Physics in Spacetime (F4051)

Lecture notes

Linus Wulff

Spring 2025

Chapter 1

Space and Time

This course deals with the special theory of relativity introduced by Einstein in a famous 1905 paper. The traditional way of introducing special relativity is to derive it, in much the same way that Einstein did, from two basic principles:

1. **The principle of relativity**
2. **The constancy of the speed of light**

From these assumptions the notion of a spacetime with (inertial) observers being connected by Lorentz transformations follows. This is a natural way to proceed if one starts from a knowledge of classical mechanics and Maxwell's equations of electrodynamics. However, it is not the best way to understand the geometrical aspects of spacetime. This is part of the reason it took Einstein another ten years to formulate the general theory of relativity, describing gravity, where the geometry of spacetime is the key player.

In this course we will follow a different route[1] which leads more directly to a geometric picture. Rather than starting with the principles described above we will derive the same physics from what is known as

- **The principle of maximum proper time**

This approach is more in line with general relativity, and this course can be thought of as a first step towards studying general relativity.

1.1 What is space and time?

The notions of Space and Time are central to physics. In physics we are interested in answering questions like

Given some configuration of particles with given positions and momenta at some initial time t_i, what will the configuration look like at some later time t_f?

[1]The approach to relativity taken here is inspired by lecture notes and a book by B. Laurent (*Introduction To Spacetime*, World Scientific, 1994).

For such questions to make sense we must have a precise way of defining what we mean by time and also what we mean by a particle's position in space. So what are space and time? Rather than get into a philosophical discussion about the nature of space and time, a more useful approach when faced with such deep questions in physics is to try to replace it by a different, more down to earth, question. After all, in physics we deal only with things that can be measured, and therefore a better question to ask is

How do we **measure** *space and time?*

We say that we define space and time *operationally*, by declaring how we measure them.

So how do we measure distances in space? The most basic way is to take a reference object, say stick of a certain length, and use it to measure the distance between two points. We will call such a reference object a *ruler*. Of course, a good ruler should not bend or change its length with temperature etc., so we will assume it is always possible to find a sufficiently good ruler (or equivalent) so that we can measure lengths to the precision we need. How do we measure time? To measure time we need a clock. It does not have to be what we normally think of as a clock, it can be any physical process which has a known time dependence, e.g. a periodic process with definite period like a pendulum or a non-periodic process like an atom in an excited state with known half-life. Again, we will assume that there exist such clocks with good enough precision for the time measurements we need to perform.

We allow each person, or *observer*, to measure time with their own clock and spatial distances with their own ruler. We will assume that these are small enough that the observer can carry them with her, i.e. they will be assumed to be in the same state of motion as the observer and experience the same forces she experiences. But if each observer makes their own measurements using their own clock and ruler, how do we relate the measurements of two *different* observers? Newton and his contemporaries assumed that there was an absolute notion of time, so that all observers clocks would tick at the same rate. In that case it is very easy to relate the measurements of two observers. We now know that this assumption was wrong. For example, taking two synchronized atomic clocks, putting one on a plane circling the earth and leaving one on the ground, one finds when comparing them at the end that they differ (by a few hundred nanoseconds). This observation is clearly inconsistent with the Newtonian idea of an absolute time.

1.2 The principle of maximum proper time

Experiments show that time runs differently for different observers. We must therefore assign each observer their own time, their *proper time*, which is the time measured on their clock. We can now state the key principle that will allow us to compare the measurements of different observers

The principle of maximum proper time:

> *If two observers are separated and then meet again, the one that does not experience any acceleration always measures the* **longest** *proper time.*

It says that proper time is maximized for *inertial*, i.e. unaccelerated, observers. There is plenty of experimental evidence to support this principle, such as the experiments with atomic clocks on planes, or the operation of GPS satellites which requires very precise time measurements. In this course we will take this principle as the starting point from which we will derive the theory of special relativity.

1.3 Spacetime

We are familiar with the fact that to specify the position of an object in our three dimensions we need to give three numbers – the coordinates with respect to some specified coordinate system. For positions on the earth we might for example give the longitude, the latitude and the height above sea level. To specify an *event* – something happening at a certain place at a certain instant of time – we must give one more number, namely the time on a clock associated to the coordinate system. In our example this could be the time GMT.

We have argued that we must allow each observer to measure distances and times using their own coordinate system defined by their ruler and clock. Each observer will therefore associate to a given event four numbers (t, x, y, z) – the *spacetime coordinates* relative to their coordinate system. Note that we are defining an event here in an idealized way as a single point in spacetime, i.e. something that happens at a point in space at a single instant of time. The set of all events make up the four-dimensional *spacetime.* Note that each observer will (in general) assign different coordinates to the same event because they are using different coordinate systems, there is no preferred coordinate system in spacetime. One of our first tasks will therefore be to understand how to relate the observations of different observers.

1.4 Worldlines

The trajectory of an object traces out a continuous path in spacetime – a *worldline* (really "worldtube" if the object is not point-like, but this distinction won't be very important to us). In ordinary Euclidean space we are familiar with the fact that there is a shortest path between any two points. This path is called a straight line. It is the path an object follows if it is not acted upon by any external forces, i.e. it is unaccelerated. Similarly, we will assume that there is precisely one straight line connecting any two events in spacetime and that any object not acted on by external forces, i.e. not experiencing any acceleration, follows such a straight worldline. To a worldline connecting two events in spacetime we can associate a number – the proper time along that worldline. Recall that this is the time an observer traveling along the worldline measures on her clock between the two events. The principle of maximum proper time says that a straight worldline corresponds to the *longest* proper time. Therefore the analog of shortest length in Euclidean space is longest proper time in spacetime and a clock can be thought of as measuring distances in spacetime. When we draw *spacetime diagrams* we will draw the worldlines of unaccelerated objects as straight lines. Curved lines will correspond to worldlines of accelerated objects (Figure 1.1).

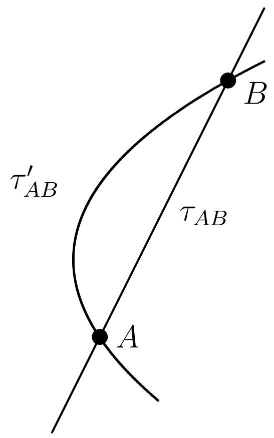

Figure 1.1: Spacetime diagram showing an accelerated (curved worldline) and an unaccelerated (straight worldline) observer meeting at events A and B. The proper time measured on their respective clocks between the meetings is τ'_{AB} and τ_{AB}. The principle of maximum proper time then says that $\tau_{AB} > \tau'_{AB}$.

An important notion in Euclidean geometry is the notion of two lines being parallel. In spacetime we can similarly have the notion of two observers being on the same course. How can two observers, e.g. two spaceships traveling in outer space, determine whether they are on the same course? One way to do this uses a construction from Euclidean space adapted to spacetime. Imagine that the two observers each send out a probe fitted with a clock, which travels freely until it is picked up by the other observer at some later time (Figure 1.2). If the two probes happen to meet halfway, i.e. after half of the proper time (from being emitted to being picked up) has elapsed on each clock, then we will say that the observers are on the same course, or that their worldlines are *parallel*. From the figure we see that this also implies that the lines AB and CD (not drawn) are parallel. Note that to carry out the experiment we really need to send clocks that also have a recording device that records the time they were sent and the time they met. We would also need to do the experiment several times to get them to meet halfway.

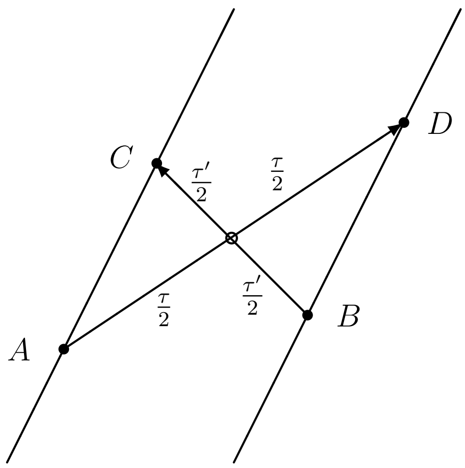

Figure 1.2: The worldlines of two observers are parallel if they can send out probes, at A and B, that meet halfway before encountering the other observer at D and C.

Notice that this construction does not refer to space or time separately, only to the full spacetime picture or the proper time measured by a particular clock. This is in sharp contrast to how we would describe such an experiment in Newtonian physics.