Lecture 5: Minkowski space and relativity

updated 7/8/18

In the first lecture I listed several structures on Euclidean space $\mathbb{R}^n$ (such as distance, open balls, inner products) that form the basis for abstractions like metric spaces, topological spaces, and inner product spaces. I gave a brief glimpse of the deep role of symmetry groups in determining which of these abstractions is appropriate for a given domain. The symmetries that we have encountered so far are rotations, reflections and translations (the last only briefly in Exercise L1-4). These are all isometries of $\mathbb{R}^n$ (and also, in the case of rotations and reflections, of the metric space (S^1, d_a)), and that's it! More precisely:

Exercise L5-1 Set $E(n) = \mathrm{Isom}(\mathbb{R}^n, d_2)$. For $\underline{x} \in \mathbb{R}^n$ define

$$T_{\underline{x}} : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad T_{\underline{x}}(\underline{y}) = \underline{x} + \underline{y}.$$

Prove that $T_{\underline{x}}$ is an isometry and that

$$T(n) := \{ T_{\underline{x}} \mid \underline{x} \in \mathbb{R}^n \}$$

is a subgroup of $E(n)$. Observe that $O(n)$ may also be identified with a subgroup of $E(n)$. Prove that

(i) every element f of $E(n)$ can be written as $f = X \circ T_{\underline{x}}$ for some $X \in O(n)$ and $\underline{x} \in \mathbb{R}^n$.

(ii) $T(n)$ is a normal subgroup and $E(n)/T(n) \cong O(n)$.

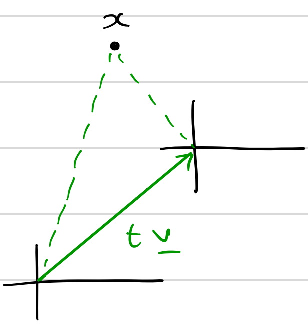
sometimes written $ISO(n)$

This group $E(n)$ is called the Euclidean group. It relates the measurements of an arbitrary pair of stationary observers in $\mathbb{R}^n$.

One way to motivate a level of abstraction <u>beyond metric spaces</u> is to find a new kind of symmetry of $\mathbb{R}^n$ beyond rotations, reflections and translations, which we are forced to care about (another motivation to move up the abstraction ladder is <u>infinite-dimensional spaces</u> which we will treat in the near future). One way we may be forced to care is by evidence that our intuitions about space, based ultimately on our perceptions, are simply <u>wrong</u> when they are extrapolated to large velocities. This is the content of special relativity, which we now discuss.

⌈ <u>Aside</u>: of course, the discovery of relativity does not invalidate the mathematical structure $(\mathbb{R}^n, d_2)$. The truth of theorems is (surely?) independent of experiment. But such experiments can help, and historically have helped, guide mathematics towards deeper truths. The Greeks believed they could figure out all there was to know about mathematics by pure introspection (i.e. damn the experiments). The Greeks <u>were wrong</u>, in practice. ⌋

<u>Moving observers</u> Let us once again consider two observers O_1, O_2 in the plane, where O_2 is moving with velocity $\underline{v}$ as measured by O_1 (and so O_1 is moving with velocity $-\underline{v}$ as measured by O_2). Then we have a commutative diagram, for each $t \in \mathbb{R}$, assuming at time $t = 0$ that O_1, O_2 are co-located and have the same orientation and their axes have no relative rotation (and they remain so):



$$\begin{tikzcd} & X \arrow[dl, "m_2"'] \arrow[dr, "m_1"] & \\ \mathbb{R}^2 \arrow[rr, "T_{t\underline{v}}"'] & & \mathbb{R}^2 \end{tikzcd}$$

Exercise L5-0 Give an analogous commuting diagram in the case where (i) O_2's coordinate axes are rotated by θ relative to O_1's, and (ii) when both the axes are rotated and the observers have different orientations.

We can incorporate the time by introducing the concept of an event which is simply a measurement $\underline{x} \in \mathbb{R}^2$ tagged with a time t (we assume the observers each have identical clocks, originally synchronised), i.e.

$$(t, \underline{x}) \in \mathbb{R} \times \mathbb{R}^2 = \mathbb{R}^3.$$

So we replace X, the "plane without coordinates" by E, the "set of events", and our observers now map events to measurements in $\mathbb{R}^3$ related by a function $F_{12} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ as in the commutative diagram

$$\begin{tikzcd}
& E \arrow[dl, "m_2"'] \arrow[dr, "m_1"] & \\
\mathbb{R}^3 \arrow[rr, "F_{12}"'] & & \mathbb{R}^3
\end{tikzcd}$$

(we re-use the notation m_1, m_2 for measurement, but to be clear these are different functions)

where we would naïvely expect F_{12} to be defined by

$$F_{12}(t, \underline{x}) = (t, T_{t\underline{v}}(\underline{x})) = (t, \underline{x} + t\underline{v}).$$

This functions "predicts" or calculates O_1's measurement of an event from prior knowledge of the relative motion $\underline{v}$ and O_2's measurement of the same event. We assume as before that m_1, m_2 are bijections.

The observers agree about linear motion a set $\mathcal{A} \subseteq E$ of events represents a finite linear motion according to observer O_2 if there exists $\alpha \le \beta$ and $\underline{u}, \underline{x}_0 \in \mathbb{R}^2$ s.t.

$$m_2(\mathcal{A}) = \{ (t, \underline{x}_0 + t\underline{u}) \mid \alpha \le t \le \beta \}$$

We call $\underline{u}$ the velocity of the motion and $\|\underline{u}\|$ its speed, according to O_2. The observer naturally divides the motion into a first half $t \in [\alpha, \frac{1}{2}(\alpha+\beta)]$ and second half $t \in [\frac{1}{2}(\alpha+\beta), \beta]$, with midpoint $t = \frac{1}{2}(\alpha+\beta)$. This corresponds to a decomposition

$$\mathcal{A} = \mathcal{A}_L \cup \mathcal{A}_R \quad \text{with} \quad |\mathcal{A}_L \cap \mathcal{A}_R| = 1 \qquad (*)$$

where $\mathcal{A}_L = \{ e \in \mathcal{A} \mid m_2(e)_1 \le \frac{1}{2}(\alpha+\beta) \}$, $\mathcal{A}_R = \{ e \in \mathcal{A} \mid m_2(e)_1 \geq \frac{1}{2}(\alpha+\beta) \}$.
We call this the midpoint decomposition of the motion. ⌜by a midpoint decomposition we mean a pair $\{\mathcal{A}_L, \mathcal{A}_R\}$ satisfying $(*)$. Not an ordered pair⌟

If O_2 records such motion then according to our def^N of F_{12}, O_1 must record

$$\begin{aligned} m_1(\mathcal{A}) = F_{12} m_2(\mathcal{A}) &= \{ F_{12}(t, \underline{x}_0 + t\underline{u}) \mid \alpha \le t \le \beta \} \\ &= \{ (t, \underline{x}_0 + t(\underline{u} + \underline{v})) \mid \alpha \le t \le \beta \}. \end{aligned}$$

so they disagree about the velocity of the motion, but they agree that it was linear motion at constant velocity, and they agree about the midpoint decomposition $(*)$.

The observation, first made in the context of studying Maxwell's equations by Lorentz, and later clarified by Poincaré, Minkowski and of course Einstein, is that for real physical observers F_{12} does not correctly predict O_1's measurements from O_2's. This is a remarkable empirical fact! The correct translation F_{12} is deduced from Einstein's postulates of special relativity. We give a specialised version of these postulates below.

First of all one needs to accept the notion of an observer making measurements $E \longrightarrow \mathbb{R}^3$ using rigid measuring rods and a system of clocks as described in §1 of Einstein's paper. There is a subtlety here about defining the time of a distant event which he cleverly handles, but we will not give the details here

The first postulate of special relativity (SR) is

> " The laws by which the states of physical systems undergo change are not affected, whether these changes of state be referred to the one or the other of two systems of co-ordinates in uniform translatory motion. "

but as we are not concerned with dynamics of physical systems (apart from the observers themselves) we will focus on the other postulates.

<u>Special relativity</u> (SR) postulates that there exists $c > 0$ s.t. for any pair of observers which are as above, i.e. their relative motion is constant and linear,

(SR-1) The observers agree on which sets $\mathcal{A}$ of events represent finite linear motion. (they may disagree on $\alpha, \beta, \underline{u}, \underline{x}_0$)

(SR-2) Given such a set $\mathcal{A}$, the observers agree on the midpoint decomposition $\mathcal{A} = \mathcal{A}_L \cup \mathcal{A}_R$ (in particular they agree which event $e \in \mathcal{A}$ <u>is</u> the midpoint of the motion). (they may disagree on which half of the motion happened first!)

(SR-3) the observers agree on which sets of events represent finite linear motion <u>of speed c</u>. (this is the shocking part)

To write this more formally, we can remove E and rewrite everything in terms of F_{12} and $\mathbb{R}^3$ (just as we did in Lecture 1 with X, R_θ and $\mathbb{R}^2$). Framed this way, the postulates identify a particular class of functions $F: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ that determine the possible conversions between pairs of observers whose relative motion is constant and linear. We fix $c > 0$.

<u>Def^n</u> A function $F: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is an "SR conversion" if it is continuous, and

(SR-2)$'$ <u>F preserves midpoints</u> given $(t, \underline{x}), (t', \underline{x}') \in \mathbb{R}^3$,

$$F\left(\frac{(t,\underline{x}) + (t',\underline{x}')}{2}\right) = \frac{F(t,\underline{x}) + F(t',\underline{x}')}{2}$$

(SR-3)$'$ Given $(t,\underline{x}), (t',\underline{x}') \in \mathbb{R}^3$ we have

$$-c^2(t'-t)^2 + \|\underline{x}'-\underline{x}\|^2 = -c^2(s'-s)^2 + \|\underline{y}'-\underline{y}\|^2$$

where $(s, \underline{y}) := F(t,\underline{x})$, $(s', \underline{y}') := F(t', \underline{x}')$.

(SR-4)$'$ $F(0,\underline{0}) = (0,\underline{0})$

<u>Remark</u> (SR-2)$'$ incorporates both (SR-1), (SR-2). An earlier version of these notes (used in the recorded lectures) contains a redundant (SR-1)$'$ which moreover contains a subtle error involving time intervals of zero length, so it is best all round to use only (SR-2)$'$.

<u>Remark</u> We have cheated somewhat in going from (SR-3) to (SR-3)$'$. What (SR-3) actually says is:

(SR-3)$''$ Given $(t,\underline{x}),(t',\underline{x}') \in \mathbb{R}^3$ we have

$$\|\underline{x}'-\underline{x}\|^2 = c^2(t'-t)^2$$

if and only if (writing $(s,\underline{y}) = F(t,\underline{x}),\ (s',\underline{y}') = F(t',\underline{x}')$)

$$\|\underline{y}'-\underline{y}\|^2 = c^2(s'-s)^2.$$

It is then another step to see this equivalent to (SR-3)$'$. The physical principle of isotropy of space is usually invoked. To derive the Lorentz transformations directly from (SR-3)$''$ see e.g.

G.C. Hegerfeldt "The Lorentz transformations: derivation of linearity and scale factor" Il Nuovo Cimento 1972.

<u>Question for physics students</u>: what is the physical basis for (SR-2)?

<u>Remark</u> Among the many things worth puzzling over in SR is that we assume observers agree on all linear motions, but later in the subject we note that no physical linear motion at speed $> c$ is possible. So really if we want our axioms to refer to <u>actually possible motions</u> we should talk about those of speed $< c$ only, and deduce e.g. linearity from this. This is a fair point, I think; see Hegerfeldt's paper for a good response to this issue.

<u>Theorem</u> Let P denote the symmetric matrix

$$P = \begin{pmatrix} -c^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and define $\langle \underline{v}, \underline{w} \rangle_P = \underline{v}^T P \underline{w}$, for $\underline{v}, \underline{w} \in \mathbb{R}^3$. Then F is an "SR conversion" if and only if it is a bijection, and

(i) a linear transformation, and

(ii) $\langle F\underline{v}, F\underline{w} \rangle_P = \langle \underline{v}, \underline{w} \rangle_P$ for all $\underline{v}, \underline{w} \in \mathbb{R}^3$

<u>Proof</u> Suppose F satisfies (i), (ii). It is certainly continuous, preserves lines and midpoints, so it only remains to show (SR-3)$'$. For $\underline{v} = (t, \underline{x}) \in \mathbb{R}^3$ we have

$$\langle \underline{v}, \underline{v} \rangle_P = (t \;\; \underline{x}) \begin{pmatrix} -c^2 & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} t \\ \\ \underline{x} \end{pmatrix} = -c^2 t^2 + \|\underline{x}\|^2.$$

so from (ii) we deduce (SR-3)$'$. Of course $F(0, \underline{0}) = (0, \underline{0})$ since F is linear.

Let us now suppose F is an SR conversion and prove (i), (ii) for F. Recall $\lambda \in [0,1]$ is <u>dyadic</u> if it may be expressed as $\frac{m}{2^n}$ for some $m, n \geq 0$ integers. Using iterated midpoints we show

$$F(\lambda(t, \underline{x}) + (1-\lambda)(t', \underline{x}')) = \lambda F(t, \underline{x}) + (1-\lambda) F(t', \underline{x}') \qquad (\dagger)$$

for any pair $(t,\underline{x}),(t',\underline{x}')\in\mathbb{R}^3$ and $0\le\lambda\le 1$ dyadic. Using that F is continuous we obtain $(\dagger)$ for all $0\le\lambda\le 1$. Then for $\underline{v},\underline{w}\in\mathbb{R}^3$

Claim 1 $F(\lambda\underline{v})=\lambda F(\underline{v})$ for all $0\le\lambda\le 1$

Proof
$$\begin{aligned}F(\lambda\underline{v}) &= F(\lambda\underline{v}+(1-\lambda)\cdot\underline{0})\\ &= \lambda F(\underline{v})+(1-\lambda)F(\underline{0})\\ &= \lambda F(\underline{v}).\end{aligned}$$
(since $F(\underline{0})=\underline{0}$)

Claim 2 $F(\underline{v}+\underline{w})=F(\underline{v})+F(\underline{w})$

Proof
$$\begin{aligned}F(\underline{v}+\underline{w}) &= F(\tfrac{1}{2}2\underline{v}+\tfrac{1}{2}2\underline{w})\\ &= \tfrac{1}{2}F(2\underline{v})+\tfrac{1}{2}F(2\underline{w})\\ &\overset{\text{Claim 1}}{=} F(\underline{v})+F(\underline{w}).\end{aligned}$$

Claim 3 $F(n\underline{v})=nF(\underline{v})$ for any $n\in\mathbb{Z}$.

Proof For $n\geqslant 0$ this follows by iterating Claim 2. Since

$$\underline{0} \overset{(SR-4)'}{=} F(\underline{0}) = F(-\underline{u}+\underline{u}) = F(-\underline{u})+F(\underline{u})$$

we get $F(-\underline{u})=-F(\underline{u})$ and the $n<0$ cases then also follow.

Claim 4 $F(\mu \underline{v}) = \mu F(\underline{v})$ for all $\mu \in \mathbb{R}$.

Proof We may write $\mu = n + \lambda$ for $n \in \mathbb{Z}$, and $0 \leq \lambda \leq 1$. Then by Claim 3 and Claim 2, and Claim 1,

$$\begin{aligned} F(\mu \underline{v}) &= F(n\underline{v} + \lambda \underline{v}) = F(n\underline{v}) + F(\lambda \underline{v}) \\ &= nF(\underline{v}) + \lambda F(\underline{v}) \\ &= (n+\lambda) F(\underline{v}) \\ &= \mu F(v). \end{aligned}$$

This completes the proof that F is linear. Now we need to use (SR-3)$'$ to see that $\langle F\underline{v}, F\underline{w} \rangle_P = \langle \underline{v}, \underline{w} \rangle_P$, for all $\underline{v}, \underline{w} \in \mathbb{R}^3$. Now, what (SR-3)$'$ tells us immediately is that for all $\underline{v} \in \mathbb{R}^3$

$$\langle F\underline{v}, F\underline{v} \rangle_P = \langle \underline{v}, \underline{v} \rangle_P .$$

But then we can use the polarisation formula (Exercise L5-2) to see

$$\begin{aligned} \langle F\underline{v}, F\underline{w} \rangle_P &= \tfrac{1}{4}\big(\langle F\underline{v} + F\underline{w}, F\underline{v} + F\underline{w} \rangle_P \\ &\qquad - \langle F\underline{v} - F\underline{w}, F\underline{v} - F\underline{w} \rangle_P \big) \\ &= \tfrac{1}{4}\big(\langle F(\underline{v}+\underline{w}), F(\underline{v}+\underline{w}) \rangle_P \\ &\qquad - \langle F(\underline{v}-\underline{w}), F(\underline{v}-\underline{w}) \rangle_P \big) \\ &= \tfrac{1}{4}\big(\langle \underline{v}+\underline{w}, \underline{v}+\underline{w} \rangle_P - \langle \underline{v}-\underline{w}, \underline{v}-\underline{w} \rangle_P \big) \\ &= \langle \underline{v}, \underline{w} \rangle_P . \quad \square \end{aligned}$$

Exercise LS-1 Fill in the following detail in the proof: (i) explain how to use iterated midpoints to deduce (†) for λ dyadic and (ii) use continuity to deduce (†) for all $0 \leq \lambda \leq 1$.

Exercise LS-2 (Polarisation identity) Prove that

$$\langle \underline{v}, \underline{w} \rangle_P = \frac{1}{4}\left(\langle \underline{v}+\underline{w}, \underline{v}+\underline{w} \rangle_P - \langle \underline{v}-\underline{w}, \underline{v}-\underline{w} \rangle_P \right).$$

It remains to analyse exactly what kind of matrices A give rise to linear transformations $F(\underline{v}) = A\underline{v}$ which satisfy (i), (ii), (iii). These matrices form the appropriate group of symmetries for special relativity.

Remark Let F be as in the theorem, and say $F(1,\underline{0}) = (\gamma, \underline{q})$. Then

$$\begin{aligned} -c^2\gamma^2 + \|\underline{q}\|^2 &= \langle (\gamma, \underline{q}), (\gamma, \underline{q}) \rangle_P \\ &= \langle (1,\underline{0}), (1,\underline{0}) \rangle_P \\ &= -c^2 \end{aligned}$$

This shows $\gamma \neq 0$, so set $\underline{r} := \frac{1}{\gamma}\underline{q}$ so this reads

$$\gamma^2 \left(\|\underline{r}\|^2 - c^2 \right) = -c^2$$

assuming $\|\underline{r}\|^2 < c$ (beware the tachyons!) this yields

$$\therefore \quad \boxed{\gamma = \frac{1}{\sqrt{1 - \|\underline{r}\|^2/c^2}}}$$

where $\underline{r}$ is the relative velocity of O_2 with respect to O_1. This number is called the Lorentz contraction factor.

Exercise L5-3 Given $\underline{x} \in \mathbb{R}^2$ and $t > 0$, define

$$\|(t,\underline{x})\|_P := \langle (t,\underline{x}), (t,\underline{x}) \rangle_P$$

Show that

(i) $\|(t',\underline{x}') - (t,\underline{x})\|_P = 0$ if and only if a particle travelling at speed c from $\underline{x}$ to $\underline{x}'$ takes $|t-t'|$ units of time. Given $\underline{v} = (t,\underline{x}) \in \mathbb{R}^3$ the set

$$C_{\underline{v}} := \{ \underline{w} \in \mathbb{R}^3 \mid \|\underline{w} - \underline{v}\|_P = 0 \}$$

is called the light cone of $\underline{v}$. Sketch the lightcone of $(0,\underline{0})$ as a set in $\mathbb{R}^3$, using the vertical axis for time.

(ii) $\|(t',\underline{x}') - (t,\underline{x})\|_P \leq 0$ if and only if a particle travelling at speed c for $|t-t'|$ units of time travels at least a distance $\|\underline{x} - \underline{x}'\|$. If $\|\underline{w} - \underline{v}\|_P \leq 0$ we say $\underline{w}$ lies inside the light cone of $\underline{v}$.

(iii) Returning to the language of observers: show using the language of this lecture (i.e. E, m_1, m_2, F) that two observers in constant relative linear motion agree on whether an event $e' \in E$ lies on, or inside, the lightcone of another event $e \in E$. This explains why the language of special relativity is organised around the concept of lightcones.

Exercise LS-4 Prove that the set of invertible matrices $A \in M_3(\mathbb{R})$ with

$$A^T P A = P$$

is a subgroup of $GL_3(\mathbb{R})$ (all invertible matrices). This is called the Lorentz group $O(2,1)$. Equivalently, this is the group of invertible $F: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ s.t. [This notation explained in Tutorial #2]

$$\langle F\underline{v}, F\underline{w} \rangle_P = \langle \underline{v}, \underline{w} \rangle_P \qquad \forall \underline{v}, \underline{w} \in \mathbb{R}^3.$$

The Poincare group is generated by $O(2,1)$ together with translations, and is the full symmetry group of SR.

The Lorentz group classifies the possible relationships between a pair of observers in SR with the same origin (i.e. $F(0,\underline{0}) = (0,\underline{0})$). Let us return to our original pair of observers who had the "same axes". We may as well assume the velocity $\underline{r}$ of observer O_2 as measured by O_1 is $\underline{r} = (0, r, 0)$ (i.e. only in the x-direction). Then one can show the appropriate group element is

$$F = \begin{pmatrix} \gamma & \frac{\gamma r}{c^2} & 0 \\ \gamma r & \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \gamma = \frac{1}{\sqrt{1 - r^2/c^2}}$$

Exercise LS-5 Explain why $F(1,0)$ must be $(\gamma, \gamma r, 0)$ and check $F^T P F = P$.

This transformation is called a Lorentz boost. Obviously we have neglected one spatial direction in the above, but with the obvious modification, i.e. $P = \mathrm{diag}(-c^2, 1, 1, 1)$ we get the Lorentz group $O(3,1)$ of special relativity in $\mathbb{R}^4$.

Exercise L5-6 With F as above, show that if O_2 measures simultaneous events (t, x_1, y_1), (t, x_2, y_2) then provided $r > 0$ and $x_1 \neq x_2$ O_1 does not measure simultaneous events, and the observers disagree about the spatial distance between the events.

Remark The full Lorentz group includes transformations which correspond to rather exotic pairs of observers, e.g. $A = \mathrm{diag}(-1, 1, 1)$ describes a pair of observers moving opposite directions in time!

Def^n Minkowski space is the pair $(\mathbb{R}^4, \langle -,- \rangle_P)$ where $P = \mathrm{diag}(-c^2, 1, 1, 1)$.

The appropriate abstraction for SR is therefore a vector space equipped with a nondegenerate bilinear form $\langle -,- \rangle_P$ of a certain signature (these terms are explained in Tutorial #2), not the notion of metric spaces (as P is not positive definite). In the passage from special to general relativity we require an additional abstraction, namely, topological spaces. We will return to this point in detail in a few lectures, but by way of foreshadowing let us give the geometric/static content of GR, which is

- Spacetime $=$ a four-dimensional connected manifold which is locally isomorphic to Minkowski space.

  i.e. "flat"

  basic concept of topological spaces

  a topological space

We will not spend any more time on relativity per se, but we will spend a lot of time on topological spaces, continuity, connectedness, ... and it is worth knowing at least one prominent example where this language is crucial to expressing the basic properties of physical space. GR is this example. The main takeaway here is:

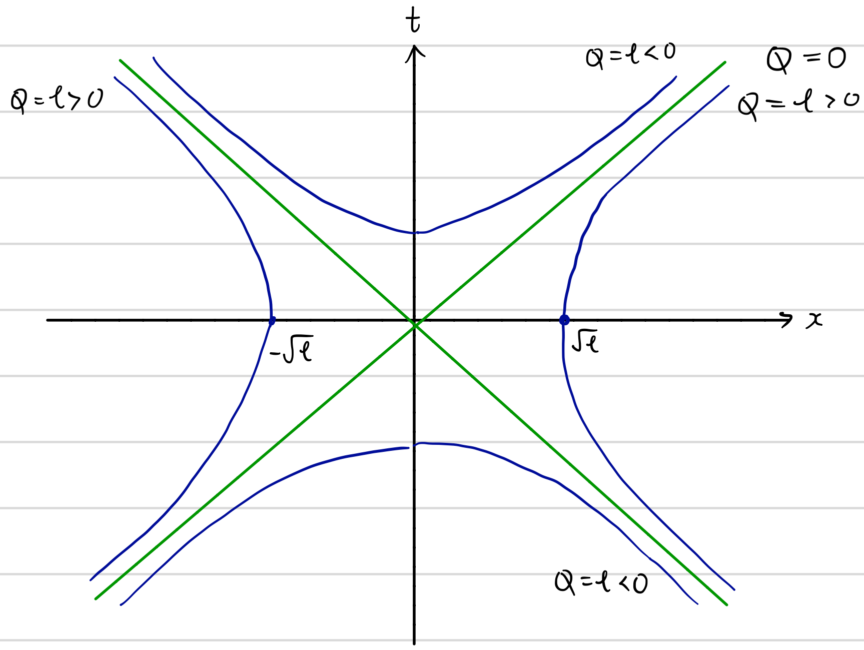
There are no global rulers

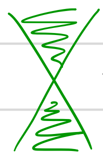
(not even pretend ones, as in Newtonian mechanics)

Hyperbolic geometry

If physics is not your thing, here is an alternative point of view on all of this.
Affine geometry in $\mathbb{R}^2$ is characterised by $E(2) = \mathrm{Isom}(\mathbb{R}^2, d_2)$, by rotations, translations and reflections, and all of this is associated to the bilinear form $\langle -,- \rangle_P$ of the matrix $P = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ which has signature $(2,0)$. The geometry of the bilinear form associated to the matrix $P = \begin{pmatrix} -1 & \\ & 1 \end{pmatrix}$ of signature $(1,1)$ is <u>hyperbolic geometry</u>. In the following we call our coordinate axes (t,x) to match with the SR situation above.

The level sets of the quadratic form $Q(t,x) = -t^2 + x^2$ are, obviously, hyperbolas




is "inside the lightcone"

The level set of $Q(t,x) = t^2 + x^2$ are circles, and since circles of a given radius are how you measure <u>distances</u>, it is no surprise we organise affine geometry around the metric.
To talk about points in hyperbolic geometry we use hyperbolic sine and cosine. There are various ways to explain this (e.g. derive DE's describing a point moving on a branch of the hyperbola at unit speed, as compared to a point moving on the circle).

Perhaps the most relevant observation for us is that with $\theta \in \mathbb{R}$, and

"hyperbolic rotation"
$$H_\theta = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}$$

$\det(F) = \cosh^2\theta - \sinh^2\theta = 1$

$H_0 = I_2$

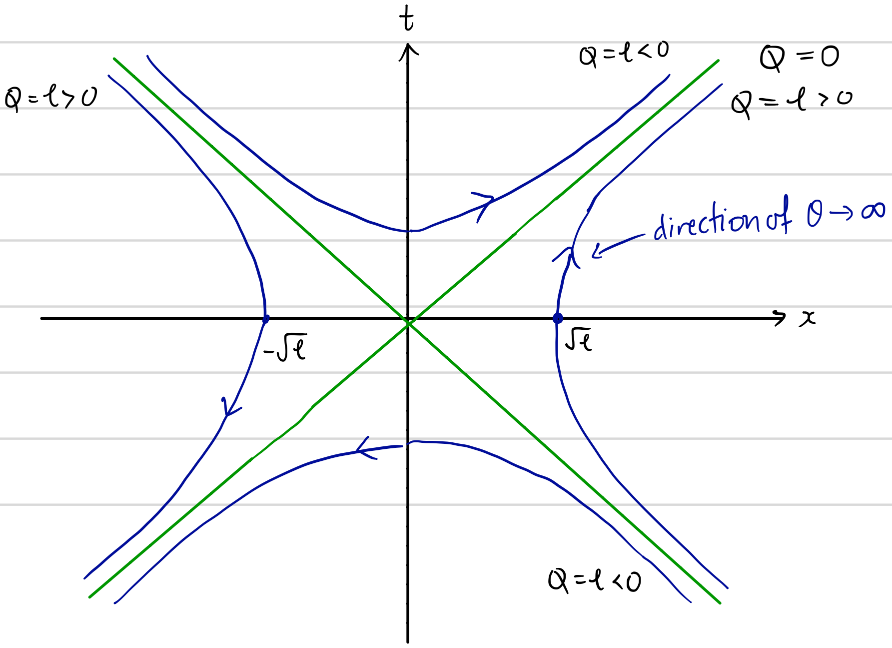
if $-t^2 + x^2 = \ell$ then

$$H_\theta \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} t\cosh\theta + x\sinh\theta \\ t\sinh\theta + x\cosh\theta \end{pmatrix}$$

and $$-\left[t\cosh\theta + x\sinh\theta \right]^2 + \left[t\sinh\theta + x\cosh\theta \right]^2$$

$$\begin{aligned} &= -t^2\cosh^2\theta - 2xt\cosh\theta\sinh\theta - x^2\sinh^2\theta \\ &\quad + t^2\sinh^2\theta + 2xt\cosh\theta\sinh\theta + x^2\cosh^2\theta \\ &= -t^2(\cosh^2\theta - \sinh^2\theta) + x^2(\cosh^2\theta - \sinh^2\theta) \\ &= -t^2 + x^2 \\ &= \ell \end{aligned}$$

So H_θ is a linear isomorphism of the plane (as $\det H_\theta = 1$) which maps each hyperbola $-t^2 + x^2 = \ell$ bijectively onto itself, according to the following schema:



$$H_\theta \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cosh\theta \\ \sinh\theta \end{pmatrix}$$

$$H_\theta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sinh\theta \\ \cosh\theta \end{pmatrix}$$

<u>Exercise L5-7</u> Determine the hyperbolic angle θ s.t. the Lorentz boost F from p.13 is a hyperbolic rotation H_θ. That is, given $0 \leq r < c$ and $\gamma = (1-r^2)^{-1/2}$, solve (here we set $c = 1$)

$$\begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} = \begin{pmatrix} \gamma & \gamma r \\ \gamma r & \gamma \end{pmatrix},$$

for θ. This shows that the geometry that we have extracted from Einstein's postulates is precisely hyperbolic geometry (at least in the (t,x) plane).