Minkowski Spacetime: the geometrical foundation of Special Relativity

Shawn Westmoreland

September 11, 2008

Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality.

- Hermann Minkowski, September 21, 1908

1 Introduction

This is the first in a series of six lectures on relativity. I've decided to start with Minkowski's important geometrical 'spacetime' approach to Einstein's Special Theory of Relativity. I can't help but point out that this month, September 2008, marks 100 years since Minkowski introduced the notion of spacetime to the scientific community ([1], pp. 75 - 91). So I feel that the timing of this talk is very fortunate.

Special Relativity reconciles the principle of relativity, which was already grasped by Galileo, with Maxwell's theory of electrodynamics. Maxwell's theory was already well-confirmed experimentally by the time Einstein arrived on the scene. One of the most astonishing aspects of Maxwell's theory was the finding that light (or more generally any form of electromagnetic radiation) always travels at a certain speed (in vacuum). This speed is *exactly* 299 792 458 meters per second [2].

On the other hand, the principle of relativity stipulates that the laws of physics are the same for all 'inertial observers.' An inertial observer is one that is neither rotating nor accelerating. More precisely, it is an observer who finds that Newton's first law of motion, the *law of inertia*, holds.[1]

Now, the speed of light is fixed by certain physical laws (i.e., Maxwell's equations). So the principle of relativity implies that, in particular, the speed of light is the same for all inertial observers. In other words, the speed of light is *invariant*. This claim must seem absurd to the uninitiated. The average person in the street would probably think that light coming at them from the headlights of an approaching car comes at them a bit faster than light coming at them from the headlights of a parked car, but that can't be the case if the speed of light is invariant. You might start to think that there just can't be any such thing as an 'invariant' speed, and that consequently Maxwell's theory of electrodynamics is incompatible with the principle of relativity, but you would be wrong. In 1905, Einstein invented Special Relativity by realizing that Maxwell's theory of electrodynamics *is* compatible with the principle of relativity ([1], pp. 37 - 65). In order to achieve this, Einstein had to give up on certain common-sense notions about space and time, such as the idea that simultaneity has an absolute meaning.

The 'spacetime' approach to physics, initiated by Minkowski in 1908, simplifies Einstein's Special Theory of Relativity in a very profound and elegant way. Moreover, the concept of 'spacetime' continues to be of essential importance in General Relativity. So it is worthwhile to study the concept.

2 The world as spacetime

The world that we perceive consists of things existing at particular places in space at particular times. A particular point in space at a particular time is called a *worldpoint*, or *event*. The collection of all worldpoints constitutes what Minkowski called 'the world' ([1], p. 76), but today, Minkowski's 'world' is better known as *spacetime*. (Of course, one cannot get away with saying that spacetime is 'just the set of worldpoints.' There is structure holding everything together. The underlying mathematical structure of Minkowski spacetime will be described soon.)

A tiny fly fluttering about in this room can be idealized as a point-like object persisting in time. If we could graph the fly's position as a function of time using four-dimensional graph paper, we would find that the fly traces

[1]Newton's First Law: Every object continues in a state of uniform motion along a straight line, or remains in a state of rest, unless a force acts on it.

out a one-dimensional curve called a *worldline* in spacetime. Similarly, an idealized string describes a two-dimensional surface called a *worldsheet* (Figure 1).

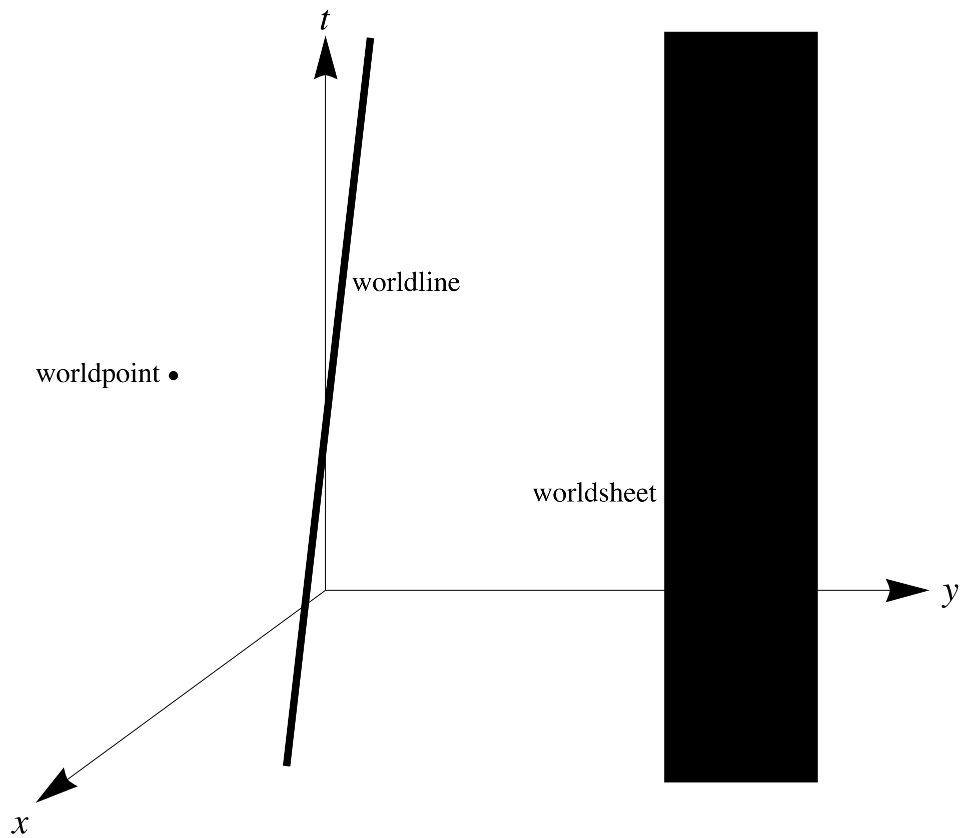

Figure 1: In this spacetime diagram, one space dimension (the 'z' dimension) has been repressed for simplicity. A *worldpoint*, or point-event, is a zero-dimensional point in spacetime. A *worldline* (which need not be 'straight') is a one-dimensional curve in spacetime. A *worldsheet* (which need not be flat) is a two-dimensional surface in spacetime. One can go on describing other objects, such as *worldtubes*, *worldvolumes*, etc. (All figures in this lecture were drawn with `Mathematica`.)

Physical interactions can be described geometrically in terms of worldlines, worldsheets, etc. For example, suppose that two billiard balls collide elastically. We can represent this in a spacetime diagram like that of Figure 2.

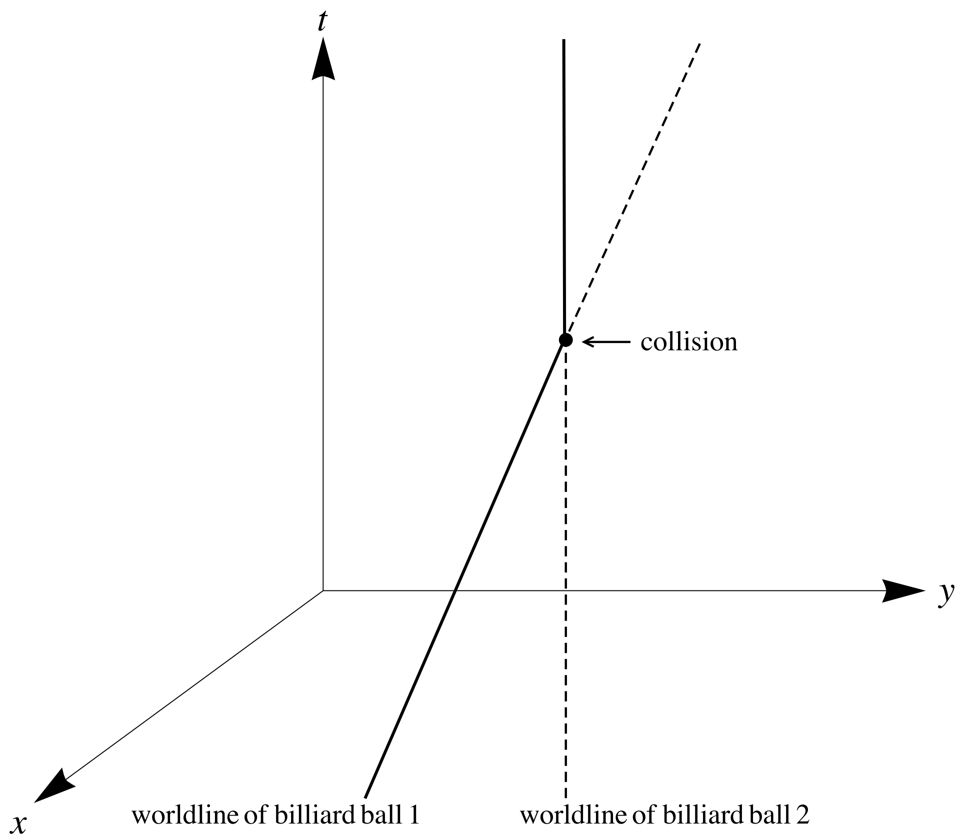

Figure 2: Two billiard balls, idealized as point-particles, collide elastically. Again, the 'z' dimension has been repressed for simplicity.

3 The spacetime interval

In Euclidean space, coordinate systems are 'bookkeeping' devices that help us keep track of points. Rectangular coordinate systems are particularly nice because the distance between two points can be expressed very simply in terms of their coordinates. We can also set up rectangular coordinate systems to do 'bookkeeping' in Minkowski spacetime.

An inertial observer can idealize herself as a point-particle resting at the origin of a vast rectangular coordinate system x, y, z that extends throughout all of three-dimensional Euclidean physical space. A good inertial observer always carries a trusty watch. Using this watch, the observer can assign a definite time-coordinate to any worldpoint that she observes in her immediate vicinity. All she has to do is use the reading on her watch at the moment the nearby worldpoint occurs. What about worldpoints that occur far away from her immediate vicinity? For example, suppose that she sees a flash of light far away, due to some kind of explosion, say. Perhaps a starcruiser was destroyed by hostile aliens. The explosion event itself marks a specific worldpoint in spacetime. (I'm pretending that the explosion is point-like.) Since the speed of light has a definite, finite value, our observer does not see the explosion until some time *after* it happens. Let us suppose that her x, y, z-coordinate system records that the explosion occurred at the spatial coordinates (x_0, y_0, z_0). She therefore determines that the explosion occurred at a distance $\delta = \sqrt{x_0^2 + y_0^2 + z_0^2}$ from her. If the reading on her watch is τ at the moment she sees the explosion, then let us have her assign the time coordinate $t_0 = \tau - \delta/c$ to the worldpoint, where c is the speed of light (Figure 3). A full set of 'spacetime' coordinates (x_0, y_0, z_0, t_0) for the explosion can therefore be worked out. I hope you are convinced that, assuming that physical space is Euclidean, *any* worldpoint can be assigned coordinates in this fashion, at least in principle. We will assume that the observer regards two worldpoints as 'simultaneous' if and only if, by following the above procedure, she assigns identical time coordinates to them.

Now that we have one coordinate system, let's have two! Suppose that another inertial observer, this one male, sets up a coordinate system x', y', z', t' in the same fashion as the original female observer. Suppose that the primed coordinate system moves at a uniform velocity v along the x-axis of the unprimed system. (We assume that $0 \leq |v| < c$.) Moreover, suppose that the x', y' and z'-axes remain parallel to the x, y and z axes respectively and suppose that the origins of the two coordinate systems, $(x, y, z, t) = (0, 0, 0, 0)$ and $(x', y', z', t') = (0, 0, 0, 0)$, coincide (Figure 4). We can ask the question: Suppose we are given that a particular worldpoint has coordinates (x, y, z, t) in the unprimed system. What are the coordinates (x', y', z', t') of the same worldpoint in the primed system? The coordinate transformations can be derived from first principles, but in the interest of time I'll just give the answer. The coordinate transformations are given by the *Lorentz transformations*:

$$\begin{bmatrix} x' \\ y' \\ z' \\ t' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 & -\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{v}{c^2\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}.$$

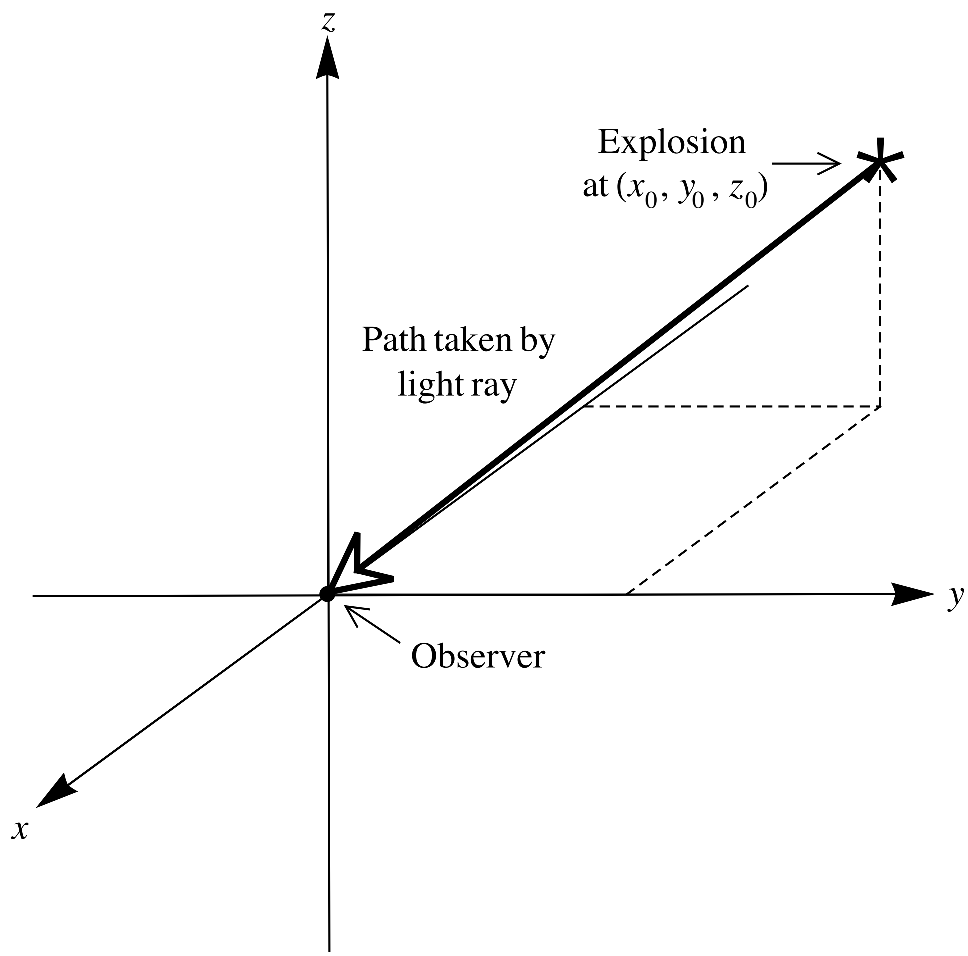

Figure 3: An explosion occurs at the spatial point with coordinates (x_0, y_0, z_0). The observer, who is located at the origin, does not see the explosion until its light reaches her. Hence the observer measures the distance to the explosion as $\delta = \sqrt{x_0^2 + y_0^2 + z_0^2}$, and she may suppose that the explosion occurred at a time δ/c before she saw it, where c is the speed of light. If the reading on the observer's watch is τ at the moment she sees the explosion, then her instructions are to assign the time coordinate $t_0 = \tau - \delta/c$ to the worldpoint corresponding to the explosion. She regards two worldpoints as 'simultaneous' if their time coordinates are identical.

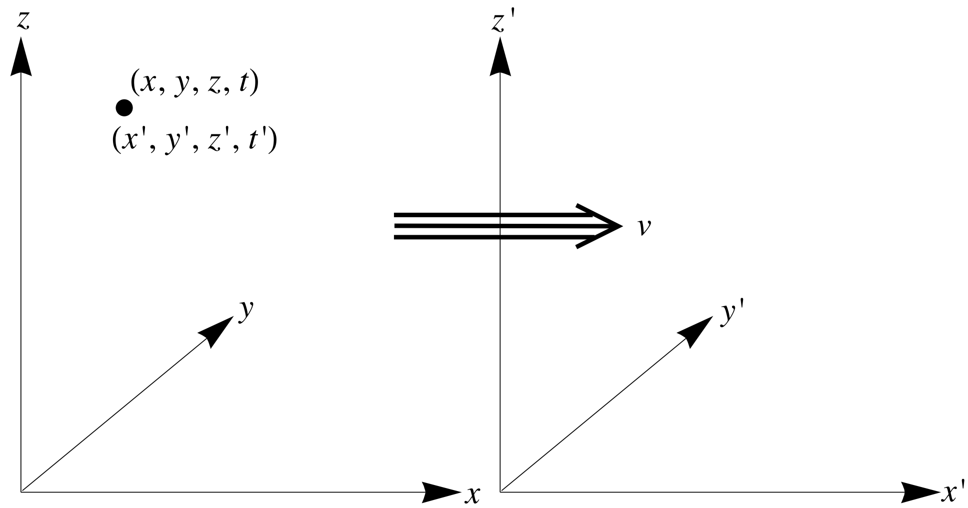

Figure 4: The primed coordinate system moves at a uniform velocity v along the x-axis of the unprimed system. (We assume that $0 \leq |v| < c$.) Moreover, the x', y' and z'-axes remain parallel to the x, y and z axes respectively. The clocks in each system can be set so that the origins of the two systems, $(x, y, z, t) = (0, 0, 0, 0)$ and $(x', y', z', t') = (0, 0, 0, 0)$, coincide.

We lose no generality only having the coordinate transformations for the case where the primed coordinate system moves along the positive x-axis and all that - because we can always rotate or translate the coordinate systems around in space and 'zero' the clocks until this particular configuration is obtained. Now, consider two worldpoints P_0 and P_1. In the unprimed system, these two worldpoints have coordinates (x_0, y_0, z_0, t_0) and (x_1, y_1, z_1, t_1) respectively. In the primed sytem, the coordinates are (x'_0, y'_0, z'_0, t'_0) and (x'_1, y'_1, z'_1, t'_1) respectively. The displacement vector between these two worldpoints can be expressed, in the unprimed coordinate system, as:

$$\langle \Delta x, \Delta y, \Delta z, \Delta t \rangle = \langle x_1 - x_0, y_1 - y_0, z_1 - z_0, t_1 - t_0 \rangle,$$

and in the primed coordinate system, as:

$$\langle \Delta x', \Delta y', \Delta z', \Delta t' \rangle = \langle x'_1 - x'_0, y'_1 - y'_0, z'_1 - z'_0, t'_1 - t'_0 \rangle.$$

Using the Lorentz transformations, one can prove the following remarkable identity:

$$(\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 - c^2(\Delta t')^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 - c^2 \Delta t^2. \qquad (1)$$

This identity is a major clue on the structure of Minkowski spacetime. We can define the *spacetime interval* between P_0 and P_1 as the square of the distance between P_0 and P_1 minus c^2 times the square of the time between P_0 and P_1, as measured by any inertial observer. In the unprimed coordinate system, the spacetime interval between P_0 and P_1 is:

$$\Delta x^2 + \Delta y^2 + \Delta z^2 - c^2 \Delta t^2.$$

In the primed coordinate system, the spacetime interval between P_0 and P_1 is:

$$\Delta(x')^2 + \Delta(y')^2 + \Delta(z')^2 - c^2 \Delta(t')^2.$$

It follows by the identity (1) that the spacetime interval between P_0 and P_1 is *independent of the observer*! In other words, all inertial observers agree on the spacetime interval between two worldpoints. Consequently, we say that the spacetime interval between two worldpoints is *invariant*.

It is customary to regard the interval as representing the square of something. Hence, in rectangular spacetime coordinates, we can write:

$$\Delta s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 - c^2 \Delta t^2.$$

For 'infinitesimal' displacements, we write:

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2.$$

4 Causal structure and the interval

The interval between two different worldpoints can be positive, negative, or zero. Suppose that an inertial observer, the pretty female one introduced previously, stretches her arms out and snaps her fingers at - what she regards as - the same time (remember, simultaneity is relative).[2] By snapping her fingers she marks two worldpoints in spacetime. Since the girl thinks that she snapped her fingers simultaneously, the time difference between these two worldpoints is zero in her coordinate system. Therefore the spacetime interval between the worldpoints is simply the square of what the girl considers to be the spatial distance between them. This is a positive quantity. Whenever the

[2] I'm borrowing an example from Geroch [4] here.

spacetime interval between two worldpoints is positive, we say that the two worldpoints are *spacelike separated*. If you want, you can prove that given two spacelike separated worldpoints, there is an inertial reference frame in which those two worldpoints are simultaneous.

Now suppose that the girl claps her hands together twice. By doing this, she has marked two new worldpoints. From her point of view, these two worldpoints occur at the same place, but at different times. Thus, the spacetime interval between these two worldpoints is minus c^2 times the square of what the girl regards as the time difference between them. This is a negative quantity. Whenever the spacetime interval between two worldpoints is negative, we say that the two worldpoints are *timelike separated*. If you want, you can prove that given two timelike separated worldpoints, there is an inertial reference frame in which those two worldpoints occur at the same place.

Suppose that it is dark and the girl switches on a flashlight. A ray of light races out of the device and hits a photomultiplier, which makes a loud 'click' upon receiving the light. We therefore have two worldpoints marked. The first worldpoint is marked by the switching on of the flashlight and the second worldpoint is marked by the 'click' of the photomultiplier. Since a ray of light travels from one worldpoint to the other, the distance between the two worldpoints must be equal to c times the time difference between the two worldpoints, in any reference frame. Thus, the spacetime interval between these two worldpoints is zero. Whenever the spacetime interval between two different worldpoints is zero, we say that the two worldpoints are *lightlike*, or *null separated*. This is because when the interval between two different wordpoints is zero, light can travel from one to the other.

The spacetime interval is to Minkowski spacetime what distance (or 'distance squared') is to Euclidean space. The expression for the spacetime interval between two worldpoints (in x, y, z, t rectangular coordinates):

$$\Delta x^2 + \Delta y^2 + \Delta z^2 - c^2 \Delta t^2,$$

bears a striking resemblance to the expression for the square of the *distance* between two points in four dimensional Euclidean space (in w, x, y, z rectangular coordinates):

$$\Delta x^2 + \Delta y^2 + \Delta z^2 + \Delta w^2.$$

Indeed, recall that the notion of distance between two points in Euclidean space can be expressed in terms of an inner product. It therefore makes sense to ask if the analogous idea holds in Minkowski spacetime. In four dimensional Euclidean space, we can form a displacement vector, an element of $\mathbb{R}^4$, by subtracting the w, x, y, z rectangular coordinates of two points component-wise. The inner product of two such displacement vectors in four dimensional Euclidean space is then given by:

$$\langle w_0, x_0, y_0, z_0 \rangle \cdot \langle w_1, x_1, y_1, z_1 \rangle = w_0 w_1 + x_0 x_1 + y_0 y_1 + z_0 z_1.$$

Now, the inner product of two displacement vectors in Euclidean space is actually equal to the product of the lengths of the displacement vectors times the cosine of the angle between them (once they are brought together by parallel transport). Thus, the Euclidean inner product is an invariant, coordinate independent quantity. After all, lengths and angles are what they are regardless of what coordinate system is being used.

In Minkowski spacetime, we can form a displacement vector, an element of $\mathbb{R}^4$, by subtracting the x, y, z, t rectangular coordinates of two worldpoints component-wise. Following the analogy with Euclidean space, the inner-product of two such displacement vectors in Minkowski spacetime ought to be given by:

$$\langle x_0, y_0, z_0, t_0 \rangle \cdot \langle x_1, y_1, z_1, t_1 \rangle = x_0 x_1 + y_0 y_1 + z_0 z_1 - c^2 t_0 t_1. \tag{2}$$

In order to have physical significance, Equation (2) must be invariant. That is, if the displacement vectors are expressed in terms of a different rectangular coordinate system, the numerical value of the inner product should be the same. This is indeed the case, as one can confirm using Lorentz transformations.

Note that the spacetime interval between two worldpoints P and Q is given by $PQ \cdot PQ$, where PQ is the displacement vector from P to Q. Hence, we say that a vector PQ is *spacelike* if $PQ \cdot PQ > 0$, *timelike* if $PQ \cdot PQ < 0$, and *lightlike* or *null* if $PQ \neq \mathbf{0}$ and $PQ \cdot PQ = 0$. We define the *magnitude* of a vector PQ in Minkowski spacetime as $\sqrt{|PQ \cdot PQ|}$.

The structure of Minkowski spacetime does not by itself supply an *intrinsic* distinction between the 'future' and the 'past.' However, an invariant global sense of 'time flow' is fixed once we *choose* a distinction between the future and past at a particular worldpoint. This means that Minkowski spacetime is *time-orientable.* To see how this works, suppose that you choose a pair of timelike separated worldpoints P and Q and you declare (by fiat)

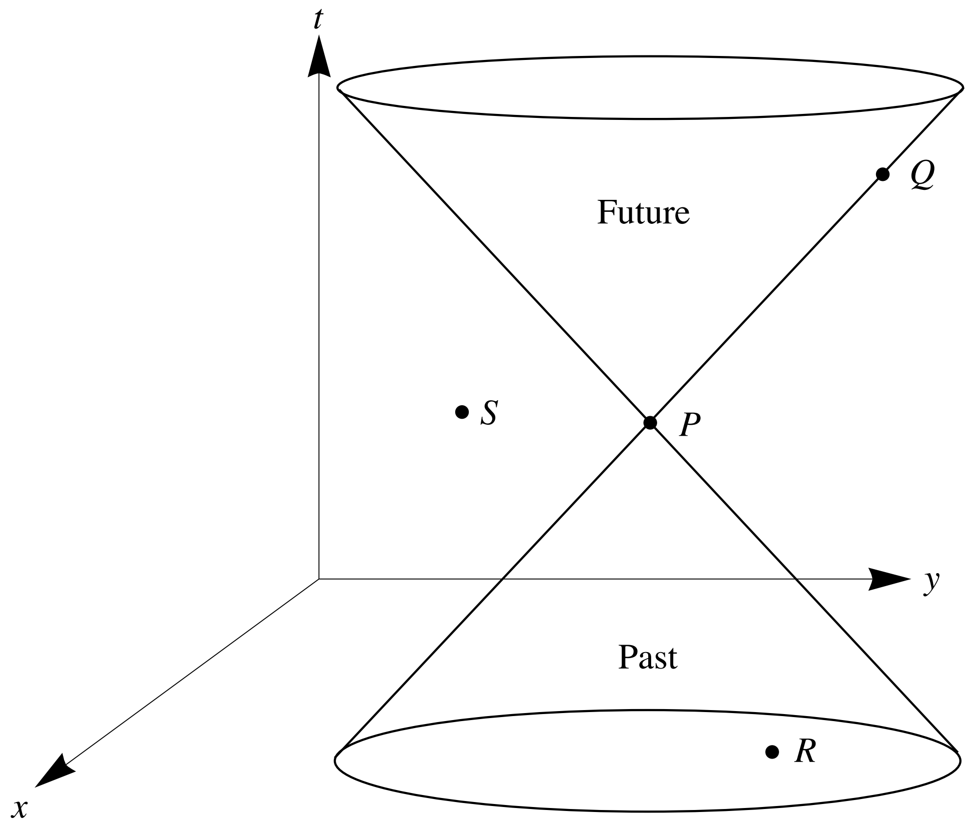

Figure 5: This is a schematic representation of the lightcone through a worldpoint P. The worldpoint Q lies on the future lightcone of P, R lies in the causal past of P, and S is not causally connectible with P in any way. In this figure, the 'z' dimension has been repressed for simplicity.

that Q is in the 'future' of P. Then an arbitrary null or timelike vector RS is said to be *future-directed* if $PQ \cdot RS < 0$, and *past-directed* if $PQ \cdot RS > 0$.

The *lightcone* through a worldpoint P is the set of all worldpoints Q such that the interval between P and Q is 0 (Figure 5). It is often thought of as consisting of two parts: the *future lightcone* and the *past lightcone*. The future lightcone through P consists of P together with all worldpoints Q such that PQ is a future-directed null vector. The past lightcone is defined similarly. Notice that the future and past lightcones through P separate spacetime into three regions. The *causal future* of P is the set of all worldpoints Q such that PQ is a future-directed timelike or null vector. The *causal past* of P is defined similarly. The region outside of P's lightcone is the set of worldpoints having no possibility of causal relationship with P. Any causal influence connecting P and a worldpoint in the region outside of P's lightcone would correspond to information traveling faster than the speed of light. The current scientific consensus is that this is physically impossible. If information can be transmitted faster than the speed of light, then messages can be sent into the past.[3]

5 Proper time and the aging hypothesis

Suppose that we have some rectangular coordinate system for Minkowski spacetime in which the worldline of a point-like particle is given by the parametric equations $(x(\lambda), y(\lambda), z(\lambda), t(\lambda))$, with parameter λ. We assume that the component functions are differentiable so that we can regard the tangent vector to the worldline as given by the function $\langle \frac{d}{d\lambda}x(\lambda), \frac{d}{d\lambda}y(\lambda), \frac{d}{d\lambda}z(\lambda), \frac{d}{d\lambda}t(\lambda)\rangle$. For physical reasons, the tangent vector of a given worldline must *always* be timelike (for a particle with mass, such as an electron) or *always* be null (for a particle without mass, such as a photon). A worldline whose tangent vector is timelike is called a *timelike worldline* and a worldline whose tangent vector is null is called a *null worldline*. The worldline of a point-like inertial observer (or a point-like particle with nonzero mass) is represented by a straight, timelike worldline.[4]

Suppose that $(x(\lambda), y(\lambda), z(\lambda), t(\lambda))$ describes the timelike worldline of some point-like particle. (This worldline need not be straight.) If we increment the parameter λ by an 'infinitesimal' amount $d\lambda$, then the corresponding displacement vector PQ along the worldline is given by:

$$\langle \frac{d}{d\lambda}x(\lambda), \frac{d}{d\lambda}y(\lambda), \frac{d}{d\lambda}z(\lambda), \frac{d}{d\lambda}t(\lambda)\rangle d\lambda = \langle dx(\lambda), dy(\lambda), dz(\lambda), dt(\lambda)\rangle$$

(see Figure 6). This displacement measures spacetime interval of:

$$ds^2 = dx(\lambda)^2 + dy(\lambda)^2 + dz(\lambda)^2 - c^2 dt(\lambda)^2. \tag{3}$$

Along an 'infinitesimal' displacement, the worldline appears to be straight. That is, the worldline of the particle can be thought of as corresponding,

[3] Exercise!

[4] That is, the worldline of a point-like inertial observer can be expressed by a collection of linear (or affine) parametric equations in rectangular x, y, z, t coordinates for spacetime. Such a worldline 'looks straight' in coordinates and describes a particle that follows a straight line trajectory in Euclidean space with a constant velocity.

at least momentarily, to the worldline of some fictitious inertial observer. Then the spacetime interval (3) is equal to minus c^2 times the square of the elapsed time, as measured by this fictitious inertial observer, between the infinitesimally close worldpoints P and Q. The elapsed time measured by the fictitious observer is therefore:

$$\begin{aligned} d\tau &= \left(-\frac{1}{c^2}ds^2\right)^{\frac{1}{2}} \\ &= \frac{1}{c}\left(-\left(\frac{d}{d\lambda}x(\lambda)\right)^2 - \left(\frac{d}{d\lambda}y(\lambda)\right)^2 - \left(\frac{d}{d\lambda}z(\lambda)\right)^2 + c^2\left(\frac{d}{d\lambda}t(\lambda)\right)^2\right)^{\frac{1}{2}} d\lambda. \end{aligned}$$

As the parameter λ varies continuously from λ_0 to λ_1, the worldline γ of the point-particle is displaced from the worldpoint $(x(\lambda_0), y(\lambda_0), z(\lambda_0), t(\lambda_0))$ to the worldpoint $(x(\lambda_1), y(\lambda_1), z(\lambda_1), t(\lambda_1))$. We define the *proper time* between these two worldpoints along γ as:

$$\tau = \frac{1}{c}\int_{\lambda_0}^{\lambda_1}\left(-\left(\frac{d}{d\lambda}x(\lambda)\right)^2 - \left(\frac{d}{d\lambda}y(\lambda)\right)^2 - \left(\frac{d}{d\lambda}z(\lambda)\right)^2 + c^2\left(\frac{d}{d\lambda}t(\lambda)\right)^2\right)^{\frac{1}{2}} d\lambda. \quad (4)$$

A clock is said to be an *ideal clock* if the amount of time it measures between two worldpoints on its worldline is equal to the proper time between those worldpoints along the clock's worldline. Since the proper time along a given worldline depends only on the tangent vector to the worldline, an ideal clock is not affected by acceleration or higher-order derivatives. Although the notion of an 'ideal clock' is merely a theoretical construct, there are natural 'clocks,' such as vibrating atoms, that are observed experimentally to behave like ideal clocks ([5], p. 43).

Consider an everyday physical clock, say, a spring-driven clock. It is clear that very high accelerations will alter the function of the clock. Such a clock is not 'ideal' (it does not accurately measure 'proper time'), but one might guess that the clock as a physical system nevertheless *ages* according to its proper time. In general, we make the hypothesis that all point-like physical systems age according to their proper time. I call this the *aging hypothesis*.[5] Does the aging hypothesis stand up to experimental scrutiny? Yep, it sure does! Observations of muons (a type of subatomic particle) show that muons

[5] Others call it the 'clock hypothesis.' For an interesting and recent philosophical treatment of this subject, see [6], pp. 103 - 107.

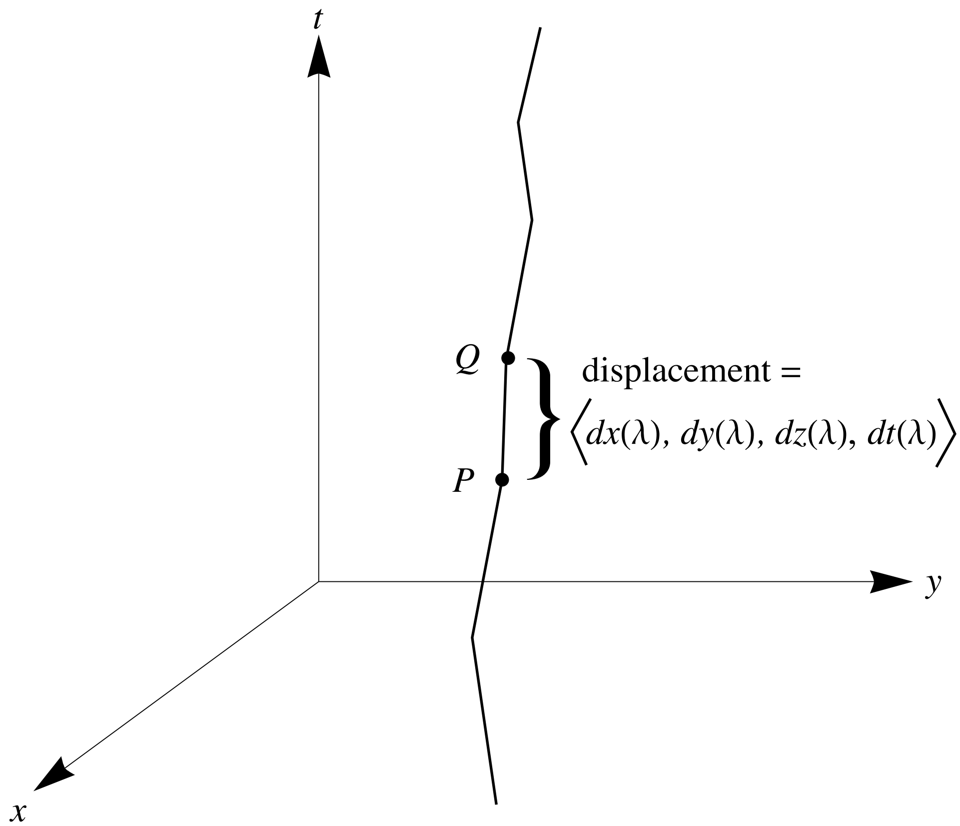

Figure 6: An arbitrary timelike worldline is thought of as being a long chain of infinitesimal linear parts. The 'proper time' along a timelike worldline can be obtained by summing up the intervals along of the infinitesimal linear parts. The 'z' dimension has been repressed for simplicity.

always decay after existing for a certain proper time. The proper lifetime of the muon is the same for all muons - this has been confirmed at CERN for muons under accelerations as high as 10^{19} g's ([5], p. 44)!

6 The twin paradox

Note that the proper time between two worldpoints along a timelike wordline depends on how the worldline threads through spacetime (Equation 4). This observation leads to the famous *twin paradox*, which is not really a paradox even though it can be made to seem like one. The so-called paradox is usually presented as a story about two twins. These twins start out quite naturally as having the same biological age. One twin takes a high speed journey through space and later reunites with his stay-at-home counterpart. Upon his return, it is found that the travelled twin is younger than the stay-at-home twin. The space-faring twin effectively winds up in the future of his counterpart!

Let's look at this more explicitly. We assume that the stay-at-home twin is an inertial point-like observer. So there exists a rectangular coordinate system x, y, z, t in which the worldline of the stay-at-home twin is given by $(x, y, z, t) = (0, 0, 0, \lambda)$. Suppose that when $\lambda = 0$, the traveling twin flies away in his starship (worldpoint A), and when $\lambda = T$, the twins are reunited (worldpoint B). The proper time during the travelled twin's journey for the *stay-at-home twin* is therefore T.

We can parameterize the worldline of the traveled twin by $(x, y, z, t) = (x(\lambda), y(\lambda), z(\lambda), \lambda)$, where $x(\lambda), y(\lambda)$, and $z(\lambda)$ are some non-constant continuously differentiable functions such that $x(0) = y(0) = z(0) = x(T) = y(T) = z(T) = 0$ and the worldline is timelike. The proper time during the journey for the *travelled twin* is therefore:

$$\begin{aligned} \tau &= \frac{1}{c}\int_0^T \left(-\left(\frac{d}{d\lambda}x(\lambda)\right)^2 - \left(\frac{d}{d\lambda}y(\lambda)\right)^2 - \left(\frac{d}{d\lambda}z(\lambda)\right)^2 + c^2\right)^{\frac{1}{2}} d\lambda \\ &< \frac{1}{c}\int_0^T \left(c^2\right)^{\frac{1}{2}} d\lambda = T. \end{aligned} \tag{5}$$

So, by the aging hypothesis, the travelled twin really will end up being younger than the stay-at-home twin.

From the spacetime perspective, this result is no more 'paradoxical' than the fact that two different paths connecting two points in space can have different lengths (Figure 7).

An important issue does remain to be resolved, however. From the traveling twin's perspective it appears that the stay-at-home twin travels away and then returns. So, by symmetry, it might seem to follow equally well that the stay-at-home twin will end up being the younger.

There is no real contradiction because the apparent symmetry is false. We assumed that the stay-at-home twin is inertial. The traveling twin has to accelerate and is therefore not inertial.[6] Geometrically, a non-inertial observer

[6]It is possible to modify the twin paradox so that acceleration is not involved and

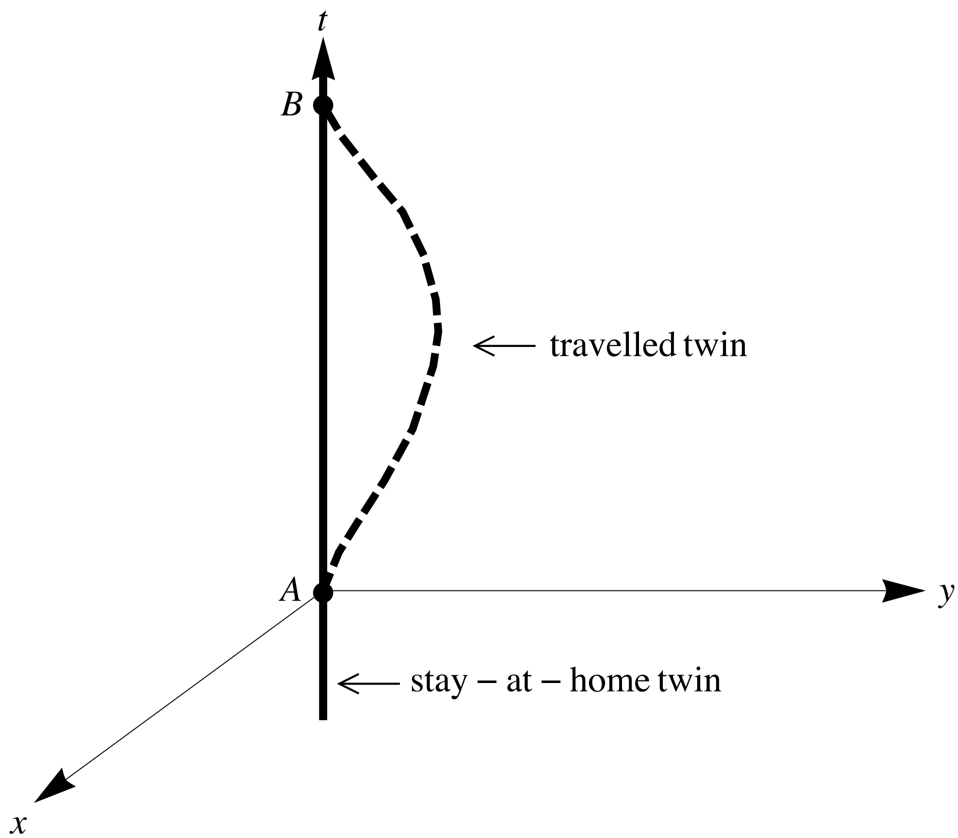

Figure 7: The so-called 'twin paradox' can be understood from the fact that two different worldlines connecting two worldpoints in spacetime can experience different amounts of proper time between those worldpoints. The stay-at-home twin (assumed to be an inertial point-like particle) is represented by the solid worldline and the travelled twin is represented by the dashed worldline. The 'z' dimension has been repressed, as usual.

differs from an inertial observer in that the worldline of an inertial observer is *straight* while the worldline of a non-inertial observer is not straight.

---

yet the two twins age asymmetrically (e.g., by rolling spacetime up into a cylinder, see [7]). In such a scenario, acceleration doesn't distinguish between the two twins but other phenomena do.

7 A maximal principle

What do we mean by a 'straight' worldline? Well, earlier I said that a straight worldline is a worldline that can be expressed by a collection of linear (or affine) parametric equations in rectangular x, y, z, t coordinates for Minkowski spacetime. I could probably get away with that definition in Minkowski spacetime, but it isn't very useful in General Relativity.

A more generalized definition of 'straightness' is already suggested by our result on the twin paradox. We agreed that the worldine of the inertial (stay-at-home) twin through the worldpoints A and B is straight. On the other hand, the worldline of the non-inertial (travelled) twin through A and B is not straight. Equation (5) tells us that the proper time along the straight worldline from A to B is necessarily *greater* than the proper time along a non-straight worldline from A to B.

Hence, we can mathematically *define* a straight timelike worldline as one that maximizes the proper time between any two worldpoints on it. This is similar to the fact that in Euclidean geometry, a straight line segment is the path of minimum length between two points. The technical jargon for 'straight' is *geodesic*. So we can say that a timelike geodesic is a curve that maximizes proper time, and the worldline of an inertial point-like particle (with nonzero mass) is a timelike geodesic.

In Minkowski spacetime, there is only one timelike geodesic through any two timelike separated worldpoints. In General Relativity however, it is sometimes possible to have more than one timelike geodesic through two worldpoints (for example, if one rolls Minkowski spacetime up into a cylinder). Thus, a timelike geodesic only maximizes proper time in a 'relative' sense: If a timelike geodesic γ between worldpoints A and B is perturbed slightly, a worldline having less proper time between A and B than γ will be produced, but there can be other geodesics through A and B that measure more or less proper time between A and B than γ.

I will have more to say about geodesics in the next lecture, which will be about pseudo-Riemannian geometry.

References

[1] A. Einstein, H. A. Lorentz, H. Weyl, H. Minkowski, A, Sommerfeld, *The Principle of Relativity*, Dover, New York, 1952.

[2] http://math.ucr.edu/home/baez/physics/Relativity/SpeedOfLight/measure_c.html A website about the speed of light.

[3] C. W. Misner, K. S. Thorne, J. A. Wheeler, *Gravitation*, W. H. Freeman and Company, San Francisco, 1973.

[4] R. Geroch, *General Relativity from A to B*, The University of Chicago Press, Chicago, 1978.

[5] W. Rindler, *Essential Relativity: Special, General, and Cosmological, 2nd ed.*, Springer-Verlag, New York, 1977.

[6] T. A. Debs and M. L. G. Redhead, *Objectivity, Invariance, and Convention: Symmetry in Physical Science*, Harvard University Press, Cambridge, 2007.

[7] C. H. Brans and D. R. Stewart, "Unaccelerated-Returning-Twin Paradox in Flat Space-Time," Phys. Rev. D, Vol. 8, No. 6, pp. 1662 - 1666, Sept., 1973.