Density of Energy Stored in the Electric Field

Department of Physics, Cornell University
© Tomás A. Arias

October 14, 2012

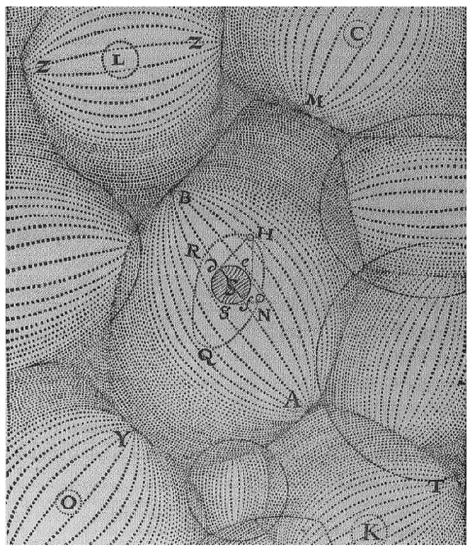

Figure 1: Diagram of "Cartesian vortices" from René Descartes' *Principia philosophiae*, published in 1644.

Contents

1 Introduction: Review of previous concepts

1.1 Electric field and action at a distance

As we have discussed since the very beginning of the course, many experiments have confirmed Coulomb's law for the force acting on charge q_2 due to the presence of a charge q_1 at a distance r,

$$\vec{F}_2 = k_C \frac{q_1 q_2}{r^2} \hat{r}. \tag{1}$$

This fact is that this law works very well when the charges are not moving. Nonetheless, a law like this presents a number of deep, fundamental questions for both Physics and Philosophy. Prime among these questions is the mystery of "action at a distance," the possibility of matter acting on other matter across reaches of empty space.

Early scientist-philosophers, including René Descartes and Gottfried Wilhelm Leibniz (who invented differential calculus independently of Sir Isaac Newton), when faced with the action-at-a-distance concept in the context of the gravitational forces keeping the earth and planets in their orbits, rejected it outright as a *metaphysical impossibility.* Indeed, modern physics also rejects this concept, but for a different reason — namely that it would imply the possibility of transmitting information at speeds faster than light, thus violating the Theory of Special Relativity. What this means is that Eq. (1) is really only an *approximation* that will have to be corrected when we study what happens as the charges begin to move.

To see how action at a distance would allow for faster-than-light communication, imagine that someone, "Person 2", has a charged object q_2 attached to a force detector and that person Person 1 has a charge q_1 under his control. Then, as soon as Person 1 moves q_1, the action at a distance inherent in Eq. (1) would imply that Person 2 could detect the slight change in the direction of the electric force *instantaneously.* Nothing says that this would not work even if Person 2 were one light-year away — so far away that it would take light one year to travel between Persons 1 and 2. If Eq. (1) really worked, the two people would be able to communicate faster than light, and the Theory of Special Relativity would be proven false! (See Figure 2.)

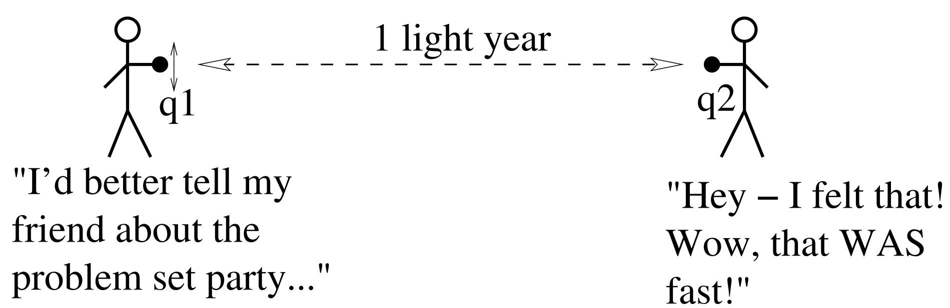

Figure 2: Possibility of faster-than-light communication enabled by action at a distance implied by Coulomb's law.

In earlier lectures, we resolved this difficulty in two steps that correspond exactly to how Descartes resolved the problem philosophically. First, we decided that the force on charge q_2 does not depend on some distant thing happening at the location of q_1. Rather, we decided that charge q_2 feels the effects of some local property of space at the location $\vec{r}_2$ where the charge q_2 sits. In Descartes's language, this corresponds to the vortex near the location of the Earth pushing on the Earth, rather than the Sun acting directly on the earth, as in Figure 1. In modern physics language, rather than calling this local property of space a "vortex", we call it the "Electric Field" $\vec{E}(\vec{r})$, and we write

$$\vec{F}_2 = q_2 \vec{E}(\vec{r}_2),$$

so that the force on q_2 really only depends on what is happening locally at the location $\vec{r}_2$. Comparing this equation to Coulomb's law, Eq. (1), this approach works so long as we define

$$\vec{E}(\vec{r}) = \frac{k_C q_1}{r^2}\hat{r}. \tag{2}$$

There, however, is still a problem, but now with Eq. (2). The problem is that we still have the question of how it is that charge q_1 can have an influence like this at points $\vec{r}$ which may be far away from q_1. The answer to this was our second key result, which corresponds to how the Cartesian vortices in Figure 1 affect only their neighbors, not distant points in space. In modern physics language, this means we want an equation for $\vec{E}(\vec{r})$ that only connects points that are right next to, *infinitesimally* close to, each other so that they are separated by *differential* distances dx, dy, and dz. Mathematically, this means we need some differential equation for $\vec{E}(\vec{r})$. In fact, we found just such a relationship in the differential form of Gauss's law,

$$\nabla \cdot \vec{E}(\vec{r}) = \rho(\vec{r})/\epsilon_0. \tag{3}$$

Indeed, it turns out that this equation, is an *exact* equation of physics that works in all circumstances, even if the charges are moving. The great advantage of avoiding action at a distance and always finding ways to express our physical laws in terms of differential equations is that we are much more likely to find equations and physical laws that always work!

1.2 Electric field and energy

The key concept we learned about connecting energy and electric fields is the *electrostatic*, or *electric*, potential $V(\vec{r})$, which is defined to be the potential energy *per unit charge* of a small positive test charge when placed at location $\vec{r}$. By definition, this has the value

$$\begin{aligned} V(\vec{r}) &\equiv \frac{1}{Q}U_Q(\vec{r}) \qquad (4)\\ &= \frac{1}{Q}\left(-\int_\infty^{\vec{r}} \vec{F}_Q(\vec{r}\,') \cdot d\vec{r}\,'\right)\\ &= -\int_\infty^{\vec{r}} \frac{\vec{F}_Q(\vec{r}\,')}{Q} \cdot d\vec{r}\,'\\ &= -\int_\infty^{\vec{r}} \vec{E}(\vec{r}\,') \cdot d\vec{r}\,', \end{aligned}$$

where in the last step we use the fact that the force per unit charge is precisely the definition of the electric field. (Note also that we use $d\vec{r}\,'$ to be more mathematically precise and avoid having our integration variable be the same as our limit of integration.) This equation for the electric potential has a similar action-at-a-distance problem. The potential at location $\vec{r}$ depends on what is happening at a who series of other points $\vec{r}\,'$, some of which are very far away because of the lower limit on the integral. This again could lead to fast-than-light communications: have Person 2 measure the potential $V(\vec{r})$ while Person 1 does things to change the electric field at the points near him, and Person 2 again could know *instantaneously* that something is happening.

A differential equation that we found for $V(\vec{r})$ solves this problem also, giving a description where $V(\vec{r})$ depends on what is happening at nearby points. Specially, the differential equation we found for $V(\vec{r})$ was

$$\vec{E}(\vec{r}) = -\nabla V(\vec{r}). \tag{5}$$

The final energy-related quantity we had was the total electric potential energy. To compute this, we divided the total charge distribution of the system into tiny little chunks of charge dq. We then used the electric potential to compute the potential energy of each of these charges by multiplying the potential energy per change $V(\vec{r})$ by each charge dq to find the energy associated with each chunk charge, $dU = V(\vec{r})\,dq$. Finally, we added up all of these contributions, taking into account the famous double-counting correction factor of 1/2, to find

$$U^{(tot)} = \frac{1}{2}\int V(\vec{r})\,dq, \tag{6}$$

where the formula for dq depends upon the type of distribution we are looking at,

$$dq = \begin{cases} \lambda\,d\ell & \text{linear distribution} \\ \sigma\,dA & \text{surface distribution} \\ \rho\,dV & \text{volume distribution} \end{cases}.$$

The main problem with Eq. (6) as a fundamental result is that it involves $V(\vec{r})$, and so it still may have vestiges of action at a distance. Also, we would like to think of the electric field as some sort of real, physical disturbance at each point in space (like one of Descartes' vortices), and not just a convenient mathematical fiction. If the electric field were a true disturbance in space, we would expect there to be an energy associated with it and, thus, that the electric energy should be related somehow directly to the electric field itself. As we shall see in the next section, the answer to both of these concerns is the same: it is possible to eliminate $V(\vec{r})$ from Eq. (6) and in the process write the stored energy $U^{(tot)}$ entirely in terms of contributions from the electric field at each point in space $\vec{r}$. The result is an equation which is not only correct in the Theory of Special Relativity, but which is also used without any further corrections in the modern Theory of Quantum Mechanics!

2 Electric potential energy in terms of the electric field

2.1 Basic strategy

To re-express Eq. (6) in terms of the electric field alone, we begin by considering the one form for dq which can describe *any* charge distribution. Specifically, this is the form for a volume charge distribution, $dq = \rho(\vec{r})\,dV$. This form ensures that we integrate over all of space and so include all charges in *any* system. Also, even if some of the charges are arranged into point, line or surfaces charges, these charges can be viewed as locations with very high volume densities of charges $\rho(\vec{r}) \to \infty$ concentrated in tiny regions near the corresponding points, lines or surfaces. Our starting point is thus

$$U^{(tot)} = \frac{1}{2}\iiint V(\vec{r})\,\rho(\vec{r})\,dV, \tag{7}$$

where the triple integral reminds us that we are doing a three-dimensional integral over all of space.

Now, keeping in mind our goal of expressing this in terms of the electric field, we note right away that, rearranging Eq. (3), we have $\rho(\vec{r}) = \epsilon_0 \nabla\cdot\vec{E}(\vec{r})$, so that Eq. (7) becomes

$$\begin{aligned} U^{(tot)} &= \frac{1}{2}\iiint V(\vec{r})\,\epsilon_0\nabla\cdot\vec{E}(\vec{r})\,dV \\ &= \frac{\epsilon_0}{2}\iiint V(\vec{r})\,\nabla\cdot\vec{E}(\vec{r})\,dV, \end{aligned} \tag{8}$$

where we have factored out the constant factor ϵ_0.

The result Eq. (8) has a very interesting mathematical structure. We would rather not have the derivative $\nabla\cdot$ acting on $\vec{E}$ — we would rather just have the electric field $\vec{E}$ to think about. Also, we really would like the derivative to be acting on $V(\vec{r})$, especially in the form $\nabla V(\vec{r})$, which we can see from Eq. (5), is $-\vec{E}(\vec{r})$. Thus, if we could just "move" the derivative from the factor $\vec{E}(\vec{r})$ over to the factor $V(\vec{r})$, we could have the total potential energy $U^{(tot)}$ entirely and directly in terms of the electric field, which is our goal!

This kind of maneuver of moving a derivative from one factor to another inside of an integral is done all the time for one-dimensional integrals in second semester calculus classes. This procedure is known as integration by parts. We now only have to learn how to generalize the procedure to multidimensional integrals, and we will have the result we want.

2.2 Integration by parts

The whole point of this subsection is to derive the formula for integration by parts in multiple dimensions, Eq. (12). If you are happy enough to accept that result, you can skip now directly to Section 2.3.

It turns out that integration by parts works in the same basic way for multidimensional integrals as it does for one-dimensional integrals. Recall that one-dimensional integration by parts begins with the product rule for derivatives,

$$\frac{d}{dx}\left(f(x)\,g(x)\right) = \frac{df}{dx}\,g(x) + f(x)\,\frac{dg}{dx},$$

and then integrating both sides,

$$\begin{aligned}\int_a^b \frac{d}{dx}\left(f(x)\,g(x)\right)\,dx &= \int_a^b \frac{df}{dx}\,g(x)\,dx + \int_a^b f(x)\,\frac{dg}{dx}\,dx \\ \left(f(x)\,g(x)\right|_a^b &= \int_a^b \frac{df}{dx}\,g(x)\,dx + \int_a^b f(x)\,\frac{dg}{dx}\,dx, \end{aligned} \tag{9}$$

where, in the second step, we use the fact that the integral of the derivative of any function returns back exactly that function. Finally, we rearrange the terms in Eq. (9) to get the final famous formula for integration by parts,

$$\int_a^b f(x)\,\frac{dg}{dx}\,dx = \left(f(x)\,g(x)\right|_a^b - \int_a^b \frac{df}{dx}\,g(x)\,dx. \tag{10}$$

To generalize the above to the case we have in mind (Eq. (8)), note that the derivation begins by looking at the derivative of the product of the two functions $f(x)$ and $g(x)$. Comparing Eqs. (10) and (8), we see that $V(\vec{r})$ plays the role of $f(x)$ and $\vec{E}(\vec{r})$ plays the role of $g(x)$. Thus, we should begin by considering the derivative of the product $V(\vec{r})\vec{E}(\vec{r})$. Because this product is a *vector* quantity, the natural type of derivative to consider is the divergence $\nabla\cdot$ — simply because you cannot take the gradient of a vector. Fortunately, it turns out that there is a simple product rule for the divergence of the product of a scalar quantity with a vector quantity,

$$\nabla\cdot\left(V\vec{E}\right) \equiv \left(\nabla V\cdot\vec{E}\right) + V\left(\nabla\cdot\vec{E}\right). \tag{11}$$

You can prove this by using the definition of the divergence, applying the ordinary product rule, and rearranging terms,[1]. (You can remember this result by noting it works just like the ordinary product rule, as long as you keep track of what kinds of derivatives and products you can take for different combinations of scalars and vectors.)

As with the one-dimensional case, the next step after identifying and taking the proper derivative is to integrate both sides of the resulting equation. Integrating Eq. (11) over all of space gives

$$\iiint \nabla\cdot\left(V\vec{E}\right)\,dV = \iiint \left(\nabla V\cdot\vec{E}\right)\,dV + \iiint V\left(\nabla\cdot\vec{E}\right)\,dV,$$

[1] Using the definitions of divergence and of the product of a scalar and a vector, we have

$$\begin{aligned}\nabla\cdot\left(V\vec{E}\right) &\equiv \frac{\partial}{\partial x}\left(VE_x\right) + \frac{\partial}{\partial y}\left(VE_y\right) + \frac{\partial}{\partial z}\left(VE_z\right) \\ &= \left(\frac{\partial V}{\partial x}E_x + V\frac{\partial E_x}{\partial x}\right) + (\text{same terms for y and z}) \\ &= \left(\frac{\partial V}{\partial x}E_x + (\text{same terms for y and z})\right) + \left(V\frac{\partial E_x}{\partial x} + (\text{same terms for y and z})\right) \\ &= \left(\nabla V\cdot\vec{E}\right) + V\left(\nabla\cdot\vec{E}\right),\end{aligned}$$

where in the last step, we make two key identifications. For the first term, we note that we have the sum of the product of each component of ∇V with the corresponding component of $\vec{E}$, thus giving the dot product of these two quantities. For the second term, we factor out the common factor of V, and what is left is the definition of $\nabla\cdot\vec{E}$.

which rearranges to what we want,

$$\iiint V\left(\nabla\cdot\vec{E}\right)\,dV=\iiint \nabla\cdot\left(V\vec{E}\right)\,dV-\iiint\left(\nabla V\cdot\vec{E}\right)\,dV, \tag{12}$$

which is the generalization of integration by parts to multidimensional integrals.

2.3 Fundamental result

We are now in a position to quickly derive our fundamental result. There are three terms in Eq. (12). The left-hand side of the equation, apart from a simple constant factor of $\epsilon_0/2$, is exactly the total energy that we want from Eq. (8). The last term on the right side of the equation will ultimately give us something very simple in terms of the electric field because $-\nabla V=\vec{E}$. This leaves only one term to consider in detail, the first term on the right side of the equation.

The first term on the right of Eq. (12) is a triple integral of a divergence and, thus, can be simplified by the divergence theorem,

$$\iiint_V \nabla\cdot\left(V\vec{E}\right)\,dV=\iint_{\partial V}\left(V\vec{E}\right)\cdot d\vec{A}, \tag{13}$$

where V represents the volume of integration and ∂V represents the surface bounding this volume. In the present case, this volume V represents all of space because we are trying to calculate the total potential energy of all of the charges in our system. The "boundary" of this volume ∂V, then represents any very large surface that contains the whole universe! Mathematically, we can represent this as the surface of a sphere of radius R where we take the limit $R\to\infty$. The point of choosing such a surface is that we know that, very far away from any charge distribution of total charge Q, the electric field (and thus also potential) from that charge distribution will look just like that coming from a point charge of charge Q. Thus, for the surface integral in Eq. (13) we can take $\vec{E}=(k_C Q/R^2)\hat{r}$ and $V=k_C Q/R$. Moreover, because the surface is spherical, we have $d\vec{A}=\hat{r}\,dA$ and our "far away" forms for $\vec{E}$ and V are constant over the surface of this sphere. Thus, when we include the surface area of the sphere $4\pi R^2$ we find that this entire second term in Eq. (12) becomes for us

$$\iint_{\partial V}\left(V\vec{E}\right)\cdot d\vec{A}=(k_C Q/R)\left(k_C Q/R^2\right)\left(4\pi R^2\right)=4\pi k_C^2 Q^2/R\to 0,$$

where in the last step we use the fact that $R\to 0$ because our surface is supposed to contain all of space. The conclusion of all of this analysis, then, is the simple fact that the first term on the right of Eq. (12) can be completely ignored!

Following this result, we drop the first term on the right side of Eq. (12). Then, multiplying the resulting equation by $\epsilon_0/2$ and using the fact that $-\nabla V(\vec{r})=\vec{E}$, we have our final result

$$\begin{aligned}
U^{(tot)} &= \frac{\epsilon_0}{2}\iiint V(\vec{r})\left(\nabla\cdot\vec{E}(\vec{r})\right)\,dV\\
&= -\frac{\epsilon_0}{2}\iiint\left(\nabla V(\vec{r})\cdot\vec{E}(\vec{r})\right)\,dV\\
&= \frac{\epsilon_0}{2}\iiint E(\vec{r})\cdot\vec{E}(\vec{r})\,dV\\
&= \frac{\epsilon_0}{2}\iiint |E(\vec{r})|^2\,dV\\
&\Rightarrow\\
U^{(tot)} &= \iiint \frac{\epsilon_0}{2}|E(\vec{r})|^2\,dV.
\end{aligned} \tag{14}$$

The way we interpret this final, beautifully simple result, is that potential energy associated with each volume of space dV that we are adding up to form our total energy $U^{(tot)}$ must be $\frac{\epsilon_0}{2}|E(\vec{r})|^2\,dV$, so that the *energy* *density*, the energy per unit volume of space associated with the electric field at each point in space $\vec{r}$ is

$$u(\vec{r}) \equiv \frac{\epsilon_0}{2}|E(\vec{r})|^2, \tag{15}$$

and we can write the total potential energy as

$$U^{(tot)} = \iiint u(\vec{r})\, dV. \tag{16}$$

3 Application and double-check

We can double-check our final result (15) by applying it to parallel plate capacitor, where we had the key results that $U^{(tot)} = (1/2)CV^2$ and $C = \epsilon_0 A/d$, where V is the voltage between the plates, A is the area of the plates and d the separation between them. To calculate the energy density, the energy per unit volume, we take the total energy $U^{(tot)}$ and divide by the volume of the capacitor, which is Ad. This gives

$$\begin{aligned} u &\equiv \frac{U^{(tot)}}{Ad} \\ &= \frac{\frac{1}{2}CV^2}{Ad} \\ &= \frac{\frac{1}{2}\frac{\epsilon_0 A}{d}V^2}{Ad} \\ &= \frac{\epsilon_0 AV^2}{2Ad^2} \\ &= \frac{\epsilon_0}{2}\left(\frac{V}{d}\right)^2. \end{aligned}$$

But, because we know that the magnitude of the electric field between the plates is constant and has value $E = V/d$, we see that the capacitor indeed exactly obeys Eq. (15)!

In most courses at this level, as you can see from the textbook, one simply does this calculation for the energy density in the capacitor and then "guesses" that the energy density formula Eq. (15) works for all cases. But, because we have learned the differential form for Gauss's law, Eq. (3), and the value avoiding equations which include action at a distance, we are actually able to prove that Eq. (15) works for all cases!!!