Chapter 1
Pre-Quantum Theories

In this introductory chapter we review two theories from classical physics – Newtonian mechanics and Maxwellian electrodynamics – and use them to introduce a number of concepts (such as determinism, locality, ontology, measurement, and configuration space) that we will explore in the context of quantum mechanics in subsequent chapters.

1.1 Newtonian Mechanics

As a first example of a "pre-quantum theory" let's consider the picture of the universe formulated by Isaac Newton. The theory, in a nutshell, says that the physical world consists of *particles* interacting by means of *forces* which the particles exert on one another and which influence the particles' motions. About the particles, Newton wrote:

> ...it seems probable to me, that God in the Beginning form'd Matter in solid, massy, hard, impenetrable, moveable Particles, of such Sizes and Figures, and with such other Properties, and in such Proportion to Space, as most conduced to the End for which he form'd them; and that these primitive Particles being Solids, are incomparably harder than any porous Bodies compounded of them; even so very hard, as never to wear or break in pieces.... [A]ll material Things seem to have been composed of the hard and solid Particles above-mention'd, variously associated.... [1, pp. 400–2]

Newton's endorsement of the idea that observable macroscopic objects are composed of invisibly small, indestructible particles is a kind of bridge between the speculative notion of atomism that had been introduced by Ancient Greek philosophers such as Democritus, and the more scientific atomic theory of matter that grew out of chemistry and physics in the centuries following Newton.

Regarding the forces that these particles exert on one another, Newton wrote that

> Bodies act one upon another by the Attractions of Gravity, Magnetism, and Electricity; and these Instances shew the Tenor and Course of Nature, and make it not improbable but that

T. Norsen, *Foundations of Quantum Mechanics*, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-65867-4_1

> there may be more attractive Powers than these. .... [W]e must learn from the Phaenomena of Nature what Bodies attract one another, and what are the Laws and Properties of the Attraction.... The Attractions of Gravity, Magnetism, and Electricity, reach to very sensible distances, and so have been observed by vulgar Eyes, and there may be others which reach to so small distances as hitherto escape Observation.... [1, p. 376]

Although he did not have any particular detailed theories about them, Newton thus anticipated the empirical quest to understand the short-range attractions and repulsions between particles that we now think of as responsible for micro-physical, chemical, and even biological processes. But of course Newton *did* have a rather well-worked-out theoretical account of the long-range *gravitational* interactions between particles.

According to Newton's law of universal gravitation, the gravitational force exerted *on* a particle of mass m_i located at position $\vec{r}_i$, *by* another particle of mass m_j located at position $\vec{r}_j$, is given by

$$\vec{F}_{i,j} = \frac{G m_i m_j}{r_{ij}^2} \hat{r}_{ij} \tag{1.1}$$

where $r_{ij} = |\vec{r}_i - \vec{r}_j|$ is just the distance between the two particles and

$$\hat{r}_{ij} = \frac{\vec{r}_j - \vec{r}_i}{r_{ij}} \tag{1.2}$$

is a unit vector pointing along the line from $\vec{r}_i$ back toward $\vec{r}_j$. The gravitational force between two elementary particles, that is, is proportional to the product of the masses of the particles, inversely proportional to the square of the distance between them, and is directed back toward the particle exerting the force. The proportionality constant, G, which we now call "Newton's constant", was first measured by Cavendish about a century after Newton.

The total or net force on the i^{th} particle is then

$$\vec{F}_i^{\text{net}} = \sum_{j \neq i} \vec{F}_{i,j}. \tag{1.3}$$

(Note that here we ignore the existence of other, short-range forces and pretend for simplicity that the particles *only* interact gravitationally.) And of course it is this net force that influences the particle's trajectory through space in accordance with Newton's second law of motion:

$$\vec{F}_i^{\text{net}} = m_i \vec{a}_i. \tag{1.4}$$

Note that Newton's inverse square law, Eq. (1.1), also embodies Newton's third law: for every action there's an equal and opposite reaction. Or more precisely: if j exerts a force on i, then i necessarily also exerts a force on j, and these two forces (that they exert on each other) have equal magnitudes but precisely opposite directions. That is:

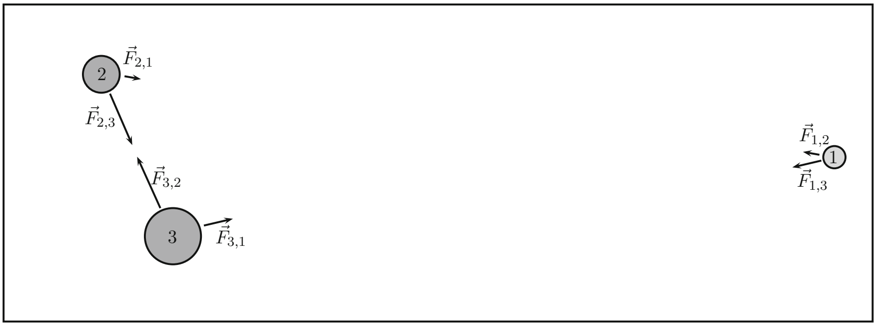


Fig. 1.1 Three massive bodies and the gravitational forces they exert on one another

$$\vec{F}_{i,j} = -\vec{F}_{j,i}. \tag{1.5}$$

It is nice to have some pictures to go along with all the equations, so in Fig. 1.1 I've illustrated some of these ideas by showing three particles (which one might think of as two stars forming a binary star system plus an orbiting planet) and the forces they exert on one another.

Note that the basic laws of Newtonian mechanics (both the expressions for the forces and also Newton's second law which describes how the particles respond to forces) are postulated as applying fundamentally to the elementary, microscopic "Particles" that Newton spoke of in the first block quote. It is perhaps not terribly surprising, but important and interesting nevertheless, that these same laws (properly understood) *also* turn out to apply to large macroscopic objects like stars and planets and apples. That is, in Newtonian mechanics, the applicability to macroscopic objects of (for example) the gravitational inverse square law and Newton's second law, are *theorems* which can be *derived* from the basic laws (understood as applying to the elementary Particles) rather than postulates. You are invited to consider this point further in some of the end-of-chapter Projects.

It is perhaps worth making more explicit that the long-range gravitational forces exerted on each particle depend, according to Eq. (1.1), on the *instantaneous* positions of the distant particles exerting the forces. There is nothing like a delay, for example, associated with some finite-speed propagation of the gravitational influence. The Newtonian gravitational forces, as described by Eq. (1.1), are thus *non-local*, by which we simply mean that they embody what Einstein would describe as a kind of "spooky action at a distance." Interestingly, though, Newton himself did not believe that this apparent non-locality should be taken seriously, as accurately capturing the true nature of gravitational interactions. In a famous 1693 letter to Richard Bentley, Newton wrote:

> It is inconceivable that inanimate brute matter should, without the mediation of something else which is not material, operate upon and affect other matter without mutual contact... That gravity should be innate, inherent, and essential to matter, so that one body may act upon another at a distance through a vacuum, without the mediation of anything else, by and

> through which their action and force may be conveyed from one to another, is to me so great an absurdity that I believe no man who has in philosophical matters a competent faculty of thinking can ever fall into it [2].

Newton thus evidently did not regard what we are here calling "Newtonian mechanics" as providing a complete and final description of gravitational interactions. Instead he seems to have regarded it as merely a starting point, justified by its success in accounting for the observed motions of planets (and comets and tides and falling apples and so on). And as we will see in the next section, physical theory did ultimately develop along the lines suggested here by Newton, with the "field" concept (and the associated removal of the troubling, if merely apparent, non-locality) that Faraday and Maxwell introduced into electro-magnetic theory (and then Einstein introduced into gravitational theory with his general theory of relativity).

Let us close this section with one last figure, which serves two functions: first, introducing the idea of a "space-time diagram" and, second, illustrating one last implication of the (perhaps merely apparent) non-locality of Newton's account of gravity.

A "space-time diagram" is just a graph of position versus time, but (by convention) with the time axis running vertically upward in the figure. Thus, a horizontal slice through the diagram represents the configuration of objects at some particular moment in time.[1] The curves representing the paths of objects "through space-time" are called "world-lines". This probably sounds fancier and deeper than it is; remember this is just a graph of position versus time, but turned sideways! Fig. 1.2 shows a space-time diagram for the same three-object system illustrated before. One sees the orbits of the two "stars" in the "binary star system" (about their mutual center of mass) as the double-helical world lines, with the world line for the (much lighter) "planet" suggesting a longer-period orbit around the "stars".

The particular time slice shaded in gray is meant to correspond to the situation depicted earlier, in Fig. 1.1. Note that, because of the dynamical non-locality mentioned before, there is a sense in which we must think of this particular "slicing" of space-time as somehow "objectively real" according to this theory. That is, in order for the equations of motion of the theory to be well-defined, there must be a real fact of the matter about which points, on the world lines of the other distant particles, are simultaneous with the point in question where we want to know the force.

To make this concrete, imagine a tilted slice through the same point on the world line of the "planet" that the slice in the Figure passes through. The tilted slice would intersect the world lines of the two "stars" at different points on their world lines, and therefore the magnitudes and directions of the gravitational forces exerted on the "planet" at that moment would be different. But the planet is going to move in some particular way, and this requires that some one of the possible ways of slicing up the

[1] That is: a particular moment in time from the point of view of some particular reference frame. Readers who have studied special relativity will recall that due to the relativity of simultaneity, a slice of space-time representing events occurring at the same time, but for a differently-moving observer, will appear tilted with respect to the one drawn in the Figure. This will be important shortly.

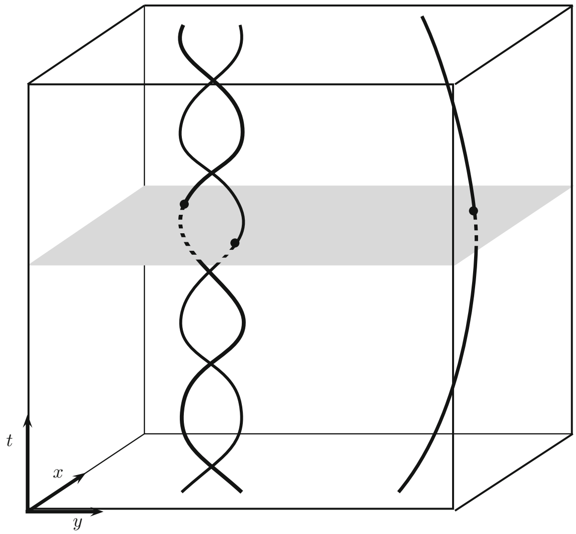


Fig. 1.2 A space-time diagram depicting the evolution, through time, of the binary-star-plus-planet system described earlier

space-time into "simultaneity slices" – the one that gives the forces that generate the planet's actual trajectory – is dynamically privileged, objectively correct.

But this idea of a special, dynamically privileged reference frame contradicts the fundamental principle of relativity theory, that all reference frames are equally valid. Or, as Einstein expressed this point,

> There is [according to relativity] no such thing as simultaneity of distant events; consequently there is also no such thing as immediate action at a distance in the sense of Newtonian mechanics [3, p. 61].

Newton's philosophical instincts may rightly have bristled at the idea of instantaneous action-at-a-distance, but Einstein's relativity theory provided, for the first time, a strong physics-based reason for denying the sort of non-locality suggested by (a naive reading of) Newton's law of universal gravitation.

1.2 Maxwellian Electrodynamics

Let us then turn to a second example "pre-quantum" theory – the theory of electrically charged particles interacting with electric and magnetic *fields*. We begin for simplicity with the case of electrostatics, which is basically parallel to Newtonian gravitation. According to *Coulomb's law*, electrically charged particles exert forces on each other, with the force exerted on charge i by charge j being

$$\vec{F}_{i,j} = -\frac{kq_i q_j}{r_{ij}^2}\hat{r}_{ij} \tag{1.6}$$

which is of course just like Newton's inverse square law for gravity but with the masses replaced by charges. (The minus sign out front is also different: whereas all masses are positive and gravitational forces are always attractive, the electrostatic force between two like charges – either both positive or both negative – is instead *repulsive*.) The net electric force on charge i is then

$$\vec{F}_i^{\text{net}} = \sum_{j \neq i} \vec{F}_{i,j}. \tag{1.7}$$

The concept of electric *field* is often first introduced as a kind of calculational tool: we define the electric field at some location $\vec{r}$ as the (net electric) force that a charged particle would feel if it were located at $\vec{r}$, divided by the charge q of that hypothetical charged particle. That is:

$$\vec{E}(\vec{r}) = \frac{\vec{F}^{\text{net}}(\vec{r})}{q} \tag{1.8}$$

(or, perhaps, to take care of a certain technical detail that need not concern us here, the same thing but in the limit as $q \to 0$). Thus, the electric field at each point $\vec{r}$ can be written as a sum over contributions from all the charged particles in the universe:

$$\vec{E}(\vec{r}) = \sum_{i=1}^{N} \frac{kq_i}{|\vec{r} - \vec{r}_i|^2}\hat{r}_i \tag{1.9}$$

where $\hat{r}_i$ is a unit vector pointing from the point $\vec{r}_i$ to the point $\vec{r}$ where we are calculating the electric field. This equation embodies what is called the principle of *superposition*, which basically means that if charged particle 1 produces a field $\vec{E}_1$ at some point (i.e., if *only* charged particle 1 were around, the electric field at that point would be simply $\vec{E}_1$) and charged particle 2 produces a field $\vec{E}_2$ at that point (i.e., if *only* charged particle 2 were around, the electric field at that point would be simply $\vec{E}_2$), then – with *both* 1 and 2 around – the field is just the sum $\vec{E}_1 + \vec{E}_2$. The concept of "superposition" plays an important role in quantum mechanics, so we highlight it here.

It follows from the definition of the electric field above that, if a particle with charge q is located at a point $\vec{r}$ where the electric field is $\vec{E}(\vec{r})$, it will feel a force

$$\vec{F} = q\vec{E}. \tag{1.10}$$

So far, the electric field should seem like a kind of pointless calculational middle-man: we say the force on a charged particle is determined by its charge and the electric field at its location, and then the electric field at its location is just defined as the superposition of a bunch of inverse-square-law contributions from all the other charged particles. Why not just eliminate the middle man and return to Eq. (1.6) with its apparent "spooky action at a distance"?

The answer is that there turns out to be compelling evidence to take the electric field $\vec{E}(\vec{r})$ – as well as the related magnetic field $\vec{B}(\vec{r})$ – *seriously*, not as mere calculational devices to help us compute the forces that the particles exert on each other, but as geniune physically-real *things* that can (for example) carry energy and momentum and other physical properties and so must actually *exist* in addition to the charged particles.

It will be helpful to remind ourselves about Maxwell's equations, which can be understood as telling us how the electric and magnetic fields change in response to the charged particles (and one another). To begin with we have "Gauss' Law"

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \tag{1.11}$$

where ρ is the electric charge density. For example, for a set of point charges q_i at positions $\vec{r}_i$, we would have $\rho(\vec{r}) = \sum_i q_i \delta^3(\vec{r} - \vec{r}_i)$. Gauss' Law should be understood as a re-formulation of Coulomb's Law, which tells us that the electric field around a point charge is radially outward and falls off in magnitude as the inverse square of the distance. (The ϵ_0 is just a constant, related to the constant k that appeared earlier in Coulomb's Law according to $k = \frac{1}{4\pi\epsilon_0}$.)

The second of Maxwell's equations is sometimes called "Gauss' Law for Magnetism." It reads

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{1.12}$$

which can be understood as saying that there are no "magnetic charges" (which produce radially-outward magnetic fields in the same way that electric charges produce radially-outward electric fields). Sometimes this is stated with the assertion that "there are no magnetic monopoles" or by noting that "magnetic field lines never begin or end but instead make closed loops".

But if there are no magnetic charges, why do magnetic fields exist at all? What produces them? The answer turns out to be: *moving* electric charges, i.e., electric *currents*. This is captured in Ampere's Law:

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{j} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t}. \tag{1.13}$$

The $\vec{j}$ on the right hand side is the electric current density. Again, for example, for a set of electric point charges moving with velocities $\vec{v}_i$, we can write the simple expression $\vec{j} = \sum_i q_i \vec{v}_i \delta^3(\vec{r} - \vec{r}_i)$. The μ_0 is another fundamental constant. So, actually, it is only the first term on the right hand side which corresponds to what I wrote just above, that magnetic fields are produced by moving charges. The second term (the so-called "displacement current" term that was famously added to Ampere's original expression by Maxwell) says that, in addition, *changing* electric fields can produce (or here sometimes one says "induce") magnetic fields. Changing electric fields, that is, in some sense act just like electric current in so far as they are able to give rise to magnetic fields.

The fourth and final Maxwell equation is Faraday's law, which says that changing magnetic fields can also give rise to ("induce") electric fields:

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \tag{1.14}$$

These last two equations taken together imply the existence of propagating ("electromagnetic") waves, in which a changing electric field at some point stimulates the appearance of a magnetic field at neighboring points, but the coming-into-existence of this magnetic field in turn induces the coming-into-existence of further electric fields at still-further neighboring points, and so on. It can be shown that these waves propagate with speed $1/\sqrt{\epsilon_0 \mu_0}$ which we identify as the speed of light, c:

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 3 \times 10^8 \text{ m/s}. \tag{1.15}$$

Together, as we have said, Maxwell's equations tell us how the electric and magnetic fields respond to (themselves, each other, and) the charged particles. To complete the formulation of the theory we also need to know how the charged particles respond to the fields. This information is contained in the so-called Lorentz Force Law, which says that a particle of charge q moving with velocity $\vec{v}$ at position $\vec{r}$ feels a force

$$\vec{F} = q\vec{E}(\vec{r}) + q\vec{v} \times \vec{B}(\vec{r}). \tag{1.16}$$

which determines the particle's trajectory according to Newton's second law.

To summarize, the overall picture of the universe according to Maxwellian electrodynamics involves particles moving through a kind of background "soup" (the fields). The fields influence the motion of the particles and can be thought of, just as Newton had suggested, as a kind of space-filling means or intermediary by which the particles influence one another. But the fields are physically real dynamical objects in their own right as well.

1.3 Locality

We have already described the (perhaps merely apparently) non-local character of Newton's theory of gravity and hinted at the idea that the situation is different in Maxwellian electrodynamics. Let us explore this further. I mentioned above that Maxwell's equations imply the existence of electromagnetic waves that propagate at the speed of light c. It is hardly obvious just looking at Maxwell's equations, but actually *all* electromagnetic interactions as such propagate at speed c (or, in some situations/senses, slower).

Let's try to extract this from the equations. To begin with, let's see how the wave equations for $\vec{E}$ and $\vec{B}$ follow from Maxwell's equations. Taking the curl of Eq. (1.14) and using the vector identity $\vec{\nabla} \times (\vec{\nabla} \times \vec{V}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{V}) - \nabla^2 \vec{V}$ gives

$$\vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} = -\frac{\partial}{\partial t}\left(\vec{\nabla} \times \vec{B}\right). \tag{1.17}$$

We may then use Eqs. (1.11), (1.13), and (1.15) to arrive at

$$\nabla^2 \vec{E} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = \vec{\nabla}\left(\frac{\rho}{\epsilon_0}\right) + \frac{\partial}{\partial t}(\mu_0 \vec{j}). \tag{1.18}$$

In empty space, where $\rho = 0$ and $\vec{j} = 0$, this is simply the wave equation for $\vec{E}$:

$$\nabla^2 \vec{E} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0. \tag{1.19}$$

So far so good.

A similar process – beginning by taking the curl of Eq. (1.13) – gives

$$\nabla^2 \vec{B} - \frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} = -\mu_0 \vec{\nabla} \times \vec{j}. \tag{1.20}$$

So the magnetic field $\vec{B}$ also satisfies the wave equation

$$\nabla^2 \vec{B} - \frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} = 0 \tag{1.21}$$

in empty space. And note that the two kinds of waves (electric and magnetic) are necessarily coupled. So in empty space we have electromagnetic waves that propagate at the speed of light, c.

To understand exactly how electric charges (and moving charges, i.e., electric currents) affect the surrounding electric and magnetic fields, however, we should study Eqs. (1.18) and (1.20) in their full glory, including the source terms on the right hand sides. To begin with, think of these two equations as *six* equations, one for each Cartesian component of the fields. These six equations all have the same basic structure, which, for simplicity, we re-write here as follows:

$$\nabla^2 \psi(\vec{x}, t) - \frac{1}{c^2}\frac{\partial^2 \psi(\vec{x}, t)}{\partial t^2} = f(\vec{x}, t) \tag{1.22}$$

where the function $f(\vec{x}, t)$ represents the source term. The source term is in fact some time- or space- derivative of the charge density ρ or the current density $\vec{j}$ (or some combination of such things) but as it turns out the details don't matter so we will work in terms of the generic f.

Working out the detailed solution to Eq. (1.22) gets a little bit technical.[2] Here we will try to explain the overall idea and invite you to work through some of the details in the Projects.

First, it is possible to show that if the source f is concentrated at a single point ($\vec{x} = \vec{x}'$) in space and only pops into existence for an instant at $t = t'$, i.e., if

$$f(\vec{x}, t) = \delta^3(\vec{x} - \vec{x}\,')\,\delta(t - t'), \tag{1.23}$$

then

$$\psi_{\vec{x}',t'}(\vec{x}, t) = -\frac{1}{4\pi}\frac{\delta\left(t - \left[t' + \frac{|\vec{x}-\vec{x}'|}{c}\right]\right)}{|\vec{x} - \vec{x}'|} \tag{1.24}$$

is a solution of Eq. (1.22).[3]

What this says, in words, is that the point source at $(\vec{x}\,', t')$ gives rise to a non-zero field ψ only where the argument of the δ-function on the right hand side vanishes – i.e., only at positions $\vec{x}$ and times t satisfying

$$t = t' + \frac{|\vec{x} - \vec{x}'|}{c} \tag{1.25}$$

– i.e., only at positions $\vec{x}$ and times t which could be reached by a signal propagating out at the speed of light from the source at $\vec{x}'$ and t'. This set of events should be thought of as a growing spherical shell that propagates outward from the source point. But it is often referred to as the "future light cone" of the source point $(\vec{x}', t')$ because, when plotted in a space-time diagram (with one of the 3 spatial dimensions suppressed to make room for time!) the points where the field ψ is affected form a cone. See Fig. 1.3.

Now, of course, realistic sources $f(\vec{x}, t)$ are not concentrated at individual points in space-time. But any realistically distributed source can always be written as an integral (think: sum) over such point-sources. And then, by the principle of superposition, the total field that these sources produce is just the sum over the fields that would be produced by each of the point sources taken individually. That means, if you think about it, that the total field produced at some point $(\vec{x}, t)$ will involve contributions from all the sources present on the *past light cone* of $(\vec{x}, t)$, i.e., from all the locations in space-time that could "broadcast" an influence outward at the speed of light that just reaches $(\vec{x}, t)$.

See Fig. 1.4 for an illustration of the implied picture, of electric charges interacting in a locally causal way by means of influences propagating through the fields. In

[2] It is explained, for example, in Chap. 6 of Jackson [4].

[3] Note also that there is a second solution proportional to $\delta\left(t - \left[t' - |\vec{x} - \vec{x}'|/c\right]\right)$. Equation (1.24) is called the "causal" solution because it describes charges and currents affecting the surrounding fields in the future; we set aside here the other solution, which describes charges and currents affecting the surrounding fields in the past. Finally, if you want to pursue the mathematics in a little more detail, it will help to know that Eq. (1.24), a solution to the wave equation for a delta-function source, is called the "Green's function" for the wave equation.

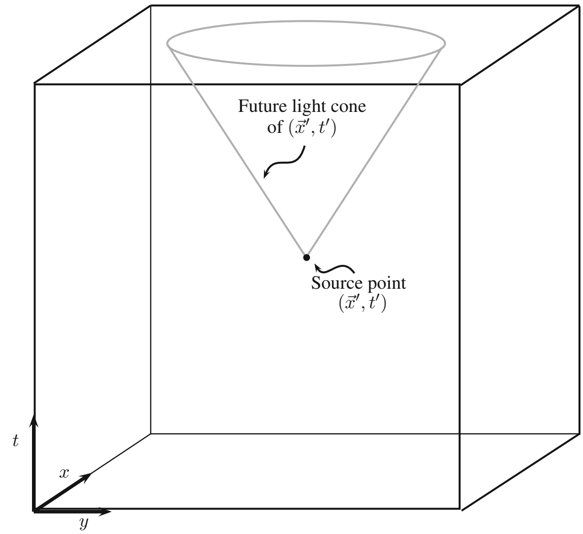


Fig. 1.3 A spatio-temporal point source $f(\vec{x}, t) = \delta^3(\vec{x} - \vec{x}')\delta(t - t')$ affects the field ψ at points on the future light cone of the source point $(\vec{x}', t')$

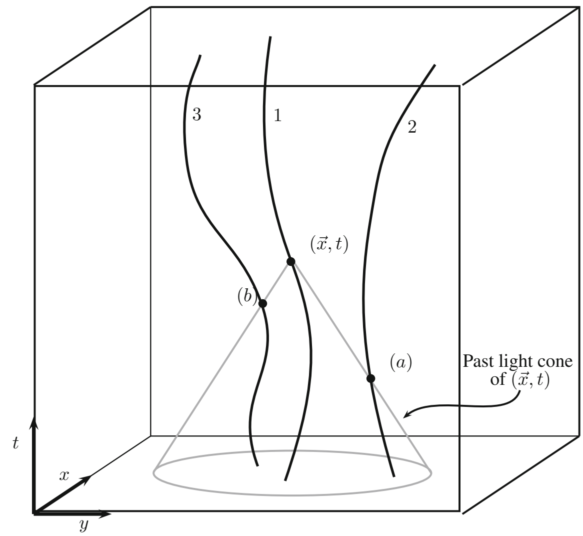


Fig. 1.4 What happens to particle 1 at the point $(\vec{x}, t)$ is determined, according to the Lorentz Force Law, by the fields $\vec{E}$ and $\vec{B}$ at that point. These, in turn, are determined by the sources at points on the past light cone of $(\vec{x}, t)$ – for example, as shown here, what particles 2 and 3 were doing at the points marked **a** and **b**

particular, if for example you want to know what particle 1 does at the point marked $(\vec{x}, t)$ in the Figure, this depends on the fields $\vec{E}(\vec{x}, t)$ and $\vec{B}(\vec{x}, t)$ at this point, which in turn are influenced by source terms on the past light cone, i.e., what particles 2 and 3 were doing at the points marked (a) and (b).

So far so good. Note, though, that in the language of differential equations, we have so far only been discussing the "particular solution" of Eq. (1.22) – that is, the contributions to the fields ψ which arise specifically from nonzero source terms f. But there is always in addition the so-called "complementary solution" – i.e., the solution of the corresponding homogeneous problem. But this part of the general solution is simple, because the corresponding homogeneous problem is just the ordinary wave equation whose solutions are electromagnetic waves that propagate at the speed of light.

Anyway, the point here is that the "general solution" of Eq. (1.22) can be understood as the sum of two things: first, contributions from electric charges and currents at points on the past light cone of the point in question, and then second, contributions from "freely propagating" electromagnetic waves (which perhaps were themselves created in the more distant past by wiggling charges, or which perhaps instead have just always been around since the beginning of time).

To summarize, then, Maxwellian electrodynamics is a completely *local* theory: what *happens* at a given point in space-time depends exclusively on events lying on (or inside) the past light cone of that event. This is really just a fancy and formal way of saying that causal influences always propagate, according to this theory, at the speed of light (or slower[4]). And this is in contrast to Newtonian gravity (at least naively interpreted) in which, as we saw, what *happens* at a given point in spacetime depends on things that are happening *at that same moment* arbitrarily far away.

Note that the local character of Maxwellian electrodynamics makes it compatible with Einstein's relativity theory in a way that Newton's theory of gravity is not. According to relativity, simultaneity is relative; that is, according to relativity, there is simply no objective fact of the matter about what set of events are *simultaneous* with a given space-time point. And so, from the point of view of relativity, the Newtonian gravitational idea that the force on a certain particle at a certain moment in time depends on the *instantaneous* configuration of all the other particles in the universe, is literally meaningless. Whereas the idea in Maxwellian electrodynamics, that the force on a certain particle at a certain moment depends on what other particles were doing *earlier*, by exactly the amount required for a signal to propagate at the speed of light to the particle in question, *is* perfectly compatible with relativity because the speed of light is, for relativity, an invariant quantity.

And note that it is essentially the fixing of this problem with Newtonian gravitational theory – namely, removing the nonlocality and thereby making it compatible with relativity – that Albert Einstein accomplished in his *general* theory of relativity in 1915.

As a way of summarizing all of this discussion, here is a nice statement by Einstein:

> The success of the Faraday-Maxwell interpretation of electromagnetic action at a distance resulted in physicists becoming convinced that there are no such things as instantaneous

[4]How could a causal influence in electrodynamics ever go slower than light, given what we've been saying here? Well in some fundamental sense it can't. But it can zig-zag back and forth through space-time in such a way that each individual "zig" or "zag" goes at the speed of light, but the overall average speed between the beginning and the end is much slower. As an analogy, if you throw a ball at a wall and wait for it to bounce back and hit you on the head, there is a sense in which the causal influence always propagates at whatever speed the ball was moving... but another sense in which the average speed, from your throw to your getting bonked in the head, is zero. The same kind of thing is possible with charged particles replacing you and the wall, and electromagnetic influences propagating at the speed of light replacing the ball.

> action at a distance (not involving an intermediary medium) of the type of Newton's law of gravitation. According to the theory of relativity, action at a distance with the velocity of light always takes the place of instananeous action at a distance or of action at a distance with an infinite velocity of transmission. This is connected with the fact that the velocity c plays a fundamental role in this theory [5, p. 47].

1.4 Bell's Formulation of "Locality"

In the last section, we sketched the process by which one can extract – from the fundamental equations of Maxwellian electrodynamics – the theory's "relativistic local causality" (or just "locality" for short). The success of relativity theory, though, has strongly suggested that this sort of relativistic local causality (i.e., the idea that causal influences should never propagate faster than the speed of light) is a necessary property of *any* acceptable physical theory. So it will be useful, going forward, to have a "generic" formulation of this idea, i.e., a concept of relativistic local causality that is not tied to any particular theory like Maxwellian electromagnetism. We therefore discuss here the formulation provided by John Stewart Bell, whose contributions to the foundations of quantum theory will show up throughout this book.

Bell begins by noting that Maxwellian electrodynamics has the following property: a complete description of the state of all the fields and charges, on a time-like "slice" (Σ) across the past light cone of some event at $(\vec{x}, t)$, will *determine* what happens at $(\vec{x}, t)$. The image of a time-like "slice" is illustrated in Fig. 1.5. One should remember, though, that this language – "slice", suggesting a two-dimensional region

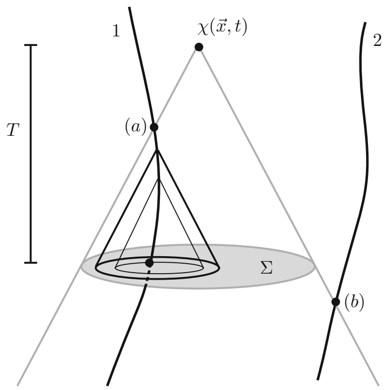


Fig. 1.5 In Maxwellian electrodynamics, a physical event χ at the point $(\vec{x}, t)$ – for example, the value of an $\vec{E}$ or $\vec{B}$ field, or the velocity of some particle that arrives there, or the electric charge density there – is uniquely determined by a complete specification ($\mathcal{C}_\Sigma$) of everything that's happening (i.e., the complete state of both fields and any charges and currents) in a horizontal "slice" through the backwards light cone of $(\vec{x}, t)$. The "slice" – Σ – is shown here as a shaded circle, although of course this really represents a three-dimensional spherical region of radius cT

that looks like the shaded circle in the Figure – is an artifact of the suppression of the third spatial dimension in these space-time diagrams. In fact, the shaded region in question represents a sphere of radius cT (where T is the time interval between the "slice" and the event in question at time t) centered at the event's spatial location $\vec{x}$.

Since, as we have argued in the previous section, causal influences in the theory always propagate at c (or slower), it's clear that all the causes of whatever happens at $(\vec{x}, t)$ must lie in this sphere. Nothing outside the sphere (at this earlier time) could *get* to $(\vec{x}, t)$ without propagating faster than light! But perhaps it is worth saying a little more about this to clarify how this idea connects with our earlier discussion.

To begin with, it is straightforward to see that there will be a contribution (corresponding to the complementary solution part of our earlier general solution) to fields at $(\vec{x}, t)$ from the electric and magnetic fields at the edge of Σ. In the three-dimensional picture, this corresponds to inward-propagating electromagnetic waves that will arrive at position $\vec{x}$ at time t. And then there will also be the contributions (corresponding to the particular solution from above) from electric charges and currents – the source terms – along the past light cone of $(\vec{x}, t)$. So, for example, in Fig. 1.5, what particle 1 is doing at the point marked (a) will influence what is happening at $(\vec{x}, t)$.

One might worry that "what particle 1 is doing at ... (a)" is not part of $\mathcal{C}_\Sigma$, our complete specification of events in Σ. But what particle 1 is doing at (a) is determined by $\mathcal{C}_\Sigma$. Think about it this way: what particle 1 is doing at (a) depends on where particle 1 was just prior to (a) and on the fields that influenced it at this earlier moment; these fields in turn depend on fields on the intersection of Σ with the past light cone of this earlier moment. See the thick black past-light-cone in the Figure. And then we can continue to step our way back along the world line of particle 1 in this same way: what it was doing at the earlier moment depends on where it was and what it was doing at an even-earlier moment, which in turn depends on the fields on the intersection of Σ with the past light cone of this even-earlier moment. (See the thin black past-light cone in the Figure.) And so on. You can then see how the whole structure of the world-line of particle 1 is determined, ultimately, by the state of particle 1 on Σ as well as the fields in a certain *part* of Σ that surrounds particle 1. And so, at the end of the day, *all* of the influences arriving at $(\vec{x}, t)$ – both direct ones and indirect ones – can indeed be traced back to physical facts about the states of the particles and fields on Σ. This is the sense in which any physical event $\chi(\vec{x}, t)$ is determined by $\mathcal{C}_\Sigma$ in Maxwellian electrodynamics.

(Note, by the way, that particles like 2 in the figure, whose world lines cross the past light cone of $(\vec{x}, t)$ prior to Σ are no problem: the influence of particle 2 on happenings at $(\vec{x}, t)$ coming from point (b) in the Figure are just "incoming electromagnetic waves" that we have already captured by including the state of the fields on the "edge" of Σ.)

In sum, by specifying the complete state of both fields and charged particles on the spacetime "slice" Σ, we have complete information about everything that is relevant to point $(\vec{x}, t)$ and therefore any physical fact pertaining to point $(\vec{x}, t)$ should be determined. We will formalize this by writing

$$\chi(\vec{x}, t) = f(\mathcal{C}_\Sigma) \tag{1.26}$$

where, again, χ is our generic name for some physical fact (like the value of some field or the velocity of some charged particle) at $(\vec{x}, t)$ and the right hand side means: some function of a complete specification of all the physical facts on Σ.

Equation (1.26) seems like it captures the idea of "relativistic local causality" for any *deterministic* theory. But remember, our goal here is to be completely generic. And as you probably already know, it has often been suggested that one of the lessons of quantum theory will turn out be that strict determinism must be abandoned. Bell proposed that we could modify the definition of locality, as follows, and arrive at something that still captures the idea of "relativistic local causality" but which is now applicable to both deterministic and indeterministic theories:

$$P\left[\chi_1|\mathcal{C}_\Sigma\right] = P\left[\chi_1|\mathcal{C}_\Sigma, \chi_2\right]. \tag{1.27}$$

This requires a bit of explanation. First of all, the left hand side means: the *probability* for some physical event χ_1 to occur at point 1, given a complete specification $\mathcal{C}_\Sigma$ of events on Σ (the "slice" through the past light cone). This is the same idea as before, but we just now speak of the probability of a certain event, rather than the event itself, since we don't want to presuppose that everything that happens is uniquely *determined* by events in the past light cone. (Of course, determinism is still allowed as a special case: in a deterministic theory, all of the probabilities will be either 1 or 0.) The right hand side is then meant to denote the probability assigned to the same event, but now with *both* $\mathcal{C}_\Sigma$ *and* some other event, χ_2 (which is at a point 2 which could not have been causally influenced by events in Σ) specified. See Fig. 1.6 for a picture.

After proposing (what we have here written as) Eq. (1.27), Bell writes:

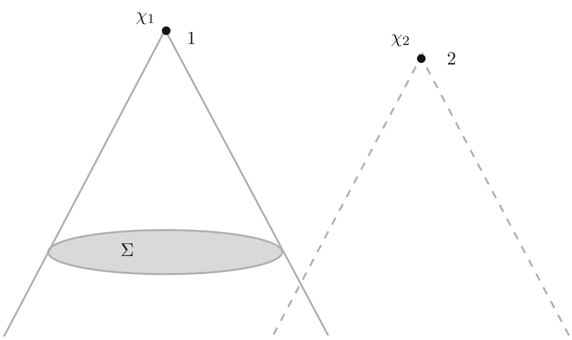


Fig. 1.6 According to Bell's generic definition of relativistic local causality, the probability of an event χ_1 at point 1, conditioned on a complete specification $\mathcal{C}_\Sigma$ of events in the region Σ (a "slice" across the backwards light cone of 1), should not be changed by specifying, in addition, some fact like χ_2 which cannot have been locally influenced by anything in Σ

> It is important that [Σ] completely shields off from 1 the overlap of the backward light cones of 1 and 2. And it is important that events in [Σ] be specified completely. Otherwise the traces in region 2 of causes of events in 1 could well supplement whatever else was being used for calculating probabilities about 1. The hypothesis is that any such information about 2 becomes redundant when [the state of things in Σ] is specified completely [6].

The importance of specifying events in Σ completely is pretty straightforward. If something is left out, then causal influences coming from the more distant past – influences which produce *correlations* between χ_1 and χ_2 – could result in χ_2 implying useful supplementary information (beyond that contained in the *incomplete* description of events in Σ).

But why is it so important that Σ should "shield" from 1 the overlapping past light cones of 1 and 2? In other words, why is it so important that the "other" event in Bell's formulation – χ_2 – be so far away that it could not have been causally influenced by events in Σ? See Fig. 1.7 and its caption for an explanation.

I hope that gives you a sense of why Eq. (1.27) seems to provide a good way of defining "relativistic local causality" in a completely general, a completely generic, way. It should come as no surprise that we will have occasion to use this formulation in later chapters.

Actually, a slight modification of Bell's formulation will also come in handy. To motivate this, one should understand that Bell's definition of locality basically amounts to the assertion that, once you provide a complete specification of events in a slice across its past-light-cone, the probability assigned to some physical event

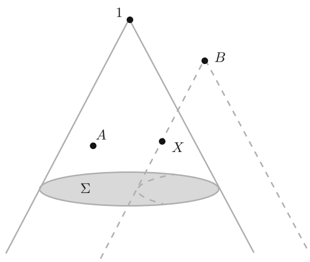


Fig. 1.7 In a non-deterministic theory, an event A may happen, despite not being determined to happen even by a complete specification of events in Σ, and then causally influence events at point 1. So specifying, in addition to $\mathcal{C}_\Sigma$, events like A that are in the past light cone of 1 but to the future of Σ, may indeed allow us to improve our predictions for events at 1. It might appear that, by contrast, events like B – which could be influenced by despite not being determined by $\mathcal{C}_\Sigma$ but which are outside the past light cone of 1 and hence could not locally influence events at point 1 – would *not* allow us to improve our predictions for events at 1. But this is incorrect: in an indeterministic theory, there might be an event X which is influenced (but not determined) by events in Σ, which then influences both B and happenings at 1. Specification of B can thus imply things about X which can in turn imply things about 1 that weren't already implied just on the basis of $\mathcal{C}_\Sigma$. This is why, in Bell's formulation of local causality, it is crucial that the other event (specification of which is not supposed to change the probability assigned to events at 1) should be outside the future light cone of Σ, i.e., this is why "[i]t is important that [Σ] completely shields off from 1 the overlap of the backward light cones of 1 and 2" [6]

should be *independent* of things happening outside the future-light-cone of that slice. Bell's formulation captures this independence by saying that the probability assigned to the event in question should not change depending on whether you do, or don't, specify such things.

But another way to capture this independence is by requiring that the probability be the same for any two different things that might be happening at this distant point:

$$P\left[\chi_1|\mathcal{C}_\Sigma, \chi_2\right] = P\left[\chi_1|\mathcal{C}_\Sigma, \chi_2'\right] \tag{1.28}$$

where χ_2 and χ_2' are two different possible versions of events at point 2 (in Fig. 1.6).

Let us illustrate all this with an example. It is not so interesting to take an example from Maxwellian electrodynamics, since that theory is manifestly local and hence respects both of our locality conditions: Bell's Equation (1.27) and our modification, Eq. (1.28). (You can try your hand at exploring this in the Projects if you are interested.) More interesting is seeing how our conditions can be used to diagnose the non-local character of Newtonian gravity. So let's take the following setup. There is one particle, of mass M, floating freely in empty space. There are two other particles, both of mass m, being held in place nearby (by some external non-gravitational forces) – one a certain distance to the left and the other the same distance to the right of our mass-M particle of interest. The symmetry implies that, if both of the mass m particles remain at rest, the net force on M will be zero and it will remain at rest. Whereas, if one of the mass-m particles is moved at the last second, the forces on M will no longer balance, and M will accelerate. See Fig. 1.8.

Thus, the probability for M to have an acceleration of zero at point 1 is different, depending on whether the mass-m particle on the right is moved, or not: Newtonian gravity explicitly violates Eq. (1.28). This is of course no surprise, but is a nice confirmation that Eq. (1.28) can be used to diagnose non-locality when it is present.

What about Eq. (1.27)? The problem is that, in a non-local theory, the left hand side, $P\left[\chi_1|\mathcal{C}_\Sigma\right]$, is simply not defined! Since the theory says the acceleration of the mass M particle depends on the locations of other distant particles at the moment in question, if we don't specify where the other distant particles are, there is no way for the theory to tell us what will (or might) happen to the particle in question. So although we expect that a genuinely local theory should respect both Eqs. (1.27) and (1.28), it is difficult to use Bell's formulation, Eq. (1.27), to explicitly diagnose the presence of non-locality in a theory. Our alternative formulation, Eq. (1.28), is nicer in this respect. It allows us to explicitly identify non-locality by seeing that the numbers on the two sides of the equation are different (rather than needing to try to compare something which isn't mathematically well-defined at all, to something that is). But whichever formulation we use, the non-locality of Newtonian gravity should be clear: the force on a particle (and hence its acceleration) depends on the location of distant particles at the exact moment in question.

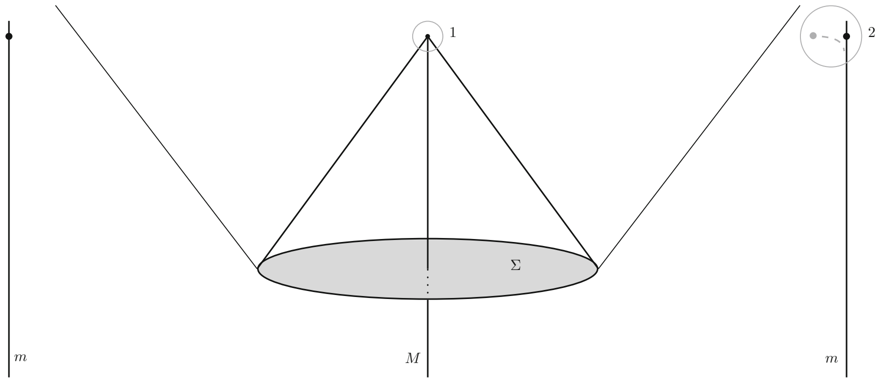


Fig. 1.8 A particle of mass M sits at rest between two particles of mass m that begin at equal distances from it in opposite directions. Let 1 be a space-time point through which the worldline of the mass-M particle will pass if everything remains as described. The complete specification, $\mathcal{C}_\Sigma$, of events on a slice, Σ, across the past-light-cone of 1, is rather simple: the particle is at rest at a certain point, and nothing else is going on! Take χ_1 to be the statement that the mass-M particle has an acceleration of zero at point 1. Suppose the mass-m particle on the *left* remains permanently fixed, but the mass-m particle on the *right* can either be *left* in place (call that χ_2), or pushed to the *left* (χ_2'). Then we have that $P\,[\chi_1|\mathcal{C}_\Sigma, \chi_2] = 1$. That is, if the mass-$m$ particle on the *right* remains fixed, the acceleration of the mass-M particle at point 1 will be zero with certainty. However, if the particle on the *right* is pushed to the *left*, the two gravitational forces on M will no longer add to zero, and M's acceleration at 1 will definitely not be zero: $P\left[\chi_1|\mathcal{C}_\Sigma, \chi_2'\right] = 0$. So we have a violation of Eq. 1.28

1.5 Ontology

"Ontology" is a fancy philosopher's word for "what really exists." In general, for the pre-quantum sorts of theories we've been looking at, "what really exists" according to the theory is fairly obvious and non-controversial. For example, according to Newtonian mechanics, the world is made of *particles* which move under the influence of (gravitational and probably other) forces they exert on one another. The picture is similar according to Maxwellian electrodynamics: the world is made of *particles* which interact with one another by means of the electric and magnetic *fields*. The ontology of Maxwellian electrodynamics, that is, includes particles and fields.

But there are often mathematically equivalent ways to formulate the basic laws of a theory, and this can sometimes raise questions about which formulation we should take seriously, as telling us in some relatively direct sense what really exists, physically, according to the theory.

For example, as you probably know, it is possible in electrodynamics to work with the so-called *potentials* (the scalar or electrostatic potential, and then the perhaps-less-familiar magnetic vector potential) instead of the *fields*. Let us briefly review some of this.

One of Maxwell's equations – the so-called "Gauss' Law for Magnetism" – reads $\vec{\nabla} \cdot \vec{B} = 0$. Since the divergence of any curl is identically zero, we can ensure that this equation is automatically satisfied if we introduce a magnetic vector potential $\vec{A}$ related to the magnetic field by

$$\vec{B} = \vec{\nabla} \times \vec{A}. \tag{1.29}$$

This allows us to re-write Faraday's Law as

$$\vec{\nabla} \times \left(\vec{E} + \frac{\partial \vec{A}}{\partial t} \right) = 0. \tag{1.30}$$

But then, because the curl of a gradient is identically zero, we can ensure that this equation is automatically satisfied if we introduce a "scalar potential" ϕ satisfying

$$\vec{E} + \frac{\partial \vec{A}}{\partial t} = -\vec{\nabla}\phi. \tag{1.31}$$

(The minus sign is just for convention/convenience.) Equivalently, if we write the electric field in terms of the scalar and vector potentials as

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \tag{1.32}$$

we guarantee the satisfaction of Faraday's Law. That then leaves the other two (non-homogeneous) Maxwell equations, which can now be re-written, in terms of ϕ and $\vec{A}$ as, respectively,

$$\nabla^2 \phi + \frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{A}) = -\frac{\rho}{\epsilon_0} \tag{1.33}$$

and

$$\nabla^2 \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} = -\mu_0 \vec{j} + \vec{\nabla}(\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t}). \tag{1.34}$$

These both look a little messy and complicated, but actually we have not yet taken full advantage of the freedom afforded by the potentials. In particular, since all that is really required of the potentials is that their various *derivatives* give the fields $\vec{E}$ and $\vec{B}$, there is an element of arbitrariness that we can leverage to make the equations look nicer.

For example, it's clear from the defining relation $\vec{B} = \vec{\nabla} \times \vec{A}$ that (since, again, the curl of a gradient is identically zero!) we could change $\vec{A}$ by the gradient of an arbitrary scalar function without affecting $\vec{B}$. That is,

$$\vec{A} \rightarrow \vec{A} + \vec{\nabla}\lambda \tag{1.35}$$

leaves $\vec{B}$ unchanged.

"But" (I can hear you saying) "changing $\vec{A}$ in this way *would* affect the electric field!" That's true, but we can "undo" that change by *also* requiring that, when $\vec{A}$ is shifted as in Eq. (1.35), we also shift the scalar potential ϕ as follows:

$$\phi \to \phi - \frac{\partial \lambda}{\partial t}. \tag{1.36}$$

It is easy to see that, if both ϕ and $\vec{A}$ are shifted in these ways, the fields $\vec{B}$ and $\vec{E}$ are unaffected. These "shifts" in the potentials are called "gauge transformations" and the idea is that there is a whole *equivalence class* of potentials (corresponding to all possible scalar functions λ) that correspond to the same field configurations.

This means that we are free to *choose* a particular set of potentials that makes some of the terms in Eqs. (1.33) and (1.34) disappear. We will briefly discuss two of these possible choices.

The first choice is to choose potentials satisfying the so-called "Lorentz gauge" condition:

$$\vec{\nabla} \cdot \vec{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} = 0. \tag{1.37}$$

With this choice, the wave equations satsified by ϕ and $\vec{A}$ take on the following particularly simple forms:

$$\nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0} \tag{1.38}$$

and

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = -\mu_0 \vec{j}. \tag{1.39}$$

These are nice because they are precisely of the general form, Eq. (1.22), that we investigated earlier. It's clear, that is, that in Lorentz gauge, the effects of charges on the potentials propagate outward at the speed of light.

But here is another perfectly valid choice of gauges, the so-called "Coulomb gauge" in which

$$\vec{\nabla} \cdot \vec{A} = 0. \tag{1.40}$$

This turns out to imply the following dynamical equations (with source terms) for ϕ and $\vec{A}$:

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = -\mu_0 \vec{j} + \frac{1}{c^2} \vec{\nabla} \frac{\partial \phi}{\partial t} \tag{1.41}$$

and

$$\nabla^2 \phi = -\rho / \epsilon_0. \tag{1.42}$$

Look in particular at the latter equation. It is just like Eq. (1.38), but with the propagation speed c set to infinity so that the term involving the second derivative with respect to time vanishes. So it implies that the scalar potential $\phi(\vec{x}, t)$ is determined by the *instantaneous* configuration of charges $\rho(\vec{x}, t)$ – in just the same way that the gravitational force on a given particle in Newtonian mechanics is determined by the instantaneous configuration of other particles. So the object ϕ – in Coulomb gauge

– is nonlocal! It changes *instantaneously*, without any speed-of-light delay, if some distant charge is wiggled.[5]

But so what? We should only be *bothered* by this kind of nonlocality, from the point of view of consistency with relativity, if we take ϕ as corresponding to something that is physically real. So does it? The usual answer is: no! The potentials ϕ and $\vec{A}$ are mere calculation devices – it is instead the fields $\vec{E}$ and $\vec{B}$ which we should take as directly corresponding to "physical stuff that really exists" according to the theory. To put this in terms of our notation from the previous section, it would be appropriate to let χ stand for (say) the value of $\vec{E}$ or $\vec{B}$ at some point, but it would not be appropriate to let χ represent the value of ϕ or $\vec{A}$ at some point. In order to function as intended, it is important that the χs in Bell's formulation of locality represent quantities that are endorsed, by the theory in question, as physically real.

Bell gives a memorable analogy here:

> The situation is further complicated by the fact that there *are* things which *do* go faster than light. British sovereignty is the classical example. When the Queen dies in London (long may it be delayed) the Prince of Wales, lecturing on modern architecture in Australia, becomes *instantaneously* King. (Greenwich Mean Time rules here.) And there are things like that in physics. In Maxwell's theory, the electric and magnetic fields in free space satisfy the wave equation
>
> $$\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} - \nabla^2 \mathbf{E} = 0,$$
> $$\frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2} - \nabla^2 \mathbf{B} = 0$$
>
> ...corresponding to propagation with velocity c. But the scalar potential, if one chooses to work in 'Coulomb gauge', satisfies Laplace's equation
>
> $$-\nabla^2 \phi = 0$$
>
> ...corresponding to propagation with infinite velocity. Because the potentials are only mathematical conveniences, and arbitrary to a high degree, made definite only by the imposition of one convention or another, this infinitely fast propagation of the Coulomb-gauge scalar potential disturbs no one. Conventions can propagate as fast as may be convenient. But then we must distinguish in our theory between what is convention and what is not [6].

This is a point that will occupy us considerably in the coming weeks. In quantum theory, which objects in the formalism are we supposed to take seriously, as corresponding to things that are physically real, and which are (like the scalar potential and British sovereignty!) bound up in some way with human knowledge or conventions? In particular, what is the ontological status of the quantum mechanical wave function?

---

[5] For a nice analysis of how, in Coulomb gauge, the electric field retains its local character even though the scalar potential is nonlocal, see Ref. [7].

1.6 Measurement

The two example theories we've been discussing are obviously regarded as good (if now slightly dated) scientific theories which have extremely favorable records of correctly predicting measured or observed phenomena in nature. It will be helpful to think a little bit about how, exactly, these theories achieve this status.

In the case of Newtonian mechanics, the situation is pretty straightforward. Newtonian mechanics is a theory about the motion of Particles – and, consequently, macroscopic assemblages of Particles which can in appropriate situations be treated as particles. Certain such particles are directly visible to us, so we can just check – by literally looking – to see if the particle is located where the theory says it will be located. If so then we say that the theory's prediction has been confirmed by observation.

For example, suppose you throw a ball up in the air from some initial height and with some initial speed. If you know something about the (say, gravitational and air drag) forces that will act on it, you can solve $\vec{F}_{\text{net}} = m\vec{a}$ for the ball and calculate, according to the theory, things like the maximum height the ball will reach and the time it will spend in the air before hitting the ground. In the case of the maximum height, the theory's prediction can perhaps be compared against a literal, perceptual observation: you just look and see how high the ball in fact goes before turning around and heading back down.

But this is, at best, pretty rough. A careful *measurement* of the ball's maximum height will require some additional sophistication. For example, you might set up some meter sticks and adjust your viewing perspective so that you can read off the maximum height to some precision by seeing exactly which mark on the meter stick lines up with the top of the ball at the moment it reaches its maximum height. One might for example also film the motion of the ball, with the meter stick in the background; looking through the individual frames later, one could identify the specific frame in which the ball reaches its maximum height and then make a more precise determination of that height using the image of the meter stick in the background. And of course more sophisticated techniques are also possible, but this indicates the overall pattern.

A measurement of, for example, the ball's time aloft will follow a similar pattern. Direct, unaided visual inspection perhaps gives some rough indication of the duration of the ball's flight, but a more precise measurement would involve additional care and equipment. For example, one might arrange for the ball's launch to trigger a stopwatch, whose second hand at that moment begins a rapid, steady rotation which is then triggered to halt when the ball strikes the ground. The final location of the second hand (which one may inspect at a convenient later time) then indicates the time the ball spent in the air.

The maximum height and time aloft of a launched ball are both things that are, in some sense, directly observable – we can "measure" them in a kind of rough and qualitative way by literally just watching the process unfold in real time. And then we have been discussing ways in which those rough perceptual observations can be improved upon by using more sophisticated measuring equipment. There are, however, things that theories talk about which are not even in principle directly observable. In such cases, "measuring" or "observing" the fact in question (to, say, test a prediction of the theory) *requires* more sophisticated measuring equipment.

For example, suppose you want to test the Maxwellian electrodynamics prediction for what the electric field is between the two plates of a charged capacitor. You work out, based on the amount of charge that's present, etc., what the electric field should be according to the theory. How do you then measure this to see if the theory's prediction is correct? You certainly can't "just look and see".

But you can arrange for the electric field to leave its mark on something that you *can* just see. For example, you might stick a particle of mass m and electric charge q, at rest, at the place where you want to know the electric field. The particle will experience an electric force $q\vec{E}$ and will hence accelerate. If we can measure the particle's acceleration $\vec{a}$, we can then infer that the electric field was given by

$$\vec{E} = \frac{m}{q}\vec{a}. \tag{1.43}$$

Measuring the acceleration is, in turn, straightforward. For example, you could set a ruler next to the particle and observe the change $\Delta\vec{x}$ in its position after a short period of time Δt. Basic kinematics then tells us that

$$\vec{a} = \frac{2\Delta\vec{x}}{\Delta t^2}. \tag{1.44}$$

So, by measuring the position of a particle at two different times (and knowing the mass and charge of the particle) we are able to indirectly measure the electric field.

This puts us in a position to step back and see the point of this discussion. On the one hand, a theory that makes no testable predictions is in some sense clearly worthless. If you can't measure or observe *anything* that the theory says things about, the theory is fundamentally cut off from the structure of empirically-grounded knowledge and really has no meaningful content at all. But on the other hand, it is not the case that everything a theory says must somehow be directly observable. It is perfectly reasonable for theories to postulate the existence of "invisible" things (like electric and magnetic fields, atoms, neutrinos, etc.). But then the theory should provide a consistent account of the interactions of those postulated invisible things with other things (perhaps made of the invisible things!) that are visible, so that overall the theory connects with the given world of direct perceptual experience and thereby makes testable predictions.

As a kind of paradigmatic concretization of this idea, we will often speak of measurements having their outcomes registered in the position of a "pointer". One should think here of a sort of black-box measuring instrument, along the lines indicated in Fig. 1.9, whose internal mechanism provides a causal link between some physical quantity that is being (perhaps indirectly) measured (e.g., the time aloft of a ball, or the magnitude of the electric field at a certain point, or the energy of a neutrino, or ...)

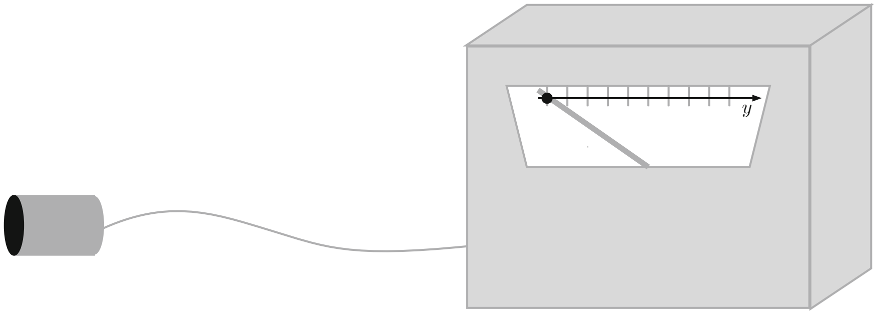


Fig. 1.9 A schematic measuring device whose probe end (on the *left*) can be arranged to interact with some (perhaps microscopic/invisible) system of interest. The outcome of the measurement is then registered by the position y of the device's pointer. This is a relatively accurate picture of how some real measuring devices work, but also captures in essentialized terms an important point about *any* kind of measurement: at the end of the day, the outcome is registered in the configuration of some directly-observable (macroscopic) object (e.g., the position of the hands of a stopwatch, the distribution of ink on a printout, etc.)

and the position of a pointer or needle that swings back and forth against a calibrated background to indicate the value in question.

This may seem like a very specialized kind of case, but actually it captures the essential idea of measurements quite generally. For example, the position of a ball (whose maximum height one wants to measure) is its own pointer. The second hand on the stopwatch functions as the pointer for measuring the ball's time aloft. The charged particle (whose final position allows one to determine its acceleration and hence the field that caused that acceleration) functions as the pointer for the measurement of the electric field that we discussed before. In general, *any* measurement must produce some effect in the configuration of some directly-observable macroscopic object, from whose final configuration we "read off" the result of the measurement. That is, any measurement must involve something that can be interpreted as a pointer.

From the point of view of assessing candidate theories, this discussion suggests several criteria. On the one hand, we should not demand too much from our theories. For example, we should not insist that everything postulated by a theory be somehow directly observable or measureable. Direct perception gives us *some* information about the structure of the world, but it does not give us *everything*; we should thus expect that, as science develops, theories should increasingly need to postulate invisible, microscopic objects, the empirical justification for which lies in the role of those postulated invisible objects in correctly accounting for the behavior of things (pointers!) that are directly visible.[6]

[6] Another way we might demand too much of a theory would be to demand that it not only account for the behavior of directly visible things like pointers, but that it somehow account for our conscious experiences of such things. The truth is that nobody understands how consciousness emerges and in particular how specific conscious experiences arise from the interactions among the external objects one perceives, one's perceptual apparatus (including eyes, brain, etc.), and the faculty of

On the other hand, we should also not demand too little from our theories. There *are* things – like the positions of "pointers" on lab equipment, whether or not a bomb has exploded, and the vitality of a certain cat – which are directly available in ordinary sense perception, and which any candidate account of the microscopic world should, in principle, be able to account for. Theories might validly postulate all kinds of crazy-sounding and counter-intuitive things, but at the end of the day, if a theory predicts the wrong thing for where pointers should point (or somehow cannot account for the existence of pointers that point at all) it cannot be correct.

1.7 Abstract Spaces

Let us raise one final point about the pre-quantum theories we've been discussing. In the case of both Newtonian mechanics and Maxwellian electrodynamics, everything we have talked about so far can be understood in terms of ordinary three-dimensional space (and/or four-dimensional space-time). The particles, for example, that are a central part of the ontology for these theories, "live" in three-dimensional space. As do the fields of Maxwellian electrodynamics. But it is worth pointing out that there are various mathematical re-formulations of some of these ideas, in which various sorts of "abstract spaces" are used.

For example, it is sometimes useful to use "phase space" to talk about the kinematics and dynamics of particles. For a single particle moving in three dimensional physical space, the phase space is a *six*-dimensional space whose axes correspond to the x, y, and z coordinates of the particle and also the three components of the particle's momentum: p_x, p_y, p_z. It is of course difficult to visualize a six-dimensional space, but if we consider a toy model system in which a particle is confined to move only along a single spatial dimension (say, x), then the phase space is two-dimensional (axes x and p_x) and we can easily visualize it.

Take, as a simple example, the case of a particle that is just moving inertially: $p_x =$ constant, so then $x = p_x t/m + x_0$. We can plot the "trajectory" of the particle through phase space; see Fig. 1.10.

A slightly more interesting case is a one-dimensional harmonic oscillator. The spatial coordinate x oscillates sinusoidally (say, about $x = 0$) and the momentum p_x also oscillates sinusoidally but out of phase with the position: the momentum is big (either positive or negative) when the position is zero, and vice versa. The "trajectory" of the particle through phase space is shown in Fig. 1.11. The "trajectory" is a closed orbit – an elliptical curve. See if you can figure out which direction the

(Footnote 6 continued)
consciousness that somehow emerges from brain structure and function (if, indeed, it is separable from these things at all). These are deep and difficult questions that are largely outside the scope of physics. As far as physics is concerned, if a theory accounts for the existence of macroscopic material objects which possess gross, coarse-grained properties consistent with what we are given in ordinary perception – i.e., if a theory gets the pointer positions right – we should regard it as perfectly acceptable and empirically adequate.

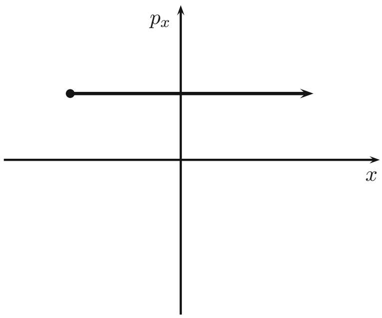
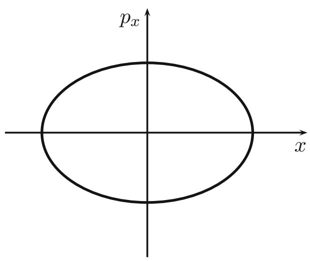


Fig. 1.10 The "trajectory" of a free particle (moving with constant momentum) through phase space: at $t = 0$ the particle begins, at a negative value of x, with a positive momentum which stays constant as the particle moves. So its "path" through phase space looks like the solid line and would continue indefinitely to the right as long as the particle continues moving with unchanged momentum

Fig. 1.11 The "trajectory" of a one-dimensional harmonic oscillator through phase space is a closed elliptical curve



phase space point moves around the ellipse. (Hint: when p_x is positive, is x increasing, or decreasing?)

Another type of "abstract space" that is sometimes useful to think about in the context of classical physics – and which we will confront again in the context of quantum mechanics – is "configuration space". For a single particle, configuration space is just the same as physical space: it is all the places the particle could be. But in a situation involving (say) two particles (moving, for simplicity, in one spatial dimension), there is the set of all places (call them x_1) where the first particle might be, and then also the set of all places (call them x_2) where the second particle might be. The configuration space is then the set of all possible configurations of the two particles jointly – that is, configuration space is a (here) two-dimensional space whose axes are x_1 and x_2.

Let's again consider two examples, one really simple and one at least slightly more interesting. Suppose, for example, that we have two particles moving in (the same) one-dimensional space. Except, let's say, they aren't actually moving, but are instead both just sitting there. In particular, particle 1 is sitting at a small positive value of x, and particle 2 is sitting at a somewhat larger negative value of x. Figure 1.12 shows this situation, both in (regular) physical space, and then again in configuration space.

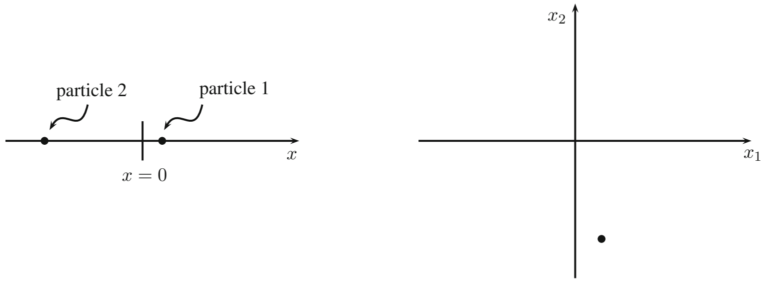


Fig. 1.12 On the *left* are two particles ("particle 1" and "particle 2") sitting at different points along the x-axis. This same situation is represented in the abstract configuration space as a single dot whose coordinates correspond to the positions (in physical space) of the particles

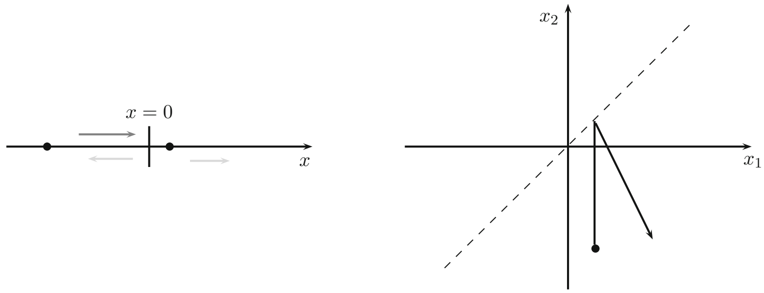


Fig. 1.13 The same two particles as before, but now particle 2 (which starts on the *left*) is given a kick at $t = 0$ so it moves to the *right* (its initial velocity is indicated by the *dark gray arrow*) until it collides with particle 1. The post-collision velocities are indicated with the *light gray arrows*. This same process is shown, on the *right*, in the abstract configuration space. Note the *dashed line* at $x_1 = x_2$ corresponding to the particles being at the same location, i.e., colliding

As a slightly more interesting example, suppose now that somebody comes in and gives particle 2 a kick so that it starts moving to the right. Eventually it runs into particle 1 and, let's suppose, they collide elastically. Suppose particle 1 has a greater mass than particle 2, so that after the collision particle 2 recoils back out to the left, whereas particle 1 drifts off slowly to the right. This process is depicted – both in physical space and in the abstract configuration space – in Fig. 1.13.

What is the point of discussing these abstract spaces? Well, as I said, they are sometimes useful ways to depict a certain physical process to gain some intuition. And there are even complete reformulations of Newtonian mechanics where things are formulated in terms of one of these abstract spaces – for example, if you've taken a course in classical mechanics, you have probably encountered "Hamiltonian mechanics" which is basically a way of re-writing $\vec{F}_{net} = m\vec{a}$ in terms of energy quantities. The basic dynamical equations in this Hamiltonian formulation of mechanics are first-order (in time) differential equations for the coordinates of a system in phase space. There is also something called the "Hamilton-Jacobi" formulation of mechanics which involves something like a time-dependent "field" on configuration space, which influences or guides the actual configuration point (representing the positions of all the particles composing the system) through that space. We won't go into these things in any detail here, but it is good to be aware that they exist.

At this stage, the main take-home lesson from this discussion is this: don't confuse any of these abstract spaces with regular old physical space! For example, if you are studying a one-dimensional harmonic oscillator, and sketch its phase space trajectory as in Fig. 1.11, you should not ask: "What force provides the centripetal acceleration which holds the particle in this elliptical orbit?" Or similarly, you should not ask, about the particle collision process depicted in Fig. 1.13, "That dotted line in the picture that the particle bounced off of... what's it *made of*??" Those sorts of questions don't actually make any sense, and would seem to be based on simply forgetting that the space in question is an abstract one. If, in these kinds of situations, you want to know what is really going on, physically, you need to translate the abstract representation back into direct, literal, physical-space terms. For example, the dotted line in the previous Figure is not really a "thing" at all, but instead a kind of abstract representation of the strong repulsive forces that the two particles exert on each other if they get too close together in physical space. And the only force present in the case of the simple harmonic oscillator is the force exerted on the particle by a spring (or whatever) as it moves back and forth along a line – there is nothing like an elliptical orbit at all, if that means some literal material particle moving along a certain curved path through a two-dimensional physical space.

This of course all seems so clear and obvious as to be almost embarrassing to have to say. But as you'll see (especially in Chap. 5) confusion will arise around these kinds of issues when we get to quantum mechanics – to which we will turn very soon!

Projects:

(1.1) Show that a spherically-symmetric distribution of mass (e.g., a thin spherical shell composed of innumerable massive Particles) exerts the same gravitational force on an outside particle as the force that would be exerted by a single Particle with the same total mass as the shell and located at the center of the shell. Explain how this is an example of the theory explaining how and why a large (spherically symmetric) assemblage of Particles can be treated as a particle.

(1.2) Show that a rigid body (like a bunch of Particles glued together) will obey Newton's 2nd law ($\vec{F} = m\vec{a}$) if the individual particles do. Note that a rigid body can *rotate* in addition to moving translationally, so you should *not* assume that all the individual Particles have the same acceleration $\vec{a}$. Indeed, the main problem here is to figure out precisely how "the acceleration of the rigid body" can be defined to make something like Newton's 2nd law true. Hint: it might be helpful to consider the *center of mass* coordinate, $\vec{R} = \frac{1}{M_{\text{total}}} \sum_i m_i \vec{r}_i$, and its various time-derivatives. Explain how this is an example of the theory explaining how and why a large assemblage of Particles can be treated as a particle.

(1.3) Consider two equal-mass stars in a binary star system, each making a circular orbit about their mutual center-of-mass point. Now suppose that gravitational forces are given by Eq. (1.1), but with $\vec{r}_j$ being the position of the distant mass at a slightly earlier time (such that a signal emitted by it at that earlier time would just arrive at the mass in question now). Draw a careful diagram showing the gravitational forces acting on each star at some particular moment. Do the forces respect Newton's third law? Will the total momentum of the two-star system be conserved (assuming no forces act from the outside)? How about the total angular momentum? Can you think of another similar situation in which the total translational momentum would not be conserved? What do you make of all this?

(1.4) Give a simple example of a system of masses interacting via Newtonian gravitational forces, and show that/how the motion of the masses would be different if one used a different slicing of space-time into simultaneity slices. (That is, show that Newtonian mechanics with instantaneous, action-at-a-distance gravitational interactions is incompatible with relativity, since it requires a dynamically privileged notion of simultaneity.)

(1.5) The mathematical parallel between Newton's law of gravity and Coulomb's law of electrostatics suggests that a relativistic theory of gravity could be developed by, in effect, copying Maxwell's equations. Play around with this and see how far you can get. (Hint: the gravitational analog of the electric field $\vec{E}$ is the gravitational field $\vec{g}$, which has units of acceleration. So the gravitational analog of Gauss' Law should be something like $\vec{\nabla} \cdot \vec{g} \sim \rho_m$ where ρ_m is the mass density. What should the proportionality constant be in order to reproduce Newtonian gravity? After you figure out the gravitational analog to Gauss' Law, you can try to work out consistent gravitational analogs to the three other Maxwell equations as well!)

(1.6) Flesh out the equivalence between Coulomb's law and Gauss' law by explaining in detail how to solve for $\vec{E}$ in Gauss' law when $\rho = q\,\delta^3(\vec{x})$.

(1.7) Work through the details of deriving wave equations for $\vec{E}$ and $\vec{B}$ from Maxwell's equations. (Assume empty space, i.e., $\rho = 0$ and $\vec{j} = 0$.) Show that there are plane-wave solutions of the form $\vec{E}(\vec{x}, t) = \vec{E}_0 \sin(\vec{k} \cdot \vec{x} - \omega t)$, and similarly for $\vec{B}$. Are the waves transverse, or longitudinal? How do you know? What is the relationship between $|\vec{k}|$ and ω? What are the phase and group velocities of the waves in terms of ϵ_0 and μ_0?

(1.8) Let's try to understand better how Eq. (1.24) is a solution of the wave equation with a delta-function source. For simplicity, suppose the source point is at $\vec{x}' = 0$ and $t' = 0$ so that the differential equation in question reads

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \delta^3(\vec{x})\delta(t). \tag{1.45}$$

The solution ψ should be spherically symmetric, i.e., should be a function of r and t only. (a) Show that, for $r \neq 0$, any function of the general form

$$\psi(r,t) = \frac{g(t - r/c)}{r} \tag{1.46}$$

solves Eq. (1.45). (b) The correct function g should also satisfy Eq. (1.45) at $r = 0$. To ensure this, integrate both sides of Eq. (1.45) over a spherical volume of radius R centered at $r = 0$. Show that taking the $R \to 0$ limit then gives $-4\pi g(t) = \delta(t)$, i.e.,

$$\psi(r,t) = -\frac{1}{4\pi}\frac{\delta(t - r/c)}{r} \tag{1.47}$$

as claimed in the text.

(1.9) When he is presenting his formulation of locality, Bell gives an analogy to the boiling of an egg. You put the egg into the boiling water and set the timer for (say) 5 min. Then, 5 min later, "[t]he ringing of the alarm establishes the readiness of the egg." That is, the two events are *correlated*. However, the ringing of the alarm does not *affect* the egg. Bell explains: "if it is already given that the egg was nearly boiled a second before, then the ringing of the alarm makes the readiness no more certain." Draw a spacetime diagram; connect "the ringing of the alarm", "the readiness of the egg", and its being "already given that the egg was nearly boiled a second before" with the terminology χ_1, χ_2, and $\mathcal{C}_\Sigma$; and explain how the last sentence captures the locally causal character of the physical processes involved in this example.

(1.10) An unstable particle is heading for a particle detector which will "click" if the particle hits it. Given the state of the particle at some earlier time, suppose there is a 50% probability of its *not* decaying first and hence hitting the detector: $P\,[\text{click}|\mathcal{C}_\Sigma] = 1/2$. On the other hand, if the particle does decay before arriving at the detector, the decay products might themselves be detected and hence indicate that the original particle will not be detected by the original detector. So, for example, $P\,[\text{click}|\mathcal{C}_\Sigma, \chi_2] = 0$, where χ_2 denotes the successful detection of one of the decay products. See Fig. 1.14. Does the non-equality of the two conditional probabilities here imply a violation of Bell's locality condition? (One would hope not, since there is clearly nothing nonlocal happening here. On the other hand, this seems to be a case where information from outside the past light cone of the event in question, does affect the probabilities assigned to that event.) Explain.

(1.11) Make up an example, maybe along the lines of the example involving Newtonian gravity from the main text, to show that Maxwellian electrodynamics gets (correctly) diagnosed as "local" both by Bell's formulation and the modified formulation.

(1.12) Is there a way of choosing a gauge such that the vector potential $\vec{A}$ propagates with infinite speed, the way ϕ does in the Coulomb gauge? Explain.

(1.13) A standard introductory physics problem involves analyzing the "ballistic pendulum" in which a block of mass M hanging from a string of length L absorbs an incoming bullet of mass m moving at some unknown initial speed

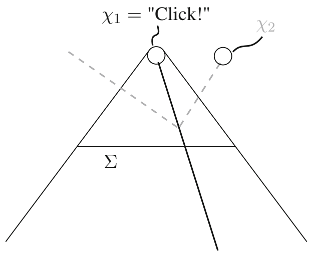


Fig. 1.14 Space-time diagram for the events described in Project 1.10

v_0. The bullet-and-block then swing up together, with the string eventually making some maximum angle θ with respect to the vertical. By observing θ one can thereby determine the initial speed of the bullet. Work out this relationship (i.e., solve for v_0 as a function of θ, m, M, and L) and explain how this method of measuring the bullet's speed fits into the general scheme introduced in the text in which the outcome is registered in some directly observable "pointer".

(1.14) Pick a measuring apparatus of interest to you (maybe something that you've used in a physics lab course or research, maybe the accelerometer in your iPhone, or just something else you're interested in) and learn more about how it actually works. Explain whether (and, assuming so, how) the actual device fits into the general scheme introduced in the text in which the outcome is registered in some directly observable "pointer".

(1.15) Two equal-mass gliders are floating on a (nearly) frictionless air track in an introductory physics classroom. The track is equipped with elastic bumpers on both ends, and the two gliders have elastic-collision attachments so they will bounce when they collide. One glider is initially at rest, while the other is given an initial velocity. Draw a space-time diagram showing world lines for both gliders as well as the two bumpers on the ends of the track. Now draw the "trajectory" of the two-glider system through the two-dimensional configuration space.

References

1. I. Newton, *Opticks* (Dover Publications Inc, New York, 1952)
2. J. Andrew, Newton's Philosophy, in *The Stanford Encyclopedia of Philosophy*, ed. by E.N. Zalta (2009), http://plato.stanford.edu/archives/win2009/entries/newton-philosophy/
3. A. Einstein, Autobiographical Notes, in *Albert Einstein: Philosopher-Scientist*, ed. by P.A. Schilpp (Harper and Row, New York, 1959)
4. J.D. Jackson, *Classical Electrodynamics*, 2nd edn. (Wiley, New York, 1975)
5. A. Einstein, *Relativity: The Special and the General Theory* (Penguin Classics, New York, 2006)
6. J.S. Bell, La Nouvelle Cuisine, reprinted in *Speakable and Unspeakable in Quantum Mechanics*, 2nd edn. (Cambridge University Press, Cambridge, 2004)
7. O.L. Brill, B. Goodman, Am. J. Phys. **35**, 832–837 (1967)