Chapter 3

Electromagnetism and Gravitation in Various Dimensions

As a candidate theory of all interactions, string theory includes Maxwell electrodynamics and its nonlinear cousins, as well as gravitation. We review the relativistic formulation of four-dimensional electrodynamics and show how it facilitates the definition of electrodynamics in other dimensions. We give a brief description of Einstein’s gravity and use the Newtonian limit to discuss the relation between Planck’s length and the gravitational constant in various dimensions. We study the effect of compactification on the gravitational constant and explain how large extra dimensions could escape detection.

3.1 Classical electrodynamics

Unlike Newtonian mechanics, classical electrodynamics is a relativistic theory. In fact, Einstein was led by considerations of electrodynamics to formulate the special theory of relativity. Electrodynamics has a particularly elegant formulation in which the relativistic character of the theory is manifest. This relativistic formulation allows a natural extension of the theory to higher dimensions. Before we discuss the relativistic formulation we must review Maxwell’s equations. These equations describe the dynamics of electric and magnetic fields.

Although most undergraduate and graduate courses in electrodynamics nowadays use the international system of units (SI units), the Heaviside-Lorentz system of units is far more convenient for discussions that involve relativity and extra dimensions. In this system
of units, Maxwell’s equations take the following form:

\[
\begin{align*}
\nabla \times \vec{E} &= -\frac{1}{c} \frac{\partial \vec{B}}{\partial t}, \\
\nabla \cdot \vec{B} &= 0, \\
\n\nabla \cdot \vec{E} &= \rho, \\
\n\nabla \times \vec{B} &= \frac{1}{c} \vec{j} + \frac{1}{c} \frac{\partial \vec{E}}{\partial t}.
\end{align*}
\]

(3.1.1) (3.1.2) (3.1.3) (3.1.4)

The above equations imply that \( \vec{E} \) and \( \vec{B} \) are measured with the same units. The first two equations are the source-free Maxwell equations. The second two involve sources: the charge density \( \rho \), with units of charge per unit volume, and the current density \( \vec{j} \), with units of current per unit area. The Lorentz force law, which gives the rate of change of the relativistic momentum of a charged particle in an electromagnetic field, takes the form

\[
\frac{dp}{dt} = q \left( \vec{E} + \frac{\vec{v}}{c} \times \vec{B} \right).
\]

(3.1.5)

Since the magnetic field \( \vec{B} \) is divergenceless, it can be written as the curl of a vector, the well-known vector potential \( \vec{A} \):

\[
\vec{B} = \nabla \times \vec{A}.
\]

(3.1.6)

In electrostatics the electric field \( \vec{E} \) has zero curl, and it is therefore written as (minus) the gradient of a scalar, the well-known scalar potential \( \Phi \). In electrodynamics, as equation (3.1.1) indicates, the curl of \( \vec{E} \) is not always zero. Substituting (3.1.6) into (3.1.1), we find a linear combination of \( \vec{E} \) and of the time derivative of \( \vec{A} \) that has zero curl:

\[
\nabla \times \left( \vec{E} + \frac{1}{c} \frac{\partial \vec{A}}{\partial t} \right) = 0.
\]

(3.1.7)

The object inside parenthesis is set equal to \(-\nabla \Phi\), and the electric field \( \vec{E} \) can be written in terms of the scalar potential and the vector potential:

\[
\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \nabla \Phi.
\]

(3.1.8)

Equations (3.1.6) and (3.1.8) express the electric and magnetic fields in terms of potentials. By doing so, the source-free Maxwell equations (3.1.1) and (3.1.2) are automatically satisfied. Equations (3.1.3) and (3.1.4) contain additional information. They are used to derive equations for \( \vec{A} \) and \( \Phi \).
3.2 Electromagnetism in three dimensions

What is electromagnetism in three spacetime dimensions? One way to produce a theory of electromagnetism in three dimensions is to begin with the four-dimensional theory and eliminate one spatial coordinate. This procedure is called dimensional reduction.

In four spacetime dimensions, both electric and magnetic fields have three spatial components: \((E_x, E_y, E_z)\) and \((B_x, B_y, B_z)\), respectively. It may seem likely that a reduction to a world without a \(z\) coordinate would require dropping the \(z\) components from the two fields. Surprisingly, this does not work! Maxwell’s equations and the Lorentz force law make it impossible.

In order to construct a consistent three-dimensional theory, we must ensure that the dynamics does not depend on the \(z\) direction, the direction that we want to eliminate. If there is motion, it must remain restricted to the \((x, y)\) plane. It is thus natural to require that no quantity should have \(z\)-dependence. This does not necessarily mean dropping quantities with a \(z\) index.

The Lorentz force law (3.1.5) is a useful guide to the construction of the lower-dimensional theory. Suppose that there is no magnetic field. Then, in order to keep the \(z\) component of momentum equal to zero we must have \(E_z = 0\); the \(z\) component of the electric field must go. The case of the magnetic field is more surprising. Assume that the electric field is zero. If the velocity of the particle is a vector in the \((x, y)\) plane, a component of the magnetic field in the plane would generate, via the cross-product, a force in the \(z\) direction. On the other hand, a \(z\) component of the magnetic field would generate a force in the \((x, y)\) plane!

We conclude that \(B_x\) and \(B_y\) must be set equal to zero, while we can keep \(B_z\). All in all,

\[
E_z = B_x = B_y = 0. 
\] (3.2.1)

The left-over fields \(E_x, E_y,\) and \(B_z\) can only depend on \(x\) and \(y\). In the three-dimensional world with coordinates \(t, x,\) and \(y\), the \(z\) index of \(B_z\) is not a vector index. Therefore, in this reduced world, \(B_z\) behaves like a Lorentz scalar (more precisely, it is an object called a pseudo-scalar). In summary, we have a two-dimensional vector \(\vec{E}\) and a scalar field \(B_z\).

We can test the consistency of this truncation by taking a look at the \(x\) and \(y\) components of (3.1.1):

\[
\begin{align*}
\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} &= \frac{1}{c} \frac{\partial B_x}{\partial t}, \\
\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} &= \frac{1}{c} \frac{\partial B_y}{\partial t}.
\end{align*}
\] (3.2.2)

Since the right-hand sides are set to zero by our truncation, the left-hand sides should vanish as well. Indeed, they do. Each term on the left-hand sides equals zero, either because it contains an \(E_z\), or because it contains a \(z\) derivative. You may examine the consistency of the remaining equations in Problem 3.3.

While setting up three-dimensional electrodynamics was not too difficult, it is much harder to guess what five-dimensional electrodynamics should be. As we will see next, the
manifestly relativistic formulation of Maxwell’s equations immediately gives the appropriate generalization to other dimensions.

### 3.3 Manifestly relativistic electrodynamics

In the relativistic formulation of Maxwell’s equations neither the electric field nor the magnetic field becomes part of a four-vector. Rather, a four-vector is obtained by combining the scalar potential \( \Phi \) with the vector potential \( \vec{A} \):

\[
A^\mu = (\Phi, A^1, A^2, A^3) .
\]  

(3.3.1)

The corresponding object with down indices is

\[
A_\mu = (-\Phi, A^1, A^2, A^3) .
\]  

(3.3.2)

From \( A_\mu \), we create an object known as the electromagnetic field strength \( F_{\mu\nu} \):

\[
F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu .
\]  

(3.3.3)

Here \( \partial_\mu \equiv \frac{\partial}{\partial x^\mu} \). Equation (3.3.3) implies that \( F_{\mu\nu} \) is antisymmetric:

\[
F_{\mu\nu} = -F_{\nu\mu} .
\]  

(3.3.4)

It follows from this property that all diagonal components of \( F_{\mu\nu} \) vanish:

\[
F_{00} = F_{11} = F_{22} = F_{33} = 0 .
\]  

(3.3.5)

Let us calculate a few entries in \( F_{\mu\nu} \). Let \( i \) denote a spatial index, that is, an index that can take the values 1, 2, and 3. Making use of (3.3.3) and (3.1.8), we find

\[
F_{0i} = \frac{\partial A_i}{\partial x^0} - \frac{\partial A_0}{\partial x^i} = \frac{1}{c} \frac{\partial A_i}{\partial t} + \frac{\partial \Phi}{\partial x^i} = -E_i .
\]  

(3.3.6)

Similarly, we can calculate \( F_{12} \):

\[
F_{12} = \partial_1 A_2 - \partial_2 A_1 = \partial_x A_y - \partial_y A_x = B_z ,
\]  

(3.3.7)

since \( \vec{B} = \nabla \times \vec{A} \). Continuing in this manner, we can compute all the entries in the matrix \( F_{\mu\nu} \):

\[
F_{\mu\nu} = \begin{pmatrix}
0 & -E_x & -E_y & -E_z \\
E_x & 0 & B_z & -B_y \\
E_y & -B_z & 0 & B_x \\
E_z & B_y & -B_x & 0
\end{pmatrix} .
\]  

(3.3.8)

We see that the electric and magnetic fields \( \vec{E} \) and \( \vec{B} \) are encoded in the field strength \( F_{\mu\nu} \).
The potentials \( A_\mu \) can be changed by \textit{gauge transformations}. Two sets of potentials \( A_\mu \) and \( A'_\mu \) that are related by gauge transformations are physically equivalent. A necessary (but not always sufficient) condition for physical equivalence of the potentials \( A_\mu \) and \( A'_\mu \) is that they give identical electric and magnetic fields or, on account of (3.3.8), identical field strengths. Gauge transformations take the form

\[
A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \epsilon .
\]  

(3.3.9)

Here \( A_\mu \) and \( A'_\mu \) are the gauge related potentials, and \( \epsilon(x) \) is an arbitrary function of the spacetime coordinates. The field strength \( F_{\mu\nu} \) is gauge \textit{invariant}, as we readily verify:

\[
F_{\mu\nu} \rightarrow F'_{\mu\nu} \equiv \partial_\mu A'_\nu - \partial_\nu A'_\mu ,
\]

\[
= \partial_\mu (A_\nu + \partial_\nu \epsilon) - \partial_\nu (A_\mu + \partial_\mu \epsilon) ,
\]

\[
= F_{\mu\nu} + \partial_\mu \partial_\nu \epsilon - \partial_\nu \partial_\mu \epsilon ,
\]

\[
= F_{\mu\nu} .
\]  

(3.3.10)

In the last step we noted that partial derivatives commute. We can write the gauge transformations more explicitly in component form. Using (3.3.9) and (3.3.2), we find

\[
\Phi \rightarrow \Phi' = \Phi - \frac{1}{c} \frac{\partial \epsilon}{\partial t} ,
\]

\[
\vec{A} \rightarrow \vec{A}' = \vec{A} + \nabla \epsilon .
\]  

(3.3.11)

The gauge transformation of \( \vec{A} \) is familiar; adding a gradient to a vector does not change its curl, so \( \vec{B} = \nabla \times \vec{A} \) is unchanged. The scalar potential \( \Phi \) also changes under gauge transformations. This is necessary to keep \( \vec{E} \) unchanged.

\textit{Quick Calculation 3.1.} Verify that \( \vec{E} \), as given in (3.1.8), is invariant under the gauge transformations (3.3.11).

Recall that the use of potentials to represent \( \vec{E} \) and \( \vec{B} \) automatically solves the source-free Maxwell equations (3.1.1) and (3.1.2). How are these equations written in terms of the field strength \( F_{\mu\nu} \)? They must be written so that they hold when (3.3.3) holds. Consider the following combination of field strengths:

\[
T_{\lambda\mu\nu} \equiv \partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} .
\]  

(3.3.12)

\( T_{\lambda\mu\nu} \) vanishes identically on account of (3.3.3):

\[
\partial_\lambda (\partial_\mu A_\nu - \partial_\nu A_\mu) + \partial_\mu (\partial_\nu A_\lambda - \partial_\lambda A_\nu) + \partial_\nu (\partial_\lambda A_\mu - \partial_\mu A_\lambda) = 0 ,
\]  

(3.3.13)

using the commutativity of partial derivatives. The vanishing of \( T_{\lambda\mu\nu} \),

\[
\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 ,
\]  

(3.3.14)
is a set of differential equations for the field strength. These equations are precisely the source-free Maxwell equations. To make this clear, first note that \( T_{\lambda\mu} \) satisfies the antisymmetry conditions

\[
T_{\lambda\mu} = -T_{\mu\lambda}, \quad T_{\lambda\mu} = -T_{\lambda\nu}.
\]  

These two equations follow from (3.3.12) and the antisymmetry property \( F_{\mu\nu} = -F_{\nu\mu} \) of the field strength. They state that \( T \) changes sign under the transposition of any two adjacent indices.

**Quick Calculation 3.2.** Verify the equations in (3.3.15).

Any object, with however many indices, that changes sign under the transposition of every pair of adjacent indices, will change sign under the transposition of any two indices: to exchange any two indices you need an odd number of transpositions of adjacent indices (do you see why?). An object that changes sign under the transposition of any two indices is said to be totally antisymmetric. Therefore \( T \) is totally antisymmetric.

Since \( T \) is totally antisymmetric, it vanishes when any two of its indices take the same value. \( T \) is non-vanishing only when each of its three indices takes a different value. In such case, different orderings of these three fixed values will give \( T \) components that can differ at most by a sign. Since we are setting \( T \) to zero these various orderings do not give new conditions. Because we have four spacetime coordinates, selecting three different indices can only be done in four different ways – leaving out a different index each time. Thus the vanishing of \( T \) gives four nontrivial equations. These four equations are the three components of equation (3.1.1) and equation (3.1.2). The vanishing of \( T_{012} \), for example, gives us

\[
\partial_0 F_{12} + \partial_1 F_{20} + \partial_2 F_{01} = \frac{1}{c} \frac{\partial B_z}{\partial t} + \frac{\partial E_y}{\partial x} - \frac{\partial E_z}{\partial y} = 0.
\]  

(3.3.16)

This is the \( z \) component of equation (3.1.1). The other three choices of indices lead to the remaining three equations (Problem 3.2).

How can we describe Maxwell equations (3.1.3) and (3.1.4) in our present framework? Since these equations have sources, we must introduce a *current* four-vector:

\[
j^\mu = (c\rho, j^1, j^2, j^3),
\]  

(3.3.17)

where \( \rho \) is the charge density and \( \vec{j} = (j^1, j^2, j^3) \) is the current density. In addition, we raise the indices of the field tensor to obtain the field tensor with upper indices:

\[
F^{\mu\nu} = \eta^{\mu\alpha} \eta^{\nu\beta} F_{\alpha\beta}.
\]  

(3.3.18)

**Quick Calculation 3.3.** Show that

\[
F^{\mu\nu} = -F^{\nu\mu}, \quad F^{0i} = -F_{0i}, \quad F^{ij} = F_{ij}.
\]  

(3.3.19)

Equation (3.3.18), together with the definition of \( F_{\mu\nu} \), gives

\[
F_{\mu\nu} = \eta^{\mu\alpha} \eta^{\nu\beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha) = \eta^{\mu\alpha} \partial_\alpha (\eta^{\nu\beta} A_\beta) - \eta^{\mu\beta} \partial_\beta (\eta^{\nu\alpha} A_\alpha),
\]  

(3.3.20)
where the constancy of the metric components was used to move them across the derivatives. It is customary to apply the rules for raising and lowering indices to partial derivatives, so we write $\partial^{\mu} \equiv \eta^{\mu \alpha} \partial_{\alpha}$. As a result,

$$F^{\mu \nu} = \partial^{\mu} A^{\nu} - \partial^{\nu} A^{\mu}.$$  

(3.3.21)

It follows from (3.3.19) and (3.3.8) that

$$F^{\mu \nu} = \begin{pmatrix} 0 & E_{x} & E_{y} & E_{z} \\ -E_{x} & 0 & B_{y} & -B_{x} \\ -E_{y} & -B_{z} & 0 & B_{y} \\ -E_{z} & B_{y} & -B_{x} & 0 \end{pmatrix}.$$  

(3.3.22)

Using this equation and the current vector (3.3.17), we can encapsulate Maxwell’s equations (3.1.3) and (3.1.4) as (Problem 3.2)

$$\frac{\partial F^{\mu \nu}}{\partial x^{\nu}} = \frac{1}{c} j^{\mu}.$$  

(3.3.23)

In the absence of sources this equation becomes

$$\partial_{\nu} F^{\mu \nu} = 0 \quad \rightarrow \quad \partial_{\nu} \partial^{\mu} A^{\nu} - \partial^{2} A^{\mu} = 0,$$  

(3.3.24)

where we have written $\partial^{2} = \partial^{\mu} \partial_{\mu}$.

Equations (3.3.3) together with equations (3.3.23) are equivalent to Maxwell’s equations in four dimensions. We will take these equations to define Maxwell theory in arbitrary dimensions. In $d$ spatial dimensions the Lorentz vector $A^{\mu}$ has components $(\Phi, \vec{A})$ where $\vec{A}$ is a $d$-dimensional spatial vector.

In three-dimensional spacetime, for example, the matrix $F_{\mu \nu}$ is a 3-by-3 antisymmetric matrix, obtained from (3.3.8) by discarding the last row and the last column:

$$F_{\mu \nu} = \begin{pmatrix} 0 & -E_{x} & -E_{y} \\ E_{x} & 0 & B_{z} \\ E_{y} & -B_{z} & 0 \end{pmatrix}.$$  

(3.3.25)

This immediately reproduces the main result of §3.2; $B_{x}, B_{y},$ and $E_{z}$ are to be set to zero.

Motivated by (3.3.22), in arbitrary dimensions we will call $F_{0i}$ the electric field $E_{i}$:

$$E_{i} \equiv F_{0i} = -F_{0i}.$$  

(3.3.26)

The electric field is a spatial vector. Equation (3.3.6) implies that, in any number of dimensions,

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \nabla \Phi.$$  

(3.3.27)
The magnetic field is identified with the $F^{ij}$ components of the field strength. In four-dimensional spacetime $F^{ij}$ is a 3-by-3 antisymmetric matrix. Its three independent entries are the components of the magnetic field vector (see (3.3.22)). In dimensions other than four, the magnetic field is no longer a spatial vector. In three spacetime dimensions the magnetic field is a single-component object. In five spacetime dimensions the magnetic field has as many entries as a 4-by-4 antisymmetric matrix, six entries. That many components do not fit into a spatial vector.

Our next goal is the determination of the electric field produced by a point charge in a spacetime with an arbitrary but fixed number of spatial dimensions. To this end, we must first learn how to calculate the volumes of higher-dimensional spheres. We turn to this subject now.

### 3.4 An aside on spheres in higher dimensions

Since we want to work in various numbers of dimensions we should be precise when speaking about spheres and their volumes. When we speak loosely we tend to confuse spheres and balls, at least in the precise sense in which they are defined in mathematics. When you say that the volume of a sphere of radius $R$ is $\frac{1}{2} \pi R^3$, you should really be saying that this is the volume of the three-ball $B^3$ – the three-dimensional space enclosed by the two-dimensional two-sphere $S^2$. In three-dimensional space $\mathbb{R}^3$ with coordinates $x_1, x_2,$ and $x_3$, we write the three-ball as the region defined by

$$B^3(R) : x_1^2 + x_2^2 + x_3^2 \leq R^2.$$  \hspace{1cm} (3.4.1)

This region is enclosed by the two-sphere:

$$S^2(R) : x_1^2 + x_2^2 + x_3^2 = R^2.$$  \hspace{1cm} (3.4.2)

The superscripts in $B$ or $S$ denote the dimensionality of the space in question. When we drop the explicit argument $R$, we mean that $R = 1$. Lower-dimensional examples are also familiar. $B^2$ is a two dimensional disk – the region enclosed in $\mathbb{R}^2$ by the one-dimensional unit radius circle $S^1$. In arbitrary dimensions we define balls and spheres as subspaces of $\mathbb{R}^d$:

$$B^d(R) : x_1^2 + x_2^2 + \ldots + x_d^2 \leq R^2.$$  \hspace{1cm} (3.4.3)

This is the region enclosed by the sphere $S^{d-1}(R)$:

$$S^{d-1}(R) : x_1^2 + x_2^2 + \ldots + x_d^2 = R^2.$$  \hspace{1cm} (3.4.4)

One last piece of terminology: to avoid confusion we will always speak of volumes. If a space is one-dimensional we take volume to mean length. If a space is two-dimensional we take volume to mean area. All higher dimensional spaces have just volumes. The volumes of the one- and two-dimensional spheres are

$$\text{vol} (S^1(R)) = 2\pi R,$$

$$\text{vol} (S^2(R)) = 4\pi R^2.$$  \hspace{1cm} (3.4.5)
3.4. AN ASIDE ON SPHERES IN HIGHER DIMENSIONS

Unless you have worked with other spheres before, you probably do not know what the volume of \( S^3 \) is.

Since volume has units of length to the power of the space dimension, the volume of a sphere of radius \( R \) is related to the volume of a sphere of unit radius by

\[
\text{vol} \left( S^{d-1}(R) \right) = R^{d-1} \text{vol}(S^{d-1}) .
\] (3.4.6)

Since the radius dependence of the volume is easily recovered, it suffices to record the volumes of unit spheres:

\[
\begin{align*}
\text{vol} (S^1) &= 2\pi , \\
\text{vol} (S^2) &= 4\pi . 
\end{align*}
\] (3.4.7)

Let’s now begin our calculation of the volume of the sphere \( S^{d-1} \). For this purpose consider \( \mathbb{R}^d \) with coordinates \( x_1, x_2, \ldots, x_d \), and let \( r \) be the radial coordinate:

\[
r^2 = x_1^2 + x_2^2 + \ldots + x_d^2 .
\] (3.4.8)

We will find the desired volume by evaluating in two different ways the following integral:

\[
I_d = \int_{\mathbb{R}^d} dx_1 dx_2 \ldots dx_d e^{-r^2} .
\] (3.4.9)

First we proceed directly. Using (3.4.8) in the exponential factor, the integral becomes a product of \( d \) gaussian integrals:

\[
I_d = \prod_{i=1}^{d} \int_{-\infty}^{\infty} dx_i e^{-x_i^2} = (\sqrt{\pi})^d = \pi^{d/2} .
\] (3.4.10)

Now we proceed indirectly. We do the integral by breaking \( \mathbb{R}^d \) into thin spherical shells. Since the space of constant \( r \) is the sphere \( S^{d-1}(r) \), the volume of a shell lying between \( r \) and \( r + dr \) equals the volume of \( S^{d-1}(r) \) times \( dr \). Therefore,

\[
I_d = \int_{0}^{\infty} dr \text{vol}(S^{d-1}(r)) e^{-r^2} = \text{vol}(S^{d-1}) \int_{0}^{\infty} dr r^{d-1} e^{-r^2}
\]

\[
= \frac{1}{2} \text{vol}(S^{d-1}) \int_{0}^{\infty} dt e^{-t} t^{\frac{d}{2}-1} ,
\] (3.4.11)

where use was made of (3.4.6), and in the final step we changed the variable of integration to \( t = r^2 \). The last integral on the right-hand side can be expressed in terms of the gamma function, a very useful special function. For positive \( x \) the gamma function \( \Gamma(x) \) is defined by

\[
\Gamma(x) = \int_{0}^{\infty} dt e^{-t} t^{x-1} , \quad x > 0 .
\] (3.4.12)
Unless \( x > 0 \) the integral does not converge near \( t = 0 \). With this definition, equation (3.4.11) becomes

\[
I_d = \frac{1}{2} \text{vol} (S^{d-1}) \Gamma \left( \frac{d}{2} \right).
\]

(3.4.13)

Comparing with the earlier evaluation (3.4.10), we get our final result:

\[
\text{vol} (S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma \left( \frac{d}{2} \right)}.
\]

(3.4.14)

It now remains to calculate the value of \( \Gamma(d/2) \). Since \( d \) is an integer, we must determine the values of the gamma function both for integer and half-integer arguments. To find \( \Gamma(1/2) \) we use the definition (3.4.12) and let \( t = u^2 \):

\[
\Gamma \left( \frac{1}{2} \right) = \int_0^\infty dt \ e^{-t} \ t^{-1/2} = 2 \int_0^\infty du \ e^{-u^2} = \sqrt{\pi}.
\]

(3.4.15)

Similarly,

\[
\Gamma(1) = \int_0^\infty dt \ e^{-t} = 1.
\]

(3.4.16)

For larger arguments the calculation of the gamma function is simplified using a recursion relation. To obtain this relation, begin with

\[
\Gamma(x + 1) = \int_0^\infty dt \ e^{-t x}, \quad x > 0,
\]

(3.4.17)

which can be rewritten as

\[
\Gamma(x + 1) = -\int_0^\infty dt \left( \frac{d}{dt} e^{-t} \right) t^x = -\int_0^\infty dt \left( \frac{d}{dt} (e^{-t} x) - x e^{-t} t^{x-1} \right).
\]

(3.4.18)

The boundary terms vanish for \( x > 0 \) and we find that

\[
\Gamma(x + 1) = x \Gamma(x), \quad x > 0.
\]

(3.4.19)

Using this recursion relation we find, for example,

\[
\Gamma \left( \frac{3}{2} \right) = \frac{1}{2} \cdot \Gamma \left( \frac{1}{2} \right) = \frac{1}{2} \sqrt{\pi}, \quad \Gamma \left( \frac{5}{2} \right) = \frac{3}{2} \cdot \Gamma \left( \frac{3}{2} \right) = \frac{3}{2} \sqrt{\pi}.
\]

(3.4.20)

For integer arguments the gamma function is related to the factorial:

\[
\Gamma(5) = 4 \cdot \Gamma(4) = 4 \cdot 3 \cdot \Gamma(3) = 4 \cdot 3 \cdot 2 \cdot \Gamma(2) = 4 \cdot 3 \cdot 2 \cdot 1 \cdot \Gamma(1) = 4!.
\]

Therefore, for \( n \in \mathbb{Z} \) and \( n \geq 1 \), we have

\[
\Gamma(n) = (n - 1)!,
\]

(3.4.21)
3.5. ELECTRIC FIELDS IN HIGHER DIMENSIONS

where we recall that 0! = 1. We can now test our formula (3.4.14) in the familiar cases:

\[
\begin{align*}
\text{vol} (S^1) &= \text{vol} (S^{2-1}) = \frac{2\pi}{\Gamma(1)} = 2\pi, \\
\text{vol} (S^2) &= \text{vol} (S^{3-1}) = \frac{2\pi^{3/2}}{\Gamma(\frac{3}{2})} = 4\pi, \\
\end{align*}
\]

(3.4.22)
in agreement with the known values. For the less familiar \( S^3 \) we find

\[
\begin{align*}
\text{vol} (S^3) &= \text{vol} (S^{4-1}) = \frac{2\pi^2}{\Gamma(2)} = 2\pi^2. \\
\end{align*}
\]

(3.4.23)

Quick Calculation 3.4. Show that \( \text{vol} (B^d) = \pi^{d/2}/\Gamma(1 + \frac{d}{2}) \).

3.5 Electric fields in higher dimensions

In this section we calculate the electric field due to a point charge in a world with \( d \) spatial dimensions. Here \( d \) could be three, in which case the answer is familiar, or less than three, but we are particularly interested in \( d > 3 \). To do this calculation we will use the general version of Maxwell’s equations appropriate for arbitrary number of spatial dimensions. As you may imagine, the electric field of a point charge is radial. Our calculation will give the radial dependence and the normalization of the electric field. With minor modifications, this result will also inform us about the gravitational fields of point particles in \( d \) spatial dimensions.

Our computation is based on the zeroth component of equation (3.3.23):

\[
\frac{\partial F^{0i}}{\partial x^i} = \rho.
\]

(3.5.1)

Since \( F^{0i} = E_i \) (see (3.3.26)), this equation is just Gauss’s law:

\[
\nabla \cdot \vec{E} = \rho.
\]

(3.5.2)

Gauss’s law is valid in all dimensions! Equation (3.5.2) can be used to determine the electric field of a point charge. Let’s first review how this is done in the familiar setting of three spatial dimensions.

Consider a point charge \( q \), a two-sphere \( S^2(r) \) of radius \( r \) centered on the charge, and the three-ball \( B^3(r) \) whose boundary is the two-sphere. We integrate both sides of equation (3.5.2) over the three-ball to find

\[
\int_{B^3} d\text{(vol)} \nabla \cdot \vec{E} = \int_{B^3} d\text{(vol)} \rho.
\]

(3.5.3)

We use the divergence theorem on the left hand side and note that the volume integral on the right hand side gives the total charge:

\[
\int_{S^2(r)} \vec{E} \cdot d\vec{a} = q.
\]

(3.5.4)
Since the magnitude $E(r)$ of $\vec{E}$ is constant over the two-sphere, we get

$$\text{vol}(S^2(r)) E(r) = q.$$ \hspace{1cm} (3.5.5)

The volume of the two-sphere is just its area $4\pi r^2$, so

$$E(r) = \frac{q}{4\pi r^2}. \hspace{1cm} (3.5.6)$$

This is the familiar result for the electric field of a point charge in three spatial dimensions. The electric field magnitude falls off like $1/r^2$.

For dimensions higher than three the starting point (3.5.2) is good, so we must ask if the divergence theorem also holds. It turns out that it does. We will first state the theorem in $d$ spatial dimensions, and then we will give some justification for it.

Consider a $d$-dimensional subspace $V^d$ of $\mathbb{R}^d$ and let $\partial V^d$ denote the boundary of $V^d$. Moreover, let $\vec{E}$ be a vector field in $\mathbb{R}^d$. The divergence theorem states that

$$\int_{V^d} d(\text{vol}) \nabla \cdot \vec{E} = \text{Flux of } \vec{E} \text{ across } \partial V^d = \int_{\partial V^d} \vec{E} \cdot d\vec{v}. \hspace{1cm} (3.5.7)$$

The last right-hand side requires some explanation. At any point on $\partial V^d$, the space $\partial V^d$ is locally approximated by the $(d - 1)$-dimensional tangent hyperplane. For a small piece of $\partial V^d$ around this point, the associated vector $d\vec{v}$ is a vector orthogonal to the hyperplane, pointing out of the volume, and with magnitude equal to the volume of the small piece under consideration. Note that this explanation is in accord with your experience in $\mathbb{R}^3$, where $d\vec{v}$ corresponds to the area vector element $d\vec{a}$.

Figure 3.1: An attempt at a representation of a four-dimensional hypercube. The two faces of constant $x$ are shown (shaded) together with their outgoing normal vectors.

Let us justify the divergence theorem for the case of four space dimensions. Following a strategy used in elementary textbooks, it suffices to prove the divergence theorem for
3.5. ELECTRIC FIELDS IN HIGHER DIMENSIONS

a small hypercube – the result for general subspaces follows by breaking such spaces into many small hypercubes. Because it is not easy to imagine a four-dimensional hypercube, we might as well use a three-dimensional picture with four-dimensional labels (Figure 3.1). We use cartesian coordinates \(x, y, z, w\), and consider a cube whose faces lie on hyperplanes selected by the condition that one of the coordinates is constant. Let one face of the cube and the face opposite to it lie on hyperplanes of constant \(x\) and constant \(x + dx\), respectively. The outgoing normal vectors are \(e_x\), for the face at \(x + dx\), and \((-e_x)\), for the face at \(x\). The volume of each of these two faces equals \(dydzdw\), where \(dy, dz\), and \(dw\), together with \(dx\) are the lengths of the edges of the cube. For an arbitrary electric field \(\vec{E}(x, y, z, w)\), only the \(x\) component contributes to the flux through these two faces. The contribution is

\[
\left[ E_x(x + dx, y, z, w) - E_x(x, y, z, w) \right] dydzdw \approx \frac{\partial E_x}{\partial x} dxdydzdw.
\]  

(3.5.8)

Analogous expressions hold for the flux across the three other pairs of faces. The total net flux from the little cube is just

\[
\text{Flux of } \vec{E} = \left( \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} + \frac{\partial E_w}{\partial w} \right) dxdydzdw = \nabla \cdot \vec{E} d(\text{vol}).
\]  

(3.5.9)

This result is precisely the divergence theorem (3.5.7) applied to an infinitesimal hypercube. This is what we wanted to show.

We can now return to the computation of the electric field due to a point charge in a world with \(d\) spatial dimensions. Consider a point charge \(q\), the sphere \(S^{d-1}(r)\) of radius \(r\) centered on the charge (this is the sphere that surrounds the charge), and the ball \(B^d(r)\) whose boundary is the sphere \(S^{d-1}(r)\). Again, we integrate both sides of equation (3.5.2) over the ball \(B^d(r)\):

\[
\int_{B^d} d(\text{vol}) \nabla \cdot \vec{E} = \int_{B^d} d(\text{vol}) \rho.
\]  

(3.5.10)

The volume integral on the right-hand side gives the total charge, and the divergence theorem (3.5.7) relates the left-hand side to a flux integral:

\[
\text{Flux of } \vec{E} \text{ across } S^{d-1}(r) = q.
\]  

(3.5.11)

The flux equals the magnitude of the electric field times the volume of \(S^{d-1}(r)\), so

\[
E(r) \text{ vol } (S^{d-1}(r)) = q.
\]  

(3.5.12)

Making use of (3.4.14) we find

\[
E(r) = \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{d/2}} \frac{q}{r^{d-1}}.
\]  

(3.5.13)

This is the value of the electric field for a point charge in a world with \(d\) spatial dimensions. For \(d = 3\) we recover the inverse-squared dependence of the electric field. In higher dimensions the electric field falls off faster at large distances. For each additional spatial dimension we get an additional factor of \(1/r\) in the radial dependence of the electric field.
Quick Calculation 3.5. Verify that for $d = 3$ equation (3.5.13) coincides with (3.5.6).

Quick Calculation 3.6. The force $\vec{F}$ on a test charge $q$ in an electric field $\vec{E}$ is $\vec{F} = q\vec{E}$. What are the units of charge in various dimensions?

The electrostatic potential $\Phi$ is also of interest. For time independent fields (3.3.27) gives

$$\vec{E} = -\nabla \Phi.$$  

(3.5.14)

This equation, together with Gauss’s law, gives the Poisson equation:

$$\nabla^2 \Phi = -\rho,$$  

(3.5.15)

which can be used to calculate the potential due to a charge distribution. The two equations above hold in all dimensions using, of course, the appropriate definitions of the gradient and the Laplacian.

### 3.6 Gravitation and Planck’s length

Einstein’s theory of general relativity is a theory of gravitation. In this very elegant theory the dynamical variables encode the geometry of spacetime. When gravitational fields are sufficiently weak and velocities are small, Newtonian gravitation is accurate enough, and one need not work with the more complex machinery of general relativity. We can use Newtonian gravity to understand the definition of Planck’s length in various dimensions and its relation to the gravitational constant. These are interesting issues that we will explain here and in the rest of the present chapter. Nevertheless, when gravitation emerges in string theory, it does so in the language of Einstein’s theory of general relativity. To be able to recognize the appearance of gravity among the quantum vibrations of the relativistic string you need a little familiarity with the language of general relativity. In here you will take a first look at the concepts involved in this remarkable theory.

Most physicists do not expect general relativity to hold at truly small distances nor for extremely large gravitational fields. This is a realm where string theory, the first serious candidate for a quantum theory of gravitation, is necessary. General relativity is the large-distance/weak-gravity limit of string theory. String theory modifies general relativity; it must do so to make it consistent with quantum mechanics. The conceptual framework which underlies these modifications is not clear yet. It will no doubt emerge as we understand string theory better in the years to come.

The spacetime of special relativity, Minkowski spacetime, is the arena for physics in the absence of gravitational fields. The geometrical properties of Minkowski spacetime are encoded by the metric formula (2.2.15), which gives the invariant interval separating two nearby events:

$$-ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu.$$  

(3.6.1)

Here the Minkowski metric $\eta_{\mu\nu}$ is a constant metric, represented as a matrix with entries $(-1,1,...,1)$ along the diagonal. Minkowski space is said to be a flat space. In the presence
of a gravitational field, the metric becomes dynamical. We then write
\[ -ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu, \]  
(3.6.2)
where the constant \( \eta_{\mu\nu} \) is replaced by the metric \( g_{\mu\nu}(x) \). If there is a gravitational field, the metric is in general a nontrivial function of the spacetime coordinates. The metric \( g_{\mu\nu} \) is defined to be symmetric
\[ g_{\mu\nu}(x) = g_{\nu\mu}(x). \]  
(3.6.3)
It is also customary to define \( g^{\mu\nu}(x) \) as the inverse of the \( g_{\mu\nu}(x) \) matrix:
\[ g^{\mu\alpha}(x) g_{\alpha\nu}(x) = \delta^\mu_\nu. \]  
(3.6.4)

For many physical phenomena gravity is very weak, and the metric \( g_{\mu\nu}(x) \) can be chosen to be very close to the Minkowski metric \( \eta_{\mu\nu} \). We then write,
\[ g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x), \]  
(3.6.5)
and we view \( h_{\mu\nu}(x) \) as a small fluctuation around the Minkowski metric. This expansion is done, for example, to study gravity waves. Those waves represent small “ripples” on top of the Minkowski metric. Einstein’s equations for the gravitational field are written in terms of the spacetime metric \( g_{\mu\nu}(x) \). These equations imply that matter or energy sources curve the spacetime manifold. For weak gravitational fields, Einstein’s equations can be expanded in powers of \( h_{\mu\nu} \) using (3.6.5). In the absence of sources, the resulting linearized equation for \( h_{\mu\nu} \) is
\[ \partial^2 h_{\mu\nu} - \partial_\alpha (\partial^\mu h^{\alpha\nu} + \partial^\nu h^{\alpha\mu}) + \partial^\mu \partial^\nu h = 0. \]  
(3.6.6)
In here \( h^{\mu\nu} \equiv \eta^{\mu\alpha} \eta^{\nu\beta} h_{\alpha\beta} \) and \( h \equiv h^{\mu\nu} h_{\mu\nu} = -h_{00} + h_{11} + h_{22} + h_{33} \). Equation (3.6.6) is the gravitational analog of equation (3.3.24), which describes Maxwell fields in the absence of sources. While (3.3.24) is exact, (3.6.6) is only valid for weak gravitational fields. It is the linearized approximation to a nonlinear equation.

The analogy with electromagnetism extends to the existence of gauge transformations. Einstein’s gravity has gauge transformations. They arise because the use of different systems of coordinates yields equivalent descriptions of gravitational physics. In learning string theory in this book you will get to appreciate the freedom to choose coordinates on the surfaces generated by moving strings. In general relativity, an infinitesimal change of coordinates
\[ x^{\mu'} = x^\mu + \epsilon^\mu(x), \]  
(3.6.7)
can be viewed as an infinitesimal change of the metric \( g_{\mu\nu} \) and, using (3.6.5), as an infinitesimal change of the fluctuating field \( h_{\mu\nu} \). One can show that the change is given as
\[ \delta h^{\mu\nu} = \partial^\mu \epsilon^\nu + \partial^\nu \epsilon^\mu. \]  
(3.6.8)
The linearized equation of motion (3.6.6) is invariant under the gauge transformation (3.6.8). We will check this explicitly in Chapter 10. In Maxwell theory the gauge parameter has no indices, but in general relativity the gauge parameter has a vector index.
As we mentioned before, Newtonian gravitation emerges from general relativity in the approximation of weak gravitational fields and motion with small velocities. For many purposes Newtonian gravity suffices. Starting now, and for the rest of this chapter, we will use Newtonian gravity to understand the definition of Planck’s length in various dimensions, and to investigate how gravitational constants behave when some spatial dimensions are curled up. The results that we will obtain hold also in the full theory of general relativity.

Newton’s law of gravitation in four-dimensions states that the force of attraction between two masses $m_1$ and $m_2$ separated by a distance $r$ is given by

$$|\vec{F}^{(4)}| = \frac{G m_1 m_2}{r^2},$$

(3.6.9)

where $G$ denotes the four-dimensional Newton constant. It follows that the units of the gravitational constant $G$ are

$$[G] = \text{[Force]} = \frac{L^2}{M^2} = \frac{ML^2}{T^2} = \frac{L^3}{MT^2}.$$  (3.6.10)

The numerical value for the constant $G$ is experimentally determined:

$$G = 6.67 \times 10^{-11} \frac{m^3}{\text{kg} \cdot \text{s}^2}.$$  (3.6.11)

Since $[c] = L/T$ and $[\hbar] = ML^2/T$, the three fundamental constants $G$, $c$, and $\hbar$ can be written as

$$G = 6.67 \times 10^{-11} \frac{m^3}{\text{kg} \cdot \text{s}^2}, \quad c = 3 \times 10^8 \frac{\text{m}}{\text{s}}, \quad \hbar = 1.06 \times 10^{-34} \frac{\text{kg} \cdot \text{m}^2}{\text{s}}.$$  (3.6.12)

In the study of gravitation it is sometimes convenient to use a “natural” system of units. Since we have three basic units, those of length, time, and mass, we can find new units of length, time, and mass such that the three fundamental constants, $G$, $c$, and $\hbar$ take the numerical value of one in those units. These units are called the Planck length $\ell_P$, the Planck time $t_P$, and the Planck mass $m_P$, respectively. In those units

$$G = 1 \cdot \frac{\ell_P^3}{m_P t_P^2}, \quad c = 1 \cdot \frac{\ell_P}{t_P}, \quad \hbar = 1 \cdot \frac{m_P \ell_P^2}{t_P},$$

(3.6.13)

without additional numerical constants – as opposed to equation (3.6.12). The above equations allow us to solve for $\ell_P$, $t_P$, and $m_P$ in terms of $G$, $c$, and $\hbar$. One readily finds

$$\ell_P = \sqrt{\frac{G \hbar}{c^3}} = 1.61 \times 10^{-33} \text{ cm},$$

(3.6.14)

$$t_P = \frac{\ell_P}{c} = \sqrt{\frac{\hbar G}{c^5}} = 5.4 \times 10^{-44} \text{ s},$$

(3.6.15)

$$m_P = \sqrt{\frac{\hbar c}{G}} = 2.17 \times 10^{-5} \text{ g}.$$  (3.6.16)
These numbers represent scales at which relativistic quantum gravity effects can be important. Indeed, the Planck length is an extremely small length, and the Planck time is an incredibly short time – the time it takes light to travel the Planck length! While Einstein’s gravity can be used down to relatively small distances and back to relatively early times in the history of the universe, a quantum gravity theory (such as string theory) is needed to study gravity at distances of the order of the Planck length or to investigate the universe when it was Planck-time old.

There is an equivalent way to characterize the Planck length: \( \ell_P \) is the unique length that can be constructed using only powers of \( G, c, \) and \( \hbar \). One thus sets

\[
\ell_P = (G)^\alpha (c)^\beta (\hbar)^\gamma,
\]

and fixes the constants \( \alpha, \beta, \) and \( \gamma \) so that the right-hand side has units of length.

**Quick Calculation 3.7.** Show that this condition fixes uniquely \( \alpha = \gamma = 1/2, \) and \( \beta = -3/2, \) thus reproducing the result in (3.6.14).

It may appear that \( m_P \) is not a very large mass, but it is, in fact, a spectacularly large mass from the viewpoint of elementary particle physics. The mass \( m_P \) is roughly \( 10^{19} \) times larger than the mass of the proton. If the fundamental theory of nature is based on the basic constants \( G, c, \) and \( \hbar, \) it is then a great mystery why the masses of the elementary particles are so much smaller than the ‘obvious’ mass \( m_P \) that can be built from the basic constants. This puzzle is usually called the hierarchy problem.

For an additional perspective on the Planck mass, consider the following question: What should be the mass \( M \) of the proton so that the gravitational force between two protons cancels the electric repulsion force between them? Equating the magnitudes of the electric and gravitational forces we get

\[
\frac{GM^2}{r^2} = \frac{e^2}{4\pi r^2} \quad \longrightarrow \quad GM^2 = \frac{e^2}{4\pi},
\]

(3.6.18)

It is convenient to divide both sides of the equation by \( \hbar c \) to find

\[
\frac{GM^2}{\hbar c} = \frac{e^2}{4\pi \hbar c} \approx \frac{1}{137} \quad \longrightarrow \quad \frac{M^2}{m_p^2} \approx \frac{1}{137},
\]

(3.6.19)

where use was made of (3.6.16). We thus find \( M \approx m_P/12, \) or about one-tenth of the Planck mass. The dimensionless ratio \( e^2/(4\pi \hbar c) \) is called the fine structure constant. It was evaluated above using the Heaviside-Lorentz definition of electric charge where \( e = \sqrt{4\pi} \times 4.8 \times 10^{-10} \) esu (see Problem 2.1(c)).

**Quick Calculation 3.8.** The mass of the electron is \( m_e = 0.911 \times 10^{-27} \) g, and its energy equivalent is \( m_e c^2 = 0.511 \) MeV. Show that the energy equivalent of the Planck mass is \( m_P c^2 = 1.22 \times 10^{19} \) GeV (1 GeV = 10^9 eV). This energy is called the Planck energy.
3.7 Gravitational potentials

We want to learn what happens with the gravitational constant $G$ when we attempt to describe gravitation in spacetimes of other dimensionalities. To find out, we will examine gravitational potentials in Newtonian gravity. In this section we obtain the equation that relates the gravitational potential to the mass distribution in a spacetime with arbitrary but fixed number of spatial dimensions. In doing so we will learn how to define the relevant gravitational constant. This result will be used in the following section to define, in any dimension, the Planck length in terms of the appropriate gravitational constant.

We introduce a gravity field $\vec{g}$ with units of force per unit mass. The definition is similar to that of an electric field in terms of the force on a test particle: the force on a given test mass $m$ at a point where the gravity field is $\vec{g}$ is given by $m\vec{g}$. We set $\vec{g}$ equal to minus the gradient of a gravitational potential $V_g$:

$$\vec{g} = -\nabla V_g.$$  \hspace{1cm} (3.7.1)

We will take this equation to be true in all dimensions. Equation (3.7.1) has content: if you move a particle along a closed loop in a static gravitational field, the net work that you do against the gravitational field is zero.

Quick Calculation 3.9. Prove the above statement.

What are the units for the gravitational potential? Equation (3.7.1) gives

$$[\vec{g}] = \left[\frac{\text{Force}}{M}\right] = \left[\frac{V_g}{L}\right] \rightarrow [V_g] = \left[\frac{\text{Energy}}{M}\right].$$  \hspace{1cm} (3.7.2)

The gravitational potential has units of energy per unit mass in any dimension. The gravitational potential $V_g^{(4)}$ of a point mass in four dimensions is

$$V_g^{(4)} = -\frac{GM}{r}.$$  \hspace{1cm} (3.7.3)

We can use the electromagnetic analogy to find the equation satisfied by the gravitational potential. In electromagnetism, we found an equation for the electrostatic potential which holds in any dimension. This is equation (3.5.15):

$$\nabla^2 \Phi = -\rho.$$  \hspace{1cm} (3.7.4)

The four-dimensional scalar potential for a point charge $q$ is

$$\Phi^{(4)} = \frac{q}{4\pi r},$$  \hspace{1cm} (3.7.5)

and it satisfies (3.7.4) where $\rho$ is the charge density for the point charge. It follows by analogy that the four-dimensional gravitational potential in (3.7.3) satisfies

$$\nabla^2 V_g^{(4)} = 4\pi G \rho_m,$$  \hspace{1cm} (3.7.6)
where $\rho_m$ is the matter density. While this equation is correct in four dimensions, a small modification is needed for other dimensions. Note that the left-hand side has the same units in any number of dimensions: the units of $V_g$ are always the same, and the Laplacian always divides by length squared. The right-hand side must also have the same units in any number of dimensions. Since $\rho_m$ is mass density, it has different units in different dimensions, and, as a consequence, the units of $G$ must change when the dimensions change. We therefore rewrite the above equation more precisely as

$$\nabla^2 V_g^{(D)} = 4\pi G^{(D)} \rho_m,$$

(3.7.7)

when working in $D$-dimensional spacetime. The superscripts shown in parenthesis denote the dimensionality of spacetime. In particular, we identify $G^{(4)}$ as the four-dimensional Newton constant $G$. In general, we will use $D$ to denote the dimensionality of spacetime, and $d$ to denote the number of spatial dimensions. Clearly, $D = d + 1$.

Equation (3.7.7) defines Newtonian gravitation in arbitrary number of dimensions. Just as the electric field of a point charge does, the gravitational field of a point mass falls off like $1/r$ in a world with $d$ spatial dimensions. As a result, the force between two point masses separated by a distance $r$ falls off like $1/r^{d+1}$. For three spatial dimensions, this is the familiar inverse-squared dependence of the gravitational force. If $D = 6$ (a world with two extra dimensions) the gravitational force falls of like $1/r^4$.

### 3.8 The Planck length in various dimensions

We define the Planck length in any dimension just as we did in four dimensions: the Planck length is the unique length built using only powers of the gravitational constant $G^{(D)}$, $c$, and $\hbar$. To compute the Planck length we must determine the units of $G^{(D)}$. This is easily done if we recall that the units of $G^{(D)} \rho_m$ (the right-hand side of (3.7.7)) are the same in all dimensions.

Comparing the cases of five and four dimensions, for example,

$$[G^{(5)}] \frac{M}{L^4} = [G] \frac{M}{L^3} \rightarrow [G^{(5)}] = L [G].$$

(3.8.1)

The units of $G^{(5)}$ carry one more factor of length than the units of $G$. We use (3.6.14) to read the units of $G$ in terms of units of length and units of $c$ and $\hbar$:

$$[G] = \frac{[c]^3 L^2}{[\hbar]}.$$  

(3.8.2)

Equation (3.8.1) then gives

$$[G^{(5)}] = \frac{[c]^3 L^3}{[\hbar]}.$$  

(3.8.3)

Since the Planck length is constructed uniquely from the gravitational constant, $c$, and $\hbar$, we can remove the brackets in the above equation and replace $L$ by the five-dimensional...
Planck length $\ell_{P}^{(5)}$:

$$\left(\ell_{P}^{(5)}\right)^3 = \frac{\hbar G^{(5)}}{c^3}. \quad (3.8.4)$$

Reintroducing the four-dimensional Planck length:

$$\left(\ell_{P}^{(5)}\right)^3 = \left(\frac{\hbar G_{(5)}}{c^3} \frac{G^{(5)}}{G} \right) \rightarrow \left(\ell_{P}^{(5)}\right)^3 = \left(\ell_{P}\right)^2 \frac{G^{(5)}}{G}. \quad (3.8.5)$$

Since they do not have the same units, the gravitational constants in four and five dimensions cannot be directly compared. Planck lengths, however, can be compared. If the Planck length is the same in four and in five dimensions, then $\frac{G^{(5)}}{G} = \ell_{P}$; the gravitational constants differ by one factor of the common Planck length.

It is not hard to generalize the above equations to $D$ spacetime dimensions:

**Quick Calculation 3.10.** Show that (3.8.4) and (3.8.5) are replaced by

$$\left(\ell_{P}^{(D)}\right)^{D-2} = \frac{\hbar G^{(D)}}{c^3} = \left(\ell_{P}\right)^2 \frac{G^{(D)}}{G}. \quad (3.8.6)$$

### 3.9 Gravitational constants and compactification

If string theory is correct, our world is really higher dimensional. The fundamental gravity theory is then defined in the higher-dimensional world, with some value for the higher-dimensional Planck length. Since we observe only four dimensions, the additional dimensions may be curled up to form a compact space with small volume. We can then ask: what is the effective value of the four-dimensional Planck length? As we shall show here, the effective four-dimensional Planck length depends on the volume of the extra dimensions, as well as on the value of the higher-dimensional Planck length.

These observations raise the possibility that the Planck length in the effectively four-dimensional world – the famous number equal to about $10^{-33}$ cm – may not coincide with the fundamental Planck length in the original higher-dimensional theory. Is it possible that the fundamental Planck length is much bigger than the familiar, four-dimensional one? We will answer this question in the following section. In this section we will work out the effect of compactification on gravitational constants.

How do we calculate the gravitational constant in four dimensions if we are given the gravitational constant in five? First, we recognize the need to curl up one spatial dimension, otherwise there is no effectively four-dimensional spacetime. As we will see, the size of the extra dimension enters the relationship between the gravitational constants. To explore these questions precisely, consider a five-dimensional spacetime where one dimension forms a small circle of radius $R$. We are given $G^{(5)}$ and we would like to calculate $G^{(4)}$.

Let $(x^1, x^2, x^3)$ denote three spatial dimensions of infinite extent, and $x^4$ denote a compactified dimension of circumference $2\pi R$ (Figure 3.2). We place a uniform ring of total
3.9. GRAVITATIONAL CONSTANTS AND COMPACTIFICATION

Figure 3.2: A world with four space dimensions, one of which, $x^4$, is compactified into a circle of radius $R$. A ring of total mass $M$ wraps around this compact dimension.

Mass $M$ all around the circle at $x^1 = x^2 = x^3 = 0$. This is a mass distribution which is constant along the $x^4$ dimension. We are interested in the gravitational potential $V_g^{(5)}$ that emerges from such a mass distribution. We could have alternatively placed a point mass at some fixed $x^4$, but this makes the calculations more involved (Problem 3.9). In the present case the gravitational potential $V_g^{(5)}$ does not depend on $x^4$. The total mass $M$ can be written as

$$\text{Total Mass} = M = 2\pi R m,$$  \hspace{1cm} (3.9.1)

where $m$ is the mass per unit length.

What is the mass density in the five-dimensional world? It is only nonzero at $x^1 = x^2 = x^3 = 0$. To represent such a mass density we use delta functions. Recall that the delta function $\delta(x)$ can be viewed as a singular function whose value is zero except for $x = 0$ and such that the integral $\int_{-\infty}^{\infty} dx \delta(x) = 1$. This integral implies that if $x$ has units of length, then $\delta(x)$ has units of inverse length. Since the five-dimensional mass density is concentrated at $x^1 = x^2 = x^3 = 0$, it is reasonable to include in its formula the product $\delta(x^1)\delta(x^2)\delta(x^3)$ of three delta functions. We claim that

$$\rho^{(5)} = m \delta(x^1)\delta(x^2)\delta(x^3).$$ \hspace{1cm} (3.9.2)

We first check the units. The mass density $\rho^{(5)}$ must have units of $M/L^4$. This works out since $m$ has units of mass per unit length, and the three delta functions supply an additional factor of $L^{-3}$. The ansatz in (3.9.2) could still be off by a constant dimensionless factor, a factor of two, for example. As a final check, we integrate $\rho^{(5)}$ over all space. The result should be the total mass:

$$\int_{-\infty}^{\infty} dx^1 dx^2 dx^3 \int_0^{2\pi R} dx^4 \rho^{(5)} = m \int_{-\infty}^{\infty} dx^1 \delta(x^1) \int_{-\infty}^{\infty} dx^2 \delta(x^2) \int_{-\infty}^{\infty} dx^3 \delta(x^3) \int_0^{2\pi R} dx^4 = m 2\pi R.$$ \hspace{1cm} (3.9.3)
This is indeed the total mass on account of (3.9.1). For the effectively four-dimensional observer the mass is point-like, and it is located at $x^1 = x^2 = x^3 = 0$. So this observer writes

$$\rho^{(4)} = M \delta(x^1) \delta(x^2) \delta(x^3).$$

Note the relation

$$\rho^{(5)} = \frac{1}{2\pi R} \rho^{(4)}. \quad (3.9.5)$$

Let us now use this information in the equations for the gravitational potential. Using the five-dimensional version of (3.7.7) and (3.9.5), we find

$$\nabla^2 V_g^{(5)}(x^1, x^2, x^3) = 4\pi G^{(5)} \rho^{(5)} = 4\pi \frac{G^{(5)}}{2\pi R} \rho^{(4)}. \quad (3.9.6)$$

As we have noted before, $V_g^{(5)}$ is independent of $x^4$ so the Laplacian above is actually the four-dimensional one. Since the effective four-dimensional gravitational potential is $V_g^{(5)}$, the above equation takes the form of the gravitational equation in four dimensions, where the constant in-between the $4\pi$ and the $\rho^{(4)}$ is the four-dimensional gravitational constant. We have therefore shown that

$$G = \frac{G^{(5)}}{2\pi R} \quad \implies \quad \frac{G^{(5)}}{G} = \frac{2\pi R}{\ell_C},$$

where $\ell_C$ is the length of the extra compact dimension. This is what we were seeking: a relationship between the strength of the gravitational constants in terms of the size of the extra dimension.

The generalization of (3.9.7) to the case where there is more than one extra dimension is straightforward. One finds that

$$\frac{G^{(D)}}{G} = (\ell_C)^{D-4}, \quad (3.9.8)$$

where $\ell_C$ is the common length of each of the extra dimensions. When the various dimensions are curled up into circles of different lengths, the above right-hand side must be replaced by the product of the various lengths. This product is, in fact, the volume of the extra dimensions.

### 3.10 Large extra dimensions

We are now done with all the groundwork. In §3.8 we found the relation between the Planck length and the gravitational constant in any dimension. In §3.9 we determined how gravitational constants are related upon compactification. We are ready to find out how the fundamental Planck length in a higher dimensional theory with compactification is related to the Planck length in the effectively four-dimensional theory.
3.10. LARGE EXTRA DIMENSIONS

To begin with, consider a five-dimensional world with Planck length $\ell_p^{(5)}$ and a single spatial coordinate curled up into a circle of circumference $\ell_C$. What then is $\ell_P$? From (3.8.5) and (3.9.7) we find that

$$\left(\ell_p^{(5)}\right)^3 = (\ell_P)^2 \frac{G^{(5)}}{G} = (\ell_P)^2 \ell_C. \quad (3.10.1)$$

Solving for $\ell_P$, we get

$$\ell_P = \ell_p^{(5)} \sqrt{\frac{\ell_p^{(5)}}{\ell_C}}. \quad (3.10.2)$$

This relation enables us to explore the possibility that the world is actually five-dimensional with a fundamental Planck length $\ell_p^{(5)}$ that is much larger than $10^{-33}$ cm. Of course, we must have $\ell_P \sim 10^{-33}$ cm. After all, this is the four-dimensional Planck length, whose value is given in (3.6.14).

Present day accelerators explore physics down to distances of the order of $10^{-16}$ cm. If this distance, or a somewhat smaller one, is the fundamental length scale, we may choose $\ell_p^{(5)} \sim 10^{-18}$ cm. What would $\ell_C$ have to be? With $\ell_p^{(5)} \sim 10^{-18}$ cm and $\ell_P \sim 10^{-33}$ cm, equation (3.10.2) gives $\ell_C \sim 10^{12}$ cm $\sim 10^7$ km. This is more than twenty times the distance from the earth to the moon. Such a large extra dimension would have been detected long time ago.

Having failed to produce a realistic scenario in five dimensions, let us try in six spacetime dimensions. For arbitrary $D$, equations (3.8.6) and (3.9.8) give

$$\left(\ell_p^{(D)}\right)^{D-2} = (\ell_P)^2 \frac{G^{(D)}}{G} = (\ell_P)^2 (\ell_C)^{D-4}. \quad (3.10.3)$$

Solving for $\ell_C$ we find

$$\ell_C = \ell_p^{(D)} \left(\frac{\ell_p^{(D)}}{\ell_P}\right)^{\frac{2}{D-4}}. \quad (3.10.4)$$

For $D = 6$ and $\ell_p^{(6)} \sim 10^{-18}$ cm this formula gives

$$\ell_C = \frac{(\ell_p^{(6)})^2}{\ell_P} \sim 10^{-3} \text{ cm}. \quad (3.10.5)$$

This is a lot more interesting! Could there be extra dimensions $10^{-3}$ cm long? You might think that this is still too big, since even microscopes probe smaller distances. Moreover, as we indicated before, accelerators probe distances of the order of $10^{-16}$ cm. Surprisingly, it is possible that “large extra dimensions” exist and that we have not observed them yet.

The existence of additional dimensions may be confirmed by testing the force law which gives the gravitational attraction between two masses. For distances much larger than the compactification scale $\ell_C$ the world is effectively four-dimensional, so the dependence of the force between two masses on their separation must follow accurately Newton’s inverse-squared law. On the other hand, for distances smaller than $\ell_C$, the world is effectively
higher-dimensional, and the force law will change. A force between two masses that goes like $1/r^4$, where $r$ is the separation, is consistent with the existence of two compact extra dimensions.

It turns out to be very difficult to test gravity at small distances; the force of gravity is extremely weak and spurious electrical forces must be cancelled very precisely. Largely motivated by the possible existence of large extra dimensions, physicists set out to test the inverse-squared law at distances smaller than one millimeter. The experiments that have been carried out to date have found no departure from the inverse-squared law down to distances of about one-tenth of a millimeter. This means that extra dimensions, if they exist, must be smaller than one-tenth of a millimeter. Compact dimensions the size of one-hundredth of a millimeter, as we found in (3.10.5), are still consistent with experiment.

You might ask: What about forces other than gravity? Electromagnetism has been tested to much smaller distances, and we know that the electric force obeys an inverse-squared law very accurately. Rutherford scattering of alpha particles off nuclei, for example, confirms that the inverse-squared law holds down to $10^{-11}\text{cm}$. Since the separation dependence of the electric force would change at distances smaller than the size of the extra dimensions, this seems to rule out large extra dimensions. The possibility of large extra dimensions, however, survives in string theory, where our spatial world could be a three-dimensional hyperplane transverse to the extra dimensions. This hyperplane is called a D3-brane. A D3-brane is a D-brane with three spatial dimensions.

Open strings have the remarkable property that their endpoints must remain attached to the D-branes. In many phenomenological models built in string theory, it is the fluctuations of open strings that give rise to the familiar leptons, quarks, and gauge fields, including the Maxwell gauge field. It follows that these fields are bound to the D3-brane and do not feel the extra dimensions. If the Maxwell field lives on the D-brane, the electric field lines of a charge remain on the D-brane and do not go off into the extra dimensions. The force law is not changed at any distance scale. Closed strings are not bound by D-branes, and therefore gravity, which arises from closed strings, is affected by the extra dimensions.

Although the Planck length $\ell_p$ is an important length scale in four dimensions, if there are large extra dimensions, the truly fundamental Planck length would be much bigger than the effective four-dimensional one. The possibility of large extra dimensions is slightly unnatural – why should the extra dimensions be much larger than the fundamental length scale? This is not a new problem, however, but rather the problem of a large hierarchy in another guise. We noted earlier that particle physics faces a puzzling hierarchy between the Planck mass and the masses of elementary particles. In the large-extra-dimensions scenario, the hierarchy is postulated to arise from extra dimensions that are much larger than the fundamental length scale. At any rate, the truly exciting fact is that present experimental constraints do not rule out large extra dimensions. The discovery of large extra dimensions would be revolutionary.
Problems

Problem 3.1. Lorentz covariance for motion in electromagnetic fields.†

The Lorentz force equation (3.1.5) can be written relativistically as

\[
\frac{dp_\mu}{ds} = \frac{q}{c} F_{\mu\nu} \frac{dx^\nu}{ds},
\]

(1)

where \( p_\mu \) is the four-momentum. Check explicitly that this equation reproduces (3.1.5) when \( \mu \) is a spatial index. What does (1) give when \( \mu = 0? \) Does it make sense? Is (1) a gauge invariant equation?

Problem 3.2. Maxwell equations in four dimensions.

(a) Show explicitly that the source-free Maxwell equations emerge from \( T_{\mu\lambda\nu} = 0.\)

(b) Show explicitly that the Maxwell equations with sources emerge from (3.3.23).

Problem 3.3. Electromagnetism in three dimensions.

(a) Find the reduced Maxwell’s equations in three dimensions by starting with Maxwell’s equations and the force law in four dimensions, using the ansatz (3.2.1), and assuming that no field can depend on the \( z \) direction.

(b) Repeat the analysis of three-dimensional electromagnetism starting with the Lorentz covariant formulation. Take \( A^\mu = (\Phi, A^1, A^2), \) examine \( F_{\mu\nu}, \) the Maxwell equations (3.3.23), and the relativistic form of the force law derived in Problem 3.1.

Problem 3.4. Electric fields and potentials of point charges.

(a) Show that for time-independent fields, the Maxwell equation \( T_{0ij} = 0 \) implies that \( \partial_i E_j - \partial_j E_i = 0. \) Explain why this condition is satisfied by the ansatz \( \vec{E} = -\nabla \Phi. \)

(b) Show that with \( d \) spatial dimensions, the potential \( \Phi \) due to a point charge \( q \) is given by

\[
\Phi(r) = \frac{\Gamma(d/2 - 1)}{4\pi d/2} \frac{q}{r^{d-2}}.
\]

Problem 3.5. Calculating the divergence in higher dimensions.

Let \( \hat{f} = f(r) \hat{r} \) be a vector function in \( \mathbb{R}^d. \) Here \( \hat{r} \) is a unit radial vector, and \( r \) is the radial distance to the origin. Derive a formula for \( \nabla \cdot \hat{f} \) by applying the divergence theorem to a spherical shell of radius \( r \) and width \( dr. \) Check that for \( d = 3 \) your answer reduces to \( \nabla \cdot \hat{f} = f'(r) + \frac{2}{r} f(r). \)
CHAPTER 3. ELECTROMAGNETISM AND GRAVITATION

Problem 3.6. Analytic continuation for gamma functions.†

Consider the definition of the gamma function for complex arguments $z$ whose real part is positive:

$$\Gamma(z) = \int_0^\infty dt \ e^{-tz} \, t^{z-1}, \quad \Re(z) > 0.$$ 

Use this equation to show that for $\Re(z) > 0$

$$\Gamma(z) = \int_1^1 dt \ t^{z-1} \left( e^{-t} - \sum_{n=0}^N \frac{(-t)^n}{n!} \right) + \sum_{n=0}^N \frac{(-1)^n}{n!} \frac{1}{z + n} + \int_1^\infty dt \ e^{-tz} \, t^{z-1}.$$ 

Explain why the above right-hand side is well-defined for $\Re(z) > -n - 1$. It follows that this right-hand side provides the analytic continuation of $\Gamma(z)$ for $\Re(z) > -n - 1$. Conclude that the gamma function has poles at $0, -1, -2, \ldots$, and give the value of the residue at $z = -n$ (with $n$ a positive integer).

Problem 3.7. Simple quantum gravity effects are small.

(a) What would be the “gravitational” Bohr radius for a hydrogen atom if the attraction binding the electron to the proton was gravitational? The standard Bohr radius is $a_0 = \frac{\hbar^2}{m e^2} \approx 5.29 \times 10^{-9}$ cm.

(b) In “units” where $G, c, \text{ and } \hbar$ are set equal to one, the temperature of a black hole is given by $kT = \frac{1}{8\pi M}$. Insert back the factors of $G, c, \text{ and } \hbar$ into this formula. Evaluate the temperature of a black hole of a million solar masses. What is the mass of a black hole whose temperature is room temperature?

Problem 3.8. Planetary motion in four and higher dimensions.

Consider the motion of planets in planar circular orbits around heavy stars in our four-dimensional spacetime and in spacetimes with additional spatial dimensions. We wish to study the stability of these orbits under perturbations that keep them planar. Such a perturbation would arise, for example, if a meteorite moving on the plane of the orbit hits the planet and changes its angular momentum.

Show that while planetary circular orbits in our four-dimensional world are stable under such perturbations, they are not so in five or higher dimensions. [Hint: You may find it useful to use the effective potential for motion in a central force field.]

Problem 3.9. Gravitational field of a point mass in a compactified five-dimensional world.

Consider a five-dimensional spacetime with space coordinates $(x, y, z, w)$ not yet compactified. A point mass $M$ is located at the origin $(x, y, z, w) = (0, 0, 0, 0)$.

(a) Find the gravitational potential $V_g^{(5)}(r)$. Write your answer in terms of $M, G^{(5)}, \text{ and } r = (x^2 + y^2 + z^2 + w^2)^{1/2}$. [Hint: Use $\nabla^2 V_g^{(5)} = 4\pi G^{(5)} \rho_m$ and the divergence theorem.]

Now let $w$ become a circle with radius $a$ while keeping the mass fixed, as shown in Figure 3.3.

(b) Write an exact expression for the gravitational potential $V_g^{(5)}(x, y, z, 0)$. This potential is a function of $R = (x^2 + y^2 + z^2)^{1/2}$ and can be written as an infinite sum.
Problems for Chapter 3

(c) Show that for $R \gg a$ the gravitational potential takes the form of a four-dimensional gravitational potential, with Newton’s constant $G^{(4)}$ given in terms of $G^{(5)}$ as in (3.9.7). [Hint: Turn the infinite sum into an integral].

These results confirm both the relation between the four- and five-dimensional Newton constants in a compactification and the emergence of a four-dimensional potential at distances large compared to the size of the compact dimension.

Problem 3.10. Exact answer for the gravitational potential.

The infinite sum in Problem 3.9 can be evaluated exactly using the identity

$$\sum_{n=-\infty}^{\infty} \frac{1}{1 + (\pi n x)^2} = \frac{1}{x} \coth \left( \frac{1}{x} \right).$$

(a) Find an exact closed-form expression for the gravitational potential $V^{(5)}_g(x, y, z, 0)$ in the compactified theory.

(b) Expand this answer to calculate the leading correction to the gravitational potential in the limit when $R \gg a$. For what value of $R/a$ is the correction of order 1%?

(c) Use the exact answer in (a) to expand the potential when $R \ll a$. Give the first two terms in the expansion. Do you recognize the leading term?
Chapter 4

Non-relativistic Strings

A full appreciation for the subtleties of relativistic strings requires an understanding of the basic physics of non-relativistic strings. These strings have mass and tension. They can vibrate both transversely and longitudinally. We study the equations of motion for non-relativistic strings and develop the Lagrangian approach to their dynamics.

4.1 Equations of motion for transverse oscillations

We will begin our study of strings with a look at the transverse fluctuations of a stretched string. The direction along the string is called the longitudinal direction, and the directions orthogonal to the string are called the transverse directions. We consider, for notational simplicity, the case when there is only one transverse direction – the generalization to additional transverse directions is straightforward.

Working in the \((x, y)\) plane, let the classical non-relativistic string have its endpoints fixed at \((0, 0)\), and \((a, 0)\). In the static configuration the string is stretched along the \(x\)-axis between these two points. In a transverse oscillation, the \(x\)-coordinate of any point on the string does not change in time. The transverse displacement of a point is given by its \(y\)-coordinate. The \(x\) direction is longitudinal, and the \(y\) direction is transverse. To describe the classical mechanics of a homogeneous string, we need two pieces of information: the tension \(T_0\) and the mass per unit length \(\mu_0\). The total mass of the string is then \(M = \mu_0a\).

Let us look briefly at the units. Tension has units of force, so

\[
[T_0] = [\text{Force}] = \frac{[\text{Energy}]}{L}.
\]

(4.1.1)

If you stretch a string an infinitesimal amount \(dx\), its tension remains approximately constant through the stretching, and the change in energy equals the work done \(T_0dx\). The total mass of the string does not change. If we were considering relativistic strings, however, a static string with more energy would have a larger rest mass. Using (4.1.1), noting that energy has units of mass times velocity squared, and that \(\mu_0\) has units of mass per unit
length, we have

$$[T_0] = \frac{M}{L}[v]^2 = [\mu_0][v]^2.$$  \hspace{1cm} (4.1.2)

For a non-relativistic string, both $T_0$ and $\mu_0$ are adjustable parameters, and the velocity on the right-hand side above will turn out to be the velocity of transverse waves. The above equation suggests that the string tension $T_0$ and the linear mass density $\mu_0$ in a relativistic string might be related by $T_0 = \mu_0 c^2$, since $c$ is the canonical velocity in relativity. We will see in Chapter 6 that this is indeed the correct relation for a relativistic string.

![Figure 4.1: A short piece of a classical non-relativistic string vibrating transversely. With different slopes at the two endpoints there is a net vertical force.](image)

Returning to our classical non-relativistic string, let’s figure out the equation of motion. Consider a small portion of the static string that extends from $x$ to $x + dx$, with $y = 0$. This piece is shown in transverse oscillation in Figure 4.1. At time $t$, the transverse displacement of the string is $y(t, x)$ at $x$ and $y(t, x + dx)$ at $x + dx$. We will assume that the oscillations are small, and by this we will mean that at all times

$$\frac{\partial y}{\partial x} \ll 1,$$  \hspace{1cm} (4.1.3)

at any point on the string. This guarantees that the transverse displacement of the string is small compared to the length of the string. The length of the string changes little, and we can assume that the tension $T_0$ is unchanged.

The slope of the string is a bit different at the points $x$ and $x + dx$. This change of slope means that the string tension changes direction and the portion of string under consideration feels a net force. For transverse oscillations we need only calculate the net vertical force; the net horizontal force is negligible (Problem 4.1). The vertical force at $(x + dx, y + dy)$ is accurately given by $T_0$ times $\partial y/\partial x$ evaluated at $x + dx$ and is pointing up; similarly, the vertical force at $(x, y)$ is $T_0$ times $\partial y/\partial x$ evaluated at $x$ and is pointing
down. Therefore the net vertical force $dF_v$ is

$$dF_v = T_0 \frac{\partial y}{\partial x} \bigg|_{x+dx} - T_0 \frac{\partial y}{\partial x} \bigg|_x \simeq T_0 \frac{\partial^2 y}{\partial x^2} dx.$$ (4.1.4)

The mass $dm$ of this piece of string, originally stretched from $x$ to $x + dx$, is given by the mass density $\mu_0$ times $dx$. By Newton’s law, the net vertical force equals mass times vertical acceleration. So we can simply write

$$T_0 \frac{\partial^2 y}{\partial x^2} dx = (\mu_0 dx) \frac{\partial^2 y}{\partial t^2}.$$ (4.1.5)

We cancel $dx$ on each side and rearrange terms to get

$$\frac{\partial^2 y}{\partial x^2} - \frac{\mu_0}{T_0} \frac{\partial^2 y}{\partial t^2} = 0.$$ (4.1.6)

This is just a wave equation! Recall that for the wave equation

$$\frac{\partial^2 y}{\partial x^2} - \frac{1}{v_0^2} \frac{\partial^2 y}{\partial t^2} = 0,$$ (4.1.7)

the parameter $v_0$ is the velocity of the waves. Thus for the transverse waves on our stretched string, the velocity $v_0$ of the waves is

$$v_0 = \sqrt{\frac{T_0}{\mu_0}}.$$ (4.1.8)

The higher the tension or the lighter the string, the faster the waves move.

### 4.2 Boundary conditions and initial conditions

Since equation (4.1.6) is a partial differential equation involving space and time derivatives, in order to fix solutions we must in general apply both boundary conditions and initial conditions. Boundary conditions (B.C.) constrain the solution at the boundary of the system, and initial conditions constrain the solution at a given starting time. The most common types of boundary conditions are Dirichlet and Neumann boundary conditions.

For our string, Dirichlet boundary conditions specify the positions of the string end-points. For example, if we attach each end of the string to a wall (Figure 4.2, left side), we are imposing the Dirichlet boundary conditions

$$y(t, x = 0) = y(t, x = a) = 0,$$ (4.2.1)

Alternatively, if we attach a massless loop to each end of the string and the loops are allowed to slide along two frictionless poles, we are imposing Neumann boundary conditions.

For our string, Neumann boundary conditions specify the values of the derivative $\partial y / \partial x$...
at the endpoints. Since the loops are massless and the poles are frictionless, the derivative \( \frac{\partial y}{\partial x} \) must vanish at the poles \( x = 0, a \) (Figure 4.2, right side). If this were not the case, then the slope of the string at a pole would be nonzero, and a component of the string tension would accelerate the rings in the \( y \)-direction. Since each ring is massless, their acceleration would be infinite. This is not possible, so, in effect, we are imposing the Neumann boundary conditions

\[
\frac{\partial y}{\partial x}(t, x = 0) = \frac{\partial y}{\partial x}(t, x = a) = 0, \quad \text{Neumann B.C. (4.2.2)}
\]

These Neumann boundary conditions apply to strings whose endpoints are free to move along the \( y \) direction.

Let’s see how we can solve the wave equation for a particular set of initial conditions. The general solution of equation (4.1.6) is of the form

\[
y(t, x) = h_+(x - v_0 t) + h_-(x + v_0 t), \quad (4.2.3)
\]

where \( h_+ \) and \( h_- \) are arbitrary functions of a single variable. This solution represents a superposition of two waves, \( h_+ \) moving to the right and \( h_- \) moving to the left. Suppose that the initial values of \( y \) and \( \frac{\partial y}{\partial t} \) are known at time \( t = 0 \). Using equation (4.2.3) we see that this information yields the equations

\[
y(0, x) = h_+(x) + h_-(x), \quad (4.2.4)
\]

\[
\frac{\partial y}{\partial t}(0, x) = -v_0 h'_+(x) + v_0 h'_-(x), \quad (4.2.5)
\]

where the left-hand sides are known functions, and the primes denote derivatives with respect to arguments. Using (4.2.4) we can solve for \( h_- \) in terms of \( h_+ \). Substituting into (4.2.5), we get a first-order ordinary differential equation for \( h_+ \). Once we have solved for \( h_+ \) (using appropriate boundary conditions), we can use (4.2.4) again, this time to find the explicit form of \( h_- \). With \( h_+ \) and \( h_- \) known, the full solution of the equations of motion is given by (4.2.3).
4.3 Frequencies of transverse oscillation

Suppose that we have a string where each point is oscillating in the $y$-direction sinusoidally and in phase. This means that $y(t, x)$ is of the form

$$y(t, x) = y(x) \sin(\omega t + \phi),$$  \hfill (4.3.1)

where $\omega$ is the angular frequency of oscillation and $\phi$ is the constant common phase. Our aim is to find the allowed frequencies of oscillation. Substituting (4.3.1) into (4.1.6) and cancelling the common time dependence, we find

$$\frac{d^2 y(x)}{dx^2} + \omega^2 \frac{\mu_0}{T_0} y(x) = 0.$$  \hfill (4.3.2)

This is an ordinary second-order differential equation for the profile $y(x)$ of the oscillations. The allowed frequencies are selected by this equation, together with the boundary conditions. Since $\omega, \mu_0,$ and $T_0$ are constants, the differential equation is solved in terms of trigonometric functions. With Dirichlet boundary conditions (4.2.1) we have the nontrivial solutions

$$y_n(x) = A_n \sin\left(\frac{n \pi x}{a}\right), \quad n = 1, 2, \ldots ,$$  \hfill (4.3.3)

where $A_n$ is an arbitrary constant. The value $n = 0$ is not included above because it represents a motionless string. Plugging $y_n(x)$ into (4.3.2), we find the allowed frequencies $\omega_n$:

$$\omega_n = \sqrt{\frac{T_0}{\mu_0}} \frac{n \pi}{a}, \quad n = 1, 2, \ldots .$$  \hfill (4.3.4)

These are the frequencies of oscillation for a Dirichlet string. The strings on a violin are Dirichlet strings. To tune a violin to the right frequency one must adjust the string tension. The higher the tension is, the higher the pitch, as predicted by (4.3.4).

For the case of Neumann boundary conditions (4.2.2), we obtain the spatial solutions

$$y_n(x) = A_n \cos\left(\frac{n \pi x}{a}\right) \quad n = 0, 1, 2, \ldots .$$  \hfill (4.3.5)

This time the $n = 0$ solution is a little less trivial: the string does not oscillate, but it is rigidly translated to $y(t, x) = A_0$. The oscillation frequencies, found by plugging (4.3.5) into (4.3.2), are the same as those in (4.3.4). Therefore, the oscillation frequencies are the same in the Neumann and Dirichlet problems. The Neumann case admits one extra solution not included in our oscillatory ansatz (4.3.1): the string can translate with constant velocity. Indeed, $y(t, x) = at + b$, with $a$ and $b$ arbitrary constants, satisfies both the boundary conditions and the original wave equation (4.1.7).
4.4 More general oscillating strings

Let us discuss briefly some problems that are closely related to the ones considered thus far. For example, we can take the mass density of the string to be a function $\mu(x)$ of position. The form (4.1.6) of the wave equation does not change since it is derived from local considerations: the examination of a little piece of string that can be chosen to be sufficiently small so that the mass density is approximately constant. We therefore get

$$\frac{\partial^2 y}{\partial x^2} - \frac{\mu(x)}{T_0} \frac{\partial^2 y}{\partial t^2} = 0. \quad (4.4.1)$$

For normal oscillations, we use the ansatz in (4.3.1) and find

$$\frac{d^2 y}{dx^2} + \frac{\mu(x)}{T_0} \omega^2 y(x) = 0. \quad (4.4.2)$$

This equation is no longer simple to solve, and it can only be studied in detail once the function $\mu(x)$ is specified. In Problems 4.4 and 4.5 you will consider some specific mass distributions, and you will explore a variational approach that gives an upper bound for the lowest oscillation frequency.

So far we have only considered strings that oscillate transversally. Strings also admit longitudinal oscillations, although the relativistic string does not. Imagine a string stretched along the $x$-axis, and consider the infinitesimal segment which at equilibrium extends from $x$ to $x+dx$. Suppose now that at time $t$ the ends of this infinitesimal segment are longitudinally displaced from their equilibrium positions by distances $\eta(t, x)$ and $\eta(t, x+dx)$, respectively. If these two quantities are not the same, the piece of string is being compressed or stretched. An equation of motion can be obtained for this system, much as we did for transverse motion. It is not possible, however, to assume that the tension is constant throughout the string. For transverse oscillations the net force acting on a little piece of string arose from the different angles at which the same tension was applied on opposite ends of the piece. If the string always lies along the $x$-axis then a net force can act on a segment only if the tension is different on its two ends. Therefore the waves on a longitudinally-oscillating string are accompanied by tension waves (Problem 4.2).

4.5 A brief review of Lagrangian mechanics

The Lagrangian $L$ of a system is defined by

$$L = T - V, \quad (4.5.1)$$

where $T$ is the kinetic energy of the system and $V$ is the potential energy of the system. For a point particle of mass $m$ moving along the $x$ axis under the influence of a time-independent potential $V(x)$, the non-relativistic Lagrangian takes the form

$$L(t) = \frac{1}{2} m (\dot{x}(t))^2 - V(x(t)), \quad \dot{x}(t) \equiv \frac{dx(t)}{dt}. \quad (4.5.2)$$
We must emphasize that the above Lagrangian is implicitly a function of time, but it has no explicit time dependence. All the time dependence arises from the time dependence of the position \( x(t) \). The action \( S \) is defined as

\[
S = \int_{\mathcal{P}} L(t)dt ,
\]

(4.5.3)

where \( \mathcal{P} \) is a path \( x(t) \) between an initial position \( x_i \) at an initial time \( t_i \), and a final position \( x_f \) at a final time \( t_f > t_i \). One such path is shown in Figure 4.3.

Figure 4.3: A path \( \mathcal{P} \) representing a possible one-dimensional motion \( x(t) \) of a particle during the time interval \([t_i, t_f]\).

The action is a functional. Whereas a function of a single variable takes one number – the argument – as input and gives another number as output, a functional takes a function as the input, and gives a number as output. Since a function is usually defined by its values at infinitely-many points, we can think of a functional as a function of infinitely-many variables. In our present application, the input for the action functional is the function \( x(t) \) which determines the path \( \mathcal{P} \). We can emphasize the argument of \( S \) by using the notation \( S[x] \). Here \([x]\) represents the full function \( x(t) \). It is potentially confusing to write \( S[x(t)] \), since it suggests that \( S \) is ultimately a function of \( t \), which it is not.

More explicitly, for any path \( x(t) \), the action is given by

\[
S[x] = \int_{t_i}^{t_f} \left\{ \frac{1}{2} m (\dot{x}(t))^2 - V(x(t)) \right\} dt .
\]

(4.5.4)

It is very important to emphasize that the action \( S \) can be calculated for any path \( x(t) \) and not only for paths that represent physically-realized motion. It is because \( S \) can be calculated for all paths that it is a very powerful tool to find the paths that can be physically realized.

Hamilton’s principle states that the path \( \mathcal{P} \) which a system actually takes is one for which the action \( S \) is stationary. More precisely, if this path \( \mathcal{P} \) is varied infinitesimally, the
action does not change to first order in the variation. In terms of the function \( x(t) \) which specifies the path, the perturbed path takes the form \( x(t) + \delta x(t) \), as shown in Figure 4.4. For any time \( t \), the variation \( \delta x(t) \) is the vertical distance between the original path and the varied path. As in the figure, we consider variations where the initial and final positions \( x_i = x(t_i) \) and \( x_f = x(t_f) \) are unchanged:

\[
\delta x(t_i) = \delta x(t_f) = 0. \tag{4.5.5}
\]

We now calculate the action \( S[x + \delta x] \) for the perturbed path \( x(t) + \delta x(t) \):

\[
S[x + \delta x] = \int_{t_i}^{t_f} \left\{ \frac{m}{2} \left( \frac{d}{dt} (x(t) + \delta x(t)) \right)^2 - V(x(t) + \delta x(t)) \right\} dt,
\]

\[
= S[x] + \int_{t_i}^{t_f} \left\{ m\ddot{x}(t) \frac{d}{dt} \delta x(t) - V'(x(t)) \delta x(t) \right\} dt + O((\delta x)^2). \tag{4.5.6}
\]

In passing to the last right-hand side we expanded \( V \) in a Taylor series about \( x(t) \). The terms of order \( (\delta x)^2 \) and higher are unnecessary to determine whether or not the action is stationary. We have thus left them undetermined and indicated them by \( O((\delta x)^2) \). We can write the new action as \( S + \delta S \), where \( \delta S \) is linear in \( \delta x \). From the equation above we see that \( \delta S \) is given by

\[
\delta S = \int_{t_i}^{t_f} \left\{ m\ddot{x}(t) \frac{d}{dt} \delta x(t) - V'(x(t)) \delta x(t) \right\} dt. \tag{4.5.7}
\]

To find the equations of motion, the variation \( \delta S \) must be rewritten in the form \( \delta S = \int dt \delta x(t) \{ \ldots \} \). In particular, no derivatives must be acting on \( \delta x \). This can be achieved
using integration by parts:

\[
\delta S = \int_{t_i}^{t_f} \left\{ \frac{d}{dt} \left( m\dot{x}(t)\delta x(t) \right) - m\ddot{x}(t)\delta x(t) - V'(x(t))\delta x(t) \right\} dt ,
\]

\[= m\dot{x}(t_f)\delta x(t_f) - m\dot{x}(t_i)\delta x(t_i) + \int_{t_i}^{t_f} \delta x(t)(-m\ddot{x}(t) - V'(x(t))) dt .\]  

(4.5.8)

Making use of (4.5.5), the variation reduces to

\[
\delta S = \int_{t_i}^{t_f} \delta x(t)(-m\ddot{x}(t) - V'(x(t))) dt .
\]

(4.5.9)

The action is stationary if \(\delta S\) vanishes for every variation \(\delta x(t)\). For this to happen, the factor multiplying \(\delta x(t)\) in the integrand must vanish:

\[m\ddot{x}(t) = -V'(x(t)).\]

(4.5.10)

This is Newton’s second law applied to the motion of a particle in a potential \(V(x)\). We have recovered the expected equation of motion by requiring that the action be stationary under variations.

Suppose that we have determined the path that the particle takes while going from \(x_i\) to \(x_f\). As we have seen, the action is then stationary under variations that vanish at the initial and final times. Is the action also stationary under variations that change the initial position at \(t_i\) or the final position at \(t_f\)? In general, the answer is no. This can be seen from equation (4.5.8). The integral term vanishes by assumption, but if \(\delta x(t_f) \neq 0\), the first term on the right-hand side would not vanish unless \(m\dot{x}(t_f)\), the final momentum of the particle, happens to vanish. The situation is analogous for \(\delta x(t_i) \neq 0\).

### 4.6 The non-relativistic string Lagrangian

Let’s return now to our string with constant mass density \(\mu_0\), constant tension \(T_0\), and ends located at \(x = 0\) and \(x = a\). The kinetic energy is simply the sum of the kinetic energies of all the infinitesimal segments that comprise the string. So it can be written as

\[
T = \int_0^a \frac{1}{2} (\mu_0 dx) \left( \frac{\partial y}{\partial t} \right)^2 .
\]

(4.6.1)

The potential energy arises from the work which must be done to stretch the segments. Consider an infinitesimal portion of string which extends from \((x, 0)\) to \((x + dx, 0)\) when the string is in equilibrium. If the string element is momentarily stretched from \((x, y)\) to \((x + dx, y + dy)\), as in Figure 4.1, then the change in length \(\Delta l\) of the infinitesimal segment is given by

\[
\Delta l = \sqrt{(dx)^2 + (dy)^2} - dx = dx \left( \sqrt{1 + \left( \frac{\partial y}{\partial x} \right)^2} - 1 \right) \approx dx \frac{1}{2} \left( \frac{\partial y}{\partial x} \right)^2 ,
\]

(4.6.2)
where we have used the small oscillation approximation (4.1.3) to discard higher order terms in the expansion of the square root. Since the work done in stretching each infinitesimal segment is $T_0 \Delta l$, the total potential energy $V$ is

$$V = \int_0^a \frac{1}{2} T_0 \left( \frac{\partial y}{\partial x} \right)^2 dx.$$  (4.6.3)

The Lagrangian for the string is given by $T - V$:

$$L(t) = \int_0^a \left[ \frac{1}{2} \mu_0 \left( \frac{\partial y}{\partial t} \right)^2 - \frac{1}{2} T_0 \left( \frac{\partial y}{\partial x} \right)^2 \right] dx = \int_0^a \mathcal{L} \, dx ,$$  (4.6.4)

where $\mathcal{L}$ is referred to as the Lagrangian density:

$$\mathcal{L} \left( \frac{\partial y}{\partial t}, \frac{\partial y}{\partial x} \right) = \frac{1}{2} \mu_0 \left( \frac{\partial y}{\partial t} \right)^2 - \frac{1}{2} T_0 \left( \frac{\partial y}{\partial x} \right)^2 .$$  (4.6.5)

The action for our string is therefore

$$S = \int_{t_i}^{t_f} L(t) \, dt = \int_{t_i}^{t_f} dt \int_0^a \left[ \frac{1}{2} \mu_0 \left( \frac{\partial y}{\partial t} \right)^2 - \frac{1}{2} T_0 \left( \frac{\partial y}{\partial x} \right)^2 \right] .$$  (4.6.6)

In this action the “path” is the function $y(t, x)$. To find the equations of motion, we must examine the variation of the action as we vary: $y(t, x) \rightarrow y(t, x) + \delta y(t, x)$. Performing the variation as before, we get

$$\delta S = \int_{t_i}^{t_f} \int_0^a \left[ \mu_0 \frac{\partial y}{\partial t} \frac{\partial (\delta y)}{\partial t} - T_0 \frac{\partial y}{\partial x} \frac{\partial (\delta y)}{\partial x} \right] .$$  (4.6.7)

**Quick Calculation 4.1.** Prove equation (4.6.7).

We must have no derivatives acting on the variations, so we rewrite each of the two terms above as a full derivative minus a term in which the derivative does not act on the variation:

$$\delta S = \int_{t_i}^{t_f} \int_0^a \left[ \mu_0 \frac{\partial y}{\partial t} \frac{\partial (\delta y)}{\partial t} - \mu_0 \frac{\partial^2 y}{\partial t^2} \delta y + \frac{\partial}{\partial x} \left( -T_0 \frac{\partial y}{\partial x} \delta y \right) + T_0 \frac{\partial^2 y}{\partial x^2} \delta y \right] .$$  (4.6.8)

The time derivative on the first line reduces to evaluations at $t_f$ and $t_i$, while the space derivative on the second line gives evaluations at the string endpoints:

$$\delta S = \int_0^a \left[ \mu_0 \frac{\partial y}{\partial t} \right]_{t=0}^{t=f} dx + \int_{t_i}^{t_f} \left[ -T_0 \frac{\partial y}{\partial x} \right]_{x=0}^{x=a} dt$$

$$- \int_{t_i}^{t_f} \int_0^a \left( \mu_0 \frac{\partial^2 y}{\partial t^2} - T_0 \frac{\partial^2 y}{\partial x^2} \right) \delta y .$$  (4.6.9)
Our final expression for $\delta S$ contains three terms. Each one must vanish independently. The third term, for example, is determined by the motion of the string for $x \in (0,a)$ and $t \in (t_i,t_f)$. The boundary conditions do not restrict $\delta y(t,x)$ here, so we set to zero the coefficient of $\delta y$, and recover our original equation (4.1.6). The first term in (4.6.9) is determined by the configuration of the string at times $t_i$ and $t_f$. If we specify these configurations, we are in effect setting $\delta y(t_i,x)$ and $\delta y(t_f,x)$ to zero. This causes the first term to vanish. We encountered an analogous situation in our study of the free particle.

The second term in (4.6.9) is new: it concerns the motion of the string endpoints $y(t,0)$ and $y(t,a)$. We can make this term vanish by specifying either Dirichlet or Neumann boundary conditions. Suppose we impose the Dirichlet boundary conditions (4.2.1). Then the positions of our endpoints are fixed throughout time, so we require that the variation $\delta y(t,x)$ vanishes for $x = 0$ and $x = a$. This will cause the second term to vanish. If, on the other hand, we assume that the endpoints are free to move, then the variations $\delta y(t,x)$ are unconstrained at $x = 0$ and $x = a$. In this case the second term will vanish if we impose the conditions

$$
\frac{\partial y}{\partial x}(t,0) = \frac{\partial y}{\partial x}(t,a) = 0, \quad \text{Neumann B.C.} \tag{4.6.10}
$$

These are the Neumann boundary conditions (4.2.2). Dirichlet boundary conditions can be written in a form where the similarity to Neumann boundary conditions is more apparent. If the string endpoints are fixed, the time derivatives of the endpoint coordinates must vanish

$$
\frac{\partial y}{\partial t}(t,0) = \frac{\partial y}{\partial t}(t,a) = 0, \quad \text{Dirichlet B.C.} \tag{4.6.11}
$$

The similarity with (4.6.10) is quite striking. The only change is that spatial derivatives were turned into time derivatives. If we write Dirichlet boundary conditions in this form, we must still specify the values of the coordinates at the fixed endpoints.

In order to appreciate further the physical import of boundary conditions, we consider the momentum $p_y$ carried by the string. There is no other component to the momentum, because we have assumed that the motion is restricted to the $y$-direction. This momentum is simply the sum of the momenta of each infinitesimal segment along the string:

$$
p_y = \int_0^a \mu_0 \frac{\partial y}{\partial t} \, dx. \tag{4.6.12}
$$

Let us see if this momentum is conserved:

$$
\frac{dp_y(t)}{dt} = \int_0^a \mu_0 \frac{\partial^2 y}{\partial t^2} \, dx = \int_0^a T_0 \frac{\partial^2 y}{\partial x^2} \, dx = T_0 \left[ \frac{\partial y}{\partial x} \right]_{x=0}^{x=a}, \tag{4.6.13}
$$

where we used the wave equation (4.1.6). We see that momentum is conserved for Neumann boundary conditions (4.6.10), but for Dirichlet boundary conditions momentum is not generally conserved! Indeed, when the endpoints of a string are attached to a wall, the wall is
constantly exerting a force on the string. In the lowest normal mode of a Dirichlet string, for example, the net momentum constantly oscillates between the $+y$- and $-y$-directions.

Why is this important for string theory? For a long time string theorists did not take seriously the possibility of Dirichlet boundary conditions. It seemed unphysical that the string momentum could fail to be conserved. Moreover, what could the endpoints of open strings be attached to? The answer is that they are attached to D-branes – a new kind of dynamical extended object. If a string is attached to a D-brane then momentum can be conserved – the momentum lost by the string is absorbed by the D-brane. A detailed analysis of the spatial boundary term induced by variation is crucial to recognize the possibility of D-branes in string theory.

We conclude this chapter with a more general derivation of the equation of motion for the string. For this, we use (4.6.5) to write the action as

$$S = \int_{t_i}^{t_f} dt \int_0^a dx \mathcal{L} \left( \frac{\partial y}{\partial t}, \frac{\partial y}{\partial x} \right).$$

(4.6.14)

We also define the quantities

$$\mathcal{P}^t \equiv \frac{\partial \mathcal{L}}{\partial y}, \quad \mathcal{P}^x \equiv \frac{\partial \mathcal{L}}{\partial y},$$

(4.6.15)

with $y' = \partial y/\partial x$. These are simply the derivatives of $\mathcal{L}$ with respect to its first and second arguments, respectively. Explicitly, they are

$$\mathcal{P}^t = \mu_0 \frac{\partial y}{\partial t}, \quad \mathcal{P}^x = -T_0 \frac{\partial y}{\partial x}.$$  

(4.6.16)

When we vary the motion by $\delta y$, the variation of the action is given by

$$\delta S = \int_{t_i}^{t_f} dt \int_0^a dx \left[ \frac{\partial \mathcal{L}}{\partial y} \delta y + \frac{\partial \mathcal{L}}{\partial y} \delta y \right] = \int_{t_i}^{t_f} dt \int_0^a dx \left[ \mathcal{P}^t \delta y + \mathcal{P}^x \delta y \right].$$

(4.6.17)

Using the standard manipulations we find

$$\delta S = \int_0^a \left[ \mathcal{P}^t \delta y \right]_{t=t_i}^{t=t_f} dx + \int_{t_i}^{t_f} dt \left[ \mathcal{P}^x \delta y \right]_{x=0}^a dt - \int_{t_i}^{t_f} dt \int_0^a dx \left( \frac{\partial \mathcal{P}^t}{\partial t} + \frac{\partial \mathcal{P}^x}{\partial x} \right) \delta y.$$  

(4.6.18)

**Quick Calculation 4.2.** Derive equation (4.6.18).

**Quick Calculation 4.3.** Match in detail equations (4.6.18) and (4.6.9).

The variation in (4.6.18) gives the equation of motion

$$\frac{\partial \mathcal{P}^t}{\partial t} + \frac{\partial \mathcal{P}^x}{\partial x} = 0.$$  

(4.6.19)
Using (4.6.16) we readily see that this is the wave equation (4.1.6).

Note that $P^t$, as given in (4.6.16), coincides with the momentum density in equation (4.6.12). This is not an accident. In Lagrangian mechanics, the derivative of the Lagrangian with respect to a velocity is the conjugate momentum. For the string, $\dot{y}$ plays the role of a velocity, so $P^t$, the derivative of the Lagrangian density with respect to $\dot{y}$, is a momentum density.

In addition, note that for string endpoints that are free to move the vanishing of $\delta S$ requires $P^x = 0$. As we can see from (4.6.16), this is a Neumann boundary condition. Furthermore, $P^t$ vanishes at the string endpoints for a Dirichlet boundary condition (4.6.11). A more detailed analysis of these facts will be given in Chapter 8, where $P^t$ and $P^x$ will be shown to have an interesting two-dimensional interpretation.

Problems

Problem 4.1. *Consistency of small transverse oscillations.*

Reconsider the analysis of transverse oscillations in §4.1. Calculate the horizontal force $dF_h$ on the little piece of string shown in Figure 4.1. Show that for small oscillations this force is much smaller than the vertical force $dF_v$ responsible for the transverse oscillations.

Problem 4.2. *Longitudinal waves on strings.*

Consider a string with uniform mass density $\mu_0$ stretched between $x = 0$ and $x = a$. Let the equilibrium tension be $T_0$. Longitudinal waves are possible if the tension of the string varies as it stretches or compresses. For a piece of this string with equilibrium length $L$, a small change $\Delta L$ of its length is accompanied by a small change $\Delta T$ of the tension where

$$\frac{1}{\tau_0} = \frac{1}{L} \frac{\Delta L}{\Delta T}.$$

Here $\tau_0$ is a tension coefficient with units of tension. Find the equation governing the small longitudinal oscillations of this string. Give the velocity of the waves.

Problem 4.3. *Evolving an initial string configuration.*

A string with tension $T_0$, mass density $\mu_0$, and wave velocity $v_0 = \sqrt{T_0/\mu_0}$, is stretched from $(x, y) = (0, 0)$ to $(x, y) = (a, 0)$. The string endpoints are fixed, and the string can vibrate in the $y$ direction.

(a) Write $y(t, x)$ as in (4.2.3), and prove that the above Dirichlet boundary conditions imply

$$h_+(u) = -h_-(-u) \quad \text{and} \quad h_+(u) = h_+(u + 2a). \quad (1)$$

Here $u \in (-\infty, \infty)$ is a dummy variable that stands for the argument of the functions $h_\pm$.

Now consider an initial value problem for this string. At $t = 0$ the transverse displacement is identically zero, and the velocity is

$$\frac{\partial y}{\partial t}(0, x) = v_0 \frac{x}{a} \left(1 - \frac{x}{a}\right), \quad x \in (0, a). \quad (2)$$
(b) Calculate $h_+(u)$ for $u \in (-a, a)$. Does this define $h_+(u)$ for all $x$?

(c) Calculate $y(t, x)$ for $x$ and $v_0 t$ in the domain $D$ defined by the two conditions

$$D = \{(x, v_0 t) \mid 0 \leq x \leq v_0 t < a\}.$$ 

Exhibit the domain $D$ in a plane with axes $x$ and $v_0 t$.

(d) At $t = 0$ the midpoint $x = a/2$ has the largest velocity of all points in the string. Show that the velocity of the midpoint reaches the value of zero at time $t_0 = a/(2v_0)$ and that $y(t_0, a/2) = a/12$. This is the maximum vertical displacement of the string.

Problem 4.4. A configuration with two joined strings.

A string with tension $T_0$ is stretched from $x = 0$ to $x = 2a$. The part of the string $x \in (0, a)$ has constant mass density $\mu_1$, and the part of the string $x \in (a, 2a)$ has constant mass density $\mu_2$. Consider the differential equation (4.4.2) that determines the normal oscillations.

(a) What boundary conditions should be imposed on $y(x)$ and $\frac{dy}{dx}(x)$ at $x = a$?

(b) Write the conditions that determine the possible frequencies of oscillation.

(c) Calculate the lowest frequency of oscillation of this string when $\mu_1 = \mu_0$ and $\mu_2 = 2\mu_0$.

Problem 4.5. Variational problem for strings.

Consider a string stretched from $x = 0$ to $x = a$, with a tension $T_0$ and a position-dependent mass density $\mu(x)$. The string is fixed at the endpoints and can vibrate in the $y$-direction. Equation (4.4.2) determines the oscillation frequencies $\omega_i$ and associated profiles $\psi_i(x)$ for this string.

(a) Set up a variational procedure that gives an upper bound on the lowest frequency of oscillation $\omega_0$. (This can be done roughly as in quantum mechanics, where the ground state energy $E_0$ of a system with Hamiltonian $H$ satisfies $E_0 \leq \langle \psi, H\psi \rangle / \langle \psi, \psi \rangle$). As a useful first step consider the inner product

$$\langle \psi_i, \psi_j \rangle \equiv \int_0^a \mu(x)\psi_i(x)\psi_j(x)dx$$

and show that it vanishes when $\omega_i \neq \omega_j$. Explain why your variational procedure works.

(b) Consider the case $\mu(x) = \mu_0 x$. Use your variational principle to find a simple bound on the lowest oscillation frequency. Compare with the answer $\omega_0^2 \simeq (18.956)\frac{T_0}{\mu_0 a^2}$ obtained by a direct numerical solution of the eigenvalue problem.

Problem 4.6. Deriving Euler-Lagrange equations.

(a) Consider an action for a dynamical variable $q(t)$:

$$S = \int dt \, L(q(t), \dot{q}(t); t) .$$

Calculate the variation $\delta S$ of the action under a variation $\delta q(t)$ of the coordinate. Use the condition $\delta S = 0$ to find the equation of motion for the coordinate $q(t)$ (the Euler-Lagrange equation).
(b) Consider an action for a dynamical field variable $\phi(t, \vec{x})$. As indicated, the field is a function of space and time, and is briefly written as the spacetime function $\phi(x)$. The action is obtained by integrating the Lagrangian density $\mathcal{L}$ over spacetime. The Lagrangian density is a function of the field and the spacetime derivatives of the field:

$$S = \int d^D x \, L(\phi(x), \partial_\mu \phi(x)).$$  \hspace{1cm} (2)

Here $d^D x = dt \, dx^1 \cdots dx^d$, and $\partial_\mu \phi = \partial \phi/\partial x^\mu$. Calculate the variation $\delta S$ of the action under a variation $\delta \phi(x)$ of the field. Use the condition $\delta S = 0$ to find the equation of motion for the field $\phi(x)$ (the Euler-Lagrange equation).