Lecture 1. Classical space-time

1.1 The problem

This course describes Einstein's resolution of the conflict between:

1) The *principle of relativity* in classical mechanics; and

2) Maxwell's *electromagnetic theory*.

The first two lectures will explain why they are in conflict. We then turn to the resolution, which does not involve changing (1) or (2), but rather the introduction of a new way of fitting space and time together into a single geometric structure.

1.2 Galileo's principle of relativity

Imagine two spaceships passing each other in empty space. In each, the passengers think that they are at rest and that it is the other spaceship that is moving. Is there some physical test that will determine who is right? Certainly there is if one is accelerating and the other is not since it is possible to *feel* acceleration. But if both are moving uniformly at constant speed, then there is not: no device within one of the spaceships will detect uniform motion, only motion relative to some external standard, such as the earth or the 'fixed stars'. The *principle of relativity* turns this from a negative to a positive statement:

Principle of Relativity

In classical mechanics, all frames of reference in uniform motion are equivalent.

We shall consider how to give this a precise formulation within the context of Newton's laws.

1.3 Inertial Frames

In order to describe the motion of a system of particles, we must first choose an inertial frame R; that is, an origin together with a (right-handed) set of Cartesian axes. So as not to prejudge the issue, we must avoid the temptation to think of R as 'fixed in space'.

The *position*, *velocity*, and *acceleration* of a particle relative to R are the vectors $\boldsymbol{r}$, $\boldsymbol{v}$, and $\boldsymbol{a}$ with components

$$(x, y, z), \quad (\dot{x}, \dot{y}, \dot{z}), \quad (\ddot{x}, \ddot{y}, \ddot{z})\,, \tag{1}$$

where x, y, z are the coordinates of the particle and dot is the derivative with respect to time.

DEFINITION 1 (INERTIAL FRAMES) *The frame R is* inertial *if the motion of any system of particles relative to R is governed by Newton's laws.*

1.4 Coordinate transformations

Suppose that R and R' are frames of reference (which may be moving or rotating relative to each other).

The coordinates of a particle in the two frames are related by an affine linear transformation—that is, a combination of a linear transformation and a translation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = H \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + T \tag{2}$$

where

$$H = \begin{pmatrix} H_{ij} \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \tag{3}$$

is a proper orthogonal matrix, representing a rotation and

$$T = \begin{pmatrix} T_i \end{pmatrix} = \begin{pmatrix} T_1 \\ T_2 \\ T_3 \end{pmatrix} \tag{4}$$

is a column vector, representing a translation. The H_{ij}s and T_is are functions of time t.

If vector $\boldsymbol{A}$ has components (a, b, c) in R and components (a', b', c') in R', then

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = H \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} . \tag{5}$$

Note that this is a *linear transformation*—there is no T-term.

1.5 Galilean transformations

By differentiating twice, we find that the components $(\ddot{x}, \ddot{y}, \ddot{z})$ of the acceleration $\boldsymbol{a}$ of the particle relative to R are related to the components $(\ddot{x}', \ddot{y}', \ddot{z}')$ of the acceleration $\boldsymbol{a}'$ relative to R' by

$$\begin{pmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{pmatrix} = H \begin{pmatrix} \ddot{x}' \\ \ddot{y}' \\ \ddot{z}' \end{pmatrix} + 2\dot{H} \begin{pmatrix} \dot{x}' \\ \dot{y}' \\ \dot{z}' \end{pmatrix} + \ddot{H} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \ddot{T}\,, \tag{6}$$

where $\dot{H} = (\dot{H}_{ij})$ and so on. Thus $\boldsymbol{a} = \boldsymbol{a}'$ for every possible motion of the particle if and only if $\dot{H} = 0$ and $\ddot{T} = 0$. That is, if and only if R' is moving relative to R without rotation or acceleration. If $\dot{H} = 0$, $\ddot{T} = 0$, then Newton's laws hold in R' if and only if they hold in R.

We can also change the origin of the time coordinate t as well as transforming x, y, z coordinates. We then we have a transformation that changes the space and time coordinates from t, x, y, z to t', x', y', z', where $t = t' +$ constant, and x, y, z and x', y', z' are related by (2), with H a constant matrix and (since $\ddot{T} = 0$),

$$T = Vt + C \tag{7}$$

for some constant column vectors V and C. In four-dimensional form, this becomes the following.

Definition 2 *A* Galilean transformation *is a coordinate transformation*

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ v_1 & H_{11} & H_{12} & H_{13} \\ v_2 & H_{21} & H_{22} & H_{23} \\ v_3 & H_{31} & H_{32} & H_{33} \end{pmatrix} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} + \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} \tag{8}$$

where v, H, and C are constant, v is a column vector of length 3, H is is a 3×3 proper orthogonal matrix, and C is a column vector of length 4.

Examples

Rotations. If $C = 0$, $v = 0$, then

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & H_{11} & H_{12} & H_{13} \\ 0 & H_{21} & H_{22} & H_{23} \\ 0 & H_{31} & H_{32} & H_{33} \end{pmatrix} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} .$$

In this case, $t = t'$ and the x, y, z and $x'y', z'$ coordinates are related by a rotation of the axes. The frames are at rest relative to each other.

Boosts. If $H = 1$, $C = 0$, then $t = t'$ and

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \begin{pmatrix} v_1 t \\ v_2 t \\ v_3 t \end{pmatrix} .$$

The axes are parallel and the origins coincide at $t = 0$. The frame R' moves with constant velocity (v_1, v_2, v_3) relative to R.

Translations. If $v = 0$, $H = 1$, then coordinate systems are related by a translation of the origin and a resetting of the zero of t.

Every Galilean transformation is a combination of a rotation, a boost, and a translation (exercise).

In the classical picture, the coordinate systems of any two inertial frames are related by a Galilean transformation, and any frame related to an inertial frame by a Galilean transformation is itself inertial. There is nothing in the laws of classical mechanics that picks out a particular inertial frame. The principle of relativity is that all inertial frames are equivalent.

1.6 Newton's fourth law

The principle of relativity was first clearly formulated by Galileo, who used as illustrations not the relative motion of spaceships, but the behaviour of butterflies and fish in the cabin of a sailing ship. The principle also played an important part in Newton's thinking: in a manuscript [1] that he wrote two and half years before his *Principia*, he had not three, but six laws of motion. The fourth was the principle of relativity.

Lex 4

The motion of bodies amongst themselves in a given space is the same whether that space is absolutely at rest or moves in a straight line without circular motion.

Newton realised that the fourth law was a consequence of the first three (the three laws we know today); but he had other reasons for believing in an absolute standard of rest, which remains 'always similar and immovable'. In the *Principia*, the fourth law is reduced to the status of a corollary to the laws of motion.

1.7 Space-time

We must get used to thinking of space and time together as a four-dimensional space, rather than as separate.

DEFINITION 3 Space-time *is the set of all events. An* event *is a particular point at a particular time.*

Events are labelled by t (time) and by the three Cartesian coordinates x, y, z. Thus we have a system of four coordinates t, x, y, z on space-time. A coordinate system in which Newton laws hold is called an *inertial coordinate system* (ICS). In classical mechanics, two ICSs are related by a Galilean transformation.

Let A, B be two events in space-time and consider the following statements.

1) A and B are simultaneous;
2) B happens time t after A;
3) A and B are simultaneous and separated by a distance D;
4) A and B happen in the same place (at different times);
5) A and B are separated by distance D, but happen at different times.

[1] *De motu corporum in mediis regulariter cedentibus*

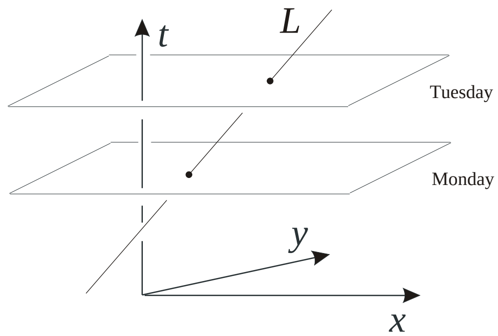

Figure 1: The space-time of Galilean relativity

The first three of these are *invariant*. If one of them is true in one system of inertial coordinates, then it is true in every system. The last two, however, are not invariant. Suppose, for example, that A is '2:00 pm in Oxford' and B is '3 pm in Oxford '. Then in an ICS in which the earth is fixed, (4) is true; but in an ICS in which the sun is at rest, A and B are separated by some 70,000 miles, since the earth is moving relative to the sun at some 19 miles per second.

With one space dimension suppressed, we can picture space-time as in figure 1, which is an example of a *space-time diagram*. Here the time axis points upwards, and the horizontal planes represent space at different times. The straight line L represents the history of a particle moving in a straight line at constant speed (the greater the slope, the *lower* the speed). In general, a curve in space-time representing the history of a particle is called a *worldline*.

With only one space dimension, a Galilean transformation takes the form

$$\begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix} \begin{pmatrix} t' \\ x' \end{pmatrix} + \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} .$$

This is illustrated in figure 2. Note that the lines of constant t and the lines of constant t' coincide, reflecting the invariance of statement (1) above. The central difference between the space-time of classical physics and Einstein's space-time is that (1) is *not* invariant in Einstein's theory: simultaneity is relative.

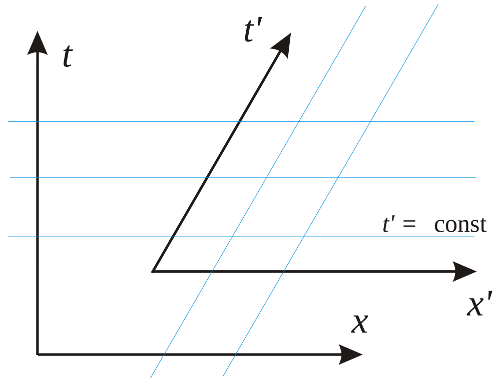

Figure 2: A Galilean transformation in two-dimensional space-time

Lecture 2. Einstein's special theory of relativity

2.1 Maxwell's theory and Galilean relativity

Maxwell's equations and the Lorentz force law give a complete and remarkably accurate description of the dynamical behaviour of charges and electromagnetic fields over an enormous range of physical phenomena, from elementary particle physics up to processes involving astronomical distances. The field equations require modification only where quantum mechanical effects are important.

This much of 'Maxwell's theory'—the field equations and the force law—is the accepted basis of modern electrodynamics. But the theory originally contained much more: it included a 'mechanical' model in which electrodynamic waves were transmitted through a medium called the *ether*. The ether hypothesis was central to Maxwell's own thinking, but it became increasingly implausible as various attempts to detect the ether ended in failure. It was finally removed from the theory by Einstein, for reasons that we shall now consider in detail.

To understand the consequences of the ether hypothesis and the reasons for rejecting it, we must consider the description of electromagnetic phenomena in moving frames of reference.

We might expect a good electromagnetic theory to be consistent with the principle of relativity. In other words, if we are given two inertial frames in relative motion, we might expect to be able to combine the Galilean transformation between them with a transformation of the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields in such a way that the basic electromagnetic equations remained invariant: if they held for the original fields in the first frame, then they would also hold for the transformed fields in the second frame.

Consider, for example, a charge e moving (slowly) with velocity $\boldsymbol{v}$ through a constant magnetic field $\boldsymbol{B}$. It is acted on by a force $e\boldsymbol{v} \wedge \boldsymbol{B}$. An observer moving through the field with the same velocity $\boldsymbol{v}$ will see the same force, but now acting on a charge at rest. He will conclude that there is an electric field $\boldsymbol{v} \wedge \boldsymbol{B}$ present. This suggests that the Galilean transformation to the observer's inertial frame should be accompanied by a transformation of the electric and magnetic fields which sends a pure magnetic field to a combination of a magnetic field and an electric field $\boldsymbol{E} = \boldsymbol{v} \wedge \boldsymbol{B}$.

This idea very nearly works, but not quite. It is possible to write down a transformation law for $\boldsymbol{E}$ and $\boldsymbol{B}$ which makes Maxwell's equations invariant under Galilean transformations provided that we ignore terms of order v^2/c^2 and consider only slowly varying fields; but if we allow Galilean transformations between two frames moving relative to each other at a velocity comparable to the speed of light, then the invariance is lost, as one can see by considering the behaviour of photons.

2.2 Photons and the Michelson-Morley experiment

It follows from Maxwell's equations that photons (wave packets of light) move with velocity c. Relative to a frame R in which the equations hold, therefore, a photon has velocity (u_1, u_2, u_3) where $u_1^2 + u_2^2 + u_3^2 = c^2$.

Relative to a second frame R', related to R by the Galilean transformation

$$t = t', \quad x = x' + vt', \quad y = y', \quad z = z', \tag{9}$$

its velocity has components

$$(u_1', u_2', u_3') = (u_1 - v, u_2, u_3) . \tag{10}$$

If $u_2' = u_3' = 0$, then $u_1 = \pm c$ and $u_1' = \pm c - v$. If $u_1' = u_2' = 0$, then $u_1 = v$ and

$$u_3 = u_3' = \pm\sqrt{c^2 - v^2} .$$

Thus a photon moving in the positive (negative) x' direction has velocity $c - v$ $(c + v)$ relative to R'; and a photon moving perpendicular to the x' direction has velocity $\sqrt{c^2 - v^2}$ relative to R'.

The time taken for a photon to make the round trip from the origin to the point $x' = D$ on the x'-axis is

$$\frac{D}{c+v} + \frac{D}{c-v} = \frac{2D}{c}\left(1 + \frac{v^2}{c^2}\right) + O(v^4/c^4) .$$

For the round trip from the origin to $y' = D$ on the y'-axis and back, it is

$$\frac{2D}{\sqrt{c^2 - v^2}} = \frac{2D}{c}\left(1 + \frac{v^2}{2c^2}\right) + O(v^4/c^4) .$$

In R', the velocity of photons is direction-dependent and so Maxwell's equations cannot hold in both R and R'. (We shall see later that the false step in this argument is the use of the Galilean transformation (9), rather than the Lorentz transformation that we shall meet later).

This is, of course, consistent with the ether hypothesis, according to which Maxwell's equations hold only in the rest frame of the ether; in other frames the velocity of photons is

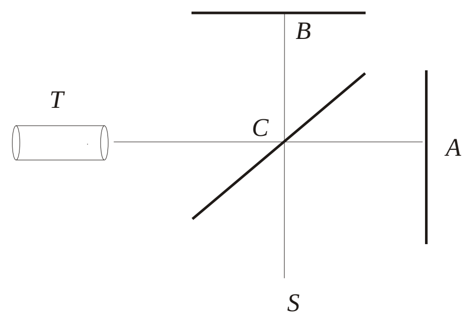

Figure 3: The Michelson-Morley experiment

direction-dependent, just as the velocity of sound is direction-dependent in a frame which is not at rest relative to the air. The problem is that it is inconsistent with experiment and with common sense, for the following reasons.

1. Mechanical forces (for example, between two colliding bodies) are electromagnetic in origin. It is unreasonable that the principle of relativity should apply to mechanics, but not to the underlying electromagnetic processes.

2. It is strange that an inconsistency between Galilean relativity and electrodynamics should emerge for rapidly changing fields and rapidly moving frames of reference, when many simple electromagnetic phenomena show a striking indifference to motion through the ether. Einstein gave the example of the current induced in a conductor by a moving magnet: it is the same whether the conductor is at rest and the magnet moving or the magnet at rest and the conductor moving.

3. In 1887, Michelson and Morley attempted to detect the direction-dependence of the velocity of light in a frame moving relative to the ether. Simplifying a little, their apparatus consisted of a slit source S of light of wavelength λ, two plane mirrors A and B, a half-silvered mirror C and a telescope T (figure 3). The light from S was split into two beams at C. One arrived at T after reflection from B and C, the other after reflection at C and A. The two beams interfered at T, producing a pattern of fringes which could be seen in the telescope.

The whole apparatus was attached to a stone disc, which could be rotated. If the earth had been moving through the ether, then the velocity of light in the frame of the earth would have been direction-dependent.

If in the original configuration the earth had velocity v relative to the ether in the direction CA, then the time taken for light to travel the route CAC would have decreased by

$$\frac{v^2 CA}{c^3} + O(v^4/c^4)$$

when the stone was then rotated through 90^o; similarly, the time for route CBC would have increased by

$$\frac{v^2 CB}{c^3} + O(v^4/c^4)\,.$$

Thus, to the second order in v/c, the rotation should have produced an effect equivalent to increasing the difference in the total distances travelled by the two routes by

$$n = \frac{v^2}{c^2\lambda}(CA + CB) \tag{11}$$

wavelengths. There should have been a corresponding shift in the interference pattern.

No fringe shift was observed, suggesting either that the earth dragged the ether along with it in its orbit around the sun (a possibility that was rapidly ruled out) or that the assumptions underlying the derivation of eqn (11) were incorrect.

H.A. Lorentz suggested a way out: he argued that a rigid body (in the case of the Michelson-Morley experiment, the stone) should contract by a factor $\sqrt{1 - v^2/c^2}$ along the direction of its motion through the ether as the result of a supposed effect of the motion on the electromagnetic forces between the particles making up the body. The contraction would reduce the fringe shift to the extent that it would be unobservable. The idea was extended by Lorentz after the failure of other ether detection experiments; and was put by Poincaré into the form of a fundamental principle that no experiment could detect motion through the ether: any effect that might be detectable would be exactly cancelled by an equal compensating effect (there was an analogy with Newton's third law: to every action there is an equal and opposite reaction).

2.3 Operational definitions of distance and time

According to Poincaré's point of view, the ether exists, but cannot be detected. This is not a good starting point for a physical theory. It also raises a still more awkward problem: it raises doubts about the physical meaning of distance measurements. A moving measuring rod contracts by a factor $\sqrt{1 - v^2/c^2}$. Thus only measurements made in a frame at rest relative to the ether are valid. But if the ether is undetectable, how is distance to be measured?

Einstein saw that the problem lay not with Maxwell's equations nor with the principle of relativity, but with the uncritical acceptance of intuitive ideas about the nature of space and time. He saw that one cannot resolve the issues raised by the Lorentz contraction without first explaining what is meant by 'length' in terms of the process of measurement of distance: that is, 'distance' must be given an *operational definition*—a definition in terms of the operations required to measure it.

Before tackling distance, however, it is necessary first to give an operational definition of 'simultaneity': even according to the classical view of space and time, distance is only defined between simultaneous events (see the first lecture: what is the distance between Oxford at 2 pm and London at 3 pm. measured in frame fixed relative to the sun?)

Einstein began his analysis of electrodynamics in moving frames by giving operational definitions of 'distance' and 'simultaneity' which were consistent with the principle of relativity and with the validity of Maxwell's equations in all uniformly moving frames of reference. Of course, the time and distance coordinates in different frames could not be then related by Galilean transformations.

We shall use very slightly different, but essentially equivalent, operational definitions due to Milne. In Milne's approach, one takes as fundamental 'clocks' and 'light signals'. Every observer carries a clock with which he can measure the time of events in his immediate vicinity; and observers can send out and receive light signals, which are carried by photons (particles of light). The vibrations of a single atom can be used to measure time, so a clock is intrinsically a much simpler object than a measuring rod made up of a very large number of atoms.

If Maxwell's equations are to hold in every frame then the definitions of distance and simultaneity must be such that the following holds.

> $(*)$ The velocity of photons is the same irrespective of the motion of their source or of the observer.

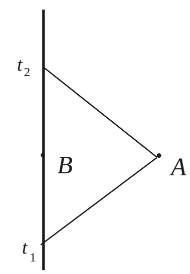

Figure 4: Milne's definitions

Definition of simultaneity

Suppose that a non-accelerating observer sends out a light signal at time t_1 (measured on his clock). This is received at an event A and immediately transmitted back to the observer, arriving at time t_2 (again measured on the observer's clock). Which event B at the observer is simultaneous with A? If $(*)$ is to hold, then the journeys of the outgoing and return photons will be reckoned by the observer to have equal duration, and so the observer will take B to be the event at his location which happens at time $\frac{1}{2}(t_1 + t_2)$ and he will assign this value of t to A. This is the *radar* definition of simultaneity. (See figure 4.)

Definition of distance

The observer similarly assigns a distance $\frac{1}{2}c(t_2 - t_1)$ from B to A. Here c is a constant which is chosen arbitrarily (but is given the same value by all observers). If t is measured in seconds and c is chosen to be 3×10^8 , then the unit of distance is called the meter. It t is measured in years and $c = 1$, then the unit of distance is the light-year; and so on.

By defining distance and simultaneity in this way and by observing the direction from which light signals arrive, a non-accelerating and non-rotating observer can in principle

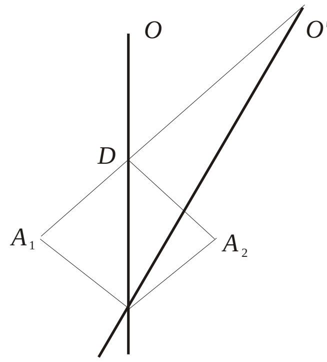

Figure 5: The relativity of simultaneity

set up a Cartesian coordinate system and label each event by x, y, z, and t. These labels are called inertial coordinates. We shall always assume that the observer takes his own location as the origin of the spatial coordinates.

2.4 The relativity of simultaneity

The advantage in accepting these definitions is that $(*)$ automatically holds, so the null result of the Michelson-Morley experiment is no longer a problem. The disadvantage is that we must also accept various consequences which are contrary to intuition. In particular, we must accept that simultaneity is relative. Two events which are reckoned to be simultaneous by one observer O may not be simultaneous according to a second observer O' moving relative to O.

This is easiest to see from the space-time diagram , figure 5 Here each point represents an event and time increases up the page. The two solid lines are the histories of O and O' and the dotted lines are the histories of photons (more prosaically, one can think of the various lines as graphs of time as functions of position).

Two light signals are transmitted from the event C where O' passes O. They are reflected at A_1 and A_2 and both arrive back at 0 at the event D. The observer O therefore judges the two events A_1 and A_2 to be simultaneous. However, photon from A_2 reaches O' before that from A_1, so O' assigns an earlier time to A_2 than to A_1: he judges that the two events are not simultaneous.

Lecture 3. Two-dimensional Lorentz transformations

3.1 Bondi's k-factor

Consider two observers O and O' travelling along a line with constant speed. They pass each other at event E and then are move directly away from each other. They both set their clocks to zero at E.

By using the radar definitions, they both set up inertial coordinate systems on two-dimensional space-time. We shall derive the relationship between these systems by making two assumptions.

(i) They both reckon that the velocity of light is c.

(ii) Only their relative motion is observable.

. Consider a photon emitted by O towards O' at time t (measured by the O clock). Suppose that it is received by O' at time $t' = kt$ (measured by the O' clock). The quantity k is *Bondi's k-factor*. Since neither observer is accelerating, k is constant; and, as a consequence of (ii), *k depends only on the relative velocity of O and O'*. It is in the this last innocuous looking statement that we depart from classical ideas.

3.2 Time dilation

Because k depends only on the relative motion, we have:

1) A photon sent by O towards O' at time t (O clock) arrives at O' at time $t' = kt$ (O' clock); and

2) A photon sent by O' towards O at time t' (O' clock) arrives at O at time $t = kt'$ (O clock).

Consider the space-time diagram figure 6. Here a photon sent by O at time t (O clock) arrives at O' at event B (time kt on O' clock); it is then sent back to O, arriving at time k^2t (O clock).

Hence O measures the distance to B and the time of B to be

$$d_B = \tfrac{1}{2}c(k^2-1)t, \qquad t_B = \tfrac{1}{2}(1+k^2)t\,.$$

Thus O reckons that the speed of O' is

$$u = \frac{d_B}{t_B} = \frac{c(k^2-1)}{k^2+1}\,.$$

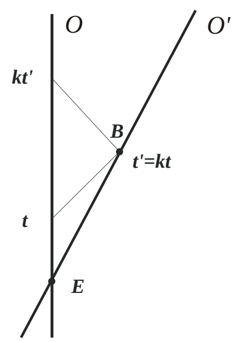

Figure 6: The k-factor

Solving for k, we have

$$k = \sqrt{\frac{c+u}{c-u}} > 1, .$$

It follows that

$$\frac{\text{Time } E \text{ to } B \text{ measured by } O}{\text{Time } E \text{ to } B \text{ measured by } O'} = \frac{t_B}{kt} = \frac{(k^2+1)t}{2kt} = \frac{1}{\sqrt{1-u^2/c^2}}\,.$$

Consider, for example, an astronaut travelling directly away from earth with speed $u = c\sqrt{3}/2$, so that $1/\sqrt{1-u^2/c^2} = 2$. So if the astronaut reckons that one hour passes between two events in his spaceship, then an observer on earth reckons that two hours pass. This is the *time dilation effect*: the time between the two events depends on the observer.

This result only becomes paradoxical if one insists on talking about 'time' independently of the process of measurement of time. It is also important to realise that the apparent asymmetry between the observer on earth and the astronaut arises because both are measuring the time interval between events on the astronaut's worldline. If we consider two events on earth (say at $t = 0$ and at $t = 1$), then the situation is reversed: O assigns a unit time separation; and the astronaut O' assigns a time separation $\gamma > 1$.

3.3 The two-dimensional Lorentz transformation

We shall now consider how the coordinate systems set up by two observers are related. We shall consider first the case of observers moving uniformly on the same straight line, so

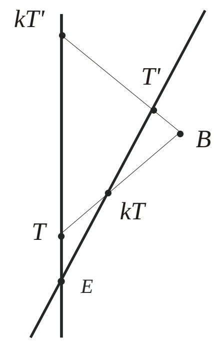

Figure 7: The derivation of the coordinate transformation

we have a two-dimensional space-time (events are points on the line at particular times). By using the radar method, each observer can assign inertial coordinates x, t to an event: x is the distance from the observer (with an appropriate sign) and t is the time of the simultaneous event on the observer's own worldline.

How are the inertial coordinate systems x, t and x', t' of two observers O and O' related. For simplicity, we shall assume that both set their clocks to zero at the event E at which they pass. Then E will be the common origin of the two coordinate systems.

Proposition The inertial coordinate systems set up by O and O' are related by

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \gamma(u) \begin{pmatrix} 1 & u/c \\ u/c & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}, \tag{12}$$

where u is the relative velocity and $\gamma(u) = 1/\sqrt{1 - u^2/c^2}$.

Proof. Let k denote Bondi's factor. Consider the space-time diagram figure 7. A photon is sent out from O at time T (O clock), passes O' at time kT (O' clock), is reflected at the event B, passes O' again at time T' (O' clock) and returns to O at time kT' (O clock). The coordinates of B are:

1) In the ICS of observer O:

$$t = \tfrac{1}{2}(kT' + T), \qquad x = \tfrac{1}{2}c(kT' - T)\,.$$

2) In the ICS of observer O':

$$t' = \tfrac{1}{2}(T' + kT), \qquad x' = \tfrac{1}{2}c(T' - kT)\,.$$

Hence we have

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \frac{c}{2}\begin{pmatrix} 1 & k \\ -1 & k \end{pmatrix}\begin{pmatrix} T \\ T' \end{pmatrix}, \qquad \begin{pmatrix} ct' \\ x' \end{pmatrix} = \frac{c}{2}\begin{pmatrix} k & 1 \\ -k & 1 \end{pmatrix}\begin{pmatrix} T \\ T' \end{pmatrix},$$

and therefore

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \frac{1}{2k}\begin{pmatrix} 1 & k \\ -1 & k \end{pmatrix}\begin{pmatrix} 1 & -1 \\ k & k \end{pmatrix}\begin{pmatrix} ct' \\ x' \end{pmatrix} = \frac{1}{2}\begin{pmatrix} k+k^{-1} & k-k^{-1} \\ k-k^{-1} & k+k^{-1} \end{pmatrix}\begin{pmatrix} ct' \\ x' \end{pmatrix}.$$

But we showed above that $k = \sqrt{(c+u)/(c-u)}$. Therefore

$$k + k^{-1} = \frac{2c}{\sqrt{c^2-u^2}}, \qquad k - k^{-1} = \frac{2u}{\sqrt{c^2-u^2}}\,,$$

and the result follows. (Note: some sign choices have been made in relating x and x', which can be positive or negative, to the distances from O or O', which are necessarily positive.) □

3.4 Properties of the two-dimensional Lorentz transformation

The relationship between the two coordinate systems is shown in figure 8. If we put $x' = 0$, then $x = ut$: so, with the sign choices we have made, O' is moving relative to O in the positive x direction with speed u. We also have $t = \gamma(u)t'$ when $x' = 0$, again the time dilation formula for events on the worldline of O'.

The inverse transformation is

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \gamma(u)\begin{pmatrix} 1 & -u/c \\ -u/c & 1 \end{pmatrix}\begin{pmatrix} ct \\ x \end{pmatrix},$$

so O is moving relative to O' in the *negative* x' direction with the same speed.

The transformation reduces to the Galilean transformation when $u \ll c$, since we have

$$\begin{pmatrix} t \\ x \end{pmatrix} = \gamma(u)\begin{pmatrix} 1 & u/c^2 \\ u & 1 \end{pmatrix}\begin{pmatrix} t' \\ x' \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}\begin{pmatrix} t' \\ x' \end{pmatrix}$$

as $c \to \infty$.

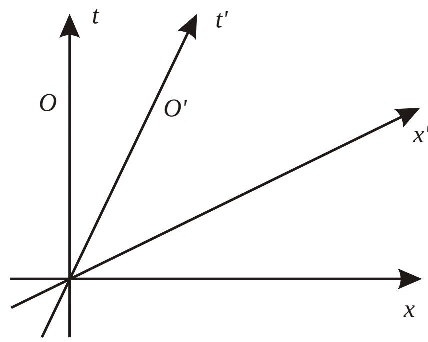

Figure 8: The two-dimensional Lorentz transformation

3.5 Transformation of velocity

Consider a non-accelerating particle moving with speed v relative to O in the negative x direction, so that $x = -vt + a$ for some constant a. Then we have

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -u/c \\ -u/c & 1 \end{pmatrix} \begin{pmatrix} ct \\ -vt + a \end{pmatrix},$$

($\gamma = \gamma(u)$ so that

$$t' = \gamma t - \frac{\gamma u}{c^2}(-vt + a) \qquad x' = -\gamma u t + \gamma(-vt + a).$$

Therefore the speed w of the particle relative to O' is

$$w = -\frac{\mathrm{d}x'}{\mathrm{d}t'} = \frac{\gamma(v + u)}{\gamma(1 + uv/c^2)} = \frac{v + u}{1 + uv/c^2}.$$

We note that this is not the same as the classical formula $w = v + u$, but reduces to it when $u, v \ll c$. We also note that if $v = c$, then $w = c$ irrespective of the value of u: photons move at speed c relative to all observers. The velocity of the observer is not added to c in a moving frame.

We also note that $|u| < c$ and $|v| < c$ implies that $|w| < c$ since

$$(c - u)(c - v) > 0 \quad \Leftrightarrow \quad u + v < c(1 + uv/c^2) \tag{13}$$

$$(c + u)(c + v) > 0 \quad \Leftrightarrow \quad u + v > -c(1 + uv/c^2). \tag{14}$$

Exercise

Show that

$$\gamma(u)\gamma(v)\begin{pmatrix} 1 & u/c \\ u/c & 1 \end{pmatrix}\begin{pmatrix} 1 & v/c \\ v/c & 1 \end{pmatrix}\begin{pmatrix} ct \\ x \end{pmatrix} = \gamma(w)\begin{pmatrix} 1 & w/c \\ w/c & 1 \end{pmatrix}$$

where $w = (u+v)/(1+uv/c^2)$. [Hint: $w^2 = c^2(\gamma(w)^2 - 1)/\gamma(w)^2$.]

Lecture 4. Lorentz transformations

4.1 Rapidity

The Lorentz transformation and velocity addition formula take on a more familiar look if we put $\phi(u) = \log k = \tanh^{-1}(u/c)$. Then

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \cosh\phi & \sinh\phi \\ \sinh\phi & \cosh\phi \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix},$$

so a Lorentz transformation is a *hyperbolic rotation*. The quantity ϕ is called the *rapidity* or *pseudo-velocity* of the transformation. It is analogous to the angle of a rotation in the plane.

In terms of rapidity, the velocity addition formula takes the more suggestive form

$$\phi(w) = \phi(u) + \phi(v)\,.$$

4.2 The Lorentz contraction

Consider two observers O and O' whose inertial coordinate systems are related by (12). Suppose that a rod lies along the x'-axis between $x' = 0$ and $x' = L$ and is at rest relative to O'. Then according to O', its length is L. What is its length as measured by O?

We must first be clear about what the question means. In the ICS of O', the worldlines of the ends of the rod are given by $x' = 0$ and by $x' = L$. In the ICS of O, therefore, they are given parametrically (with t' as parameter) by

$$(1) \qquad \begin{pmatrix} ct \\ x \end{pmatrix} = \gamma \begin{pmatrix} 1 & u/c \\ u/c & 1 \end{pmatrix} \begin{pmatrix} ct' \\ 0 \end{pmatrix} = \gamma \begin{pmatrix} ct' \\ ut' \end{pmatrix} \tag{15}$$

$$(2) \qquad \begin{pmatrix} ct \\ x \end{pmatrix} = \gamma \begin{pmatrix} 1 & u/c \\ u/c & 1 \end{pmatrix} \begin{pmatrix} ct' \\ L \end{pmatrix} = \gamma \begin{pmatrix} ct' + Lu/c \\ ut' + L \end{pmatrix}. \tag{16}$$

The question is: what is the distance measured by O between two events E and B, one on each worldline, which are simultaneous according to O? If we take E to be the event $t = 0$, $x = 0$, then B must be as in (16), with t' chosen so that $t = 0$. That is $t' = -Lu/c^2$, which implies that B is the event

$$t = 0, \qquad x = \gamma(-Lu^2/c^2 + L) = L(1 - u^2/c^2)^{1/2}\,. \tag{17}$$

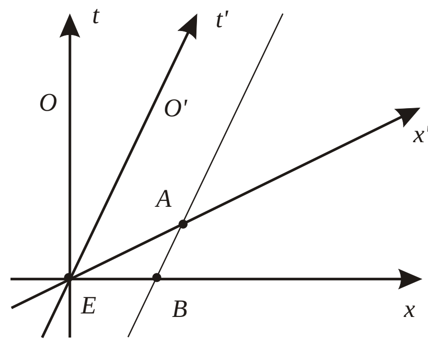

Figure 9: The Lorentz contraction

So according to O, the rod is shorter by a factor $\sqrt{1-u^2/c^2}$. This is the same as the factor proposed by Lorentz in the context of the ether hypothesis (hence the term *Lorentz contraction*, or sometimes *FitzGerald-Lorentz contraction* since the Irish mathematician George FitzGerald put forward the same idea a few years before Lorentz). However, the interpretation here is very different. In particular, it is important to remember that when O measures the length, he is assigning a 'distance' to the spatial separation between E and B; but when O' measures the length, he is assigning a 'distance' to the separation between E and A, the event at the other end of the rod simultaneous with E according to his own definition of simultaneity (see figure 9).

4.3 Minkowski space

By using the radar method and by observing the directions from light signals arrive, an observer can set up *inertial coordinates* t, x, y, z in four-dimensional space-time, taking his own location as the origin of the space coordinates ($x = y = z = 0$). We shall not spell out exactly how this is done. Instead, we shall derive the relationship between the inertial coordinate systems t, x, y, z and t', x', y', z' of two non-accelerating ('inertial') observers O and O' in relative motion by making certain assumptions about the transformation, derived from the principle of relativity and the two-dimensional derivation.

A1 The transformation is affine linear. That is, it is of the form

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = L \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} + C\,, \tag{18}$$

where L is a 4×4 matrix and C is a column vector.

A2 Photons travel in straight lines with velocity c relative to any inertial coordinate system.

A3 Nothing travels faster than light.

A4 The principle of relativity applies to all physical phenomena (i.e. only the relative motion of non-accelerating observers can be detected by physical experiments).

(Al) is one property of Galilean transformations that we do not want to throw away. It is equivalent to the assertion that Newton's first law should hold in any inertial coordinate system (i.e. that worldlines of free particles—particles not acted on by any force—are straight lines in space-time, given by linear equations in any inertial coordinate system). (A2) is essentially the same as ($*$); and (A3) is needed for consistency—as we shall see. The restriction of (A4) to non-accelerating observers is the origin of the 'special' in 'special relativity'.

Time dilation factor

We denote the top left entry in L by γ: this is the *time dilation factor* for the motion of O' relative to O: along the worldline of O', which is given by $x' = y' = z' = 0$, we have

$$t = \gamma t' + \text{constant}.$$

So γ relates the time measurements of events on the worldline of O' in the two coordinate systems. Similarly, if γ' is the top left entry in L^{-1}, then along the worldline of O ($x = y = z = 0$) we have

$$t = \gamma' t + \text{constant}.$$

So γ' is the time dilation factor for the motion of O relative to O'. It follows from the relativity assumption (A4) that $\gamma = \gamma'$.

The light cone

It follows from (A2) at the worldlines of photons through an event A form a cone in space-time, called the *light-cone* of the event. If A is the event $x = y = z = t = 0$, then the light-cone of A has equation

$$c^2t^2 - x^2 - y^2 - z^2 = 0 \tag{19}$$

(any event whose coordinates satisfy this is at distance $|ct|$ from $x = y = z = 0$ at time t in the inertial coordinate system of O).

The $t =$ constant > 0 sections of the light-cone are spherical 'wave-fronts' $x^2+y^2+z^2 = c^2t^2$ spreading out from the origin with speed c. By (A3), all particle worldlines through A must lie inside the light-cone.[2] In Galilean relativity, all observers agree on which events are simultaneous with A In Einstein's relativity, they do not, but agree on which events lie on the light-cone of A. As $c \to 0$, the light-cone degenerates into a hyperplane of constant t through A, and so we recover the Galilean picture in the limit.

The structure of space-time in Einstein's theory was first made clear by H. Minkowski. The space-time of special relativity is called *Minkowski space*.

4.4 The Lorentz transformation 2

Consider two events A and B with coordinates t_A, x_A, y_A, z_A and t_B, x_B, y_B, z_B in the coordinate system of O; and t'_A, x'_A, y'_A, z'_A and t'_B, x'_B, y'_B, z'_B in the coordinate system of O'. Put

$$T = c(t_B - t_A), \quad X = x_B - x_A, \quad Y = y_B - y_A, \quad Z = z_B - z_A$$

and

$$T' = c(t'_B - t'_A), \quad X = x'_B - x'_A, \quad Y = y'_B - y'_A, \quad Z = z'_B - z'_A\,.$$

Then

$$\begin{pmatrix} T \\ X \\ Y \\ Z \end{pmatrix} = L \begin{pmatrix} T' \\ X' \\ Y' \\ Z' \end{pmatrix}\,.$$

[2] We shall always draw space-time diagrams so that the generators of the light-cone are at 45^o to the vertical.

Now A and B lie on the worldline of a photon if and only if

$$T^2 - X^2 - Y^2 - Z^2 = 0$$

(the condition that the distance-squared $X^2 + Y^2 + Z^2$ should be c^2 times the square of the time separation $t_B - t_A$). In the second coordinate system, the condition is

$$T'^2 - X'^2 - Y'^2 - Z'^2 = 0\,.$$

These must be equivalent. So we conclude that

$$T^2 - X^2 - Y^2 - Z^2 = 0 \quad \text{if and only if} \quad T'^2 - X'^2 - Y'^2 - Z'^2 = 0\,. \tag{20}$$

If we introduce the diagonal matrix g with diagonal entries $1, -1, -1, -1$, then

$$T^2 - X^2 - Y^2 - z^2 = \begin{pmatrix} T \; X \; Y \; Z \end{pmatrix} g \begin{pmatrix} T \\ X \\ Y \\ Z \end{pmatrix} = 0\,.$$

From (20) and from the fact that L and L^{-1} have the same time dilation factor, we shall prove the following *pseudo-orthogonality* property of L.

Proposition 1 *$g = L^{\mathrm{t}} g L$.*

Lecture 5. Lorentz Transformations (cont)

From (20) and from the fact that L and L^{-1} have the same time dilation factor, we shall prove the following *pseudo-orthogonality* property of L.

PROPOSITION 2 *$g = L^{\mathrm{t}} g L$.*

LEMMA 1 *Let G be a symmetric nonzero 4×4 matrix. Suppose that*

$$\begin{pmatrix} T\ X\ Y\ Z \end{pmatrix} G \begin{pmatrix} T \\ X \\ Y \\ Z \end{pmatrix} = 0 \,,$$

whenever $T^2 - X^2 - Y^2 - Z^2 = 0$. Then $G = \alpha g$ for some nonzero $\alpha \in \mathbb{R}$

Proof. Write

$$G = \left(\begin{array}{c|ccc} \alpha & p & q & r \\ \hline p & & & \\ q & & S & \\ r & & & \end{array} \right) ,$$

where S is a 3×3 symmetric matrix. Put $T = 1$. Then we have that

$$\alpha + 2(pX + qY + rZ) + (X\ Y\ Z) S \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = 0 \quad \text{whenever} \quad X^2 + Y^2 + Z^2 = 1 \,.$$

But the second equation is the equation of the unit sphere with centre at the origin; the first must therefore be as well. It follows that $p = q = r = 0$ and that $-S/\alpha$ is the identity. The lemma follows.

We can now prove the the propostion.

Proof. We have

$$T^2 - X^2 - Y^2 - Z^2 = \begin{pmatrix} T\ X\ Y\ Z \end{pmatrix} g \begin{pmatrix} T \\ X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} T'\ X'\ Y'\ Z' \end{pmatrix} L^{\mathrm{t}} g L \begin{pmatrix} T' \\ X' \\ Y' \\ Z' \end{pmatrix} .$$

It follows from the lemma and from (20) that $L^{\mathrm{t}}gL = \alpha g$ for some nonzero $\alpha \in \mathbb{R}$, and hence that $L^{-1} = \alpha^{-1}gL^{\mathrm{t}}g$ since $g^{-1} = g$. Therefore the top left entry in L^{-1} is γ/α, where γ is the top left entry in L. But we deduced from our relativity assumption (A4) that the top left entries in L and L^{-1} are equal. Hence $\alpha = 1$.

It follows from the proposition that $L^{-1} = gL^{\mathrm{t}}g$ and hence that $LgL^{\mathrm{t}} = g$ since $g^2 = 1$.

5.1 The standard Lorentz transformation

Suppose that O is moving along the x'-axis in the coordinates of O' and that O' is moving along the x-axis in the coordinates of O; and suppose further that they both take the origin of their coordinate systems to be the event at which they pass each other. Then $C = 0$ and the t, x and t', x' coordinates are related by (12). Hence

$$L = \begin{pmatrix} \gamma & \gamma u/c & p & q \\ \gamma u/c & \gamma & r & s \\ P & Q & a & b \\ R & S & c & d \end{pmatrix}.$$

From $L^{\mathrm{t}}gL = g$,we obtain

$$\begin{aligned} \gamma^2 - \gamma^2u^2/c^2 - P^2 - R^2 &= 1 \\ \gamma^2u^2/c^2 - \gamma^2 - Q^2 - S^2 &= -1\,. \end{aligned}$$

But $\gamma^2 - \gamma^2u^2/c^2 = 1$. Hence $P = Q = R = S = 0$. Similarly, from $L^{\mathrm{t}}gL = g$, we get $p = q = r = s = 0$, and then that

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is an orthogonal matrix. By making an orthogonal transformation in y, z-plane about the by A^{-1}, we can arrange without loss of generality that $A = 1$. We then have

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & \gamma u/c & 0 & 0 \\ \gamma u/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix}, \tag{21}$$

where $\gamma = \gamma(u) = 1/\sqrt{1 - u^2/c^2}$. This is the *standard Lorentz transformation*.

5.2 The general Lorentz transformation

In deriving the standard Lorentz transformation, we made assumptions about the relative orientations of the spatial axes of the two coordinate systems. If we drop these (but still assume that O' is moving directly away from O), then we must combine (21) with (1) a rotation of the x, y, z coordinates and (2) a rotation of the x', y', z' coordinates. The result is

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = L \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} \tag{22}$$

with

$$L = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} L(u) \begin{pmatrix} 1 & 0 \\ 0 & Q^{\mathrm{t}} \end{pmatrix} \tag{23}$$

where P and Q are 3×3 proper orthogonal matrices and $L(u)$ is the standard Lorentz transformation (for some u). A transformation of the form (22) is called a *proper orthochronous Lorentz transformation*

It is conventional to label the entries in L as

$$L = \begin{pmatrix} L^0{}_0 & L^0{}_1 & L^0{}_2 & L^0{}_3 \\ L^1{}_0 & L^1{}_1 & L^1{}_2 & L^1{}_3 \\ L^2{}_0 & L^2{}_1 & L^2{}_2 & L^2{}_3 \\ L^3{}_0 & L^3{}_1 & L^3{}_2 & L^3{}_3 \end{pmatrix} \tag{24}$$

using upper and lower indices for reasons that derive from tensor analysis.

The coordinate systems of our two observer O and O' are related by (22) The worldline of the second observer is given by $x' = y' = z' = 0$, if we assume he takes his own location as the origin of his coordinate system. In the coordinate system of O, the worldline is given by

$$t = L^0{}_0 t', \quad x = cL^1{}_0 t', \quad y = cL^2{}_0 t', \quad z = cL^3{}_0 t'\,.$$

So if we treat t' as a parameter along the worldline,

$$(c\dot{t}, \dot{x}, \dot{y}, \dot{z}) = c(L^0{}_0, L^1{}_0, L^2{}_0, L^3{}_0)\,,$$

where the dot is the derivative with respect to t'. On eliminating t', we get

$$\left(\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t}\right) = \frac{(L^1{}_0, L^2{}_0, L^3{}_0)}{L^0{}_0}\,.$$

So the entries in the first column of L give the velocity of O' relative to O. Similary the entries in the first row give minus the velocity of O relative to O'.

Proposition 3 *If L is a proper orthochronous Lorentz transformation, then*

(1) $L^{-1} = gL^{\mathrm{t}}g$

(2) $\det(L) = 1$

(3) $L^0{}_0 > 0$.

Proof. (3) follows from the fact $L^0{}_0 = \gamma(u)$. Note that $g^2 = I$ and that (1) and (2) hold for the standard Lorentz transformation and for any L of the special form

$$L = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}; \qquad P \in \mathrm{SO}(3)\,. \tag{25}$$

Hence they hold for any proper orthochronous transformation.

The positivity of $L^0{}_0$ ensures that t is an increasing function of t' for fixed x', y', z' (hence 'orthochronous'); and the positivity in addition of $\det(L)$ ensures that the 'handedness' of the two sets of spatial axes are the same (hence 'proper'). The converse proposition is also true: that is, if L satisfies (1), (2), and (3), then there exists $P, Q \in \mathrm{SO}(3)$ such that (23) holds, with $L(u)$ the standard Lorentz transformation for some v ($|v| < c$). The proof is left as an exercise.

Finally, if we drop the condition that O' should be moving directly away from O, then we can combine (22) with a spatial translation and a change in the origin of the time coordinate. The result is

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = L \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} + T \tag{26}$$

where L is a proper orthochronous Lorentz transformation matrix and T is a constant column vector. Eqn (26) is an *inhomogeneous Lorentz transformation*.

Examples

Rotations. If
$$L = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix},$$
where P is a 3×3 matrix, then
$$gL^{\mathrm{t}}g = \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P^{\mathrm{t}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & P^{\mathrm{t}} \end{pmatrix},$$
where I is the 3×3 identity matrix. So if $L^{-1} = gL^{\mathrm{t}}g$, then $P^{-1} = P^{\mathrm{t}}$, so P must be orthogonal. Since $\det L = \det P$, we must also have that P is a rotation.

The standard Lorentz transformation. In the case of the standard Lorentz transormation,
$$\begin{aligned} gL^{\mathrm{t}}g &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma & \gamma u/c & 0 & 0 \\ \gamma u/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma & -\gamma u/c & 0 & 0 \\ -\gamma u/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$
A direct calculation shows that this is the inverse matricx of L.

5.3 Euclidean space and Minkowski space

The four-dimensional space-time of special relativity has a geometry that is analogous in many ways to that of three-dimensional Euclidean space. In Euclidean space, for example, two systems of Cartesian coordinates are related by the transformation
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = H \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + C \tag{27}$$
where $H^{-1} = H^{\mathrm{t}}$ and C is a constant column vector. In Minkowski space, the analogous transformation is the inhomogeneous transformation relating two inertial coordinate

systems

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = L \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} + T \, . \tag{28}$$

The orthogonality property of H ($H^{-1} = H^{\mathrm{t}}$) is replaced by the corresponding *pseudo-orthogonality* property of L

$$L^{-1} = (gLg)^{\mathrm{t}} \, . \tag{29}$$

That is, the inverse of L is obtained by changing the signs of the entries $L^0{}_1$, $L^0{}_2$, $L^0{}_3$, $L^1{}_0$, $L^2{}_0$, and $L^3{}_0$, and then transposing the resulting matrix.

Similarly, in Euclidean geometry, the squared distance

$$(x_{(1)} - x_{(2)})^2 + (y_{(1)} - y_{(2)})^2 + (z_{(1)} - z_{(2)})^2 \tag{30}$$

between $(x_{(1)}, y_{(1)}, z_{(1)})$ and $(x_{(2)}, y_{(2)}, z_{(2)})$ is invariant (it is the same in any Cartesian coordinate system); in Minkowski space, the corresponding quantity is the invariant

$$c^2(t_{(1)} - t_{(2)})^2 - (x_{(1)} - x_{(2)})^2 - (y_{(1)} - y_{(2)})^2 - (z_{(1)} - z_{(2)})^2 \tag{31}$$

associated with two space-time events.

5.4 Four-vectors

The analogies are not complete (there is nothing in Euclidean geometry corresponding to the distinction between orthochronous ($L^0{}_0 > 0$) and non-orthochronous ($L^0{}_0 < 0$) Lorentz transformations); and they are sometimes misleading. But they do suggest that it would be useful to consider space-time vectors, defined by adapting the transformation law for the components of a vector in space.

In three-dimensional Euclidean space, a vector has $\boldsymbol{X}$ has three components X_1, X_2, X_3 that transform under (27) by

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = H \begin{pmatrix} X'_1 \\ X'_2 \\ X'_3 \end{pmatrix} ,$$

where X'_1, X'_2, X'_3 are the components in the x', y', z' coordinate system. Although it is not usual to do so, one cane use this as the *definition* of a vector: it is an object that associates a set of componets with each Cartesian coordinate system, subject to this transfromation rule.

To avoid confusion, vectors in space-time will be called four-vectors and ordinary vectors in Euclidean space will be called three-vectors (although both will be shortened to 'vector' when there is no danger of confusion).

DEFINITION 4 *A* four-vector *is an object X that associates an element (X^0, X^1, X^2, X^3) of $\mathbb{R}^4$ to each inertial coordinate system. The X^as ($a = 0, 1, 2, 3$) are called the* components *of X. They have the following property. If two inertial coordinate systems are related by (28) then the components X^a in the first (unprimed) system are related to components X'^a in the second (primed) by*

$$\begin{pmatrix} X^0 \\ X^1 \\ X^2 \\ X^3 \end{pmatrix} = L \begin{pmatrix} X'^0 \\ X'^1 \\ X'^2 \\ X'^3 \end{pmatrix} \tag{32}$$

This rather awkward definition says no more than that a four-vector is an object with four components X^0, X^1, X^2, X^3 and that the components transform under an inhomogeneous Lorentz transformation of the coordinates by the corresponding homogeneous transformation (i.e. the same transformation as the coordinates, but without the constant column vector). The important point to remember is which way round the transformation goes (look carefully at the relationship between (28) and (32)).

As with three-vectors, one can add four-vectors and take scalar multiples. Thus

$X + Y$ has components $(X^0 + Y^0, X^1 + Y^1, X^2 + Y^2, X^3 + Y^3)$

λX ($\lambda \in \mathbb{R}$) has components $(\lambda X^0, \lambda X^1, \lambda X^2, \lambda X^3)$.

The set of all four-vectors is a four-dimensional vector space.

Lecture 6. Four-vectors

6.1 Temporal and spatial parts

If two inertial observers O and O' are at rest relative to each other, then their time axes in space-time will point in the same direction and their inertial coordinate systems will be related by

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & H \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} + T\,, \tag{33}$$

where H is a 3×3 proper orthogonal matrix. The components of a four-vector X in the two systems are related by

$$X^0 = X'^0, \qquad \begin{pmatrix} X^1 \\ X^2 \\ X^3 \end{pmatrix} = H \begin{pmatrix} X'^1 \\ X'^2 \\ X'^3 \end{pmatrix} . \tag{34}$$

So under this restricted class of transformations (which consists of rotations and translations of the spatial axes together with changes in the origin of the time coordinate) the three spatial components X^l, X^2, X^3 behave as the components of a three-vector $\boldsymbol{X}$ and the time component X^0 is invariant. The decomposition of X into a *temporal part* X^0 and a *spatial part* $\boldsymbol{X}$ depends only on the direction of the time axis and not on the particular choice of origin and orientation of the spatial coordinate axes. (Under a general transformation between the ICSs of two observers in relative motion, the direction of the time axis changes and the temporal and spatial parts are mixed up.)

We shall write $X = (X^0, X^1, X^2, X^3)$ as shorthand for 'X has components X^0, X^1, X^2, X^3' in a particular inertial coordinate system and $X = (X^0, \boldsymbol{X})$ for 'X has temporal part X^0 and spatial part $\boldsymbol{X}$' relative to a particular direction for the time axis.

6.2 The inner product

Proposition 4 *(Analogue of the dot product). If X and Y are four-vectors, then*

$$\begin{aligned} g(X,Y) : &= \begin{pmatrix} X^0 \, X^1 \, X^2 \, X^3 \end{pmatrix} g \begin{pmatrix} Y^0 \\ Y^1 \\ Y^2 \\ Y^3 \end{pmatrix} \\ &= X^0 Y^0 - X^1 Y^1 - X^2 Y^2 - X^3 Y^3 \\ &= X^0 Y^0 - \boldsymbol{X} \, . \, \boldsymbol{Y} \end{aligned} \tag{35}$$

is an invariant.

Proof. If L is the matrix of a Lorentz transformation, then $L^{\mathrm{t}} g L = g$. Hence if

$$\begin{pmatrix} X^0 \\ X^1 \\ X^2 \\ X^3 \end{pmatrix} = L \begin{pmatrix} X'^0 \\ X'^1 \\ X'^2 \\ X'^3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} Y^0 \\ Y^1 \\ Y^2 \\ Y^3 \end{pmatrix} = L \begin{pmatrix} Y'^0 \\ Y'^1 \\ Y'^2 \\ Y'^3 \end{pmatrix} \tag{36}$$

then

$$\begin{aligned} \begin{pmatrix} X^0 \, X^1 \, X^2 \, X^3 \end{pmatrix} g \begin{pmatrix} Y^0 \\ Y^1 \\ Y^2 \\ Y^3 \end{pmatrix} &= \begin{pmatrix} X'^0 \, X'^1 \, X'^2 \, X'^3 \end{pmatrix} L^{\mathrm{t}} g L \begin{pmatrix} Y'^0 \\ Y'^1 \\ Y'^2 \\ Y'^3 \end{pmatrix} \\ &= \begin{pmatrix} X'^0 \, X'^1 \, X'^2 \, X'^3 \end{pmatrix} g \begin{pmatrix} Y'^0 \\ Y'^1 \\ Y'^2 \\ Y'^3 \end{pmatrix} . \end{aligned} \tag{37}$$

Note that $g(\, . \, , \, . \,)$ is a symmetric bilinear form on the vector space of four-vectors. It is called the *inner product* and it has the same properties as the inner (or dot) product of vectors in Euclidean space, except that it is not positive definite.

Definition 5 *Two four-vectors X and Y are* orthogonal *if $g(X,Y) = 0$.*

6.3 Index notation

In some contexts, it is useful to put $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$, and so on, and write (26) as

$$x^a = \sum_{b=0}^{3} L^a{}_b x'^b + T^a \qquad (a = 0, 1, 2, 3) \tag{38}$$

The positioning of the indices takes a little getting used to; it is important in tensor analysis.

The transformation rules can be written more compactly by using the *Einstein conventions*.

Summation convention

When an index in a term is repeated, once as an upper index and once as a lower index, a sum over $0, 1, 2, 3$ is implied.

Range convention

An index which is not repeated is a *free index*. Any equation is understood to hold for all values of the free indices over the range $0, 1, 2, 3$.

With these conventions, we have

$$X^a = L^a{}_b X'^b \quad \text{when} \quad x^a = L^a{}_b x'^b + T^a \ ,, \tag{39}$$

where in both equations the repetition of the index b implies a sum eover $b = 0, 1, 2, 3$ (see eqn 38). The free index in both is a, and the equations are understood to hold for $a = 0, 1, 2, 3$, by the range convention. Similarly, the invariant $g(x, Y)$ can be written

$$g(X, Y) = g_{ab} X^a Y^b \tag{40}$$

where the g_{ab}s are the entries in the matrix g (i.e. $g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$, and $g_{ab} = 0$ when $a \neq b$). Here there are two summations over the repeated indices $a = 0, 1, 2, 3$ and $b = 0, 1, 2, 3$.

A further notational device in common use is to put $X_a = g_{ab}X^b$ (again with a summation). Then $X_0 = X^0$, $X^1 = -X_1$, $X_2 = -X^2$, $X_3 = -X^3$ and

$$g(X,Y) = X_a Y^a = X_0 Y^0 + X_1 Y^1 + X_2 Y^2 + X_3 Y^3 \,. \tag{41}$$

The operation of forming the X_as from the components X^a of X is called 'lowering the index'. The conventions for the positioning of indices are such that summations are always over one lower index and one upper index.

6.4 Classification of four-vectors

Definition 6 *A four-vector X is said to be* timelike *(TL),* spacelike *(SL) , or* null *(N) as $g(X,X) > 0$, $g(X,X) < 0$, or $g(X,X) = 0$.*

Examples

The four-vectors with components

$$(1,0,0,0), \quad (0,1,0,0), \quad \text{and} \quad (1,1,0,0)$$

(in some ICS) are respectively timelike, spacelike, and null. Note that a null four-vector need not be the zero four-vector.

A four-vector whose spatial part vanishes in some ICS must be timelike; and a four-vector whose temporal part vanishes in some ICS must be spacelike. The converses of these statements are the following two propositions.

Proposition 5 *If X is timelike, then there exists an ICS in which $X^1 = X^2 = X^3 = 0$.*

Proof. Consider the components X'^0, X'^1, X'^2, X'^3 of X in an ICS x'^a. By rotating the spatial axes to make the x'-axis parallel to the spatial part of X, we can ensure that $X'^2 = X'^3 = 0$.

We can then make the third spatial component vanish by making a standard Lorentz transformation chosen so that

$$\begin{pmatrix} \gamma(u) & \gamma(u)u/c \\ \gamma(u)u/c & \gamma(u) \end{pmatrix} \begin{pmatrix} X'^0 \\ X'^1 \end{pmatrix} = \begin{pmatrix} X^0 \\ 0 \end{pmatrix} . \tag{42}$$

That is, we choose u such that $|u| < c$ and

$$uX'^0/c + X'^1 = 0\,. \tag{43}$$

This is possible because X is timelike, and so $|X'^1/X'^0| < 1$.

By a similar argument, we also have the following proposition.

Proposition 6 *If X is spacelike, then there exists and ICS in which $X^0 = 0$.*

In the case of timelike and null vectors (but *not* spacelike vectors), the sign of the time component X^0 is invariant.

Proposition 7 *Suppose that X is timelike or null. If $X^0 > 0$ in some ICS, then $X^0 > 0$ in every ICS.*

Proof. Since rotations do not alter X^0, it is sufficient to consider what happens to X^0 under a standard Lorentz transformation. So suppose

$$\begin{pmatrix} X^0 \\ X^1 \\ X^2 \\ X^3 \end{pmatrix} = \begin{pmatrix} \gamma & \gamma u/c & 0 & 0 \\ \gamma u/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X'^0 \\ X'^1 \\ X'^2 \\ X'^3 \end{pmatrix}; \quad \gamma = \gamma(u)\,. \tag{44}$$

Then

$$X^0 = \gamma(X'^0 + uX'^1/c) > 0 \tag{45}$$

since $|u/c| < 1$ and $|X'^1| \le |X'^0|$.

Definition 7 *A timelike or null vector X is said to be* future-pointing *(FP) if $X^0 > 0$ in some (and hence every) ICS, and* past-pointing *(PP), if $X^0 < 0$.*

We cannot make a similar definition for spacelike vectors since the sign of the time component of a spacelike vector is different in different ICSs.

The space of four-vectors is illustrated in figure 10, where the x^0 axis is vertical and one spatial dimension is suppressed. The null vectors lie on the cone

$$(X^0)^2 - (X^1)^2 - (X^2)^2 - (X^3)^2 = 0 \tag{46}$$

which has its vertex at the origin.

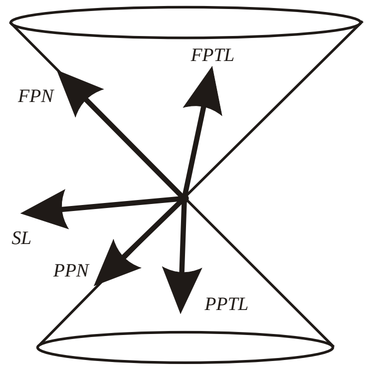

Figure 10: The space of 4-vectors

6.5 Displacement vectors

Suppose that A and B are two events with coordinates x^a and y^a in some ICS. Put $X^a = y^a - x^a$. Then the X^as transform as the components of a four-vector X, called the *displacement vector* from A to B (it is analogous to the vector from one point to another in Euclidean space).

The temporal part of X in an ICS is the time from A to B multiplied by c; the spatial part is the three-vector from the point where A happens to the point where B happens. There are various possibilities.

(1) X is spacelike. It is impossible to get from A to B without travelling faster than light, so B lies outside the light cone of A (and vice versa). By the results of §6.4, there exists an inertial coordinate system in which $X^0 = c(t_{(2)} - t_{(1)}) = 0$ that is in which A and B are simultaneous. The invariant $g(X, X)$ is equal to $-D^2$, where D is the distance from A to B measured in this system.

There exist ICSs in which A happens before B and ICSs in which A happens after B (it is for this reason that the prohibition on faster-than-light travel is required for the consistency of the theory with commonsense ideas about causality).

(2) X is null. Then A and B lie on the worldline of a photon. If X is future-pointing (past-pointing), then B happens after (before) A_l in every ICS.

(3) X is timelike. There exists an ICS in which $X^1 = X^2 = X^3 = 0$; that is in which A_l and B happen at the same place. The invariant $g(X, X)$ is equal to $c^2\tau^2$ where $\tau = t_{(2)} - t_{(1)}$

is the time from A_l to B in this ICS. If X is future-pointing (past-pointing), then $\tau > 0$ ($\tau < 0$) and B happens after (before) A in every ICS. $\square$

Note that if A, B, C are events and that if X and Y are the displacement vectors from A to B and from B to C (respectively), then $X + Y$ is the displacement vector from A to C.

Lecture 7. Relativistic motion

7.1 Proper time and four-velocity

Consider the worldline of a nonaccelerating particle (or observer). This is a straight line L in space-time which lies inside the light cone of any event on the line.

There is a natural parameter τ along L called proper time.

Definition 8 *The* proper time *along the worldline of a particle in uniform motion is the time measured in a frame in which the particle is at rest.*

It is analogous to the distance parameter s along a line in Euclidean space (a correct, but rather eccentric, definition of s would be that s is equal to the z-coordinate along the line in a coordinate system in which the line is parallel to the z-axis).

If the particle is at rest in the ICS t', x', y', z' then its worldline is given by

$$t' = \tau, \quad x = \text{constant}, \quad y' = \text{constant}, \quad z' = \text{constant}\,. \tag{47}$$

Hence if t, x, y, z is a second ICS related to t', x', y', z' by the standard Lorentz transformation (with speed v), then along L

$$t = \gamma(v)\tau + \text{constant}\,. \tag{48}$$

It follows that in a general ICS t, x, y, z,

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma(v) \tag{49}$$

where v is the speed of the particle.

All observers agree on τ (to within an arbitrary additive constant). Put $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$ (as before) . Then L can be represented in parametric form by

$$x^a = x^a(\tau); \qquad a = 0, 1, 2, 3\,.$$

Put

$$V^a = \frac{\mathrm{d}x^a}{\mathrm{d}\tau}; \qquad a = 0, 1, 2, 3. \tag{50}$$

Proposition 8 *The V^as are the components of a four-vector V.*

Proof. Under a change of inertial coordinate system we have

$$x^a = L^a{}_b x'^b + T^a \ , .$$

Since the $L^a{}_b$s are constant,

$$V^a = \frac{\mathrm{d}x^a}{\mathrm{d}\tau} = L^a{}_b \frac{\mathrm{d}x'^b}{\mathrm{d}\tau} = L^a{}_b V'^b \,,$$

which is the four-vector transformation rule.

DEFINITION 9 *The four-vector V with components (V^0, V^1, V^2, V^3) is called the* four-velocity *of the particle.*

Suppose that the particle has velocity $\boldsymbol{v}$ relative to the ICS t, x, y, z. Then

$$\frac{\mathrm{d}x^0}{\mathrm{d}\tau} = c\frac{\mathrm{d}t}{\mathrm{d}\tau} = c\gamma(v) \tag{51}$$

($v = |\boldsymbol{v}|$) by eqn (48); and

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}x}{\mathrm{d}t} = \gamma(v)v_1, \quad \frac{\mathrm{d}y}{\mathrm{d}\tau} = \gamma(v)v_2, \quad \frac{\mathrm{d}z}{\mathrm{d}\tau} = \gamma(v)v_3\,, \tag{52}$$

where v_1, v_2, v_3 are the components of $\boldsymbol{v}$. Hence the components of the four-velocity V are

$$(V^0, V^1, V^2, V^3) = \gamma(v)(c, v_1, v_2, v_3)\,. \tag{53}$$

In other words V decomposes into temporal and spatial parts as

$$V = \gamma(v)(c, \boldsymbol{v}) \tag{54}$$

PROPOSITION 9 *$g(V, V) = c^2$.*

Proof. In an ICS in which the particle is at rest, $V^0 = c$, $V^1 - V^2 - V^3 = 0$. Hence in this system

$$g(V, V) = c^2 \tag{55}$$

But $g(V, V)$ is an invariant. Hence (55) is valid in any ICS.

An alternative proof is the following. In a general ICS

$$\begin{aligned} g(V,V) &= (V^0)^2 - (V^1)^2 - (V^2)^2 - (V^3)^2 \\ &= \gamma(v)^2(c^2 - \boldsymbol{v}\,.\,\boldsymbol{v}) \\ &= c^2 \end{aligned} \tag{56}$$

Example

Addition of velocities. Relative to some ICS, an observer has velocity $\boldsymbol{u}$ and a particle has velocity $\boldsymbol{v}$. Find the speed w of the particle relative to the observer in terms of $\boldsymbol{u}$ and $\boldsymbol{v}$.

Solution. Let U be the four-velocity of the observer and let V be the four-velocity of the particle. In the given ICS,

$$U = \gamma(u)(c, \boldsymbol{u}) \quad \text{and} \quad V = \gamma(v)(c, \boldsymbol{v})\,. \tag{57}$$

Hence

$$g(U,V) = \gamma(u)\gamma(v)(c^2 - \boldsymbol{u}\,.\,\boldsymbol{v})\,. \tag{58}$$

Now consider an inertial coordinate system in which the observer is at rest. In this system

$$U = (c, 0), \qquad V = \gamma(w)(c, \boldsymbol{w})\,, \tag{59}$$

where $\boldsymbol{w}\,.\,\boldsymbol{w} = w^2$. Hence

$$g(U,V) = \gamma(w)\,. \tag{60}$$

But $g(U,V)$ is invariant. Therefore

$$c^2\gamma(w) = \gamma(u)\gamma(v)(c^2 - \boldsymbol{u}\,.\,\boldsymbol{v})\,. \tag{61}$$

On solving for w from $\gamma(w) = (1 - w^2/c^2)^{-1/2}$, one finds that

$$w = \frac{c\sqrt{c^2(\boldsymbol{u}-\boldsymbol{v})\,.\,(\boldsymbol{u}-\boldsymbol{v}) - u^2v^2 + (\boldsymbol{u}\,.\,\boldsymbol{v})^2}}{c^2 - \boldsymbol{u}\,.\,\boldsymbol{v}} \tag{62}$$

which reduces to the classical formula $w = |\boldsymbol{u} - \boldsymbol{v}|$ when $u, v \ll c$.

Example

A non-accelerating observer O has four-velocity U. Let A and B be two events. Show that O reckons that A and B are simultaneous if and only if the displacement vector X from A to B is orthogonal to U (i.e. $g(U,X) = 0$).

Solution. Pick an ICS in which O is at rest. Then U and X have components

$$U = (c, 0, 0, 0) \qquad X = (X^0, X^1, X^2, X^3)\,, \tag{63}$$

where X^0 is the time separation of A and B, multiplied by c. Thus in this ICS.

$$g(U, X) = cX^0 \tag{64}$$

Now O reckons that A and B are simultaneous if and only if $X^0 = 0$; that is if and only if $g(U, X) = 0$.

Example

Two observers O and O' are travelling in straight lines at constant speeds. Show that there is a pair of events A and A', with A on the worldline of O and A' on the worldline of O', which O and O' agree are simultaneous.

Solution. Let the four-velocities of O and O' be V and V', respectively. Pick two events B (on the worldline of O) and B' (on the worldline of O). Then if A is any other event on the worldline of O, the displacement vector from B to A is τV for some $\tau \in \mathbb{R}$; and if A' is any other event on the worldline of O', the displacement vector from B' to A' is $\tau' V'$ for some $\tau' \in R$. The displacement vector from A to A' is therefore

$$Y = X - \tau V + \tau' V' \tag{65}$$

where X is the displacement vector from B to B'.

By the previous example, it is sufficient to find A and A' such that

$$\begin{aligned} 0 = g(Y, V) &= g(X, V) - c^2\tau + g(V', V)\tau' \\ 0 = g(Y, V') &= g(X, V') - g(V, V')\tau + c^2\tau' \end{aligned} \tag{66}$$

From above, $g(V, V') = c^2\gamma(w)$, where w is the speed of O' relative to O. Hence either $w \neq 0$, in which case $g(V, V') > c^2$ and eqns (66) have a unique solution for τ and τ', or $w = 0$ and $g(V, V') = c^2$, in which case there are infinitely many solutions.

7.2 Acceleration four-vector

Consider now an accelerating particle. Its worldline is a curve in space-time given by

$$x = x(t), \quad y = y(t), \quad z = z(t) \tag{67}$$

in some ICS (figure ??). Let E (at time t) and E' (at time $t + \delta t$) be two nearby events on the world-line. We define the *proper time* from E to E' to be the time $\delta\tau$ measured in a second ICS in which the particle is instantaneously at rest at the event E.

In the t, x, y, z coordinate system, the displacement vector X from E to E' is

$$X = (c, \boldsymbol{u})\delta t \tag{68}$$

where $\boldsymbol{u}$ is the velocity of the particle. Therefore,

$$c^2\delta\tau^2 = g(X, X) = (c^2 - \boldsymbol{u} \,.\, \boldsymbol{u})\delta t^2\,. \tag{69}$$

Hence, as in the case of a non-accelerating particle,

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = \gamma(u)^{-1} \tag{70}$$

where u is the speed of the particle

Definition 10 *The parameter τ defined up to the addition of a constant by eqn (70) is called the* proper time *along the particle worldline.*

Clock hypothesis

Proper time is the time measured by an accelerating ideal clock travelling with the particle.

'Ideal' means that the physical opperation of the clock is not affected by the acceleration. A pendulum clock, for example, is not 'ideal'.

Exactly as in the definition of the four-velocity of a non- accelerating particle, the quantities

$$V^a = \frac{\mathrm{d}x^a}{\mathrm{d}\tau} \quad \text{and} \quad A^a = \frac{\mathrm{d}^2x^a}{\mathrm{d}\tau^2}\,, \tag{71}$$

(where $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$) are the components of two four-vectors V and A, called the *four-velocity* and *four-acceleration* of the particle. Note that V and A depend on τ and need not be constant along the worldline.

In a general ICS, V and A have spatial and temporal parts

$$V = \gamma(u)(c, \boldsymbol{u}) \tag{72}$$

$$A = \gamma(u)\frac{\mathrm{d}}{\mathrm{d}t}\Big(\gamma(u)(c, \boldsymbol{u})\Big) \tag{73}$$

$$= c^{-2}\gamma^4(c, \boldsymbol{u})u\frac{\mathrm{d}u}{\mathrm{d}t} + \gamma^2\left(0, \frac{\mathrm{d}^2\boldsymbol{u}}{\mathrm{d}t^2}\right) . \tag{74}$$

In an ICS in which the particle is instantaneously at rest,

$$V = (c, 0), \qquad A = (0, \boldsymbol{a})$$

where $\boldsymbol{a} = \mathrm{d}^2\boldsymbol{u}/\mathrm{d}t^2$ is the ordinary acceleration, measured in this ICS. It follows that

$$g(A, V) = 0, \quad g(V, V) = c^2, \quad g(A, A) = -a^2 ,$$

where a is the magnitude of the acceleration measured in the ICS in which the particle is instantaneously at rest; that is, the acceleration 'felt' by an observer moving with the particle.

Lecture 8. Relativistic Motion (Cont)

Example

Constant acceleration along a line. Suppose that $y = z = 0$ along the worldline (in some fixed ICS) and that $a =$ constant. Find the worldline of the particle.

Solution.

The components of V and A in the fixed ICS are

$$(c\dot{t}, \dot{x}, 0, 0) \quad \text{and} \quad (c\ddot{t}, \ddot{x}, 0, 0)\,, \tag{75}$$

where the dot denotes $\mathrm{d}/\mathrm{d}\tau$ (differentiation with respect to proper time).

Hence

$$c^2\dot{t}^2 - \dot{x}^2 = c^2 \quad \text{and} \quad c^2\ddot{t}^2 - \ddot{x}^2 = -a^2\,. \tag{76}$$

By differentiating the first equation and substituting into the second

$$c\ddot{t} = a(\dot{t}^2 - 1)^{1/2}, \qquad c\ddot{x} = a(\dot{x}^2 + c^2)^{1/2}\,. \tag{77}$$

Hence (with a suitable origin for τ)

$$\dot{t} = \cosh(a\tau/c)) \quad \text{and} \quad \dot{x} = c\sinh(a\tau/c) \tag{78}$$

and (with a suitable origin for t and x)

$$t = \frac{c}{a}\sinh\left(\frac{a\tau}{c}\right) \quad \text{and} \quad x = \frac{c^2}{a}\cosh\left(\frac{a\tau}{c}\right)\,. \tag{79}$$

The worldline of the particle is a hyperbola with asymptotes $ct = \pm x$ (figure 11): even though the acceleration (measured in any ICS in which the particle is instantaneously at rest) is constant, the velocity of the particle never exceeds c.

Consider the events on the worldline A at which $x = c^2/a$ and $t = 0$ and B at which $x = c^2/a\cosh(a\tau/c)$ and $t = c/a\sinh(a\tau/c)$. The time between these measured by a clock travelling with the particle is τ. Measured in the fixed ICS t, x, y, z, the time between them is

$$t = \frac{c}{a}\sinh\left(\frac{a}{\tau}c\right) \tag{80}$$

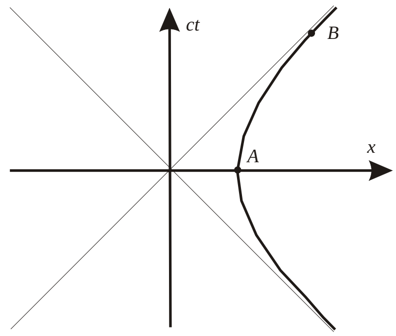

Figure 11: A constant acceleration worldline

and the distance between them is

$$x = \frac{c^2}{a}\cosh\left(\frac{a}{\tau}c\right) \tag{81}$$

If $a = g$ (the acceleration due to gravity) and $\tau = 10$ years, then

$$t = 11,000\,\text{years} \quad \text{and} \quad x = 11,000\text{light-years}. \tag{82}$$

One could exploit this result for inter-stellar travel. Suppose that a spaceship sets out from earth to travel to a star 44,000 light-years from earth. It could do this by accelerating for 10 years with $a = g$ and then decelerating for 10 years with $a = g$, arriving at the star after a passage of 20 years (measured on the spaceship). It could then return in the same way, arriving back at earth after the passage of a further 20 years (again measures on the spaceship).

The catch is that the time between departure and return measured on earth is 44,000 years (the interest on the loan that financed the expedition would by then have exceeded the GNP of the entire universe).

Two points should be noted. First the fact that the spaceship can complete a round trip of 44,000 light-years in 40 years does not involve a violation of the prohibition on faster-than-light travel since the distance and the time are measured in different frames: the distance is measured on earth and the time is measured on the spaceship.

Second, the fact that the time between departure and return measured on earth is not the same as the time measured on the spaceship does not violate the principle of relativity.

There is an asymmetry between the two: the spaceship is accelerating, but the earth is not.

8.1 Operational definition of mass

The concept of mass enters Newtonian mechanics in two ways, as inertial mass (the constant m is the second law $F = ma$) and as gravitational mass (the constants m and m' in the inverse-square law $F = Gmm'/r^2$).

In this chapter, we shall look at how the idea of inertial mass must be modified in the context of special relativity. We shall not attempt to consider gravitational mass and gravitational interactions since they would take us into the realm of the general theory of relativity.

We shall look at dynamical systems consisting of point particles which interact only in collisions, so we shall avoid having to face up to the awkward conflict between the classical idea of 'action at a distance' and the axiom of Einstein's theory that no influence can be transmitted from one particle to another at a speed faster than that of light.

8.2 Collisions in Newtonian mechanics

In Newtonian mechanics, the behaviour of particles in a collision is governed by two laws.

Conservation of mass. If the (inertial) masses of the incoming particles are $m_l, m_2, \ldots, m_k$ and those of the outgoing particles are $m_{k+1}, m_{k+2}, \ldots, m_n$, then

$$\sum_1^k m_i = \sum_{k+1}^n m_i \,.$$

(The number of incoming particles need not be the same as the number of outgoing particles—some of the particles may break up or coalesce.)

Conservation of three-momentum. If the velocities of the incoming particles are $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ and those of the outgoing particles are $\boldsymbol{v}_{k+1}, \boldsymbol{v}_{k+2}, \ldots, \boldsymbol{v}_n$, then

$$\sum_1^k m_i \boldsymbol{v}_i = \sum_{k+1}^n m_i \boldsymbol{v}_i \,.$$

Energy is also conserved, but kinetic energy need not be (unless the collision is elastic) since there is generally an interchange between different forms of energy. In an inelastic collision, kinetic energy is transformed into heat; in an explosion, chemical energy is transformed into kinetic energy.

8.3 Rest mass

The Newtonian conservation laws hold to a very high degree of accuracy where the velocities of the particles are much less than the velocity of light, but as they stand they are not consistent with special relativity since they are not invariant under Lorentz transformations. (The inconsistency will be removed by dropping the implicit assumption that all observers assign the same inertial mass to a particle, irrespective of its velocity).

We need to formulate relativistic collision laws which reduce to the Newtonian laws in the limit of low velocities. The first step is to adopt an operational definition of mass.

Given a standard mass M, an observer can assign a mass m to any other particle by colliding it at low speed with the standard mass, measuring the resulting velocities, and applying the Newtonian law of conservation of momentum. Since this becomes exact as the velocities go to zero, the observer can in principle use a limiting procedure to measure m when the particle is at rest.

Definition. The *rest-mass* of a particle is the mass measured by low speed collisions in an inertial coordinate system in which the particle is at rest.

(The 'limiting procedure' is required to get around the awkwardness of the requirement that the particle should simultaneously be at rest and involved in collisions.)

8.4 Conservation of four-momentum

Associated with any particle, we have a scalar m (its rest mass) and a four-vector V (its four-velocity). The four-vector $P = mV$ is called the *four-momentum* of the particle.

The four-momentum has temporal and spatial parts

$$P = (m\gamma(v)c, m\gamma(v)\boldsymbol{v}), \tag{83}$$

where $\boldsymbol{v}$ is the three-velocity.

As $\boldsymbol{v} \to 0$, $\gamma(v) = 1 + O(v^2/c^2)$ and

$$P = (mc, m\boldsymbol{v}) + O(v^2/c^2). \tag{84}$$

Thus if all the velocities are so small that terms in v^2/c^2 can be neglected, then the Newtonian laws of conservation of mass and momentum are equivalent to the conservation of the temporal and spatial parts of four-momentum.

A straightforward way to arrive at laws for high speed collisions which are consistent with the Lorentz transformation is to adopt the following.

Four-momentum hypothesis If the incoming particles in a collision have four-momenta

$$P_1, P_2, \ldots, P_k$$

and the outgoing particles have four-momenta

$$P_{k+1}, P_{k+2}, \ldots, P_n\,,$$

then

$$\sum_1^k P_i = \sum_{k+1}^n P_i\,. \tag{85}$$

By taking the temporal and spatial parts of (85) (for arbitrary velocities) we have

$$\sum_1^k m_i\gamma(v_i) = \sum_{k+1}^n m_i\gamma(v_i) \tag{86}$$

$$\sum_1^k m_i\gamma(v_i)\boldsymbol{v}_i = \sum_{k+1}^n m_i\gamma(v_i)\boldsymbol{v}_i\,, \tag{87}$$

where the m_i are the rest masses of the particles. These take the same form as the Newtonian laws of mass and momentum conservation when we identify $m\gamma(v)$ with inertial mass and $m\gamma(v)\boldsymbol{v}$ with three-momentum.

Definition. Suppose that a particle of rest mass m has velocity $\boldsymbol{v}$ relative to some ICS. The quantities $m_I = m\gamma(v)$ and $\boldsymbol{p} = m_I\boldsymbol{v}$ are the *inertial mass* and *three-momentum* of the particle relative to the ICS.

Four-momentum conservation is equivalent to conservation of inertial mass and of three-momentum (in every ICS) . The new feature of the relativistic theory is that the inertial mass of a particle increases with its velocity (albeit only very slightly for velocities much less than that of light).

Note that rest mass is a scalar (all observers agree on its value) but inertial mass is different in different inertial coordinate systems. Rest mass and inertial mass are equal for a particle at rest.

Example A particle of rest mass M is at rest when it splits into two particles, each of rest mass m, which move with velocities $(u, 0, 0)$ and $(-u, 0, 0)$. Show that $M = 2m\gamma(u)$.

Solution. By conservation of four-momentum

$$M(c, 0, 0, 0) = m\gamma(u)(c, u, 0, 0) + m\gamma(u)(c, -u, 0, 0)\,, \tag{88}$$

Hence $M = 2m\gamma(u)$. Note that $M > 2m$; and that $m \to 0$ as $u \to c$ (for fixed M).

Lecture 9. Relativistic mechanics (cont)

Example *Elastic collision*. An elastic collision is one in which the rest masses of the particles are unchanged, so there is no exchange between kinetic and 'stored' energy.

A particle of rest mass m is moving with velocity $\boldsymbol{u}$ (relative to some ICS) when it collides elastically with a second particle, also of rest mass m, which is at rest. After the collision, the particles have velocities $\boldsymbol{v}$ and $\boldsymbol{w}$. Show that if θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$, then

$$\cos\theta = \frac{c^2}{vw}\left(1-\sqrt{1-v^2/c^2}\right)\left(1-\sqrt{1-w^2/c^2}\right). \tag{89}$$

Solution. By conservation of four-momentum

$$m(c,0) + m\gamma(u)(c,\boldsymbol{u}) = m\gamma(v)(c,\boldsymbol{v}) + m\gamma(w)(c,\boldsymbol{w}). \tag{90}$$

Therefore

$$\begin{aligned} \gamma(u)+1 &= \gamma(v)+\gamma(w) && (91)\\ \gamma(u)\boldsymbol{u} &= \gamma(v)\boldsymbol{v}+\gamma(w)\boldsymbol{w} && (92)\end{aligned}$$

By taking the modulus-squared of both sides of (92),

$$\gamma(u)^2u^2 = \gamma(v)^2v^2 + \gamma(w)^2w^2 + 2\gamma(v)\gamma(w)\boldsymbol{v}\,.\,\boldsymbol{w}\,. \tag{93}$$

But $\gamma(u)^2u^2 = c^2(\gamma(u)^2-1)$ (a very useful identity). Hence

$$2c^{-2}\gamma(v)\gamma(w)\boldsymbol{v}\,.\,\boldsymbol{w} = \gamma(u)^2 - \gamma(v)^2 - \gamma(w)^2 + 1\,. \tag{94}$$

But by squaring both sides of (91),

$$\gamma(u)^2 + 2\gamma(u) + 1 = \gamma(v)^2 + \gamma(w)^2 + 2\gamma(v)\gamma(w)\,, \tag{95}$$

from which we obtain

$$\begin{aligned} 1+\gamma(u)^2-\gamma(v)^2-\gamma(w)^2 &= 2(\gamma(v)\gamma(w)-\gamma(u))\\ &= 2(\gamma(v)\gamma(w)+1-\gamma(v)-\gamma(w))\\ &= 2(\gamma(v)-1)(\gamma(w)-1)\,.\end{aligned}$$

Hence

$$\frac{\boldsymbol{v}\,.\,\boldsymbol{w}}{vw} = \frac{c^2(\gamma(v)-1)(\gamma(w)-1)}{vw\gamma(v)\gamma(w)}\,, \tag{96}$$

from which (89) follows.

Note that $\cos\theta > 0$ whenever $v, w > 0$, so the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$ is always acute. Also $\cos\theta \to 1$, so $\theta \to 0$ as $v, w \to c$.

This is in contrast with the Newtonian theory, where kinetic energy is conserved as well as momentum since the collision is elastic. The Newtonian theory gives

$$\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w} \qquad \text{and} \qquad u^2 = v^2 + w^2\,, \tag{97}$$

from which it follows that $\boldsymbol{v}\,.\,\boldsymbol{w} = 0$ and $\theta = \pi/2$ for all values of $\boldsymbol{v}$ and $\boldsymbol{w}$.

The fact that θ can be seen in bubble chamber photographs to be very much less than $\pi/2$ for high speed collision offers clear confirmation of the relativistic collision laws.

9.1 Equivalence of mass and energy

In a general collision, it is not rest mass which is conserved, but the temporal part P^0 of the four-momentum. If we neglect terms of order v^4/c^4, but keep terms of order v^2/c^2, then

$$P^0 = c^{-1}(mc^2 + \tfrac{1}{2}mv^2)\,, \tag{98}$$

where m is the rest mass (since $\gamma(v) = 1 + \frac{1}{2}v^2/c^2 + O(v^4/c^4)$). Thus cP^0 is the sum of the Newtonian kinetic energy and a much larger term mc^2, which also has the dimensions of energy.

Definition For a particle of rest mass m, the quantity mc^2 is called the *rest energy* of the particle.

Any collision that involves a gain or loss of kinetic energy (such as an explosion or an inelastic collision) must involve a corresponding loss or gain in the total rest energies of the particles: kinetic energy can be traded for rest mass, and vice-versa.

When v is comparable with c, we must include the higher order terms in v/c. We then have

$$cP^0 = m_I c^2\,, \tag{99}$$

where $m_I = \gamma(v)m$ is the inertial mass; cP^0 is called the total energy of the particle (relative to the ICS). It is usually denoted E (hence $E = m_I c^2$). It is conserved in collisions.

The total energy of a particle is frame-dependent: it is different in different coordinate systems and is minimum in the ICS in which the particle is at rest, where it is equal to the rest energy.

In classical mechanics, the energy that can be stored in a given mass in the form of heat or chemical or nuclear energy is, in principle, unlimited. In special relativity, by contrast, mass and energy are equivalent and any stored energy contributes to inertial mass. If a body is heated, for example, then its rest mass increases (usually by a negligible amount). Conversely the maximum stored energy that can be extracted from a stationary body of rest mass m is its rest energy mc^2. This upper limit is very large. Even in the explosion of an atomic bomb, only about 0.1% of the rest mass is released as other forms of energy.

9.2 The four-momentum of a photon

According to quantum theory, a photon of (angular) frequency ω has energy $E = \hbar\omega$. The relativistic version of this formula associates a four-momentum

$$P = \frac{\hbar\omega}{c}(1, \boldsymbol{e}) \tag{100}$$

with a photon with frequency ω trevelling in the direction of the unit vector $\boldsymbol{e}$.

The rest mass of an ordinary particle can be found from its four-momentum by using the relation $m^2c^2 = g(P, P)$. In the case of a photon, however, $g(P, P) = 0$ since L is null. Photons are therefore examples of 'zero-rest-mass' particles, although this is a misleading term since photons do not have rest frames (they have speed c in all inertial coordinate systems). It is possible that neutrinos are also zero-rest-mass particles.

The law of conservation of four-momentum extends to collisions involving photons.

Example 5.5.1 *Compton Scattering.* A photon of frequency ω collides with an electron of rest-mass m, which is initially at rest. After the collision, the photon has frequently ω'. Show that

$$\hbar\omega\omega'(1 - \cos\theta) = mc^2(\omega - \omega'), \tag{101}$$

where θ is the angle between the initial and final trajectories of the photon.

Solution. In the ICS in which the electron is initially at rest, the four-momenta of the electron before and after the collision are

$$P = m(c, 0) \quad \text{and} \quad Q = m\gamma(u)(c, \boldsymbol{u}), \tag{102}$$

where $\boldsymbol{u}$ is the velocity after the collision.

The four-momenta of the photon before and after the collision are

$$L = \frac{\hbar\omega}{c}(1, \boldsymbol{e}) \quad \text{and} \quad L' = \frac{\hbar\omega'}{c}(1, \boldsymbol{e}'), \tag{103}$$

where $\boldsymbol{e}$ and $\boldsymbol{e}'$ are unit vectors along the initial and final trajectories.

By conservation of four-momentum

$$P + L = Q + L' \tag{104}$$

Now $g(P,P) = g(Q,Q) = m^2c^2$ and $g(L,L) = g(L',L') = 0$. Hence, since $g(L+P, L+P) = g(Q+L', Q+L')$, we have

$$g(L,P) = g(L',Q)\,. \tag{105}$$

Therefore, since $g(L', L'+Q) = g(L', L+P)$, we have

$$g(L,P) = g(L',Q) = g(L',L) + g(L',P)\,. \tag{106}$$

It follows that

$$g(L',L) = g(P, L-L')\,. \tag{107}$$

But from (103) the left-hand side of this is

$$\frac{\hbar^2\omega\omega'}{c^2}(1 - \boldsymbol{e}\,.\,\boldsymbol{e}') = \frac{\hbar^2\omega\omega'}{c^2}(1-\cos\theta) \tag{108}$$

and the right-hand side is

$$m\hbar(\omega - \omega')\,. \tag{109}$$

The result follows.

Lecture 10. Vector fields and invariant operators

10.1 The four-gradient and four-divergence

In Euclidean space, the three partial derivatives $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$ transform as if they were the components of a vector $\boldsymbol{\nabla}$. In other words, if we make the coordinate transformation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = H \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + C \tag{110}$$

where H is a proper orthogonal matrix, then

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} = H \begin{pmatrix} \partial'_x \\ \partial'_y \\ \partial'_z \end{pmatrix}, \tag{111}$$

where $\partial_x = \partial/\partial_x$, $\partial'_x = \partial/\partial x'$ and so on. By making $\boldsymbol{\nabla}$ act on a scalar field f or a vector field $\boldsymbol{X}$, we can form the familiar differential operators

$$\operatorname{grad} f = \boldsymbol{\nabla} f \tag{112}$$

$$\operatorname{div} \boldsymbol{X} = \boldsymbol{\nabla} \,.\, \boldsymbol{X} \tag{113}$$

$$\operatorname{curl} \boldsymbol{X} = \boldsymbol{\nabla} \wedge \boldsymbol{X} \,. \tag{114}$$

These are invariant in the sense that they are the same regardless of which Cartesian coordinate system is used to evaluate them.

We shall now look at the four-vector versions of grad and div (the four-vector version of curl is more complicated). These are the *four-gradient*, which sends a function on space-time to a four-vector field (that is a map from space-time to the space of four-vectors), and the *four-divergence*, which sends a four-vector field to a scalar function. To define them, we need the following.

Let t, x, y, z be an inertial coordinate system and put

$$\nabla^0 = c^{-1}\partial_t, \quad \nabla^1 = -\partial_x, \quad \nabla^2 = -\partial_y, \quad \nabla^3 = -\partial_z \,. \tag{115}$$

Proposition 10 *Under the inhomogeneous Lorentz transformation*

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = L \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} + T\,, \tag{116}$$

the ∇^as behave as the components of a four vector:

$$\begin{pmatrix} c^{-1}\partial_t \\ -\partial_x \\ -\partial_y \\ -\partial_z \end{pmatrix} = L \begin{pmatrix} c^{-1}\partial'_t \\ -\partial'_x \\ -\partial'_y \\ -\partial'_z \end{pmatrix}, \tag{117}$$

where $\partial_t = \partial/\partial\tilde{t}$, $\partial'_t = \partial/\partial t'$, and so on.

Proof. Put $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$, $x'^0 = ct'$, $x'^1 = x'$, $x'^2 = y'$, $x'^3 = z'$. Then

$$x^a = \sum_{b=0}^{4} L^a{}_b x'^b + T^a. \tag{118}$$

Hence by the chain rule

$$\partial'_a = \sum_{b=0}^{3} \frac{\partial x^b}{\partial x'^a}\,\partial_b = \sum_{b=0}^{3} L^b{}_a \partial_b \tag{119}$$

That is

$$\begin{pmatrix} \partial'_0 \\ \partial'_1 \\ \partial'_2 \\ \partial'_3 \end{pmatrix} = L^{\mathrm{t}} \begin{pmatrix} \partial_0 \\ \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix}. \tag{120}$$

But $g(L^{\mathrm{t}})^{-1} = Lg$ because L is the matrix of a Lorentz transformation. Therefore, since $g^{-1} = g$,

$$\begin{pmatrix} c^{-1}\partial_t \\ -\partial_x \\ -\partial_y \\ -\partial_z \end{pmatrix} = g \begin{pmatrix} \partial_0 \\ \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} = g(L^{\mathrm{t}})^{-1} g \begin{pmatrix} \partial'_0 \\ \partial'_1 \\ \partial'_2 \\ \partial'_3 \end{pmatrix} = L \begin{pmatrix} c^{-1}\partial'_t \\ -\partial'_x \\ -\partial'_y \\ -\partial'_z \end{pmatrix} \tag{121}$$

By making the ∇^as act on a function $f(t,x,y,z)$ on space-time, we can form the components $\nabla^a f$ $(a = 0,1,2,3)$ of a four-vector field. This is the four-gradient of f, and is denoted ∇f. It has components

$$c^{-1}\frac{\partial f}{\partial t}, \quad -\frac{\partial f}{\partial x}, \quad -\frac{\partial f}{\partial y}, \quad -\frac{\partial f}{\partial z}\,.$$

The four-divergence of a four-vector field X is the function $\operatorname{Div} X$ formed by taking the 'scalar product' of X with the operator ∇. In other words

$$\begin{aligned}\operatorname{Div} X &= \nabla^0 X^0 - \nabla^1 X^1 - \nabla^2 X^2 - \nabla^3 X^3 \\ &= \frac{1}{c}\frac{\partial X^0}{\partial t} + \frac{\partial X^1}{\partial x} + \frac{\partial X^2}{\partial y} + \frac{\partial X^3}{\partial z}\,. \end{aligned} \tag{122}$$

Proposition 11 *Let X be a four-vector field. Then* $\operatorname{Div} X$ *is invariant.*

Proof. On making the inhomogeneous Lorentz transformation (116), we have

$$\begin{aligned}\operatorname{Div} X &= \begin{pmatrix}\nabla^0 & \nabla^1 & \nabla^2 & \nabla^3\end{pmatrix} g \begin{pmatrix} X^0 \\ X^1 \\ X^2 \\ X^3 \end{pmatrix} \\ &= \begin{pmatrix}\nabla'^0 & \nabla'^1 & \nabla'^2 & \nabla'^3\end{pmatrix} L^{\mathrm{t}} g L \begin{pmatrix} X'^0 \\ X'^1 \\ X'^2 \\ X'3 \end{pmatrix} \\ &= \begin{pmatrix}\nabla'^0 & \nabla'^1 & \nabla'^2 & \nabla'^3\end{pmatrix} g \begin{pmatrix} X'^0 \\ X'^1 \\ X'^2 \\ X'3 \end{pmatrix} \end{aligned} \tag{123}$$

Example

The wave operator. Let $u(t,x,y,z)$ be a function on space-time and put

$$\Box u = \mathrm{Div}\,\nabla u\,, \tag{124}$$

so $\Box u$ is the four-divergence of the four-gradient of u. Then in any inertial coordinate system

$$\Box u = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2}\,. \tag{125}$$

The operator $\Box$ is the *d'Alembertian* (it is a four-dimensional version of the Laplacian ∇^2). Since the four-divergence and four-gradient are invariant, $\Box$ is also invariant: the right-hand side of (125) is the same in any ICS.

10.2 The current four-vector

The charge and current density of a distribution of charges form a four-vector called the *current four vector*. To show this we need to make an assumption.

Assumption. The charge of a particle is the same in all inertial coordinate systems.

One piece of physical evidence for this is the overall neutrality of matter. When at rest, electrons and protons have equal and opposite charges. In an atom the electrons are generally moving much faster than the protons in the nucleus. If the charge of a particle depended on its velocity, then one would not expect an exact balance between the charges of the electrons and the protons.

Proposition 12 *$J^0 = \rho c$ and $\boldsymbol{J}$ are the temporal and spatial parts of a four-vector $\boldsymbol{J}$.*

Proof. Consider first the case of uniform density and current. Suppose that in some ICS, there are n (a constant) particles per unit volume and that each has charge e and velocity $\boldsymbol{u}$ (a constant vector) . Then $\rho = en$ and

$$(J^0, \boldsymbol{J}) = \rho(c, \boldsymbol{u})\,. \tag{126}$$

We know that $\gamma(u)(c, \boldsymbol{u})$ transforms as a four-vector (the four-velocity of the particles) . So our task is to show that $\rho/\gamma(u)$ is the same in all inertial coordinate systems.

Let t', x', y', z' be a coordinate system in which the particles are at rest. Suppose that there are $n'L$ particles in the cube, four of whose vertices are

$$A : (0,0,0) \quad B : (L,0,0) \quad C : (0,L,0) \quad D : (0,0,L)\,. \tag{127}$$

Let t, x, y, z be a second ICS related to t', x', y', z' by the standard Lorentz transformation with velocity u. In the unprimed coordinates, A, B, C, D have worldlines

$$\begin{array}{lcl} A & : & x = ut, \quad y = z = 0 \\ B & : & x = ut + L\sqrt{1-u^2/c^2}, \quad y = z = 0 \\ C & : & x = ut, \quad y = L, \quad z = 0 \\ D & : & x = ut, \quad y = 0, \quad z = L. \end{array} \tag{128}$$

Therefore the cube appears in the second ICS to have volume $L^3/\gamma(u)$ and to be moving with velocity $\boldsymbol{u} = u\boldsymbol{i}$; so there are $n = \gamma(u)n'$ particles per unit volume in the second ICS. Hence, since the charges of the individual particles are the same in both coordinate systems, the charge density $\rho' = en'$ in the rest system is related to the charge density $\rho = en$ in the system t, x, y, z by

$$\rho = en'\gamma(u) = \rho'\gamma(u)\,. \tag{129}$$

Thus $\rho/\gamma(u) = \rho'$, independently of u.

This argument extends to arbitrary Lorentz transformations. Hence $\rho/\gamma(u)$ is the same in all inertial coordinate systems and so $(J^0, \boldsymbol{J})$ transforms as a four-vector in the uniform case.

If $\boldsymbol{u}$ and n are functions of position and time, then we can apply the same argument after first restricting attention to a small neighbourhood of an event in which $\boldsymbol{u}$ and n can be treated as approximately constant.

Finally, we can deal with the general case by regarding a general distribution as a superposition of different streams of charged particles with velocities $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots$, applying the argument to each stream separately, and adding the resulting four-vectors.

We note that the *continuity equation* is $\mathrm{Div}\, J = 0$.

Lecture 11. Poisson's equation and the wave equation

11.1 Poisson's equation

Poisson's equation is

$$\nabla^2 u = -\rho$$

where ∇^2 is the three-dimensional Laplacian and ρ is some given function on x, y, z. It arises in gravitation theory, where u is the gravitational potential and ρ is (a multiple of) the matter density, and also in electrostatics, where ρ is (a multiple of) the charge density.

If V vanishes outside some volume V, then Poisson's equation has a solution

$$u(\boldsymbol{r}) = \frac{1}{4\pi} \int_{\boldsymbol{r}' \in V} \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \, \mathrm{d}v' \,,$$

where $\mathrm{d}V' = \mathrm{d}x' \, \mathrm{d}y' \, \mathrm{d}z'$. Under suitable boundary conditions at infinity (u falls off like $1/r$ and its first derivatives like $1/r^2$), the solution is unique. One can think of it as expressing u as a superposition of the potentials of point masses or charges.

11.2 The wave equation

The scalar wave equation

$$\Box = \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = 0 \tag{130}$$

determines the propagation of waves in three-dimensional space. It generalizes the one-dimensional equation considered in Mods

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0 \tag{131}$$

(in both cases, c is the velocity of the waves).

Eqn (2.6.2) has general solution

$$u = f(x + ct) + g(x - ct) \tag{132}$$

where f and g are arbitrary functions of a single variable. These are also solutions of eqn (130); but it is not possible to write down the general solution of the three-dimensional wave equation in the same simple way. However, for any constant unit vector $\boldsymbol{e}$ and any suitable function f of a single variable,

$$u = f(\boldsymbol{e}\,.\,\boldsymbol{r} - ct)$$

is a solution, since $\nabla^2 u = \boldsymbol{e}\,.\boldsymbol{e} f''$ and $u_{tt} = c^2 f''$. Such a solution is constant on the planes orthogonal to $\boldsymbol{e}$ and progresses in the direction of $\boldsymbol{e}$ with speed c.

Eqn (132) has special solutions called *harmonic waves*, of the form

$$u = A \sin[\omega(c^{-1}x - t) + \epsilon] \tag{133}$$

where A (the *amplitude*), $\omega > 0$ (the *angular frequency*), and ϵ are constants.

The graph of u at fixed t is a sine curve, which moves to the right or left with constant velocity c as t increases. It follows from the theory of the Fourier transform that every solution of (2.6.2) can be written as a 'sum' (in fact, an integral) of harmonic waves.

This result does generalise to three dimensions. A plane (harmonic) wave is a solution of eqn (130) of the form

$$u = \alpha \sin \Omega + \beta \cos \Omega\,. \tag{134}$$

Here

$$\Omega = \omega(c^{-1}\boldsymbol{r}\,.\,\boldsymbol{e} - t)\,,$$

where $\omega > 0$, α, β and $\boldsymbol{e}$ are constant; $\boldsymbol{e}$ is a unit vector which gives the direction of propagation (adding T to t and $cT\boldsymbol{e}$ to $\boldsymbol{r}$ leaves u unchanged).

Example

The 4-vector $K = -c\nabla\Omega = \omega(1, \boldsymbol{e})$ is called the freqeuncy four-vector. It is future-pointing and null. An observer moving through the wave with 4-velocity $V = \gamma(v)(c, \boldsymbol{v})$ will see a harmonic wave in his rest frame with frequency $g(K, V)/c$.

Again it follows by Fourier analysis that every solution of eqn (130) is a combination of plane waves.

11.3 Advanced and Retarded Solutions

We shall now consider the *inhomogeneous* wave equation, which is anlogouse to Poisson's equation:

$$\Box u = \rho\,,$$

where $\rho(t, \boldsymbol{r})$ is a given function of the space-time coordinates, which vanishes for all t outside some bounded region in space. Define ψ and χ as functions of $\boldsymbol{r}$ by replacing t by $-r/c$ in $u(t, \boldsymbol{r})$ and $\partial_t u(t, \boldsymbol{r})$, respectively. We shall solve the equation under the boundary condition

$$\psi = O(r^{-1}), \quad \boldsymbol{\nabla}\psi = O(r^{-2}), \quad \chi = O(r^{-2})\,, \tag{135}$$

as $r \to \infty$ for $t < t_0$, for some fixed t_0.

Proposition 13 *Suppose that u satisfies*

$$\Box u = \rho$$

together with the boundary condition (135). Then

$$u(0,0,0,0) = \frac{1}{4\pi}\int \rho(-r/c, \boldsymbol{r})\,\frac{\mathrm{d}V}{r}\,,$$

where the integral is over all space.

Proof. We have

$$\begin{aligned}
\frac{\partial\psi}{\partial x} &= \frac{\partial u}{\partial x} - \frac{x}{cr}\frac{\partial u}{\partial t} \\
\frac{\partial^2\psi}{\partial x^2} &= \frac{\partial^2 u}{\partial x^2} - \frac{2x}{cr}\frac{\partial^2 u}{\partial x \partial t} + \frac{x^2}{c^2 r^2}\frac{\partial^2 u}{\partial t^2} - \frac{1}{cr}\frac{\partial u}{\partial t} + \frac{x^2}{cr^3}\frac{\partial u}{\partial t} \\
\mathrm{div}\left(\frac{\chi \boldsymbol{r}}{r^2}\right) &= \frac{1}{r^2}\frac{\partial u}{\partial t} + \frac{\boldsymbol{r}}{r^2}\,.\,\boldsymbol{\nabla}\left(\frac{\partial u}{\partial t}\right) - \frac{1}{cr}\frac{\partial^2 u}{\partial t^2}\,,
\end{aligned}$$

where the right-hand sides are evaluated at $t = -r/c$, with similar expressions for the other derivatives. Therefore

$$\frac{1}{r}\nabla^2\psi + \frac{2}{c}\mathrm{div}\left(\frac{\chi\boldsymbol{r}}{r^2}\right) = -\frac{1}{r}\left(\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \nabla^2 u\right) = -\frac{\rho(-r/c, \boldsymbol{r})}{r}\,.$$

and so by the solution of Poisson's equation

$$4\pi\psi(0) = \int_V \frac{\rho(-r/c, \boldsymbol{r})}{r}\,\mathrm{d}V + \int_{S_2}\left[\frac{2\chi\boldsymbol{r}}{cr^2} + \frac{\boldsymbol{\nabla}\psi}{r} - \psi\boldsymbol{\nabla}\left(\frac{1}{r}\right)\right].\,\mathrm{d}\boldsymbol{S}\,.$$

As $r_2 \to \infty$, the surface integral goes to zero by (135).

Proposition 14 *The volume element* $\mathrm{d}N = \mathrm{d}V/r$ *on the past light-cone of the event* $(t, \boldsymbol{r}) = 0$ *is invariant under Lorentz transformations.*

Proof. Since it is clearly invariant under rotations, we have only to show that it is invariant under the standard Lorentz transformation. That is, we must show that if

$$\begin{pmatrix} -r \\ x \\ y \\ z \end{pmatrix} = L_u \begin{pmatrix} -r' \\ x' \\ y' \\ z' \end{pmatrix}$$

where $r' = |\boldsymbol{r}'|$, $\boldsymbol{r}' = (x', y', z')$, and L_u is the standard Lorentz transformation matrix in (21), then

$$\begin{vmatrix} \partial_{x'}x & \partial_{y'}x & \partial_{z'}x \\ \partial_{x'}y & \partial_{y'}y & \partial_{z'}y \\ \partial_{x'}z & \partial_{y'}z & \partial_{z'}z \end{vmatrix} = \frac{r}{r'} .$$

But $y = y'$, $z = z'$, and $x = \gamma(u)(x' - ur'/c)$. Therefore

$$\partial_{x'}x = \gamma(u)\left(1 - \frac{ux'}{r'c}\right) = \frac{r}{r'} .$$

The proposition follows.

Under the boundary condition (135), therefore, we can write the four-potential at an event in the form of an integral over the past light-cone N of the event:

$$u = \frac{1}{4\pi}\int_N \rho \, \mathrm{d}N .$$

This is called the *retarded solution* since the potential is determined by the behaviour of ρ at earlier times. There is an equally valid *advanced solution*, where the integral is over the *future* light cone. Here the boundary condition is the rather more artificial condition that (135) holds, but with $\psi(\boldsymbol{r}) = u(r/c, \boldsymbol{r})$ and $\chi(\boldsymbol{r}) = u_t(r/c, \boldsymbol{r})$.

Lecture 12. Index notation and tensors

To put the space-time coordinates on an equal footing, we write

$$ct = x^0, \quad x = x^1, \quad y = x^2, \quad \text{and} \quad z = x^3 .$$

They then all have dimensions of length. The coordinates are labelled by upper indices. This is important, if unfamiliar: a lot of information will be stored by making a distinction between upper and lower indices.

With this notation, we can write the transformation between inertial coordinate systems in the compact form

$$x^a = \sum_{b=0}^{3} L^a{}_b x'^b + T^a \quad (a = 0, 1, 2, 3). \tag{136}$$

Note that we keep track of the order of the indices on L. The upper index a comes first; it labels the rows of the matrix. The lower index b labels the columns, and comes second. By differentiating, we have that

$$L^a{}_b = \frac{\partial x^a}{\partial x'^b}$$

and that

$$(L^{-1})^a{}_b = \frac{\partial x'^a}{\partial x^b} .$$

Further notational economies are achieved by the adopting the following conventions and special notations.

The summation and range conventions

When an index is repeated in an expression (a *dummy* index), a sum over 0,1,2,3 is implied. An index that is not summed is a *free index*. Any equation is understood to hold for all possible values of its free indices. To apply the conventions consistently, an index must never appear more than twice in any term in an expression, once as an upper index and once as a lower index.

The metric coefficients and the Kronecker delta

We define the quantities g_{ab}, g^{ab} and δ^a_b by

$$g_{ab} = g^{ab} = \begin{cases} 1 & a = b = 0 \\ -1 & a = b \neq 0 \\ 0 & \text{otherwise} \end{cases} \qquad \delta^a_b = \begin{cases} 1 & a = b \\ 0 & \text{otherwise} \end{cases}$$

The notation is very efficient; without it, calculations in relativity tend to be overwhelmed by a mass of summation signs. It does, however, have to be used with care and strict discipline. Free indices—indices for which there is no summation—must balance on the two sides of an equation. Excessive repetition can lead to ambiguous expressions in which it is not possible to restore the summation signs in a unique way. The following illustrate some of the uses and pitfalls of the notation.

Example

We can now omit the summation sign in (136). It becomes

$$x^a = L^a{}_b x'^b + T^a. \tag{137}$$

Repetition of b implies summation over $0, 1, 2, 3$, while the range convention means that the equation is understood to hold as the free index a runs over the values $0, 1, 2, 3$.

We shall also use $\tilde{x}^a$ to denote a second inertial coordinate system: $\sim$ is less likely than $'$ to be confused with an index.

Example

If two events have coordinates x^a and y^a in the first system and x'^a and y'^a in the second system, then

$$x^a - y^a = L^a{}_b(x'^b - y'^b) = L^a{}_b x'^b - L^a{}_b y'^b\,. \tag{138}$$

This illustrates that one must take care about what is meant by a 'term in an expression'. In principle, you should multiply out all the brackets before applying the summation rule—otherwise the three-fold repetition of b in the middle expression could cause confusion. In practice, however, the meaning is clear, and the mild notational abuse in taking the summation through the brackets is accepted without causing difficulty.

Example

The Lorentz condition $L^{\mathrm{t}}gL = g$ becomes

$$L^c{}_a L^d{}_b g_{cd} = g_{cd}\frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b} = g_{ab}\,.$$

Note that it does not matter in which order one writes the Ls and gs so long as the indices are 'wired up' correctly. In this equation a, b are free, while c, d are *dummy indices*—like dummy variables in an integral. The sum over c is the sum in the matrix product $L^{\mathrm{t}}g$, while the sum over d is the sum in the matrix product gL.

Similarly, $L^{-1}g^{-1}(L^{\mathrm{t}})^{-1} = g^{-1}$ becomes

$$g^{cd}\frac{\partial \tilde{x}^a}{\partial x^c}\frac{\partial \tilde{x}^b}{\partial x^d} = g^{ab}\,. \tag{139}$$

Example

If one combines two coordinate transformations

$$x^a = K^a{}_b\tilde{x}^b, \qquad \tilde{x}^a = L^a{}_b x'^b + T^a \tag{140}$$

then the result is

$$x^a = K^a{}_b L^b{}_c x'^c + K^a{}_b T^b\,. \tag{141}$$

To avoid ambiguity, it is necessary to change the dummy index in the second equation before making the substitution. It is then clear that there are two sums, over $b = 0, 1, 2, 3$ and over $c = 0, 1, 2, 3$. If you did not do this, then you would end up with the ambiguous expression $K^a{}_b L^b{}_b$, which could mean $\sum_{b=0}^{3} K^a{}_b L^b{}_b$.

Example

Written in full, the equation $A_aC^a = B_aC^a$ is

$$A_0C^0 + A_1C^1 + A_2C^2 + A_3C^3 = B_0C^0 + B_1C^1 + B_2C^2 + B_3C^3\,.$$

In the compact form, there is a temptation to cancel C^a to deduce that $A_a = C_a$. The full form shows that this temptation must be resisted.

Example

We note the identity

$$g_{ab}g^{bc} = \delta^c_a. \tag{142}$$

This is the matrix identity $g_{..}g^{..} = 1$, where $g_{..}$ and $g^{..}$ are the 4×4 matrices with, respectively, entries g_{ab} and g^{ab}. Here c, a are the free indices and b is a dummy index.

Example

A four-vector is an object with components V^a which transform by

$$V^a = L^a{}_b V'^b$$

under change of inertial coordinates.

Example

The four-velocity: if $x^a = x^a(\tau)$ is the worldline of a particle, parametrized by proper time τ, then the four-velocity has components $V^a = \mathrm{d}x^a/\mathrm{d}\tau$. Under coordinate change $x^a \mapsto x'^a$,

$$V^a = \frac{\mathrm{d}x^a}{\mathrm{d}\tau} = \frac{\partial x^a}{\partial x'^b}\frac{\mathrm{d}x'^b}{\mathrm{d}\tau} = L^a{}_b \frac{\mathrm{d}x'^b}{\mathrm{d}\tau}\,, \tag{143}$$

so the four-vector transformation rule is a consequence of the chain rule.

12.1 Tensors in Minkowski Space

Other objects in special relativity have a similar transformation rule. *Tensor algebra* draws the various rules together into a common framework. The basic idea is that a set of physical quantities measured by one observer can be put together as the components of a single *tensor* in space-time. A four-vector is an example of a tensor. There is then a standard transformation rule that allows one to calculate the components in another coordinate system, and hence the same quantities as measured by a second observer. For example, the energy (divided by c) and momentum of a particle form the time and space components of a four-vector. If they are known in one frame, then the transformation rule gives their values in another.

Example

The *gradient covector* of a function $f(x^a)$ of the space-time coordinates has components $\partial_a f$, where $\partial_a = \partial/\partial x^a$. These transform by the chain rule

$$\partial_a f = \frac{\partial x'^b}{\partial x^a}\tilde{\partial}_b f. \tag{144}$$

Note that it is $\partial x'/\partial x$ on the right-hand side, not $\partial x/\partial\tilde{x}$, so this is *not* the four-vector transformation rule, but rather a dual form of the rule. Hence the term 'covector'.

DEFINITION 11 *A tensor of type (p, q) is an object that assigns a set of components $T^{a\ldots b}{}_{c\ldots d}$ (p upper indices, q lower indices) to each inertial coordinate system, with the transformation rule under change of inertial coordinates from x^a to $\tilde{x}^a$*

$$T^{a\ldots b}{}_{c\ldots d} = \frac{\partial x^a}{\partial\tilde{x}^e}\cdots\frac{\partial x^b}{\partial\tilde{x}^f}\frac{\partial\tilde{x}^h}{\partial x^c}\cdots\frac{\partial\tilde{x}^k}{\partial x^d}\tilde{T}^{e\ldots f}{}_{h\ldots k}.$$

A tensor can be defined at a single event, or along a curve, or on the whole of space-time, in which case the components are functions of the coordinates and we call T a tensor field. If $q = 0$ then there are only upper indices and the tensor is said to be contravariant*; if $p = 0$, then there are only lower indices and the tensor is said to be* covariant*.*

It should be checked that the transformation rule is consistent: that is that in passing from coordinate system x^a to $\tilde{x}^a$ to $\hat{x}^a$, one gets the same transformation as by the direct route from x^a to $\hat{x}^a$. In fact, this follows from the product rule for Jacobian matrices

$$\frac{\partial x^a}{\partial\hat{x}^c} = \frac{\partial x^a}{\partial\tilde{x}^b}\frac{\partial\tilde{x}^b}{\partial\hat{x}^c}\,.$$

Example

A four-vector V^a is a tensor of type $(1, 0)$, also called a vector or contravariant vector.

Example

The gradient covector $\partial_a f$ is a tensor of type $(0,1)$. A tensor α_a of type $(0,1)$ is generally called a covector or covariant vector.

Example

The *Kronecker delta* is a tensor of type $(1,1)$ since

$$\delta^c_d \frac{\partial x^a}{\partial \tilde{x}^c} \frac{\partial \tilde{x}^d}{\partial x^b} = \frac{\partial x^a}{\partial \tilde{x}^c} \frac{\partial \tilde{x}^c}{\partial x^b} = \delta^a_b \,, \tag{145}$$

by the chain rule.

Example

The *contravariant metric* has components g^{ab} and is a tensor of type $(2,0)$, by 139. The *covariant metric* has components g_{ab} and is a tensor of type $(0,2)$.

Both the Kronecker delta and the metric in Minkowski space are special in that they have the same components in every inertial frame. For a general tensor, the components in different frames are not the same.

Lecture 13. Tensors (continued)

13.1 Operations on Tensors

Addition

For S, T of the same type: $S + T$ has components

$$S^{a\ldots b}{}_{c\ldots d} + T^{a\ldots b}{}_{c\ldots d}.$$

Scalars

A scalar at an event is simply a number. A scalar field is a function on space-time. The value of a scalar is unchanged by coordinate transformations. We can multiply a tensor T by a scalar f to get a tensor of the same type with components $fT^{a\ldots b}{}_{c\ldots d}$. The operations of addition and multiplication by constant scalars make the space of tensors of type (p, q) into a vector space of dimension 4^{p+q}.

Tensor product

It S, T are tensors of types (p, q), (r, s), respectively, then the *tensor product* is the tensor of type $(p + r, q + s)$ with components $S^{a\ldots b}{}_{c\ldots d}T^{e\ldots f}{}_{g\ldots h}$. It is denoted by ST or $S \otimes T$.

Differentiation

If T is a tensor field of type (p, q), then ∇T is defined to be the tensor of type $(p, q + 1)$ with components

$$\nabla_a T^{b\ldots c}{}_{d\ldots e} = \partial_a T^{b\ldots c}{}_{d\ldots e}, \quad \partial_a = \frac{\partial}{\partial x^a}.$$

Under change of *inertial* coordinates,

$$\begin{aligned}
\partial_a T^{b\dots}{}_{d\dots} &= \frac{\partial \tilde{x}^t}{\partial x^a}\tilde{\partial}_t \left[\frac{\partial x^b}{\partial \tilde{x}^r}\dots\frac{\partial \tilde{x}^s}{\partial x^d}\dots\tilde{T}^{r\dots}{}_{s\dots}\right] \\
&= \frac{\partial \tilde{x}^t}{\partial x^a}\frac{\partial x^b}{\partial \tilde{x}^r}\dots\frac{\partial \tilde{x}^s}{\partial x^d}\dots\tilde{\partial}_t\tilde{T}^{r\dots}{}_{s\dots}\,,
\end{aligned}$$

which is the correct transformation rule for tensor components of type $(p, q+1)$.

The calculation only works because $\partial x/\partial\tilde{x}$ is constant. We have to work harder to define differentiation in non-inertial coordinates.

Contraction

If T is of type $(p+1, q+1)$, then we can form a tensor S of type (p, q) by *contracting on the first upper index and first lower index* of T:

$$S^{b\dots c}{}_{e\dots f} = T^{ab\dots c}{}_{ae\dots f}\,.$$

Note that there is a sum over a. Under change of coordinates

$$\begin{aligned}
S^{b\dots c}{}_{e\dots f} &= T^{ab\dots c}{}_{ae\dots f} \\
&= \frac{\partial x^a}{\partial \tilde{x}^k}\frac{\partial x^b}{\partial \tilde{x}^l}\dots\frac{\partial x^c}{\partial \tilde{x}^m}\frac{\partial \tilde{x}^s}{\partial x^a}\frac{\partial \tilde{x}^t}{\partial x^e}\dots\frac{\partial \tilde{x}^u}{\partial x^f}\tilde{T}^{kl\dots m}{}_{st\dots u} \\
&= \frac{\partial x^b}{\partial \tilde{x}^l}\dots\frac{\partial x^c}{\partial \tilde{x}^m}\frac{\partial \tilde{x}^t}{\partial x^e}\dots\frac{\partial \tilde{x}^u}{\partial x^f}\tilde{S}^{l\dots m}{}_{t\dots u}
\end{aligned}$$

since

$$\frac{\partial x^a}{\partial \tilde{x}^k}\frac{\partial \tilde{x}^s}{\partial x^a} = \delta^s_k.$$

One can also contract on other pairs of indices, one upper and one lower.

Raising and lowering

If α is a covector and $U^a = g^{ab}\alpha_b$, then U is a four-vector, formed by tensor multiplication combined with contraction. We write α^a for U^a and call the operation 'raising the index'. Raising the index changes the signs of the 1,2,3 components, but leaves the first component unchanged. So

$$(\alpha^0, \alpha^1, \alpha^2, \alpha^3) = (\alpha_0, -\alpha_1, -\alpha_2, -\alpha_3)\,.$$

The reverse operation is 'lowering the index': $V_a = g_{ab}V^b$. One similarly lowers and raises indices on tensors by taking the tensor product with the covariant or contravariant metric and contracting, for example $T^a{}_b = g_{bc}T^{ac}$. One must be careful to keep track of the order of the upper and lower indices since $T^a{}_b$ and $T_b{}^a$ are generally distinct. Do not risk confusion by writing either as T^a_b.

Example

If f is scalar field, then $\nabla^a f$, where

$$(\nabla^a f) = (\partial_t f, -\partial_x f, -\partial_y f, -\partial_z f)$$

is a four-vector field. It is the 'gradient four-vector'.

Example

If X^a is a four-vector field, then $\operatorname{Div} X = \nabla_a X^a$ is the four-divergence. It is an invariant because is it is constructed by tensor operations (differentiation and contraction).

Example

If U and V are four-vectors, then

$$g(U, V) = g_{ab}U^aV^b = U^aV_a = U_aV^a.$$

Example

Raising one index on g_{ab} or lowering one index on g^{ab} gives the Kronecker delta since $g^{ab}g_{bc} = \delta^a_c$.

Example

Suppose that $(S^a) = (1,0,0,0)$ and $(T^a) = (1,1,0,0)$. Then $S \otimes T$ and $T \otimes S$ have respective components

$$(S^aT^b) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad (T^aS^b) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Note that $S \otimes T \neq T \otimes S$, but when written as matrices, as above, the components of $S \otimes T$ and $T \otimes S$ are related by transposition.

13.2 The four-potential

We now return to Maxwell's equations and their consistency with the principle of relativity. We address the question: is there a plausible transformation rule for $\boldsymbol{B}$ and $\boldsymbol{E}$ under change of inertial coordinates that leaves Maxwell's equations invariant? Our main tools will be the electromagnetic potentials ϕ and $\boldsymbol{A}$, in tems of which the fields are given by

$$\boldsymbol{B} = \operatorname{curl} \boldsymbol{A}, \qquad \boldsymbol{E} - \boldsymbol{\nabla}\phi - \partial_t \boldsymbol{A}\,.$$

We can write Maxwell's equations as

$$\Box\phi = \rho/\epsilon_0, \qquad \Box\boldsymbol{A} = \boldsymbol{J}/c\epsilon_0 \tag{146}$$

where ρ and $\boldsymbol{J}$ are the charge and current densities and ϕ and $\boldsymbol{A}$ are the electric and magnetic potentials, in the Lorenz gauge

$$\frac{1}{c^2}\frac{\partial\phi}{\partial t} + \operatorname{div}(\boldsymbol{A}) = 0\,. \tag{147}$$

If we write

$$\Phi = (\phi, c\boldsymbol{A}), \qquad J = (c\rho, \boldsymbol{J}),$$

then the equations become

$$\Box\Phi = \frac{1}{c\epsilon_0} J. \tag{148}$$

We have shown that the right-hand side of (148) transforms as a four-vector. The form of the equation suggests that Φ should as well. In fact under suitable boundary conditions we have the retarded solution

$$\Phi = \frac{1}{4\pi\epsilon_0 c}\int_N J \, \mathrm{d}N.$$

It can be shown that this satisfies (148), and also the Lorenz condition $\nabla_a \Phi^a = 0$ provided that J satisfies the continuity equation $\nabla_a J^a = 0$.

Assumption

ϕ and $c\boldsymbol{A}$ are the temporal and spatial parts of a four-vector Φ (called the *four-potential*).

13.3 The transformation of the electric and magnetic fields

Consider the skew-symmetric tensor F with entries

$$F_{ab} = \nabla_a \Phi_b - \nabla_b \Phi_a \qquad (a, b = 0, 1, 2, 3), \tag{149}$$

where $\Phi_a = g_{ab}\Phi^b$. This is the *electromagnetic field*. We also consider the Contravariant field $F^{ab} = \nabla^a \Phi^b - \nabla^b \Phi^a$ obtained by raising the indices on F_{ab}

Proposition 15 *F^{ab} is given in terms of the components of $\boldsymbol{B}$ and $\boldsymbol{E}$ by*

$$(F^{ab}) = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -cB_3 & cB_2 \\ E_2 & cB_3 & 0 & -cB_1 \\ E_3 & -cB_2 & cB_1 & 0 \end{pmatrix}. \tag{150}$$

Proof. The diagonal entries in F vanish since F is skew-symmetric. The second entry in the first row of F is

$$\begin{aligned} F^{01} &= \nabla^0\Phi^1 - \nabla^1\Phi^0 \\ &= \frac{1}{c}\frac{\partial}{\partial t}(cA_1) + \frac{\partial\phi}{\partial x} \\ &= -E_1\,. \end{aligned} \tag{151}$$

The third entry in the second row is

$$\begin{aligned} F^{12} &= \nabla^1\Phi^2 - \nabla^2\Phi^1 \\ &= -\frac{\partial}{\partial x}(cA_2) + \frac{\partial}{\partial y}(cA_1) \\ &= -c(\operatorname{curl}\boldsymbol{A})_3 \\ &= -cB_3 \end{aligned} \tag{152}$$

The other entries are dealt with in a similar way.

Under the inhomogeneous Lorentz transformation

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = L \begin{pmatrix} c\tilde{t} \\ \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} + C\,, \tag{153}$$

$F = F^{ab}$ transforms according to

$$F = L\tilde{F}L^{\mathrm{t}}\,,$$

where $\tilde{F}$ is defined in the same way as F but with Φ^a and ∇^a replaced by $\tilde{\Phi}^a$ and $\tilde{\nabla}^a$ (the components of Φ and ∇ in the tilde coordinate system). This allows us to relate the components E_1, E_2, E_3, and B_1, B_2, B_3 of the electric and magnetic fields in the original coordinate system to the components $\tilde{E}_1, \tilde{E}_2, \tilde{E}_3$, and $\tilde{B}_1, \tilde{B}_2, \tilde{B}_3$ in the tilde system.

13.4 Standard Lorentz transformations

If L is a standard Lorentz transformation, then written out explicitly, the relationship is

$$\begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -cB_3 & cB_2 \\ E_2 & cB_3 & 0 & -cB_1 \\ E_3 & -cB_2 & cB_1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \gamma & \gamma u/c & 0 & 0 \\ \gamma u/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -\tilde{E}_1 & -\tilde{E}_2 & -\tilde{E}_3 \\ \tilde{E}_1 & 0 & -c\tilde{B}_3 & c\tilde{B}_2 \\ \tilde{E}_2 & c\tilde{B}_3 & 0 & -c\tilde{B}_1 \\ \tilde{E}_3 & -c\tilde{B}_2 & c\tilde{B}_1 & 0 \end{pmatrix} \begin{pmatrix} \gamma & \gamma u/c & 0 & 0 \\ \gamma u/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{154}$$

where $\gamma = \gamma(u)$. (Warning: the standard Lorentz transformation is symmetric, but this is not true of an arbitrary Lorentz transformation matrix)

By equating entries on either side, we have

$$E_1 = \tilde{E}_1, \quad E_2 = \gamma(\tilde{E}_2 + u\tilde{B}_3), \quad E_3 = \gamma(\tilde{E}_3 - u\tilde{B}_2) \tag{155}$$

and

$$B_1 = \tilde{B}_1, \quad B_2 = \gamma(\tilde{B}_2 - u\tilde{E}_3/c^2), \quad B_3 = \gamma(\tilde{B}_3 + u\tilde{E}_2/c^2) \tag{156}$$

so the transformation mixes electric and magnetic fields. If $u << c$, then $\boldsymbol{E} = \tilde{\boldsymbol{E}} - \boldsymbol{u} \wedge \tilde{\boldsymbol{B}}$, where $\boldsymbol{u} = u\boldsymbol{i}$, and $\boldsymbol{B} = \tilde{\boldsymbol{B}}$. Thus an observer moving slowly with velocity $\boldsymbol{u}$ through a pure magnetic field $\boldsymbol{B}$ with $\boldsymbol{E} = 0$ sees an electric field $\tilde{\boldsymbol{E}} = \boldsymbol{u} \wedge \boldsymbol{B}$.

Example

Invariants of the electromagnetic field. It follows from (155) and (156) that

$$\boldsymbol{E} \,.\, \boldsymbol{B} = \tilde{E}_1\tilde{B}_1 + \gamma^2(\tilde{E}_2 + u\tilde{B}_3)(\tilde{B}_2 - \tilde{E}_3/c^2) + \gamma(\tilde{E}_3 - u\tilde{B}_2)(\tilde{B}_3 + u\tilde{E}_2/c^2) = \tilde{\boldsymbol{E}} \,.\, \tilde{\boldsymbol{B}}\,. \tag{157}$$

Hence $\boldsymbol{E} \,.\, \boldsymbol{B}$ is invariant under standard Lorentz transformations. Clearly it is also invariant under rotations. Hence it is an invariant of the electromagnetic field (it is the same in every ICS) It follows, for example, that if $\boldsymbol{E}$ and $\boldsymbol{B}$ are not orthogonal in some ICS, then $\boldsymbol{E} \neq 0$ and $\boldsymbol{B} \neq 0$ in every ICS.

Another invariant is $\boldsymbol{E} \,.\, \boldsymbol{E} - c^2\boldsymbol{B} \,.\, \boldsymbol{B}$.

Lecture 14. Variational principles

14.1 An example

Consider the *Dirichlet problem* of finding a solution to Laplace's equation

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

on the unit disc $D = \{x^2 + y^2 \le 1\}$, subject to the boundary condition $u = f$ on the circle

$$\partial D = \{x^2 + y^2 = 1\}\,,$$

where f is some given function on ∂D. A key idea is that this is equivalent to the problem of *minimizing* the *action integral*

$$S(v) = \tfrac{1}{2}\int_D \boldsymbol{\nabla} v\,.\,\boldsymbol{\nabla} v\,\mathrm{d}A$$

over all functions v on D that are equal to f on ∂D. We shall not be too precise about what is meant by 'all functions'—there is a long story there—but shall simply assume throughout that the functions concerned are 'sufficiently continuous and differentiable' on D, including its boundary circle.

The proof is easy. Suppose that u, v are both equal to f on ∂D, and that $\nabla^2 u = 0$ on the interior of D. Then

$$\int_D \boldsymbol{\nabla} u\,.\,\boldsymbol{\nabla}(u-v)\,\mathrm{d}A = \int_D \mathrm{div}\,((u-v)\,\boldsymbol{\nabla} u)\,\mathrm{d}A = \int_{\partial D}(u-v)\boldsymbol{\nabla} u\,.\,\boldsymbol{n}\,\mathrm{d}s = 0\,. \tag{158}$$

At the first step, we have used $\nabla^2 u = \mathrm{div}\,(\boldsymbol{\nabla} u) = 0$; at the second, we have used the divergence theorem; at the third that $u - v = 0$ on ∂D. Therefore

$$\int_D \boldsymbol{\nabla} u\,.\,\boldsymbol{\nabla} v\,\mathrm{d}S = 2S(u)\,.$$

It follows that

$$0 \le \tfrac{1}{2}\int_D (\boldsymbol{\nabla} u - \boldsymbol{\nabla} v)\,.\,(\boldsymbol{\nabla} u - \boldsymbol{\nabla} v)\,\mathrm{d}A = S(v) - S(u)\,.$$

Therefore $S(v) \ge S(u)$, with equality only if $\boldsymbol{\nabla} u = \boldsymbol{\nabla} v$ everywhere in D. That is, with equality, only if $u - v$ is constant, and therefore zero since it vanishes on the boundary.

We have shown that *the solution to the Dirichlet problem is the unique minimum of the action*.

There is another way to interpret (158) which makes it possible to extend the idea to a wider context. That is, to make the weaker deduction that u is a *critical point* of the action. We suppose that v is close to u and put $\delta u = v - u$. Then (158) implies that at $v = u$

$$\delta S = \int_D \boldsymbol{\nabla} u \,.\, \boldsymbol{\nabla} \delta u \, \mathrm{d}A = 0\,.$$

In other words, our result is that if we make a small variation of u, then the action is unchanged, just as a function of several variables is unchanged to the first order at a critical point under small changes in the variables. In this example, we have $\delta S = 0$ without approximation. In general, we shall try to re-express field equations, like Maxwell's equations, by *variational principles*—that δS should vanish to the first order at a solution under small changes in the field quantities. In relativity, the solutions are generally not minima. Indeed, the worldline of an inertial observer between two given events is singled out amongst all worldlines joining the events by the fact that it *maximizes* proper time. By contrast, in Euclidean space, a straight line *minimizes* distance between two points.

Variational principles are important in mathematical physics for several reasons.

(1) As a technical tool for proving existence of solutions of systems of differential equations. It is often simpler to show by topological arguments that a quantity must achieve its minimum than to show directly that a solution exists.

(2) As a means of finding approximate solutions. Here one looks for critical points under some restricted set of variations labelled by a finite number of parameters. The approximate solution is found by looking for critical points in the ordinary sense of a function of the parameters.

(3) As a link between classical and quantum theory through *Feynman's path integral formalism*. You will have to wait for next year to see that.

(4) As a means of understanding the invariance of field equations under coordinate and other transformations.

(5) As a means of deriving and understanding conservation laws, such as conservation of energy.

It is the last two points that will particularly concern us.

14.2 The divergence theorem

In order to deal with variational principles in four dimensions in the same way as in our example, we need a four-dimensional version of the divergence theorem. We have neither the time nor the need to develop fully the theory of integration in higher dimensions—there is a very elegant theory of *differential forms* that allows the formulation and proof of a general version of *Stokes' theorem* in any dimension in a way that includes the divergence theorem in two, three and four dimensions, as well as the more familiar form of Stokes' theorem for integrals over surfaces.

We shall just need the following.

PROPOSITION 16 Divergence theorem. *Suppose that $X = (\xi, \boldsymbol{x})$ is a four-vector field and that V is a fixed volume in space in some inertial coordinate system t, x, y, z. Let $\mathrm{d}V$ denote the volume element $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ and let t_1, t_2 be two values of t. Then*

$$\int_{t_1}^{t_2}\int_V \operatorname{Div} X\,\mathrm{d}V\,\mathrm{d}t = \frac{1}{c}\int_V \xi(t_2)\mathrm{d}V - \frac{1}{c}\int_V \xi(t_1)\mathrm{d}V + \int_{t_1}^{t_2}\int_{\partial V} \boldsymbol{x}\,.\,\mathrm{d}\boldsymbol{S}\,\mathrm{d}t\,,$$

where ∂V is the boundary of V.

Proof. . Since $\operatorname{Div} X = c^{-1}\partial_t\xi + \operatorname{div}\boldsymbol{x}$, the proposition follows from the three-dimensional divergence theorem, applied to $\boldsymbol{x}$.

COROLLARY 1 *Let D be a bounded region in Minkowski space. That is, it is bounded in space and time—all four coordinates are bounded. Suppose that $X = 0$ outside of D. Then*

$$\int \operatorname{Div} X\mathrm{d}\nu = 0$$

where $\mathrm{d}\nu$ is the space-time 'volume element' $\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ and the integral is over the whole of Minkowksi space.

Proof. This follows from the proposition by choosing t_1, t_2 and V so that D is contained in V between t_1 and t_2. All the terms on the right vanish because $X = 0$ outside of D.

The corollary is still true if X does not actually vanish outside D, but merely falls off sufficiently fast as $t, x, y, z \to \infty$.

14.3 Two space-time examples

We now look at a 'space-time' analogue of the variational principle for Laplace's equation. Let us consider the '*action*'

$$S(u) = \tfrac{1}{2}\int_D (\nabla_a u)(\nabla^a u)\,\mathrm{d}\nu\,,$$

where, u is a scalar function of the space-time coordinates and, as above, D is a bounded region of space-time and $\mathrm{d}\nu$ is the space-time 'volume' element. Let δu be a small variation that vanishes on the boundary of D. Then

$$\delta S(u) = S(u+\delta u) - S(u) = \int_D (\nabla_a \delta u)(\nabla^a u)\,\mathrm{d}\nu\,,$$

where we have ignored terms in the square of δu and its derivatives. However

$$(\nabla_a \delta u)(\nabla^a u) = \mathrm{Div}\,(\delta u \nabla u) - \delta u \Box u\,,$$

where ∇u is the gradient four-vector, with components $\nabla^a u$ and $\Box$ is the d'Alembertian. So by the corollary above,

$$\delta S(u) = -\int_D \delta u \Box u\,\mathrm{d}\nu\,.$$

If the right-hand side is to vanish for all such δu, then $\Box u = 0$ on D (this is not completely obvious, but the issue was dealt with in the second-year calculus of variations course). We conclude that the solutions to the wave equation are the critical points of the action.

Let us look at a second example before we consider the general theory. Suppose that Φ_a is a covector field, and denote by F_{ab} the corresponding tensor field $\nabla_a\Phi_b - \nabla_b\Phi_a$. Put

$$S(\Phi) = -\tfrac{1}{4}\epsilon_0 \int_D F_{ab}F^{ab}\,\mathrm{d}\nu\,,$$

and consider variations $\delta\Phi$ that vanish on the boundary. This time we use

$$F_{ab}\nabla^a\delta\Phi^b = \nabla_a(\delta\Phi_b\, F^{ab}) - \delta\Phi_b\,\nabla_a F^{ab}\,,$$

and apply the corollary with $X^a = \delta\Phi_b\, F^{ab}$. We conclude that the critical points of the action of the solutions to Maxwell's equations $\nabla_a F^{ab} = 0$.

Lecture 15. Relativistic electrodynamics: examples

15.1 Tensor form of Maxwell's equations

The use of tensor formalism and variational prinicples enable us to understand the invariance of electromagnetic theory under Lorentz transformations. To explore the application of the new tools, we first consider the tensor form of Maxwell's equations, as equations on the elctromagnetic field

$$F^{ab} = \nabla^a \Phi^b - \nabla^b \Phi^a \,,$$

which is related to $\boldsymbol{E}$ and $\boldsymbol{B}$ by

$$(F^{ab}) = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -cB_3 & cB_2 \\ E_2 & cB_3 & 0 & -cB_1 \\ E_3 & -cB_2 & cB_1 & 0 \end{pmatrix} .$$

Note that F_{ab} is unchanged if Φ_a is replaced by $\Phi_a + \nabla_a f$; that is, it is invariant under *gauge transformations*. Also

$$\nabla_a F_{bc} + \nabla_b F_{ca} + \nabla_c F_{ab} = 0 \,. \tag{159}$$

since

$$\begin{aligned} & \nabla_a F_{bc} + \nabla_b F_{ca} + \nabla_c F_{ab} \\ & \quad = \nabla_a \nabla_b \Phi_c - \nabla_a \nabla_c \Phi_b + \nabla_b \nabla_c \Phi_a - \nabla_b \nabla_a \Phi_c + \nabla_c \nabla_a \Phi_b - \nabla_c \nabla_b \Phi_a \\ & \quad = 0 \,. \end{aligned}$$

In vector form, Maxwell's equations are

$$\operatorname{div} \boldsymbol{E} = \rho/\epsilon_0, \qquad \operatorname{curl} \boldsymbol{B} - c^{-2}\partial_t \boldsymbol{E} = \mu_0 \boldsymbol{J} \tag{160}$$

together with

$$\operatorname{div} \boldsymbol{B} = 0, \qquad \operatorname{curl} \boldsymbol{E} + \partial_t \boldsymbol{B} = 0 \,. \tag{161}$$

Here $\mu_0 \epsilon_0 = c^{-2}$ and $J = (\rho c, \boldsymbol{J})$. In terms of the electromagnetic field tensor, the first pair are

$$\nabla_a F^{ab} = J^b/\epsilon_0 c \tag{162}$$

and the second pair is equivalent to (159). For example, with $b=0$, eqn (162) is

$$\partial_t F^{00} + \partial_x F^{10} + \partial_y F^{20} + \partial_z F^{30} = \operatorname{div} \boldsymbol{E} = \rho/\epsilon_0$$

while with $b=1$ it is

$$\partial_t F^{01} + \partial_x F^{11} + \partial_y F^{21} + \partial_z F^{31} = -\partial_t E_1 + c\partial_y B_3 - c\partial_z B_2 = J_1/c\epsilon_0$$

which gives the first component of the second equation in (160).

With $a=1, b=2, c=3$, eqn (159) gives

$$\partial_x F_{23} + \partial_y F_{31} + \partial_z F_{12} = c \operatorname{div} \boldsymbol{B} = 0$$

while with $a=0, b=1, c=2$, we have

$$\partial_t F_{12} + \partial_x F_{20} + \partial_y F_{01} = -c\partial_t B_3 - \partial_x E_2 + \partial_y E_1 = 0$$

(note change of sign from lowering indices on F^{ab}), which is the third component of the second equation in (161).

We can now understand the invariance of Maxwell's equations without sources under Lorentz transformations as being a conequence of the fact that they are derived from an invariant Lagrangian density

$$L = -\tfrac{1}{4}\epsilon_0 F_{ab} F^{ab} \,.$$

With sources, we have to add a multiple of $\Phi_a J^a$.

Example

Linearly polarised plane waves. Consider the equations for the four-potential in the Lorenz gauge in the absence of sources

$$\Box\Phi = 0 \qquad \operatorname{Div}\Phi = 0 \,. \tag{163}$$

Let us look for a solution of the form

$$\Phi = P\cos\Omega \tag{164}$$

where P is a constant four-vector, ϵ is a constant, and $\Omega \neq 0$ is an affine linear function of the coordinates. That is

$$\Omega = \omega(\boldsymbol{e}\,.\,\boldsymbol{r}/c - t) + k \tag{165}$$

for some constants ω, $\boldsymbol{e}$, and k. Without loss of generality, $\omega > 0$. Put $K = -c\nabla\Omega$. Then K is a constant four-vector and eqns (163) are equivalent to

$$g(K, K)P \cos\Omega = 0, \quad g(K, P)P \sin\Omega = 0\,. \tag{166}$$

So for Φ to be a solution, K must be a future-pointing null vector, and P must be a spacelike four-vector orthogonal to K. Each component of Φ is a solution of the wave equation with frequency four-vector K. In tensor notation, we have

$$\Phi_a = P_a \cos\Omega, \qquad -c\nabla_a\Omega = K_a$$

The corresponding electromagnetic field tensor is

$$F_{ab} = c^{-1}(K_a P_b - K_b P_a) \sin\Omega$$

If we decompose P and K by putting

$$K = \omega(1, \boldsymbol{e}), \qquad P = (0, \boldsymbol{p}) + \beta\omega(1, \boldsymbol{e}) \tag{167}$$

for some $\beta \in \mathbb{R}$ (chosen so that $P - \beta K$ has vanishing temporal part), then

$$\boldsymbol{E} = \frac{\omega}{c}\boldsymbol{p} \sin\Omega, \qquad \boldsymbol{B} = \frac{1}{c}\boldsymbol{e} \wedge \boldsymbol{E}\,, \tag{168}$$

so the corresponding solution of Maxwell's equation is a linearly polarised plane wave with frequency ω (the temporal part of K).

Note two points. First, adding a constant multiple of βK_a to P_a leaves F_{ab} unchanged, but makes Φ_a undergo the *gauge transformation*

$$\Phi_a \mapsto \Phi_a + \nabla_a(\beta c^{-1}\Omega).$$

Second, the wave appears as a linearly polarized plane wave to all observers. The frequency is given by $\omega = g(K, V)/c$.

Definition 12 *K is the* frequency four-vector *of the electromagnetic wave.*

A general plane electromagnetic wave can be generated by a four-potential of the form $\Phi = P \cos\Omega + Q \sin\Omega$ where P and Q are constant four-vectors.

In the ICS of an observer with four-velocity U, we have

$$K = \omega(1, \boldsymbol{e}), \qquad U = (c, 0)\,. \tag{169}$$

Therefore the frequency measured by the observer is

$$\omega = c^{-1} g(K, U)\,. \tag{170}$$

In some other ICS,

$$K = \omega'(1, \boldsymbol{e}') \qquad \text{and} \qquad U = \gamma(u)(c, \boldsymbol{u}) \tag{171}$$

where ω' and $\boldsymbol{e}'$ are the frequency and direction of propagation of the wave in new ICS. By evaluating the right-hand side of (171) in the new system,

$$\omega = c^{-1}\omega'\gamma(u)(c - \boldsymbol{e}' \,.\, \boldsymbol{u})\,. \tag{172}$$

This is the relativistic *Doppler formula*: it gives the frequency measured by an observer who is moving with velocity $\boldsymbol{u}$ through a plane wave of frequency ω' propagating in the direction of the unit vector $\boldsymbol{e}$.

15.2 The Lorentz force law

We now turn to the problem of finding the equation of motion of a fast moving charge in an electromagnetic field. All we shall assume is that in an inertial coordinate system in which the charge is instantaneously at rest,

$$m\boldsymbol{a} = e\boldsymbol{E}\,, \tag{173}$$

where m is the (constant) rest mass, $\boldsymbol{a}$ is the acceleration relative to the ICS, e is the charge, and $\boldsymbol{E}$ is the electric field in the ICS.

At the event at which the particle is at rest, the components of its four-velocity and four-acceleration are

$$(V^0, V^1, V^2, V^3) = (c, 0, 0, 0), \quad (A^0, A^1, A^2, A^3) = (0, a_1, a_2, a_3) \tag{174}$$

where a_1, a_2, a_3 are the components of $\boldsymbol{a}$. We can therefore write eqn (173) in the form

$$m \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} = \frac{e}{c} \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -cB_3 & cB_2 \\ E_2 & cB_3 & 0 & -cB_1 \\ E_3 & -cB_2 & cB_1 & 0 \end{pmatrix} g \begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix} \tag{175}$$

where F is the electromagnetic field matrix in the ICS. That is

$$mA^a = eF^{ab}V_b/c\,.$$

But this is a tensor equation, so it is valid in any frame. In a general coordinate system, therefore, we have

$$m\frac{\mathrm{d}V^a}{\mathrm{d}\tau} = c^{-1}eF^{ab}V_b\,. \tag{176}$$

That is

$$m\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\gamma(u)\begin{pmatrix} c \\ u_1 \\ u_2 \\ u_3 \end{pmatrix}\right] = e\gamma(u)\begin{pmatrix} \boldsymbol{E}\,.\,\boldsymbol{u}/c \\ E_1 + u_2B_3 - u_3B_2 \\ E_2 + u_3B_1 - u_1B_3 \\ E_3 + u_1B_2 - u_2B_1 \end{pmatrix} \tag{177}$$

where u_1, u_2, u_3 are the components of the velocity relative to the ICS.

Definition 13 *The four-vector $eF^{ab}V_b/c$ is called the* Lorentz four-force. *It has temporal part $e\gamma(u)\boldsymbol{E}\,.\,\boldsymbol{u}/c$ and spatial part $e\gamma(u)(\boldsymbol{E}+\boldsymbol{u}\wedge\boldsymbol{B})$ (i.e. the Lorentz (three-) force multiplied by $\gamma(u)$).*

The temporal and spatial parts of eqn (177) are

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(m\gamma(u)c^2) &= \frac{\mathrm{d}}{\mathrm{d}t}(m_Ic^2) = e\boldsymbol{E}\,.\,\boldsymbol{u} \\ \frac{\mathrm{d}}{\mathrm{d}t}(m\gamma(u)\boldsymbol{u}) &= \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = e(\boldsymbol{E}+\boldsymbol{u}\wedge\boldsymbol{B}) \end{aligned} \tag{178}$$

where $m_I = m\gamma(u)$ is the inertial mass and $\boldsymbol{p} = m\boldsymbol{u}$ is the three-momentum.

The second equation is what we expect to get by combining relativistic mechanics with the Lorentz force law. The first equation (which follows from the second) states that the rate of change of the energy of the particle is $e\boldsymbol{E}\,.\,\boldsymbol{u}$, which is also not surprising.