is represented by a certain *world-line* with end-points A and B. There is an absolute orientation of such world-line, which can be called its "time-arrow": its physical meaning is that one of the end-point events, e.g. B, is in the future of the other one A. The pair of events (A, B) is also said to be *causally separated*; it is not the case for all pairs of events. The limits of causality are determined by the world-lines of *light-rays* passing by each event: the Minkowski spacetime is thus basically equipped with a *light-webbed structure*. In that geometrical representation, one is thus led to distinguish radically the "absolute properties", also called "relativistic invariant properties" from the properties which are "relative to a reference frame" and thereby comparable with the effects of spatial perspective in the usual Euclidean geometry. The basic absolute property of Minkowski spacetime is the fact that it is a mathematical space equipped with a *pseudo-distance*, which is closely linked with the existence of the light-webbed structure of the universe: along the world-lines of light-rays, this pseudo-distance vanishes ! The most striking feature of this absolute pseudo-distance is the *inverse triangular inequality*, which is responsible for the overwhelming phenomenon of "Langevin twins": The "length" of one side (e.g. the aging of the twin at rest) is *longer* than the sums of the "lengths" of the other two sides of the triangle (namely the aging of the travelling twin). As a matter of fact, eventhough the full spacetime is (in mathematical terms) an abstract four-dimensional manifold, such an overwhelming property as the aging difference for twins with different motions can be visualized in terms of planar geometry. It is in fact sufficient to consider two-dimensional sections of spacetime in which a single dimension of space is involved for having a fully correct and intuitive geometrical picture of the Minkowskian triangular inequality. Similarly, one can easily visualize in such a planar section of spacetime the phenomenon of relativistic perspective called "the contraction of lengths". Of course, the last important step for our understanding of spacetime concerns the way in which the usual three-dimensional Euclidean geometry is embedded in the Minkowskian four-dimensional spacetime. The fact that different embeddings hold for observers in relative uniform motion is implied by the notion of Lorentz frame; there appears the relevance of the group of Poincaré transformations. All these aspects of elementary Minkowskian geometry following from an axiomatic Euclid-type construction will be covered in our part 2; a short preliminary part is devoted to the use of geometry in mathematical physics, as an introduction to the concept of spacetime.

At that point, one might have the feeling that nothing more has to be added for understanding the kinematics of special relativity, but this is not so. In fact, the conceptual revolution that it represents is so rich that after the basic articles of 1905 and 1908 in which it was delivered, several aspects of it deserved to be deepened and clarified: this was performed around 1960 in two directions.

a) If the parallel between the Euclidean geometry of our usual three-dimensional space and the Minkowskian geometry of four-dimensional spacetime is actually complete in the physical world, this parallel has to be checked not only for the geometry of straight world-lines, namely for uniform motions, but for arbitrary (smooth) curved world-lines, namely for accelerated motions. The interpretation of Minkowskian pseudo-length as a *proper time* measured by a clock along the world-line of the motion and the geometrical property asserting that such a pseudo-length is always smaller than that of the corresponding uniform motion originating and terminating at the same events had to be tested experimentally. This basic property of Minkowskian geometry, which can be nicely summarized by saying that "In proper-time distances, the straight-line is the *longest* distance between two points (namely two events)", was already present in Einstein's article [1] under the physical terminology of "clock slowing-down phenomenon". However, it remained to be checked experimentally that clocks submitted to accelerated motions were as unsensitive to the accelerations as graduated ribbons were unsensitive to curvature for measuring Euclidean curvilinear distances. What was in question in such investigations had to do with the physical nature of the clocks, considered as trustful measuring instruments, whose robustness with regard to the accelerations had to be quantitatively estimated. Thanks to the progress of physics during the twentieth century, the set of traditional clocks (called "dynamical") was enriched by a new class of clocks, based on microphysics phenomenons and called *"atomic clocks"*, whose precision degree and robustness were far higher. Around 1960 (see in particular Sherwin's paper[S]), this property of unsensitivity to accelerations has been established (and confirmed since then with higher and higher precision) for various types of atomic clocks. These results then exclude radically the last objections of the opponents to the "twin paradox" (see [S]). In particular, they allow one to present a completely acceptable version of the twin phenomenon in uniformly accelerated motions, namely a version which is *biologically bearable* by human experimentalists, even though for technical reasons it remains presently a "Gedanken experiment". Moreover, these manifestations of the Minkowskian geometrical structure in accelerated motions gives an opportunity to state clearly that they must not be confused with possible effects of general relativity. In fact, the latter occur substantially when the accelerations are caused by the presence of large masses of matter, which produces an additional curvature effect on the Minkowskian geometry of spacetime.

b) Since 1959 with the articles of Terrell [6] and of V. Weisskopf [7], problems of relativistic perspective have been reconsidered. Progresses have been made on the problem of what should be the real optical appearance of a fast-moving *extended* object with respect to an observer linked to a given Lorentz frame. The understanding of the phenomenon of "contraction of lengths" was thus revisited and corrected for the case of extended objects. Much more recently, impressive visualizations of moving objects with relativistic velocities have been given thanks to the help of computer technique (see [8] and references therein).

An account of the previous developments a) and b) will be given below respectively in parts 3 and 4. Part 5 and the companion paper by Ugo Moschella will illustrate the fundamental role played by the conceptual framework of Minkowski spacetime in two domains of physics whose order of magnitudes of spacetime distances differ by 10^{40}; we mean respectively particle physics and cosmology. A short final part 6 will serve as a bridge between the two papers.

It is at the scale of particle physics phenomenons that the validity of special relativity and of its expression in the Minkowski spacetime framework appears with its full strength. In fact, the second revolutionary discovery which can be found in the second Einstein's paper [2] on special relativity in 1905, namely the equivalence relation of mass and energy $E = mc^2$, provides the relevant kinematical framework for understanding the energy-balance of all the nuclear and electromagnetic reactions. In geometrical terms, this framework corresponds to supplement Minkowski's spacetime by the introduction of another identical Minkowskian space, interpreted as the space of energy-momentum vectors of material points. This framework gives a remarkably good description of the kinematics of high-energy particle physics. In the Minkowskian energy-momentum space, Einstein's relation $E = mc^2$ is visualized under the form of the mass hyperboloid, called the *mass shell* of the particles: it is the surface which represents the set of all possible states of a free relativistic particle with mass m. This description includes the case of photons: for these "massless particles", the mass shell coincides with the "light-cone". In the energy-momentum space, the law of conservation of total energy-momentum admits a simple geometrical formulation. In that space, the Minkowskian triangular inequality accounts for the production of any number of particles in high-energy collisions of two particles (including the massless case of photons). All that constitutes the basic background for the formulation of high-energy particle scattering in the general framework of quantum field theory. In particular, the world-line representation of free particles and of their multiple collisions in Minkowski's spacetime obeying the rules of relativistic kinematics plays a basic role in the corresponding quantum field-theorerical treatment of particle physics: it explains the so-called *Landau singularities* of the multiparticle scattering functions.

At cosmological scales, the concept of spacetime introduced by Minkowski is still valid, provided one includes as a new revolutionary ingredient the notion of *curvature*: here is the geometrical content of *general relativity*. There are two reasons for this curvature phenomenon: while the first one is the local density of matter (or "gravific mass") which is present near each event in the universe, the other one is linked with the expansion of the universe; it is encoded in the so-called *cosmological constant* in the equations of tentative geometrical models of the universe, whose simplest one (with zero mass density) is the de Sitter universe (1917) presented in the companion paper. Under this respect, the role of Minkowskian geometry for the *local* description of the universe throughout its evolution parallels the role of planar Euclidean geometry for the *local* description of the surface of the earth. In mathematical terms, the latter is a two-dimensional Euclidean manifold: the straight-line distance of planar geometry is replaced by the *geodesical distance* between two points of the surface of the earth, *which is the shortest one* with respect to all possible paths joining these two points on the surface. Similarly, the universe (considered throughout its evolution) appears as a four-dimensional Minkowskian (one also says "Lorentzian") manifold: between two causally-separated events, there is a *geodesical time-like distance, which is the longest one* with respect to all possible world-lines joining these two events. For instance, when one estimates the age of the universe to be of the order of 14.5 billions of years, one has in mind the value of such a geodesical time-like distance between an event that can be called "the big bang" (in the most currently accepted cosmological models) and the event called "here and now" by the inhabitants of the earth in the year 2005. However, it is philosophically questioning to remain conscious of the following: according to the structure of Minkowskian manifold of the universe, any other world-line that relates those two events is covered in a *shorter* time-like distance. According to the motion which is associated with that world-line, it can be ... one century, one year, one day, one second ... or even zero, if one considers a light-ray trajectory, namely a world-line which is composed of pieces of light-like geodesics ....

1 On the use of geometry in mathematical physics and the concept of spacetime

1.1 Geometry of description and geometry of representation

As we all know it, Euclidean geometry (in two or three dimensions) corresponds to an idealized description of the space which surrounds us, as it is felt by our visual and tactile perceptions. The etymology of the word "geometer" (and for instance in France its standard meaning as a profession...) is still reminding us of the fact that, since very ancient times, this branch of mathematics was progressively elaborated from the consideration of practical physical problems, such as the measurement and sharing of ground pieces; the description of the trajectories of celestial bodies also provided another powerful motivation for the development of geometry. It is not a triviality, but a subject of wondering and of philosophical questioning that the idealized notions of "elementary geometry" (points, lines etc...) equipped with logical relations called axioms or postulates, allow us to construct "rigorous proofs" of nontrivial properties of the geometrical pictures. While their experimental checking in physical space is fully satisfactory, these properties also appear to us with the strength of evidence as elements of an "absolute reality of the mind", namely of a very special "world of Platonician ideas": the world of geometrical concepts. One can then say that, *as a geometry of description*, Euclidean geometry appears as the oldest manifestation of the spirit of mathematical physics.

Another considerable achievement in the history of mathematics is the fundamental correspondence between numbers and geometrical concepts which started from the length measurement procedure and resulted in the elaboration of Cartesian coordinates and of the so-called "analytic geometry". As it may be already familiar to pupils at the terminal level of high-school, this implies a relationship between algebra and geometry whose interest is two-fold. On the one-hand, the properties of geometrical curves can be equivalently represented by algebraic equations relating the coordinates of their points. This representation is unique, once the choice of a system of coordinates has been specified. For example in orthogonal coordinates, the equation of the unit circle $x^2 + y^2 - 1 = 0$ makes use of the standard Pythagore theorem for characterizing the points $M = (x, y)$ of that curve. On the other hand, any numerical relation between two quantities x and y (always representable by an equation of the form $f(x, y) = 0$) admits a pictorial representation by a curve in a plane equipped with given coordinate axes; this pictorial representation is specially interesting when x and y denote physical quantities related by a physical law. In fact, the curve which one thus constructs represents all the "states" of the observed phenomenon, each state being characterized by a pair of values of the quantities x and y which are simultaneously observed and thus associated with a particular point $M = (x, y)$ of the curve. The geometrical constructions which may be associated with the pictorial representation of a physical phenomenon in a plane or in a three-dimensional space equipped with coordinates pertain to what we shall call a *geometry of representation*. By using such a terminology, we adopt typically a viewpoint of mathematical physicist: while geometry presents all its mathematical characteristics, in particular the fact that its logical arguments are immediately perceived by a special type of global visual intuition, all its elements are here given a physical interpretation in terms of a certain category of phenomenons; in other words, these phenomenons are actually represented in terms of geometrical concepts.

1.2 The use of geometry in more than three dimensions

From a purely mathematical viewpoint, the correspondence between numbers and geometrical concepts can be extended to $n-$dimensional abstract spaces $\mathbf{R}^n$, with n larger than three. The concept of "point in $\mathbf{R}^n$" is now introduced as a $n-$tuplet of coordinates $M = (x_1, \ldots, x_n)$. The concept of "surface of dimension p" with $2 \leq p \leq n-1$ (called "curve" for $p = 1$ and "hypersurface" for $p = n - 1$) is then introduced as a subset of points of $\mathbf{R}^n$ whose coordinates satisfy $n - p$ independent equations; correspondingly, these coordinates can also be expressed by parametric equations involving p independent parameters. If one wishes, one can equip the space $\mathbf{R}^n$ with a Eulidean distance, which is obtained by an obvious extrapolation from the usual one, two and three-dimensional cases: by definition, the squared length of a linear segment $MN = (a_1, a_2, \ldots, a_n)$ is $[MN]^2 = a_1^2 + a_2^2 + \cdots + a_n^2$, which implies the usual triangular inequality $[ON] \leq [OM] + [MN]$. The equation $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$ is represented geometrically by the "unit hypersphere". In any two-dimensional or three-dimensional section of $\mathbf{R}^n$ defined by linear equations in terms of the coordinates, one recovers respectively a plane or a three-dimensional space equipped with the usual Euclidean distance. So one can develop a set of geometrical concepts, relations and constructions which generalize those of the usual geometry; this can be done at will either in terms of equations or in a purely geometrical language.

From the viewpoint of mathematical physics, the use of geometry in more than three dimensions turns out to be necessary, if one wishes to represent phenomenons whose description necessitates more than three independent quantities. A typical example is the six dimensional space $\mathbf{R}^6_{\mathbf{x}_1,\mathbf{x}_2} = \mathbf{R}^3_{\mathbf{x}_1} \times, \mathbf{R}^3_{\mathbf{x}_2}$ of the positions $(\mathbf{x}_1, \mathbf{x}_2)$ of pairs of material points (or pointlike particles) in mutual interaction. Trajectories of such pairs are represented by curves in $\mathbf{R}^6$, described in terms of a time parameter t by equations of the form $\mathbf{x}_1 = \mathbf{x}_1(t)$, $\mathbf{x}_2 = \mathbf{x}_2(t)$. Another type of geometrical representation which is also often used in physics with strong motivations is *complex geometry*: for example the extension of functions of the real *frequency* variable to (analytic) functions of the corresponding complex variable in a domain of the complex plane $\mathbf{C}$ is of current use. It is in fact a basic property of structural functions describing linear response phenomenons, which provides a convenient visual representation of *resonance* phenomenons by real or complex *poles*. In particle physics a similar use of complex geometry in spaces $\mathbf{C} \times \cdots \times \mathbf{C} = \mathbf{C}^n$ of various variables (positions, time, momenta, energies) plays an important conceptual role.

In the following, we shall be concerned with a very special type of geometry of representation, called *spacetime*, whose purpose is to provide a visualization of the motion phenomenons *throughout their whole history*. If we consider motions in the Euclidean space $\mathbf{R}^3$, providing as usual a geometry of description of the world which surrounds us, we need an additional time-coordinate and therefore an affine space $\mathbf{R}^4$ for representing geometrically all the events of the world. Such a map is intended to picture in an idealistic way the whole history of the world: the motion of any material point (or of any observer) will be represented as a curve, called a *world-line*, which describes all its history from the remote past to the far future. The usual notion of trajectory will then appear as the *projection* of the world-line onto the Euclidean space $\mathbf{R}^3$. The world-line is a geometrical concept which contains all the information on the motion, which is not the case for the trajectory: two different worldlines (i.e. motions) may project onto the same trajectory.

1.3 Galilean spacetime as a geometry of representation of motion phenomenons

In its simplest form, which we shall call *Galilean spacetime*, the concept of spacetime appears as a geometry of representation for the phenomenons of motion, *as they are perceived by a privileged observer called* $\mathcal{O}_0$, submitted to the following prejudice: *the time interval that elapses between two events A and B is an absolute quantity*; its value is the same for observers moving in an arbitrary way between A and B, provided they are equipped with identical clocks.

Keeping the previous notations, $x = \mathbf{x}$ now denotes a point, or equivalently three coordinates called *space coordinates*, in the usual Euclidean space $\mathbf{R}^3$ in which we are living, while $y \equiv t$ denotes a time coordinate. A point $X = (\mathbf{x}, t)$ in $\mathbf{R}^4$ represents the *event* which takes place at time t at the point $\mathbf{x}$ of Euclidean space $\mathbf{R}^3$. In particular, the origin O represents the event called "here and now" (at a certain instant…) by the observer $\mathcal{O}_0$, who stands "at rest" at $\mathbf{x} = 0$; by definition, this means that the observer's worldline is the time-axis with equation $\mathbf{x} = 0$. For $\mathcal{O}_0$, the coordinate hyperplane with equation $t = 0$ represents the set of all *simultaneous events* which constitute the "present". Similarly, for every fixed value t_0 of t, the hyperplane with equation $t = t_0$ is a complete set of simultaneous events, which we call *set of simultaneity* and which belongs to the future or to the past according to whether t is positive or negative. The whole *future* and the whole *past* are represented respectively by the open half-spaces $t > 0$ and $t < 0$ of $\mathbf{R}^4$. In such a representation of the events, one says that the time-axis associated with the Euclidean space $\mathbf{R}^3$ of "present events" constitute the *reference frame* of the observer $\mathcal{O}_0$ (the choice of the "present time" $t = 0$ is of course a matter of convention for $\mathcal{O}_0$).

Let $\mathcal{O}_{\mathbf{v}_0}$ be an observer in uniform motion with vector velocity $\mathbf{v}_0$ with respect to $\mathcal{O}_0$ and passing by O, which means that he shares with $\mathcal{O}_0$ the same and unique event that we called "here and now". The time-axis $\Delta_{\mathbf{v}_0}$ for this observer is defined by the corresponding worldline, namely the straightline with (vector) equation $\mathbf{x} = \mathbf{v}_0 t$ (see fig. 1).

For any such observer, the sets of simultaneity $t = t_0$ are the same as for the observer $\mathcal{O}_0$. More precisely, every event $M = (\mathbf{x}, t)$ of spacetime is perceived by the observer $\mathcal{O}_{\mathbf{v}_0}$ as having coordinates $(\mathbf{x}', t')$ such that $\mathbf{x}' = \mathbf{x} - \mathbf{v}_0 t$ and $t' = t$. This change of coordinates from $\mathcal{O}_0$ to $\mathcal{O}_{\mathbf{v}_0}$ is also called a *Galilean transformation*; it implies the basic property of *additivity of velocities*: a uniform motion with worldline $\mathbf{x} = \mathbf{v}t$ is seen by $\mathcal{O}_{\mathbf{v}_0}$ as a uniform motion with equation $\mathbf{x}' = \mathbf{v}'t$, with velocity vector $\mathbf{v}' = \mathbf{v} - \mathbf{v}_0$. For example, in a train whose velocity is v_0 =100 kmh, a passenger walking longitudinally with velocity v' =5 kmh has a velocity with respect to the earth which is v =105 kmh or 95 kmh according to whether the forward or backward direction of the train has been chosen by that passenger…

We note that the Galilean changes of coordinates do not preserve the notion of orthogonality in $\mathbf{R}^4$. If for convenience we choose to represent the simultaneity sets as "horizontal spaces" (the dimension of space being unfortunately reduced to two in our visual perception…) and the time-axis of the observer at rest $\mathcal{O}_0$ by a vertical line, the reference frame for $\mathcal{O}_{\mathbf{v}_0}$ will associate the

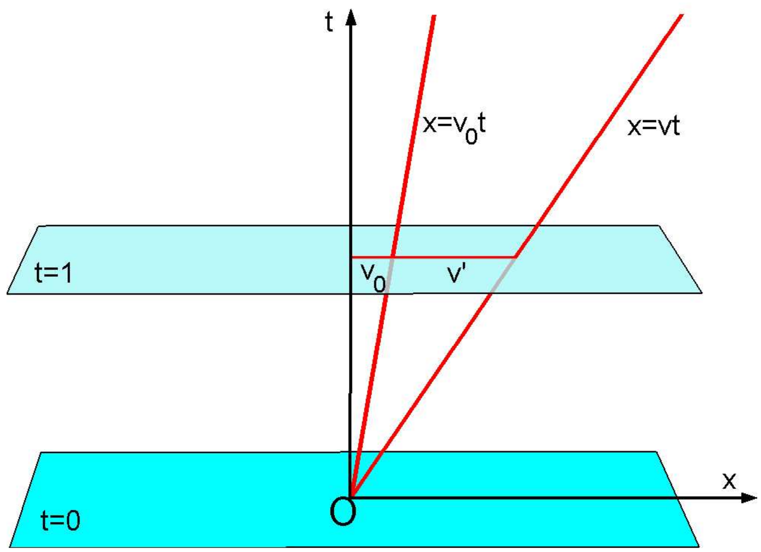

Figure 1: The Galilean spacetime

oblique time-axis $\Delta_{\mathbf{v}_0}$ with the horizontal space. But the observer at rest enjoys no special physical properties with respect to any other observer in uniform motion (that's the *"Galilean principle of relativity"* which follows from the law of inertia). So the verticality of the time-axis could have been chosen for representing the worldline of any given uniform motion: there is nothing deep in that choice. One can also say that the Galilean spacetime is defined for $\mathcal{O}_0$ *up to the arbitrariness in the choice of the time-axis* or in mathematical terms *up to a Galilean transformation*: it is the equivalence class of all these representations. But the *same* representation of spacetime is then also acceptable by any observer $\mathcal{O}_{\mathbf{v}_0}$ in uniform motion, which expresses precisely in geometrical terms the content of the Galilean relativity principle.

Here it is also worthwhile to point out that, in contrast with the "horizontal" Euclidean subspaces $\mathbf{R}^3$, the Galilean spacetime $\mathbf{R}^4$ is only an *affine* space; it is not equipped with any physically sensible global notion of orthogonality and distance. But this is consistent with our standard perception: why would space and time strangely mix each other in some supergeometry? Galilean spacetime is just a geometry of representation in a very poor sense: it has no global geometrical structure. But let us now incorporate the strange properties of light velocity and then discover that such a phantasmic supergeometry holds in the realistic spacetime of physics, namely in the four-dimensional world called *Minkowski's spacetime* !!

2 Postulates and construction of Minkowski's spacetime

Preliminary Remark The postulates and the construction which we propose do not pretend to be the most economical ones from the viewpoint of formal logics. In particular, we must draw the attention of the reader to the important mathematical article by E.C. Zeeman entitled *"Causality Implies the Lorentz Group"* [9]. We shall briefly indicate at the end of Sec.2-1 how the latter can be interpreted in our approach, which is much more pedestrian since making use of the basic physical concept of uniform motion and of the familiar representations of Euclid's geometry.

We shall introduce five postulates for our construction of the *spacetime of special relativity*. The first two postulates introduce a representation of spacetime conceived by the observers at rest, while the third and fourth postulates express minimal properties to be shared by all the observers in uniform motion. The contents of the first and third postulates are easily accepted as being already satisfied in the Galilean spacetime, but the second and fourth postulates introduce the *world-lines of light* as playing a fundamental role in spacetime. In fact, these postulates express in a geometrical way the revolutionary result obtained at first by the experiments of Michelson and Morley: *For all observers, either at rest or in uniform motion, the velocity of light in the vacuum is a universal constant c; neither it depends from the motion and from the nature of the light-emitter, nor from the direction of emission and the various changes of direction of the light rays considered (e.g. obtained by the interposition of mirrors), nor from the wave-length of the light.* Renewed experiments which make use of a variety of experimental devices and whose range extends to electromagnetic waves outside the spectrum of visible light (including in particular the propagation of radiowaves) have been repeatedly performed throughout the twentieth century. They all have confirmed the universality property of c, even if its precise value ($c = 299,776 \cdots$ km/sec as measured in 1940 by Anderson) is now thought to be possibly fluctuating with time at astronomical scales and also depending on the type of clocks (atomic or dynamical) for time measurements. The overwhelming fact about the universality property of c is that light does not satisfy the usual (Galilean) property of additivity of velocities: by switching on a lamp on a train, it is impossible to make its light travel at the velocity c plus the velocity of the train !!

Finally, it is pertinent and (as we shall see) useful for our construction to add a fifth postulate: the latter requires that, *in the limit of very low velocities* (those which we perceive in our life), Galilean spacetime has to be an excellent approximation of the new spacetime. Here lies the wisdom of all revolutions in the domain of science: the old theory is not thrown away as completely perverse, it is honestly recognized as a good first-order approximation of the new theory when the order of magnitude of certain variables lies within certain limits.

NOTATION: In all the following, the symbol $A \doteq B$ will be used when this equality serves as a definition either of A in terms of B or of B in terms of A. Examples: a vector $\mathbf{x} \doteq (x_1, x_2, x_3)$; $x_1^2 + x_2^2 + x_3^2 \doteq \mathbf{x}^2$, the squared norm of $\mathbf{x}$; the norm $|\mathbf{x}| \doteq (\mathbf{x^2})^{\frac{1}{2}}$.

2.1 The postulates and the light-cone structure of spacetime

First postulate: the spacetime representation

All the observers at rest in the Euclidean space $\mathbf{R}^3_{\mathbf{x}}$ (where $\mathbf{x} = (x_1, x_2, x_3)$) agree on the existence of a geometrical representation of all "events" of the universe by points in a space $\mathbf{R}^4_{\mathbf{x},t} = \mathbf{R}^3_{\mathbf{x}} \times \mathbf{R}_t$, with the same notions of simultaneity sets $t = t_0$ as in the Galilean spacetime. The time-axis is the worldline of the observer $\mathcal{O}_0$; the time-axis together with the "present hyperplane" $t = 0$ constitute the reference frame of observers at rest, its origin O being the "present event" ("here and now") of the observer $\mathcal{O}_0$.

This postulate calls for three remarks:

i) The *events*, and thereby their representation by points in $\mathbf{R}^4$ are conceived as *"absolute elements of reality"*; however, the given choice of coordinates $(\mathbf{x}, t)$ priviledges the class of observers at rest, whose worldlines are all the parallels to the time-axis. The basic problem of our construction will be to determine the corresponding choices of coordinates for any observer in uniform motion. As in the Galilean spacetime, the worldline of any observer in uniform motion is a straight-line. For example $\Delta_{\mathbf{v}}$ is the worldline of the observer $\mathcal{O}_{\mathbf{v}}$ whose motion is defined as in the Galilean case.

ii) All the observers at rest are supposed to be equipped with identical devices for measuring lengths (i.e. *graduated rods*) and for measuring time-intervals (i.e. *clocks*). The fact that all observers at rest agree on their Euclidean representation of *space* is trivial for us (after more than 2000 years of cartographical techniques...). The fact that they agree on the simultaneity of two events requires a procedure of "*synchronization of clocks*" through the emission of light-signals, which we shall be led to specify later. For the moment, we just postulate that this notion of simultaneity for all the observers at rest, which is basic for our geometrical representation of spacetime, is physically meaningful.

iii) In all our pictorial representations, the time-axis will be represented as vertical and upward-oriented; the ascending arrow indicates the future. The Euclidean space $\mathbf{R}^3_{\mathbf{x}}$ with equation $t = 0$ is then considered as horizontal. In many arguments, it will be sufficient to consider a *single* space variable $x = x_1$, namely the planar section (Ox_1, Ot) of spacetime, with the axis Ox horizontal and rightward-oriented as usual.

Second postulate: the light-cone

All the world-lines of light rays emitted from the event O by any *(moving or at rest) light-emitter are represented in $\mathbf{R}^4_{\mathbf{x},t}$ by the linear generatrices of the cone C^+ with equation $|\mathbf{x}| = ct,\ t > o$, which is called the "future light-cone of O". Similarly, all the light rays emitted in the past of O by any (moving or at rest) light-emitter and which are detected at O have worldlines which are carried by the generatrices of the cone C^- with equation $|\mathbf{x}| = -ct,\ t < o$, which is called the "past light-cone of O". The whole set of light world-lines passing at O is the set of generatrices of the "light-cone C of O" (see fig.2), with*

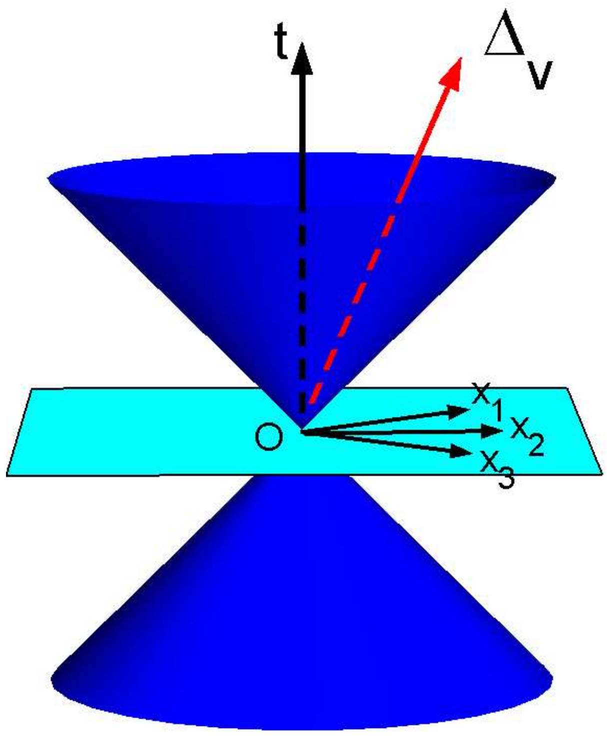

Figure 2: The light-cone

quadratic equation

$$c^2t^2 - (x_1^2 + x_2^2 + x_3^2) = 0.$$

Similarly, with each event $X = (\mathbf{x}, t)$ of $\mathbf{R}^4_{\mathbf{x},t}$, one can associate the "light-cone $C(X)$ of X, which is obtained from C by the action of the translation with vector $\vec{[OX]}$ in $\mathbf{R}^4_{\mathbf{x},t}$.

It is worthwhile to emphasize that the *absolute localization on the cone C* of the world-lines of light rays passing at O did not hold in the usual Galilean spacetime representation, since light was treated there as any other motion and therefore obeyed the principle of additivity of velocities. To be more illustrative, let us consider light-propagation along a single direction of space Ox represented as our horizontal axis, but with the two possible orientations of light rays emitted from O, namely the rightward light ray (towards positive x's) and the leftward light ray (towards negative x's). The world-lines of these two light rays in the planar section (Ox, Ot) of spacetime are respectively the half-lines C_R and C_L with equations $x = ct,\ t > 0$ and $x = -ct,\ t > 0$ (see fig.3): they are the traces of the future light-cone C^+ in the planar section (Ox, Ot). If the light rays emitted from O are emitted from a train with velocity v in the direction Ox, its propagation is still observed by an observer at rest as having the velocity c and not $c+v$ or $c-v$, which would have been the case according to the Galilean viewpoint. In the planar section (Ox, Ot) of Galilean spacetime, the worldlines of the light rays emitted from O would have had equations of the form $x = (\pm c + v)t$ (resp. $x = (\pm c - v)t$), depending on the velocity v (resp. $-v$) along Ox of the light-emitter at O. Therefore, the Galilean world-lines of these light rays would cover the whole half-plane of positive t's, namely "the absolute Galilean future". It is therefore crucial to understand that in the new "relativistic spacetime" that we construct, the half-lines C_R and C_L and more generaly the cone C^+ are *new absolute data*.

Remark on the choice of units: Instead of using the very large value of c expressed in km/sec which would make unpracticable the geometrical representation of spacetime, we can choose time and space units in such a way that $c = 1$. For example, we can adopt the choice of *year* and *light-year* which is standard in astronomy. The light-cone C is then well-represented as the cone with equation $t^2 - \mathbf{x}^2 \doteq t^2 - (x_1^2 + x_2^2 + x_3^2) = 0$ and the light world-lines C_R and C_L are then well-pictured along the diagonals of the axes (Ox, Ot) (fig.3). Another possible convention whose advantage is also to keep the same geometrical representation but without fixing the value of c consists in considering that one plots the variable ct instead of t. Here it is relevant to notice that the variable ct has the "physical dimension" of a distance, which prepares us to understand why it can be treated on the same footing as the space coordinates $\mathbf{x}$ in the following.

Third postulate: isochronousness in all uniform motions

For every observer $\mathcal{O}$ in uniform motion, let $t_{\mathcal{O}}$ be its time-variable , measured by a clock which is identical with that of $\mathcal{O}_0$. Its world-line is a straight line denoted by Δ which carries the time-axis of $\mathcal{O}$. Let then X_1, X_2, X_3 be three events in Δ. We postulate that it is equivalent that their time-coordinates t_1, t_2, t_3 satisfy the equality $t_2 - t_1 = t_3 - t_2$, namely that X_2 be the middle of the segment $[X_1X_3]$, or that the corresponding times $(t_{\mathcal{O}})_1, (t_{\mathcal{O}})_2, (t_{\mathcal{O}})_3$ measured by $\mathcal{O}$ satisfy the equality $(t_{\mathcal{O}})_2 - (t_{\mathcal{O}})_1 = (t_{\mathcal{O}})_3 - (t_{\mathcal{O}})_2$.

This postulate is of course trivially satisfied in the *absolute time* viewpoint of Galilean spacetime. Here one only requires that the flow of time measured via a regular sequence of events by an observer $\mathcal{O}$ is also perceived as regular *up to a change in the scale*, when the same successive events linked to $\mathcal{O}$ are detected (with an identical clock) by an observer in uniform motion with respect to $\mathcal{O}$.

Fourth postulate: "Physical" uniform motions and the universality of c

a) The only uniform motions considered as having a physical meaning are those whose velocity v is smaller than c. For such motions whose world-line $\Delta_{\mathbf{v}}$ contains the event O, $\Delta_{\mathbf{v}} \setminus O$ is made up of two half-lines $\Delta_{\mathbf{v}}^+$ and $\Delta_{\mathbf{v}}^-$ which are respectively contained in the convex conical volumes V^+ and V^-:

V^+ is the set of all events $(\mathbf{x}, t)$ such that $|\mathbf{x}| < ct,\ t > 0$, called "the absolute future of O";

V^- is the set of all events $(\mathbf{x}, t)$ such that $|\mathbf{x}| < -ct,\ t < 0$, called "the absolute past of O".

Similarly, for each event X one can introduce the convex conical volumes $V^+(X)$ and $V^-(X)$, namely respectively the absolute future and past of X, whose union contains all the worldlines Δ of the uniform motions passing at X. The future light-cone $C^+(X)$ (resp. past light-cone $C^-(X)$) thus appears as the boundary of the corresponding future cone $V^+(X)$ (resp. past cone $V^-(X)$).

b) For every observer $\mathcal{O}_{\mathbf{v}}$ with worldline $\Delta_{\mathbf{v}}$ graduated by the time-variable $t_{\mathbf{v}}$, there exist coordinates $\mathbf{x}_{\mathbf{v}}$ of the space perceived at rest *by $\mathcal{O}_{\mathbf{v}}$, such that any event $X = (\mathbf{x}, t)$ of the light-cone C is detected by $\mathcal{O}_{\mathbf{v}}$ as having coordinates $(\mathbf{x}_{\mathbf{v}}, t_{\mathbf{v}})$ satisfying the relation $|\mathbf{x}_{\mathbf{v}}| = c|t_{\mathbf{v}}|$.*

Part a) of the postulate, which requires that the light-velocity is an absolute limit to the propagation velocity of any physical system to which an observer can be linked, will appear as deeper than a pure physical requirement. It will in fact be seen below that the lines of spacetime which could be interpreted as worldlines of motions with velocity larger than c (or "*superluminal motions*") are necessarily given another interpretation, which is of purely *spatial* nature. So the requirement a) is deeply involved in the self-consistency of the relativistic spacetime representation.

Part b) again pertains to the basic statement about the constancy of the velocity of light. It can also be seen as contained in the principle of relativity which claims that all the physical laws, and therefore in particular the velocity of light, are the same for all observers in uniform motion: no rest frame is physically priviledged as it was presupposed in the old concept of ether.

Fifth postulate: validity of the Galilean approximation

For every observer $\mathcal{O}$ in uniform motion or at rest, there is a Galilean representation of spacetime which is an excellent approximation of the exact spacetime for the description of motions whose relative velocity with respect to $\mathcal{O}$ is very small with respect to c.

The precise mathematical formulation of this postulate will appear clearly in the following.

Remark In the present approach, the interpretation of the basic result of [9] seems to be the following. Let us assume that the light-cone structure of the spacetime $\mathbf{R}^4$ holds for the observer at rest $\mathcal{O}_0$ as in our first and second postulates. Let us now consider observers in *unspecified* motion, for which the spacetime $\mathbf{R}^4$ is also perceived with a lightcone structure (implying the same universal velocity of light c). Let us assume that for such observers the causality order of events X, Y (denoted $X < Y$) is defined by the fact that Y belongs to the future cone $V^+(X)$ of X, and that *this order coincides with the one perceived by the observer at rest*. Then it is proven that such observers are *necessarily in uniform motion* and that *their scales of time and length are linear functions of those of the observer at rest* so that the whole structure of Minkowski's spacetime follows. In particular, our postulate three concerning the "isochronousness property" of uniform motions would then be redundant. However, as it has been pointed out in [9], the result does not hold in two-dimensional spacetime; a nontrivial use of the dimension larger than two has been made in that work. Our approach is rather opposite: in view of its pedagogical nature, it aims to exhibit already in two-dimensional spacetime (which is much simpler to describe) how the construction of Minkowski's spacetime can be worked out. In fact, this will be made in detail from Sec.2-2 to Sec.2-6. It is only in Sec.2-7 that we shall be ready to tackle the four-dimensional spacetime equipped with the group of general Lorentz transformations. This subsection, which is more mathematical, may be skipped by the reader more interested in the physical or philosophical aspects of special relativity.

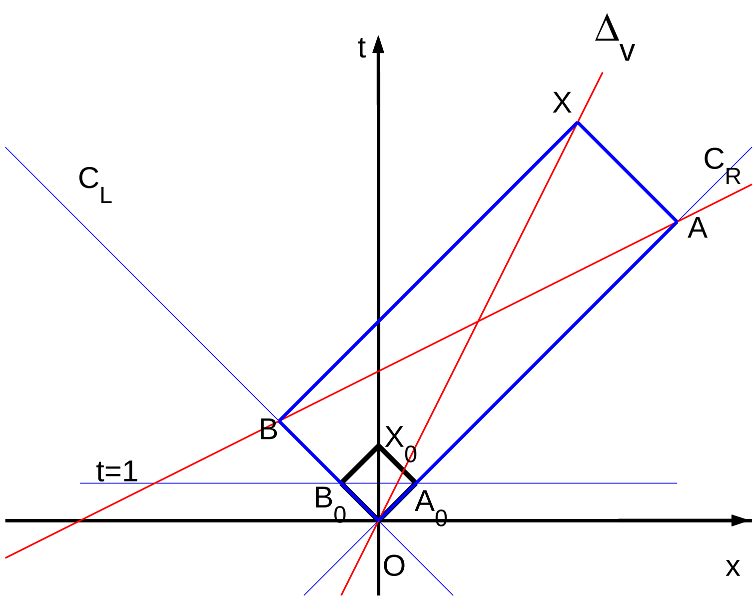

Figure 3: Simultaneous events

2.2 Simultaneousness revisited

The notion of *absolute simultaneousness*, namely the identity of every simultaneity space $t = t_0$ for all the observers (at rest or in motion) is encoded in the Galilean spacetime representation. However this viewpoint is purely idealistic, because for each observer the property of simultaneity of two events is a physical property which has to be checked via some procedure implying the use of lengths and time measurements. Now in view of the universality of the velocity of light, the use of light-signals will be particularly helpful for clarifying the notion of *simultaneousness relatively to each observer at rest or in uniform motion*.

We shall describe a physical procedure for characterizing simultaneous events whose geometrical representation in spacetime is quite simple. It only requires observers and light-signals moving in a single space dimension Ox, which allows one to represent phenomenons in the two-dimensional section (Ox, Ot) of spacetime. We are led to use the geometrical representation of light worldlines as being all parallel either to C_R or to C_L (according to our first and second postulates). For simplicity, chosen units are years and lightyears so that $c = 1$.

For the observer at rest $\mathcal{O}_0$, the procedure must of course confirm that (for instance) the events $A_0 \doteq (x = 1, t = 1)$ and $B_0 \doteq (x = -1, t = 1)$ are simultaneous. To that purpose, one considers rightward and leftward lightrays emitted from O and reflected (by mirrors) at the respective points $x = 1$ and $x = -1$. The worldlines of these reflected lightrays are respectively parallel to C_L and C_R and therefore converge at the event $X_0 \doteq (x = 0, t = 2)$ of the worldline of $\mathcal{O}_0$, which allows the latter to conclude that the "mirror events" A_0 and B_0 are simultaneous: since the velocity of light is the same in right and left directions, the mirror events have been simultaneously produced at half of the time of X_0 (namely $t = 1$). As seen on fig.3, the geometrical representation of the previous light-signal procedure exhibits that the quadrilateral $(OA_0X_0B_0)$ is a parallelogram. We also notice that this procedure is useful for allowing all the observers at rest to synchronize their clocks with respect to $\mathcal{O}_0$'s clock and therefore to agree on the same representation of spacetime. For instance the observer situated at $x = 1$ (i.e. whose worldline has the equation $x = 1$) will be warned by $\mathcal{O}_0$ that he should assign the time $t = 1$ to the event A_0, at which he receives the light signal coming from O.

Now we can repeat the same construction for any given observer $\mathcal{O}_v$ in uniform motion, with $|v| < 1$. We use again two rightward and leftward lightrays emitted from O and therefore represented along C_R and C_L, but we now set the mirrors (at some points $x_A > 0$ and $x_B < 0$) in such a way that the worldlines of the two reflected lightrays intersect at an event X which belongs to the worldline Δ_v of $\mathcal{O}_v$. Here again the two mirror events A and B are such that the quadrilateral $(OAXB)$ formed by the four light worldlines is a parallelogram, and it then follows that, except when $v = 0$, the linear segment AB is *not* parallel to the axis Ox (fig.3).

Now in view of b) of the fourth postulate, the forward and backward travels of light corresponding to the worldline segments OA and AX (resp. OB and BX) are performed during the same time for $\mathcal{O}_v$, since performed at the same universal velocity. Therefore if $t_v(X)$ denotes the time of the event X measured by $\mathcal{O}_v$, the times of the mirror events A and B measured by $\mathcal{O}_v$ will be both equal to $\frac{t_v(X)}{2}$: these two events are therefore to be considered as *simultaneous* by $\mathcal{O}_v$. Moreover (in view of the same postulate), the events A and B will be produced at spatial coordinates $x_v = \pm\frac{t_v(X)}{2}$.

We shall now use our third postulate for proving that *all the points of the straight line* (AB) *represent the events which appear to be simultaneous to* A *and* B *for the observer* $\mathcal{O}_v$.

We consider at first the event G at the intersection of OX and AB. Since (in the parallelogram $(OAXB)$) one has $OG = GX$, the observer $\mathcal{O}_0$ perceives the event G at the time $t(G) = \frac{t(X)}{2}$. Then in view of the third postulate, the event G is also perceived by the observer $\mathcal{O}_v$ at the time $t_v(G) = \frac{t_v(X)}{2}$, which shows that G is simultaneous to A and B for $\mathcal{O}_v$.

Let now P be any point on the half-line with origin G and containing A, and let E and F be the intersections of the straight line (OX) respectively with the parallels to C_R and C_L by P. Thales property then yields (fig.4):

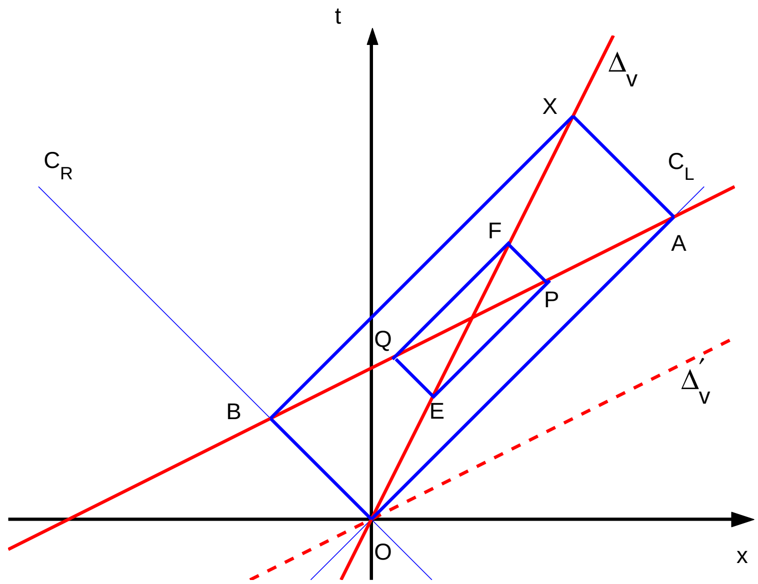

Figure 4: Conjugate axes

$$\frac{GF}{GX} = \frac{GP}{GA} = \frac{EG}{OG}, \quad \text{and therefore} \quad EG = GF.$$

By introducing the point Q, symmetric of P with respect to G one then gets a parallelogram $(EPFQ)$. Therefore the same argument as above applies to the lightrays emitted at E, reflected at P and Q and converging at F: it shows that P,Q and G are simultaneous with respect to $\mathcal{O}_v$. Since the symmetric pair (P,Q) may vary arbitrarily on the straight line (AB), this line is a line of simultaneity for $\mathcal{O}_v$ (corresponding to the time $\frac{t_v}{2}$).

Since the choice of t_v was arbitrary in the previous argument, one concludes that the lines of simultaneity for the observer $\mathcal{O}_v$ in the plane (Ox, Ot) are all the parallels to (AB); *in particular the straight line Δ'_v parallel to (AB) and containing O represents the "present events" $(t_v = 0)$ for $\mathcal{O}_v$* . As seen on fig.4, half of the line Δ'_v (on the right of O for the choice $v > 0$) contains events at $t > 0$, which are therefore perceived as belonging to the future by $\mathcal{O}_0$ together with all the observers at rest, while the other half (on the left of O) contains events at $t < 0$, perceived as belonging to the past by the same observers.

The direction Δ'_v, obtained from Δ_v by the previously described *parallelogram construction*, is said to be *conjugate of Δ_v with respect to the (light world)lines C_R and C_L*. Points $X = (x,t)$ and $X' = (x',t')$ of Δ_v and Δ'_v satisfy the equations

$$x = vt, \quad x' = \frac{1}{v}t', \quad \text{and therefore} \quad tt' - xx' = 0.$$

This calls for two remarks:

i) *conjugacy or pseudo-orthogonality relation:*

The relation $tt' - xx' = 0$ (or in unit-independent form $(ct)(ct') - xx' = 0$) can be called a pseudo-orthogonality relation between the vectors $[OX]$ and $[OX']$, by analogy with the orthogonality relation $xx' + yy' = 0$ in a Euclidean plane. Such a relation, which expresses the geometrical property of *conjugacy* of the pair (Δ_v, Δ'_v) with respect to the pair (C_R, C_L), introduces *a joint geometrical structure of space and time*, which will appear still stronger in the analysis of Sec. 2-3.

For the moment, we can simply notice the following properties of conjugate pairs (Δ_v, Δ'_v):

a) when v varies, Δ_v and Δ'_v are turning in opposite ways (one clockwise and one anticlockwise) in the plane (Ox, Ot).

b) when v tends to 1 (resp. -1), both lines tend together to C_R (resp. C_L).

c) there is a single conjugate pair which is orthogonal, namely the supports of the axes of coordinates Ox, Ot.

Here, however, one must stress that the choice of orthogonal space and time axes Ox, Ot for the observers at rest is a pure convention, as it was already the case for the Galilean spacetime representation. A more general, but equivalent choice *which does not ascribe a special role to observers at rest* would be the following. One first gives oneself the pair of light worldlines (C_R, C_L) and one chooses for (Ox, Ot) any pair of straight lines which are conjugate with respect to (C_R, C_L) (defined intrinsically through the parallelogram construction). The analysis above would have given the same result, namely that the time and space axes for any observer $\mathcal{O}_v$ are carried by conjugate pairs (Δ_v, Δ'_v) with respect to (C_R, C_L). Among them, the special pair which is orthogonal (namely the bisectors of (C_R, C_L)) would then be associated with *a certain* uniform motion having no special physical properties: in fact, it was one of the primary ideas of special relativity theory that systems in uniform motions are physically undistinguishable . So, as in the Galilean case, we keep the idea that the orthogonality of the rest system is only a convenient convention, but there is a whole class of equivalent representations of the planar relativistic spacetime in which the following notions have an *absolute* meaning: *i) the light lines* (C_R, C_L) *and ii) the systems of conjugate pairs* (Δ_v, Δ'_v) *for the coordinate axes of uniform motions, including the rest system.*

ii) *"superluminal motions":*

For $\mathcal{O}_0$, the line Δ'_v might be interpreted as the worldline of a superluminal motion with velocity $\frac{1}{v}(=\frac{c^2}{v})$...But this would be very strange, since all the events of that line are perceived as simultaneous by $\mathcal{O}_v$: for the latter, a hypothetic observer $\mathcal{O}'_v$ with worldline Δ'_v would then have the "ubiquity property" $(t_v = 0, x_v$ arbitrary)! The interpretation of this motion would become even more paradoxical for an observer $\mathcal{O}_w$ with velocity w such that $v < w < c$. In fact, one can easily check geometrically (by using the property a) of conjugate pairs in the previous remark) that for $\mathcal{O}_w$ the line Δ'_v is parametrized by a time-coordinate t_w which is negative decreasing, while t is positive increasing: for $\mathcal{O}_w$, the hypothetic observer $\mathcal{O}'_v$ would be travelling back to the past !

The latter remark strenghtens the meaning of part a) of our fourth postulate and justifies that the cones V^+ and V^- be considered respectively as the *absolute future and past* of the event O. It can now be fully understood that all events represented by points X *outside the union of* V^+ and V^- (like the points of any line Δ'_v) are in *acausal relation* with the event O: no physical signal can propagate either from O to X or from X to O.

2.3 Space-ships' flight: the anniversary curve

So far, we have discovered the conjugate directions of the space and time coordinate axes of all observers in uniform motions, but what remain unknown are the *scales* of time and length along these axes. As a matter of fact, we already see that only the scale of time remains a problem, since once it is known, the scale of length immediately follows from the knowledge of the velocity of light (universal for all uniform motions).

To set this problem of time scaling in an illustrative way, we consider a set of space-ships flying away simultaneously from the same place, let us say at the event O, along the unique horizontal direction Ox, but with various velocities v_i either rightwards $(0 < v_i < 1)$ or leftwards $(-1 < v_i < 0)$ (with units such that $c = 1$); one of them remains at rest $(v_0 = 0)$. All space-ships contain observers $\mathcal{O}_{v_i}$ equipped with identical clocks, and all these observers are invited to celebrate the anniversary of their common departure by representing these events (each anniversary event in the corresponding space-ship) by points correctly situated in spacetime. *On what curve H of the plane (Ox, Ot) will all these points be situated ?*

In the case of Galilean spacetime where time is absolute, the answer to that question is trivial, namely the straight line with equation $t = 1$. In the present framework of spacetime, governed by the five postulates stated in Sec. 2-1, one determines the curve H by showing that it must satisfy the following property

Theorem: For each point X of H, there exists a tangent to H at X whose direction is conjugate of (OX) with respect to the pair (C_R, C_L).

Proof: This result follows directly from the conjugacy property of space and time axes established in Sec. 2-2 together with our *fifth* postulate. In fact, we know that for a given observer $\mathcal{O}_v$ whose world-line $\Delta_v = (OX)$ contains the anniversary event X $(x_v = 0, t_v = 1)$, the straight line of simultaneous events $(t_v = 1)$ is the parallel by X to the conjugate direction of Δ_v; in view of the parallelogram construction, this parallel intersects C_R and C_L in two points M and N such that X is the middle of MN. Now the fifth postulate asserts that for observers $\mathcal{O}'_v$ with velocity v' very close to v (this is what means "with very small relative velocities with respect to $\mathcal{O}_v$") the corresponding anniversary event $X_{v'}$ should be represented *with an excellent approximation* by the Galilean representation of $\mathcal{O}_v$, namely by the point at the intersection of the world-line $\Delta_{v'}$ and of the straight line with equation $t_v = 1$, i.e. (MN). This means that, in mathematical language, the straight line (MN) has to be the *tangent* to the unknown curve H at the point X (see fig.5).

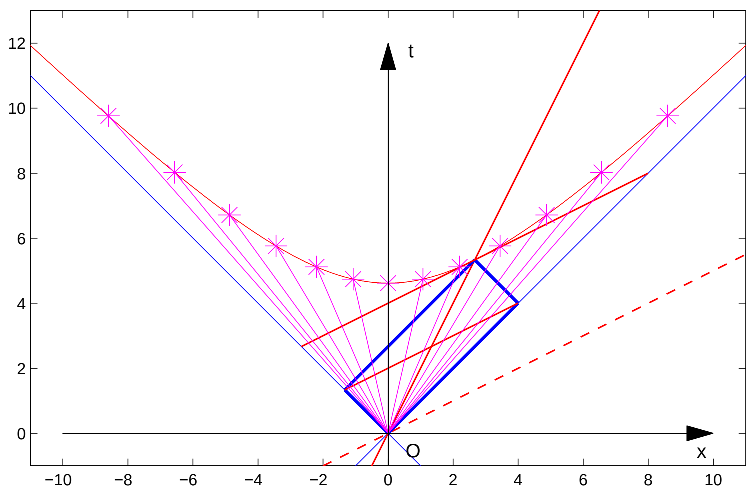

Figure 5: The anniversary curve

Now it is well-known in elementary geometry that *every curve H, whose tangent at each point X intersects two given (nonparallel) straight lines C_R, C_L at two points M, N such that X is the middle of MN is a branch of hyperbola with asymptotes C_R and C_L.*

Since it must contain the anniversary event at rest $X_0 = (x = 0, t = 1)$, the anniversary curve H is therefore uniquely determined as the branch of hyperbola whose equation is $t^2 - x^2 = 1,\ \ t > 0$ (fig.5). The anniversary point $X = X_v$ of any observer $\mathcal{O}_v$ is thus given by the formulae

$$t(v) = \frac{1}{\sqrt{(1-v^2)}}, \quad x(v) = \frac{v}{\sqrt{(1-v^2)}} \quad \text{(where } |v| < c = 1).$$

It is convenient to introduce instead of the velocity v the parameter χ called the *rapidity* which is defined by $v = \tanh\chi$; χ is a *"hyperbolic angle"* which takes all possible values from $-\infty$ to $+\infty$. The previous formulae can then be rewritten equivalently in the following form, which is similar to the angular parametrisation of the circle:

$$t(v) = \cosh\chi, \quad x(v) = \sinh\chi.$$

2.4 Minkowskian (pseudo-)distance and the inverse triangular inequality: the twin "paradox"

From the algebraic viewpoint, the hyperbolae with equation $t^2 - x^2 = a^2$ present strong similarities with the circles with center O and radius R, whose equations are $x^2 + y^2 = R^2$ in orhonormal coordinates. They are *"level curves" of a certain "quadratic form"* $X \rightarrow Q(X)$ (with $X = (x,t)$ or $X = (x,y)$) specified by a second-degree homogeneous polynomial ($Q(X) = t^2 - x^2$ or $Q(X) = x^2 + y^2$)

This mathematical analogy between the hyperbola and the circle admits here a physical counterpart which is very striking. In fact, after the analysis of Sec. 2-3 we naturally come to the idea that our problem of space-ship travellers and its solution are quite comparable to the following very elementary situation in Euclid's planar geometry. Consider walkers equipped with identical graduated rods who start from the same point O along various straight lines and cover the same distance R: they all have reached the circle with center O and radius R. While the latter statement appears trivial to us because of our visual perception of geometry, the former result concerning the "anniversary curve" H tells us that individual time-measurements made by observers in uniform motion or, as one says, *"proper-time measurements"* inform us about the existence of a certain kind of "time-like distance" in spacetime between events related by physical causality. For that "time-like distance" which we shall also call *"Minkowskian distance"*, H appears as a unit level-curve with starting point O and *in the future of* O. Of course all the level-curves of that Minkowskian distance will appear as homothetic hyperbolae centered at O with equation $t^2 - x^2 = a^2$; they are obtained from H by either a dilatation or a contraction scale factor and completion by the "past branches". In fact, each of these hyperbolae contains *two branches* which are distinguished by the sign of t: the branch on which t remains positive (as the anniversary curve H) is contained in the (absolute) future V^+ of O, while the branch on which t remains negative is in the (absolute) past V^- of O: this is the case for the "negative anniversary curve" which is the set of all past events X from which a one-year travel until O is possible via a uniform motion.

In Euclidean space the notion of distance $d(A,B)$ between two points is characterized by the validity of the triangular inequality: $d(A,B) \leq d(A,C) + d(B,C)$, the equality being obtained if and only if the points A, B, C are on the same straight line. This fact is illustrated geometrically by constructing such triangles (ABC) with given side-lengths a, b and c: one just has to check the intersection property of circles with centers A and B, whose sum of radii b and a is larger than $d(A,B) = c$ (fig.6).

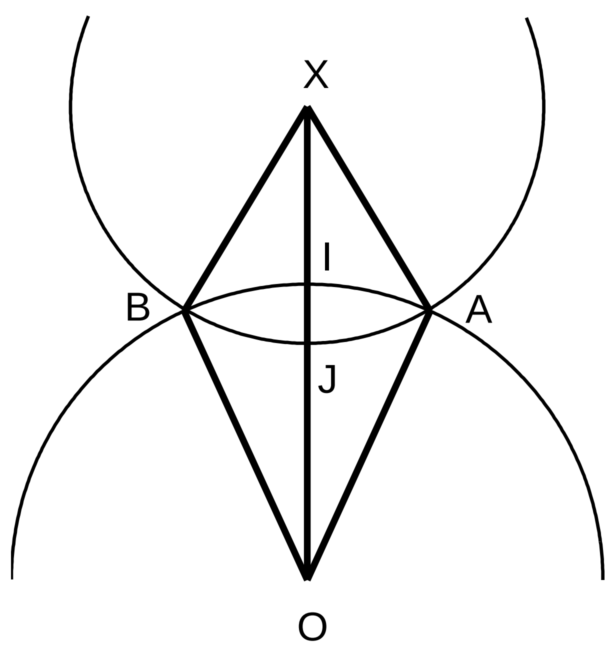

Figure 6: $d(O,X) \leq d(O,A) + d(A,X)$

In the spacetime plane (Ox, Ot), which we shall now properly call *the Minkowskian plane*, a similar geometrical construction shows that there exists again a triangular inequality for the Minkowskian distance d_M, but *with the inverse sign*, namely we have:

Minkowskian triangular inequality: Let three points O, A, X be such that A and X be in V^+, with X in the future of A ($X \in V^+(A)$, then the corresponding Minkowskian distances satisfy the inequality:

$$d_M(O,X) \geq d_M(O,A) + d_M(A,X),$$

the equality being obtained if and only if the points O, A, X belong to the same straight line.

The fact that $d_M(O,X) = d_M(O,A) + d_M(A,X)$ when O, A and X are aligned just expresses the additivity of the corresponding proper time intervals measured by an observer whose world-line is (OAX). Let us now consider the general case when O, A and X form a (non-flattened) triangle. We then consider two branches of hyperbola containing the point A: the first one, called H_O^+ is centered at O and lies in the future cone of O, while the other one, called H_X^- is centered at X and lies in the past cone of X (see fig.7).

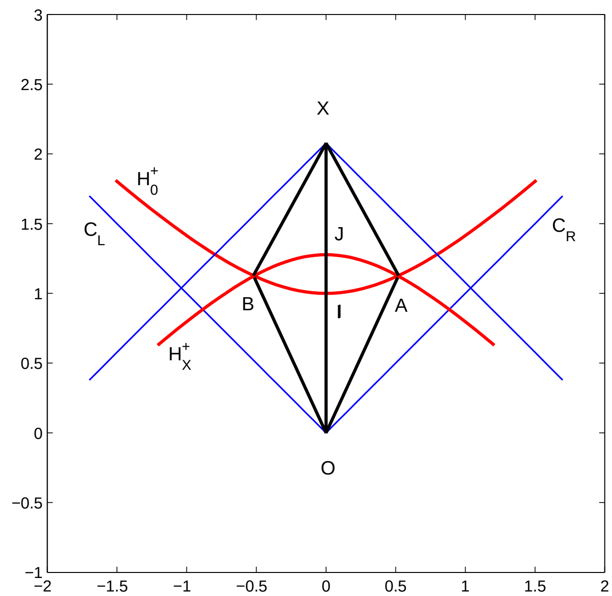

Figure 7: $d_M(O,X) \geq d_M(O,A) + d_M(A,X)$

H_O^+ and H_X^- intersect each other at A and at another point B (such that the straight lines (AB) and (OX) have conjugate directions with respect to (C_R, C_L)). Now the straight line (OX) intersects H_O^+ and H_X^- respectively in two points I and J such that the order of increasing times for the events along (OX) is: O, I, J, X. We therefore have

$$d_M(O,X) = d_M(O,I) + d_M(I,J) + d_M(J,X) \geq d_M(O,I) + d_M(J,X).$$

But since H_O^+ and H_X^- are level-curves for Minkowskian distances we have:

$$d_M(O,I) = d_M(O,A) = d_M(O,B) \text{ and } d_M(J,X) = d_M(A,X) = d_M(B,X),$$

which implies the Minkowskian triangular inequality.

We notice that what makes the difference between the Euclidean and the Minkowskian cases is the *concavity* of the region between one branch of hyperbola and its asymptotes, to be compared with the *convexity* of the region inside a circle.

The "twin paradox"

The physical interpretation of this inverse triangular inequality is the famous "*twin paradox*" of special relativity, which actually is not a paradox once one has got rid of the concept of absolute time, since it expresses in a very illustrative way the content of the Minkowskian geometrical structure of spacetime.

One compares the aging of two persons between two events such as O and X at which they meet together. X can be chosen on the time-axis Ot and one of these persons is supposed to stay on the earth, namely on the world-line (OX). During that time, the other person (which we can imagine in O as the twin of the former) is submitted to a *one-year* travel in uniform motion (with a velocity v which is not small with respect to c) until the event A is reached; then this traveller comes back to the earth with the opposite uniform motion, namely with the opposite velocity $-v$. So two years have past between O and X for the traveller, while the aging of the twin at rest was two years *plus* the time represented by the Minkowskian distance $d_M(I,J)$.

Exercise: Compute $d_M(I,J)$ in terms of $\frac{v}{c}$. In terms of the rapidity χ, one finds that

$$d_M(I,J) = 2(\cosh\chi - 1).$$

What should the value of $\frac{v}{c}$ be equal to in order to produce a shift of one year between the ages of the twins ?

2.5 Spatial equidistance and the "Lorentz contraction" of lengths

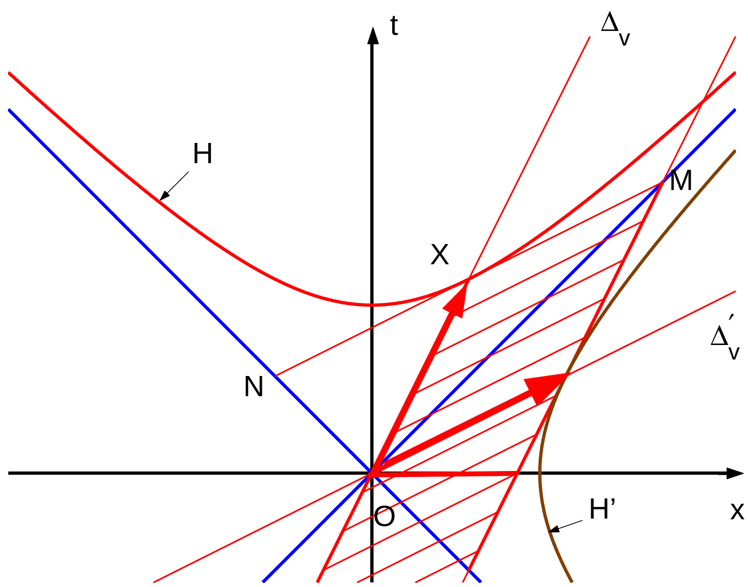

Figure 8: Equidistance curve and "contraction of lengths"

In order to complete the coordinatization of spacetime associated with an observer $\mathcal{O}_v$, we reconsider the anniversary event $X = X_v$ of such an observer, situated at the intersection of the curve H and of the world-line Δ_v. Since the points M and N of the tangent to H at X belong respectively to the light world-lines C_R and C_L and represent events which are simultaneous for $\mathcal{O}_v$ with the time $t_v = 1$, they also define the spatial-distance unit for $\mathcal{O}_v$ in view of our fourth postulate (part b)). One can thus write (with a standard choice of orientation) $M = (x_v = 1, t_v = 1)$, $N = (x_v = -1, t_v = 1)$. This defines the spatial unit vector $[OX_v']$ of $\mathcal{O}_v$ to be such that the quadrilateral $(OX_v'MX_v)$ is a parallelogram (fig.8). OX_v' is thus the unit vector of the space-axis Δ_v' of $\mathcal{O}_v$, conjugate to Δ_v with respect to (C_R, C_L).

The curve of spatial equidistance H': It is clear that the point X_v' is the transform of X_v by the symmetry S_R with axis C_R which exchanges the rest-frame coordinate axes Ox and Ot. This means that if one puts $X_v = (x,t)$ and $X'_v = (x',t')$, then $x' = t, t' = x$. Therefore X'_v belongs to the curve H' with equation

$$t'^2 - x'^2 = -1, \quad x' > 0,$$

which is a branch of hyperbola with asymptotes C_R and C_L, obtained from H by applying the symmetry S_R.

As we shall see in Sec.3-2, the curve H' can be physically interpreted as *the world-line of a uniformly accelerated motion*. What is remarkable is the fact that an observer submitted to that motion always remains spatially equidistant from the fixed event O, since the latter is the center of the hyperbola H'. It even remains *perpetually contemporaneous* of the event O (as this will be fully explained in Sec.3).

Let us now consider the curve H' *completed* by its opposite (from the side $x' < 0$), together with all the homothetic hyperbolae $H'(a)$ with equations $t^2 - x^2 = -a^2$ (taken for all values of a). These are level curves of the Minkowskian quadratic form $Q(X) = t^2 - x^2$ which cover the two regions of spacetime defined by $|t| < |x|$, and respectively $x > 0$ and $x < 0$. These two regions in which $Q(X)$ remains *negative* are called *space-like regions*. Any point X in either one of these regions represents an event which is in acausal relation with respect to O.

The spatial triangular inequality:

The previous construction shows that for any spacelike event X' in a hyperbola $H'(a)$, the usual Euclidean spatial distance $d(O,X')$ between O and X (measured in the system (Δ, Δ') such that $\Delta' = (OX)$) is given by

$$d(O,X)^2 = -Q(X).$$

Let now $(OA'X')$ be a triangle whose three sides have spacelike directions. Then the corresponding (spatial) lengths of these sides satisfy the following Minkowskian triangular inequality

$$d(O,X') \geq d(O,A') + d(A',X').$$

The proof of the latter is immediate by noticing that the symmetric of the triangle $(OA'X')$ with respect to the axis C_R (or C_L) is a triangle (OAX) whose all sides have time-like directions; moreover by construction, the spatial lengths of the sides of the triangle $(OA'X')$ are equal to the Minkowskian (proper-time) distances of the corresponding sides of (OAX). Therefore the triangular inequality for (OAX) (see Sec.2-4) can be transported for $(OA'X')$.

The contraction of lengths:

Another surprising property which results from the Minkowskian geometry of spacetime is the famous apparent contraction of lengths. Here is the argument, which can easily be understood geometrically with the help of fig.8. Consider a one-dimensional rigid body in uniform motion linked with the observer $\mathcal{O}_v$; at the time $t_v = 0$, it can be represented for example as the linear segment $[OX'_v]$ (with unit length for $\mathcal{O}_v$). Then the set of world-lines of all the points of that rigid body generate a strip (in hatchings on fig.8) which is bordered by Δ_v and by the parallel to Δ_v at X'_v. The latter is the tangent to the curve H' at X'_v, which intersects Ox at the point A whose abscissa is $\frac{1}{\cosh \chi} < 1$. It is clear that the passage of the rigid body at time $t = o$ in the rest system occupies the segment $[OA]$: the apparent contraction of length of the moving rigid body is therefore equal to

$$\delta(v) = 1 - \frac{1}{\cosh \chi} = 1 - \sqrt{1 - v^2}.$$

2.6 Lorentz transformations in the Minkowskian plane and two-dimensional Lorentz frames

To summarize the previous constructions, we can say that the light world-lines C_R and C_L separate the Minkowskian (vector) plane with origin O into four angular regions: the future and past time-like regions V^+, V^- are characterized by the *positivity* of the quadratic form $Q(X) \doteq t^2 - x^2$; the spacelike regions by the *negativity* of $Q(X)$. Up to a sign, $Q(X)$ gives the square of the distance between O and X, but this distance is either time-like (measured by a clock) or spatial (measured by a rod). Here is the full meaning of the *non-positive-definite character* of the Minkowskian quadratic form $Q(X)$. In contrast with the Euclidean case, the set defined by the equation $Q(X) = 0$ does not reduce to O but is the union of the light world-lines C_R and C_L: two events separated by the propagation of a light ray have a mutual Minkowskian distance equal to zero.

We are now going to transfer to the Minkowskian plane some basic notions of the Euclidean plane: there is a dictionary between the languages of these two worlds, but also big differences due to the priviledged role of the pair of straight lines (C_R, C_L) in the Minkowskian case. (Note however that in the Euclidean case, a similar structure would also be recovered by a *complexification* of the coordinates: the pair of "isotropic lines" with equations $x = \pm iy$ then plays the same role as the pair (C_R, C_L)).

In the Euclidean vector plane, the elementary notion of angle is complementary to the notion of norm (or distance) in the following sense. The circles centered at the origin O are invariant under the rotations with center O and arbitrary angle θ. These rotations $R(\theta)$ form a commutative group: $R(\theta')R(\theta) = R(\theta + \theta')$. Each system of orthonormal axes (Δ, Δ') is transformed by any rotation $R(\theta)$ into another orthonormal system $(\Delta_{(\theta)}, \Delta'_{(\theta)})$. The corresponding two coordinatizations of the Euclidean plane, denoted respectively by $[OX] = (x, y)$ and $[OX] = (x_{(\theta)}, y_{(\theta)})_\theta$, are such that the Euclidean quadratic form $Q(X)$, identified with the *squared norm* of the vector $[OX]$, is invariant:

$$Q(X) \doteq [OX]^2 = x^2 + y^2 = \ x^2_{(\theta)} + y^2_{(\theta)}.$$

In the Minkowskian vector plane, it is the notion of "rapidity " or "hyperbolic angle" χ which plays the role of the angle θ. One can in fact also introduce a *commutative group* of transformations $L(\chi)$ called *"the Lorentz group in the plane"*; in the spirit of this paper, it is also suggestive to call it "the group of hyperbolic rotations". It acts in such a way that all the branches of hyperbolae centered at the origin with asymptotes (C_R, C_L) are invariant under all the transformations $L(\chi)$. Moreover all the previous statements of the Euclidean case remain valid, if one replaces the pairs of orthonormal axes by *pairs of conjugate axes* (normalized by the curves H and H' as it has been explained above) and if $Q(X)$ now denotes the non-positive-definite Minkowskian quadratic form, or *"squared (pseudo)norm"* of the vector $[OX]$.

The Lorentz group in the plane

One can give an elementary presentation of the action of the transformations $L(\chi)$. These transformations of the plane are linear; so it is sufficient to know their action on two independent vectors OM, ON and convenient to choose the latter *lightlike*, namely along the lines C_R and C_L. We put:

$$L(\chi)[OM] = e^{\chi}[OM], \quad \text{for} \quad M \quad \text{in} \quad C_R,$$

$$L(\chi)[ON] = e^{-\chi}[ON], \quad \text{for} \quad N \quad \text{in} \quad C_L.$$

The lines C_R and C_L (and thereby the set with equation $Q(X) = 0$) are separately conserved by these transformations: in fact, they provide two one-dimensional representations of the multiplicative group ($e^{\chi}e^{\chi'} = e^{\chi+\chi'}$). Now every vector $[OX]$ can be decomposed in the form $[OX] = [OM] + [ON]$, with respect to the pair (C_R, C_L), so that we can define by linearity:

$$[OX_{(\chi)}] \doteq L(\chi)[OX] = \ e^{\chi}[OM] + e^{-\chi}[ON].$$

That means that the coordinates $u(\chi) = e^{\chi} > 0, \ \ v(\chi) = e^{-\chi} > 0$ of the point $X_{(\chi)}$ with respect to the ("light"-)basis $([OM], [ON])$ satisfy the equation

$$u(\chi) \times v(\chi) = 1,$$

which represents a branch of hyperbola with asymptotes $(C_R,\ C_L)$.

Therefore one sees that *all the level curves of $Q(X)$, either in the time-like or in the spacelike regions and including also the light-like world-lines ($Q(X)=0$), are left invariant by the action of all the transformations $L(\chi)$*. This is what we call *the basic geometrical property of the Lorentz transformations.*

One also checks the commutativity property of this group, namely the validity of the relation $L(\chi')L(\chi) = L(\chi+\chi')$ for all χ, χ', which is built-in in the previous definition.

Transforms of conjugate axes

Let us now consider the pair of unit vectors $[OX_0] = (0,1), [OX_0'] = (1,0)$ of the coordinate axes at rest. We will show that each transformation $L(\chi)$ transports this pair into the corresponding pair of unit vectors $[OX_v], [OX_v']$ of conjugate coordinate axes (Δ_v, Δ_v') such that $v = \tanh\chi$. To see this, we introduce the two lightlike vectors $[OM_0] = (\frac{1}{2}, \frac{1}{2})$ and $[ON_0] = (-\frac{1}{2}, \frac{1}{2})$ such that $[OX_0] = [OM_0] + [ON_0]$ and $[OX_0'] = [OM_0] - [ON_0]$. In view of the previous definition of the action of $L(\chi)$, we thus have

$$L(\chi)[OX_0] = \ e^{\chi}[OM_0] + e^{-\chi}[ON_0] = (\sinh\chi, \cosh\chi) = [OX_v],$$

$$L(\chi)[OX_0'] = \ e^{\chi}[OM_0] - e^{-\chi}[ON_0] = (\cosh\chi, \sinh\chi) = [OX_v'].$$

One can also compute similarly the action of another transformation $L(\chi')$ on the pair $([OX_v], [OX_v'])$; it gives another conjugate pair $([OX_w], [OX_w'])$ where $w = \tanh(\chi+\chi')$. In fact one has

$$L(\chi')[OX_v] = (\sinh(\chi+\chi'), \cosh(\chi+\chi')) = L(\chi+\chi')[OX_0] = [OX_w],$$

$$L(\chi')[OX_v'] = (\cosh(\chi+\chi'), \sinh(\chi+\chi')) = L(\chi+\chi')[OX_0'] = [OX_w'].$$

Additivity of rapidities:

The previous computation shows that the action of the commutative group of "hyperbolic rotations" $L(\chi)$ on pairs of conjugate axes (Δ_v, Δ_v') (normalized by H and H') is similar to the action of the group of rotations $R(\theta)$ on pairs of orthonormal axes. A physical interpretation of the latter concerns the composition law of velocities: *the Galilean "law of additivity of velocities" is replaced by the Minkowskian "law of additivity of rapidities"*. If a relativistic particle **A** has the rapidity χ with respect to the earth and emits in the forward direction a particle **B** with rapidity χ' in its center of mass system, then **B** has the rapidity $\chi+\chi'$ with respect to the earth. The corresponding composition law for velocities is therefore:

$$w = \tanh(\chi+\chi') = \frac{\tanh\chi + \tanh\chi'}{1+\tanh\chi\ \tanh\chi'} = \frac{v+v'}{c(1+\frac{vv'}{c^2})}.$$

Lorentz frames and Lorentz invariance of $Q(X)$:

Every vector $[OX] = t[OX_0] + x[OX_0']$ of the Minkowskian plane can be rewritten as

$$[OX] = t_v[OX_v] + x_v[OX_v']$$

for any choice of congugate axes (Δ_v, Δ_v') with unit vectors $([OX_v], [OX_v'])$. We shall also write in short: $[OX] = (x,t) = (x_v, t_v)_v$. Choosing such a coordinatization is also called *"choosing a Lorentz frame with velocity v (or rapidity χ)"* in the Minkowskian plane.

The last point to be checked for completing the parallel between the Lorentz group in the Minkowskian plane and the rotation group in the Euclidean plane is the *"invariance property of the Minkowskian quadratic form by changes of Lorentz frame"*, namely the fact that for any Lorentzian coordinatization $X = (x,t) = (x_v, t_v)_v$, one has the invariance relation

$$Q(X) = t^2 - x^2 = \ t_v^2 - x_v^2.$$

To show this, we associate with $v = \tanh\chi$ the Lorentz transformation $L(-\chi)$ (namely the inverse of $L(\chi)$) which pulls the pair $[OX_v], [OX'_v]$ back to the pair at rest $[OX_0], [OX'_0]$. With every vector $[OX] = (x,t) = t_v[OX_v] + x_v[OX'_v]$ we can then associate its transform $[OX(-v)] \doteq L(-\chi)[OX] = t_v[OX_0] + x_v[OX'_0] = (x_v, t_v)$. Then according to the basic geometrical property of Lorentz transformations, the points $X(-v)$ and X belong to the same level-curve of $Q(x)$, which proves the invariance relation written above.

Change of Lorentz frame in the light-cone coordinatization

For the rest-frame as well as for the Lorentz frame with rapidity χ, it is convenient to introduce the corresponding *light-cone coordinates* of the point $X = (x,t) = (x_v, t_v)_v$, namely

$$(U \doteq t + x,\ V \doteq t - x), \quad (U_v \doteq t_v + x_v,\ V_v \doteq t_v - x_v).$$

In fact if one puts

$$x = a\sinh\psi,\ t = a\cosh\psi; \quad x_v = a\sinh\psi_v,\ t = a\cosh\psi_v,$$

one obtains:

$$U = ae^{\psi},\ V = ae^{-\psi}; \quad U_v = ae^{\psi_v},\ V_v = ae^{-\psi_v}.$$

But we know that $\psi = \psi_v + \chi$ (this is the action of the "hyperbolic rotation with rapidity" χ that has been presented above). One thus obtains the very simple relations

$$\frac{V_v}{U_v} = \frac{V}{U} \times e^{2\chi}.$$

$$(\text{or} \quad \frac{t_v - x_v}{t_v + x_v} = \frac{t - x}{t + x} \times e^{2\chi}.)$$

2.7 The four-dimensional Minkowski's spacetime; tetrads, Lorentz group and Poincaré group

Up to now we have concentrated on relativistic motions along a single direction of space (Ox), which allowed us to construct a two-dimensional section (Ox, Ot) of Minkowski's spacetime and to introduce the corresponding group of Lorentz transformations in this Minkowskian plane.

We shall now show how the geometrical exploitation of the five postulates (stated in Sec.2-1) can be extended so as to construct the full four-dimensional Minkowski's spacetime. This can be performed in three steps:

i) *Use of the rotational symmetry for the observer* $\mathcal{O}_0$*:*

According to our first postulate, the observers at rest can represent each event X as follows:

$$X \doteq (x_1, x_2, x_3, ct) \doteq (\mathbf{x}, ct) \doteq (|\mathbf{x}|\ \mathbf{j},\ ct),$$

where $\mathbf{j}$ denotes a spatial unit vector ($|\mathbf{j}| = \mathbf{1}$) which may serve to indicate a direction of motion. In fact, if we consider uniform motions passing at O with velocity $\mathbf{v} \doteq v\mathbf{j}$ oriented in a given spatial direction $\mathbf{j}$, we can reproduce all the previous considerations (from Sec.2-2 to Sec.2-6) for representing these motions in a Minkowskian plane generated by the axis with unit vector $\mathbf{j}$ and Ot. By analogy with geographical representations of space, such planes can be called *meridian planes of spacetime with respect to the observers at rest*.

So we can say that by rotational symmetry (all the directions $\mathbf{j}$ being equivalent), the union of anniversary curves in all meridian planes generate an "anniversary hypersurface", still denoted by H. This is the set of events $X_{\mathbf{v}}$ reached by all observers $\mathcal{O}_{\mathbf{v}}$ starting together from O towards all possible directions $\mathbf{j}$ of space, after one year has elapsed *at their own clock*. H is a *sheet of hyperboloid* whose equation is

$$(ct)^2 - (x_1^2 + x_2^2 + x_3^2) \doteq (ct)^2 - \mathbf{x}^2 = c^2; \quad t > 0$$

Here we have restored the unit-independent notation including c. We shall generally keep it also in the next sections in order to always exhibit explicitly the physical dimensionality of the quantities involved.

This anniversary hypersurface H can be seen as providing by itself a geometrical characterization of all the uniform motions. In fact, one can say that any pointlike object in uniform motion is characterized by the Minkowskian vector $[OX_{\mathbf{v}}] \doteq cu$ whose tip $(X_{\mathbf{v}})$ belongs to H. Putting $u \doteq (u_1, u_2, u_3, u_0) \doteq (\mathbf{u}, u_0)$, one then has:

$$u^2 \doteq u_0^2 - u_1^2 - u_2^2 - u_3^2 \doteq u_0^2 - \mathbf{u}^2 = 1, \quad \text{with } u_0 > 0.$$

In the latter u^2 denotes what we call *the squared Minkowskian pseudonorm of* u, and u is also called a *timelike unit vector* (Note that the anniversary hypersurface H is now normalized at c).

Equivalently u can be characterized by the pair $(\chi, \mathbf{j})$, where χ is the rapidity (such that $v = c \tanh \chi$) and $\mathbf{j}$ specifies the direction of the motion, according to the following formulae

$$u_0 = \cosh \chi = \frac{1}{[1 - \frac{v^2}{c^2}]^{\frac{1}{2}}},$$

$$\mathbf{u} = \sinh \chi \; \mathbf{j} = \frac{\frac{v}{c}}{[1 - \frac{v^2}{c^2}]^{\frac{1}{2}}} \; \mathbf{j}.$$

This leads one to call *relativistic velocity vector* the Minkowskian vector $cu = (c\mathbf{u}, cu_0)$, since its space-component admits a small-velocity expansion

$$c\mathbf{u} = \mathbf{v}(1 + \frac{v^2}{2c^2}) + \cdots ,$$

which reproduces the velocity vector $\mathbf{v}$ in the first-order Galilean (or "non-relativistic") approximation. The unit vector u can then be called the *"dimensionless" relativistic velocity vector* of the uniform motion.

The same considerations of rotational symmetry lead us to introduce the one-sheeted hyperboloid with equation

$$(ct)^2 - (x_1^2 + x_2^2 + x_3^2) \doteq (ct)^2 - \mathbf{x}^2 = -c^2,$$

which is obtained as the union of all branches of hyperbolae H' in the meridian planes generated by a space axis with unit vector $\mathbf{j}$ and Ot. This hypersurface, still denoted by H' is the set of all points $X'_{\mathbf{v}}$ such that the pair of axes $(\Delta_{\mathbf{v}}, \Delta'_{\mathbf{v}})$ are conjugate with respect to the light world-lines inside the corresponding meridian plane. The (hyper)surfaces H and H' are represented on fig.9.

The Minkowskian quadratic form $Q(x)$

In view of the fundamental role played by H and H', we are led to introduce the following quadratic form on the four-dimensional Minkowski's spacetime:

$$Q(X) \doteq (ct)^2 - x_1^2 - x_2^2 - x_3^2,$$

whose level (hyper)surfaces are described as follows:

a) all the sheets of hyperboloids centered at O which are homothetic to H and lie either in V^+ or in V^-. They correspond to $Q(X) > 0$.

b) all the one-sheeted hyperboloids centered at O which are homothetic to H'. They correspond to $Q(X) < 0$.

c) the light-cone C whose equation is $Q(X) = 0$.

We shall denote by $\hat{H}$ anyone of these level hypersurfaces of $Q(X)$.

Remark We shall use in the following the fact that the restriction of $Q(X)$ to any spacelike hyperplane has level surfaces which are the sections of the previous hyperboloids by that hyperplane:

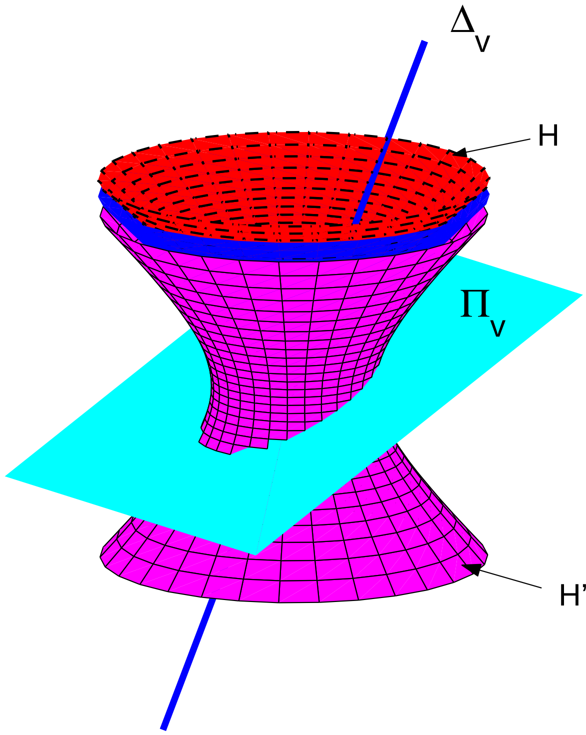

Figure 9: A representation of the four-dimensional Minkowski's spacetime: Level surfaces H, H' of $Q(X)$ and a conjugate pair $(\Delta_{\mathbf{v}}, \Pi_{\mathbf{v}})$

these level surfaces are therefore ellipsoids (exceptionally spheres when the hyperplane is parallel to (Ox_1, Ox_2, Ox_3).)

ii) *Conjugacy properties: the space hyperplanes* $\Pi_{\mathbf{v}}$

The analysis of all simultaneous events with respect to any given observer $\mathcal{O}_{\mathbf{v}}$ can be performed along the same line as in Sec.2-2, even if the geometry is a bit more complicated than in the Minkowskian plane. In fact, the principle is always the same, being based on the second postulate which settles the light-cone C as the primary absolute element of spacetime.

Being given an observer $\mathcal{O}_{\mathbf{v}}$ with world-line $\Delta_{\mathbf{v}}$ (inside the light-cone C) and a space-direction Δ' (i.e. by definition outside the light-cone C), these two straight lines determine a plane P which intersects C along a pair of light-lines. Now we can say that Δ' is *a direction of simultaneity* for $\mathcal{O}_{\mathbf{v}}$ if, *in the plane* P, $\Delta_{\mathbf{v}}$ and Δ' are conjugate with respect to the light-lines of P : that means that by performing the parallelogram construction of Sec.2-2 (fig.3), in the plane P, with $\Delta_{\mathbf{v}}$ as the given diagonal, one obtains Δ' as the direction of the second diagonal. In view of the universality of the light-velocity (fourth postulate) completed again by "isochronousness" (third postulate), this geometrical construction remains the *universal criterion of simultaneity with respect to* $\mathcal{O}_{\mathbf{v}}$. We shall now show the following:

Linearity property: The set of all directions of simultaneity Δ' *for* $\mathcal{O}_{\mathbf{v}}$ *is a three-dimensional linear subspace. This hyperplane* $\Pi_{\mathbf{v}}$ *is physically interpreted as providing the space-slices at constant time* $t_{\mathbf{v}}$ *for* $\mathcal{O}_{\mathbf{v}}$.

Let us show that if Δ'_1 and Δ'_2 are directions of simultaneity for $\mathcal{O}_{\mathbf{v}}$, then any direction Δ' in the plane determined by these two directions is also a direction of simultaneity for $\mathcal{O}_{\mathbf{v}}$. Given the planes P_1 and P_2 determined respectively by $(\Delta_{\mathbf{v}}, \Delta'_1)$ and $(\Delta_{\mathbf{v}}, \Delta'_2)$ and given any point X of $\Delta_{\mathbf{v}}$ in V^+ , one can construct the corresponding parallelograms (OA_1XB_1) and (OA_2XB_2)

whose all sides are light-like segments (as in fig.3 of Sec.2-2) and whose diagonals A_1B_1 and A_2B_2 are respectively parallel to Δ'_1 and Δ'_2 and intersect at the middle of OX. Since the four-points A_1, A_2, B_1, B_2 all belong to the future light-cone C^+, as well as to the past light-cone $C^-(X)$ with apex X, they belong to their intersection which is an ellipse E: A_1B_1 and A_2B_2 are diameters of this ellipse. If we now consider any direction Δ' in the plane determined by Δ'_1 and Δ'_2, which is parallel to the plane of E, we see that the diameter of E parallel to Δ' intersects E in two points A and B such that $(OAXB)$ is a lightlike-sided parallelogram: therefore Δ' is a direction of simultaneity for $\mathcal{O}_{\mathbf{v}}$. This proves that the set of directions of simultaneity for $\mathcal{O}_{\mathbf{v}}$ is a linear subspace of the spacetime. The fact that this subspace is three-dimensional is easy to see: Assuming that it were two-dimensional, it would determine with $\Delta_{\mathbf{v}}$ a three-dimensional subspace S of spacetime outside which no spacelike direction Δ' could be a direction of simultaneity for $\mathcal{O}_{\mathbf{v}}$. But let us then pick up any spacelike direction Δ' outside S. It determines with $\Delta_{\mathbf{v}}$ a plane P' which intersects C along two light-lines and therefore allows one to construct a direction of simultaneity Δ" for $\mathcal{O}_{\mathbf{v}}$ inside P'. Since P' can intersect S only along $\Delta_{\mathbf{v}}$ (if not, it would be contained in S and Δ' would be contained in S), the assumption cannot be true.

To summarize, we have associated with each world-line $\Delta_{(u)}$ with timelike unit vector u , a corresponding spacelike hyperplane $\Pi_{\mathbf{v}} \doteq \Pi_{(u)}$ which can be called the *conjugate hyperplane to* $\Delta_{\mathbf{v}}$. The intersection of $\Pi_{(u)}$ with the one-sheeted hyperboloid H' is an *ellipsoid* $\mathcal{E}_{\mathbf{v}} \doteq \mathcal{E}_{(u)}$ which represents *the set of all events X experienced as simultaneous at zero time and at (lightyear) unit distance from the origin by the observer $\mathcal{O}_{\mathbf{v}}$ with world-line* $\Delta_{(u)}$. This hyperplane and the corresponding ellipsoid are tentatively illustrated on fig.9 (as a plane and an ellipse represented in perspective). It is worthwhile to emphasize that the ellipsoid $\mathcal{E}_{(u)}$ (as well as all the homothetic ellipsoids having their centers on the axis $\Delta_{(u)}$) *are perceived as spheres centered at the origin* by the corresponding observer $\mathcal{O}_{(u)}$.

iii) *Four-dimensional Lorentz transformations, tetrads and the invariant forms of* $Q(X)$

We consider any given conjugate pair $(\Delta_{(u)}, \Pi_{(u)})$ associated with a certain observer $\mathcal{O}_{(u)}$ in uniform motion with relativistic velocity vector u; $[OX_{(u)}]$ is the unit vector of the time-axis $\Delta_{(u)}$ of $\mathcal{O}_{(u)}$. We are looking for coordinatizations of the Minkowskian spacetime *adapted to that observer*. Such coordinatizations can be defined by choosing triplets of unit spatial vectors $[OX'_{(u),1}]$, $[OX'_{(u),2}]$, $[OX'_{(u),3}]$ in the hyperplane $\Pi_{(u)}$, (namely vectors whose tips belong to the ellipsoid $\mathcal{E}_{(u)}$), and by decomposing any vector $[OX]$ of spacetime under the following form

$$[OX] = (ct_{(u)})[OX_{(u)}] + x_{(u),1}[OX'_{(u),1}] + x_{(u),2}[OX'_{(u),2}] + x_{(u),3}[OX'_{(u),3}].$$

However the remaining problem consists in determining all possible triplets $[OX'_{(u),1}]$, $[OX'_{(u),2}]$, $[OX'_{(u),3}]$ such that the Minkowskian quadratic form $Q(X)$ still has the same invariant form with respect to these new coordinates, namely:

$$Q(X) \doteq (ct)^2 - (x_1^2 + x_2^2 + x_3^2) = (ct_{(u)})^2 - (x_{(u),1}^2 + x_{(u),2}^2 + x_{(u),3}^2).$$

In fact, since the restriction of $Q(X)$ to any hyperplane parallel to $\Pi_{(u)}$ has *ellipsoidal level surfaces which must be perceived as spheres by the observer* $\mathcal{O}_{(u)}$, the previous diagonal "Pythagoreanlike" form of $Q(X)$ characterizes the corresponding triplet $[OX'_{(u),1}]$, $[OX'_{(u),2}]$, $[OX'_{(u),3}]$ *an orthonormal system for* $\mathcal{O}_{(u)}$. If this is the case, we shall say that the linear transformation $L_{(u)}$ which transforms the unit vectors of the rest-frame $[OX'_{0,1}]$, $[OX'_{0,2}]$, $[OX'_{0,3}]$, $[OX_0]$, into the *"tetrad"* ($[OX'_{(u),1}]$, $[OX'_{(u),2}]$, $[OX'_{(u),3}]$, $[OX_{(u)}]$), is a *Lorentz transformation of Minkowski's spacetime*. One can also say that *this tetrad is affiliated to the conjugate pair* $(\Delta_{(u)}, \Pi_{(u)})$ or also that it is *admissible for the observer* $\mathcal{O}_{(u)}$. In the usual terminology, each tetrad is also called a *Lorentz frame* when one refers to the corresponding coordinatization of spacetime.

The construction of general Lorentz transformations rely on two basic classes of such transformations which it is easy to describe.

a) *The group $\mathcal{L}_{ort}$ of orthogonal transformations at rest:*

We consider the group of transformations which transform the initial rest-frame into another rest-frame whose spatial axes form a new orthonormal (positively oriented) system of the space (Ox_1, Ox_2, Ox_3), while the time-axis Ot is preserved. Since these transformations preserve the form of the spatial Euclidean distance

$$x_1^2 + x_2^2 + x_3^2 = {x'_1}^2 + {x'_2}^2 + {x'_3}^2,$$

they obviously leave $Q(x)$ invariant.

b) *The group $\mathcal{L}_{hyp}$ of "pure Lorentz transformations":*

Let us fix $\mathbf{j}$ along Ox_1 and the vector $\Delta_{(u)}$ with unit vector $u \doteq u_{(1)}$ in the Minkowskian plane (Ox_1, Ot). Then it is easily checked that the conjugate hyperplane $\Pi_{(u_{(1)})}$ is generated by the conjugate axis $\Delta'_{(u_{(1)})}$ in the plane (Ox_1, Ot) (see fig.8) together with the spatial plane (Ox_2, Ox_3). We then consider the linear transformations which keeps all the vectors in the plane (Ox_2, Ox_3) fixed, and acts as a two-dimensional hyperbolic rotation with rapidity χ in the plane (Ox_1, Ot). This transformation is called a *pure Lorentz transformation* of Minkowski's spacetime. The corresponding change of coordinates is of the form

$$(x_1, x_2, x_3, ct) \to (x'_1, x'_2 = x_2, x'_3 = x_3, ct'),$$

where the passage from (x_1, ct) to (x'_1, ct') has been given in Sec.2-6. It then follows from the invariance property presented at the end of Sec.2-6 that one has:

$$Q(X) \doteq c^2t^2 - x_1^2 - x_2^2 - x_3^2 = c^2{t'}^2 - {x'_1}^2 - {x'_2}^2 - {x'_3}^2.$$

It also results from the study of Sec.2-6 that these transformations form a commutative group.

The most general Lorentz transformations:

In order to construct the most general Lorentz transformation, we shall compose special transformations of the previous groups $\mathcal{L}_{ort}$ and $\mathcal{L}_{hyp}$. We also keep in mind that when such special Lorentz transformations act on any point X of spacetime, the transform remains on the correponding level hypersurface $\hat{H}_X$ of $Q(X)$ passing at X: either on a spherical horizontal section of $\hat{H}_X$ in the former case, or in a hyperbolic section of $\hat{H}_X$ parallel to the plane (Ox_1, Ot) in the latter case.

Now we proceed as follows. Being given any conjugate pair $(\Delta_{(u)}, \Pi_{(u)})$, one can find a transformation L_1 in $\mathcal{L}_{ort}$ which transforms that pair into a pair $(\Delta_{(u_{(1)})}, \Pi_{(u_{(1)})})$, with $u_{(1)}$ in the plane (Ox_1, Ot). (It must transform the unit vector $\mathbf{j}$ of the horizontal component of u into the unit vector of Ox_1). Then there exists a unique transformation L_2 in $\mathcal{L}_{hyp}$ which transforms the pair $(\Delta_{(u_{(1)})}, \Pi_{(u_{(1)})})$ into the pair at rest $(Ot, (Ox_1, Ox_2, Ox_3))$.

Let us now consider an arbitrary transformation L_0 in $\mathcal{L}_{ort}$ and define the composition product

$$L_{(u)} \doteq L_1^{-1} \circ L_2^{-1} \circ L_0.$$

We call $([OX'_{(u),1}], [OX'_{(u),2}], [OX'_{(u),3}], [OX_{(u)}])$ the image by $L_{(u)}$ of the orthonormal system (or "reference tetrad") $([OX'_{0,1}], [OX'_{0,2}], [OX'_{0,3}], [OX_0])$, the last vector $[OX_{(u)}]$ being (by construction) the time unit vector for the given observer $\mathcal{O}_{(u)}$. We then claim that this image is a general admissible tetrad affiliated with the given pair $(\Delta_{(u)}, \Pi_{(u)})$. This can be seen by an argument similar to the one given at the end of Sec.2-6 for the two-dimensional case. With every vector

$$[OX] = (ct_{(u)})[OX_{(u)}] + x_{(u),1}[OX'_{(u),1}] + x_{(u),2}[OX'_{(u),2}] + x_{(u),3}[OX'_{(u),3}],$$

one associates its "pull-back transform"

$$[OX_{pb}] \doteq L_{(u)}^{-1}[OX] = (ct_{(u)})[OX_0] + x_{(u),1}[OX'_{0,1}] + x_{(u),2}[OX'_{0,2}] + x_{(u),3}[OX'_{0,3}].$$

Then since $L_{(u)}^{-1} = L_0^{-1} \circ L_2 \circ L_1$, one can make use of the fact that the successive images X_1, X_2 and finally X_{pb} of X by the sequence of transformations L_1, L_2 and L_0^{-1} remain on the same level hypersurface of $Q(X)$. This entails that

$$Q(X) = Q(X_{pb}) = (ct_{(u)})^2 - x_{(u),1}^2 - x_{(u),2}^2 - x_{(u),3}^2.$$

Conversely, one sees by the same argument that any tetrad admissible for $\mathcal{O}_{(u)}$ is transformed by $L_2 \circ L_1$ into a tetrad admissible for $\mathcal{O}_0$, which is thereby the image by some transformation L_0 in $\mathcal{L}_{ort}$ of the reference tetrad defined by the coordinate axes.

Pseudoorthogonality and the group property of Lorentz transformations

Being given any pair of events X, X' in spacetime, let us define the following symmetric expression

$$[OX].[OX'] \doteq \frac{1}{2}[Q(X+X') - Q(X) - Q(X')] = (ct)(ct') - x_1x_1' - x_2x_2' - x_3x_3',$$

in which the event $X + X'$ denotes the tip of the vector $[OX] + [OX']$. This algebraic expression is similar to the one which defines the scalar product of two vectors $\mathbf{x}$, $\mathbf{y}$ in terms of the squared norms of $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{x} + \mathbf{y}$ in Euclidean space. By analogy, We shall say that the vectors $[OX]$ and $[OX']$ are *pseudoorthogonal* if

$$[OX].[OX'] \doteq (ct)(ct') - x_1x_1' - x_2x_2' - x_3x_3' = 0.$$

It is easy to check that the vectors of the reference tetrad are mutually pseudo orthogonal.

We know that the images of any event X by the transformations L in $\mathcal{L}_{ort}$ or in $\mathcal{L}_{hyp}$ remain on the level hypersurfaces of $Q(X)$. Then it follows from the previous definition that the images of all pseudoorthogonal pairs by all these transformations are pseudoorthogonal pairs. This is therefore also true for all the Lorentz transformation $L_{(u)}$ constructed in the previous paragraph. So by applying this result to the reference tetrad, we conclude that in every tetrad affiliated with any possible conjugate pair $(\Delta_{(u)}, \Pi_{(u)})$, all the vectors of the tetrad are mutually pseudoorthogonal: so for the spacelike triplet in the tetrad, pseudoorthogonality coincides with the Euclidean notion of orthogonality inside $\Pi_{(u)}$, while the pseudoorthogonality of this triplet with respect to $[OX_{(u)}]$ is identical with the property of conjugacy introduced earlier. Taking into account the fact that all the vectors $[OX]$ of a tetrad are unit timelike or spacelike vectors (i.e. such that either $Q(X) = 1$ or $Q(X) = -1$), we can say that all tetrads are *systems of pseudoorthonormal vectors with respect to* Q.

In view of this characteristic property of tetrads, we can thereby conclude that *the action of any Lorentz transformation* $L_{(u)}$ *on any tetrad gives another tetrad.*

It follows that the composition product of two Lorentz transformation $\mathcal{L}_{(u_1)} \circ \mathcal{L}_{(u_2)}$ is another Lorentz transformation (since it transforms the reference tetrad into a tetrad). The definition of inverse transformations being obvious, we conclude that all the transformations $L_{(u)}$ form a group, called the *Lorentz group of the four-dimensional Minkowski's spacetime.*

By adjunction of the translations of space and time, one obtains the more general *"inhomogeneous Lorentz transformations"* which act on any vector $[OX]$ as follows:

$$[OX] \ \rightarrow (L_{(u)}, a)[OX] = L_{(u)}([OX]) + a;$$

in the latter, a denotes a given four-vector which specifies a translation T_a of spacetime. The set of all the inhomogeneous Lorentz transformations form a group which is called the *Poincaré group.*

Remark on the rest-frame and on the distorted appearance of the general Lorentz frames:

We note that among all the conjugate pairs $(\Delta_{\mathbf{v}}, \Pi_{\mathbf{v}})$, one and only one is orthogonal in the usual sense. The familiar choice of this orthogonal pair (e.g. vertical-horizontal) for representing the rest-frame is a manifestation of our biased geometrical perception which priviledges orthogonality and sedentarity. But as in the Galilean case, the observer at rest enjoys no special physical properties with respect to any other observer in uniform motion (that's again the *"principle of relativity"*). So the verticality of the time-axis and the horizontality of space could have been chosen for representing the Lorentz frame of any given uniform motion: there is nothing deep in that choice. One can also say that the Minkowskian representation of the spacetime of special relativity is defined for $\mathcal{O}_0$ (as well as for any observer $\mathcal{O}_{\mathbf{v}}$) *up to the arbitrariness in the choice of the Lorentz frame* or in short *up to a Lorentz transformation*: it is the equivalence class of all these representations. But any chosen representation provides an absolute and faithful description of the events of the universe. Another aspect of all that which deserves to be pointed out again concerns the unavoidable "distorted visual perception" introduced by the conjugacy property. We mean the fact that we have an *ellipsoidal representation* of the surfaces which are actually perceived as *spheres* by observers in uniform motion. Probably the best way for becoming familiar with that strange aspect of the Minkowskian representation consists again in using the metaphor of geographical maps. One can always represent a land on a map equipped with oblique coordinates and different scales of length on the two coordinate axes. That's awkward for our perception, but it remains an absolute and faithful description of the land. In the Minkowskian representation of spacetime, this is the price to pay for having a global geometrical description of the all the "spatial slices", corresponding to all possible observers in uniform motion !!

3 Accelerated motions and curved world-lines

The only motions that have been considered for stating the postulates of special relativity and for constructing Minkowski's spacetime are uniform motions. Their world-lines are oriented straight lines whose direction belongs to the cone V^+ and one also call them *inertial motions* by referring to the fact that no force is acting on a pointlike object whose motion is of that type. Under the name of *accelerated* (or *noninertial) motions* we shall denote the most general type of motion; such a motion is geometrically represented by a curved world-line in Minkowski's spacetime. A curved world-line is smooth if it is an oriented smooth curve admitting at each point a tangent whose direction belongs to V^+. A general world-line can be considered as an oriented union of smooth curved world-line segments. From the physical viewpoint, objects endowed with motions of such a general type are submitted to the action of a time-dependent force and to additional shocks which produce possible discontinuities in the direction of the tangent to the corresponding world-line. Here we shall keep outside the treatment of dynamical problems of special relativity (except for the special case of *uniformly accelerated motions* considered in Sec.3-2). In fact, we shall only concentrate on the kinematical aspects of these motions, which can be presented in terms of the Minkowskian geometry of curved world-lines by pursuing our analogy with Euclid's geometry.

3.1 Curvilinear distances and the slowing down of clocks

Recall on Euclidean space: Let γ be any curved path with end-points A and B in Euclidean space $\mathbf{R}^3$; we suppose it to be smooth or composed of a finite succession of smooth paths. Mathematically, the *length* $d_\gamma(A, B)$ *of the path* γ is defined by the theory of curvilinear integrals as

$$d_\gamma(A, B) = \int_\gamma ds,$$

where ds denotes the Euclidean length element

$$ds = [dx_1^2 + dx_2^2 + dx_3^2]^{\frac{1}{2}}.$$

This theory involves the following ideas:

i) *conceptually*, $d_\gamma(A, B)$ appears as the limit for N tending to infinity of the length d_N of an approximate polygonal path composed of successive small linear paths of equal lengths $\frac{1}{N}$, whose end-points A_j all belong to γ, with $A_1 = A$ and $d(A_{j_N}, B) \leq \frac{1}{N}$. The points A_j can be constructed recursively by the following rule: A_j is at the intersection of γ with the sphere of radius $\frac{1}{N}$ centered at A_{j-1} (and such that $A_j \neq A_{j-2}$).

ii) *physically*, the length of the path γ can be measured by using a flexible graduated ribbon.

iii) *numerically*, the previous curvilinear integral can be computed by introducing any parametrization of the form $\mathbf{x} \doteq (x_1, x_2, x_3) = \mathbf{x}(t)$ of γ, where t is a parameter varying between t_A and t_B, such that $\mathbf{x}(t_A) = A$ and $\mathbf{x}(t_B) = B$. One then has:

$$d_\gamma(A, B) = \int_{t_A}^{t_B} \frac{ds}{dt} dt.$$

The Minkowskian length or "proper time" of a curved world-line:

The previous Euclidean considerations admit a close parallel for curved world-lines in Minkowski's space.

Let γ be any general curved world-line with initial and final events A and B in Minkowski's spacetime $\mathbf{R}^4$: the event B lies in the future of A (namely in the future cone $V^+(A)$). Mathematically, the *Minkowskian length* $d_\gamma(A, B)$ *of the world-line* γ is again defined by the theory of curvilinear integrals as

$$d_\gamma(A, B) = \int_\gamma ds,$$

but ds now denotes the Minkowskian length element or *"proper-time element"*

$$ds = [(c\ dt)^2 - dx_1^2 - dx_2^2 - dx_3^2]^{\frac{1}{2}}.$$

This theory involves the same ideas as in the Euclidean case, but their physical interpretation in terms of time-measurements must now be kept in mind:

i) *conceptually*, $d_\gamma(A, B)$ again appears as the limit for N tending to infinity of the Minkowskian length d_N of an approximate polygonal path. This path is composed of successive small linear paths of equal Minkowskian lengths or *time=like distances* $\frac{1}{N}$, whose end-points A_j all belong to γ, with $A_1 = A$ and $d(A_{j_N}, B) \leq \frac{1}{N}$. The points A_j can now be constructed recursively by the following rule: A_j is at the intersection of γ with the sheet of hyperboloid $H^+_{A_{j-1}}(\frac{1}{N})$ centered at A_{j-1} and whose all points lie in the future of A_{j-1} and at the time-like distance $\frac{1}{N}$ from A_{j-1}: this sheet of hyperboloid is homothetic of the anniversary surface of A_{j-1} with the scaling ratio $\frac{1}{N}$.

ii) *physically*, the *(time-like) length* of the path γ can be measured by using a *clock* which has to be as much unsensitive to accelerations as possible. The fact that atomic clocks satisfy such requirements with a high degree of robustness against strong accelerations has been established experimentally in various works around 1960 (see in particular the article by Sherwin [S]).

iii) *numerically*, the previous curvilinear integral can again be computed by introducing any relevant parametrization of the path γ, but a specially significant parametrization results in a very nice formula due to Einstein.

Einstein's formula for the slowing down of clocks:

One assumes that the events A and B occur at the same point $\mathbf{x}_A = \mathbf{x}_B$ in the rest system, so that physically the path γ may represent any motion starting from $\mathbf{x}_A$ at time t_A and coming back to the same point at time t_B.

Let us now choose precisely the time-coordinate t in the rest system as a relevant parameter for the description of γ; the latter is thus given by a parametrization of the following form:

$$(\mathbf{x}, ct) \doteq (x_1, x_2, x_3, ct) = (\mathbf{x}(t), ct), \quad \text{with} \quad t_A \leq t \leq t_B.$$

One then has:

$$\frac{ds}{dt} = c\left[1 - \left(\frac{dx_1}{cdt}\right)^2 - \left(\frac{dx_2}{cdt}\right)^2 - \left(\frac{dx_3}{cdt}\right)^2\right]^{\frac{1}{2}} = c\left[1 - \left(\frac{\mathbf{v(t)}}{c}\right)^2\right]^{\frac{1}{2}},$$

where $\frac{d\mathbf{x}(t)}{dt} \doteq \mathbf{v}(t)$ represents the *instantaneous velocity* of the motion in the rest-frame at the rest-time t. By plugging the latter expression of $\frac{ds}{dt}$ in the curvilinear integral for $d_\gamma(A,B)$, one thus obtains:

$$d_\gamma(A,B) = c\int_{t_A}^{t_B}\left[1 - \left(\frac{\mathbf{v(t)}}{c}\right)^2\right]^{\frac{1}{2}} dt \leq c(t_B - t_A).$$

This formula thus exhibits the general phenomenon of "slowing down of the clock attached to the world-line γ" with respect to the clock at rest. It provides a quantitative physical formulation of the following geometrical statement (namely the most general form of the Minkowskian triangular inequality):

"IN MINKOWSKI'S SPACETIME, ANY TIME-LIKE STRAIGHT-LINE SEGMENT IS *LONGER* THAN ANY CURVED SEGMENT WITH THE SAME END-POINTS."

Remark The previous computation provides an expression for the *slowing down*

$$\sigma_\gamma \doteq (t_B - t_A) - \frac{1}{c}d_\gamma(A,B)$$

which exhibits a *very simple* first-order approximation at low velocities ($\frac{v}{c}$ small). One gets:

$$\sigma_\gamma = \int_{t_A}^{t_B}\frac{1}{2}\frac{\mathbf{v(t)}^2}{c^2}dt = (t_B - t_A)\frac{v_M^2}{2c^2},$$

where v_M^2 denotes the *mean squared velocity* of the motion with world-line γ between the initial and final times. This formula is remarkably interesting for performing experimental checks of the slowing-down phenomenon, since v_M may for example be related to the temperature of atoms in thermal motion (see [5] and references therein).

3.2 Minkowski's description of accelerations

The instantaneous relativistic velocity vector for a general motion

We have seen in Sec.2-7 that any pointlike object in uniform motion is intrinsically characterized by its *normalized relativistic velocity vector* u, which is a unit vector in the Minkowskian sense: $u^2 \doteq u_0^2 - \mathbf{u}^2 = 1$. We can then pursue the parallel between smooth Euclidean curved lines and Minkowskian world-lines by considering in both cases the notion of *unit tangent vector* $u(X_0)$ at any point X_0 of the line. If the line is parametrized by the length parameter s via a vector equation of the form $X = X(s)$, one then defines $u(X_0)$ at $X_0 = X(s_0)$ by the equation:

$$u(X_0) = \frac{d}{ds}X(s)_{|s=s_0}.$$

In both cases the squared norm or pseudonorm of $u(X_0)$ is equal to 1, since one has in view of the definition of ds^2:

a) in three-dimensional Euclidean space (as an example)

$$u(X_0)^2 = \left(\frac{dx_1}{ds}\right)^2 + \left(\frac{dx_2}{ds}\right)^2 + \left(\frac{dx_3}{ds}\right)^2 = 1.$$

b) similarly in Minkowskian spacetime:

$$u(X_0)^2 = \left(c\frac{dt}{ds}\right)^2 - \left(\frac{dx_1}{ds}\right)^2 - \left(\frac{dx_2}{ds}\right)^2 - \left(\frac{dx_3}{ds}\right)^2 = 1 \quad \text{with} \quad \frac{dt}{ds} > 0.$$

In the latter case, $cu(X_0)$ will be called the *instantaneous relativistic (or Minkowskian) velocity vector* of the motion $(X = X(s))$ at the event X_0. $u(X_0)$ can be called the *dimensionless* instantaneous velocity vector.

The acceleration vector:

According to Minkowski, one defines the acceleration vector $\gamma(X_0)$ at X_0 as

$$\gamma(X_0) \doteq c^2 \frac{du(X(s))}{ds}|s = s_0.$$

In the latter, the normalization factor c^2 ensures the right dimensionality LT^{-2} of acceleration. Then by taking the derivative with respect to s of the equation $u(X(s))^2 = 1$, we obtain the pseudoorthogonality relation

$$\gamma(X).u(X) \doteq \gamma_0 u_0 - \gamma_1 u_1 - \gamma_2 u_2 - \gamma_3 u_3 = 0$$

which is valid for all points $X = X(s)$ of the world-line. In other words:

The Minkowskian acceleration $\gamma(X)$ *is always a spacelike vector which is conjugate to* $u(X)$. The physical interpretation of the latter is that at any event X_0 of the world-line, the vector $u(X_0)$ indicates the corresponding time-axis $\Delta_{(u(X_0))}$ of the traveller, while the acceleration vector $\gamma(X_0)$ is contained in the conjugate hyperplane $\Pi_{(u(X_0))}$, interpreted by the traveller as the Euclidean space at time zero. Then the Euclidean norm of this vector defines the intensity of the acceleration which is felt by the traveller at the event X_0. In view of the sign convention for defining the squared Minkowskian pseudonorm of $\gamma(X_0)$, which is negative, it is given by

$$|\gamma(X_0)| = (-\gamma(X_0)^2)^{\frac{1}{2}}.$$

Uniformly accelerated motions

We shall now present the Minkowskian treatment of *one-dimensional uniformly accelerated motions*. Under this name, we now mean the motions represented by a world-line in a Minkowskian two-dimensional plane (Ox, Ot), whose acceleration's intensity $|\gamma(X)|$ is a constant γ. That means that the tip of the spacelike vector $\gamma(X)$ varies on a branch of hyperbola centered at O and homothetic either to the curve H' or to its opposite in that plane (see Sec.2-5).

We will check that *all such branches of hyperbolae* together with those obtained from the latter by spacetime translations *are themselves the world-lines of uniformly accelerated motions*. (For simplicity, we shall skip the proof of the fact that they represent *all* the one-dimensional uniformly accelerated motions). We introduce such hyperbolic world-lines by the following parametrization in which the parameter τ will be seen to be the proper time of the motion (the notation τ being thus substituted to the length notation $s = c\tau$ of the previous paragraph).

$$X = X(\tau) \doteq \quad (x(\tau), ct(\tau)) :$$
$$x(\tau) = a \cosh \frac{c\tau}{a} + x_0, \quad ct(\tau) = a \sinh \frac{c\tau}{a} + ct_0.$$

We just have to compute successively:

$$u(X(\tau)) = \frac{d}{d(c\tau)} X(\tau) = (u_1(\tau), u_0(\tau)) :$$

$$u_1(\tau) = \sinh \frac{c\tau}{a}, \quad u_0(\tau) = \cosh \frac{c\tau}{a},$$

which shows that $u(X(\tau))^2 = 1$.

$$\gamma(X(\tau)) = c^2 \frac{d}{d(c\tau)} u(X(\tau)) = (\gamma_1(\tau), \gamma_0(\tau) :$$

$$\gamma_1(\tau) = \frac{c^2}{a}\cosh\frac{c\tau}{a}, \quad \gamma_0(\tau) = \frac{c^2}{a}\sinh\frac{c\tau}{a},$$

from which it follows that $\gamma(X(\tau))^2 = -\frac{c^4}{a^2}$ is constant and yields the value $\gamma = \frac{c^2}{a}$ for the acceleration. So one can say that the acceleration is proportional to the "time-curvature" $\frac{1}{a}$ of the world-line.

Remarks

a) Non-relativistic (or Galilean) approximation: It is clear that the hyperbolic world-line with equation $x^2 - (ct)^2 = a^2$ or $x = [(ct)^2 + a^2]^{\frac{1}{2}}$ admits as a second-order approximation near the event $x = a, t = 0$ the familiar parabola with equation

$$x = a + \frac{c^2}{2a}t^2 = a + \frac{1}{2}\gamma t^2$$

,

b) In Euclidean geometry, the "osculating circle" at a point X of a Euclidean curve is obtained as the limit of the circle containing three neighbouring points of the curve, when these three points tend together to X. Minkowski introduced similarly (in[M]) a notion which can be called the "osculating uniformly accelerated motion" of a general motion at the event X: its world-line is the limit of a hyperbolic world-line containing three neighbouring events of the general motion, in the limit when these three events tend to X.

3.3 A comfortable trip for the "Langevin traveller"

The standard presentation of the "twin paradox" (or "Langevin traveller"), which amounts to a direct trip with return between a point of the earth and some far-distant space station S, with large uniform velocity v in both directions, is remarkable by its beautiful pedagogical simplicity. In fact, we have seen in Sec.2-4 that it exactly illustrates what we called in geometrical terms the Minkowskian triangular inequality. However, since it appeared in the literature, various objections have been raised whose point was generally to conclude that this was a school example, which was probably physically incorrect or at best unrealistic. This type of opinion has also been often endorsed by vulgarizers of special relativity, as a reassuring thought with respect to what looks like a scandal for the common sense.

The main objection was about the instantaneous passage from velocity v to velocity $-v$ when reaching the term of the travel. Such passage had to be produced by a shock, or even if smoothened by some decelerating device, it seemed to involve so large accelerations that certainly the biological organisms and maybe the clocks themselves could not stand such constraints. Now in view of Minkowski's study of uniformly accelerated motions (presented above in Sec.3-2), one can actually show the possibility of organizing a more comfortable trip for the Langevin traveller, in which the latter would be submitted to a *constant acceleration (or deceleration)* We even impose (for making the acceleration biologically normal) that its value be precisely equal to the value of the gravity acceleration g on the earth. Of course, we admit that the whole travel will take place in the vacuum, far from any gravitational source, in such a way that the flat Minkowskian spacetime remains a reasonably good approximation to the real spacetime.

After having specified an appropriate class of world-lines for that space-traveller, the problem, which is purely geometrical, consists in comparing the length of proper time τ (namely the timelike Minkowskian length) of the traveller's world-line with the corresponding time t that will have elapsed on the earth between the traveller's departure and return. A table of the corresponding values of τ and t will be given below. Its result is striking: while the maximal value of τ fits with a reasonably long life-time for a human being (let us say eighty-six years), the corresponding value of t reaches about five billions of years, namely the age of the earth !!

Of course, a second problem (which has a touch of dream as in anticipation novels...) concerns the production of the constant acceleration for the spaceship on which the traveller is going to live. If the acceleration is produced by either expelling or disintegrating a mass of matter aboard the spaceship, as in conventional rockets, one can make a simple computation of the minimal mass consumption based on the relativistic law of energy-momentum conservation (see Sec.5 below). Assuming that all the disintegrated mass is transformed into photons (which is the most favourable process) it is possible to compute the ratio between the remaining mass $M(\tau)$ at proper time τ and the initial mass M_0 of the spaceship. The set of values which are listed in the table indicate that that for τ larger than twenty years, the procedure becomes radically unrealistic. In fact, the mass to be loaded aboard the spaceship then becomes a non-negligible fraction of the mass of the earth (which also means that gravitational effects have to be taken into account; the use of flat Minkowski's spacetime is no longer justified). But the limitations of this procedure do not exclude the consideration of other types of possible propulsions, which could make use for instance of energies available in the cosmic medium.

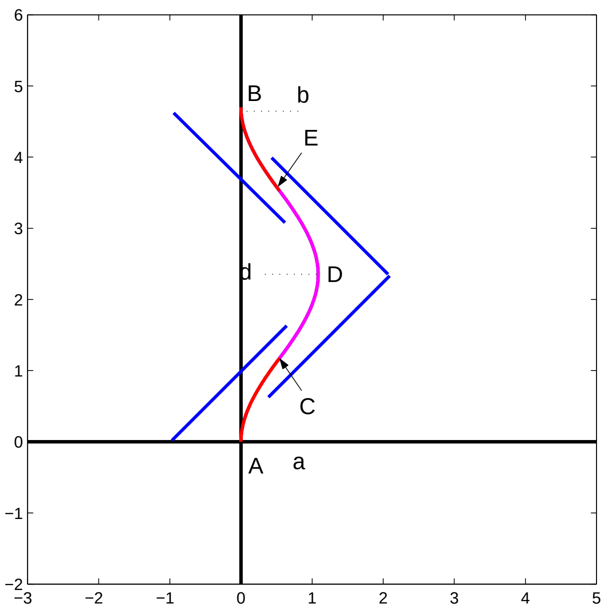

Figure 10: A comfortable motion for the Langevin traveller

Choice of the motion:

The trajectory is along a straight line joining the earth, denoted by S_0, and the space station S considered as at rest with respect to the earth. The travel which is proposed is composed of

i) a uniformly accelerated motion with acceleration g from S_0 to the middle M of S_0S;

ii) a uniformly accelerated motion with acceleration $-g$ from M to S (namely a phase of deceleration);

iii) the acceleration $-g$ is maintained as in ii) and produces half of the returning trip from S to M;

iv) the acceleration is shifted from $-g$ to g for producing a uniformly decelerated motion from M to S_0.

It is clear that the discontinuity of the acceleration (from g to $-g$) produced at M is bearable by the physical and biological systems in the spaceship: if g is equal to the value of the gravity acceleration on the earth, it is just felt as a sudden inversion of the direction of gravity.

The spacetime representation of this motion is a worldline composed of three successive arcs of hyperbolae with centers a, d and b (see fig.10), namely:

i) an arc AC joining the departure event A on S_0 to the end of the acceleration phase C at the point M; this arc is parametrized by the proper time τ of the spaceship according to the following equations:

$$x = \frac{c^2}{g}(\cosh \frac{g}{c}\tau \ - 1), \quad t = \frac{c}{g} \sinh \frac{g}{c}\tau.$$

ii) an arc CDE where D denotes the passage on S (no stop being expected there) and the end-point E denotes the passage at M in the way back.

iii) an arc EB representing the last deceleration phase whose end-event B represents the arrival on S_0.

As it is visualized on fig.10, the arcs CD, DE, and EB of the traveller's worldline are obtained from the arc AC by obvious symmetries and it is clear that the total traveller-time length τ_B of the travel as well as the corresponding earth-time length t_B are respectively equal to four times the traveller-time length τ_C and the corresponding earth-time length t_C that have elapsed between A and C. In view of the equations of AC this yields the following relation between t_B and τ_B:

$$t_B = 4\frac{c}{g} \sinh \frac{g}{c}\frac{\tau_B}{4}.$$

It is pleasant to notice that with our choice of units (i.e. years and lightyears) not only $c = 1$ but also the earth's value of g is very close to 1. We thus obtain the very simple formula

$$t_B = 4 \sinh \frac{\tau_B}{4}$$

whose numerical application can be found in the table.

We notice that for small values of the travel's length of time τ_B, namely between zero and four years, the corresponding values of the earth-time length t_B is not very different; this is because τ_B is the first-order approximation of $4 \sinh \frac{\tau_B}{4}$ at small τ_B. But for larger travel's lengths of time, the exponential character of the sinh function becomes preponderous, which yields such overwhelming discrepancies as two-thousand years of earth's time for twenty-eight years of travel's time and... geologicallike ages for seventy years of travel's time !

Mass decrease required for the spaceship's propulsion

The equation for the rate of mass decrease will be fully justified in Sec.5 on the basis of the relativistic law of conservation of the total energy-momentum of the system. This equation is

$$\frac{d}{d\tau}M(\tau) = -\frac{g}{c}M(\tau) = -M(\tau),$$

which therefore yields the formula

$$M(\tau_B) = M_0 e^{-\tau_B}$$

illustrated numerically in the table.

4 On the visual appearance of rapidly moving objects: Lorentz contraction revisited

Although being valid as a two-dimensional geometrical property of Minkowski's spacetime in a plane (Ox, Ot) , the property of "contraction of lengths" described in Sec.2-5 differs from what would actually be seen by an observer (or a camera) at the passage of a rapidly moving object. As a matter of fact, according to the original Terrell's work [6] (see also [**?**, W1] the analysis of the actual physical phenomenon can be summarized as follows.

i) Even if the moving object S is one-dimensional, namely is an infinitely thin rod alined with the motion trajectory Ox (as considered in Sec.2-5), one must consider the observer at rest $\mathcal{O}$ as situated *at a certain distance* d of Ox. Therefore the actual visual appearance of the rod for such an observer at a certain time $t = t_0$ is obtained by determining the set of light world-lines

The "Langevin traveller" in uniformly accelerated motion

traveller's proper time τ (in years)	earth's proper time t (in years)	$\frac{M(\tau)}{M_0}$
1	1 and 4 days	0.37
2	2 and 1 month	0.13
4	4.7	0.02
8	14.5	4×10^{-4}
12	40.1	8×10^{-6}
16	104	1.6×10^{-7}
20	297	3.2×10^{-9}
28	$2,200$	1.3×10^{-12}
32	$5,960$	
40	$44,000$	
48	$326,000$	
60	6.54×10^6	
72	131×10^6	
84	2.64×10^9	
86	5 billions	

which have been emitted from all the points of the rod in the past of t_0 and which converge at the corresponding "reception event" $\mathcal{O}(d, t_0)$ of the observer $\mathcal{O}$. This determines the "photograph" of the rod at time t_0. When the value of t_0 varies, the geometrical construction of the relevant light world-lines results in modifications of the direction of observation and of the apparent length of the rod; these modifications of the visual appearance of the object for the observer $\mathcal{O}$ at rest will thus accompany the motion of the object. In other words, the aspect of the rod on the photograph will vary with time *by combining the relativistic property of contraction of lengths together with perspective effects*; the latter are comparable to those which occur in ordinary space when changing the direction of observation in order to catch successive situations of the moving object (in a purely Galilean treatment with infinite light-velocity).

ii) The previous type of analysis being taken into account, a more realistic study still has to be done for the case of a three-dimensional object. For instance, it is interesting to consider a cube-shaped or spherical object S whose center moves along the axis Ox and whose size may be considered as small with respect to the distance d from the observer to Ox. It turns out that the visual appearance of such thick objects *never exhibits* the phenomenon of contraction of lengths in the direction Ox as it was pictured in Gamov's famous book ("The adventures of Mr Tompkins in the land of special relativity"). As a matter of fact, the observed appearance of an object at successive times exhibits a perspective effect whose corresponding ("virtual") direction of observation is *shifted* with respect to the real direction of observation, as though the perspective were accompanied by an "anomalous rotation effect". This apparent change of direction of observation is a typical geometrical effect of Minkowski's spacetime: it is characterized by an angle called *"the relativistic aberration"*. It is interesting to note that for the special case of a spherical object, the disk-shaped appearance remains for all the directions of observation which accompany the object's motion.

The relativistic aberration:

Let S and O represent two given events of spacetime corresponding respectively to the emission of a light beam by a pointwise object and to the reception of this light beam: O belongs to the future light-cone $C^+(S)$ of S. The object is in uniform motion with respect to the rest-frame of an observer $\mathcal{O}$ who will observe the reception event at O. This uniform motion is characterized by its world-line $\Delta_{(u)}$ which we choose to belong to the plane (Sx, St) (the point S is contained in $\Delta_{(u)}$; it now plays the role of the origin of Minkowski's spacetime, called O in Sec.2). χ and $v = \tanh \chi$ will denote the rapidity and velocity of the motion; d denotes the distance from the observer $\mathcal{O}$ to the motion's line Sx of the object. At O, the light beam coming from the object is received by the observer $\mathcal{O}$ from a direction which includes the angle θ with the axis Sx in the coordinate-plane (Sx, Sy) and t_O denotes the corresponding reception time.

From these data, we can express the coordinates of the reception event O in the rest frame as follows

$$(x = ct_O \ \cos\theta,\ y = d = ct_O \ \sin\theta,\ z = 0,\ t = t_O).$$

(Note that in all the argument the scenario remains in the three-dimensional spacetime (Sx, Sy, St)).

With [6] we now introduce another observer $\mathcal{O}'$ who is *at rest in the frame of the moving object* and whose world-line (parallel to $\Delta_{(u)}$) contains the point O: that means that this moving observer $\mathcal{O}'$ "shares with $\mathcal{O}$" the reception event O of the light beam emitted by the object at S, although the latter is now seen as "at rest" by $\mathcal{O}'$. At this event O, $\mathcal{O}'$ receives the light beam from a direction which includes the angle θ' with the corresponding space-axis Sx' of the object's Lorentz frame: this axis Sx' is conjugate of $\Delta_{(u)}$ in the plane (Sx, St). The space hyperplane $\Pi_{(u)}$ of $\mathcal{O}'$ is in fact generated by the three axes Sx', Sy, Sz, the coordinates y and z being unchanged with respect to those of the rest-frame of $\mathcal{O}$.

We can now express the coordinates of the reception event O in the Lorentz frame of $\mathcal{O}'$ as follows

$$(x' = ct'_O \ \cos\theta',\ y = d = ct'_O \ \sin\theta',\ z = 0,\ t' = t'_O).$$

It is the difference $\alpha \doteq \theta' - \theta$ which is called the *relativistic aberration* and the basic computation which remains to be done is to compute θ' and thereby α as a function of θ and of the rapidity χ (or velocity v) of the object.

Comparing the two representations of O leads one to introduce at first the ratio

$$M \doteq \frac{t_O}{t'_O} = \frac{\sin\theta'}{\sin\theta}.$$

We now make use of the formulae for the change of Lorentz frames in the light-cone coordinates (see the end of Sec.2-6). Applying this formula to the event O (or more properly to its projection onto the plane (Sx, St)) by putting

$$U = ct + x,\ V = ct - x,\ \ U' = ct' + x',\ V' = ct' - x',$$

$$\text{we get}: \ \frac{V}{U} = \tan^2\frac{\theta}{2},\ \ \frac{V'}{U'} = \tan^2\frac{\theta'}{2},$$

$$\text{which yields}: \ \ \tan\frac{\theta'}{2} = \tan\frac{\theta}{2} \times e^{\chi}.$$

The latter relation defines a function $\theta' = \theta'(\theta, \chi)$ which enjoys the following properties:

a) for fixed χ, θ and θ' tend together either to zero or to infinity;

b) for $\theta = \frac{\pi}{2}$ (resp. $\theta' = \frac{\pi}{2}$), one has $\sin\theta' = \frac{1}{\cosh\chi}$ (resp. $\sin\theta = \frac{1}{\cosh\chi}$), where $\frac{1}{\cosh\chi} = (1 - \frac{v^2}{c^2})^{\frac{1}{2}}$ is the *Lorentz contraction factor* (see Sec.2-5).

The visual appearance of extended objects

Let us now suppose that the moving object is *extended* instead of being pointwise, but that *its extension is small* with respect to the distance d at which the observer $\mathcal{O}$ is standing, and to begin with, that it is *"flat for the observer $\mathcal{O}$"* (and therefore also for $\mathcal{O}'$): that means that the set of its world-lines form a small cylinder parallel to $\Delta_{(u)}$ in the subspace (Sx, Sy, St); there is no extension in the third direction Sz.

We now consider the small angles $d\theta$ and $d\theta'$ subtended by the object, as they are seen from O respectively by the observers $\mathcal{O}$ and $\mathcal{O}'$, namely in the planes respectively parallel to (Sx, Sy) and (Sx', Sy). It is clear that the relation between these two angles is obtained by differentiating (at fixed χ) the previous relation between θ, θ' and χ. The result is:

$$\frac{d\theta'}{d\theta} = \frac{\sin\theta'}{\sin\theta} = M$$

This ratio M of the subtended angles, or of the apparent dimensions of the object when passing from the observer $\mathcal{O}$ to the observer $\mathcal{O}'$, can thus be called the *magnification*. What is remarkable in that relation between $d\theta$ and $d\theta'$ is that (eventhough θ' is a function of θ and χ), it does not depend explicitly of the rapidity χ.

As a matter of fact, one can even give a still nicer interpretation of it by introducing the distances r and r' at which the (small) object is seen respectively by $\mathcal{O}$ and $\mathcal{O}'$. Concerning r' it is of course a fixed distance, since the object is at rest for $\mathcal{O}'$ and one has (in the plane parallel to (Sx', Sy) by O)

$$d = r'\sin\theta'.$$

Concerning r, it is the distance from $\mathcal{O}$ (in the plane (Sx, Sy)) of the position occupied by the object at the emission event S and one thus also has

$$d = r\sin\theta.$$

It then immediately follows from these relations that one has:

$$rd\theta = r'd\theta',$$

which means that *the dimensions of the object transversally to the directions of observation of $\mathcal{O}$ and $\mathcal{O}'$ are equal.* One can now see very simply that *a similar result is valid for general small objects having also an extension in the direction Sz.* In fact, the component along Sz of the object is the same in the rest frame as in the Lorentz frame where the object is at rest; it therefore has equal transversal extensions $dz = dz'$ along Oz for both observers $\mathcal{O}$ and $\mathcal{O}'$, which entails:

$$r d\theta dz = r' d\theta' dz'$$

This means that the surface transversal dimensions of the object with respect to the directions of observation of $\mathcal{O}$ and $\mathcal{O}'$ are equal: the *perspectival shapes and dimensions* of the object are the same when the object is moving as when it is fixed, *provided one replaces the actual direction of observation of the moving object, namely the angle θ, by the "virtual" direction $\theta' = \theta'(\theta, \chi)$ corresponding to its observation as a fixed object.*

However, in view of the different distances r and r' from $\mathcal{O}$ and $\mathcal{O}'$ to the object, this identity of the perspectival shapes and dimensions is *modified from the angular viewpoint by the magnification factor*

$$M = \frac{d\theta'}{d\theta} = \frac{r}{r'},$$

whose expression as a function of θ and χ is:

$$M(\theta, \chi) = \frac{\sin\theta'(\theta, \chi)}{\sin\theta}.$$

This transformation can be seen as a certain *conformal mapping* on the unit sphere for regions of small subtended solid angle.

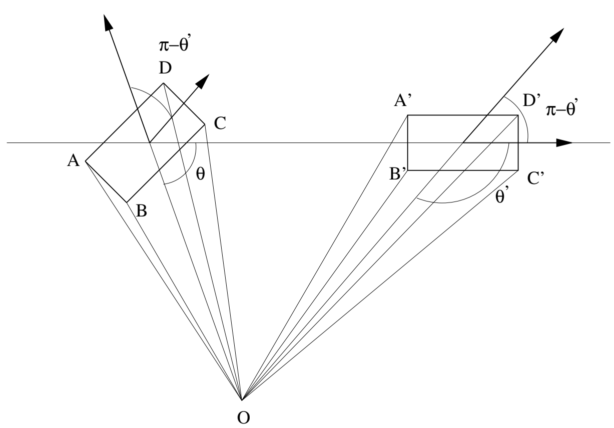

Figure 11: Passage of a "relativistic bus": The relativistic aberration and the apparent rotation

Practical Geometrical Construction:

In order to represent how the object with rapidity χ is seen by the observer $\mathcal{O}$ from the direction with angle θ, one determines the direction with angle $\theta' = \theta'(\theta, \chi)$ from which it is seen by $\mathcal{O}'$ as a fixed object. One then applies to the fixed object (with its true dimensions) a rotation with angle $\alpha = \theta' - \theta$ (i.e. the relativistic aberration) before settling it at the point where $\mathcal{O}$ expects to see it from the direction θ. This is the *correct perspective* under which $\mathcal{O}$ will see the object from that direction. This procedure has been illustrated on fig.11 by taking the example of a "relativistic bus". It is now clear that if the object is spherical, its disk-shaped appearance and dimension is preserved for all possible directions of observation.

What about the "hidden Lorentz contraction" ?

Having obtained the previous general result, let us come back to our very first case of an infinitely thin rod with length l, oriented along Sx and moving along Sx with rapidity χ. Assume that one fixes $\theta = \frac{\pi}{2}$, which means that the observer $\mathcal{O}$ at rest looks at the rod from the direction Sy where he or she is sitting. By refering to the geometrical argument of Sec.2-5, one easily checks that in that case the observer *does* observe a Lorentz contracted rod with apparent length $\frac{l}{\cosh\chi}$. Now let us look at it from the viewpoint of the general result. This rod is seen by $\mathcal{O}'$ as a fixed rod from a direction defined by the angle θ' such that $\sin\theta' = \frac{1}{\cosh\chi}$ (see above the property b) of the function $\theta'(\theta,\chi)$ Then by applying the previous Practical Geometrical Construction, one sees that the observer $\mathcal{O}$ must see the rod as if it were rotated by the angle $\alpha = \theta' - \frac{\pi}{2}$, so that its perspectival length is

$$l \times \sin\theta' = \frac{l}{\cosh\chi},$$

the corresponding angle subtended by the object being equal to $\frac{l}{d}\frac{1}{\cosh\chi}$.

The rotation has exactly reproduced the Lorentz contraction !!

Observing the object without perspective effect

In the Galilean treatment (with infinite velocity of light), the object is observed by $\mathcal{O}$ without perspective effect when the direction of observation is perpendicular to the line of motion, namely when $\theta = \frac{\pi}{2}$. In the case of Minkowski's spacetime, the corresponding phenomenon is obtained when $\theta' = \frac{\pi}{2}$, namely when the observer $\mathcal{O}'$ sees the object without perspective effect. Then the identical effect is obtained by $\mathcal{O}$ provided his or her direction of observation includes an angle θ_0 with the motion's axis. According to property b) of the function $\theta'(\theta,\chi)$, this angle θ_0 is such that

$$\sin\theta_0 = \frac{1}{\cosh\chi}.$$

For the case of the infinitely thin rod, we see that it appears to the observer $\mathcal{O}$ *with its exact length* l when looked at in that direction, but from the angular viewpoint the subtended angle remains (because of the "magnification factor") $\frac{l}{d}\frac{1}{\cosh\chi}$ ... i.e. the same as for the Lorentz contracted appearance at $\theta = \frac{\pi}{2}$!

In conclusion, the effects of *perspective modified by the relativistic aberration, which acts as a rotation,* are clearly defined for describing the visual appearance of moving objects of general shape. The concept of "Lorentz contraction", although perfectly clear in two-dimensional spacetime, then becomes hidden as far as the observation of three-dimensional objects is concerned; it may be restored in the special case of thin objects, but the term is of subtle use and semantically confusing...

5 The Minkowskian energy-momentum space: E=Mc^2 and particle physics

In the Newtonian dynamics, based on the Galilean conception of spacetime, one introduces for each massive pointlike object with mass m and constant velocity $\mathbf{v}$ its momentum $\mathbf{p} = m\mathbf{v}$. For any isolated dynamical system composed of such objects, their velocities and momenta depend on time, but the total momentum, namely the vector sum $\mathbf{P}$ of all the corresponding momenta, must be conserved at all times. The other quantities which have to be conserved at all times are a) the total energy E of the system, and b) the masses of the various objects, since the latter are supposed to conserve their individualities for all times.

In the framework of special relativity, each massive pointlike object with mass m in uniform motion is now characterized by its relativistic (or Minkowskian) velocity vector cu. According to Einstein"s beautiful idea, one can now associate with it a relativistic four-momentum vector $p = mcu$, which can be represented in the coordinates of the rest-frame as follows:

$$p = (\mathbf{p}, p_0); \quad \mathbf{p} = mc\sinh\chi\,\mathbf{j} = m\mathbf{v}[1-\frac{v^2}{c^2}]^{-\frac{1}{2}}, \quad p_0 = mc\cosh\chi = mc[1-\frac{v^2}{c^2}]^{-\frac{1}{2}}.$$

The (tip of the) vector p thus belongs to the upper sheet of hyperboloid H_m^+ with equation

$$p_0^2 - p_1^2 - p_2^2 - p_3^2 = m^2c^2, \quad p_0 > 0.$$

The space-component $\mathbf{p}$ of p admits a small-velocity expansion of the following form

$$\mathbf{p} = m\mathbf{v}(1 + \frac{v^2}{2c^2}) + \cdots ,$$

which therefore reproduces the Newtonian momentum $m\mathbf{v}$ at the first-order approximation. As for the time-component p_0, its small-velocity expansion gives

$$p_0 = mc[1 - \frac{v^2}{c^2}]^{-\frac{1}{2}} = mc(1 + \frac{v^2}{2c^2}) + \cdots .$$

Multiplying both sides of the latter by c in order to get the dimensionality of an *energy*, i.e. $\mathrm{ML^2T^{-2}}$, one then obtains:

$$p_0 c = mc^2 + \frac{1}{2}mv^2 + \cdots .$$

While the second term of this expansion is clearly identified as the kinetic energy of the massive object in the Newtonian formalism, the first term $E_0 = mc^2$ *is the "internal energy at rest" of the massive object, identified (up to the dimensionality factor c^2) with its mass m.* In fact, when the velocity v vanishes, the four-vector pc is along the time-axis and reduces to its time-component $E_0 = mc^2$. One can then also say that for an arbitrary uniform motion with velocity $\mathbf{v}$, the time-component p_0c of the four-vector pc is the *complete relativistic energy of the moving object*, whose value is

$$E \doteq p_0 c = mc^2[1 - \frac{v^2}{c^2}]^{-\frac{1}{2}} = |\mathbf{p}|\frac{\mathbf{c^2}}{\mathbf{v}}.$$

This is why the four-momentum vector pc or p is also called the *energy-momentum vector* of the object (the identification being often made, in view of the convenient choice of units such that $c = 1$).

Remark It is very important to note that in units where $c = 1$, the squared mass $m^2 = (mc)^2$ of the object is equal to the squared pseudonorm of the four-momentum vector p. It is therefore (like the proper time of a motion) a *relativistic invariant*: its value is independent of the Lorentz frame which has been chosen for describing the object.

Massive and massless free particles

In microphysics, the theoretical treatment of particles requires a quantum-mechanical framework. However, this framework makes use basically of the Minkowskian space of four-momenta of point-like massive objects that we have just described. As a matter of fact, the quantum elementary particles *with mass m* are described as "wave-packets" which are probabilistic superpositions of *"classical" four-momentum configurations* $p = (\mathbf{p}, p_0)$ satisfying the so-called *"mass shell" condition*:

$$p \text{ belongs to } H_m^+, \text{ i.e. } p_0^2 - \mathbf{p}^2 = (mc)^2 \text{ with } p_0 > 0.$$

Photons are similarly treated as *massless particles* ($m = 0$). The latter are thereby characterized by a four-momentum vector p which belongs to the light-cone C^+:

$$p_0 = |\mathbf{p}|.$$

The concept of massive pointlike object and of relativistic four-momentum thus keep some meaning for describing the *free particles* of microphysics, namely non-interacting particles. However, it becomes meaningless for describing particles in mutual interaction, in contrast with the case of Newtonian objects, whose momenta and energies keep their meaning as functions of the time during the interaction.

The simplest thing that can be done *a priori* for describing the mutual interactions in particle physics is to describe the relations between the four-momentum configurations of free particles *before* the interaction and those which occur after interaction; in fact, for the interactions of nuclear type, the latter always takes place in a very short time. Then there is a basic relativistic law, which generalizes the Newtonian laws of conservation of the total momentum and of the total energy of the system. This law is

The law of conservation of the total energy-momentum vector of the system of free particles

This law states that if a set of several (let us say n) free particles with initial four-momentum vectors $p^{(1)}, p^{(2)}, \dots, p^{(n)}$ meet together in some region of Minkowski's spacetime where they interact, then another set of free particles will emerge in the future of that region and their number n' is not necessary equal to n. However, the four-momentum vectors ${p'}^{(1)}, {p'}^{(2)}, \dots, {p'}^{(n')}$ of these final particles are such that the following vector equality holds in the Minkowskian four-momentum space:

$$p^{(1)} + p^{(2)} + \cdots + p^{(n)} = {p'}^{(1)} + {p'}^{(2)} + \cdots + {p'}^{(n')}.$$

Of course, this implies that in contrast with the case of Newtonian pointlike objects, the particles of microphysics do not conserve their individualities throughout the interaction. However the vector conservation law which they obey puts some strong constraints which are *consequences of the Minkowskian triangular inequality* .

Let us consider for example the case of two initial particles with four-momenta $p^{(1)}$, $p^{(2)}$ (which is physically the generic case for the collisions produced in the accelarators). Let us call m_1 and m_2 the masses of these particles; one thus has:

$${p^{(1)}}^2 = m_1^2,\ {p^{(2)}}^2 = m_2^2.$$

Then the total four-momentum is

$$P = p^{(1)} + p^{(2)},$$

whose squared pseudonorm $P^2 \doteq M^2$ is interpreted as the *squared total mass of the system.* M is of course a relativistic invariant, independent of the Lorentz frame. In practice one often chooses a frame in which P is along the time-axis, which one calls the *center-of-mass frame.* Now , we see that because of the Minkowskian triangular inequality applied to the triangle whose sides are

$$[OQ_1] = p^{(1)},\ [Q_1Q_2] = p^{(2)},\ [OQ_2] = P,$$

one has necessarily

$$M \geq m_1 + m_2,$$

the equality being valid if and only if $p^{(1)}$ and $p^{(2)}$ are collinear; this means that the two particles are both at rest in the center-of-mass frame. If they are not, the difference $Mc^2 - m_1c^2 - m_2c^2$ represents *the (relativistic) kinetic energy of the system.*

Let us consider for example the case of equal masses $m_1 = m_2 \doteq m$. Then one has $M \geq 2m$. Now, let us ask ourselves what can be the constraints on the number of final particles emerging from the interaction. By iterating the previous geometrical argument with Minkowskian triangles, one gets the following result.

For $M < 3m$, only two final particles can be produced; one will then speak of an *"elastic" collision of two particles.* For $3m \leq M < 4m$, either two or three can be produced; both processes are geometrically possible. More generally, if $(n-1)m \leq M < nm$, all processes including the production of any number of final particles *smaller than or equal to* $n-1$ are possible. For the production of three or more particles, one also speaks of *"inelastic" collision of two particles.*

One can of course generalize the previous geometrical argument to the case of particles of different masses: note that the values of the masses of the existing particles of microphysics is a discrete set whose determination requires the treatment of quantum relativistic dynamical theories such as *Quantum Field Theories* (a very hard program which is by far outside the scope of this paper).

Inclusion of the photons

It is important to note that massless particles such as photons can be included in the previous geometrical arguments. In particular one can check (by drawing the corresponding triangles) that

i) From the collision of two photons, one can obtain a total momentum whose mass M can be arbitrarily large, so that any numer of final massive particles can a priori be produced throughout the interaction of these two photons: *"pure light can create matter"*

ii) Together with the elastic collision of two massive particles, one can always expect *a priori* the additional production of any number of photons, even if the total mass M ($> 2m$) of the system is not very much larger than $2m$.

An exercise on four-momentum conservation: "the propulsion of the Langevin-traveller's spaceship" (see Sec.3-3)

Let us assume that at time τ (in its proper time), the spaceship's mass is $M(\tau)$ and that, in its restframe, it is submitted to a constant field-strength whose intensity equal to g. Since its velocity is equal to zero in this frame, Newton's fundamental principle of dynamics applies and gives:

$$\frac{d\mathbf{P}}{d\tau} = M(\tau)g.$$

This field strength, which ensures the propulsion of the spaceship in uniformly accelerated motion, is produced by the expulsion of a part of the mass by unit of time, namely $\frac{dM(\tau)}{d\tau}$ whose associated momentum component must be equal in intensity and opposite to $\frac{d\mathbf{P}}{d\tau}$.

From a relativistic viewpoint, this loss of mass must in fact be identified (up to the factor c^2) with an emission of energy $\frac{1}{c^2}\frac{dE}{d\tau}$ under either form of a mass of matter (with relativistic velocity $v < c$) or of light (i.e. photons with velocity c).

In the case of matter, these energy and momentum losses are related to the velocity by the relativistic formula (given previously):

$$\left|\frac{d\mathbf{P}}{d\tau}\right| = \frac{v}{c^2}\left|\frac{dE}{d\tau}\right|,$$

which therefore yields the differential equation

$$\frac{dM}{d\tau} = -\frac{g}{v}M(\tau).$$

In the case of photons, one has a similar relation (with $v = c$):

$$\left|\frac{d\mathbf{P}}{d\tau}\right| = \frac{1}{c}\left|\frac{dE}{d\tau}\right|,$$

which yields

$$\frac{dM}{d\tau} = -\frac{g}{c}M(\tau).$$

One concludes that the loss of mass is minimized when $v = c$, namely if one can dispose of an engine which transforms matter into pure radiation.

6 Toward simple geometries of curved spacetimes

In spite of its non positive-definite distance, Minkowski's spacetime still shares with Euclidean space the property of being "flat", namely an affine space. But in the same way as the Euclidean plane must be replaced by a sphere (as a first approximation) for the observer who wishes to represent the surface of the earth, the four-dimensional Minkowskian spacetime must be replaced by a *curved* spacetime for the observer of the universe who wishes to describe the inclusion of matter submitted to gravitational attraction and the evolutional properties of the universe at astronomical scales of lengths and times. What we are mentioning here concerns the second big revolution of theoretical physics in the twentieth century: according to the principles of general relativity introduced by Einstein in 1916 (and also independently by Hilbert in a more mathematical formulation), local curvature of spacetime around an event X is caused by the presence of a density of matter at that point. But there is also another type of global curvature which is linked to expansion or contraction properties of spatial sections of spacetime; this type of curvature is characterized by what is called a "cosmological constant".

The general mathematical theory of curved spacetimes is outside the scope of the present pedagogical essay and we shall only indicate here some hint about the primary concepts involved in that theory. In mathematics, the notion of *manifold* introduces an additional abstraction to geometry. In the same way as the two-dimensional surface of the earth is perceived by us as "embedded in the ambient three-dimensional spacetime", a model of curved spacetime can reasonably be conceived as a "surface of dimension four embedded in a flat space of larger dimension" (for example five). As a matter of fact, this type of geometrical representation in terms of an "ambient space of higher dimension" is not necessary for defining the relevant mathematical notion of "manifold", which has been inspired by the geographical notion of *"atlas"*. In a world atlas, one is given a set of *planar* representations of various regions of the surface of the earth, in such a way that: a) each region is represented by precise geometrical rules encoded in a lattice of level curves representing parallels and meridians which constitute a map of that region; b) whenever two regions overlap, there are consistent geometrical rules which exhibit the correspondence between the two corresponding maps in their representations of the overlapping region; c) the union of all maps cover the whole surface of the earth. Such a type of collection of local data, which provides a faithful representation of a curved surface without requiring an embedding in a higher-dimensional ambient space, is used in the general mathematical definition of "abstract manifolds". The concept of atlas is thus often used for representing various models of curved spacetimes, thereby defined as *"abstract Minkowskian (or Lorentzian) manifolds"*. Ie such an atlas, each map is then specified by what one calls *a system of local coordinates of space and time*. The *Minkowskian local structure* is specified in each given map, by prescribing in terms of the corresponding local coordinates what are the cones of light world-lines passing at each given event X: these light world-lines will in general appear as curved lines, constituting a *"light-conoid" C_X with apex X, composed of the union of a future conoid C_X^+ and of a past conoid C_X^-*.

From the physical viewpoint, one can say that the conceptual advantage of this "atlas-representation" of a curved space or spacetime is to make the economy of an "ambient space", which has a priori no physical interpretation. As a matter of fact, the problem of the physical interpretation of additional dimensions introduced for mathematical reasons currently appears in various up-to-date investigations of theoretical physics.

However for certain models, a representation making use of an embedding of spacetime in a five-dimensional flat ambient space can be very illustrative and useful. Here of course, the word "ambient space" is of pure mathematical nature. These models correspond to "quadratic spacetimes" represented by appropriate *quadrics* (i.e. second-degree hypersurfaces) which enjoy the following simple geometrical property with respect to the ambient space. At each event X of the spacetime, the light-conoid C_X is *the cone of all linear generatrices of the quadrics* passing at X. These models of quadratic spacetimes have in common to be "pure-cosmological-constant models", which means that no density of matter is incorporated there. They enter in two classes with rather different mathematical properties and physical interest, which are called "de Sitter" and "anti-de Sitter spacetimes": they are presented in Ugo Moschella's paper.

References

[1] A. Einstein, *On the electrodynamics of moving bodies*, Annalen der Physik **17**, 891 (1905).

[2] A. Einstein, *Does the inertia of a body depend on its energy content?* Annalen der Physik, **18** pp. 639-41 (1905).

[3] H. Minkowski *Space and Time*, Cologne Conference, sept. 1908

[4] H. Poincaré, *Sur la dynamique de l'électron*, Note a l"Académie des Sciences, Paris, et Rendiconti del Circolo Matematico di Palermo (1905)

[5] C. W. Sherwin, *Some Experimental Tests of the "Clock Paradox"*, Phys. Rev. **120** 17-21, (1960).

[6] J. Terrell, *Invisibility of the Lorentz Contraction*, Phys. Rev. **116** 1041-1045, (1959).

[7] V. F. Weisskopf, *The visual appearance of rapidly moving objects*, Physics Today, Sept. 1960, p.24-27.

[8] D. Weiskopf, *An Immersive Virtual Environment for Special Relativity*

[9] E. C. Zeeman, *Causality Implies the Lorentz Group*, Journal Math. Phys. **5** 490-493, (1964).