Physics 162: Answers to End-of-Class Questions
February 5, 2015

The almost continuous 2 hours of lecture can be a little too much. Even though the examples are helpful, if we could intersperse part lectures and working on problems into class, that might be better.

This is a good point and I appreciate your making it. I will try to avoid talking so much on Thursday's and give the class more time to work on problems, either homework problems or problems related to the homework. It is hard for me to judge when it is better for me to work through some examples or for you to work out problems in groups, but I do agree the balance has tilted too much the wrong way.

Can we review properties of insulators and conductors? When looking at problem 4 of Assignment 4, a classmate and I realized we knew how two conducting metals interact, but we didn't know if the rules apply to the scenario of problem 4, a non-conducting charged slab parallel to a conducting metal slab.

For problem 4 and for most of this semester, a non-conductor has the property that charges in or on the non-conductor can not move in response to external electric fields from other objects, while a conductor (usually made of a metal but there are non-metallic conductors such as salt water) has the property that charges are mobile (actually, just the electrons for metals) and so charges in or on a conductor can move about in response to the electric fields produced by other charged objects.

So in problem 4 of Assignment 4, the non-conducting slab with volume charge density ρ has the property that the charge does not move about and rho stays constant, even when other objects are brought near the slab like a polarizing metal slabs. In contrast, the metal slab parallel and to the right of the non-conducting slab has mobile electrons that respond to any electric field (e.g., that produced by the non-conducting slab) such that it polarizes (some parts of its surface become positively charged, some parts negatively charged), and such that the total electric field vector anywhere inside the conducting slab is zero.

Those are the practical details you need to know. At a deeper level, it is interesting to ask what are the rules that determine whether some substance is a metal or insulator. These rules are more complicated than saying something like "any material made of atoms that have a few free valence electrons (like Al, Ag, K, Na, Fe) will form a metal", it turns out that metals are deeply quantum mechanical substances and can not be understood in terms of simple chemistry (counting valence electrons) or classical physics. As one example, gaseous atomic hydrogen can be squeezed under pressure and at low temperatures until it forms a crystalline solid that initially is a non-conductor. But if you squeeze the hydrogen crystal hard enough, the crystal suddenly acts like a metal and electrons become highly mobile, the same substance has two phases. The phase transition from non-conducting to conducting phase with increased pressure is called a "Mott transition" (see the Wikipedia article by the same name) and cannot be understood without using quantum mechanics.

If you have a proton and an electron separated by 10^{-15} m, would they combine to form a neutron or any other neutral subatomic particle, or will they start as they were, that distance apart?

The answer to your question is known from experiment: an isolated neutron is unstable (radioactive) and decays with a half life of about 10.3 minutes into three particles: a proton, an electron, and an electron anti-neutrino:

$$n \to p + e + \bar{\nu}_e. \tag{1}$$

This naively implies that a neutron is a not a bound state of a proton and electron, and further that the inverse reaction $p + e \rightarrow n$ can not occur by itself, at least if the proton and electron do not have too much kinetic energy when they interact. Either a third particle like the neutrino (which is unlikely since it has an extremely weak interaction with matter) or the proton and electron having enough initial kinetic energy would be needed to form a neutron from a proton and electron.

I say "naively" because the decay of a neutron or the formation of a neutron from a proton and electron is a quantum mechanical process that can not be understood with a simple picture of classical balls coming together or coming apart. There is a simple argument using the Heisenberg uncertainty principle that an electron can not exist as a particle inside a neutron, otherwise its location and energy would be known to such accuracy that the uncertainty principle would be violated. So the electron and neutrino do not exist as separate entities in a neutron, they come into existence through the decay of the neutron, which makes no sense classically but is typical of quantum mechanical events.

An important astronomical event in which many protons combine with electrons into neutrons is when a large mass star (star with mass greater than about 5 times the mass of our Sun) runs out of fusion fuel in its core, and then collapses inwards because the heat from the core can no longer hold back the immense mass of the star. The resulting supernova can compress the hydrogen-electron plasma of the core so greatly that the endothermic reverse reaction takes place and a huge ball of neutrons is formed with a mass of about 2 Sun's in a region just a few kilometers in diameter. This is a so-called neutron star (badly named because it no longer generates heat via fusion so it is definitely not a star), and is one of the neater and stranger objects astronomers have discovered in space.

Why do we start from infinity when calculating sums or integrals for the electric potential V?

This is a rather subtle point, and is discussed briefly in the Knight book in regard to Eq. (28.10) on page 812 and in the sentence right after Eq. (28.13) on page 814.

The subtlety is that the potential energy itself does not have any useful physical meaning, it is only the difference $V(B) - V(A)$ of the potential energies between two different locations that has physical meaning; this corresponds to the energy per charge needed to move a charged particle from point A to point B in the presence of a given electric field. Since only differences in V show up in physical problems, one is free to add an arbitrary constant to $V(x, y, z)$, say writing it as $V + c$ since the constant will subtract out: $(V(B) + c) - (V(A) + c) = V(B) - V(A)$.

If you recall, in today's class (also discussed in the Knight book, see Section 28.2 on page 814), I derived the difference in electric potential for two points near a point charge and got the result

$$V(B) - V(A) = V(r_B) - V(r_A) = \frac{KQ}{r_B} - \frac{KQ}{R_A}, \tag{2}$$

and then said that this suggested that the electric potential must be the expression

$$V(r) = \frac{KQ}{r}. \tag{3}$$

But this was not the most general case, I should have said that Eq. (2) suggests that

$$V(r) = \frac{KQ}{r} + c, \tag{4}$$

for c some arbitrary constant, where this constant doesn't matter since it gets subtracted out whenever I calculate a difference $V(B) - V(A)$.

Since the constant is arbitrary, one might choose the value $c = 0$ which then simplifies the math (just a tiny bit but every bit helps). This choice also seems reasonably physically in that two charged particles that are far away from each other have a weak influence on one another and so one might expect their potential energy to be close to zero.

There is one qualification and it is unfortunate that Knight does not mention this clearly: you can not use the formula $V = KQ/d$ for the electric potential at a point d from a point charge Q to calculate the potential of objects that have infinite extent, e.g., the potential of an infinite straight charged line or of an infinite charged plane. The reason is that these objects extend all the way to infinity and so it is wrong to assume that the constant c is zero there. So for infinitely extended objects, you generally end up with a nonzero constant.

I'll try to give an example in class tomorrow (Friday). But as one quick example, let's say you want to calculate the potential $V(r)$ at some point P that is a distance d from an infinite straight line of linear charge density λ. One might be tempted to break the line segment into little infinitesimal pieces of length dx and infinitesimal charge $dq = \lambda dx$ and then calculate the potential at P with this kind of integral:

$$V(d) = \int_{-\infty}^{\infty} \frac{K(\lambda\, dx)}{\sqrt{x^2 + d^2}} = K\lambda \int_{-\infty}^{\infty} \frac{du}{\sqrt{u^2 + 1}}, \tag{5}$$

and this integral turns out to be divergent (unbounded) so one gets a nonsensical answer. (In more detail, the integral over a finite charged line segment of length L, so with bounds from $-L/2$ to $L/2$ would give a potential $K\lambda \sinh^{-1}(L/2) = K\lambda \ln\left(1 + (L/2)\sqrt{1 + (L/2)^2}\right)$ and this expression diverges in the limit $L \to \infty$. The reason why this expression diverges is that we are using the expression $K\,dq/\sqrt{d^2 + x^2}$ for the potential of the charge dq, and this expression assumes that the potential at infinity is zero which is not the case since an infinite line extends all the way out to infinity.

So the right way to calculate the potential $V(d)$ at the point P a distance d from an infinite line is first to use symmetry and Gauss's law to deduce the electric field everywhere in space:

$$\mathbf{E} = \frac{2K\lambda}{d}\hat{\mathbf{r}}, \tag{6}$$

and then choose some arbitrary reference point P_0 that is a finite distance r_0 from the line, and then calculate the potential difference with respect to P_0:

$$V(r) - V(r_0) = -\int_{r_0}^{r} \frac{2K\lambda}{r}\,dr = 2K\lambda \ln\left(\frac{r_0}{r}\right), \tag{7}$$

and you end up with an arbitrary constant r_0 that disappears when you calculate some potential difference $V(r_1) - V(r_2)$. Note that the $\ln(r_0/r)$ makes sense physically compared to $\ln(r/r_0)$ in that, if the wire has positive line density $\lambda > 0$, moving closer to the wire, which means making the radial distance r smaller, should increase the potential (raise the potential energy for a positive charge), and indeed $\ln(r_0/r)$ increases with decreasing r since the argument r_0/r increases with decreasing r and the ln function is monotonic increasing.

Does an electric field penetrate a non-conducting solid with a charge density, e.g., does the electric field of a nearby slab pass into the interior of the non-conducting slab?

Yes. The principle of superposition holds generally: if you have several charged objects that each produce a certain electric field by themselves, then the total electric field vector at any point in space, even if that point is inside some other object, that is due to all of the objects is the sum of the electric field vectors from each object. This implies that one object can not block the electric field of the other objects.

A simple example would be three parallel infinite charged non-conducting planes, say labeled 1, 2, and 3. The plane in the middle, plane 2, does not block the electric field of plane 1 from influencing plane 3 and vice versa. In the problem 1 of Quiz 2, in which a charged plane was attached to the side of a charged slab, the electric field of the charged plane fills all parts of the charged slab.

This is even true for the interior of conductors. If you have a charged object outside a solid metal conductor, the electric field of that charged object is non-zero everywhere, even at points inside the metal. However,

3

surface charges on the metal will slide around until, when a static equilibrium is attained, the additional electric field inside the conductor created by the surface charges on the metal, added to the electric field created by an external charged object, will yield a **total** electric field that is zero at every point inside the conductor. So it is not true that each electric field from each separate object is itself zero inside a conductor, it is only the sum of all the non-zero electric fields produced by all charges that add up to zero at each point inside the conductor. This point should be made clear by problem 7 of Assignment 3.

I know Gauss's law is only usefully applicable for certain cases, is this also true for the electric potential?

The situation is similar but more general for electric potential: only for rather simple geometric objects, often having a lot of symmetry (but less symmetry than the three cases for which Gauss's law lets you calculate $\mathbf{E}$), can one find concise useful mathematical formulas for the electric potential V along some axis or for all points in space.

For example, it is fairly easy to calculate the electric potential for general points in space produced by a finite line segment of length L and uniform linear charge density λ, and easy to calculate the potential on the axis of a ring, of a disk, of a cone, of a cylinder, and of a sphere (all uniformly charged). It is not so easy to calculate the electric potential for these objects when the charge density is non-uniform or for irregular geometries.

Fortunately, there are enough simple geometric shapes for which answers can be computed that one can develop some good intuition, and for more complex objects one would turn to a computer code.

Why does conservation of energy exist? Is it simply a fundamental portion of the laws of physics or is there a reason why energy is conserved?

Energy conservation is a consequence of a deep non-obvious symmetry of nature, which is that the mathematical rules that govern how physical systems evolve over time, such as Newton's laws of motion, Maxwell's equations of electricity and magnetism, Schrodinger's equation of quantum mechanics, and Einstein's theory of gravity, all have the property that they don't depend on when the physics is occurring, that is the time t does not show up in the equations explicitly as some parameter.

Thus Newton's equation of motion $ma = F$ for a harmonic oscillator

$$m\frac{d^2x}{dt^2} = -kx, \tag{8}$$

involves the variable t through the second derivative, but it does not involve t explicitly as some parameter, an example of which would be writing the right side as $-k(1 + bt)x$ where b is some constant. This means that if I shift the origin of my time coordinate by the substitution $t \to t + t_0$ in Eq. (8), the equation has the same mathematical form, as you should verify.

As one simple demonsration of why energy is conserved, consider the total mechanical energy of the harmonic oscillator

$$E_{\mathrm{mech}} = \frac{1}{2}mv^2 + \frac{1}{2}kx^2, \tag{9}$$

and differentiate this with respect to time. This gives

$$\frac{dE_{\mathrm{mech}}}{dt} = \frac{1}{2} \times 2mv\frac{dv}{dt} + \frac{1}{2} \times 2kx\frac{dx}{dt}. \tag{10}$$

But if you recognize that $dv/dt = d^2x/dt^2 = (-k/m)x$ from Newton's second law Eq. (2) and substitute this into Eq. (10), and recognize that $dx/dt = v$, you see that $dE_{\mathrm{mech}}/dt = 0$, i.e., the total mechanical energy is constant, which is what is meant by conservation of energy. This works only because there are no explicit time variables t appearing in Eq. (8).

A cute non-calculus few-minute explanation of energy conservation is given in this "1-minute physics video"

A more general proof would go something like this. First, one needs to know that a conservative force $\mathbf{F}$, which is the only kind of force for which a potential energy can be defined, can be written as a 'derivative' or gradient of a potential energy function ϕ like this: $\mathbf{F} = -\nabla\phi$ and then the total mechanical energy has the form

$$E_{\text{mech}} = \frac{1}{2}mv^2 + \phi(x,y,z) \tag{11}$$

and Newton's second law of motion can be written in the vector form:

$$m\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F} = -\nabla\phi. \tag{12}$$

Now differentiate Eq. (11) with respect to time and use the chain rule to get:

$$\frac{dE_{\text{mech}}}{dt} = m\,\mathbf{v}\bullet\frac{d\mathbf{v}}{dt} + \nabla\phi\bullet\frac{d\mathbf{x}}{dt}. \tag{13}$$

(You should be able to do this easily if you have had a course in multivariate calculus, otherwise this math may be unclear.) Next substitute $d\mathbf{v}/dt = d^2\mathbf{x}/dt^2 = -(1/m)\nabla\phi$ from Newton's second law and write $d\mathbf{x}/dt = \mathbf{v}$ and you see that the total mechanical energy is conserved, i.e., has zero time derivative. This all works because the potential $\phi(x,y,z)$ does not depend on time t explicitly, i.e., the physics is the same no matter when you carry out an experiment, the laws of nature are time invariant.

So time and energy are deeply related quantities, with energy being conserved because the equations of motion don't depend explicitly on time. Similarly, one can show that momentum $\mathbf{p}$ is conserved because the equations of motion don't depend explicitly on where an experiment is carried out (the equations are not changed by shifting the origin of the coordinates), and angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is conserved because the equations of motion don't depend on the orientation of an experiment (the equations are unchanged by an arbitrary rotation about some axis passing through the origin of the coordinate system).