A Field Theory Primer

With a Comparison between Electrostatics and Newtonian Gravity

M. A. Reynolds
Embry-Riddle University

In the 1830s and 1840s, Michael Faraday discovered electric induction (where moving magnets can cause electric currents to flow in nearby circuits), and accounted for the "action-at-a-distance" character of the electric and magnetic forces in terms of *lines of force*, and eventually in terms of the concept of a FIELD. He perceived the space around a magnet to be permeated by a magnetic field, and it was this field that caused the electric current to flow. Maxwell described Faraday's thinking in the following way

> "[Faraday] saw lines of force traversing all space where the mathematicians saw centers of force acting at a distance. Faraday saw a medium where they saw nothing but distance: Faraday sought the seat of the phenomena in real actions going on in the medium, they were satisfied that they had found it in a power of action at a distance impressed on the electric fluids."[1]

In this primer, I will develop a field theory for both electrostatics (the electric field) alongside gravitation (the gravitational field), and show how the familiar notion of *forces* that act at a distance can be replaced by the unfamiliar notion of *fields*. Because the electric and gravitational forces are mathematically identical, a comparison between the two makes the concept of an electric field easier to assimilate because gravitation is so familiar.

I will only cover *static* field theory, that is, electrostatics and Newtonian gravity, and I will not extend the discussion to electromagnetics, general relativity, or quantum field theory. It turns out that in these more advanced topics, field theory is essential, but for the static forces discussed here, the two descriptions (forces and fields) are equivalent.[2] In many situations, though, the use of the field concept makes the mathematics more tractable. In addition, a thorough understanding of static field theory provides a solid foundation for the future study of electromagnetics, general relativity, and quantum field theory.

[1] *Treatise on Electricity and Magnetism*, James Clerk Maxwell, 1861.

[2] As McMullin (Ref. 1) states, "Historians of science disagree ... on whether a theory consistent with the notion of action at a distance ought qualify as a 'field' theory."

1 Preliminaries

There are two axioms (or "facts of experience") which underly both electrostatics and gravitation (in fact, all of the formalism in this primer follows mathematically from these two axioms). They are (1) the INVERSE SQUARE FORCE LAW between two objects, and (2) the principle of SUPERPOSITION. In the mathematical development that follows, the equations describing electrostatics will be in a column on the left side of the page, and the corresponding equations describing gravitation will be in a column on the right. Any equation that applies to both in general will be placed in the center. In addition, the text will describe both theories at the same time, with the words and symbols pertaining to gravitation in parentheses immediately following those pertaining to electrostatics.

<u>**Electricity**</u>	<u>**Gravitation**</u>

Force The first experimental fact is expressed by Coulomb's Law for electricity, and Newton's Universal Law of Gravitation, which can be stated as follows:

> "The force exerted by a charge(mass) $Q(M)$ that is fixed at the origin on a charge(mass) $q(m)$ at position $\vec{r}$ (a distance r from the origin) is
>
> $$\vec{F}_e = \left(\frac{1}{4\pi\epsilon_o}\right) \frac{Qq}{r^2}\,\hat{r} \qquad\qquad \vec{F}_g = \Big(-G\Big) \frac{Mm}{r^2}\,\hat{r} \tag{1}$$
>
> where $\hat{r} = \vec{r}/r$ is the unit vector in the radial direction."[3]

The quantities in the parentheses indicate the strengths of the forces, and the signs indicate that the Coulomb(Newton) force is repulsive(attractive). Gravity, of course, is always attractive. It appears that the electric force is repulsive, but since there exists the possibility of negative charge, electrostatics admits an attractive force between charges of opposite sign.

Superposition We can consider $q(m)$ as a "test" charge(mass). That is, it's an object that we use to determine how other objects affect it, or in other words, we are interested in the forces exerted on it by other objects with charge(mass). If there are many charges(masses) "near" the origin,[4] it is found experimentally that the net force on our test charge(mass) will simply be the sum of all the forces (see Fig. 1)

$$\vec{F}_e = \sum_i \left(\frac{1}{4\pi\epsilon_o}\right) \frac{Q_i q}{|\vec{r}-\vec{r}_i|^3}\,(\vec{r}-\vec{r}_i) \qquad\qquad \vec{F}_g = \sum_i \Big(-G\Big) \frac{M_i m}{|\vec{r}-\vec{r}_i|^3}\,(\vec{r}-\vec{r}_i) \tag{2}$$

where $\vec{r}_i$ is the vector pointing from the origin to the charge(mass) $Q_i(M_i)$, and the test charge(mass) is still at position $\vec{r}$. Equation (2) is all that is needed (along with Newton's

[3] The notation here is standard: the radius vector is $\vec{r} = x\hat{x} + y\hat{y} + z\hat{z} = (x, y, z)$, and its magnitude is $r = \sqrt{x^2 + y^2 + z^2}$

[4] The meaning of "near" will be made clear below.

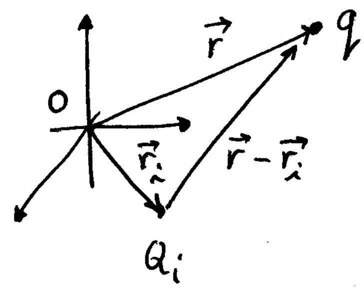

Figure 1: Geometry for the calculation of the force on a test charge q at position $\vec{r}$ due to the discrete superposition of infinitesimal charges Q_i.

Second Law of dynamics, of course) to determine the motion of our test charge(mass). Even though all the necessary physics is contained in Eq. (2), solving for the motion analytically may prove tedious or even impossible.

To help solve difficult problems, there are two abstract ideas that can be developed. The first is the concept of potential energy U, which you have seen in mechanics, and the second is the concept of a field. As Ernst Mach (yes, he's *the* Mach of Mach number fame) said, "... it is at once recognized as a highly convenient and *economical* course to investigate in the place of the forces themselves the function U."[5]

Potential Energy Recall that the gravitational force is *conservative*, and hence by virtue of its similarity, the electric force must be also. What does conservative mean? It means that the work done by the force on an object is independent of the path taken. Or, the work done around a closed path is zero,

$$\oint \vec{F} \cdot d\vec{\ell} = 0.$$

This allows us to define a POTENTIAL ENERGY for the charge(mass) $q(m)$ in the following way,

$$U_f - U_i \equiv -\int_i^f \vec{F} \cdot d\vec{\ell}.$$

Note that only *differences* in potential energy are meaningful, and that we are free to choose the location where the potential energy is zero. I can therefore express the potential energy as a function of position $\vec{r}$

$$U(\vec{r}) \equiv -\int_{\vec{r}_o}^{\vec{r}} \vec{F} \cdot d\vec{\ell}$$

where $U(\vec{r}_o) = 0$.[6] For the Coulomb and Newton forces, it is convenient (though not necessary) to set the potential energy equal to zero infinitely far away from the charges(masses)

[5] *The Science of Mechanics*, Ernst Mach, 1893.

[6] The notation here simply means that $U(\vec{r}) = U(x, y, z)$, i.e., U is a function of all three position coordinates.

$Q_i(M_i)$, where the electric(gravitational) force goes to zero, and our test charge(mass) feels no effect due to the objects that are near the origin.[7]

Since the integral is path independent, we can choose to integrate along a radial line so that $d\vec{\ell} = d\vec{r}$. For one charge(mass) that is fixed at the origin, U is a function of r only (not of the other spherical coordinates θ and ϕ), and we have

$$U_e(r) = -\int_\infty^r \left(\frac{1}{4\pi\epsilon_o}\right) \frac{Qq}{r'^{\,2}}\, dr' \qquad\qquad U_g(r) = -\int_\infty^r \left(-G\right) \frac{Mm}{r'^{\,2}}\, dr' \tag{3}$$

where r' is a dummy integration variable. After evaluating these simple integrals (make sure that you can do them), I get the well-known results

$$U_e(r) = \left(\frac{1}{4\pi\epsilon_o}\right) \frac{Qq}{r} \qquad\qquad U_g(r) = \left(-G\right) \frac{Mm}{r} \tag{4}$$

Of course, if there are several objects near the origin, the principle of superposition states that the total potential energy of our test charge(mass) is equal to the sum of the potential energies due to each object, and we have the logical result of Eq. (2)

$$U_e(\vec{r}) = \sum_i \left(\frac{1}{4\pi\epsilon_o}\right) \frac{Q_i q}{|\vec{r} - \vec{r}_i|} \qquad\qquad U_g(\vec{r}) = \sum_i \left(-G\right) \frac{M_i m}{|\vec{r} - \vec{r}_i|} \tag{5}$$

Note that now the potential energy is not a function only of r because the other objects are only *near* the origin, but not *at* it.

> NOTE: For a system of particles, we can really only define a potential energy for the *entire* system. There is no meaningful way to assign a portion of the potential energy to one object, as we have just done. However, since only *differences* in potential energy are meaningful, and since we are only interested in the case where the objects near the origin are stationary (and are the sources of the force on our test object), what I really mean by Eq. (5) is the difference in potential energy of the system when the test object is at position $\vec{r}$ compared to when it is infinitely far away from the rest of the objects.

2 Basic Field Theory

The concept embodied in Eq. (2) is that objects exert forces on other objects *without touching them!* This is sometimes called "action-at-a-distance," and it concerned physicists from Descartes to Newton to Faraday to Maxwell to Einstein. Newton conceived of space as

[7] Here, then, is an operational definition of what is meant by "near" the origin: All the charges (masses), except for the test charge(mass), are localized in one region of space, and we choose the origin to be near those objects. In other words, it must be possible to move our test charge(mass) as far away from these other objects as we desire. If the charges(masses) are not localized but are spread out infinitely far, as is often the case for simple examples in textbooks, the mathematics is slightly tricky, but the physical situation is also unrealistic so I won't consider such a case here.

absolute, and force as something that changes an object's momentum (his second law), but he was not able to explain *how* a force such as gravity could act across empty space. In fact, Newton thought of his gravitational theory as "mathematical," in that it gave the proper prediction of how the planets moved, but it did not *explain* anything in the sense that it supplied no "physical" interpretation of *how* the gravitational force acted as it did, nor *why* it should do so. In a letter, Newton wrote

> 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 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.[8]

Gravity was one thing, involving planets separated by very large distances, but magnets in the laboratory were another. Wasn't direct contact between objects necessary for cause and effect? How can an effect be felt at large distances instantaneously? This kind of thinking is partially what led Faraday to invent the concept of a FIELD.[9] In this conception, one object is the source of an electric(gravitational) field, and a second object does not experience a force due to the first object directly, as required by Eq. (2), but experiences a force due to the field. It is the field, therefore, that extends throughout space and exerts its influence on other objects.

Field Ultimately, since we are interested in the force exerted on our test charge(mass) at position $\vec{r}$, we need to know the electric(gravitational) field at that point. The field at position $\vec{r}$ due to the charges(masses) near the origin is defined as

$$\vec{E}(\vec{r}) \equiv \frac{\vec{F}_e}{q} \qquad\qquad \vec{g}(\vec{r}) \equiv \frac{\vec{F}_g}{m} \tag{6}$$

where $\vec{E}$ is the ELECTRIC FIELD, $\vec{g}$ is the GRAVITATIONAL FIELD, and $\vec{F}_e$ and $\vec{F}_g$ are given by Eq. (2). Notice that I have taken the force acting on the test charge(mass), which is proportional to $q(m)$, and divided that force by $q(m)$ in order to eliminate any mention of the test object. Thus the field at $\vec{r}$ depends solely on the objects that are localized near the origin.[10] This means that the field can be written

$$\vec{E}(\vec{r}) = \sum_i \left(\frac{1}{4\pi\epsilon_o}\right) \frac{Q_i}{|\vec{r}-\vec{r}_i|^3}\ (\vec{r}-\vec{r}_i) \qquad\qquad \vec{g}(\vec{r}) = \sum_i \Big(-G\Big) \frac{M_i}{|\vec{r}-\vec{r}_i|^3}\ (\vec{r}-\vec{r}_i) \tag{7}$$

[8] Letter to Richard Bentley, 1692, Ref. 1.

[9] This notion of space being permeated by a medium was not original with Faraday, but went back at least to Descartes, who thought that space without matter was unthinkable. Descartes felt that all space must be occupied by matter of one sort or another, regardless of whether we could perceive it or not. It was this medium that supported the electric and magnetic fields of Faraday, and it was not until 1905 that Einstein finally showed that this notion was unnecessary. Electric and magnetic fields *can* exist in the "vacuum" of space, and no "luminiferous ether" (as the medium was known then) was needed to support those fields.

[10] You can think of the field as the "force-per-unit-charge(mass)."

Now that we know the value of the field at the position $\vec{r}$, how do we calculate the force that would be exerted on a test charge(mass) if we placed it at $\vec{r}$? Simply invert the definition in Eq. (6)!

$$\vec{F}_e = q\vec{E}(\vec{r}) \qquad\qquad \vec{F}_g = m\vec{g}(\vec{r}) \tag{8}$$

Near the Earth's surface, the vertical component of the gravitational force vector in Eq. (8) should be familiar to you. It is just $F_g = -mg$, where $g \approx 9.8$ m/s^2. This value is due to a sum over all the particles that comprise the Earth — see Eq. (7). Some authors take Eq. (7) as the definition of the electric(gravitational) field, rather than Eq. (6), but in either case, the measurable quantity — force — is given by Eq. (8).

> NOTE: From a purely mathematical point of view, the definition of a field is straightforward: a field is simply a function of position. That is, at every point $(x,\ y,\ z)$ in space, a value is assigned. In the case of the electric(gravitational) field $\vec{E}$ ($\vec{g}$), we have a VECTOR FIELD. More rigorously, a vector field is a map $\vec{\mathcal{B}} : R^3 \mapsto R^3$ that assigns to each $\vec{r}$ a vector function $\vec{\mathcal{B}}(\vec{r})$. From a physical point of view, the function $\vec{\mathcal{B}}$ must represent some physical quantity.

With this concept of a field, it appears that we have made our job more complicated. We have taken a one-step process for determining the force on a test object — Eq. (2) — and replaced it with a two-step process — Eqs. (7) and (8) — where we first must take the intermediate step of calculating the field at the position of our test object. If we were only interested in the static case, where the fields are not functions of time, these two processes are mathematically equivalent, and it would make no sense to double our work in this manner. However, once we include electric fields that vary with time (and their accompanying magnetic fields), or once we investigate general relativity, the field point of view is required.

Potential The electric(gravitational) field is not the only field (in the mathematical sense) that can be defined. In tandem with the concept of potential energy, we can define a "potential energy per-unit-charge(mass)"

$$V(\vec{r}) \equiv \frac{U_e}{q} \qquad\qquad \Phi(\vec{r}) \equiv \frac{U_g}{m} \tag{9}$$

where V is the ELECTRIC POTENTIAL, Φ is the GRAVITATIONAL POTENTIAL, and U_e and U_g are given by Eq. (5). When it is clear which force I am talking about, I will refer to these as just the potential. Using Eq. (5), the potentials can be expressed as

$$V(\vec{r}) = \sum_i \left(\frac{1}{4\pi\epsilon_o}\right) \frac{Q_i}{|\vec{r} - \vec{r}_i|} \qquad\qquad \Phi(\vec{r}) = \sum_i \Big(-G\Big) \frac{M_i}{|\vec{r} - \vec{r}_i|} \tag{10}$$

Since the potential energy is a scalar, not a vector, the potential is what mathematicians call a SCALAR FIELD; that is, it's a scalar function of position.[11]

[11] For those of you familiar with quantum mechanics, the "wave function" ψ is a scalar field. It is a function of position $(x,\ y,\ z)$. In fact, in his first paper on the wave equation, Schrodinger called it the "field scalar," and only later did it become known as the wave function.

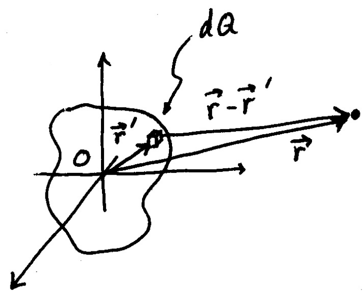

Figure 2: Geometry for the calculation of the field at position $\vec{r}$ due to the continuous superposition of infinitesimal charges dQ.

But be careful! The quantities V and Φ are *not* potential energies. They differ from potential energies by only a multiplicative factor, but conceptually they are distinct. So just as with the forces, if you want to know the potential energy of your test charge(mass), you must multiply the electric(gravitational) potential at that location by the test charge(mass)

$$U_e = qV(\vec{r}) \qquad\qquad U_g = m\Phi(\vec{r}) \tag{11}$$

Again, it appears that we have doubled our workload, but again, if you wish to describe electromagnetic waves (light) or black holes (general relativity), for example, the field and potential concepts are required.

3 Intermediate Field Theory

The previous section contains all the basic physics of static field theory, and it can be used in that form for calculations. However, even though matter is made up of discrete objects (protons, neutrons, and electrons), it can be more useful, and mathematically more powerful, to treat matter as a continuous distribution of charge(mass). Because the discrete particles of matter are so small, it appears from our perspective as human beings (i.e., collections of a large number of these particles) that the charge(mass) is smoothly distributed, with perhaps a density that varies with position. In this approximation, when calculating forces and potentials, the sums in Eqs. (7) and (10) must become integrals.

For example, if we are interested in the field at position $\vec{r}$, we must integrate over the all charge(mass) at position $\vec{r}\,'$ that is near the origin. The infinitesimal field $d\vec{E}(d\vec{g})$ at position $\vec{r}$ due to the infinitesimal charge(mass) $dQ(dM)$ at position $\vec{r}\,'$ is (see Fig. 2)

$$d\vec{E} = \left(\frac{1}{4\pi\epsilon_o}\right)\frac{(\vec{r}-\vec{r}\,')}{|\vec{r}-\vec{r}\,'|^3}\,dQ \qquad\qquad d\vec{g} = \Big(-G\Big)\frac{(\vec{r}-\vec{r}\,')}{|\vec{r}-\vec{r}\,'|^3}\,dM \tag{12}$$

where the charge(mass) at $\vec{r}\,'$ is $dQ = \rho_Q dV$ ($dM = \rho_M dV$). Here, $\rho_Q(\rho_M)$ is the charge(mass) density at $\vec{r}\,'$, and dV is the infinitesimal volume element. Note that $\vec{r}_i$ in Eq. (7) is replaced by $\vec{r}'$ in Eq. (12). Therefore, the total field is an integral over the infinitesimal field

$$\vec{E}(\vec{r}) = \left(\frac{1}{4\pi\epsilon_o}\right) \int \frac{(\vec{r}-\vec{r}')\rho_Q(\vec{r}')}{|\vec{r}-\vec{r}'|^3}\, d^3r' \qquad \vec{g}(\vec{r}) = \Big(-G\Big) \int \frac{(\vec{r}-\vec{r}')\rho_M(\vec{r}')}{|\vec{r}-\vec{r}'|^3}\, d^3r' \qquad (13)$$

where d^3r' is another way of writing the volume element $dV = d^3r' = dx'dy'dz'$ that explicitly states which coordinates are being integrated over. The integral in Eq. (13) is really a triple integral over the entire volume for which $\rho \neq 0$.[12] Similarly, the potential in Eq. (10) due to discrete objects becomes an integral over all the charge(mass) near the origin

$$V(\vec{r}) = \left(\frac{1}{4\pi\epsilon_o}\right) \int \frac{\rho_Q(\vec{r}')}{|\vec{r}-\vec{r}'|}\, d^3r' \qquad \Phi(\vec{r}) = \Big(-G\Big) \int \frac{\rho_M(\vec{r}')}{|\vec{r}-\vec{r}'|}\, d^3r' \qquad (14)$$

Equation (13) is a *vector* integral, which means that it really is *three* (triple) integrals, one for each component. On the other hand, Eq. (14) is only a *scalar* (triple) integral (i.e., just one), and is usually easier to evaluate. Why is this distinction important? Our goal (remember our goal?) is to calculate the force on a test object. The force is the real, physical quantity, and the one that we need to plug into Newton's Second Law of Dynamics if we want to determine the motion of our test object. So, given a distribution of charge(mass), continuous or discrete, we first need to calculate the field, and then the force. However, because vector integrations can sometimes be tedious, in those cases it is simpler to first calculate the potential (another intermediate step!) and then the field, and then the force.[13] Unfortunately, we are missing one step: what is the relationship between the field and the potential? Look at our definition of potential energy on page 3: it's the integral of the force. It must be, then, that the force is the derivative of the potential! This is true, but what kind of derivative? Since force is a vector and potential energy is a scalar, it must be a vector derivative — in fact, it's the gradient[14]

$$\vec{F} = -\nabla U,$$

where

$$\nabla \equiv \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}$$

is an operator that turns a scalar field into a vector field. In the simplest case of one object located at the origin, it is easy to show that this differential operator allows you to obtain Eq. (1) from Eq. (4). (Make sure that you can do this calculation![15]) Since the field and potential are related to the force and potential energy by a constant factor, it follows that

$$\vec{E} = -\nabla V \qquad \vec{g} = -\nabla\Phi \qquad (15)$$

We now have a three-step algorithm for calculating the force on a test charge(mass) that is located at $\vec{r}$ due to localized charges(masses) near the origin. Since derivatives are

[12]If I am discussing both fields in general, I will leave off the subscript on ρ.
[13]In the discrete case, it's simpler to evaluate Eq. (10) than Eq. (7).
[14]Daniel Bernoulli, in 1738, was the first to realize that a force can be derived from a potential function.
[15]That is, make sure that you can show that $\nabla(1/r) = -\hat{r}/r^2$.

straightforward to obtain, it is usually easier to evaluate a scalar integral and then take a derivative than to evaluate a vector integral.

1. Use Eq. (10) or Eq. (14) to calculate the potential;
2. Use Eq. (15) to calculate the field; and
3. Use Eq. (8) to calculate the force on your test charge(mass).

For complicated distributions of charge(mass), it is only step 1 that might be difficult.

4 Advanced Field Theory

You now have some fairly sophisticated mathematical tools, and I have developed expressions that enable you to calculate the potential and field due to any arbitrary charge(mass) distribution (if you can evaluate the integrals, that is). However, field theory is not just a set of tools to calculate forces, it's an aesthetically beautiful, logical structure, and we can gain a deeper understanding of how nature works by investigating how the pieces of the puzzle fit together. This deeper understanding, however, comes with a price: advanced vector calculus.

Consider the puzzle so far. We really have *three* mathematical fields (i.e., functions of position), two scalar and one vector. There is the charge(mass) distribution $\rho_Q(\rho_M)$; the potential $V(\Phi)$; and the field $\vec{E}(\vec{g})$. So far we have treated ρ as the *cause* (i.e., the source of the field and potential) and the field and potential as the *effects*, and this is the correct physical reasoning. However, it is also valid to think of them separately, and choose cause and effect depending on our knowledge. For example, if you know the electric(gravitational) field $\vec{E}(\vec{g})$ at all positions, can you calculate the charge(mass) distribution that is consistent with that field? How about the potential? Or more generally, given any one of the fields, can you calculate the other two? The answer is yes, but it requires the advanced vector calculus mentioned above.

Equations (13) and (14) allow us to calculate the field and potential directly from ρ, and Eq. (15) allows us to calculate the field directly from the potential. We wish to develop formulæ that will go the other way: to calculate ρ from either the field or the potential, and also to calculate the potential from the field.

Gauss's Law The place to start is with GAUSS'S LAW, which is a mathematically equivalent statement of the Coulomb(Newton) force law as embodied in Eqs. (7) or (13). It is usually written as

$$\oint_S \vec{E}\cdot d\vec{A} = \frac{1}{\epsilon_o}\int_V \rho_Q\, dV \qquad\qquad \oint_S \vec{g}\cdot d\vec{A} = -4\pi G\int_V \rho_M\, dV \tag{16}$$

which states that the average value of the component of a field normal to a closed surface S (called the *flux*) is proportional to the amount of charge(mass) that is in the volume V defined by the closed surface.[16] The integral on the right-hand-side, $\int \rho dV$ is just equal to the total charge(mass) in the volume V. This is consistent with the principle of superposition, and hence Eq. (7), because if you double the amount of charge(mass), you double the strength of the field.[17] However, Gauss's Law does not give us what we are looking for, because all field quantities are inside integrals. But we can turn this *integral form* of Gauss's Law into its differential form using the DIVERGENCE THEOREM, which states that for any vector field $\vec{\mathcal{B}}(\vec{r})$

$$\oint_S \vec{\mathcal{B}} \cdot d\vec{A} = \int_V \nabla \cdot \vec{\mathcal{B}}\, dV.$$

That is, the flux of $\mathcal{B}$ through a closed surface is equal to the divergence of $\mathcal{B}$ integrated throughout the volume enclosed by the surface. The proof of this theorem is difficult, but standard fare in advanced calculus courses, so I'll omit the proof here, but use the result. Using the divergence theorem, we can replace the left-hand-side of Eq. (16) to obtain

$$\int_V \nabla \cdot \vec{E}\, dV = \frac{1}{\epsilon_o} \int_V \rho_Q\, dV \qquad \qquad \int_V \nabla \cdot \vec{g}\, dV = -4\pi G \int_V \rho_M\, dV \tag{17}$$

It appears that we have not achieved anything, as all our desired quantities are still inside integrals, and I would like to make a statement about the integrands. But note that because the integration limits in Eq. (17) are arbitrary, it follows that the integrands are equal

$$\nabla \cdot \vec{E} = \frac{1}{\epsilon_o} \rho_Q \qquad \qquad \nabla \cdot \vec{g} = -4\pi G \rho_M \tag{18}$$

This is the *differential form* of Gauss's Law, and is one of the relations we are looking for — it allows us to calculate ρ directly from a knowledge of the field.

That last step, obtaining Eq. (18) from Eq. (17), where I asserted the equality of the integrands from the equality of the integrals, is an important tool in mathematics, so I'll take a moment now to look at it more closely. In general, if two integrals are equal, it does not mean that the integrands are also equal. For example, it is easy to show that

$$\int_0^3 2x\, dx = \int_0^3 x^2\, dx \qquad \text{and} \qquad \int_0^4 2x\, dx \neq \int_0^4 x^2\, dx.$$

Just because the ranges of integration are equal, as in Eq. (17), it does not imply equality of the integrands. But if the limits are *arbitrary*, then the integrands must be equal. This can be seen from the following one dimensional proof (the extension to three dimensions is straightforward).

[16] This is the mathematical proof of the assertion that the Earth, a spherically symmetric object, exerts a force on other objects that is the same as if all the Earth's mass were concentrated at its center. It is only true for an inverse square force.

[17] The proportionality constant, however, seems to be off by a factor of 4π. This factor comes back when you realize that if the surface integral is performed over a sphere, the angular integration (in spherical coordinates) is exactly 4π

$$\int_0^{2\pi} \int_0^{\pi} \sin\theta\, d\theta\, d\phi = 4\pi.$$

In other words, there are 4π steradians in a sphere. Make sure you understand this integration.

Theorem 1. *If $f(x)$ and $g(x)$ are both continuous on $[a,d]$, and if*

$$\int_b^c (f-g)dx = 0 \tag{1.1}$$

for all b and c in $[a,d]$; Then $f(x)=g(x)$ for all x in $[a,d]$.

Proof. The Mean Value Theorem for Integrals states that

$$\int_x^{x+h} F(t)dt = F(z)h$$

for some z such that $x < z < x+h$. In the limit $h \to 0^+$, we also have $z \to x^+$, and therefore

$$\lim_{h\to 0^+} \frac{1}{h}\int_x^{x+h} F(t)dt = F(x).$$

Since the integral in (1.1) holds for *all* b and c, let $c = b+h$, and in the limit $h \to 0^+$, (1.1) becomes

$$\lim_{h\to 0^+} \frac{1}{h}\int_b^{b+h} (f-g)dx = f(b)-g(b) = 0.$$

The last equality holds for all b in $[a,d]$. □

Poisson's and Laplace's Equations Another useful relation follows from eliminating the field from Eqs. (15) and (18) to give

$$\nabla^2 V = -\frac{1}{\epsilon_o}\rho_Q \qquad\qquad \nabla^2\Phi = 4\pi G\rho_M \tag{19}$$

This is called Poisson's Equation and allows us to calculate ρ directly from a knowledge of the potential. A special cases arises if you are interested in the potential in a region of space where $\rho = 0$. Equation (19) then becomes Laplace's Equation

$$\nabla^2 V = 0 \qquad\qquad \nabla^2\Phi = 0 \tag{20}$$

which is one of the most powerful equations in all physics. It is applicable not only to electrostatics and gravitation, but also to incompressible fluid flow ($\nabla^2 S = 0$, where S is the "velocity potential"), and to steady state heat flow ($\nabla^2 T = 0$, where T is the temperature). In general, the study of Eq. (20) is called "potential theory." In the 18th century, Euler, Lagrange, Legendre, and Laplace obtained many results related to the solution of Laplace's equation in different situations. In particular, they were able to deduce the shape of the Earth (an oblate spheroid) from a knowledge of Φ above the Earth's surface. Today, Laplace's equation is an indispensable tool for determining the gravitational force on satellites.

Path Independence We need just one more relation, one that allows a direct calculation of the potential from a knowledge of the field. It turns out we already have stated this relation on page 3 in the discussion of potential energy. There, I claimed that if the work done by a conservative force was path independent, it allowed for the definition of a potential energy function. Because these concepts of conservative force and potential energy are central to all of physics, I present here a formal proof.[18]

Theorem 2. *If $\vec{F}(\vec{r})$ is continuous on an open connected region, then the line integral*

$$\int_C \vec{F} \cdot d\vec{\ell}$$

is independent of path if and only if

$$\vec{F}(\vec{r}) = \nabla f(\vec{r})$$

for some function f.

Proof. 1. Let f be defined by

$$f(\vec{r}) \equiv \int_{\vec{r}_o}^{\vec{r}} \vec{F} \cdot d\vec{\ell}$$

where $\vec{r}_o = (x_0, y_0, z_0)$ is a fixed point and $\vec{r} = (x, y, z)$ is an arbitrary point. Suppose the integral is independent of path. Then f depends only on $\vec{r}$ and not on the path C from $\vec{r}_o$ to $\vec{r}$.
2. Break the path C into two pieces C_1 and C_2, where C_1 is an (arbitrary) path from (x_0, y_0, z_0) to (x_1, y, z), and C_2 is the straight line segment from (x_1, y, z) to (x, y, z), where x_1 is a fixed value. The function f can be written as the sum of two integrals

$$f = \int_{C_1} \vec{F} \cdot d\vec{\ell} + \int_{C_2} \vec{F} \cdot d\vec{\ell} = \int_{\vec{r}_o}^{(x_1,y,z)} \vec{F} \cdot d\vec{\ell} + \int_{(x_1,y,z)}^{(x,y,z)} \vec{F} \cdot d\vec{\ell}.$$

Since the first integral does not depend on x, taking a partial derivative results in

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} \int_{(x_1,y,z)}^{(x,y,z)} \vec{F} \cdot d\vec{\ell}.$$

3. The vector function $\vec{F}$ can be written

$$\vec{F}(\vec{r}) = M(x, y, z)\hat{x} + N(x, y, z)\hat{y} + P(x, y, z)\hat{z},$$

and so

$$\vec{F} \cdot d\vec{\ell} = M dx + N dy + P dz.$$

[18]See any calculus textbook, e.g., *Calculus*, by Earl W. Swokowski, Boston, 1979.

4. Using the fact that $dy = 0 = dz$ on the path C_2 we have

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\int_{(x_1,y,z)}^{(x,y,z)} M(x', y, z)dx'.$$

The Fundamental Theorem of Calculus states that integration and differentiation are inverse operations, which allows us to write the previous integral as

$$\frac{\partial f}{\partial x} = M(x, y, z).$$

5. A similar argument can be made for y and z to obtain

$$\frac{\partial f}{\partial y} = N(x, y, z) \qquad \text{and} \qquad \frac{\partial f}{\partial z} = P(x, y, z).$$

This proves that $\nabla f = \vec{F}$. $\square$

Of course, physicists use the convention that forces try to accelerate objects toward regions of low potential energy, and therefore the function f is defined with a minus sign. With this convention, the inverse of Eq. (15) is

$$V(\vec{r}) = -\int_{\vec{r}_o}^{\vec{r}} \vec{E}\cdot d\vec{\ell} \qquad\qquad \Phi(\vec{r}) = -\int_{\vec{r}_o}^{\vec{r}} \vec{g}\cdot d\vec{\ell} \tag{21}$$

which is our desired result. Of course, to make it practical we need to choose a reference point $\vec{r}_o$, which, again, is usually chosen to be infinitely far away from any charge or mass.

Helmholtz's Theorem As a final result, I will introduce a concept needed for the study of electromagnetism. Equation (18) can be thought of in two different ways. First, it's a direct way to calculate ρ from a knowledge of the field. On the other hand, it is a partial differential equation for the field if ρ is known. Of course, we already have Eq. (13) which directly gives us the field in terms of an integral over ρ, but often that integral is not trivial, and there are techniques that make solving Eq. (18) straightforward, especially in situations when ρ is symmetrically distributed. However, Eq. (18) does not determine the field uniquely. It is similar to the situation is one dimension, where even if the derivative of a function is known, the function itself is only known up to an arbitrary additive constant. In this case we wish to know the exact value of the field so we can calculate the force on a test object. For that we need not only the function's divergence, but its curl as well, because there is a theorem (which I won't prove) that states

> *A vector field is uniquely specified by giving its divergence and curl within a volume and its normal component over the boundary of that volume.*

What is the curl of a conservative field? Zero!

$$\nabla\times\vec{E} = 0 \qquad\qquad \nabla\times\vec{g} = 0 \tag{22}$$

Why is this? Because the field can be expressed as the gradient of the potential, Eq. (15), and a standard result of vector calculus states that the curl of a gradient is identically zero[19]

$$\nabla \times (\nabla f) = 0.$$

(Challenge: can you prove this?) This extra piece of information is what allows us to solve Eq. (18) for the field. It is an extra condition that determines the vector field uniquely.

Helmholtz's theorem, again one which I won't prove, is a more powerful theorem that is important when magnetic fields and time variation are included in the field theory. It states that

> *A vector field $\vec{\mathcal{B}}$ whose divergence and curl both vanish at infinity may be written as the sum of two parts, one of which is irrotational, and the other solenoidal.*[20]

An IRROTATIONAL vector field is one whose curl vanishes everywhere (so obviously the static electric field $\vec{E}$ and the Newtonian gravitational field $\vec{g}$ are both irrotational, by virtue of Eq. (22)), and a SOLENOIDAL vector field is one whose divergence vanishes everywhere. The electric(gravitational) field satisfies the conditions of Helmholtz's theorem because we require that ρ is localized near the origin, and Eq. (18) shows that the divergence therefore vanishes infinitely far away. Of course, Eq. (22) implies that the curl vanishes infinitely far away. This theorem means that our vector field $\vec{\mathcal{B}}$ can be written as the sum of a gradient and a curl

$$\vec{\mathcal{B}} = \nabla\varphi + \nabla \times \vec{\mathcal{C}}.$$

The first term is irrotational because of the result stated above, while the second term is solenoidal because of another standard result of vector calculus that states that the divergence of a curl is identically zero

$$\nabla \cdot \left(\nabla \times \vec{\mathcal{C}}\right) = 0.$$

(Challenge: can you prove this?) Therefore, when you study advanced electromagnetism, you will find that it is no longer possible to express the electric field $\vec{E}$ just as the gradient of the potential, but another term, the curl of a "vector potential," must be added, and Eq. (15) must be modifed

$$\vec{E} = -\nabla V + \nabla \times \vec{C}.$$

[19]Maxwell knew this result in the 1870s, when vector analysis was just beginning to be used.

[20]Arfken, *Mathematical Methods*, 2nd ed., Academic, 1970, page 67.

Summary I have developed several sophisticated results in this section on advanced field theory, and I think it might be useful to collect the important ones in one place. One of our main goals is to be able to calculate any of the three quantities (ρ, $\vec{E}/\vec{g}$, or V/Φ) directly from any of the others. The six equations that allow us to do that are Eqs. (13), (14), (15), (18), (19), and (21). They are listed again here.

$$\vec{E}(\vec{r}) = \left(\frac{1}{4\pi\epsilon_o}\right)\int \frac{(\vec{r}-\vec{r}\,')\rho_Q(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|^3}\, d^3r' \qquad \vec{g}(\vec{r}) = \Big(-G\Big)\int \frac{(\vec{r}-\vec{r}\,')\rho_M(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|^3}\, d^3r' \tag{13}$$

$$V(\vec{r}) = \left(\frac{1}{4\pi\epsilon_o}\right)\int \frac{\rho_Q(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\, d^3r' \qquad \Phi(\vec{r}) = \Big(-G\Big)\int \frac{\rho_M(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\, d^3r' \tag{14}$$

$$\vec{E} = -\nabla V \qquad \vec{g} = -\nabla\Phi \tag{15}$$

$$\nabla\cdot\vec{E} = \frac{1}{\epsilon_o}\rho_Q \qquad \nabla\cdot\vec{g} = -4\pi G\rho_M \tag{18}$$

$$\nabla^2 V = -\frac{1}{\epsilon_o}\rho_Q \qquad \nabla^2\Phi = 4\pi G\rho_M \tag{19}$$

$$V(\vec{r}) = -\int_{\vec{r}_o}^{\vec{r}} \vec{E}\cdot d\vec{\ell} \qquad \Phi(\vec{r}) = -\int_{\vec{r}_o}^{\vec{r}} \vec{g}\cdot d\vec{\ell} \tag{21}$$

A Bibliography

The concepts of action-at-a-distance and field have occupied philosophers of science almost as long as scientists have used these concepts. Consequently there is a vast literature that would take several lifetimes to study. However, there are two (relatively) short articles that cover most of the basic points, and are a suitable starting point for further investigation.

1. "The Origins of the Field Concept in Physics," by Ernan McMullin, *Physics in Perspective*, volume **4**, pages 13-39, 2002.

2. "Action at a Distance in Classical Physics," by Mary Hesse, *Isis*, volume **46**, pages 337-353, 1955.

Postscript 2008

Two other relevant references have come to my attention. First, there is an analysis of the famous "Feynman disk paradox," in which it is shown that even in static situations, the electromagnetic field carries angular momentum and therefore the field concept is crucial.

3. "Field versus action-at-a-distance in a static situation," by N. L. Sharma, *American Journal of Physics*, volume **56**, pages 420-423, 1988.

A second, more mathematically advanced, paper, develops a "generalized" Helmholtz theorem.

4. "A generalized Helmholtz theorem for time-varying vector fields," by Artice M. Davis, *American Journal of Physics*, volume **74**, pages 72-76, 2006.