# 2 Classical Field Theory

These notes are designed to give a broad overview of some important topics in special relativity, classical field theory, and electrodynamics as briefly as possible, and are by no means complete. An elementary understanding of classical mechanics is assumed, as is a familiarity with vector calculus. Some basic knowledge of electromagnetism will be helpful, but is not required. I recommend the books:

- L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (1971)
- D. J. Griffiths, *Intoduction to Electrodynamics* (1999)
- J. D. Jackson, *Classical Electrodynamics* (2001)



## 2.1 The Principle of Relativity

In our study of classical mechanics, we saw that interactions between particles were described by means of potential functions, which depend on the relative positions of particles in the system, $V(\boldsymbol{x}_1 - \boldsymbol{x}_2)$. Such a description necessarily leads to the conclusion of the instantaneous propagation of interaction between particles: a change in the position of one particle immediately impacts the force on another, even if separated by very large distances. Such instantaneous interactions would appear to violate the concept of *locality* – that the motion of a particle should depend only on its immediate surrounds. More importantly, instantaneous interactions are shown by experimentation not to exist in nature. Therefore, we conclude that there must be a finite "speed of propagation" of interactions between bodies. To restore locality, we must describe interactions not as direct action-at-a-distance, but as mediated by physical *fields* whose changes propagate through space-time according to local equations of motion. The study of *field theory* will result in a rich field of physics. In particular, it is required for a full (classical) description of electrodynamics, and will also form the basis from which we can study quantum field theory.

In order to quantitatively study any physical processes, we require a *reference system*, by which we mean a set of coordinates to keep track of the positions of particles, as well as time intervals between events. We also remind the reader

that there exist certain special reference frames, termed *inertial* reference frames, in which free particles do not accelerate. A pair of inertial reference frames move with respect to each other with constant velocity.

One of the key principles upon which classical mechanics was built was the *Galilean* principle of relativity, which encodes the concept of invariance of physical law between different inertial reference frames. One of the key properties of the Galilean relativity is that time is considered absolute; all observers in all reference frames share the identical time coordinate. At the same time, however, Maxwell's theory of electrodynamics implied that the speed of light (in vacuum), $c$, was a consequence of physical law, and, by the principle of relativity, should therefore be the same in all reference frames. The constancy of the speed of light has been verified by many experiments, including the famous experiment of Michelson and Morley. As we shall see, this observation is incompatible with the Galilean principle of relativity.

We shall build field theory from Einstein's *principle of relativity*, which shall supplant the Galilean version:

1. The laws of physics are the same in all inertial reference frames,

2. The speed of light (in vacuum), $c$, is the same for all reference frames.

To see that these conditions are incompatible with the Galilean principle, consider two reference frames $K$ and $K'$, described by coordinates $t$ and $x$, and $t'$ and $x'$, respectively. The $K'$ system moves relative to $K$ (along the $x$ axis) with speed $v$. The Galilean transformations linking the two reference frames are

$$x' = x - vt, \qquad t' = t, \tag{2.1}$$

where importantly, the time coordinate is considered universal and the same in all frames. Under this transformation, the speed of light is clearly not the same in both frames. If the path of a light ray in the $K$ coordinates is $x = ct$, then in $K'$ it would be $x' = ct - vt = (c - v)t$, so the ray moves with speed $c - v$.

Our first job is to seek a set of transformations that will surpass the Galilean transforms, and preserve the constancy of $c$. It turns out that it is not possible to find such a transformation that leaves the concept of time universal between reference frames. An important consequence of this is that events which are simultaneous in one inertial reference frame may not be simultaneous in another.

### 2.1.1 The relativity of simultaneity

We shall begin by considering the canonical example: a pair of observers, $A$ and $B$, moving together along the $x$-axis at constant speed $v$ relative to a stationary observer. We use coordinates $t$ and $x$ to label points as observed in the "stationary" $K$ frame, and coordinates $t'$ and $x'$ in the $K'$ frame that is co-moving with

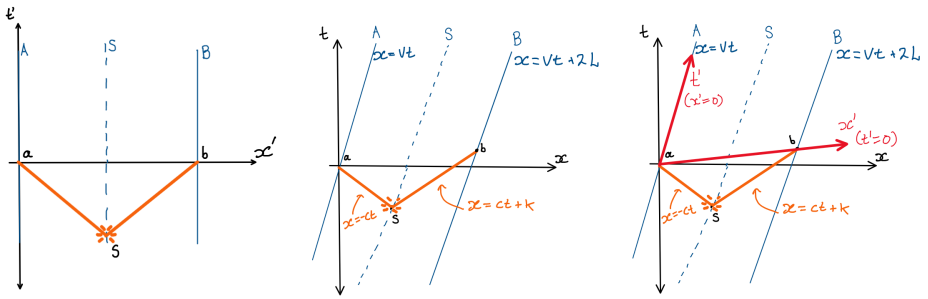


Figure 2.1: Space-time diagrams showing the relevant paths (*world lines*) in the co-moving $K'$ frame, and the "stationary" $K$ frame. The orange curves show the paths of the light rays, the blue lines show the paths of $A$ and $B$, and the dotted line shows the path of the light source, $S$. The space-time point (*event*) of light emission is labelled $s$, and the events of the light reaching $A$ and $B$ are labelled $a$ and $b$. The red lines show the $K'$ axes ($x' = 0$ and $t' = 0$).

$A$ and $B$. (Since there is no motion along the $y$ or $z$ axis, we need not consider them.) Equidistant between $A$ and $B$ is a light source, $S$, which moves along with them and emits a pulse of light:


We choose the origin of the coordinates in the each frame such that the light pulse reaches $A$ at time $t = t' = 0$ and position $x = x' = 0$. The co-moving $K'$ frame is termed the *rest frame* of observers $A$, $S$, and $B$, since in this frame they are at rest with respect to each other.

In the co-moving $K'$ frame, the light pulse clearly reaches $A$ and $B$ at the same time ($t' = 0$), since and $A$ and $B$ are equidistant from the source. Therefore the *events* $a$ and $b$ (the space-time points where the signal reaches observer $A$ and $B$) are *simultaneous*. However, in the stationary frame, it is seen that the light pulse reaches $A$ *before* it reaches $B$. This is because the speed of light, $c$, is the same in all reference frames, but in the time it takes for the light pulse to reach the observers, $A$ has moved towards the source, while $B$ has moved away from it, so the light ray has to travel a longer distance to reach $B$. As a result, the events $a$ and $b$ are *not* simultaneous when observed from the stationary frame. This is shown in Fig. 2.1. It's possible to derive the set of coordinate transformations linking the reference frames geometrically, as is most common in textbooks. Here, we will take a slightly different route by first considering an invariant property called the *spacetime interval*.

### 2.1.2 The space-time interval

As demonstrated above, time intervals as measured in different reference frames are not necessarily the same. By the assertion that the speed of light $c$ is the same in all frames, this directly implies that lengths measured in different reference frames are also not invariant. This may be somewhat disturbing, since regular geometry is built on the concept of invariant lengths; lengths are preserved under spatial rotations and the Galilean transformations. We are thus led to ask: is there an equivalent such invariant quantity within relativistic mechanics?

As is typical, we will begin our analysis by considering the paths taken by rays of light. Consider a light ray, emitted from point $\boldsymbol{x}_1$ at time $t_1$, and arriving at point $\boldsymbol{x}_2$ at time $t_2$. The distance the light ray moves is given by $|\Delta\boldsymbol{x}| \equiv |\boldsymbol{x}_2 - \boldsymbol{x}_1|$, or $\sqrt{\Delta\boldsymbol{x}^2}$.[1] At the same time, since the ray moves at the speed of light, the same distance can be expressed at $c\Delta t$. Therefore, we have the equality:

$$c^2\Delta t^2 - \Delta\boldsymbol{x}^2 = 0. \tag{2.2}$$

Clearly, since the speed of light is the same in all reference frames, the same equality holds in any reference frame:

$$c^2\Delta t'^2 - \Delta\boldsymbol{x}'^2 = 0. \tag{2.3}$$

Motivated by this equality, we define a quantity, *the interval* $\mathrm{d}s$ (or $\mathrm{d}\tau$)[2]:

$$\mathrm{d}s^2 \equiv c^2\mathrm{d}\tau^2 \equiv c^2\mathrm{d}t^2 - \mathrm{d}\boldsymbol{x}^2, \tag{2.4}$$

which is equal to zero for the interval between any two space-time points that can be connected by a light ray ("light-like" separated points), but in general will be non-zero. Note that we have written this as an *infinitesimal* interval $\mathrm{d}s$; since general intervals $\Delta s$ can always be built up from successive infinitesimal ones, there is no loss of generality here.

As we saw, if the interval $\mathrm{d}s^2$ is zero in one frame, then it will be zero in every frame. It is natural to ask if this invariance holds for general intervals. While we do no yet know the precise form of the transformations between different inertial reference frames, Einstein's postulates assert they must be linear: To ensure that inertial motion (uniform velocity) in one frame remains inertial in another, transformations must preserve straight-line trajectories, which requires linear coordinate transformations. Since the interval is quadratic in the coordinates, linearity implies the interval can change only by a scalar factor:

$$\mathrm{d}s'^2 = f\,\mathrm{d}s^2, \tag{2.5}$$

[1] In our notation, bold symbols always represent regular three dimensional vectors, e.g., $\boldsymbol{x} = (x, y, z)$, and we use the shorthand $\boldsymbol{x}^2 \equiv x^2 + y^2 + z^2$.

[2] A common alternative definition is $\mathrm{d}s^2 = \mathrm{d}\boldsymbol{x}^2 - c^2\mathrm{d}t^2 = -c^2\mathrm{d}\tau^2$, as discussed below.

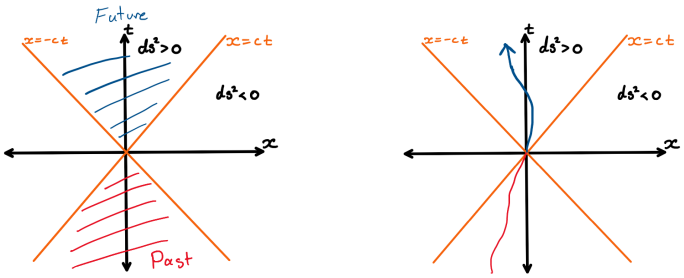


Figure 2.2: There is no frame where the events in upper region occur before the origin; they can therefore be considered strictly future points. The converse holds for the lower region. No points outside the *light cone* can reach (or be reached by) the origin; they are thus absolutely separated from the origin. The *world line* (space-time path) of any particle passing through the origin must remain inside this light cone.

where $f$ is some (yet-undetermined) function. We note that due to the homogeneity of space, $f$ cannot depend explicitly on time or coordinates. Further, due to the isotropy of space, it cannot depend on the *direction* of the velocity, but only its magnitude. Therefore, we must have $f = f(v^2)$.

To proceed, consider a set of three reference frames, where frames 2 and 3 move relative to frame 1 with velocities $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$, respectively. We then have

$$\mathrm{d}s_2^2 = f(v_2^2)\,\mathrm{d}s_1^2, \qquad \mathrm{d}s_3^2 = f(v_3^2)\,\mathrm{d}s_1^2, \quad \text{and} \quad \mathrm{d}s_3^2 = f(v_{23}^2)\,\mathrm{d}s_2^2, \tag{2.6}$$

where $\boldsymbol{v}_{23}$ is the relative velocity of the third system with respect to the second (in Galilean relativity, this would be simply $\boldsymbol{v}_{23} = \boldsymbol{v}_3 - \boldsymbol{v}_2$, though in general is not). Solving for $f(v_{23}^2)$, we find:

$$f(v_{23}^2) = \frac{f(v_3^2)}{f(v_2^2)}. \tag{2.7}$$

Note, however, that $|v_{23}|$ depends not only on the magnitudes of the $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$ vectors, but also on the angle between them: $f(|v_{23}|) = f(|v_2|, |v_3|, \theta_{13})$. Since this angle does not appear on the right-hand-side of Eq. (2.7), the function $f(v^2)$ must be simply a constant, independent of the velocity.[3] Immediately, Eq. (2.7) also tells us this constant must be 1, which implies the interval $\mathrm{d}s$ is invariant.

A pair of space-time points (*events*) connected by interval $\mathrm{d}s^2 > 0$ are said to be *time-like* separated. Events with $\mathrm{d}s^2 < 0$ are said to be *space-like* separated. Since speeds may not exceed $c$, events which are space-like separated can have

[3] To see this more clearly, imagine frame 2 and 3 move with the same velocity with respect to frame 1 – clearly both sides of (2.7) are just 1, since $\boldsymbol{v}_3 = \boldsymbol{v}_2$, and $\boldsymbol{v}_{23} = 0$. Now imagine changing the angle between the velocities of frames 2 and 3, keeping their magnitudes the same; $|v_{13}|$ and $|v_{12}|$ stay fixed by construction, but $|v_{23}|$ becomes non-zero.

no causal contact, and can be considered *absolutely separated*. For any time-like separated events, there exists a frame of reference where they occur at the same position, but no frame in which they occur at the same time. Likewise, for space-like separated events, there exists a frame where they occur at the same time, but not where they occur at the same position; see Fig. 2.2 and Problem 2.2.2.

## 2.2 Lorentz transformation

Transformations between inertial reference frames are termed *Lorentz transformations*. These include the familiar spatial rotations, as well as transformations between reference frames that move relative to each other with constant velocity, known as *Lorentz boosts* (or simply boosts). Since we understand that physics should remain invariant under global rotations of the spatial coordinates (the assumption of isotropy of space), we shall first consider boosts. Note that *translations* (in space and time) are not included in the Lorentz transformations; the generalised set of transformations which also includes translations is termed the *Poincaré* transformations.

We can use the invariance of the interval directly to work out the form of the Lorentz transformations. Consider two inertial reference systems. We label the coordinates of the first system $t$,$x$,$y$,$z$, and of the second $t'$,$x'$,$y'$,$z'$. For simplicity, we will consider the case in which the second frame moves with constant speed $v$ along the $x$-axis with respect to the first (due to the assumed invariance under spatial rotations, we are always free to choose any direction as the $x$-axis).

As we saw above, Galilean transformations, which leave the time coordinate invariant, cannot preserve the invariance of the interval. We therefore consider a general linear transformation between the two coordinates:

$$\begin{aligned} ct' &= \gamma ct + \eta x \\ x' &= \alpha x + \beta ct, \end{aligned} \tag{2.8}$$

where $\alpha$, $\beta$, $\gamma$, $\eta$ are as-yet undetermined dimensionless functions (as $v \to 0$, we have $\alpha, \gamma \to 1$, and $\beta, \eta \to 0$). The speed of light, $c$, is introduced to keep the dimensions of the equations the same. Since there is no motion along the $y$ or $z$ axes, we simply have $y' = y$ and $z' = z$. In order for the interval $\mathrm{d}s^2$ to be invariant, we require

$$\begin{aligned} c^2t^2 - x^2 &= c^2t'^2 - x'^2 \\ &= (\gamma^2 - \beta^2)c^2t^2 - (\alpha^2 - \eta^2)x^2 + 2(\gamma\eta - \alpha\beta)ctx. \end{aligned} \tag{2.9}$$

From this, we see that

$$\gamma^2 - \beta^2 = \alpha^2 - \eta^2 = 1, \quad \text{and} \quad \gamma\eta - \alpha\beta = 0, \tag{2.10}$$

which is three equations in four unknowns.

To break the degeneracy, we consider now the special case of the motion of the origin of the un-primed frame ($x = 0$). In the primed frame, this point moves with $x'/t' = -v$ by construction. By dividing the equations (2.8) and setting $x = 0$, we see that $\beta = -\gamma v/c$, which leaves three equations in three unknowns:

$$\begin{aligned} \gamma^2[1 - (v/c)^2] = \alpha^2 - \eta^2 = 1 \\ \gamma(\eta + \alpha v/c) = 0. \end{aligned} \tag{2.11}$$

From the third equation, we immediately see that $\eta = -v\alpha/c$, and from the first two, it is clear that $\alpha = \gamma$ and

$$\gamma(v) = \frac{1}{\sqrt{1 - (v/c)^2}}. \tag{2.12}$$

(The sign ambiguity can be resolved by noting that $\gamma(0) = 1$.)

Putting things together, the Lorentz transformations for a boost along the $x$ direction with speed $v$ are

$$\begin{aligned} ct' = (ct - \tfrac{v}{c}x)\gamma\,, \qquad y' = y\,, \\ x' = (x - vt)\gamma\,, \qquad z' = z. \end{aligned} \tag{2.13}$$

The generalisation to boosts along an arbitrary axis is fairly clear (we are free to choose any direction as the $x$ axis). There are several important observations to make. Firstly, the factor $\gamma$ is only real valued for $|v| < c$, which implies the principle of relativity can only hold if nothing can move faster than the speed of light. Secondly, notice that by rearranging Eq. (2.13) and solving for $x$ and $t$, we have

$$ct = (ct' + \tfrac{v}{c}x')\gamma, \quad \text{and} \quad x = (x' + vt')\gamma, \tag{2.14}$$

which is the same as Eq. (2.13) under $v \to -v$. The symmetry of this result is intuitive: we could have just as easily started from the primed system, with the unprimed system moving along the $x'$ axis with the opposite speed $-v$.

Also, just as the notion of simultaneity was seen not to be consistent between reference frames, we find that both time intervals and lengths are not necessarily the same in different reference frames. This is known as time dilation and length contraction. For example, the length between $A$ and the source $S$ was observed to be $L$ in the un-primed frame (see Fig. 2.1). It can be seen that this length is *shorter* than the length in the co-moving (or rest) frame, $L_0$, which is termed the *proper length*. In our case, we had $L = |x_S - x_A|$, while $L_0 = |x'_S - x'_A|$, leading to the relation for length contraction:

$$L = L_0/\gamma. \tag{2.15}$$

Using the same logic, and the assertion that the speed of light is the same in all frames, we also have

$$\Delta t = \Delta\tau\gamma, \tag{2.16}$$

where $\Delta\tau$ is the time interval measured in the rest frame. The quantity $\tau$ is known as the *proper time* – it is the time interval as measured in an observer's rest frame. With our convention [see Eq. (2.4)], proper time intervals are related to the space-time interval as $\mathrm{d}s = c\,\mathrm{d}\tau$.

We may now return to the important problem of the addition of velocities. Consider two frames, $K$ and $K'$, where $K'$ moves with speed $V$ along the $x$-axis with respect to $K$. A particle moves with velocity $v'$ in the $K'$ frame. What is the velocity of the particle, $v$, as observed in the $K$ frame? For the simplest case of all velocities co-linear along the $x$-axis, we have $v = dx/dt$, and $v' = dx'/dt'$. By diving the terms from Eq. (2.14), we immediately see:

$$v = \frac{v' + V}{1 + Vv'/c^2}. \tag{2.17}$$

As a final note, we mention that it is customary to work in so-called *natural units*, in which the speed of light takes the value

$$c = 1.$$

This can be done, for example, by measuring lengths in 'light seconds' (the distance covered by light in one second). This makes the equations much simpler, since we may drop the many factors of $c$ that appear all over the place. In such units, velocities become dimensionless; as a result, the units for energies and mass become equivalent. We will use these natural units sometimes, though the $c$ values will be kept in most places. You should get used to swapping between regular and natural units.

***Problem* 2.2.1:** Repeat the derivation of Eq. (2.17) for the case of general velocity $\boldsymbol{v} = d\boldsymbol{x}/dt$, $\boldsymbol{v}' = d\boldsymbol{x}'/dt'$. You may still take the two frames to move along the $x$-axis with velocity $V$ with respect to each other.

*Answer* (2.2.1): $v_x = \frac{v'_x + V}{1 + v'_x V/c^2}$, $v_y = \frac{v'_y\,\gamma(V)}{1 + v'_x V/c^2}$, $v_z = \frac{v'_z\,\gamma(V)}{1 + v'_x V/c^2}$.

***Problem* 2.2.2:** Show that, for two time-like separated events, there exists a frame of reference where they occur at the same position, but no frame where they occur at the same time. Show also the converse for space-like separated events.

*Solution* (2.2.2): Consider two events, $a$ and $b$, where, without loss of generality, we choose coordinates such that $b$ is at the origin at $t = 0$, and orient the axis such that the points lie along the $x$-axis, with $a$ being at position $x$ at time $t$. For time-like separated $a$ and $b$, we have

$$\mathrm{d}s^2 = c^2t^2 - x^2 > 0,$$

which is the same as $|x|/(c|t|) < 1$. Under a Lorentz boost,

$$ct' = (ct - vx/c)\gamma \quad \text{and} \quad x' = (x - vt)\gamma.$$

The condition that the two events occur at the same time in the new frame is $t' = 0$, or $v = c^2t/x$. Since $|x/t| < c$, from the time-like condition, we see that this is impossible, since $|v| < c$. On the other hand, the condition for the events to occur at the same place is $x' = 0$, or $v = x/t$, which is readily satisfied. Therefore, there may always be found a frame where to time-like separated events occur at the same location, but no frame where they occur simultaneously. The proof of the converse for space-like separated events follows the exact same logic.

### 2.2.1 Four vectors

We can arrange the three spatial coordinates and the time-coordinate into a kind of vector, known as a *four-vector* (sometimes called a Lorentz vector):

$$X = (ct, x, y, z)^{\mathrm{Tr}}. \tag{2.18}$$

The components of the four vector are typically denoted $X^\mu$, with $X^0 = ct$, $X^1 = x$, etc. (the convention is that coordinate four vectors have length dimension). It is standard to use Greek indices to denote the components of four-vectors, running 0 – 3, and reserve Latin indices for the spatial components (1 – 3). As is standard, we reserve bold-type for regular three-dimensional vectors, and use regular type for four-vectors. This sometimes leads to confusion. As a result, it is common to use the index notation, and refer to $X^\mu$ as a four vector.[4] It is fairly common to see four-vectors written as $X = (ct, \boldsymbol{x})$, or $x^\mu = (ct, \boldsymbol{x})$.

With this, the Lorentz transformation can be written as a matrix equation,

$$X' = \Lambda X. \tag{2.19}$$

where, for a boost of speed $v$ along the $x$-axis, we have:

$$\Lambda^{(x)}(v) = \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.20}$$

Of course, we could have written this for a boost along a general direction, but the matrix becomes much more complicated (see Appendix 2.11.1 for a derivation of the general case). We can always write $\Lambda = R^{-1}\Lambda^{(x)}R$, where $R$ is the matrix for spatial rotations. The Lorentz transform for spatial rotations are just the standard rotation matrices, leaving the $t$-component unchanged. For example, a rotation about the $z$-axis may be expressed:

$$\Lambda^{(z)}(\theta) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.21}$$

We will formalise this by taking *the definition* of a four vector to be an object that transforms between inertial reference frames according to Eq. (2.19), in much the same way that a regular three-dimensional vector is defined by its transformation under spatial rotations. In the same sense, we may define a *Lorentz scalar* (usually just called a scalar in this context), as being a quantity that is *invariant* under Lorentz transformations.

The concept of Lorentz scalars is very important. Since our goal is to construct a relativistic theory in which laws of physics are the same in all reference frames, it is clearly important to identify the invariant quantities (scalars) which are the same in all reference frames.

[4] Strictly speaking, of course, $a^\mu$ is the $\mu$ component of the vector $a$.

It's important to understand that four-vectors do not obey the rules of regular (Euclidean) geometry. For example, if we naively consider a "dot product"

$$X \cdot X = c^2t^2 + x^2 + y^2 + z^2 \stackrel{?}{=} X' \cdot X'$$

we see from an application of the Lorentz transformation (2.13) that $X \cdot X \neq X' \cdot X'$, and so this quantity is *not* a scalar. Instead, it is the *Minkowski product*, $c^2t^2 - \boldsymbol{x}^2$, that plays the role of the invariant. Our four-dimensional space-time vectors form what is called $(3+1)$-dimensional *pseudo-Euclidean* space, or *Minkowski* space, which will be explored in the next section.

***Problem* 2.2.3:** Show explicitly that $\Lambda^{(x)}(v)^{-1} = \Lambda^{(x)}(-v)$. The physical meaning is clear: a boost of $+v$ followed by one of $-v$ must return us to the original frame.

### 2.2.2 Minkowski space

We can define a quantity, called the *metric tensor*:

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{2.22}$$

which allows us to simplify the expressions. For example, we can now write:

$$\mathrm{d}s^2 = \sum_{\mu\nu} \mathrm{d}x^\mu \eta_{\mu\nu} \mathrm{d}x^\nu. \tag{2.23}$$

We may now also *define* a new type of four-vector, denoted with a lowered index:

$$a_\mu \equiv \sum_\nu \eta_{\mu\nu} a^\nu. \tag{2.24}$$

We call the regular vectors *contravariant* (or "upstairs") vectors, while the other type are called *covariant* (or "downstairs") vectors. If the contravariant vectors have components $a^\mu \doteq (a_0, \boldsymbol{a})$, then the covariant vectors have $a_\mu \doteq (a_0, -\boldsymbol{a})$. Note that the covariant vectors transform as $a'_\mu = \Lambda_\mu{}^\nu a_\nu$, see Problem 2.2.4. We may also define the inverse $\eta^{\mu\nu} = [\eta_{\mu\nu}]^{-1}$ such that:

$$a^\mu \equiv \sum_\nu \eta^{\mu\nu} a_\nu. \tag{2.25}$$

It's straight forward to see that $\eta^{\mu\nu} = \eta_{\mu\nu}$; a direct multiplication shows this is consistent with Eq. (2.24).

We will also introduce here the *Einstein summation convention*, where repeated indices are assumed to be summed over:

$$a^\mu b_\mu \equiv \sum_{\mu=0}^{3} a^\mu b_\mu. \tag{2.26}$$

Using this notation, the Lorentz transformation matrix Eq. (2.20) has the components $\Lambda^{\mu}{}_{\nu}$. It's important to keep track of the position of the indices when there are mixed upstairs and downstairs types, since, in general $T^{\mu}_{\ \nu} \neq T_{\nu}{}^{\mu}$. This notation leads to greatly simplified equations. For example, the interval formula (2.4) can be expressed compactly as

$$\mathrm{d}s^2 = \mathrm{d}x_{\mu}\mathrm{d}x^{\mu}, \tag{2.27}$$

and the Lorentz transform as

$$a'^{\mu} = \Lambda^{\mu}{}_{\nu} a^{\nu}. \tag{2.28}$$

It's easy to check explicitly that $a_{\mu}b^{\mu} = a^{\mu}b_{\mu}$.

It is possible to construct objects with more than one Lorentz index (we have already come across a few). For example, we can construct an object known as a *Lorentz tensor* (or simply a tensor) as $T^{\mu\nu} = A^{\mu}B^{\nu}$. It is clear that this tensor transforms according to

$$T'^{\mu\nu} = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}T^{\alpha\beta}. \tag{2.29}$$

Certainly, not all tensors can be written as the product of two vectors, but we take the *definition* of a tensor to be an object which transforms according to Eq. (2.29). The number of Lorentz indices required to express the components of an object is referred to as its *rank*. In that sense, a scalar is a rank-0 tensor, a vector is a rank-1 tensor, the above tensor is rank 2, and so on. As an aside, we mention that it is always possible, and often very useful, to express a rank 2 tensor $T$ as a sum of symmetric $T^{(+)}$ and anti-symmetric $T^{(-)}$ parts:

$$T_{\mu\nu} = T^{(+)}_{\mu\nu} + T^{(-)}_{\mu\nu} \quad \text{with} \quad T^{(\pm)}_{\mu\nu} \equiv \frac{1}{2}\left(T_{\mu\nu} \pm T_{\nu\mu}\right). \tag{2.30}$$

Finally, we note that a common short-hand is often used for the "square" of four vectors and tensors: $a^2 = a_{\mu}a^{\mu}$, $T^2 = T_{\mu\nu}T^{\mu\nu}$.

We notice that the metric tensor is a special tensor, in the sense that its components remain the same in all reference frames (see Problem 2.2.6). In fact, we can turn this around and use as the *condition* for a tensor $\Lambda$ to be a Lorentz transformation that is satisfies:

$$\Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\,\eta^{\alpha\beta} = \eta^{\mu\nu}. \tag{2.31}$$

This follows logically from the requirement that $a^{\mu}b_{\mu} = a'^{\mu}b'_{\mu} \implies a^{\mathrm{T}}\eta\beta = a'^{\mathrm{T}}\eta\beta'$. It is worth noting that $\eta^{\nu}_{\mu} = \eta_{\mu\alpha}\eta^{\alpha\nu}$ has the same property. This tensor is the identity, and functions as the Kronecker delta in four dimensions: $\eta^{\nu}_{\mu} = \delta^{\nu}_{\mu} = 1$ if $\mu = \nu$, while $\delta^{\nu}_{\mu} = 0$ if $\mu \neq \nu$.

It is appropriate here to note that our definition of the interval is not unique – we could just as easily have defined $\mathrm{d}s^2 = \mathrm{d}\boldsymbol{x}^2 - \mathrm{d}t^2$. In fact, both choices

are commonly made. The other choice corresponds to taking the opposite sign for the metric $\eta$ (2.22). This is usually denoted as the $(+---)$ metric (with the *metric signature* of $-2$) for our choice, or the $(-+++)$ metric (with the metric signature $+2$) for the other. The $+2$ metric is most common among those studying general relativity, while the $-2$ metric is most common in relativistic quantum mechanics. One must take particular care when mixing formulas taken from different works.

***Problem* 2.2.4:** Derive the transformation law for covariant vectors, and show they transform like contravariant ones under $v \to -v$ (i.e., by inverse transform, see Problem 2.2.3).

*Solution* (2.2.4): For the contravariant vectors, we have $A'^{\mu} = \Lambda^{\mu}{}_{\nu} A^{\nu}$, and the covariant vector is defined $A'_{\mu} \equiv \eta_{\mu\nu} A'^{\nu}$. Therefore,

$$A'_{\mu} = \eta_{\mu\rho} \Lambda^{\rho}{}_{\nu} A^{\nu} = \eta_{\mu\rho} \Lambda^{\rho}{}_{\sigma} \eta^{\sigma\nu} A_{\nu} = \Lambda_{\mu}{}^{\nu} A_{\nu}.$$

A straight-forward matrix multiplication $\eta\Lambda\eta$ shows the components of $\Lambda_{\mu}{}^{\nu} A_{\nu}$ are like those of $\Lambda^{\mu}{}_{\nu} A^{\nu}$ with $v \to -v$. (Be careful with the notation: $M_{\mu\nu} v_{\nu} \neq M_{\nu\mu} v_{\nu}$.)

***Problem* 2.2.5:** Prove explicitly that $a_{\mu} b^{\mu}$ is a scalar (i.e., the same in all frames). It is enough to consider boosts along a single axis, noting that Euclidean distances are invariant under spatial rotations.

*Answer* (2.2.5): Follows directly from Problem (2.2.4).

***Problem* 2.2.6:** Show explicitly that the metric tensor is the same in all frames.

***Problem* 2.2.7:** Show the product $S_{\mu\nu} A_{\mu\nu}$ of a symmetric and anti-symmetric tensor is zero.

### 2.2.3 Covariant derivatives and the four-velocity

In regular three dimensional space, we can form vectors from the derivative of scalar functions, $\nabla\phi$. Can we likewise form four-vectors from four-derivatives? In other words, is the quantity

$$\frac{\partial \phi}{\partial x^{\mu}} = \left( \tfrac{\partial \phi}{\partial x^0}, \tfrac{\partial \phi}{\partial x^1}, \tfrac{\partial \phi}{\partial x^2}, \tfrac{\partial \phi}{\partial x^3} \right) \tag{2.32}$$

a four vector (does it transform between frames according to Lorentz transformations)? It's easy to check, since we know how $\mathrm{d}x$ transforms. Considering again the case of boosts along the $x$-axis, from Eq. (2.14) we have

$$\frac{\partial t}{\partial t'} = \frac{\partial x}{\partial x'} = \gamma \quad \text{and} \quad \frac{c\,\partial t}{\partial x'} = \frac{\partial x}{c\,\partial t'} = \frac{v}{c}\gamma.$$

Therefore,

$$\begin{aligned} \frac{\partial \phi}{c\,\partial t'} &= \frac{\partial \phi}{c\,\partial t}\frac{\partial t}{\partial t'} + \frac{\partial \phi}{\partial x}\frac{\partial x}{c\,\partial t'} = \left( \frac{\partial \phi}{c\,\partial t} + \frac{v}{c}\frac{\partial \phi}{\partial x} \right)\gamma, \qquad \text{and} \\ \frac{\partial \phi}{\partial x'} &= \frac{\partial \phi}{\partial x}\frac{\partial x}{\partial x'} + \frac{\partial \phi}{\partial t}\frac{\partial t}{\partial x'} = \left( \frac{\partial \phi}{\partial x} + v\frac{\partial \phi}{c\,\partial t} \right)\gamma. \end{aligned} \tag{2.33}$$

Comparing with Eqs. (2.13) or (2.20), we see that this derivative *nearly* transforms as a four vector – it transforms as though through a Lorentz transformation with the opposite sign for the velocity. That is, it transforms as a *covariant* vector (see Problem 2.2.4). We introduce the standard notation for the *covariant derivative* operator,

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \frac{1}{c}\frac{\partial}{\partial t} + \nabla, \tag{2.34}$$

which makes the covariance of the derivative explicit. It is also useful to define the *contravariant derivative*, $\partial^\mu \equiv \eta^{\mu\nu}\partial_\nu$. With this, we can write simply the d'Alembertian operator $\partial^2 \equiv \partial_\mu\partial^\mu = \frac{1}{c^2}\partial_t^2 - \nabla^2$ (often denoted $\Box$).

Now that we have a solid definition for derivatives, we may be tempted to form a velocity vector as $v^\mu = \mathrm{d}x^\mu/\mathrm{d}t$. It is not difficult to see, however, that this quantity is not a valid four vector: while $\mathrm{d}x^\mu$ is a Lorentz vector, $\mathrm{d}t$ is not a Lorentz scalar. Instead, we can take the derivative of position with respect to *proper time*, $\tau$

$$u^\mu \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}, \tag{2.35}$$

which we call the *four velocity*. Since $\tau$ is a scalar, $u$ is a vector.

We can see the connection to the regular velocity, $\boldsymbol{v}$, by examining the components:

$$u^\mu \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau}. \tag{2.36}$$

Since $\mathrm{d}\tau = \sqrt{\mathrm{d}s^2/c^2} = \mathrm{d}t\sqrt{1 - \mathrm{d}\boldsymbol{x}^2/(c\,\mathrm{d}t)^2} = \mathrm{d}t\sqrt{1-(v/c)^2}$, we see the temporal and spatial components of $U$ are given by

$$u^0 = c\,\gamma, \quad \text{and} \quad u^i = v^i\gamma. \tag{2.37}$$

Note that, by its definition,

$$u^\mu u_\mu = c^2. \tag{2.38}$$

As such, the four velocity only has three independent components.

***Problem* 2.2.8:** Show that $\mathrm{d}t\,\mathrm{d}V$ is a scalar (hint: you may use the invariance of volumes $\mathrm{d}V$ under rotations to arbitrarily choose the $x$-axis).

## 2.3 Relativistic dynamics

We wish to find the rules that govern the dynamics of particles within the framework of the principle of relativity. Following the same logic from classical mechanics, the equations of motion should follow from the *principle of stationary action* (see Mechanics Sec. 1.2). For the equations of motion to be the same in all inertial reference frames, we need an action that is the same in all inertial

reference frames – i.e., the action should be a Lorentz scalar. The simplest scalar is the invariant interval $\mathrm{d}s$, so we can try an action of the form

$$S = kc \int \mathrm{d}s, \tag{2.39}$$

where $k$ is an arbitrary constant, and $c$ is introduced for dimensional convenience. To be dimensionally correct, the $k$ constant should have mass dimensions.

In Mechanics Sec. 1.3, we found based on symmetry arguments that the Lagrangian for a free particle should be a function of $v^2$ only. Those arguments should hold in the relativistic case, and indeed they do. To put things into more familiar form, note that $\mathrm{d}s = \sqrt{c^2\mathrm{d}t^2 - \mathrm{d}\boldsymbol{x}^2}$. By pulling out a factor of $\mathrm{d}t$, we may write

$$S = kc^2 \int \sqrt{1 - (v/c)^2}\, \mathrm{d}t. \tag{2.40}$$

Now, the integrand may be recognised as a Lagrangian $L = kc^2/\gamma$. Consider the non-relativistic expansion, that is, the expansion around small $v$:

$$L \approx kc^2 - k\frac{v^2}{2} - \frac{kv^4}{8c^2} - \dots \tag{2.41}$$

The first term is simply a constant, which will not affect the equations of motion. Therefore, it is the second term that is important in the non-relativistic limit. Note that if we set $k = -m$, then the action recovers the non-relativistic equations of motion (for a free particle) [Mechanics Eq. (1.16)]. Therefore, we have the relativistic Lagrangian for a free particle:

$$L = -mc\frac{\mathrm{d}s}{\mathrm{d}t} = -mc^2/\gamma. \tag{2.42}$$

The canonical momentum, $P_i$ (momentum conjugate to general coordinate $q_i$), can be determined from the Lagrangian [Mechanics (1.18)]

$$P_i = \frac{\partial L}{\partial \dot{q}_i}, \tag{2.43}$$

from which we find the relativistic expression for the linear momentum:

$$\boldsymbol{p} = m\boldsymbol{v}\gamma. \tag{2.44}$$

Likewise, the energy (Hamiltonian) of the free particle can be determined from

$$H = P_i\dot{q}_i - L \tag{2.45}$$

[see Mechanics (1.37)], from which we find:

$$\mathcal{E} = mc^2\gamma. \tag{2.46}$$

Expanding around small $v$,

$$\mathcal{E} \approx mc^2 + \frac{mv^2}{2} + \frac{3mv^4}{8c^2} + \dots \tag{2.47}$$

Therefore, in the low-velocity limit, the energy does not tend to zero, but to a constant – the famous value $mc^2$. The term $\propto v^4$ can be considered the lowest-order relativistic correction to the kinetic energy.[5]

For a particle in a potential, the Lagrangian may gain a position dependence. In analogy with the non-relativistic case, we may write

$$L = -mc^2/\gamma - V(\boldsymbol{x}). \tag{2.48}$$

A straight-forward application of the the Euler-Lagrange equations leads to

$$m\frac{\mathrm{d}(\boldsymbol{v}\gamma)}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -\nabla V, \tag{2.49}$$

which is the relativistic version of Newton's second law. We therefore define the *force* vector in the familiar way

$$\boldsymbol{F} \equiv \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}, \tag{2.50}$$

which may equivalently be written as $\boldsymbol{F}\gamma = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}\tau}$.

### 2.3.1 Four-momentum

Above, we found relativistic expressions for the linear momentum and energy of a particle. It's clear that the energy of a particle is *not* a Lorentz scalar – it depends on the relative velocity of the observer with respect to the particle. At the same time, the relativistic momentum was a three-vector, not a Lorentz four-vector. If we require sets of relativistically invariant equations, we must seek a relativistically invariant way of expressing these important quantities. In other words, is there a way to form a four vector from these quantities?

A natural stating point is to consider the four-velocity, $u^\mu$. We define a quantity $p^\mu$ by multiplying $u^\mu$ by mass

$$p^\mu = mu^\mu. \tag{2.51}$$

Since the four velocity is a Lorentz vector so is $p^\mu$. From Eq. (2.37), we see immediately that the spatial components of $p^\mu$ are exactly the relativistic momentum. Likewise, from Eqs. (2.37) and (2.46), the temporal component is the

[5] In older texts, the "rest mass" $m_0$ was commonly defined via $m = m_0\gamma$. This terminology is no longer widely used; mass $m$ is understood to be a scalar parameter of the theory. We may instead define "rest energy" as $\mathcal{E}_0 = mc^2$.

relativistic energy (divided by $c$). We call this vector the *four momentum*, which has components $p^\mu = (\mathcal{E}/c, \boldsymbol{p})$.

Just as the four velocity has only three independent components (2.38), the same is true for the four momentum. The components of the four momentum are restricted by the condition

$$p^\mu p_\mu = \mathcal{E}^2/c^2 - \boldsymbol{p}^2 = m^2c^2, \tag{2.52}$$

which follows directly from the above expressions. Rearranging, we come to the famous Einstein energy-momentum relation:

$$\mathcal{E}^2 = m^2c^4 + \boldsymbol{p}^2c^2, \tag{2.53}$$

which (of course) is the same as Eq. (2.46).

Note that the above Lagrangian formalism clearly does not work in the case of massless particles. This is a result of the fact that there exists no reference frame where massless particles are at rest. While the relations in Eq. (2.46) and (2.44) are not valid for massless particles, Eq. (2.53) *is* valid. For such particles, this relation implies $|p| = \mathcal{E}/c$.

## 2.4 Classical fields

In general, field theory is the study of continuous distributions, which can be described by continuous functions of time and coordinates: $\phi = \phi(t, \boldsymbol{x})$. While classical fields may describe (near-)continuous distributions of matter, we will focus on the case of the intrinsically continuous fields that mediate the interaction between bodies (e.g., electromagnetic or gravitational fields). Our goal, in general, is to ascertain the set of *field equations* which govern the temporal and spatial evolution of the fields.

Without any assertion for the specific form these equations may take, we can follow the exact same logic as in classical mechanics, and assert they obey a principle of stationary action, with action $S$ defined:

$$S = \int \mathrm{d}t\, L, \tag{2.54}$$

where the Lagrangian $L$ depends on the fields $\phi(t, \boldsymbol{x})$ and their derivatives, $\partial_\mu\phi$. From the principles of homogeneity and isotropy, $L$ cannot depend explicitly on coordinates [e.g., see Mechanics Sec. 1.3].

Further, if we require relativistically invariant equations of motion, then the action should also be relativistically invariant, i.e., a Lorentz scalar. Note, however, that if $S$ is a scalar, then the Lagrangian $L$ in Eq. (2.54) certainly cannot be, since $\mathrm{d}t$ is not a scalar. For this reason, we introduce the *Lagrangian density*[6],

[6]It is common to simply use the term Lagrangian to refer to the Lagrangian density in the context of field theories; it is usually clear from context what is meant.

$\mathcal{L}$, defined via $L \equiv \int \mathrm{d}^3x\, \mathcal{L}$, such that

$$S = \int \mathrm{d}t\, \mathrm{d}V\, \mathcal{L} = \frac{1}{c} \int \mathrm{d}^4x\, \mathcal{L}, \tag{2.55}$$

where $\mathrm{d}^4x \equiv c\, \mathrm{d}t\, \mathrm{d}x\mathrm{d}y\mathrm{d}z$. Since $\mathrm{d}t\, \mathrm{d}V$ *is* a scalar (see Problem 2.2.8), so is $\mathcal{L}$.

The derivation of the Euler-Lagrange equations for a field theory follows closely that in the non-relativistic case; we seek an extremum of the action such that $\delta S = 0$. Since the Lagrangian depends on the field $\phi$ and its derivatives $\partial_\mu \phi$, the general variation of the action is

$$\begin{aligned} \delta S &= \frac{1}{c} \int \mathrm{d}^4x \left[ \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\mu(\delta\phi) \right] \\ &= \frac{1}{c} \int \mathrm{d}^4x \left[ \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \right] \delta\phi, \end{aligned} \tag{2.56}$$

where in the first line we used the commutativity of the derivative to write $\delta(\partial_\mu \phi) = \partial_\mu(\delta\phi)$, and in the second line we used integration by parts (assuming the boundary term goes to zero). For $\delta S$ to be zero for general variations $\delta\phi$, the term in the square brackets must be zero, yielding the Euler-Lagrange equations for the field $\phi$

$$\frac{\partial \mathcal{L}}{\partial \phi} = \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right). \tag{2.57}$$

The generalisation to the case of multiple fields $(\phi_1, \phi_2, \ldots)$ is clear; just as in the case of multiple particles in classical mechanics, there is a separate Euler-Lagrange equation for each field. Also, while we explicitly wrote the equations as though $\phi$ were a scalar field, the derivation is exactly the same for vector (and higher rank) fields $\phi \to \phi^\mu$: there is one equation for each field component.

It's also important to realise that any term added to the Lagrangian density that is a total divergence

$$\mathcal{L} \to \mathcal{L} + \partial_\mu k^\mu, \tag{2.58}$$

will change the action by at most a boundary term (due to the divergence theorem). If the source $k^\mu$ goes to zero at infinity, which is usually the case, then such a term cannot impact the field equations.

Finally, it's instructive to notice what happens in the special case that the field has *no* spatial degrees of freedom, i.e., $\phi = \phi(t)$. The spatial derivatives are zero, and the Euler-Lagrange equations become exactly those from regular (particle) classical mechanics, with $\phi(t)$ playing the role of $x(t)$. A field theory with no spatial degrees of freedom reduces to regular particle mechanics.

### 2.4.1 Field theory as the continuum limit

As a first example of a field theory, we'll consider a simple non-relativistic one dimensional system of coupled harmonic oscillators, which you may have seen

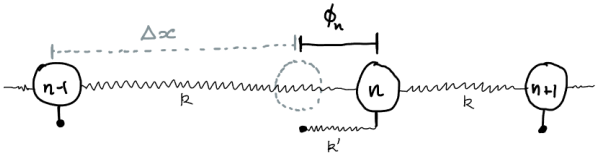


Figure 2.3: A series of masses connected to each other by springs of constant $k$, and held independently in place by springs of constant $k'$. The equilibrium distance is $\Delta x$, and the displacement from equilibrium of mass $n$ is $\phi_n$.

before. Take a large set of particles, each of mass $m$, connected by a series of springs of spring constant $k$. In equilibrium, the masses are separated by $\Delta x$. Finally, we might consider the case where each mass is also independently held in position by a separate spring with constant $k'$, as shown in Fig. 2.3.

To write the Lagrangian, the kinetic energy of each particle is $m\dot{\phi}_n^2/2$, where $\phi_n$ denotes the deviation of the $n$th particle from its equilibrium position. There is also a potential energy associated with each spring; one for each pair of particles given by Hooke's law $k(\Delta\phi_n)^2/2$, where $\Delta\phi_n = \phi_n - \phi_{n-1}$, and one for the independent springs $k'\phi_n^2/2$ (see Fig. 2.3). Then, the Lagrangian is

$$L = \sum_n \left( \frac{m}{2}\dot{\phi}_n^2 - \frac{k}{2}(\Delta\phi_n)^2 - \frac{k'}{2}\phi_n^2 \right). \tag{2.59}$$

To see how this problem can be treated using field theory, we shall take the continuum limit, where the separation between each mass goes to zero $\Delta x \to 0$. To aid in taking the limit, we factor out $\Delta x$

$$L = \Delta x \sum_n \left( \frac{m}{2\Delta x}\dot{\phi}_n^2 - \frac{k\Delta x}{2}\frac{(\Delta\phi_n)^2}{(\Delta x)^2} - \frac{k'}{2\Delta x}\phi_n^2 \right). \tag{2.60}$$

Note that we could write instead $\phi_n = \phi(x_n)$, where $x_n$ is the location of the $n^{\text{th}}$ mass. It's important to note that it is $\phi$, not $x_n$, that is the *dynamical variable*; $x_n$ is simply the *label* for which mass is being referred to.

In taking the $\Delta x \to 0$ limit, we recognise $\Delta\phi_n/\Delta x$ as the derivative $\partial_x\phi$. We further define $\mu = m/\Delta x$ (mass density), $\tau = k\Delta x$ (tension), and $\sigma = k'/\Delta x$ (stiffness), leading to

$$L = \int \mathrm{d}x \left[ \frac{\mu}{2}(\partial_t\phi)^2 - \frac{\tau}{2}(\partial_x\phi)^2 - \frac{\sigma}{2}\phi^2 \right], \tag{2.61}$$

where the term in the brackets may be identified as the Lagrangian density $\mathcal{L}$. The equations of motion are

$$\frac{\partial^2\phi}{\partial t^2} - \frac{\tau}{\mu}\frac{\partial^2\phi}{\partial x^2} + \frac{\sigma}{\mu}\phi = 0. \tag{2.62}$$

In the simplest case of $\sigma = 0$, the solutions are just plane waves travelling with speed $v = \sqrt{\tau/\mu}$. In the case where $\tau = 0$, the equations are just those of the simple harmonic oscillator with $\omega = \sqrt{\sigma/\mu} = \sqrt{k'/m}$, which is of course not surprising. The generalisation to three spatial degrees of freedom is clear: $\partial_x \to \nabla$.

### 2.4.2 Relativistic field theory

In general, we can consider Lagrangians for relativistic field theories that satisfy the following principles:

- To ensure locality, the Lagrangian density should be a *local* function of the fields and their derivatives (i.e., involve only a single spacetime coordinate)
- To ensure relativistic invariance, it should be a Lorentz scalar
- It should depend on terms that are at most first-order in time derivatives, ensuring second-order equations of motion, in analogy with classical mechanics. Lorentz invariance then also implies that only first-order spatial derivatives may appear.

To form a Lorentz invariant Lagrangian density from the fields, we must understand what happens to the fields under Lorentz transformations. In the simplest case of a scalar field, the meaning is reasonably intuitive. A scalar field is simply a function that assigns a number – a scalar – to every space-time location. By definition, scalars remain invariant under Lorentz transformations, so we have

$$\phi(x^\mu) \to \phi'(x^{\mu\prime}) = \phi(x^\mu),$$

where $\phi'$ is the field as written in the transformed coordinates, and $x'^\mu = \Lambda^\mu{}_\nu x^\nu$. That is, the value of the transformed field at the transformed coordinate should be the same as the original field at the original coordinate. It is typically more convenient to write this as

$$\phi'(x) = \phi(\Lambda^{-1}x). \tag{2.63}$$

In much the same way, a vector field is a function that assigns a *vector* to every space-time location. Vectors transform according to Eq. (2.28), so we have

$$A^\mu(x) \to A'^\mu(x') = \Lambda^\mu{}_\nu A^\nu(x), \qquad \text{or}$$

$$A'^\mu(x) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x). \tag{2.64}$$

The extension to tensor and higher-rank forms carries in the same way.

There is another constraint we may place on the Lagrangian that holds in many (but certainly not all) cases of physical interest. Many fields in nature, for example electromagnetic and gravitational fields, obey the *principle of superposition*. If there are two fields, $\phi_1$ and $\phi_2$, which independently satisfy the

field equations and obey the superposition principle, then total field is the simple composition of each, $\Phi = \phi_1 + \phi_2$, which must also obey the field equations. In other words, if $\phi_1$ and $\phi_2$ are both solutions to the field equations, then so must be $\phi_1 + \phi_2$. The field equations of motions resulting from the Euler-Lagrange condition (2.57) are differential equations. For the superposition principle to hold, these must be *linear* differential equations. That is, the field equations of motion must contain terms involving only linear functions in $\phi$. In varying the action (i.e., applying the Euler-Lagrange equations), the degree of the terms in the Lagrangian is reduced by one. This implies the further condition:

- If the field obeys the *superposition principle*, the Lagrangian density may contain terms that are at most quadratic in the field.

The superposition principle is not a universal constraint; in general the Lagrangian may contain any power of the field. It's important to note, however, that higher powers may be important at certain energy scales only. The simplest Lorentz scalar that is quadratic in the fields and containing only first-order derivatives is $\partial_\mu\phi\partial^\mu\phi$. Note this is exactly the first two terms from our simple example in Eq. (2.61) in the case $\tau = \mu$. We may then also consider possible potential terms of higher field powers:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - c_1\phi - c_2\phi^2 - c_3\phi^3 - c_4\phi^4 - \ldots, \tag{2.65}$$

An important physical consequence for cases such as these is that the behaviour of the field may change significantly with the amplitude of the field. For example, for small field perturbations, the quadratic term will dominate (the linear term can generally be eliminated, see Problem 2.4.1). For $\phi \gtrsim c_2/c_3$, the cubic term will become important, and so on. This is the general principle underlying *effective field theories*.

***Problem* 2.4.1:** Consider the scalar Lagrangian in Eq. (2.65), keeping only the derivative, $c_1$ and $c_2$ terms. Show that the linear term can be eliminated by an appropriate constant shift of the field, $\phi \to \varphi = \phi + \phi_0$, such that the equations of motion remain unchanged.

*Solution* (2.4.1): In terms of the new field, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - (c_1 - 2c_2\phi_0)\varphi - c_2\varphi^2 + \text{const.},$$

where we may safely discard any remaining constant terms. Therefore, with the choice $\phi_0 = c_1/(2c_2)$, the linear term can be made to vanish. The resulting Lagrangian is equivalent to one with $c_1 = 0$; therefore a linear term can be removed with a simple field redefinition.

### 2.4.3 Scalar field theory: Klein-Gordon equation

We shall now briefly consider an important example, the classical Klein-Gordon equation for a scalar field $\phi$. The action for a scalar field may be written

$$S = \frac{1}{c}\int \mathcal{L}(\phi, \partial_\mu\phi)\,\mathrm{d}^4x. \tag{2.66}$$

As we discussed above, one of the simplest examples for an invariant Lagrangian can be written as

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{\mu^2}{2}\phi^2, \tag{2.67}$$

where the signs and factors are arbitrary. Since the Lagrangian density has dimension energy per volume, $[\mathcal{L}] = E/L^3$, the field $\phi$ has: $[\phi] = \sqrt{E/L}$.

It is straight forward to find the field equations. From Eq. (2.57), we have

$$\frac{\partial\mathcal{L}}{\partial\phi} = -\mu^2\phi, \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = \partial^\mu\phi. \tag{2.68}$$

From the Euler-Lagrange condition (2.57), this leads to the equations of motion for the Klein-Gordon field:

$$\partial^2\phi + \mu^2\phi = 0, \tag{2.69}$$

which you may recognise as a wave equation. It is not difficult to see that solutions include those of the form of (real) combinations of oscillating plane waves:

$$\phi = e^{\pm ik_\mu x^\mu}. \tag{2.70}$$

These describe classical waves with angular frequency $\omega = ck_0$, and wave vector $\boldsymbol{k}$. On inserting Eq. (2.70) into (2.69), we find the condition for the Klein-Gordon wave four vector:

$$k_\mu k^\mu = \mu^2. \tag{2.71}$$

## 2.5 Symmetries and conservation laws

Following our discussion of symmetries in classical mechanics [Sec. 1.4], we shall consider the extension to relativistic field theories. Consider a general *continuous* transformation of the form $\phi \to \phi + \delta\phi$. Since continuous transformations can always be build from successive infinitesimal ones, we write

$$\delta\phi = \epsilon f(\phi, \partial_\mu\phi, x), \tag{2.72}$$

where $\epsilon$ is a continuous infinitesimal parameter.Another way to write the same thing is $\frac{\partial\phi}{\partial\alpha} = f$, where $\alpha$ is the continuous parameter of the transformation, and $\epsilon = \delta\alpha$. Note that, while we explicitly write the equations as though for scalar fields, the arguments hold for the general case.

This transformation is considered a *symmetry* if it leaves the equations of motion unchanged. From the principle of stationary action, this means it may change the action by at most a constant. In other words, the change in the Lagrangian $\delta\mathcal{L}$ must be either zero, or a total divergence. Therefore, in the assumption that Eq. (2.72) is a symmetry, we have

$$\delta\mathcal{L} = \epsilon\partial_\mu k^\mu, \tag{2.73}$$

for some $k^\mu = k^\mu(\phi, \partial\phi, x)$.

The general variation in the Lagrangian from the variation $\delta\phi$ is

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi). \tag{2.74}$$

If the transform is a symmetry, then the equations of motion (2.57) hold. Using them to re-write the $\partial\mathcal{L}/\partial\phi$ term, we have:

$$\begin{aligned}\delta\mathcal{L} &= \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) \\ &= \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right].\end{aligned} \tag{2.75}$$

Comparing to Eq. (2.73), we thus have

$$\partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}f - k^\mu\right] = 0, \tag{2.76}$$

where the term in the square brackets defines a conserved *current density*, $J^\mu$:

$$\partial_\mu J^\mu = 0, \quad \text{with} \quad J^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}f - k^\mu. \tag{2.77}$$

This important result is called *Noether's theorem*[7]:

> *For every continuous symmetry of the action, there is a corresponding locally-conserved charge and current.*

### 2.5.1 Continuity equations

Note that Noether's theorem (2.77) is a similar, but a much stronger, result than we found in the classical case [see Mechanics Sec. 1.4.4], where we had

$$\frac{\mathrm{d}}{\mathrm{d}t}Q = 0, \quad \text{with} \quad Q \equiv \sum_i p_i f(q_i). \tag{2.78}$$

By comparison, we could define the conserved *charge*,[8] as the volume integral of the zero component of the conserved current density:

$$Q \equiv \frac{1}{c}\int \mathrm{d}^3x\, J_0.$$

[7] E. Noether, *Invariante Variationsprobleme*, Gott. Nachr. **1918**, 235 (1918), E. Noether, *Invariant Variation Problems (Translation)*, arXiv:physics/0503066.

[8] We use the word "charge" here generally to refer to any conserved quantity.

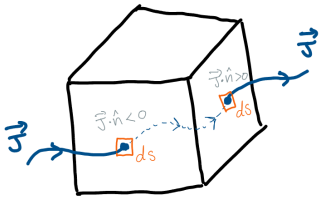


Figure 2.4: The flux of current $\boldsymbol{J}$ through surface, $S$, of a volume, $V$, is the integral of $\boldsymbol{J}\cdot\hat{\boldsymbol{n}}\,\mathrm{d}S$ over the surface, where $\hat{\boldsymbol{n}}$ in the unit vector normal to the surface element $\mathrm{d}S$. By Eq. (2.81), this is equal to the (negative) of the rate of change of the total charge $Q=\int J^0\mathrm{d}V$ in the volume.

The classical case, Eq. (2.78), implies a globally conserved charge, $Q$. The relativistic case, Eq. (2.77), implies the stronger condition of a continuously conserved current:

$$\frac{1}{c}\frac{\partial}{\partial t}J_0+\nabla\cdot\boldsymbol{J}=0, \tag{2.79}$$

which is a *continuity equation*. This implies a *locally* conserved four-current, rather than simply a globally conserved charge. To see this, integrate over an arbitrary volume, $V$:

$$\int \mathrm{d}V\left(\frac{1}{c}\frac{\partial}{\partial t}J_0+\nabla\cdot\boldsymbol{J}\right)=\frac{\partial}{\partial t}Q+\int\mathrm{d}V\,(\nabla\cdot\boldsymbol{J})=0 \tag{2.80}$$

$$\implies\quad \frac{\partial}{\partial t}Q=-\oint_S \boldsymbol{J}\cdot\boldsymbol{n}\,\mathrm{d}S, \tag{2.81}$$

where we used Gauss' theorem in the final step [see Eq. (2.248) in Appendix 2.11.3], $S$ is the surface of the volume $V$, and $\boldsymbol{n}$ is the unit vector perpendicular to the surface. This says that the change in the charge inside any arbitrary volume $V$ [left-hand-side of (2.81)], must be compensated for by a corresponding flux of current through the surface of the volume [right-hand-side of (2.81)]; see Fig. 2.4.

For the particular case of *coordinate transformations*, $x^\mu\to x^\mu+\delta x^\mu$ (e.g., translations, rotations, boosts), we have

$$\delta\phi=(\partial_\mu\phi)\,\delta x^\mu,\quad\text{ and }\quad \delta\mathcal{L}=(\partial_\mu\mathcal{L})\delta x^\mu. \tag{2.82}$$

In many cases of interest (including translations and rotations), the divergence of the coordinate transform is zero: $\partial_\mu\delta x^\mu=0$. Then, we may write the variation in the Lagrangian particularly simply

$$\delta\mathcal{L}=\partial_\mu(\mathcal{L}\delta x^\mu), \tag{2.83}$$

which allows us to write the conserved current as

$$J^\mu=\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\frac{\delta\phi}{\epsilon}-\frac{\delta x^\mu}{\epsilon}\mathcal{L}. \tag{2.84}$$

### 2.5.2 Stress-energy tensor

In Mechanics, we saw that the symmetry of a system under temporal translations led to the law of conservation of energy, and that the symmetry under spatial translations led to momentum conservation. We can now extend this idea to the relativistic scenario. Consider space-time translations of the form

$$x^\nu \to x^\nu + \delta x^\nu, \tag{2.85}$$

where $\delta x^\nu$ is a constant. At first, to get used to manipulating multi-index equations, you may consider $\nu$ to represent a single fixed (yet arbitrary) coordinate index. Under this translation, we have again

$$\delta\phi = \partial_\mu\phi\,\delta x^\mu$$

where, in comparison with the above, we set $f = \partial_\nu\phi$, and $\epsilon = \delta x^\nu$.

Therefore, using Noether's theorem, we have the conserved current:

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi - \partial_\nu x^\mu\mathcal{L}. \tag{2.86}$$

where $\partial_\nu x^\mu = \delta^\mu_\nu$ is 1 for $\mu = \nu$, and zero otherwise. We will have one such equation for each $\nu$ index. The resulting set of equations define the components of a rank two *tensor*, which we term the *stress energy tensor*:

$$T^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L}, \tag{2.87}$$

where we used Eq. (2.25) to raise the second index. Note that by construction (2.76), we have

$$\partial_\mu T^{\mu\nu} = 0, \tag{2.88}$$

so long as space-time translation is a symmetry of the system.

### 2.5.3 Energy, momentum, and pressure

As we saw in our study of classical mechanics, symmetry under time translations (i.e., the above $\delta x^\nu$ with $\nu = 0$), led to energy conservation. Motivated by this, we consider the $\nu = 0$ terms of the stress-energy tensor:

$$\begin{aligned}\partial_\mu T^{\mu 0} &= \partial_0 T^{00} + \partial_i T^{i0} \\ &= \frac{1}{c}\partial_t T^{00} + \nabla\cdot\left(\frac{1}{c}\frac{\partial\mathcal{L}}{\partial(\nabla\phi)}\dot{\phi}\right) = 0,\end{aligned} \tag{2.89}$$

which can be understood as a continuity equation for energy density. We can thus recognise $T^{00}$ as the *energy density*, or *Hamiltonian density*:

$$T^{00} \equiv \mathcal{H} = \frac{\partial\mathcal{L}}{\partial\dot{\phi}}\dot{\phi} - \mathcal{L}, \tag{2.90}$$

which is the quantity that is conserved if the system is invariant under temporal translations. Note the similarity to the classical expression for the Hamiltonian (2.45). The continuity equation (2.89) can therefore be written

$$\frac{\partial}{\partial t}\mathcal{H} = -\nabla \cdot \left( \frac{\partial \mathcal{L}}{\partial(\nabla\phi)}\dot{\phi} \right), \tag{2.91}$$

and we can recognise $cT^{i0}$ (the term in parenthesis) as the *energy density flux*.

Similarly, $g^i \equiv T^{0i}/c$ can be recognised as the *momentum density*, with

$$p^i \equiv \int g^i \mathrm{d}^3x = \int \mathrm{d}^3x \, \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\boldsymbol{x})}\partial^i\phi. \tag{2.92}$$

In analogy with the classical expression (2.43), it is customary to also define the *canonical momentum density* for the field

$$\pi(\boldsymbol{x}) \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\boldsymbol{x})}. \tag{2.93}$$

With this, we may write the expression for the momentum as

$$\boldsymbol{p} = -\int \mathrm{d}^3x \, \pi(\boldsymbol{x}) \, \nabla\phi, \tag{2.94}$$

which is the quantity that is conserved if the system is invariant under spatial translations. This can be written as $g^\mu = T^{0\mu}/c$, or[9]

$$p^\mu = \frac{1}{c}\int \mathrm{d}^3x \, T^{0\mu}. \tag{2.95}$$

We shall now consider the *pressure* exerted by a field within some volume, $V$. We first identify the force on the field as the rate of change of its momentum:

$$F^j = \frac{\mathrm{d}}{\mathrm{d}t}p^j = \frac{\mathrm{d}}{\mathrm{d}t}\int_V \frac{1}{c}T^{0j} \, \mathrm{d}^3x. \tag{2.96}$$

(The source of this force will depend on the specific form of the potential function in the Lagrangian.) Now, from the continuity equation, we have

$$\partial_\mu T^{\mu\nu} = 0 \quad \Longrightarrow \quad \frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}T^{0\nu} = -\partial_i T^{i\nu},$$

where the sum over repeated index $i$ extends over $i = 1, 2, 3$. This leads to

$$F^j = -\int_V \partial_i T^{ij} \, \mathrm{d}^3x = -\oint_S T^{ij} \, \hat{n}_i \, \mathrm{d}S, \tag{2.97}$$

[9] By convention, the stress-energy tensor has energy density units, while the four momentum has momentum units. This sometimes confusing distinction disappears in units with $c = 1$. Note also that $g^\mu$ is *not* a four vector, though $p^\mu$ is.

where we used Gauss' law in the final step, and $\hat{n}$ is the unit vector normal to the surface $S$ of the volume $V$. We therefore conclude that $T^{ij}$ encodes the flux of the $j$-component of momentum across a surface in the $i$ direction. Since, by definition, pressure is the perpendicular force per unit area, we can interpret the diagonal spatial components of the stress tensor, $T^{ii}$ (the $i,i$ component, not the trace), as the pressure exerted *by* the field in the $i$ direction.

These definitions may seems a little abstract for now. One way to justify the identification of these terms as the energy and momentum is to consider the limit of a field with zero spatial degrees of freedom: $\phi(t,\boldsymbol{x}) \to \phi(t)$. As we discussed earlier, in this limit, we recover regular particle mechanics when identifying $\phi(t) = x(t)$. It's straight forward to show that the above expressions for $T^{00}$ and $T^{0i}/c$ do reduce to the classical definitions of energy and momentum density in this limit. The physical meaning of these terms will become much clearer once we consider specific examples.

As a particularly simple example, we can calculate the Hamiltonian of a scalar field that has Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi).$$

Without much work, we find

$$\mathcal{H} = \frac{1}{2c^2}\dot{\phi}^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi) = \frac{1}{2c^2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi), \tag{2.98}$$

where, just as in classical mechanics, we express the Hamiltonian in terms of the conjugate momentum, using Eq. (2.93) to replace all $\dot{\phi}$ terms with $\pi$.

***Problem* 2.5.1:** Find the stress-energy tensor for the classical Klein-Gordon field with Lagrangian in Eq. (2.67), and show explicitly using the equations of motion (2.69) that $\partial_\mu T^{\mu\nu} = 0$.

***Problem* 2.5.2:** Find the expressions for the conserved energy and momentum of this field.

***Problem* 2.5.3:** Assuming this field and its potential are homogenous and isotropic, find the explicit expression for its pressure.

### 2.5.4 Lorentz transformations and angular momentum

As another important example, consider a coordinate transformation corresponding to a general infinitesimal Lorentz transformation (i.e., a rotation and/or boost). Since we consider an infinitesimal, we can write:

$$\Lambda^\mu{}_\nu = 1 + \epsilon\lambda^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon\lambda^\mu{}_\nu. \tag{2.99}$$

For simplicity, we consider the case of a scalar field, which is invariant under Lorentz transformations: $\phi \to \phi'(x) = \phi(\Lambda^{-1}x)$. The corresponding variation in the field and Lagrangian are

$$\delta\phi = -\epsilon(\partial_\mu\phi)\lambda^\mu{}_\nu x^\nu \quad \text{and} \quad \delta\mathcal{L} = -\epsilon(\partial_\mu\mathcal{L})\lambda^\mu{}_\nu x^\nu, \tag{2.100}$$

where we made use of the infinitesimal nature of the transform to express the inverse as $\Lambda^{-1} \approx 1 - \epsilon\lambda$, so that with $x \to \Lambda^{-1}x$, we have $\delta x^\mu = -\epsilon\lambda^\mu{}_\nu x^\nu$.

The condition that $\Lambda$ be a valid Lorentz transformation is that it leaves the metric tensor unchanged, see Eq. (2.31). This allows us to place constraints on the form of $\lambda$. Remembering how tensors transform [Eq. (2.29)], we thus have

$$\begin{aligned}\eta^{\mu\nu} = \eta'^{\mu\nu} &= (\delta^\mu_\alpha + \epsilon\lambda^\mu{}_\alpha)(\delta^\nu_\beta + \epsilon\lambda^\nu{}_\beta)\eta^{\alpha\beta} \\ &= \eta^{\mu\nu} + \epsilon(\lambda^{\nu\mu} + \lambda^{\mu\nu}) + \mathcal{O}(\epsilon^2).\end{aligned} \tag{2.101}$$

Namely, this tells us that $\lambda$ must be anti-symmetric:

$$\lambda^{\nu\mu} + \lambda^{\mu\nu} = 0. \tag{2.102}$$

To proceed, notice that we have

$$\partial_\mu(\lambda^\mu{}_\nu x^\nu) = (\partial_\mu\lambda^\mu{}_\nu)x^\nu + \lambda^\mu{}_\nu\delta^\nu_\mu = 0,$$

where the first term is zero, since, by construction, the Lorentz transform is independent of position, and the second term is zero due to the anti-symmetry of $\lambda$. Therefore, as in Eq. (2.83), we can write the variation in the Lagrangian as

$$\delta\mathcal{L} = -\epsilon\partial_\mu(\mathcal{L}\lambda^\mu{}_\nu x^\nu), \tag{2.103}$$

which is a total derivative. We can therefore apply Noether's theorem in the form of Eq. (2.84), and write the corresponding conserved current as

$$\begin{aligned}J^\alpha &= -\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\phi)}(\partial_\mu\phi)\lambda^\mu{}_\nu x^\nu + \lambda^\alpha{}_\nu x^\nu\mathcal{L} \\ &= -\lambda^\mu{}_\nu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\phi)}(\partial_\mu\phi) - \delta^\alpha_\mu\mathcal{L}\right)x^\nu.\end{aligned}$$

Notice that the term in the parentheses is just the stress-energy tensor (2.87), and so we have the conserved current $\partial_\alpha J^\alpha = 0$, with

$$J^\alpha = -\lambda^\mu{}_\nu T^\alpha{}_\mu x^\nu = -\lambda_{\mu\nu}T^{\alpha\mu}x^\nu. \tag{2.104}$$

We can see this as actually being a set of six conserved currents, one for each of the possible terms of $\lambda$ (being an anti-symmetric four-tensor, $\lambda$ has 6 independent components). This is not too surprising, since there are six independent Lorentz transformations (three rotations, and three boosts). From the anti-symmetry of $\lambda$, we can write

$$J^\alpha = -\lambda_{\mu\nu}T^{\alpha\mu}x^\nu = \frac{1}{2}\lambda_{\mu\nu}\left(x^\mu T^{\alpha\nu} - x^\nu T^{\alpha\mu}\right). \tag{2.105}$$

We may therefore define

$$\mathcal{M}^{\alpha\mu\nu} = x^\mu T^{\alpha\nu} - x^\nu T^{\alpha\mu}, \tag{2.106}$$

which satisfies $\partial_\alpha \mathcal{M}^{\alpha\mu\nu} = 0$, since $\partial\lambda = 0$. As we will see below, this can be thought of as a relativistic field-theory extension of the angular momentum.

Note that, using the definition of the momentum in terms of the stress-energy tensor (2.95), we may define

$$M^{ij} \equiv \int \mathrm{d}^3x \mathcal{M}^{0ij} = \int \mathrm{d}^3x (x^i T^{0j} - x^j T^{0i}). \tag{2.107}$$

Since, from Eq. (2.92), we recognise $T^{0j}$ as the momentum density ($j$-component), we recognise $M^{ij}$ as an expression for the $k$-component of the angular momentum of the field. This is conserved due to the symmetry under rotations ($\mu, \nu = 1, 2, 3$).

### 2.5.5 Uniqueness of the stress-energy tensor

It's important to notice that the definition of the stress-energy tensor is not unique. If we add to the tensor a derivative of the form

$$T^{\mu\nu} + \partial_\lambda K^{\mu\lambda\nu}, \quad \text{with} \quad K^{\mu\lambda\nu} = -K^{\lambda\mu\nu}, \tag{2.108}$$

(i.e., where $K$ is anti-symmetric in its first two indices), then the continuity condition $\partial_\mu T^{\mu\nu} = 0$ (2.88) will remain unchanged. To see this, note:

$$\partial_\mu \partial_\lambda K^{\mu\lambda\nu} = -\partial_\mu \partial_\lambda K^{\lambda\mu\nu} = -\partial_\lambda \partial_\mu K^{\lambda\mu\nu} = -\partial_\mu \partial_\lambda K^{\mu\lambda\nu} = 0,$$

where we used the anti-symmetry of $K$, then the commutativity of the derivative, and then relabelled the dummy indices $\mu \leftrightarrow \lambda$. Since this derivative is zero, Eq. (2.88) will not be changed. It is also possible to show that the total energy and momentum (2.95) are not impacted by this change. We shall not go through the details here, but it turns out that the stress energy tensor may be uniquely chosen so that it is symmetric (see, e.g., Landau Sec. 84[10]).

## 2.6 Particle and field interactions

In a quantum field theory, particles and fields will become unified, and will be described by a single quantum action. Classically, however, they are treated differently. If we consider a set of particles, described by coordinates $x$, as well as a set of fields $\phi$, then the total action may be expressed as the sum of terms involving only particle coordinates, terms involving only fields, and interaction terms, which involve both:

$$S = S_{\text{particle}}[x] + S_{\text{field}}[\phi] + S_{\text{interaction}}[x, \phi] \tag{2.109}$$

[10] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (1971).

(for simplicity, we write the equations for a single particle and a single field, though the generalisation is clear). It is no surprise that if there is no interaction between the particles and the fields, then they evolve independently, and may be considered separately. If the interaction term is non-zero, however, the equations of motion will not be separable; physically this means that the field impacts how the particle evolves, and likewise, the particle impacts how the field evolves.

In certain circumstances, a simplifying approximation is possible. If the interaction between the particle and the field is sufficiently weak, then the action of the particle on the field may be neglected (consider, e.g., small mass in a gravitational field, or an electron in a macroscopic electric field). In that case, supposing the values for the fields are known, we need only consider the particle and interaction terms in the problem:

$$S = -mc \int \mathrm{d}s + \frac{1}{c} \int \mathrm{d}^4x \, \mathcal{L}_{\mathrm{int}}(\phi, x). \tag{2.110}$$

We'll consider such a case in detail when discussing interactions of particles with electromagnetic fields in Sec. 2.7.

For now, let's consider briefly the simple example of an interaction between a particle and a scalar field. Since the action must be a scalar, the simplest interaction term we can write involving both particle and field terms is

$$S_{\mathrm{int}} = -\int f(\phi(x)) \, \mathrm{d}s, \tag{2.111}$$

where $f(\phi)$ is any scalar function of $\phi$. Note that the values of $\phi(x^\mu)$ are evaluated at the positions (along the world line) of the particle, and the arbitrary negative sign is by convention. In analogy with the pure particle case (2.42), the total Lagrangian is then (with $c = 1$)

$$L_{\mathrm{part+int}} = -\left[m + f(\phi)\right] \sqrt{1 - v^2}. \tag{2.112}$$

By comparison with Eq. (2.42), it's clear the effect of this term on the equation of motion of the particle is $m \to m + f(\phi)$. This is a (very crude) analogy for how the Higg's field (which is a scalar field, albeit a quantum one) may affect the masses of particles.

### 2.6.1 Interacting fields: the Yukawa potential

Above, we considered how we may form theories in which particles and fields interact. As a final note in this section, we will briefly consider how interactions between different fields may appear. Consider a very similar case to the previous scalar field example (Sec. 2.4.3), except now with two fields:

$$\mathcal{L} = \mathcal{L}_{\mathrm{free}}(\phi) + \mathcal{L}_{\mathrm{free}}(\chi) - \lambda \phi^2 \chi,$$

where $\mathcal{L}_{\text{free}}$ is the "free" Klein-Gordon Lagrangian (2.67) for the $\phi$ and $\chi$ fields (with $\mu_\phi$ and $\mu_\chi$, respectively), and the term proportional to $\lambda$ is known as an *interaction term* (whose form was chosen arbitrarily as an example). Without the interaction term, this would simply describe two independent scalar fields that evolve entirely separately. The interaction term, however, contains *both* fields, and leads to a coupling between the two equations of motions. The equations of motion can be found without too much difficulty:

$$(\partial^2 + \mu_\phi^2)\phi = -2\lambda\phi\chi, \quad \text{and} \quad (\partial^2 + \mu_\chi^2)\chi = -\lambda\phi^2. \tag{2.113}$$

These are quite similar to those we found before [Eq. (2.69)], though now they are inhomogeneous equations, containing source terms.

We shall (approximately) solve the equations for the $\chi$ field. Consider the case that there is a point-source of $\phi$ field at the origin. We can write this as

$$\phi^2(x) = \frac{A}{\mu_\phi}\delta(\boldsymbol{x}), \tag{2.114}$$

where the powers and factors are chosen to ensure the correct units. Here, $A$ is a dimensionless amplitude, which we will simply take to be 1, since it does not impact the rough analysis. Further, for simplicity, assume that there is no time-dependence in the $\chi$ field (i.e., assume the fields are evolving slowly enough for us to discard any time variations). Then, the equations of motion become

$$(\nabla^2 - \mu_\chi^2)\chi(\boldsymbol{x}) = \frac{\lambda}{\mu_\phi}\delta(\boldsymbol{x}), \tag{2.115}$$

which you might recognise as the equation for a Green's function. The simplest way to solve this is to write $\chi$ as a Fourier transform, and use the Fourier representation of the Dirac delta function (see Appendix 2.11.4):

$$\chi(\boldsymbol{x}) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\,\tilde{\chi}(\boldsymbol{k}), \quad \text{and} \quad \delta(\boldsymbol{x}) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{2.116}$$

Plugging into (2.115), we can immediately read off:

$$\left(k^2 + \mu_\chi^2\right)\tilde{\chi}(\boldsymbol{k}) = -\frac{\lambda}{\mu_\phi}. \tag{2.117}$$

Finally, we take the inverse Fourier transform to solve for $\chi(x)$:

$$\chi(\boldsymbol{x}) = -\frac{\lambda}{\mu_\phi}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\frac{e^{i\boldsymbol{k}\cdot\boldsymbol{x}}}{k^2 + \mu_\chi^2} = -\frac{\lambda}{\mu_\phi}\frac{e^{-\mu_\chi|\boldsymbol{x}|}}{4\pi|\boldsymbol{x}|}, \tag{2.118}$$

where we used a standard integral identity for the final step.[11]

[11] The integral can be performed, e.g., using complex contour integration; since it is not crucial for our current discussions, we will just take it as a given for now.

If we write the Hamiltonian as $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\text{int}}$, where $\mathcal{H}_0$ is the Hamiltonian density for $\lambda = 0$, which we calculated in Eq. (2.98), and $\mathcal{H}_{\text{int}}$ is the interaction part, we can see that

$$\mathcal{H}_{\text{int}} = \lambda\phi^2\chi,$$

where, in our example, $\chi$ is sourced from the lump of $\phi$ at the origin. Consider a second point source of $\phi$, located at point $\boldsymbol{r}$. Combining with the expressions for $\phi^2$ (2.114) and $\chi$ (2.118), we see this leads to a position-dependant interaction energy:

$$V_{\text{int}} = \int \mathrm{d}^3x' \, \mathcal{H}_{\text{int}} = -\frac{\lambda^2}{\mu_\phi^2}\frac{e^{-\mu_\chi r}}{4\pi r}, \tag{2.119}$$

where $r$ is the distance between the two point-sources ("lumps") of $\phi$. This is known as the Yukawa potential[12], and implies an attractive force between two lumps of $\phi$. In this sense, it can be interpreted as the $\chi$ field mediating the interaction between two densities of the $\phi$ field. This will of course gain a deeper meaning when we consider quantum field theory.

## 2.7 Electromagnetic fields and interactions

You are likely aware that electric and magnetic fields are described by (three-) vector, rather than scalar, quantities. As it turns out, electromagnetic fields and their interactions may be described by a (four-) vector field, denoted $A^\mu$, which is called the *electromagnetic vector potential*. We will start by considering the motion of a particle within the influence of such a vector field. At first, we will consider again the weakly interacting case, where the influence of the particle on the field may be neglected. Therefore, on top of the particle terms, we need only consider interactions terms in the action. (Of course, the full theory will describe the evolution of the fields, which will be impacted by the particles; we will return to this in the next section.) We note here that we employ the *Heaviside-Lorentz* units for electromagnetic quantities, in which the factors $\varepsilon_0$ and $\mu_0$ do not appear explicitly in the equations; see Appendix 2.11.2 for discussion and definitions.

To determine how we should change the action of the particle to account for its interaction with the field, we can ask what scalars can be made from a vector field, $A^\mu$? The simplest such term we could add to the action is proportional to

$$A_\mu \mathrm{d}x^\mu.$$

This is of course not the only option; we could also have any term of the form $f(A_\mu)\mathrm{d}s$, where $f$ is a general scalar function. Ultimately, we must be guided by experiment. It turns out that the simplest term above is sufficient, as we shall see; we will return to a discussion of further possibilities in time.

[12] H. Yukawa, *On the interaction of elementary particles*, Prog. Theor. Phys. **1**, 1 (1955).

The contribution to the action is therefore written

$$S_{\text{int}} = -\frac{q}{c}\int A_\mu \mathrm{d}x^\mu = -q\int\left(A_0 - \frac{1}{c}\boldsymbol{A}\cdot\boldsymbol{v}\right)\mathrm{d}t, \tag{2.120}$$

where $q$ is an arbitrary scalar constant (negative sign and the factor of $c$ are introduced for dimensional convenience), and $x$ refers to the particle coordinates. To be dimensionally correct, the combination $qA_\mu$ must have energy units. The constant $q$ is called the *charge* of the particle. Notice that as $q \to 0$, the interaction term will have no impact on the equations of motion; we say that the charge quantifies the strength of the interaction between the particle and the field.

Combining with Eq. (2.42), we may write the total Lagrangian as

$$L = -mc^2/\gamma - qA_0 + \frac{q}{c}\boldsymbol{A}\cdot\boldsymbol{v}, \tag{2.121}$$

where we have $A^\mu = (A_0, \boldsymbol{A})$; $A_0 \equiv \Phi$ may be called the *electric scalar potential*, and $\boldsymbol{A}$ the *magnetic vector potential*; we shall see why shortly. Directly from the Lagrangian, we may find the expression for the generalised momentum

$$\boldsymbol{P} = \frac{\partial L}{\partial \boldsymbol{v}} = m\boldsymbol{v}\gamma + \frac{q}{c}\boldsymbol{A} = \boldsymbol{p} + \frac{q}{c}\boldsymbol{A}, \tag{2.122}$$

using $\boldsymbol{p} = m\boldsymbol{v}\gamma$ from Eq. (2.44).

### 2.7.1 Lorentz force law

We shall now find the equations of motion for the system. From Mechanics [Eq. (1.19)], the equations of motion can be expressed $\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} = \frac{\partial L}{\partial \boldsymbol{x}}$. First, we find

$$\begin{aligned}\frac{\partial L}{\partial \boldsymbol{x}} &= -\frac{q}{c}\left[c\nabla\Phi - \nabla(\boldsymbol{A}\cdot\boldsymbol{v})\right] \\ &= -\frac{q}{c}\left[c\nabla\Phi - (\boldsymbol{v}\cdot\nabla)\boldsymbol{A} - \boldsymbol{v}\times(\nabla\times\boldsymbol{A})\right],\end{aligned} \tag{2.123}$$

where we made use of the vector identity for the gradient of a dot product[13] (note that the spatial derivatives are performed at fixed velocity).

For $\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t}$, use Eq. (2.122), and note that $\boldsymbol{A}$ is evaluated at the particle position $\boldsymbol{x}(t)$, and may also have an intrinsic time dependence. Therefore,

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = \frac{\partial \boldsymbol{A}}{\partial t} + \frac{\partial \boldsymbol{A}}{\partial x_i}\frac{\partial x_i}{\partial t} = \frac{\partial \boldsymbol{A}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{A}, \tag{2.124}$$

and so the equations of motion are

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -\frac{q}{c}\left[c\nabla\Phi + \frac{\partial \boldsymbol{A}}{\partial t} - \boldsymbol{v}\times(\nabla\times\boldsymbol{A})\right]. \tag{2.125}$$

[13] $\nabla(\boldsymbol{a}\cdot\boldsymbol{b}) = (\boldsymbol{a}\cdot\nabla)\boldsymbol{b} + (\boldsymbol{b}\cdot\nabla)\boldsymbol{a} + \boldsymbol{a}\times(\nabla\times\boldsymbol{b}) + \boldsymbol{a}\times(\nabla\times\boldsymbol{a})$; see Appendix 2.11.3.

Notice that the equations of motion depend only on *derivatives* of the potential $A^\mu$. We are therefore led to define the fields

$$\boldsymbol{E} \equiv -\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} - \nabla\Phi \quad \text{and} \quad \boldsymbol{B} \equiv \nabla \times \boldsymbol{A}, \tag{2.126}$$

where $\boldsymbol{E}$ is called the *electric field strength*, and $\boldsymbol{B}$ the *magnetic field strength*. With these definitions, the equations of motion may be written as

$$\boldsymbol{F} = q\boldsymbol{E} + \frac{q}{c}\boldsymbol{v} \times \boldsymbol{B}, \tag{2.127}$$

which is known as the *Lorentz force law*.

The equations for the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields (2.126) are linked through the potential, $A_\mu$. It is also possible to form a set of equations for the fields, independently from the potential. Taking the curl of $\boldsymbol{E}$, and the divergence of $\boldsymbol{B}$, we find

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \quad \text{and} \quad \nabla \cdot \boldsymbol{B} = 0, \tag{2.128}$$

where we made use of the fact that the curl of a divergence, and the divergence of a curl, is zero (see Appendix 2.11.3). You may recognise these as the homogeneous pair of Maxwell's equations (Maxwell–Faraday equation, and Gauss's law for magnetism), which follow directly from the definition of the fields. The full equations governing the dynamics of the electromagnetic fields (i.e., the second inhomogeneous pair of Maxwell's equations) will require a consideration of the action for the fields themselves; we will return to this question in the coming sections.

***Problem* 2.7.1:** Show directly that $(\boldsymbol{v}\cdot\nabla)\boldsymbol{A} + v \times (\nabla \times \boldsymbol{A}) = \nabla(\boldsymbol{A}\cdot\boldsymbol{v})$, where $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{x})$, and $\boldsymbol{v}$ is a constant.

### 2.7.2 Gauge invariance

It's important to notice that the equations of motion (the Lorentz force law) depend only on the electromagnetic fields, $\boldsymbol{E}$ and $\boldsymbol{B}$, not on the potentials themselves. Since the fields (2.126) depend on *derivatives* of the potential, it is clear that the equations of motion are invariant under the global transformation

$$A_\mu \to A'_\mu = A_\mu + \text{const.}$$

In fact, it turns out the symmetry is much deeper, and the equations of motion remain invariant under the *local* transformation:

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu\theta(x^\mu), \tag{2.129}$$

where $\theta = \theta(x^\mu)$ is *any* (continuous, differentiable) function of spacetime coordinates. In terms of the scalar and (three) vector potentials, this is

$$\Phi \to \Phi + \frac{1}{c}\frac{\partial\theta}{\partial t} \quad \text{and} \quad \boldsymbol{A} \to \boldsymbol{A} - \nabla\theta.$$

The symmetry under this type of transformation is known as *gauge invariance*.

While it's fairly straight forward to directly show that the fields in Eq. (2.126) remain invariant under this transformation, it is possible to show the equations obey this symmetry at a deeper level. The change to the term in the action (2.120) under the gauge transformation (2.129) is

$$A_\mu \mathrm{d}x^\mu \to A_\mu \mathrm{d}x^\mu + (\partial_\mu \theta)\mathrm{d}x^\mu \tag{2.130}$$

which is a total differential, and thus changes the action by at most a constant. It therefore cannot impact the equations of motion. That the electromagnetic interactions are observed to obey this gauge symmetry is the reason the $\int \mathrm{d}s A_\mu A^\mu$ term, which is not gauge invariant, may not appear in the action. As we shall see in our discussion of quantum field theory, this has a deep connection with the fact that the photon is massless.

While any gauge choice is valid, some choices may be more convenient for certain problems. A common example is the *Lorenz* gauge, for which

$$\partial_\mu A^\mu = 0. \tag{2.131}$$

This choice is widely used, and is convenient due to the explicit Lorentz invariance. The Lorenz condition is known as a *partial* gauge condition, since it does not completely determine the gauge. Any further transform of the form $A^\mu \to A^\mu + \partial^\mu \phi$ where $\partial_\mu \partial^\mu \phi = 0$ will change the gauge while leaving the Lorenz condition intact.

A particularly common choice is the *Coulomb gauge*, also known as the transverse gauge, which is explicitly not Lorentz invariant. In the Coulomb gauge we assert that the spatial derivatives are zero:

$$\nabla \cdot \boldsymbol{A} = 0. \tag{2.132}$$

This choice is particularly common in practical and non-relativistic calculations. In general, it is always possible to choose a gauge where any one component of $A_\mu$ is zero; it is often convenient to set $\Phi = 0$ (the radiation gauge).

***Problem* 2.7.2:** Find the condition on the gauge function, $\theta$, in terms of a general potential $A^\mu$, such that ${A'}_\mu = A_\mu + \partial_\mu \theta$ satisfies (i) the Lorenz, and (ii) the Coulomb gauge conditions.

### 2.7.3 Electromagnetic field tensor

In the above sections, we derived the Lorentz force law, which was the equation of motion for a charged particle in a background electromagnetic field. Despite beginning from a Lorentz invariant action, we resulted in an equation in terms of regular three-vectors. Of course, it should be possible to write the equations in four-vector (i.e., covariant) form.

Consider the action for a particle in an electromagnetic field [e.g., Eq. (2.121)]:

$$S = \int \left( -mc\,\mathrm{d}s - \frac{q}{c} A_\mu \mathrm{d}x^\mu \right), \tag{2.133}$$

where we maintain the weak interaction approximation as discussed above. Noting that $\mathrm{d}s = \sqrt{\mathrm{d}x_\mu \mathrm{d}x^\mu}$, we can write the variation in $\mathrm{d}s$ as

$$\delta(\mathrm{d}s) = \frac{\mathrm{d}x_\mu}{\mathrm{d}s}\delta(\mathrm{d}x^\mu) = \frac{1}{c} u_\mu \mathrm{d}(\delta x^\mu), \tag{2.134}$$

where $u$ is the four-velocity (2.35). At the same time, we have

$$\delta(A_\mu \mathrm{d}x^\mu) = \delta(A_\mu)\mathrm{d}x^\mu + A_\mu \mathrm{d}(\delta x^\mu),$$

where we again used $\delta(\mathrm{d}x^\mu) = \mathrm{d}(\delta x^\mu)$. Therefore, the variation in the action is

$$\delta S = \int \Big( - m\, u_\mu \mathrm{d}(\delta x^\mu) - \frac{q}{c} \left[ \partial_\nu A_\mu \delta x^\nu \mathrm{d}x^\mu + A_\mu \mathrm{d}(\delta x^\mu) \right] \Big), \tag{2.135}$$

where we made use of

$$\delta(A_\mu) = \frac{\partial A_\mu}{\partial x^\nu}\, \delta x^\nu = \partial_\nu A_\mu\, \delta x^\nu.$$

Integrating the first and third terms in (2.135) by parts and discarding the boundary terms, this becomes

$$\begin{aligned} \delta S &= \int \left( m\, \mathrm{d}u_\mu\, \delta x^\mu - \frac{q}{c} \left[ \partial_\nu A_\mu \delta x^\nu \mathrm{d}x^\mu - \partial_\nu A_\mu \mathrm{d}x^\nu\, \delta x^\mu \right] \right) \\ &= \int \left( m\, \mathrm{d}u_\mu - \frac{q}{c} \left[ \partial_\mu A_\nu \mathrm{d}x^\nu - \partial_\nu A_\mu \mathrm{d}x^\nu \right] \right) \delta x^\mu = 0, \end{aligned} \tag{2.136}$$

where we swapped the dummy indices $\mu \leftrightarrow \nu$ in the second term. Finally, since the variation must be zero for arbitrary $\delta x^\mu$, the equations of motion are found

$$m\, \mathrm{d}u_\mu - \frac{q}{c} \left[ \partial_\mu A_\nu - \partial_\nu A_\mu \right] \mathrm{d}x^\nu = 0. \tag{2.137}$$

Just as above, we find the equations of motion depend only on the derivatives of the $A$ field. We thus now introduce the *Electromagnetic field tensor* (sometimes called the Faraday tensor, or the Maxwell tensor), defined

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu, \tag{2.138}$$

which is an anti-symmetric, rank two tensor. From the definitions in Eq. (2.126), the components can be found to be[14]

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{2.139}$$

[14] Note the potentially confusing notation: $a_x = a^1$ is the $x$-component of $\boldsymbol{a}$.

It is important to notice that, with this definition, Lorentz transformations can be seen to directly mix the components of electric and magnetic fields.

With this, the Lorentz force law (2.137) may be written in covariant form:

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}s} = \frac{q}{c^2} F^{\mu\nu} u_\nu. \tag{2.140}$$

It can be convenient to notice that the components of the electromagnetic fields can be expressed as

$$E^i = F^{i0}, \quad \text{and} \quad B^i = -\frac{1}{2}\varepsilon^{ijk} F_{jk} \quad \text{or} \quad F^{ij} = -\varepsilon^{ijk} B^k = \varepsilon^{ijk} B_k, \tag{2.141}$$

where $\varepsilon^{ijk}$ is the totally anti-symmetric Levi-Civita symbol. Landau[15] introduces the compact notation $F^{\mu\nu} \equiv (-\boldsymbol{E}, \boldsymbol{B})$, with which $F_{\mu\nu} = (\boldsymbol{E}, \boldsymbol{B})$.

We may now express the homogeneous pair of Maxwell's equations (2.128) in covariant form. To do this, consider derivatives of the tensor terms of the form

$$\partial_\lambda F_{\mu\nu} = \partial_\lambda(\partial_\mu A_\nu - \partial_\nu A_\mu) = \partial_\mu \partial_\lambda A_\nu - \partial_\nu \partial_\lambda A_\mu, \tag{2.142}$$

where we commuted the derivatives. Noting that we may write $\partial_\lambda A_\nu = F_{\lambda\mu} + \partial_\mu A_\lambda$, we see that we can form combinations that sum to zero:

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} \equiv \partial_{[\lambda} F_{\mu\nu]} = 0. \tag{2.143}$$

This is called the *Bianchi identity*. Evaluating the Bianchi identity with $\{\lambda, \mu, \nu\}$ as $\{0, 1, 2\}$ and $\{1, 2, 3\}$, respectively, we eventually arrive at

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \quad \text{and} \quad \nabla \cdot \boldsymbol{B} = 0, \tag{2.144}$$

which you may recognise as the Maxwell–Faraday equation and Gauss's law for magnetism. Therefore, Eq. (2.143) is the covariant form for the homogeneous pair of Maxwell's equations. To find the other pair of Maxwell's equations, we will have to find the equations of motion, which we shall come to soon.

It is convenient to also define the *dual tensor* $\widetilde{F}^{\mu\nu}$ (sometimes denoted $\mathcal{F}$):

$$\widetilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}, \tag{2.145}$$

where $\epsilon^{\mu\nu\rho\sigma}$ is the totally anti-symmetric rank four tensor.[16] Using the Landau notation, $\widetilde{F}^{\mu\nu} = (-\boldsymbol{B}, -\boldsymbol{E})$. With the dual tensor, the homogeneous Maxwell equations can be expressed as

$$\partial_\mu \widetilde{F}^{\mu\nu} = 0. \tag{2.146}$$

[15] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (1971).

[16] $\epsilon^{\mu\nu\rho\sigma} = 1$ under any even permutation of $\epsilon^{0123}$, $= -1$ under any odd permutation, $= 0$ if any two indices are equal, and $\epsilon^{\mu\nu\rho\sigma} = -\epsilon_{\mu\nu\rho\sigma}$.

***Problem* 2.7.3:** Explicitly determine the components $F_{\mu\nu}$ of the Faraday tensor using its definition in Eq. (2.138) and the definition of the fields (2.126).

***Problem* 2.7.4:** Show explicitly that the Lorentz force law (2.127) follows directly from the spatial terms in Eq. (2.140).

***Problem* 2.7.5:** What is the physical meaning of the temporal term in Eq. (2.140)?
*Answer* (2.7.5): It is the work equation: $\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} = q\boldsymbol{E}\cdot\boldsymbol{v}$.

***Problem* 2.7.6:** Show explicitly from its definition (2.138), that the field tensor $F^{\mu\nu}$ is invariant under the gauge transformation (2.129).

***Problem* 2.7.7:** Find the scalar $F_{\mu\nu}F^{\mu\nu}$.

*Solution* (2.7.7): With matrix notation $F = F^{\mu\nu}$, and noting that $F_{\mu\nu} = \eta_{\mu\alpha}\eta_{\nu\beta}F^{\alpha\beta} = \eta F\eta$, we have $F_{\mu\nu}F^{\mu\nu} = \mathrm{tr}[(\eta F\eta)^{\mathrm{T}}F] = -\mathrm{tr}[(\eta F\eta)F] = 2(\boldsymbol{B}^2 - \boldsymbol{E}^2)$, where $\mathrm{tr}(M) = M^{\mu\mu}$ is the trace.

***Problem* 2.7.8:** Explicitly work out equations (2.144) from the Bianchi identity (2.143), and from the dual tensor equation (2.146).

### 2.7.4 Current density

Before we continue to full field theory description of electromagnetic fields, we require a way to deal properly with the concept of particle point charges in the context of a continuous field theory. We can define the *charge density*, $\rho$, for a collection of point particles as

$$\rho \equiv \sum_i q_i\,\delta(\boldsymbol{x} - \boldsymbol{x}_i), \tag{2.147}$$

where $\delta(\boldsymbol{x})$ is the Dirac delta function (see Appendix 2.11.4). The total charge in a volume, $Q \equiv \sum_i q_i$, is clearly the volume integral of the charge density:

$$Q = \int \mathrm{d}V\,\rho = \sum_i q_i. \tag{2.148}$$

This is not surprising, of course, but shows our definitions are consistent.

From the above, we may identify $\mathrm{d}q \equiv \rho\,\mathrm{d}V$ (the charge contained in an infinitesimal volume). Note that $q$ (and thus $Q$) is a scalar by its definition [see Eq. (2.120)], and therefore so is $\mathrm{d}q$. On the other hand, neither $\rho$ nor $\mathrm{d}V$ is a scalar; only the combination $\rho\mathrm{d}V$ is. If we multiply by the vector $\mathrm{d}x^\mu$, we have

$$\mathrm{d}q\,\mathrm{d}x^\mu = \rho\,\mathrm{d}^4x\,\frac{\mathrm{d}x^\mu}{c\,\mathrm{d}t}. \tag{2.149}$$

Note that $\mathrm{d}q\mathrm{d}x^\mu$ is a four vector, and $\mathrm{d}^4x = c\mathrm{d}t\,\mathrm{d}V$ is a scalar. Therefore $\rho\frac{\mathrm{d}x^\mu}{\mathrm{d}t}$ is a four vector, so we define the *current density* four vector

$$j^\mu \equiv \rho\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = (c\rho,\,\boldsymbol{j}), \tag{2.150}$$

which has components $j^\mu = (j^0, \boldsymbol{j})$. The definition as a current density is justified: $j^0$ is (up to a constant of $c$) simply the charge density, $j^0 = c\rho$, and the spatial terms are $\boldsymbol{j} = \rho\boldsymbol{v}$. It is sometimes convenient to factor out the unit charge $q$ from the definition of the current density, and define the *vector current* $\mathcal{J}^\mu$:

$$\mathcal{J}^\mu = \frac{1}{q} j^\mu. \tag{2.151}$$

With these definitions, we can re-express the interaction contribution to the action (2.120) in terms of the current density by writing $q = \rho\,\mathrm{d}V$:

$$\begin{aligned} S_{\text{int}} &= -\frac{1}{c}\int \rho\,\mathrm{d}V A_\mu \mathrm{d}x^\mu = -\frac{1}{c}\int \rho \frac{\mathrm{d}x^\mu}{\mathrm{d}t} A_\mu\,\mathrm{d}V\mathrm{d}t \\ &= -\frac{1}{c^2}\int j^\mu A_\mu\,\mathrm{d}^4x. \end{aligned} \tag{2.152}$$

We therefore identify the interaction contribution to the Lagrangian density as:

$$\mathcal{L}_{\text{int}} = -\frac{1}{c} j^\mu A_\mu, \quad \text{or} \quad \mathcal{L}_{\text{int}} = -\frac{q}{c}\mathcal{J}^\mu A_\mu, \tag{2.153}$$

which may be used for general charge distributions.

We shall now consider the law of the conservation of charge, by examining the four-divergence of the current density, $\partial_\mu j^\mu$. Integrating this divergence over some finite volume, $V$, we have

$$\begin{aligned} \int \partial_\mu j^\mu\,\mathrm{d}V &= \frac{\partial}{\partial t} Q_V + \int \nabla\cdot(\rho\boldsymbol{v})\,\mathrm{d}V \\ &= \frac{\partial}{\partial t} Q_V + \oint_S \rho\,(\boldsymbol{v}\cdot\hat{\boldsymbol{n}})\,\mathrm{d}S, \end{aligned} \tag{2.154}$$

where $Q_V$ is the total charge within the volume, and we used Gauss' theorem for the second term [see Appendix 2.11.3], with $S$ the closed surface bounding the volume and $\hat{\boldsymbol{n}}$ the unit vector normal to the surface. The first term on the right-hand-side of Eq. (2.154) is the rate of change of the total charge contained in the volume. The second term is the rate at which charge passes through the surface of the volume (see Sec. 2.5.1). If we assert that the charge must be a locally conserved quantity, then the these terms must sum to zero, leading to

$$\partial_\mu j^\mu = 0, \tag{2.155}$$

which is called the *continuity* equation. While, for now, we have simply asserted this fact based on physical arguments, we shall soon see that this law is intrinsically linked to the principle of gauge invariance.

***Problem* 2.7.9:** In the rest frame of a charge distribution ($K'$), we have $J^\mu = (\rho_0/c, \boldsymbol{0})$. Observed from another frame ($K$), the charge is observed to move along the $x$-axis with velocity $v$. (i) What is the current as observed in the $K$ frame? (ii) As we saw above, the total charge $Q = \int \mathrm{d}V j^0$ is a scalar, while the density, $\rho = j^0/c$ is not. Write the current in terms of the density $\rho$ in the current frame.

*Answer* (2.7.9): i: $(c\rho_0\gamma, v\rho_0\gamma, 0, 0)$, ii: $(c\rho, v\rho, 0, 0)$

## 2.8 Electrodynamics: Maxwell's equations

We now seek the set of equations that govern the evolution of the $A_\mu$ vector potential. To construct the required action, noting the principles we discussed above, we may consider scalar functions of the fields $A_\mu$ and their first derivatives $\partial_\nu A_\mu$. Further, since from experience, we expect the the electromagnetic fields to obey the superposition principle, we restrict our consideration to terms that are quadratic in the $A_\mu$ fields. Finally, since the equations of motion we found before were gauge invariant, we shall consider only terms which are themselves gauge invariant.

The terms proportional to $A_\mu A^\mu$ or $\partial_\mu A^\mu$ are not gauge invariant. Under the gauge transformation $A_\mu \to A_\mu + \partial_\mu \theta$ [Eq. (2.129)], they become

$$\begin{aligned} &A_\mu A^\mu \to A_\mu A^\mu + 2A_\mu \partial^\mu \theta + \partial_\mu \theta \partial^\mu \theta, \quad \text{and} \\ &\partial_\mu A_\nu \to \partial_\mu A_\nu + \partial_\mu \partial_\nu \theta. \end{aligned} \tag{2.156}$$

The anti-symmetric combination $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ (2.138), however, is gauge invariant (see Problem 2.7.6). Therefore, the scalar $F^{\mu\nu}F_{\mu\nu}$ (often written as $F^2$ for brevity) is the only scalar quantity that is both gauge invariant, and second-order in the field.

Therefore, the term in the Lagrangian involving only the fields may be

$$\mathcal{L}_\mathrm{F} = \frac{-1}{4} F^{\mu\nu}F_{\mu\nu} = \frac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{B}^2). \tag{2.157}$$

The factor $1/4$ is arbitrary, and corresponds to a choice of units. Combining with the interaction term from Eq. (2.153), the total Lagrangian including all field terms is

$$\mathcal{L} = -\frac{1}{c} j^\mu A_\mu - \frac{1}{4} F^{\mu\nu}F_{\mu\nu}. \tag{2.158}$$

This is often termed the *Maxwell Lagrangian*.

To determine the field equations, we apply the Euler-Lagrange condition (2.57) with $\phi = A_\nu$. We have (see Problem 2.8.1)

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu}, \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial A_\nu} = -\frac{1}{c} j^\nu, \tag{2.159}$$

so the field equations are

$$\partial_\mu F^{\mu\nu} = \frac{1}{c} j^\nu. \tag{2.160}$$

To write these in three-dimensional form, we evaluate (2.160) with $\nu = 0$ and $\nu = \{1, 2, 3\}$, respectively, resulting in:

$$\nabla \cdot \boldsymbol{E} = \rho, \quad \text{and} \qquad \nabla \times \boldsymbol{B} = \frac{1}{c}\left(\boldsymbol{j} + \frac{\partial \boldsymbol{E}}{\partial t}\right). \tag{2.161}$$

These are of course Gauss's law and Ampère's law; we thus recognise Eq. (2.160) as the inhomogeneous Maxwell equations. To summarise, in the covariant notation, the Lorentz force law and Maxwell's equations can be summarised with

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}s} = \frac{q}{c^2} F^{\mu\nu} u_\nu, \qquad \partial_\mu \widetilde{F}^{\mu\nu} = 0, \quad \text{and} \quad \partial_\mu F^{\mu\nu} = \frac{1}{c} j^\nu. \tag{2.162}$$

***Problem* 2.8.1:** Derive the derivative $\frac{\partial F^2}{\partial(\partial_\mu A_\nu)} = 4F^{\mu\nu}$, required in Eq. (2.159).

*Solution* (2.8.1): We can use chain rule as usual, treating $F = F(\partial_\mu A_\nu)$, where $\partial_\mu A_\nu$ is the independent variable. For brevity, write $\partial_\mu A_\nu \equiv t_{\mu\nu}$. Being careful with index labels, we find:

$$\frac{\partial(F^{\alpha\beta}F_{\alpha\beta})}{\partial t_{\mu\nu}} = 2F^{\alpha\beta}\frac{\partial F_{\alpha\beta}}{\partial t_{\mu\nu}}.$$

Noting that $\partial t_{\mu\nu}/\partial t_{\sigma\rho} = \delta^\sigma_\mu \delta^\rho_\nu$ is zero unless $\mu = \sigma$ and $\nu = \rho$, we have

$$\frac{\partial F_{\alpha\beta}}{\partial t_{\mu\nu}} = \frac{\partial t_{\alpha\beta}}{\partial t_{\mu\nu}} - \frac{\partial t_{\beta\alpha}}{\partial t_{\mu\nu}} = \delta^\mu_\alpha \delta^\nu_\beta - \delta^\mu_\beta \delta^\nu_\alpha.$$

Putting together, noting that $F^{\beta\alpha} = -F^{\alpha\beta}$, we see $\frac{\partial F^2}{\partial(\partial_\mu A_\nu)} = 4F^{\mu\nu}$.

***Problem* 2.8.2:** Derive the equations of motion (2.160) by directly varying the action. This is often easier than applying the general Euler-Lagrange equations.

*Solution* (2.8.2): We have $\delta(F^{\mu\nu}F_{\mu\nu}) = 2F^{\mu\nu}\delta(F_{\mu\nu}) = 4F^{\mu\nu}\delta(\partial_\mu A_\nu) = 4F^{\mu\nu}\partial_\mu(\delta A_\nu)$, where we used anti-symmetry of $F$ in penultimate step. The rest follows as per Eq. (2.56).

### 2.8.1 Gauge invariance and charge conservation

The gauge invariance and the charge conservation in electrodynamics are intrinsically linked. In the full quantum field theory, it turns out that the conserved Noether current (2.77) associated with the gauge invariance (2.129) is just the electric charge current (2.150). Without a concept of the gauge symmetry for the matter part of the problem (i.e., the point charges or charge density), this doesn't quite work. Even without the full quantum theory, however, we can still show that these concepts are closely connected.

Under the gauge transformation Eq. (2.129), the variation of the Maxwell Lagrangian (2.158) is

$$\delta\mathcal{L} = \frac{-1}{c} j^\mu \partial_\mu \theta, \tag{2.163}$$

which follows simply due to the explicit gauge invariance of the $F^2$ term (2.157). The corresponding contribution to the action can be written as

$$\delta S = \frac{-1}{c^2}\int j^\mu\, \partial_\mu\theta\, \mathrm{d}^4x = \frac{1}{c^2}\int (\partial_\mu j^\mu)\, \theta\, \mathrm{d}^4x, \tag{2.164}$$

where we used integration by parts, and assumed the current terms go to zero at infinity. Since, by assertion, the gauge function $\theta$ may be *any* function of time and coordinates, it follows that this term is zero only if $\partial_\mu j^\mu = 0$. Therefore, the gauge invariance implies the continuity equation (2.155), which, as discussed above, is the condition for local charge conservation.

### 2.8.2 Energy of the electromagnetic field

From Maxwell's equations (2.144) and (2.161), we can form the expressions

$$\boldsymbol{E}\cdot\nabla\times\boldsymbol{B}=\frac{1}{c}\boldsymbol{E}\cdot\left(\frac{\partial\boldsymbol{E}}{\partial t}+\boldsymbol{j}\right),\quad\text{and}\quad\boldsymbol{B}\cdot\nabla\times\boldsymbol{E}=-\frac{1}{c}\boldsymbol{B}\cdot\frac{\partial\boldsymbol{B}}{\partial t}. \tag{2.165}$$

Subtracting these, and making use of the standard vector identity,[17] we find

$$\frac{1}{2}\frac{\partial}{\partial t}\left(E^2+B^2\right)=-\boldsymbol{E}\cdot\boldsymbol{j}-\nabla\cdot[c\boldsymbol{E}\times\boldsymbol{B}]. \tag{2.166}$$

It is customary to define the term in brackets as the *Poynting vector*:

$$\boldsymbol{S}\equiv c\boldsymbol{E}\times\boldsymbol{B}. \tag{2.167}$$

To elucidate the physical meaning of this equation (2.166), we shall integrate it over a volume, $V$. Recognising the current density as being that from a collection of point particles [Eq. (2.147)], $\boldsymbol{j}(x)=\sum q_i\boldsymbol{v}_i\delta(x-x_i)$, we have

$$\frac{\partial}{\partial t}\int\left(\frac{E^2+B^2}{2}\right)\mathrm{d}V=-\sum_i q_i\boldsymbol{E}\cdot\boldsymbol{v}_i-\int\nabla\cdot\boldsymbol{S}\,\mathrm{d}V. \tag{2.168}$$

From Problem 2.7.5, the rate of change in kinetic energy of a particle of charge $q$ in an electromagnetic field is given

$$\frac{\mathrm{d}\mathcal{E}_{\text{part.}}}{\mathrm{d}t}=q\boldsymbol{E}\cdot\boldsymbol{v}.$$

Therefore, we may express Eq. (2.168) as

$$\frac{\partial}{\partial t}\left(\int\mathcal{U}\,\mathrm{d}V+\sum\mathcal{E}_{\text{part.}}\right)=-\oint\boldsymbol{S}\cdot\mathrm{d}\boldsymbol{\Omega}, \tag{2.169}$$

where we used Gauss' theorem [Eq. (2.248) in the appendix] for the term on the right-hand-side, with $\mathrm{d}\boldsymbol{\Omega}$ being the surface element. If the volume is taken to include all space, and the fields are assumed to go to zero at infinity, then the surface integral is zero, and it's clear that

$$\mathcal{U}\equiv\frac{E^2+B^2}{2} \tag{2.170}$$

can be interpreted as the *energy density* of the electromagnetic fields, assuring the energy conservation. If the volume is taken to be finite, then the term on the right-hand-side may be non-zero. In this case, it clear that the Poynting vector, $\boldsymbol{S}$, should be interpreted as the energy density flux through the surface of the volume (energy transfer per unit area, per unit time).

[17] $\nabla\cdot(\boldsymbol{a}\times\boldsymbol{b})=\boldsymbol{b}\cdot\nabla\times\boldsymbol{a}-\boldsymbol{a}\cdot\nabla\times\boldsymbol{b}$; see Appendix 2.11.3.

### 2.8.3 Electromagnetic stress-energy tensor

We will now calculate the *Maxwell stress-energy tensor*, in the source free ($j^\mu = 0$) case. From Eq. (2.87),

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial^\nu \phi - \eta^{\mu\nu} \mathcal{L},$$

where the Maxwell Lagrangian is given by Eq. (2.158). Using the derivatives we already computed (2.159), with the field being the components of the vector potential (i.e., $\phi = A_\lambda$) the expression is readily found

$$T^{\mu\nu} = -F^{\mu\lambda} \partial^\nu A_\lambda + \frac{1}{4} \eta^{\mu\nu} F^{\rho\lambda} F_{\rho\lambda}. \tag{2.171}$$

Note, however, that this expression is neither symmetric nor gauge invariant.

It's clear that we could form a symmetric tensor by adding to Eq. (2.171) a term of the form

$$F^{\mu\lambda} \partial_\lambda A^\nu. \tag{2.172}$$

But is this justified? Note that, in the absence of charges, the equations of motion read $\partial_\lambda F^{\mu\lambda} = 0$. Therefore, we can write:

$$F^{\mu\lambda} \partial_\lambda A^\nu = \partial_\lambda (F^{\mu\lambda} A^\nu). \tag{2.173}$$

Since this term is of the form of Eq. (2.108), we may safely add it to the $T^{\mu\nu}$ tensor (i.e., it does not break the condition $\partial_\mu T^{\mu\nu} = 0$). With this, we arrive at the symmetric stress-energy tensor:

$$T^{\mu\nu} = F^{\mu\lambda} F_\lambda{}^\nu + \frac{1}{4} \eta^{\mu\nu} F^{\rho\lambda} F_{\rho\lambda}. \tag{2.174}$$

From the relations in Sec. 2.7.3 [in particular, Eq. (2.141)], the temporal component, which should correspond to the energy density, is readily found:

$$T^{00} = \boldsymbol{E}^2 + \frac{1}{4} F^2 = \frac{1}{2} \left( \boldsymbol{E}^2 + \boldsymbol{B}^2 \right). \tag{2.175}$$

This matches exactly the expression from classical electromagnetism, Eq. (2.170). Similarly, it is quick to show that

$$T^{0i} = \boldsymbol{E} \times \boldsymbol{B} = \frac{1}{c} \boldsymbol{S} \tag{2.176}$$

Therefore, consistent with the discussion in Sec. 2.5.3, we can recognise $cT^{0i}$ as the energy density flux. Similarly, we may find the momentum density:

$$\boldsymbol{g} = \frac{1}{c} T^{0i} = \frac{1}{c^2} \boldsymbol{S}. \tag{2.177}$$

We can further understand the physical significance of the Maxwell stress-energy tensor by considering the case where we reintroduce a non-zero current density, $j^\mu$. Now, since we derived $T_{\mu\nu}$ from the Noether theorem in the source-free case, we can no longer expect it to be conserved. Indeed, we find

$$\begin{aligned}\partial_\mu T^{\mu\nu} &= (\partial_\mu F^{\mu\lambda})F_\lambda{}^\nu + F^{\mu\lambda}(\partial_\mu F_\lambda{}^\nu) + \frac{1}{2}F^{\rho\lambda}(\partial^\nu F_{\rho\lambda}) \\ &= \frac{1}{c}j_\mu F^{\mu\nu} + \frac{1}{2}F_{\mu\lambda}\left[2\partial^\mu F^{\lambda\nu} + \partial^\nu F^{\mu\lambda}\right] \\ &= \frac{1}{c}j_\mu F^{\mu\nu} + \frac{1}{2}F_{\mu\lambda}\left[\partial^\mu F^{\lambda\nu} + \partial^\lambda F^{\mu\nu}\right],\end{aligned} \tag{2.178}$$

where we used the equations of motion (2.160) to replace $\partial_\mu F^{\mu\lambda}$ in the second line, used the Bianchi identity (2.143) to simplify the term in the brackets in the third line, and re-labelled/re-balanced the indices. Now, the second term on the right-hand-side is in the form an anti-symmetric tensor multiplied by a symmetric tensor (in $\mu\lambda$), and thus is zero (Problem 2.2.7). Therefore, we have

$$\partial_\mu T^{\mu\nu} = \frac{1}{c}j_\mu F^{\mu\nu}. \tag{2.179}$$

From a comparison with Eq. (2.140), we can understand this equation as a generalisation of the Lorentz force law. Importantly, this shows how the total energy is conserved: the electromagnetic field exchanges energy and momentum with charged matter. Considering the case with $\nu = 0$, we find the condition for energy (density) conservation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{U} + \nabla\cdot\boldsymbol{S} = -\boldsymbol{j}\cdot\boldsymbol{E}, \tag{2.180}$$

which is Poynting's theorem [c.f. Eq. (2.166)], while taking $\nu = 1, 2, 3$ results in the conservation law for momentum density

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{g} + \nabla\cdot\boldsymbol{T} = -\left(\rho\boldsymbol{E} + \frac{1}{c}\boldsymbol{j}\times\boldsymbol{B}\right) \tag{2.181}$$

where $\boldsymbol{T}$ is the Maxwell stress tensor (spatial part of $T^{\mu\nu}$), and the right-hard-side is the negative of the usual Lorentz force on the charges (2.127), reflecting Newton's third law.

As a final note, we mention that we could have essentially *guessed* the form of the stress-energy tensor from the knowledge that the energy density is second-order in the fields. There are only two tensors we can form from the $F$ (or $A$) fields that are gauge independent, symmetric, and second-order in the fields:

$$T^{\mu\nu} = a\,F^{\mu\lambda}F_\lambda{}^\nu + b\,\eta^{\mu\nu}F^2.$$

The constants $a = 1$ and $b = 1/4$ may then be relatively easily found from the requirement that $T^{00}$ equals the well-known energy density (2.170).

### 2.8.4 Electro- and magnetostatics

From Maxwell's inhomogeneous equation (2.160) we have

$$\partial_\mu F^{\mu\nu} = \partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu = \frac{1}{c} j^\nu.$$

By imposing the Lorenz gauge choice, $\partial_\mu A^\mu = 0$ (2.131), this simplifies to:

$$\partial^2 A^\mu = \frac{1}{c} j^\mu, \tag{2.182}$$

which may be called the *inhomogeneous d'Alembert* equation. In terms of the perhaps more familiar scalar and vector potentials, we have

$$\frac{1}{c^2}\frac{\partial^2 \Phi}{\partial t^2} - \nabla^2 \Phi = \rho, \quad \text{and} \quad \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \nabla^2 \boldsymbol{A} = \frac{1}{c}\boldsymbol{j}, \tag{2.183}$$

which are general (inhomogeneous) wave equations; an important observation that we shall return to in the next section.

In the static case – by which we mean the *potentials* are static, corresponding to static charge or current distributions – this reduces to *Poisson's equation*, $\nabla^2 A^\mu = -j^\mu/c$. This can be solved, e.g., using the method of Green's functions, and has the familiar solution (see Appendix 2.11.5 for derivation):

$$A^\mu(\boldsymbol{x}) = \frac{1}{4\pi\, c} \int \frac{j^\mu(\boldsymbol{x'})}{|\boldsymbol{x} - \boldsymbol{x'}|}\, \mathrm{d}^3 x'. \tag{2.184}$$

From this, we may find expressions for the electric and magnetic fields (2.126) due to static charge and current distributions. Taking the gradient of the (static) scalar potential, we find

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{x}) = -\nabla_x \Phi(\boldsymbol{x}) &= -\frac{1}{4\pi} \int \nabla_x \frac{\rho(\boldsymbol{x'})}{|\boldsymbol{x} - \boldsymbol{x'}|}\, \mathrm{d}^3 x' \\ &= \frac{1}{4\pi} \int \rho(\boldsymbol{x'}) \frac{\boldsymbol{n}}{|\boldsymbol{x} - \boldsymbol{x'}|^2}\, \mathrm{d}^3 x', \end{aligned} \tag{2.185}$$

where $\boldsymbol{n}$ is the unit vector $(\boldsymbol{x} - \boldsymbol{x'})/|\boldsymbol{x} - \boldsymbol{x'}|$. This is *Coulomb's law* for general charge distributions. Similarly, taking the curl of the vector potential, we find[18]

$$\begin{aligned} \boldsymbol{B}(\boldsymbol{x}) = \nabla \times \boldsymbol{A}(\boldsymbol{x}) &= \frac{1}{4\pi\, c} \int \nabla_x \times \left[ \frac{\boldsymbol{j}(\boldsymbol{x'})}{|\boldsymbol{x} - \boldsymbol{x'}|} \right] \mathrm{d}^3 x' \\ &= -\frac{1}{4\pi\, c} \int \frac{\boldsymbol{n} \times \boldsymbol{j}(\boldsymbol{x'})}{|\boldsymbol{x} - \boldsymbol{x'}|^2}\, \mathrm{d}^3 x', \end{aligned} \tag{2.186}$$

which is a general statement of the *Biot-Savart* law.

[18] Using the vector identity $\nabla \times (\phi \boldsymbol{a}) = (\nabla \phi) \times \boldsymbol{a} + \phi(\nabla \times \boldsymbol{a})$

## 2.9 Electromagnetic waves and radiation

One important thing to notice about Maxwell's equations is that there are non-trivial solutions, even in the case where there are no charges. Such solutions are known as free (or vacuum) *electromagnetic waves*.

In the absence of charges, and maintaining the Lorenz gauge condition, the d'Alembert equation (2.182) becomes

$$\partial^2 A^\nu = 0,$$

which you may recognise as the *wave equation*:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) A^\mu = 0. \tag{2.187}$$

It can be seen that functions of the form

$$A^\mu(x) = \epsilon^\mu a_0 e^{\pm i k_\nu x^\nu} \tag{2.188}$$

are solutions, where $\epsilon^\mu$ is a unit four-vector termed the *polarisation vector*, $a_0$ is an arbitrary coefficient (the *amplitude*), $k_\mu$ is an arbitrary four vector, and it is understood that the real part of the solution is taken. If we write the components of this vector as $k^\mu = (\omega/c, \boldsymbol{k})$, then these solutions may be identified as plane waves of angular frequency $\omega$, and wavevector $\boldsymbol{k}$ (which are related to the frequency $f = \omega/(2\pi)$ and wavelength $\lambda = 2\pi/|\boldsymbol{k}|$). The solution describes a plane wave travelling along the direction $\boldsymbol{k}$, with speed $c$. Indeed, it was this remarkable observation that led Maxwell to propose that light itself was an electromagnetic wave, writing in 1865:

> "*This velocity is so nearly that of light, that it seems we have strong reason to conclude that light itself...is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field*"[19]

Of course, general solutions may be much more complicated than the simple plane waves of Eq. (2.188). However, general solutions may be formed from linear combinations of plane waves with different amplitudes and $k^\mu$ vectors (Fourier theorem):

$$A^\mu(x) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, a^\mu_{\boldsymbol{k}}\, e^{i k_\nu x^\nu}. \tag{2.189}$$

The condition that $A$ be real requires that $a_{\boldsymbol{k}} = a^*_{-\boldsymbol{k}}$.[20] Solutions that can be written in the form of Eq. (2.188) – i.e., that involve only a single frequency $\omega$ – are said to be *monochromatic*.

[19] J. C. Maxwell, *A dynamical theory of the electromagnetic field*, Philos. Trans. **155**, 459 (1865).

[20] The integral needs only extend over three independent components of $k^\nu$, see Eq. (2.190).

From the wave equation (2.187), we find the condition on $k_\mu$:

$$k^\mu k_\mu = 0, \tag{2.190}$$

which means there are only three independent components of $k^\mu$. Combining this with the definition of $k_\mu$, we see that $\omega^2 = c^2\boldsymbol{k}^2$, and

$$\boldsymbol{k} = \frac{\omega}{c}\hat{\boldsymbol{n}}, \tag{2.191}$$

where $\hat{\boldsymbol{n}}$ is the unit vector along the direction of propagation of the wave. Further, taking the derivative of the solution (2.188), and noting the Lorenz gauge condition $\partial_\mu A^\mu = 0$, we find

$$k_\mu \epsilon^\mu = 0, \tag{2.192}$$

so the polarisation must be perpendicular to the direction of propagation.

We shall now further impose the *Coulomb* gauge condition $\nabla\cdot\boldsymbol{A} = 0$. From the Lorenz condition $\partial_\mu A^\mu = 0$, it is immediately clear that for plane-wave solutions, this implies $\Phi = 0$. In the context of source-free monochromatic plane waves, the Coulomb gauge thus implies the *radiation gauge* (see Sec. 2.7.2). Further, taking the divergence of the vector part of $A$,

$$\nabla\cdot\boldsymbol{A} = \pm i a_0\,(\boldsymbol{k}\cdot\boldsymbol{\epsilon})\,e^{\pm i k_\mu x^\mu} = 0, \tag{2.193}$$

we see that $\boldsymbol{k}\cdot\boldsymbol{\epsilon} = 0$, which means the vector potential is transverse to the direction of propagation.

As a final note in this section, we shall work out the form of the electric and magnetic fields. Remembering the definition of the fields (2.126), we take the time derivative of the vector potential (2.188) in the Coulomb gauge to find:

$$\boldsymbol{E} = \mathrm{re}\left\{ i\frac{a_0\omega}{c}\,\boldsymbol{\epsilon}\,e^{i(\omega t - \boldsymbol{k}\cdot\boldsymbol{x})}\right\} = -\frac{a_0\omega}{c}\,\boldsymbol{\epsilon}\,\sin(\omega t - \boldsymbol{k}\cdot\boldsymbol{x}). \tag{2.194}$$

Similarly, by taking the curl, we have

$$\boldsymbol{B} = \mathrm{re}\left\{ i a_0\,\boldsymbol{k}\times\boldsymbol{\epsilon}\,e^{i(\omega/ct - \boldsymbol{k}\cdot\boldsymbol{x})}\right\} = \boldsymbol{n}\times\boldsymbol{E}. \tag{2.195}$$

Therefore, the plane wave solutions may be understood to consist of simultaneously oscillating electric and magnetic components, both of which are orthogonal to the direction of propagation, and to each other.

It's important to notice that the electromagnetic waves carry both energy and momentum. The energy density stored in the waves $\mathcal{H} \propto (E^2 + B^2)$ is non-zero. Importantly, the Poynting vector is also non-zero:

$$\boldsymbol{S} = c\boldsymbol{E}\times\boldsymbol{B} = \frac{a_0^2\omega^2}{c}\sin^2(\omega t - \boldsymbol{k}\cdot\boldsymbol{x})\,\boldsymbol{n}, \tag{2.196}$$

meaning the waves carry both energy and momentum through space, with both being along the direction of the wave propagation (i.e., along $\boldsymbol{k}$). It is often more useful to consider the time-averages:

$$\langle \mathcal{H} \rangle = \frac{a_0^2 \omega^2}{2c^2}, \quad \text{and} \quad \langle \boldsymbol{S} \rangle = \frac{a_0^2 \omega^2}{2c} \boldsymbol{n}. \tag{2.197}$$

### 2.9.1 Moving charges and the retarded potentials

We shall now consider the general case of the potentials due to time-dependent charge and current distributions. For simplicity, we will consider the scalar potential; the derivation for the vector potential follows similarly. Beginning from the inhomogeneous d'Alembert equation (2.182), and maintaining the Lorenz Gauge condition, we have

$$\partial^2 \Phi(\boldsymbol{x}, t) = \rho(\boldsymbol{x}, t). \tag{2.198}$$

Since this is a linear equation, the general solution can be expressed as the sum of solutions due to charge distributions of separate infinitesimal pieces of space. As such, we first consider just the solution due to a small point charge $\mathrm{d}q$, located at position $\boldsymbol{x}_0$. We thus write $\rho(t, \boldsymbol{x}) = \mathrm{d}q(t)\delta(\boldsymbol{x} - \boldsymbol{x}_0)$, so that

$$\partial^2 \Phi = \mathrm{d}q(t)\delta(\boldsymbol{R}), \tag{2.199}$$

where we introduce the variable $\boldsymbol{R} \equiv \boldsymbol{x} - \boldsymbol{x}_0$, noting that $\boldsymbol{x}_0$ is a constant (the $t$-dependence has been absorbed into $\mathrm{d}q$). We solve this using standard techniques; for an alternative approach using complex analysis see Appendix 2.11.6.

Since we consider just a single point charge, the problem has a clear spherical symmetry: $\Phi$ is a function only of $R = |\boldsymbol{R}|$. Everywhere except at $R = 0$, we have $\partial^2 \Phi = 0$. Writing the Laplacian in spherical coordinates [see Appendix 2.11.3]:

$$\frac{1}{c^2}\frac{\partial^2 \Phi}{\partial t^2} - \frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2 \frac{\partial \Phi}{\partial R}\right) = 0,$$

and using the common trick of defining $\chi \equiv R\Phi$, we arrive at (for $R \neq 0$):

$$\frac{1}{c^2}\frac{\partial^2 \chi}{\partial t^2} - \frac{\partial^2 \chi}{\partial R^2} = 0, \tag{2.200}$$

which is just the one-dimensional wave equation. As we saw above, this has periodic solutions of the form $e^{i(\omega t - kR)}$, with $k = \pm\omega/c$. Since we need just one particular solution, we take positive case, and thus write $\chi = \chi_1(t - R/c)$, or

$$\Phi = \frac{\chi_1(t - R/c)}{R}. \tag{2.201}$$

We may determine the specific form of $\chi_1$, and thus $\Phi$, by asserting that $\Phi$ returns the correct solution to (2.199) at $R = 0$. Notice that, for periodic $\chi$

solutions, the potential $\Phi \to \infty$ as $R \to 0$ from Eq. (2.201). As such, close to $R = 0$, the time derivatives in Eq. (2.199) are negligible compared to the spatial derivatives, and the solution has the same form as the static case we found in Eq. (2.184). Therefore, we have $\chi(t) = \frac{1}{4\pi}\mathrm{d}q(t)$, and

$$\Phi(\boldsymbol{x}, t) = \frac{1}{4\pi}\frac{\mathrm{d}q(t - R/c)}{R}. \tag{2.202}$$

The extension to general charge distributions is clear: write $\mathrm{d}q = \rho \mathrm{d}V$, and integrate over the entire space:

$$\Phi_r(\boldsymbol{x}, t) = \frac{1}{4\pi}\int \frac{\rho(\boldsymbol{x}', t_r)}{|\boldsymbol{x} - \boldsymbol{x}'|}\,\mathrm{d}^3 r', \tag{2.203}$$

where $t_r = t - |\boldsymbol{x} - \boldsymbol{x}'|/c$. Of course, the derivation for the other components of the potential follow in exactly the same way, resulting in

$$A^\mu_r(\boldsymbol{x}, t) = \frac{1}{4\pi\, c}\int \frac{j^\mu(\boldsymbol{x}', t_r)}{|\boldsymbol{x} - \boldsymbol{x}'|}\,\mathrm{d}^3 r', \tag{2.204}$$

which is known as the *retarded potential*.

The retarded potential, $A^\mu_r$, at position $\boldsymbol{x}$ and time $t$, is thus understood to be generated by time-varying current distributions located at position $\boldsymbol{x}'$ and at the *past time* (or retarded time) $t_r = t - |\boldsymbol{x} - \boldsymbol{x}'|/c$. Of course, the d'Alembert equation (2.182) may also omit homogeneous solutions, $A^\mu_0$ [see Appendix 2.11.5]. As such, total solutions will be of the form

$$A^\mu = A^\mu_0 + A^\mu_r,$$

where $A^\mu_0$ are determined from the initial/boundary conditions of the problem.

As a final note, we point out that at Eq. (2.201), we specifically chose the retarded solution, in which the potentials at $(\boldsymbol{x}, t)$ depend on the sources at earlier times $t_r = t - |\boldsymbol{x} - \boldsymbol{x}'|/c$. This ensures that changes in the potentials caused by changes in the current distributions propagate outward at the speed of light, encoding causality. The retarded potential thus make explicit the *locality* encoded in the relativistic field theory. One could, in principle, also construct the so-called *advanced potential*, which depend on sources at future times. The d'Alembert equation is time symmetric, so it's not surprising we get both kinds of solutions. The retarded potential is usually the useful one: given some distribution of charges or currents in the past, we can determine the potentials and fields they imply with the retarded solutions. Since the theory is reversible, if we instead know the "future" (or final states) of the charges, we can use the advanced potential to work out the prior potentials that would imply them.

We may also explicitly write down the retarded and advanced Green's functions, which solve

$$\partial^2 G(x - x') = \delta^{(4)}(x - x').$$

The two relevant solutions, the advanced (+) and retarded (−) Green's functions,

$$G_{\pm}(x-x') = \frac{\delta\big(t-t' \pm |\boldsymbol{x}-\boldsymbol{x}'|/c\big)}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}, \tag{2.205}$$

may be integrated with the source term in Eq. (2.182) to produce the retarded and advanced potentials, respectively, as discussed in Appendix 2.11.5.

### 2.9.2 The Liénard-Wiechert potential

We now seek the potential due to an arbitrarily moving point charge, which follows the trajectory $\boldsymbol{x}_0(t)$. This will be useful for studying the electromagnetic fields produced by accelerated charges, and in particular, for describing radiation processes. We start from the retarded potential (2.204), which shows that the potential at some spacetime point, $x^\mu$ (i.e., at time $t = x^0/c$), is determined by the current density $j^\mu(t_r)$ at the earlier *retarded* time, $t_r$. For the potential generated by a *point* particle, the time difference $t - t_r$ corresponds to the time of propagation of the light signal from the location of point charge at the retarded time $\boldsymbol{x}_0(t_r)$ to the current point, $\boldsymbol{x}$. In other words, the retarded time satisfies

$$t - t_r = \frac{|\boldsymbol{x}-\boldsymbol{x}_0(t_r)|}{c} \equiv \frac{R(t_r)}{c}, \tag{2.206}$$

where $\boldsymbol{R} \equiv \boldsymbol{x} - \boldsymbol{x_0}$ is the displacement from the point charge, and $R = |\boldsymbol{R}|$.

There are several ways to determine the form of the retarded potential for an arbitrarily moving charge. Perhaps the simplest is to begin in the frame of reference where the point charge is at rest at the retarded time $t_r$, in which case the potential is just that for the static Coulomb case, and then boost "backwards" to the frame where it moves with velocity $\boldsymbol{v}$. In the instantaneous rest frame (denoted with a prime), the potential at time $t$ has just the form of the static potential (2.184) with current density $j^\mu = (c\rho, \boldsymbol{0})$. Noting the charge density for a point charge,

$$\rho(\boldsymbol{x}) = q\,\delta(\boldsymbol{x} - \boldsymbol{x_0}),$$

we carry out the integral in Eq. (2.184) to obtain the familiar Coulomb expression:

$$A'_0(\boldsymbol{x}, t) = \frac{q}{4\pi}\frac{1}{R(t_r)} = \frac{q}{4\pi}\frac{1}{c(t-t_r)} \quad \text{and} \quad \boldsymbol{A}' = 0. \tag{2.207}$$

To determine the potential in the desired frame, we perform a Lorentz transformation corresponding to a boost of $-\boldsymbol{v}$ (so that the particle will have velocity $\boldsymbol{v}$). We discussed the transformation of a vector field in Sec. 2.4.2; we have $A'^\mu(x') = \Lambda^\mu{}_\nu A^\nu(x)$, or, writing the desired frame as the unprimed coordinates,

$$A^\mu(x) = \Lambda^\mu{}_\nu A'^\nu(\Lambda^{-1}x).$$

The form of the Lorentz transformation for a boost along an arbitrary direction is derived in Appendix 2.11.1. For $\Lambda^{-1}x$, noting this is the *inverse* boost of velocity $-\boldsymbol{v}$ (equivalent to a boost of $\boldsymbol{v}$), we have

$$c\Delta t \to \gamma\left(c\Delta t - \frac{1}{c}\boldsymbol{v}\cdot\Delta\boldsymbol{x}\right) = \frac{1}{c}u^\nu R_\nu, \tag{2.208}$$

where $u^\nu = (\gamma c,\, \gamma\boldsymbol{v})$ is the four velocity, and we defined

$$R^\nu \equiv (c[t - t_r],\, \boldsymbol{R})\,,$$

with which Eq. (2.206) can be expressed as $R^\nu R_\nu = 0$. Now, we have:

$$A'_0(\Lambda^{-1}x) = \frac{q}{4\pi}\frac{c}{u^\nu R_\nu} \quad \text{and} \quad \boldsymbol{A}' = 0.$$

Applying the Lorentz transformation [Eq. (2.236)] to $A'$, we finally arrive at

$$\begin{aligned} A^\mu(x) = \Lambda^\mu_\alpha A'^\alpha(\Lambda^{-1}x) &= \frac{q}{4\pi}\frac{c}{u^\nu R_\nu}\,(\gamma,\, \gamma v_x/c,\, \gamma v_y/c,\, \gamma v_z/c\,) \\ &= \frac{q}{4\pi}\left[\frac{u^\mu}{u^\nu R_\nu}\right]_{t_r}, \end{aligned} \tag{2.209}$$

where the values on the right-hand-side are to be evaluated at the retarded time, found from Eq. (2.206). This are known as the *Liénard-Wiechert* (LW) potential, and is the potential due to an arbitrarily moving point charge, with four velocity $u$. In terms of the scalar and three-vector potentials, these are:

$$\Phi = \frac{1}{4\pi}\frac{q}{\left[R - \frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{R}\right]_{t_r}}, \quad \text{and} \quad \boldsymbol{A} = \frac{1}{4\pi\, c}\frac{q\,\boldsymbol{v}(t_r)}{\left[R - \frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{R}\right]_{t_r}}. \tag{2.210}$$

Note that it's more common in textbooks to derive the these potentials by directly integrating the retarded potential (2.204) with the potential of a point charge (see, e.g., Jackson[21] Sec. 14.1). Care must be taken with the delta function, which depends on the retarded coordinates (see Appendix 2.11.4).

### 2.9.3 Radiation of electromagnetic waves

To investigate the physical consequences of the Liénard-Wiechert potentials (2.210), we shall now find the explicit expressions for the fields they imply. Particular care must be taken with the derivatives, since the fields (2.126) depend on derivatives with respect to the regular $t, \boldsymbol{x}$ coordinates, while the potentials are in terms of the retarded quantities (2.206). While fairly straight forward, the derivation is

[21] J. D. Jackson, *Classical Electrodynamics* (2001).

long and rather cumbersome. The main steps are sketched in Appendix 2.11.7. We eventually find:

$$\boldsymbol{E} = \frac{q}{4\pi\, R^2}\frac{\left(\boldsymbol{n} - \frac{1}{c}\boldsymbol{v}\right)}{\gamma^2\,\eta^3} + \frac{q}{4\pi\, c\, R}\frac{\boldsymbol{n}\times\left[\left(\boldsymbol{n} - \frac{1}{c}\boldsymbol{v}\right)\times\frac{1}{c}\boldsymbol{a}\right]}{\eta^3}, \tag{2.211}$$

$$\text{and} \qquad \boldsymbol{B} = \boldsymbol{n}\times\boldsymbol{E}, \tag{2.212}$$

where all quantities on the right-hand-side are evaluated at the retarded time, $\boldsymbol{v} = \partial\boldsymbol{x}_0/\partial t_r$, the acceleration is $\boldsymbol{a} = \partial\boldsymbol{v}/\partial t_r$, and we defined

$$\boldsymbol{n} = \frac{\boldsymbol{R}}{R} \quad \text{and} \quad \eta \equiv (1 - \tfrac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}). \tag{2.213}$$

Notice that the electric and magnetic fields are always perpendicular.

The first term in the expression for the electric field (2.211) has a familiar form. It depends on the velocity, diminishes like $1/R^2$, and corresponds to the relativistic correction to the static Coulomb field. Indeed, this term, denoted $\boldsymbol{E}_v$, can be seen to arise from a Lorentz transformation of the electric field due to a static point charge.

The second term in Eq. (2.211) is different. It diminishes just *linearly* with $R$, and depends only on the acceleration of the charge. This term is called the *radiation* field, denoted $\boldsymbol{E}_a$. To see why, consider the Poynting vector:

$$\boldsymbol{S} = c\boldsymbol{E}\times\boldsymbol{B} = cE^2\boldsymbol{n} - c(\boldsymbol{E}\cdot\boldsymbol{n})\boldsymbol{E}. \tag{2.214}$$

The total energy flux through the surface of a sphere of radius $R$ is

$$P = \int \boldsymbol{S}\cdot\mathrm{d}\boldsymbol{A} = \int \boldsymbol{S}\cdot\hat{\boldsymbol{n}}\, R^2\,\mathrm{d}\Omega.$$

If we take $R\to\infty$, this gives the total power radiated away by the fields, and we can interpret $\boldsymbol{S}\cdot\boldsymbol{n}$ as the *angular distribution* of the radiated energy. From Eq. (2.211), we see the $E_v^2$ and $E_v\cdot E_a$ terms fall as $R^{-4}$ and $R^{-3}$, respectively, and thus go to zero in the integral as $R\to\infty$. The $E_a^2$ term, on the other hand, falls as $R^{-2}$ and thus survives the integration. Therefore, the acceleration term leads to a loss of energy via radiation out to infinity! Notice that the acceleration field is transverse, i.e.,. it is perpendicular to $\boldsymbol{n}$, so we have $\boldsymbol{S}_a = cE^2\boldsymbol{n}$, and

$$P = c\int |E_a|^2\, R^2\,\mathrm{d}\Omega. \tag{2.215}$$

It is instructive to consider a simple example, where the charged particle oscillates with uniform sinusoidal motion: $\boldsymbol{x}_0(t) = \boldsymbol{r}\sin(\omega t)$, and consider the field at large distances (the *radiation zone*), where $x \gg r$, so that $\boldsymbol{R}\approx\boldsymbol{x}$. If we further take the charge to be moving non-relativistically, $v \ll c$, we find:

$$\boldsymbol{E} \approx \frac{\omega^2\, qr}{4\pi\, c^2\, R}\sin(\omega t)\left[\hat{\boldsymbol{r}} - (\boldsymbol{n}\cdot\hat{\boldsymbol{r}})\boldsymbol{n}\right], \tag{2.216}$$

where $\boldsymbol{n} = \boldsymbol{R}/R$ is the unit vector pointing from the charge to the observer, and $\hat{\boldsymbol{r}} = \boldsymbol{r}/r$ is the unit vector for the maximum displacement of the charge. This is known as *dipole* radiation, and the amplitude is determined by $q\boldsymbol{r}$, which may be called the *dipole moment*. Notice that the term in square brackets implies that the radiation field vanishes when the observation direction $\boldsymbol{n}$ is parallel to the displacement $\boldsymbol{r}$, meaning no radiation is emitted along the axis of oscillation. Taking the time average of $E^2$, we find

$$\langle E \rangle^2 = \frac{\omega^4 (qr)^2}{32\pi^2\, c^4\, R^2} \sin^2\theta,$$

where $\theta$ is the angle between $\hat{\boldsymbol{r}}$ and $\boldsymbol{n}$, from which we can immediately find the angular distribution and the total power radiated:

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{\omega^4 (qr)^2}{32\pi^2\, c^3} \sin^2\theta, \quad \text{and} \quad P = \frac{\omega^4 (qr)^2}{12\pi\, c^3}. \tag{2.217}$$

This formula is the *Larmor formula* for dipole radiation; the relativistic generalisation is given by Eq. (2.215).

***Problem* 2.9.1:** Derive the explicit form of the fields from the LW potentials (2.210). You may refer to the Appendix 2.11.7.

***Problem* 2.9.2:** Thomson scattering: A free electron is accelerated by an incident linearly polarised plane-wave electric field of frequency $\omega$. (i) Find the angular distribution of the radiation it subsequently emits, assuming the non-relativistic case $v/c \ll 1$. (ii) Find the *differential cross section* $\mathrm{d}\sigma/\mathrm{d}\Omega$, defined as the scattered power per unit solid angle divided by the incident intensity $I_{\text{in}} = \langle \boldsymbol{S} \cdot \boldsymbol{n} \rangle$. (iii) Integrate over angles to find the total *Thomson cross section*, $\sigma_T$, and the classical electron radius defined via $\sigma_T = \frac{8\pi}{3} r_e^2$.

*Answer* (2.9.2): $\mathrm{d}P/\mathrm{d}\Omega = \frac{q^2 a_\perp^2}{32\pi^2 m^2 c^3} \sin^2\theta$, where $a_\perp = qE_0/m$. $\mathrm{d}\sigma/\mathrm{d}\Omega = r_e^2 \sin^2\theta$, with $r_e = \frac{q^2}{4\pi m c^2}$.

## 2.10 Gravitation: motivation for General Relativity

Given our success with electrodynamics, it seems reasonable to now seek a relativistically invariant field theory for gravity. We may try, for example, a scalar field $\phi$, to describe gravity, similar to the scalar theory we considered in Sec. 2.4.3:

$$\mathcal{L} = -\kappa\, \partial_\mu \phi \partial^\mu \phi - \phi T. \tag{2.218}$$

Here, $T$ is the source term for gravitation (presumably, related to mass or energy density), and $\kappa$ is an arbitrary constant (corresponding to a choice of units for $\phi$). Note that if $T$ has mass density dimension, then $[\phi] = E/M = L^2/T^2$, the same as the usual gravitational potential. The equations of motion are:

$$\partial^2 \phi = -\frac{1}{2\kappa} T \quad \xrightarrow{\text{static limit}} \quad \nabla^2 \phi = \frac{1}{2\kappa} T. \tag{2.219}$$

If we set $T = \rho_m$, the mass density, then, with $\kappa^{-1} = 8\pi G$, this exactly recovers the Newtonian Poisson equation. We are therefore motivated to recognise $\phi$ as the gravitational potential.

However, $\rho_m$ is *not* a Lorentz scalar, and so this theory would not be relativistically invariant. We may then seek a scalar that, in the non-relativistic limit, reduces to the mass density $\rho_m$. One such quantity is the trace of the stress-energy tensor, $T = \frac{1}{c^2} T^\mu_\mu$ (non-relativistically $T^{ij} \ll T^{00}$). So we postulate

$$\mathcal{L} = \frac{1}{8\pi G} \partial_\mu \phi \partial^\mu \phi - \frac{1}{c^2} \phi \, \eta_{\mu\nu} T^{\mu\nu}, \tag{2.220}$$

which will recover Newtonian gravity in the non-relativistic limit.

We could then consider interactions between a small test particle of mass $m$ with the background gravitational field, following a similar approach to that used for electromagnetic interactions in Sec. 2.7. A scalar interaction term could then have the form (for example)

$$S_{\text{int}} = -\frac{q}{c} \int \phi(x) \mathrm{d}s,$$

where $q$ plays the role of the gravitational "charge". To be dimensionally consistent, $q$ should have mass dimension. This is consistent with our expectation that the inertial mass, $m$, *is* the gravitational charge, so we take $q = m$. It's a quick exercise (following Sec. 2.3) to show that this indeed leads to the expected Lagrangian $L = \frac{1}{2} m v^2 - m\phi$ in the non-relativistic limit.

While such a scalar theory successfully reproduces Newtonian gravity in the non-relativistic limit, it suffers from several flaws. For example, it cannot predict the gravitational bending of light, since the trace of the Maxwell stress-energy tensor (2.174) vanishes. Additionally, while it satisfies the weak equivalence principle (universality of free-fall), it does not satisfy the general Einstein equivalence principle, which requires local Lorentz and position invariance.[22] Presumably, the real theory should couple to the full stress-energy tensor, hinting at a tensor, rather than scalar, field structure. These shortcomings are resolved in the theory of *general relativity*, where gravity is described by a tensor field $g_{\mu\nu}$.

### 2.10.1 Particle in a gravitational field, the geodesic

As a motivation for general relativity, we will consider how a particle may interact with such a tensor field. Remember that to couple a particle to a scalar field, we added a term to the action $\phi \mathrm{d}s$. Similarly, to couple a vector field we added a term to the action $A_\mu \mathrm{d}x^\mu$. Assuming the interaction with gravity can be described by a tensor field, $g_{\mu\nu}$, we could include a term $g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu$, though this

[22] We don't prove or discuss this here, though it forms the fundamental basis for all of general relativity. See any GR textbook for the detailed discussion.

would be second-order in the coordinate differentials. We therefore *guess* that the symmetric tensor field $g$ enters the action as[23]

$$S = -mc \int \sqrt{g_{\mu\nu} \mathrm{d}x^\nu \mathrm{d}x^\mu}, \tag{2.221}$$

where $g = g(x)$ is a function of the spacetime coordinate. Notice that, as the gravitational field goes to zero, we must have $g_{\mu\nu} \to \eta_{\mu\nu}$ in order to return the correct action for a relativistic particle, Eq. (2.39). Just as we did in the case of electromagnetic interactions, we work in the weak coupling limit, where the effect of the motion of the particle on the gravitational field can be neglected.

To continue, we re-define $\mathrm{d}s$ as a generalisation of the Minkowski interval:

$$\mathrm{d}s \equiv \sqrt{g_{\mu\nu} \mathrm{d}x^\nu \mathrm{d}x^\mu},$$

which we can simply consider as a notational convenience for now (though, will gain deep significance within the full theory). Then, varying the action, we have

$$\begin{aligned} \delta S &= \frac{-mc}{2} \int \frac{1}{\mathrm{d}s} \delta(g_{\mu\nu} \mathrm{d}x^\nu \mathrm{d}x^\mu) \\ &= \frac{-mc}{2} \int \frac{1}{\mathrm{d}s} \left[ \delta(g_{\mu\nu}) \mathrm{d}x^\nu \mathrm{d}x^\mu + 2 g_{\mu\nu} \mathrm{d}x^\nu \delta(\mathrm{d}x^\mu) \right], \end{aligned} \tag{2.222}$$

where we used the symmetry of $g_{\mu\nu}$. We have

$$\delta(g_{\mu\nu}) = (\partial_\lambda g_{\mu\nu}) \delta x^\lambda \quad \text{and} \quad \delta(\mathrm{d}x^\mu) = \mathrm{d}(\delta x^\mu),$$

and so the condition $\delta S = 0$ becomes

$$\int \left[ (\partial_\lambda g_{\mu\nu}) \mathrm{d}x^\nu \frac{\mathrm{d}x^\mu}{\mathrm{d}s} \delta x^\lambda + 2 g_{\lambda\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}s} \mathrm{d}(\delta x^\lambda) \right] = 0, \tag{2.223}$$

where we exchanged $\lambda \leftrightarrow \mu$ in the second term, and divided through by the differential $\mathrm{d}s$. Using integration by parts for the second term, and discarding boundary terms, we arrive at

$$\begin{aligned} &\int \left[ (\partial_\lambda g_{\mu\nu}) \mathrm{d}x^\nu \frac{\mathrm{d}x^\mu}{\mathrm{d}s} \delta x^\lambda - 2\,\mathrm{d} \left( g_{\lambda\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}s} \right) \delta x^\lambda \right] \\ &= \int \left[ (\partial_\lambda g_{\mu\nu}) \mathrm{d}x^\nu \frac{\mathrm{d}x^\mu}{\mathrm{d}s} - 2(\partial_\mu g_{\lambda\nu}) \frac{\mathrm{d}x^\nu}{\mathrm{d}s} \mathrm{d}x^\mu - 2 g_{\lambda\nu} \frac{\mathrm{d}^2 x^\nu}{\mathrm{d}s^2} \mathrm{d}s \right] \delta x^\lambda = 0. \end{aligned} \tag{2.224}$$

Since the action should be minimised for arbitrary variations $\delta x$, the term in the brackets should be zero. Dividing through by $\mathrm{d}s$, the equations of motion become

$$g_{\lambda\nu} \frac{\mathrm{d}^2 x^\nu}{\mathrm{d}s^2} + \partial_\mu g_{\lambda\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}s} \frac{\mathrm{d}x^\nu}{\mathrm{d}s} - \frac{1}{2} \partial_\lambda g_{\mu\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}s} \frac{\mathrm{d}x^\mu}{\mathrm{d}s} = 0. \tag{2.225}$$

[23] See Eq. (2.30) and Prob. 2.2.7 for why we only need to consider a symmetric tensor.

To further simplify, we now *define* the raised tensor $g^{\mu\sigma}$ as the inverse of $g$:

$$g^{\mu\sigma} g_{\sigma\nu} \equiv \delta^{\mu}_{\nu}, \tag{2.226}$$

which is a departure from the Minkowski notation where $\eta$ is used to raise or lower indices. Indeed, as it turns out, in general relativity it is $g$ that plays the role of the (now position-dependent) metric, though we are getting ahead of ourselves. For now, we shall simply take this as a definition used to simplify the form of the equations. With this, we may multiply by $g^{\sigma\lambda}$, leading to

$$\frac{\mathrm{d}^2 x^{\sigma}}{\mathrm{d}s^2} + g^{\sigma\lambda}(\partial_{\mu} g_{\lambda\nu} - \frac{1}{2}\partial_{\lambda} g_{\mu\nu})\frac{\mathrm{d}x^{\mu}}{\mathrm{d}s}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}s} = 0. \tag{2.227}$$

Finally, we define

$$\Gamma^{\sigma}_{\mu\nu} \equiv \frac{1}{2} g^{\sigma\lambda}(\partial_{\mu} g_{\lambda\nu} + \partial_{\nu} g_{\mu\lambda} - \partial_{\lambda} g_{\mu\nu}), \tag{2.228}$$

after which the equations of motion become

$$\frac{\mathrm{d}^2 x^{\sigma}}{\mathrm{d}s^2} + \Gamma^{\sigma}_{\mu\nu}\frac{\mathrm{d}x^{\mu}}{\mathrm{d}s}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}s} = 0 \tag{2.229}$$

where we exploited the symmetry in (2.227) to replace $\lambda \leftrightarrow \nu$ in the middle term. The $\Gamma$ factor (2.228) is called a *Christoffel symbol*. For now, we take this simply as notation, though will discuss its significance in a moment.

To understand what this equation (2.229) is actually telling us, it is constructive to consider the non-relativistic limit, where the particle is moving slowly compared to the speed of light. This implies $\mathrm{d}x^0 = c\,\mathrm{d}t \gg \mathrm{d}x^i$, so that

$$\frac{\mathrm{d}^2 x^{\sigma}}{\mathrm{d}s^2} + \Gamma^{\sigma}_{00}\frac{\mathrm{d}x^{0}}{\mathrm{d}s}\frac{\mathrm{d}x^{0}}{\mathrm{d}s} \approx 0.$$

Further, we will consider the case where the gravitational field is weak, in which case we can write $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and keep terms only up to linear in $h$. Finally, we also consider the case where the gravitational field is independent of time. With these approximations, the Christoffel symbol becomes

$$\begin{aligned}\Gamma^{\sigma}_{00} &= \frac{1}{2} g^{\sigma\lambda}(\partial_0 g_{\lambda 0} + \partial_0 g_{0\lambda} - \partial_{\lambda} g_{00}) \\ &= -\frac{1}{2}\eta^{\sigma\lambda}\partial_{\lambda} h_{00} = +\frac{1}{2}\delta^{\sigma i}\partial_i h_{00},\end{aligned} \tag{2.230}$$

and the equation of motion reduces to

$$\frac{\mathrm{d}^2 \boldsymbol{x}}{\mathrm{d}t^2} = -\frac{c^2}{2}\nabla h_{00}.$$

If we recognise $h_{00} = \frac{2}{c^2}\phi$, with $\phi$ the gravitational potential, then this is exactly what we expect from Newton! In other words, we recognise

$$g_{00} \approx 1 + \frac{2}{c^2}\phi(\boldsymbol{x}).$$

In summary, our guessed interaction with a symmetric tensor field reproduced Newtonian gravity in the appropriate limit, and indeed, Eq. (2.229) is the correct form of the equations of motion for a test mass implied by general relativity. The remarkable insight is that this equation of motion (2.229) is equivalent to the "straight line" motion of a free particle moving through curved spacetime, known as a *geodesic equation*. The Christoffel symbol Eq. (2.228), which arose naturally from the coupling of a tensor field to the spacetime coordinates, arises exactly in this form when taking derivatives in curved spacetime.

Within general relativity, we define the *covariant derivative*,

$$\nabla_\nu A^\mu = \partial_\nu A^\mu + \Gamma^\mu_{\nu\lambda} V^\lambda, \tag{2.231}$$

as a way of differentiating tensors that respects the curved geometry of spacetime. The Christoffel symbol can be seen as a correction term that accounts for the curvature encoded in the metric $g_{\mu\nu}$. It ensures that the result transforms as a tensor under general coordinate transformations, something that the ordinary (partial) derivative does *not* guarantee in curved spacetime (we do not prove this here, but mention it as a motivation). Crucially, since $\nabla_\nu A^\mu$ transforms as a tensor but $\partial_\nu A^\mu$ does not, the Christoffel symbol $\Gamma^\mu_{\nu\lambda}$ is *not* a tensor.

As a final remark, note that we have not yet considered the equations of motion for the field $g_{\mu\nu}$ itself. These are known as the *Einstein field equations*, in which gravity is described by a position-dependent metric, $g_{\mu\nu}$, that encodes the curvature of spacetime, defines the interval, $\mathrm{d}s$, and whose evolution depends on the total stress-energy tensor. The full discussion of these equations shall await a study General Relativity.

## 2.11 Appendix

### 2.11.1 Lorentz transformations for arbitrary boosts

While, fundamentally, boosts along a single direction are all that's required to formulate the theory (due to the assumed rotational invariance), it is often useful in practical problems to have a general form for Lorentz boosts. Here, we consider the Lorentz transformation due to a boost of $\boldsymbol{v}$ along an arbitrary axis. It's crucial to note that this is *not* the same a boost $v_x$ along the $x$-axis followed by a boost $v_y$ along the $y$-axis etc. A boost can be interpreted as a rotation within the four dimensional Minkowski space, and in general, rotations do not commute. In other words, the boost $v_x$ followed by the boost $v_y$ *is not the same* as the boost $v_y$ followed by the boost $v_x$, which can be seen directly from the matrix form of the single-axis Lorentz boost, Eq. (2.20).

To determine the correct form for the general boost, we first write the coordinate three-vector as

$$\boldsymbol{r} = \boldsymbol{r}_\parallel + \boldsymbol{r}_\perp, \tag{2.232}$$

where $\boldsymbol{r}_\parallel$ is the part of $\boldsymbol{r}$ parallel to the boost direction, $\hat{\boldsymbol{n}}$:

$$\boldsymbol{r}_\parallel = (\boldsymbol{r}\cdot\hat{\boldsymbol{n}})\hat{\boldsymbol{n}} = \frac{(\boldsymbol{r}\cdot\boldsymbol{v})\boldsymbol{v}}{v^2}. \tag{2.233}$$

Since the components perpendicular to the boost velocity are invariant, the transformed quantities are [see Eq. (2.13)]:

$$\begin{aligned}\boldsymbol{r}' &= \gamma(\boldsymbol{r}_\parallel - \boldsymbol{v}t) + \boldsymbol{r}_\perp \\ &= \boldsymbol{r} + \boldsymbol{r}_\parallel(\gamma-1) - \gamma\boldsymbol{v}t, \\ &= \boldsymbol{r} + \left[\frac{(\boldsymbol{r}\cdot\boldsymbol{v})}{v^2}(\gamma-1) - \gamma t\right]\boldsymbol{v},\end{aligned} \tag{2.234}$$

where we used Eq. (2.232) to eliminate $\boldsymbol{r}_\perp$ and Eq. (2.233) to eliminate $\boldsymbol{r}_\parallel$. Since the time coordinate transformation depends only on the component of position along the direction of motion, this transformation is even simpler:

$$ct' = \gamma\left[ct - \frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{r}\right], \tag{2.235}$$

which follows directly from the rotational invariance of the dot product. Using the four-vector notation $a = (a_0, \boldsymbol{a})$, we can write this in a compact form as:

$$\Lambda(\boldsymbol{v}) = \begin{pmatrix} \gamma & -\gamma\boldsymbol{v}^T/c \\ -\gamma\boldsymbol{v}/c & \mathbb{I}_3 + (\gamma-1)\frac{\boldsymbol{v}\boldsymbol{v}^T}{v^2}\end{pmatrix}, \tag{2.236}$$

where $\boldsymbol{v}\boldsymbol{v}^T$ is the outer product resulting in matrix $[\boldsymbol{v}\boldsymbol{v}^T]_{ij} = v_i v_j$, and $\mathbb{I}_3$ is the $3\times3$ identity (note that $\boldsymbol{v}^T\boldsymbol{a} \equiv \boldsymbol{v}\cdot\boldsymbol{a}$).

### 2.11.2 Unit systems for electrodynamics

In defining the units for electrodynamics, one must make choices in linking the definitions of charges and field strengths to the usual mechanical units of length $L$, mass $M$, and time $T$. For example, in Coulomb's law:

$$F = k_e \frac{q_1 q_2}{r^2},$$

the value of the proportionality constant, $k_e$, depends on the definition of charges in terms of the usual mechanical units (and similarly for magnetic fields). In SI units, for example, charges are measured in Coulombs, and $k_e$ is taken to be $1/(4\pi\varepsilon_0)$. It's important to keep track of the unit system being used; in electrodynamics, not only do the values of constants change, but the factors appearing in equations change also. This is because switching between common systems not only changes the *units* of the fields and charges, but also their *dimension*.[24]

Within the SI system of units, the vacuum permittivity $\varepsilon_0$, and vacuum permeability $\mu_0$, are linked via[25]

$$\varepsilon_0\mu_0 = 1/c^2, \qquad \text{and} \qquad \mu_0 = \frac{4\pi\alpha\hbar}{e^2 c}, \qquad \text{or} \qquad 4\pi\varepsilon_0 = \frac{e^2}{\hbar c\,\alpha}. \tag{2.237}$$

The SI system is a so-called *rationalised* unit system, where no $4\pi$ factors appear in Maxwell's equations; instead they appear in the inverse-square force laws (i.e., Coulomb's law and the Biot-Savart law). It is also possible to choose a *non-rationalised* unit system, where the $4\pi$ factors do not appear in the inverse-square force laws, but instead in Maxwell's equations.

Particularly in theoretical physics, it is extremely common to work with units systems where $\varepsilon_0$ and $\mu_0$ do not appear in the equations (this can be thought of as by defining units in which, e.g., $\varepsilon_0 = 1$, or by absorbing $\varepsilon_0$ into the definition of the charges). Two common choices are the *Heaviside-Lorentz* system, which is a rationalised system, and the *Gaussian* system, which is non-rationalised. The Gaussian system was very common, particularly in older textbooks, and is still commonly used in atomic physics. The Heaviside-Lorentz system (rationalised) is widely used in relativistic and particle physics. In the Gaussian system, the Coulomb law constant is taken to be $k_e = 1$, meaning charges have dimension $\sqrt{M.L^3}/T$. The Heaviside-Lorentz unit system is similar, except that it is rationalised; the constant is taken to be $k_e = 1/(4\pi)$. In both the Heaviside-Lorentz and Gaussian systems, the electric and magnetic fields have the same units as each other; in the SI system, they differ by the velocity dimension $L/T$.

[24] Something of *dimension* length, for example, may have *units* meters (or inches etc.).

[25] Of these constants, only the fine structure constant, $\alpha \approx 1/137.035\,999\,084(21)$ is a measured value. The elementary charge is defined *exactly* as $e \equiv 1.602\,176\,634 \times 10^{-19}$ C. Note that with this definition, $e = |e|$ and the electron has charge $-e$ (the opposite convention is also common, unfortunately). Similarly, the speed of light is defined exactly as $c \equiv 299\,792\,458\,\mathrm{m\,s^{-1}}$, and Planck's constant as $h \equiv 2\pi\hbar \equiv 6.626\,070\,15 \times 10^{-34}$ J s.

In the different systems, Coulomb's law takes the form:

$$F = \underbrace{\frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2}}_{\text{SI}} = \underbrace{\frac{1}{4\pi}\frac{q_1 q_2}{r^2}}_{\text{HL}} = \underbrace{\frac{q_1 q_2}{r^2}}_{\text{Gaus.}}, \tag{2.238}$$

and the Lorentz force law is:

$$\boldsymbol{F} = \underbrace{q(\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B})}_{\text{SI}} = \underbrace{q(\boldsymbol{E} + \tfrac{1}{c}\boldsymbol{v}\times\boldsymbol{B})}_{\text{HL, Gaus.}}. \tag{2.239}$$

For the definition of the fields, in all systems we have $\boldsymbol{B} = \nabla\times\boldsymbol{A}$, though

$$\boldsymbol{E} = \underbrace{-\frac{\partial\boldsymbol{A}}{\partial t} - \nabla\Phi}_{\text{SI}} = \underbrace{-\frac{1}{c}\frac{\partial\boldsymbol{A}}{\partial t} - \nabla\Phi}_{\text{HL, Gaus.}}, \tag{2.240}$$

where $\Phi$ is the electric scalar potential, and $\boldsymbol{A}$ is the magnetic vector potential:

$$A^\mu \equiv (A_0,\, \boldsymbol{A}) = \underbrace{(\tfrac{1}{c}\Phi,\, \boldsymbol{A})}_{\text{SI}} = \underbrace{(\Phi,\, \boldsymbol{A})}_{\text{HL, Gaus.}}. \tag{2.241}$$

Note that in SI units, the scalar and vector potentials have different *dimension*. It's useful to note that, in all systems, $e\Phi$ has energy dimension, which is typically easier than dealing directly with units for charge.

Following from these, we have for the Poynting vector and energy density:

$$\begin{aligned}
\boldsymbol{S} &= \underbrace{\frac{1}{\mu_0}\boldsymbol{E}\times\boldsymbol{B}}_{\text{SI}} = \underbrace{c\,\boldsymbol{E}\times\boldsymbol{B}}_{\text{HL}} = \underbrace{\frac{c}{4\pi}\boldsymbol{E}\times\boldsymbol{B}}_{\text{Gaus.}}, \quad \text{and} \\
\mathcal{U} &= \underbrace{\frac{\varepsilon_0}{2}\left(\boldsymbol{E}^2 + c^2\boldsymbol{B}^2\right)}_{\text{SI}} = \underbrace{\frac{1}{2}\left(\boldsymbol{E}^2 + \boldsymbol{B}^2\right)}_{\text{HL}} = \underbrace{\frac{1}{8\pi}\left(\boldsymbol{E}^2 + \boldsymbol{B}^2\right)}_{\text{Gaus.}}.
\end{aligned} \tag{2.242}$$

The three-vector form of Maxwell's equations are:

| SI: | Heaviside-Lorentz: | Gaussian: |
|---|---|---|
| $\nabla\cdot\boldsymbol{E} = \dfrac{\rho}{\varepsilon_0}$ | $\nabla\cdot\boldsymbol{E} = \rho$ | $\nabla\cdot\boldsymbol{E} = 4\pi\rho$ |
| $\nabla\times\boldsymbol{B} = \mu_0\left(\boldsymbol{j} + \varepsilon_0\partial_t\boldsymbol{E}\right)$ | $\nabla\times\boldsymbol{B} = \dfrac{1}{c}\left(\boldsymbol{j} + \partial_t\boldsymbol{E}\right)$ | $\nabla\times\boldsymbol{B} = \dfrac{1}{c}\left(4\pi\boldsymbol{j} + \partial_t\boldsymbol{E}\right)$ |
| $\nabla\times\boldsymbol{E} = -\partial_t B$ | $\nabla\times\boldsymbol{E} = -\dfrac{1}{c}\partial_t B$ | $\nabla\times\boldsymbol{E} = -\dfrac{1}{c}\partial_t B$ |
| $\nabla\cdot\boldsymbol{B} = 0$ | $\nabla\cdot\boldsymbol{B} = 0$ | $\nabla\cdot\boldsymbol{B} = 0,$ |

(2.243)

while in the four-vector form we have:

$$\partial_\mu F^{\mu\nu} = \underbrace{\mu_0 j^\nu}_{\text{SI}} = \underbrace{\frac{1}{c}j^\nu}_{\text{HL}} = \underbrace{\frac{4\pi}{c}j^\nu}_{\text{Gaus.}}, \tag{2.244}$$

where, if $F^{\mu\nu} = (-\boldsymbol{E}, \boldsymbol{B})$ in the Heaviside-Lorentz and Gaussian systems, then $F^{\mu\nu} = (-\boldsymbol{E}/c, \boldsymbol{B})$ in the SI system. The Bianchi identity $\partial_{[\lambda} F_{\mu\nu]} = 0$, or

$$\partial_\mu \widetilde{F}^{\mu\nu} = \frac{1}{2}\partial_\mu \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} = 0,$$

remains the same in all systems.

An often confusing point, which follows from the above definitions, is that the definition of the magnetic moment changes between unit systems. In all systems, it is defined via its interaction energy:

$$\Delta U = -\boldsymbol{\mu} \cdot \boldsymbol{B}.$$

The (non-relativistic) magnetic moment induced by a particle of mass $m$ and charge $q$ in circular motion with angular momentum $\boldsymbol{l}$ is

$$\boldsymbol{\mu} = \underbrace{\frac{q}{2m}}_{\text{SI}} \boldsymbol{l} = \underbrace{\frac{q}{2mc}}_{\text{HL, Gaus.}} \boldsymbol{l} \quad . \tag{2.245}$$

### 2.11.3 Vector calculus identities and notation

The fundamental theorem of calculus

$$\int_{x_1}^{x_2} \partial_x \phi(x)\, \mathrm{d}x = \phi(x_2) - \phi(x_1), \tag{2.246}$$

may be generalised simply to vector gradients of scalar functions:

$$\int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} [\nabla \phi(\boldsymbol{x})] \cdot \mathrm{d}\boldsymbol{l} = \phi(\boldsymbol{x}_2) - \phi(\boldsymbol{x}_1), \tag{2.247}$$

where $\mathrm{d}\boldsymbol{l}$ is the line element for any path connecting $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. The extension to general vector functions are known as the divergence theorem (or Gauss' theorem), and Stokes' theorem.

For a vector function $\boldsymbol{A}(\boldsymbol{x})$ within a volume $V$ bounded by a closed surface $S$, the volume integral of the divergence of $\boldsymbol{A}$ is related to the surface integral of the component of $\boldsymbol{A}$ (outwardly) normal to the surface $S$ as:

$$\int_V (\nabla \cdot \boldsymbol{A})\, \mathrm{d}V = \oint_S \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{S} \tag{2.248}$$

$\mathrm{d}V = \mathrm{d}^3x = \mathrm{d}x\mathrm{d}y\mathrm{d}z$ is the volume element, and $\mathrm{d}\boldsymbol{S} = \hat{\boldsymbol{n}}\mathrm{d}S$ is the surface (area) element normal to the surface. This is known as the *divergence theorem*, or Gauss' theorem. This can be generalised to higher dimensions:

$$\int_\Omega \partial_\mu A^\mu\, \mathrm{d}\Omega = \oint_\omega A^\mu N_\mu \mathrm{d}\omega, \tag{2.249}$$

where $\Omega$ is the generalised (e.g., 4D) volume, $\omega$ is the closed surface bounding the volume, with generalised (e.g., 3D) surface element $N_\mu \mathrm{d}\omega$, where $N_\mu$ is the unit vector outwardly normal to the surface.

Similarly, for a vector function $\boldsymbol{A}$ within an arbitrary *open* surface $S$, the surface integral of the component of the curl of $\boldsymbol{A}$ normal to the surface is related to the line integral along $L$, the closed path which bounds the surface $S$ as

$$\int_S (\nabla \times \boldsymbol{A}) \cdot \mathrm{d}\boldsymbol{S} = \oint_L \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}, \tag{2.250}$$

where $\mathrm{d}\boldsymbol{l}$ is the line element of the path $L$. This is *Stokes' theorem.*

Another useful identity is Green's theorem,

$$\begin{aligned}\int_V \left(\phi\nabla^2\psi - \psi\nabla^2\phi\right) \mathrm{d}V &= \oint_S (\phi\nabla\psi - \psi\nabla\phi) \cdot \mathrm{d}\boldsymbol{S} \\ &= \oint_S \left(\phi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\phi}{\partial n}\right) \mathrm{d}|S|,\end{aligned} \tag{2.251}$$

which follows from Gauss' theorem, where $\mathrm{d}\boldsymbol{S} = \hat{\boldsymbol{n}}\mathrm{d}|S|$, and $\partial/\partial n$ is the (outward) normal derivative at the surface $S$.

In spherical (with $z = r\cos\theta$) and cylindrical coordinates, the Laplacian is

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}, \tag{2.252}$$

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2}. \tag{2.253}$$

A number of other identities are provided for convenience:

$$\begin{aligned}
&\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{a}\cdot\boldsymbol{c}) - \boldsymbol{c}(\boldsymbol{a}\cdot\boldsymbol{b}), && [(\nabla\cdot\boldsymbol{a})\boldsymbol{b}] = (\boldsymbol{a}\cdot\nabla)\boldsymbol{b} + (\boldsymbol{b}\cdot\nabla)\boldsymbol{a}, \\
&\nabla\cdot(\boldsymbol{a}\times\boldsymbol{b}) = \boldsymbol{b}\cdot\nabla\times\boldsymbol{a} - \boldsymbol{a}\cdot\nabla\times\boldsymbol{b}, && \nabla\cdot(\nabla\times\boldsymbol{a}) = \nabla\times(\nabla\phi) = 0, \\
&\nabla\times(\nabla\times\boldsymbol{a}) = \nabla(\nabla\cdot\boldsymbol{a}) - (\nabla\cdot\nabla)\boldsymbol{a}, && \nabla\times(\phi\boldsymbol{a}) = (\nabla\phi)\times\boldsymbol{a} + \phi(\nabla\times\boldsymbol{a}) \\
&\nabla\times(\boldsymbol{a}\times\boldsymbol{b}) = \boldsymbol{a}(\nabla\cdot\boldsymbol{b}) - (\boldsymbol{a}\cdot\nabla)\boldsymbol{b} + (\boldsymbol{b}\cdot\nabla)\boldsymbol{a} - \boldsymbol{b}(\nabla\cdot\boldsymbol{a}), && \\
&\nabla(\boldsymbol{a}\cdot\boldsymbol{b}) = (\boldsymbol{a}\cdot\nabla)\boldsymbol{b} + (\boldsymbol{b}\cdot\nabla)\boldsymbol{a} + \boldsymbol{a}\times(\nabla\times\boldsymbol{b}) + \boldsymbol{b}\times(\nabla\times\boldsymbol{a}). &&
\end{aligned} \tag{2.254}$$

With the index notation, and Einstein summation convention, we have

$$\nabla\cdot\boldsymbol{a} = \partial_i a^i \quad \text{and} \quad [\boldsymbol{a}\times\boldsymbol{b}]^i = \epsilon^{ijk}a^j b^k \tag{2.255}$$

where $\epsilon$ is the anti-symmetric Levi-Civita symbol (with $\epsilon^{123} \equiv +1$, and we avoid defining $\epsilon_{ijk}$). Finally, we use the standard convention for the indices of matrices:

$$T^{\mu\nu} = \begin{pmatrix} T^{00} & T^{01} & \dots & T^{0N} \\ T^{10} & T^{11} & \dots & T^{1N} \\ \vdots & \vdots & \ddots & \vdots \\ T^{N0} & T^{N1} & \dots & T^{NN} \end{pmatrix}. \tag{2.256}$$

### 2.11.4 Dirac delta function

The Dirac delta "function" is defined to have a value of zero everywhere except at the origin, and an integral over the entire real line equal to one. Technically, no *function* has this property, so the Dirac delta is more accurately considered a *distribution*, defined via its action on a well-behaved function, $f$:

$$\int_{-\infty}^{\infty} f(x)\delta(x)\,\mathrm{d}x = f(0). \tag{2.257}$$

The "well-behaved" function must be convergent over the domain in question.

While the definition (2.257) suffices in most circumstances, it is instructive to consider a possible functional representation for the delta. Consider the Gaussian:

$$g_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-x^2}{2\sigma^2}\right),$$

which is normalised $\int g_\sigma(x)\mathrm{d}x = 1$ for any $\sigma$. Consider the action of this function on a test function, $f$, under integration:

$$\int_{-\infty}^{\infty} g_\sigma(x)\,f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\frac{1}{\sigma}g_1(x/\sigma)\,f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} g_1(y)\,f(y\sigma)\,\mathrm{d}y, \tag{2.258}$$

where we made use of the fact that $g_\sigma(x) = \frac{1}{\sigma}g_1(x/\sigma)$, and then made the change of variables $x = y\sigma$. Now, take the limit as $\sigma \to 0$

$$\lim_{\sigma\to 0}\int_{-\infty}^{\infty} g_1(y)\,f(y\sigma)\,\mathrm{d}y = f(0)\int_{-\infty}^{\infty} g_1(y)\,\mathrm{d}y = f(0). \tag{2.259}$$

Therefore, comparing to Eq. (2.257), we may recognise the Gaussian function $g_\sigma$, in the limit of zero variance, as a *representation* of the delta function.

It is also useful to consider the Fourier transform

$$\widetilde{\delta}(k) \equiv \int \mathrm{d}x\,\delta(x)e^{-ikx} = \int \mathrm{d}x e^{-ik0} = 1.$$

Taking the inverse Fourier transform yields the Fourier (or integral) representation of the delta function:

$$\delta(x) = \int \frac{\mathrm{d}k}{2\pi}e^{ikx}. \tag{2.260}$$

Note that many different conventions for the Fourier transforms are common; some care must be taken when combining formulas from different sources.

Several useful properties of the Dirac delta function follow from the above definitions. For example, $\delta(ax) = \frac{1}{a}\delta(x)$. The derivative of a delta function

follows from integration by parts, noting the function $f$ is definitionally well behaved, meaning we may discard the $f(x)\delta(x)|_a^b$ boundary term:

$$\int \mathrm{d}x \, \frac{\mathrm{d}\delta(x)}{\mathrm{d}x} f(x) = -\int \mathrm{d}x \, \delta(x) \frac{\mathrm{d}f(x)}{\mathrm{d}x} = -f'(0). \tag{2.261}$$

Another useful relation, which may be proven in the same method as above is

$$\delta\left(f(x)\right) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|}, \tag{2.262}$$

where $x_i$ are the roots $f(x_i) = 0$. Finally, we note that it is common to use notion of the form $\delta^{(3)}(x) = \delta(\boldsymbol{x}) = \delta(x)\delta(y)\delta(z)$, etc.

### 2.11.5 Green's functions

Consider an inhomogeneous differential equation of the form

$$L_x u(x) = f(x), \tag{2.263}$$

where $L$ is a general *linear* operator, which acts on coordinate $x$ (e.g., $\partial^2$). If we could "invert" the operator, we could easily find solutions $u = L^{-1} f$.

To this end, we associate to $L$ a *Green's function*, $G$, defined via:

$$L_x G(x, x') \equiv \delta(x - x'). \tag{2.264}$$

Assuming $G$ is known, notice that we may write

$$\begin{aligned} \int \mathrm{d}x' L_x G(x, x') f(x') &= \int \mathrm{d}x' \delta(x - x') f(x') \\ L_x \int \mathrm{d}x' G(x, x') f(x') &= f(x). \end{aligned} \tag{2.265}$$

By comparing with Eq. (2.263), we recognise the remaining integral as $u(x)$. As such, presuming $G$ can be found by solving (2.264), we may easily find solutions to the inhomogeneous equation:

$$u(x) = \int \mathrm{d}x' G(x, x') f(x'), \tag{2.266}$$

called the *Green's function solution* to the differential equation.

The operator $L$ may also admit solutions $u_0$ to the *homogeneous equation*

$$L u_0(x) = 0.$$

Since the operator is linear, the general solution to Eq. (2.263) can be written as the sum $u(x) = u_0(x) + u_*(x)$, where $u_*(x)$ is a particular solution satisfying $Lu_*(x) = f(x)$, which is given by the Green's function expression (2.266). Therefore, the *general solution* to the inhomogeneous equation is:

$$u(x) = u_0(x) + \int \mathrm{d}x'\, G(x, x') f(x'). \tag{2.267}$$

The homogeneous component $u_0(x)$ typically includes arbitrary parameters that are fixed by the boundary or initial conditions.

As a concrete (and pertinent) example, we'll consider the 3D Poisson equation, $\nabla^2\phi = -\rho$, for which the corresponding Green's function satisfies

$$\nabla^2 G = \delta^{(3)}(\boldsymbol{x} - \boldsymbol{x}'). \tag{2.268}$$

Noting the spherical symmetry of the Laplacian and the delta, we conclude that $G = G(r)$, where $r = |\boldsymbol{x} - \boldsymbol{x}'|$. It is therefore also convenient to write the Laplacian in spherical coordinates. For $r \neq 0$, we have

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}G}{\mathrm{d}r}\right) = 0.$$

The simplest way to solve this is by defining $\chi = rG$, in which case the equation becomes $\chi'' = 0$. Integrating twice, we see $\chi = a + br$, or

$$G = \frac{a}{r} + b. \tag{2.269}$$

Requiring that $G \to 0$ as $r \to \infty$ allows us to set $b = 0$. We find $a$ by integrating Eq. (2.268) over a sphere of radius $R$ around $r = 0$ using Gauss' theorem (2.248):

$$\int (\nabla^2 G)\, \mathrm{d}^3 r = \oint_S (\nabla G) \cdot \mathrm{d}\boldsymbol{S} = 1,$$

noting the delta function integrates to 1. We have $\nabla\frac{1}{r} = -\frac{1}{r^3}\boldsymbol{r}$, so

$$-a \oint_S \frac{1}{r^2}(\hat{\boldsymbol{n}} \cdot \mathrm{d}\boldsymbol{S}) = -a\frac{1}{R^2}\oint_S (\hat{\boldsymbol{n}} \cdot \mathrm{d}\boldsymbol{S}) = -a\frac{1}{R^2}(4\pi R^2) = 1,$$

where the radial unit vector $\hat{\boldsymbol{n}}$ is parallel to the surface element $\mathrm{d}\boldsymbol{S}$, so integral is just the surface area of the sphere. Thus, we find the Green's function:

$$G(\boldsymbol{x}, \boldsymbol{x}') = \frac{-1}{4\pi|\boldsymbol{x} - \boldsymbol{x}'|}. \tag{2.270}$$

Finally, we have the solution to the Poisson equation:

$$\phi(\boldsymbol{x}) = \frac{1}{4\pi}\int \frac{\rho(\boldsymbol{x})}{|\boldsymbol{x} - \boldsymbol{x}'|}\mathrm{d}^3 x'. \tag{2.271}$$

It's important to note that the spherical symmetry argument that applied for the Green's function certainly does not apply (in general) for the Poisson equation.

### 2.11.6 Retarded (and advanced) Green's functions: Fourier method

Following from Sec. 2.9.1, we shall derive the retarded Green's functions, this time using the method of Fourier transforms. A familiarity with complex contour integration is assumed. We work directly in the covariant form, and set $c = 1$.

Starting from the inhomogeneous d'Alembert equation (2.182), we associate the Green's function and its Fourier representation:

$$\begin{aligned} \partial^2 G &= \delta^{(4)}(x - x'), \\ G(x - x') &= \int \frac{d^4k}{(2\pi)^4} \widetilde{G}(k) e^{-ik\cdot(x-x')}. \end{aligned} \tag{2.272}$$

from which we find $\widetilde{G}(k) = \frac{-1}{k\cdot k}$. Taking the inverse Fourier transform ($\omega \equiv k_0$):

$$G(x - x') = \frac{-1}{2\pi} \int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} \int \mathrm{d}\omega \frac{e^{-i\omega(t-t')}}{\omega^2 - \boldsymbol{k}^2}. \tag{2.273}$$

We evaluate the $\omega$ integral first, by extending into the complex plane and using the Cauchy residue theorem. There are two simple poles, at $\omega = \pm|\boldsymbol{k}|$. These are along the real axis; for the integral to converge, we must shift the contour to avoid the poles.[26] There are several choices for avoiding the poles: for example, we can go above or below them as shown in Fig. 2.5.

First, consider the red "above" contour. For $t > t'$, Jordan's Lemma tells us the integral (2.273) goes to zero around a large arc in the lower plane, so we close the contour below. Both poles are enclosed, so:

$$\oint \mathrm{d}\omega \frac{e^{-i\omega(t-t')}}{\omega^2 - \boldsymbol{k}^2} = -\frac{2\pi}{|k|} \sin(|\boldsymbol{k}|(t - t')).$$

For $t < t'$, on the other hand, we close the contour in the upper plane. Since neither pole is enclosed, the integral is zero. We eventually find ($r \equiv |\boldsymbol{x} - \boldsymbol{x}'|$):

$$\begin{aligned} G_r(x - x') &= \int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} \frac{\sin(|\boldsymbol{k}|(t - t'))}{|\boldsymbol{k}|} \theta(t - t') \\ &= \frac{\theta(t - t')}{4\pi\, r} \int_{-\infty}^{\infty} \frac{\mathrm{d}|\boldsymbol{k}|}{2\pi} \left[ e^{i|\boldsymbol{k}|(t-t'-r)} - e^{i|\boldsymbol{k}|(t-t'+r)} \right], \end{aligned} \tag{2.274}$$

where the Heaviside theta function ensures the $t < t'$ part is zero. The '$r$' subscript is to remind us that this was the solution from choosing the *r*ed contour; as we shall see in a moment, we can also realise this is will be the *r*etarded Green's function (nonzero only for times $t > t'$).

[26] It is not obvious this is a sensible thing to do. The integral does not converge, and changing the contour changes its value. In the end, we must check if the result is a valid Green's function.

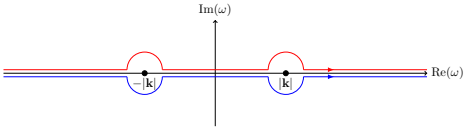


Figure 2.5: Contour plot for the $\omega$ integral in Eq. (2.273).

The above integrals are just the Fourier representation of the 1D delta functions, and only the first survives (since $\Delta t > 0$, enforced by the $\theta$ term, and $r > 0$ by definition). As such, we finally find:

$$G_r(x - x') = \frac{1}{4\pi r}\delta(t - t' - r), \tag{2.275}$$

which is the retarded Green's function (we dropped the $\theta$ term, since the delta function, which implies $t = t' + r$, already forces $\Delta t = t - t' > 0$). In exactly the same way, if we choose the blue "lower" contour, we would find the *advanced* Green's function:

$$G_a(x - x') = \frac{1}{4\pi r}\delta(t - t' + r). \tag{2.276}$$

These can be integrated with the source term in the inhomogeneous d'Alembert equation (2.182) to immediately yield the advanced and retarded potentials.

### 2.11.7 Fields from the Liénard-Wiechert potential

The Liénard-Wiechert (LW) potentials are (see Sec. 2.9.2):

$$\Phi = \frac{q}{4\pi}\frac{1}{\mathcal{D}(t_r)}, \quad \text{and} \quad \boldsymbol{A} = \frac{\boldsymbol{v}(t_r)}{c}\Phi,$$

where we defined $\boldsymbol{n} = \boldsymbol{R}/R$, and

$$\mathcal{D} \equiv R(1 - \tfrac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}) \equiv R\,\eta. \tag{2.277}$$

The fields may be derived in the usual way, though particular care must be taken with the derivatives. The fields (2.126) depend on derivatives with respect to the regular $t, \boldsymbol{x}$ coordinates, while the potentials are in terms of the retarded quantities (2.206). In other words, we have:

$$\frac{\mathrm{d}}{\mathrm{d}x} = \frac{\partial}{\partial x} + \nabla t_r \frac{\partial}{\partial t_r} \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial t_r}{\partial t}\frac{\partial}{\partial t_r}.$$

First, to determine $\partial t_r/\partial t$, we calculate the time derivative of $R = c(t - t_r)$:

$$\frac{\partial R}{\partial t} = c\left(1 - \tfrac{\partial t_r}{\partial t}\right). \tag{2.278}$$

At the same time, from the identity $R^2 = \boldsymbol{R}\cdot\boldsymbol{R}$, we have:

$$\frac{\partial R}{\partial t} = \frac{\partial R}{\partial t_r}\frac{\partial t_r}{\partial t} = \frac{1}{2R}\frac{\partial R^2}{\partial t_r}\frac{\partial t_r}{\partial t} \quad \text{with} \quad \frac{\partial R^2}{\partial t_r} = 2\boldsymbol{R}\cdot\frac{\partial \boldsymbol{R}}{\partial t_r} = -2\boldsymbol{R}\cdot\boldsymbol{v}, \tag{2.279}$$

and we remind that, since the expressions in the LW potentials are evaluated at the retarded time, $\boldsymbol{v} \equiv \partial \boldsymbol{x}_0/\partial t_r$. Combining these, we find

$$\frac{\partial t_r}{\partial t} = \frac{R}{\mathcal{D}}. \tag{2.280}$$

To find the spatial derivative $\nabla t_r$, we calculate:

$$\nabla R = \frac{1}{2R}\nabla R^2 \quad \text{using} \quad \frac{\mathrm{d}R^2}{\mathrm{d}x^i} = \frac{\partial R^2}{\partial x^i} + \frac{\partial R^2}{\partial t_r}\frac{\partial t_r}{\partial x^i},$$

with $\partial R^2/\partial x^i = 2R_i$, which leads to $\nabla R = \boldsymbol{R}/\mathcal{D}$. At the same time, from $R = c(t - t_r)$, we have $\nabla R = -c\nabla t_r$. Combining these leads to

$$\nabla t_r = -\frac{1}{c}\frac{\boldsymbol{R}}{\mathcal{D}}. \tag{2.281}$$

The final ingredients required are the derivatives of the denominator $\mathcal{D}$. Using the above relations, these can be found without too much difficulty:

$$\nabla\mathcal{D} = \nabla R - \tfrac{1}{c}\nabla(\boldsymbol{R}\cdot\boldsymbol{v}) = \frac{\boldsymbol{R}}{\mathcal{D}} - \frac{\boldsymbol{v}}{c} + \frac{1}{c^2}\frac{\boldsymbol{R}}{\mathcal{D}}\left(\boldsymbol{R}\cdot\boldsymbol{a} - v^2\right), \tag{2.282}$$

with acceleration $\boldsymbol{a} \equiv \partial\boldsymbol{v}/\partial t_r$, and

$$\frac{\partial\mathcal{D}}{\partial t_r} = -\frac{\boldsymbol{v}\cdot\boldsymbol{R}}{R} + \tfrac{1}{c}v^2 - \tfrac{1}{c}\boldsymbol{R}\cdot\boldsymbol{a}. \tag{2.283}$$

With these, we have for the derivatives of the potentials:

$$\nabla\Phi = -\frac{q}{4\pi}\frac{1}{\mathcal{D}^2}\nabla\mathcal{D}, \quad \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = \frac{1}{c}\frac{R}{\mathcal{D}}\left(\boldsymbol{v}\frac{\partial\Phi}{\partial t_r} + \Phi\boldsymbol{a}\right), \text{ with } \frac{\partial\Phi}{\partial t_r} = -\frac{q}{4\pi}\frac{1}{\mathcal{D}^2}\frac{\partial\mathcal{D}}{\partial t_r},$$

and we eventually come to

$$\begin{aligned}\boldsymbol{E} &= \frac{q}{4\pi\,R^2}\frac{\left(\boldsymbol{n} - \frac{1}{c}\boldsymbol{v}\right)}{\gamma^2\,\eta^3} + \frac{q}{4\pi\,c\,R}\frac{\left[\frac{1}{c}\boldsymbol{a}(\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n} - 1) + (\boldsymbol{n} - \frac{1}{c}\boldsymbol{v})(\frac{1}{c}\boldsymbol{a}\cdot\boldsymbol{n})\right]}{\eta^3}, \\ &= \frac{q}{4\pi\,R^2}\frac{\left(\boldsymbol{n} - \frac{1}{c}\boldsymbol{v}\right)}{\gamma^2\,\eta^3} + \frac{q}{4\pi\,c\,R}\frac{\boldsymbol{n}\left[\times(\boldsymbol{n} - \frac{1}{c}\boldsymbol{v})\times\frac{1}{c}\boldsymbol{a}\right]}{\eta^3},\end{aligned} \tag{2.284}$$

and

$$\boldsymbol{B} = \boldsymbol{n}\times\boldsymbol{E}, \tag{2.285}$$

where all quantities on the right-hand-side are evaluated at the retarded time.

# References

[1] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (1971).

[2] D. J. Griffiths, *Intoduction to Electrodynamics* (1999).

[3] J. D. Jackson, *Classical Electrodynamics* (2001).

[4] A. Einstein, *Zur Elektrodynamik bewegter Körper*, Ann. Phys. **322**, 891 (1905).

[5] A. Einstein, *On the Electrodynamics of Moving Bodies (Translation)*, `https://www.fourmilab.ch/etexts/einstein/specrel/www/`.

[6] A. Einstein, *Ist Die Trägheit Eines Körpers von Seinem Energieinhalt Abhängig?*, Ann. Phys. **323**, 639 (1905).

[7] A. Einstein, *Does the Inertia of a Body Depend upon its Energy-Content? (Translation)*, `https://www.fourmilab.ch/etexts/einstein/E_mc2/www/`.

[8] E. Noether, *Invariante Variationsprobleme*, Gott. Nachr. **1918**, 235 (1918).

[9] E. Noether, *Invariant Variation Problems (Translation)*, arXiv:physics/0503066.

[10] H. Yukawa, *On the interaction of elementary particles*, Prog. Theor. Phys. **1**, 1 (1955).

[11] J. C. Maxwell, *A dynamical theory of the electromagnetic field*, Philos. Trans. **155**, 459 (1865).