Introduction
At the beginning of the 1900's, physicists were struggling with the incompatibility between

i) The relativity principle, which says that the laws of physics must be the same in all inertial frames,

ii) Galilean transformations, that give the rules of going from one inertial system to another,

iii) Maxwell's equations, which describe electrodynamics.

The incompatibility arises because Maxwell's equations are not covariant under Galilean transformations, thus breaking the principle of relativity.

It is in fact easy to see where the problem lies: Maxwell's equations involve explicitly the speed of light $c$ as a parameter. But according to Galilean transformations, the speed of light should be different in different inertial frames that move relative to each other.

Indeed Galilean transformations between IF's with relative velocity $V$ are:

\[
\begin{align*}
\vec{x}' &= \vec{x} - \vec{V}t \\
\vec{t}' &= \vec{t}
\end{align*}
\]

But if $|\vec{V}| = c$ in one frame, the speed of light must be $|\vec{V}'| = |\vec{V}+\vec{V}'|$ in another. More formally, you can check, e.g., that the form of the wave equation for electromagnetic potentials is not invariant (i.e., changes form) when one goes from one IF to another using Galilean transformations.

In fact, it is not difficult to see that assuming Maxwell's equations in different inertial frames coupled with the invariance of distance and absolute time accepted in Newtonian mechanics...
leads to fundamental inconsistencies.

To see this, consider a simple system and let's look at it in two different IF's. The system is made of 2 cylinders, one with equal + & - charges and the other positively charged. In both cylinders, positive charges move with the same velocity,

\[ \text{charge to IF.} \quad \text{moving uniformly} \quad \text{with + charges} \]

\[ \text{using} \quad G.T. \]

Here, electric force is zero because change in upper cylinder remains, only magnetic (attractive) force due to parallel currents!

In the IF, where + charges are at rest, there is no interaction between the cylinders, where in the original frame there is, clearly a contradiction. We'll see later how this problem is fixed in S.R.

The leading explanation at the time was that the principle of relativity didn't hold for electromagnetism, that there was a special frame at rest with a mechanical medium, the "ether", which made possible light propagation. Thus, Maxwell's eqs were only applicable in this preferred frame. In another IF, one should see the speed of light vary as dictated by Galilean transformations. Therefore, one could use electromagnetic waves to find out, for any system, its velocity with respect to the ether, and thus determine its "absolute" state of motion. All such experiments failed, however.

The most famous of such experiments is perhaps the Michelson-Morley
Einstein instead took a different point of view: the pole of relativity is valid, two Maxwell eqs are valid in all I.F.'s (unaltered), but it is just the Galilean transformations that need to be replaced by Lorentz transformations (which preserve the form of Maxwell eqs when going from one I.F. to another). These transformations lead to completely different notions of how space and time are related to each other leading to the concept of spacetime. They are of course consistent with the same value of the speed of light in all inertial frames, independent of the state of motion of the emitting body.

Because the Newtonian notions of space and time are changed, the laws of mechanics need to be changed as well. And finally, as we already mentioned, changing G.T. to Lorentz transformations leads to another problem: since Newtonian gravity is only invariant under G.T., it now violates the pole of relativity! The search for a gravitation theory consistent with S.R. lead to G.R.

(> invariant under Lorentz)

The postulate that the speed of light is the same in all inertial frames lead to nontrivial modifications to the concepts of space and time (e.g., the relativity of simultaneity). Consider a simple example, an observer O at the center of a rocket that receives light signals from the two ends of the spaceship.
He concludes that the events @ A & B (departure of light signal) must be simultaneous.

The fact that the two light signals touch O at the same time is indisputable (an event is an event, everyone agrees that it happens no matter their state of motion), but looked from the Earth's point of view leads to different conclusions about the simultaneity of A & B:

For O', Light from A must depart earlier than from B, because O is moving away from the original @ departure) position of A and, on the other hand, approaching the original position of B. This is because the speed of light is the same for O' than for O.

Therefore, O' concludes the event @ A happened earlier than the event of B.

**Spacetime**

The existence of a finite invariant speed (the speed of light) leads to different notions of space and time for different observers in different inertial frames. Therefore it is convenient to introduce spacetime coordinates @ at each inertial frame (t, x, y, z) to characterize events (things that happen at given time & space). A different observer in motion w/respect to such inertial frame will have its own spacetime coordinates (t', x', y', z') "attached" to its own inertial frame.

In analogous way to newtonian mechanics, where Galilean transformations conserve the distance between points, in
special relativity the lorentz transformations that relate
\((t, x, y, z)\) to \((t', x', y', z')\) conserve a new kind of distance (called
the interval) between events, only that it mixes space with
time (that is why space-time is a useful concept). To motivate
the form of the interval consider a simple example of light
propagation between parallel mirrors:

\[
\begin{align*}
\Delta t &= \frac{2L}{c} \quad \text{(time diff between events A & C)} \\
\Delta x &= \Delta y &= \Delta z = 0
\end{align*}
\]

On the other hand in a frame where the mirrors are moving we
obtain from the invariance of the speed of light that,

\[
\begin{align*}
\Delta t' &= t_c' - t_a' \quad \Delta x' = \Delta y' = \Delta z' = 0 \\
(\Delta t')^2 &= (2L)^2 + (\Delta x')^2 \\
\text{constancy of speed of light}
\end{align*}
\]

Note: here we are assuming distances perpendicular to motion (L) are
unchanged in the frame moving with vel \(V\) — it's easy to see that
must be the case: consider 2 cylinders, one white the other gray (same
size), moving towards each other:

In this frame, when they collide, they don't go through each other:

and in rest frame of gray cylinder this situation:

Clearly, this does not make sense \(\Rightarrow\) transverse directions
are invariant in different frames.
Then we have,

\[ A_s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 - (c\Delta t)^2 = \frac{(\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 - (c\Delta t')^2}{(2L)^2} = A_s'^2 \]

In differential form, we have the line element

\[ ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \]

which is an invariant independent of inertial frame.

This is analogous to concept of distance in Newtonian mechanics, where

\[ ds^2 = dx^2 + dy^2 + dz^2 \]

is invariant,

The line element characterizes the geometry of spacetime, and it is non-Euclidean (due to the - sign in time interval) but also flat (as we will characterize later when we discuss spacetime curvature). It is known as Minkowskian spacetime or flat spacetime and can be formalized a bit more by introducing a metric tensor

\[ \eta_{\mu\nu} \quad ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \]

\[ \eta_{\mu\nu} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \]

where \( dx^\mu = (c dt, dx, dy, dz) = (c dt, d\vec{x}) \)

(some books use different signature for the interval, or \( \eta_{\mu\nu} \), i.e. \( \eta_{\mu\nu} = diag(+,-,-,-) \), that's just convention.)

We'll get back to the metric later when we discuss GR, in which the metric takes the role of a dynamical variable, not fixed constants like in SR.

The non-Euclidean character of the geometry is obvious since the "invariant distance" (the interval) between different events
can be positive, zero, or negative!

Since \((\Delta s)^2\) is an invariant, all observers agree on its sign, so it makes sense to classify a pair of events according to the sign of their \((\Delta s)^2\):

\[
\begin{align*}
\Delta s^2 > 0 & \Rightarrow \text{space-like} : \text{there is an IF where events happen at the same time: } \Delta t = 0 \\
\Delta s^2 = 0 & \Rightarrow \text{light-like, or null} : \text{events connected by light} \\
\Delta s^2 < 0 & \Rightarrow \text{time-like} : \text{there is an IF where events happen at the same position } \Delta x = \Delta y = \Delta z = 0
\end{align*}
\]

The "boundary" \(\Delta s^2 = 0\) plays an important role in S.R. as it separates events that can be simultaneous (space-like) from events that can be connected by particle motion (time-like).

All events from a given one (adjacent point \(P\)) that have \(\Delta s^2 = 0\) form the lightcone @ \(P\):

- \(A\) & \(C\) are inside the lightcone of \(P\),
- \(A\) in the future light cone, \(C\) in the past light cone.

Events inside the lightcone are causally connected to \(P\), since they are time-like separated \(\Delta s^2 < 0\) and can be connected by particle motion with speed less than \(c\).

\(B\) is outside the lightcone and is causally disconnected from \(P\). It has a space-like \(\Delta s^2 > 0\) separation. You need faster than light propagation to connect \(B\) to \(P\).

Clocks measure timelike distances, ruler measure spacelike distances.

Timelike distances \(\Delta s\) that correspond to events that occur at the same point in space are special, since \(\Delta s = 0 \Rightarrow -c^2 \Delta t^2 = \Delta s^2\).
...and then they are invariant. These are known as proper time, the time measured by a clock carried along the world line of a particle.

[Diagram 1: Light rays passing through a triangle with $ct$ and $x$ axes labeled.]

[Diagram 2: World line of a massive particle (speed < c) with $ct$ and $x$ axes labeled.]