2. The constant C is independent of k and m. So I can measure it for one spring–mass system and know it for all spring–mass systems, no matter the mass or spring constant. The constant is a universal constant.

The requirement that dimensions be valid has simplified the analysis of the spring–mass system. Without using dimensions, the problem would be to find (or measure) the three-variable function f that connects m, k, and x_0 to the period:

$$T = f(m, k, x_0).$$

Whereas using dimensions reveals that the problem is simpler: to find the function h such that

$$\frac{kT^2}{m} = h().$$

Here $h()$ means a function of no variables. Why no variables? Because the right side contains all the other quantities on which kT^2/m could depend. However, dimensional analysis says that the variables appear only through the combination kT^2/m, which is already on the left side. So no variables remain to be put on the right side; hence h is a function of zero variables. The only function of zero variables is a constant, so $kT^2/m = C$.

This pattern illustrates a famous quote from the statistician and physicist Harold Jeffreys [25, p. 82]:

> A good table of functions of one variable may require a page; that of a function of two variables a volume; that of a function of three variables a bookcase; and that of a function of four variables a library.

Use dimensions; avoid tables as big as a library!

Dimensionless groups are a kind of invariant: They are unchanged even when the system of units is changed. Like any invariant, a dimensionless group is an abstraction (**Chapter 3**). So, looking for dimensionless groups is recipe for developing new abstractions.

6.4 Hydrogen atom

Hydrogen is the simplest atom, and studying hydrogen is the simplest way to understand the **atomic theory**. Feynman has explained the importance of the atomic theory in his famous lectures on physics [, p. 1-2]:

> If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations

> of creatures, what statement would contain the most information in the fewest words? I believe it is the *atomic hypothesis* (or the atomic *fact*, or whatever you wish to call it) that *all things are made of atoms – little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another.* In that one sentence, you will see, there is an *enormous* amount of information about the world...

The atomic theory was first stated by Democritus. (Early Greek science and philosophy is discussed with wit, sympathy, and insight in Bertrand Russell's *History of Western Philosophy* [30].) Democritus could not say much about the properties of atoms. With modern knowledge of classical and quantum mechanics, and dimensional analysis, you can say more.

6.4.1 Dimensional analysis

The next example of dimensional reasoning is the hydrogen atom in order to answer two questions. The first question is how big is it. That size sets the size of more complex atoms and molecules. The second question is how much energy is needed to disassemble hydrogen. That energy sets the scale for the bond energies of more complex substances, and those energies determine macroscopic quantities like the stiffness of materials, the speed of sound, and the energy content of fat and sugar. All from hydrogen!

The first step in a dimensional analysis is to choose the relevant variables. A simple model of hydrogen is an electron orbiting a proton. The orbital force is provided by electrostatic attraction between the proton and electron. The magnitude of the force is

$$\frac{e^2}{4\pi\epsilon_0}\frac{1}{r^2},$$

where r is the distance between the proton and electron. The list of variables should include enough variables to generate this expression for the force. It could include q, ϵ_0, and r separately. But that approach is needlessly complex: The charge q is relevant only because it produces a force. So the charge appears only in the combined quantity $e^2/4\pi\epsilon_0$. A similar argument applies to ϵ_0.

Var	Dim	What	ω
T^{-1}	frequency	k	
L^{-1}	wavenumber	g	
LT^{-2}	gravity	h	
L	depth	ρ	
ML^{-3}	density	γ	
MT^{-2}	surfacetension		

Therefore rather than listing q and ϵ_0 separately, list only $e^2/4\pi\epsilon_0$. And rather than listing r, list a_0, the common notation for the Bohr radius (the radius of ideal hydrogen). The acceleration of the electron depends on the electrostatic force, which can be constructed from $e^2/4\pi\epsilon_0$ and a_0, and on its mass m_e. So the list should also include m_e. To find the dimensions of $e^2/4\pi\epsilon_0$, use the formula for force

$$F = \frac{e^2}{4\pi\epsilon_0}\frac{1}{r^2}.$$

Then

$$\left[\frac{e^2}{4\pi\epsilon_0}\right] = \left[r^2\right] \times [F] = ML^3T^{-2}.$$

The next step is to make dimensionless groups. However, no combination of these three items is dimensionless. To see why, look at the time dimension because it appears in only one quantity, $e^2/4\pi\epsilon_0$. So that quantity cannot occur in a dimensionless group: If it did, there would be no way to get rid of the time dimensions. From the two remaining quantities, a_0 and m_e, no dimensionless group is possible.

The failure to make a dimensionless group means that hydrogen does not exist in the simple model as we have formulated it. I neglected important physics. There are two possibilities for what physics to add.

One possibility is to add relativity, encapsulated in the speed of light c. So we would add c to the list of variables. That choice produces a dimensionless group, and therefore produces a size. However, the size is not the size of hydrogen. It turns out to be the classical electron radius instead. Fortunately, you do not have to know what the classical electron radius is in order to understand why the resulting size is not the size of hydrogen. Adding relativity to the physics – or adding c to the list – allows radiation. So the orbiting, accelerating electron would radiate. As radiation carries energy away from the electron, it spirals into the proton, meaning that in this world hydrogen does not exist, nor do other atoms.

The other possibility is to add quantum mechanics, which was developed to solve fundamental problems like the existence of matter. The physics of quantum mechanics is complicated, but its effect on dimensional analyses is simple: It contributes a new constant of nature $\hbar$ whose dimensions are those of angular momentum. Angular momentum is mvr, so

$$[\hbar] = \mathrm{ML}^2\mathrm{T}^{-1}.$$

Var	Dim	What
a_0	L	size
$e^2/4\pi\epsilon_0$	$\mathrm{ML}^3\mathrm{T}^{-2}$	
m_e	M	electron mass
$\hbar$	$\mathrm{ML}^2\mathrm{T}^{-1}$	quantum

The $\hbar$ might save the day. There are now two quantities containing time dimensions. Since $e^2/4\pi\epsilon_0$ has T^{-2} and $\hbar$ has T^{-1}, the ratio $\hbar^2/(e^2/4\pi\epsilon_0)$ contains no time dimensions. Since

$$\left[\frac{\hbar^2}{e^2/4\pi\epsilon_0}\right] = \mathrm{ML},$$

a dimensionless group is

$$\frac{\hbar^2}{a_0 m_e (e^2/4\pi\epsilon_0)}$$

It turns out that all dimensionless groups can be formed from this group. So, as in the spring–mass example, the only possible true statement involving this group is

$$\frac{\hbar^2}{a_0 m_e (e^2/4\pi\epsilon_0)} = \text{dimensionless constant.}$$

Therefore, the size of hydrogen is

$$a_0 \sim \frac{\hbar^2}{m_e (e^2/4\pi\epsilon_0)}.$$

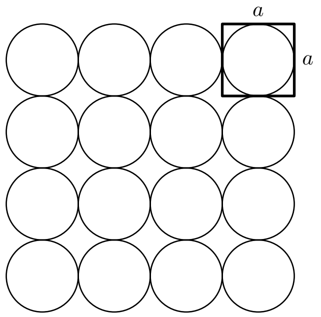
Putting in values for the constants gives

$$a_0 \sim 0.5\text{Å} = 0.5 \cdot 10^{-10}\,\mathrm{m}.$$

It turns out that the missing dimensionless constant is 1, so the dimensional analysis has given the exact answer.

6.4.2 Atomic sizes and substance densities

Hydrogen has a diameter of 1Å. A useful consequence is the rule of thumb is that a typical interatomic spacing is 3Å. This approximation gives a reasonable approximation for the densities of substances, as this section explains.



Let A be the atomic mass of the atom; it is (roughly) the number of protons and neutrons in the nucleus. Although A is called a mass, it is dimensionless. Each atom occupies a cube of side length $a \sim 3\,\text{Å}$, and has mass Am_{proton}. The density of the substance is

$$\rho = \frac{\text{mass}}{\text{volume}} \sim \frac{Am_{\text{proton}}}{(3\,\text{Å})^3}.$$

You do not need to remember or look up m_{proton} if you multiply this fraction by unity in the form of N_A/N_A, where N_A is Avogadro's number:

$$\rho \sim \frac{Am_{\text{proton}}N_A}{(3\,\text{Å})^3 \times N_A}.$$

The numerator is A g, because that is how N_A is defined. The denominator is

$$3 \cdot 10^{-23}\,\text{cm}^3 \times 6 \cdot 10^{23} = 18.$$

So instead of remembering m_{proton}, you need to remember N_A. However, N_A is more familiar than m_{proton} because N_A arises in chemistry and physics. Using N_A also emphasizes the connection between microscopic and macroscopic values. Carrying out the calculations:

$$\rho \sim \frac{A}{18}\,\text{g}\,\text{cm}^{-3}.$$

Element	$\rho_{\text{estimated}}$	ρ_{actual}
Li	0.39	0.54
H_2O	1.0	1.0
Si	1.56	2.4
Fe	3.11	7.9
Hg	11.2	13.5
Au	10.9	19.3
U	13.3	18.7

The table compares the estimate against reality. Most everyday elements have atomic masses between 15 and 150, so the density estimate explains why most densities lie between 1 and $10\,\text{g}\,\text{cm}^{-3}$. It also shows why, for materials physics, cgs units are more convenient than SI units are. A typical cgs density of a solid is $3\,\text{g}\,\text{cm}^{-3}$, and 3 is a modest number and easy to remember and work with. However, a typical SI density of a solid $3000\,\text{kg}\,\text{m}^{-3}$. Numbers such as 3000 are unwieldy. Each time you use it, you have to think, 'How many powers of ten were there again?' So the table tabulates densities using the cgs units of $\text{g}\,\text{cm}^{-3}$. I even threw a joker into the pack – water is not an element! – but the density estimate is amazingly accurate.

6.4.3 Physical interpretation

The previous method, dimensional analysis, is mostly mathematical. As a second computation of a_0, we show you a method that is mostly physics. Besides checking the Bohr radius, it provides a physical interpretation of it. The Bohr radius is the radius of the orbit with the lowest energy (the ground state). The energy is a sum of kinetic and potential energy. This division suggests, again, a divide-and-conquer approach: first the kinetic energy, then the potential energy.

What is the origin of the kinetic energy? The electron does not orbit in any classical sense. If it orbited, it would, as an accelerating charge, radiate energy and spiral into the nucleus. According to quantum mechanics, however, the proton confines the electron to a region of size r – still unknown to us – and the electron exists in a so-called stationary state. The nature of a stationary state is mysterious; no one understands quantum mechanics, so no one understands stationary states except mathematically. However, in an approximate estimate you can ignore details such as the meaning of a stationary state. The necessary information here is that the electron is, as the name of the state suggests, stationary: It does not radiate. The problem then is to find the size of the region to which the electron is confined. In reality the electron is smeared over the whole universe; however, a significant amount of it lives within a typical radius. This typical radius we estimate and call a_0.

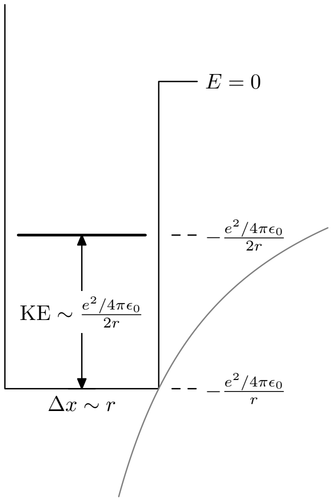
For now let this radius be an unknown r and study how the kinetic energy depends on r. Confinement gives energy to the electron according to the **uncertainty principle**:

$$\Delta x \Delta p \sim \hbar,$$

where Δx is the position uncertainty and Δp is the momentum uncertainty of the electron. In this model $\Delta x \sim r$, as shown in the figure, so $\Delta p \sim \hbar/r$. The kinetic energy of the electron is

$$E_{\text{Kinetic}} \sim \frac{(\Delta p)^2}{m_e} \sim \frac{\hbar^2}{m_e r^2}.$$

This energy is the **confinement energy** or the **uncertainty energy**. This idea recurs in the book.



The potential energy is the classical expression

$$E_{\text{Potential}} \sim -\frac{e^2}{4\pi\epsilon_0 r}.$$

The total energy is the combination

$$E = E_{\text{Potential}} + E_{\text{Kinetic}} \sim -\frac{e^2}{4\pi\epsilon_0 r} + \frac{\hbar^2}{m_e r^2}.$$

The two energies compete. At small r, kinetic energy wins, because of the $1/r^2$; at large r, potential energy wins, because it goes to zero less rapidly. Is there a minimum combined energy at some intermediate value of r? There has to be. At small r, the slope dE/dr is negative. At large r, it is positive. At an intermediate r, the slope crosses between positive and negative. The energy is a a minimum there. The location would be easy to estimate if the energy were written in dimensionless form. Such a rewriting is not mandatory in this example, but it is helpful in complicated examples and is worth learning in this example.

In constructing the dimensionless group containing a_0, we constructed another length:

$$l = \frac{\hbar^2}{m_e(e^2/4\pi\epsilon_0)}.$$

To scale any length – to make it dimensionless – divide it by l. So in the total energy the scaled radius

$$\bar{r} \equiv \frac{r}{l}.$$

The other unknown in the total energy is the energy itself. To make it dimensionless, a reasonable energy scale to use is $e^2/4\pi\epsilon_0 l$ by defining scaled energy as

$$\bar{E} \equiv \frac{E}{e^2/4\pi\epsilon_0 l}.$$

Using the dimensionless length and energy, the total energy

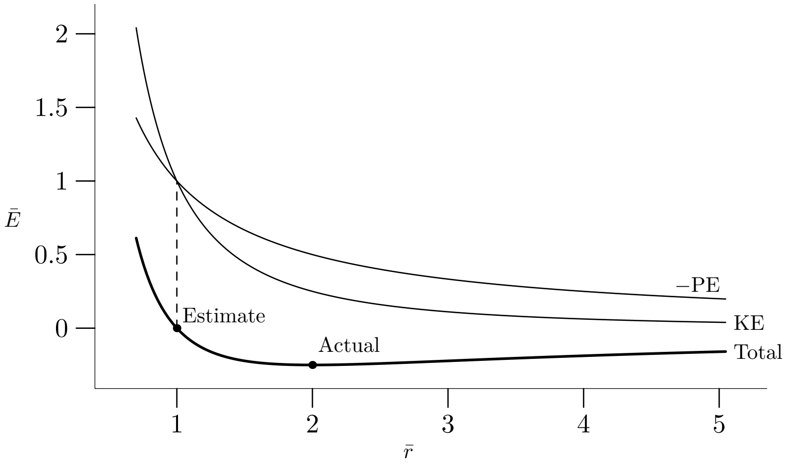
$$E = E_{\text{Potential}} + E_{\text{Kinetic}} \sim -\frac{e^2}{4\pi\epsilon_0 r} + \frac{\hbar^2}{m_e r^2}$$

becomes

$$\bar{E} \sim -\frac{1}{\bar{r}} + \frac{1}{\bar{r}^2}.$$

The ugly constants are placed into the definitions of scaled length and energy. This dimensionless energy is easy to think about and to sketch.

Simple calculus: minimizing scaled energy $\bar{E}$ versus scaled bond length $\bar{r}$. The scaled energy is the sum of potential and kinetic energy. The shape of this energy illustrates Feynman's explanation of the atomic hypothesis. At a 'little distance apart' – for large $\bar{r}$ – the curve slopes upward; to lower their energy, the proton and electron prefer to move closer, and the resulting force is attractive. 'Upon being squeezed into one another' – for small $\bar{r}$ – the potential rapidly increases, so the force between the particles is repulsive. Somewhere between the small and large regions of $\bar{r}$, the force is zero.



Calculus (differentiation) locates this minimum-energy $\bar{r}$ at $\bar{r}_{min} = 2$. An alternative method is **cheap minimization**: When two terms compete, the minimum occurs when the terms are roughly equal. This method of minimization is familiar from **Section 4.5.2**.

Equating the two terms $\bar{r}^{-1}$ and $\bar{r}^{-2}$ gives $\bar{r}_{min} \sim 1$. This result gives a scaled length. In actual units, it is

$$r_{min} = l\bar{r}_{min} = \frac{\hbar^2}{m_e(e^2/4\pi\epsilon_0)},$$

which is the Bohr radius computed using dimensional analysis. The sloppiness in estimating the kinetic and potential energies has canceled the error introduced by cheap minimization!

Here is how to justify cheap minimization. Consider a reasonable general form for E:

$$E(r) = \frac{A}{r^n} - \frac{B}{r^m}.$$

This form captures the important feature of the combined energy

$$E = E_{\text{Potential}} + E_{\text{Kinetic}} \sim -\frac{e^2}{4\pi\epsilon_0 r} + \frac{\hbar^2}{m_e r^2},$$

that two terms represent competing physical effects. Mathematically, that physical fact is shown by the opposite signs.

To find the minimum, solve $E'(r_{\text{min}}) = 0$ or

$$-n\frac{A}{r_{\text{min}}^{n+1}} + m\frac{B}{r_{\text{min}}^{m+1}} = 0.$$

The solution is

$$\frac{A}{r_{\text{min}}^n} = \frac{n}{m}\frac{B}{r_{\text{min}}^m} \qquad \text{(calculus)}.$$

This method minimizes the combined energy by equating the two terms A/r^n and B/r^m:

$$\frac{A}{r_{\text{min}}^n} = \frac{B}{r_{\text{min}}^m}.$$

This approximation lacks the n/m factor in the exact result. The ratio of the two estimates for r_{min} is

$$\frac{\text{approximate estimate}}{\text{calculus estimate}} \sim \left(\frac{n}{m}\right)^{1/(m-n)},$$

which is smaller than 1 unless $n = m$, when there is no maximum or minimum. So the approximate method underestimates the location of minima and maxima.

To judge the method in practice, apply it to a typical example: the potential between nonpolar atoms or molecules, such as between helium, xenon, or methane. This potential is well approximated by the so-called Lennard–Jones potential where $m = 6$ and $n = 12$:

$$U(r) \sim \frac{A}{r^{12}} - \frac{B}{r^6}.$$

The approximate result underestimates r_{min} by a factor of

$$\left(\frac{12}{6}\right)^{1/6} \sim 1.15.$$

An error of 15 percent is often small compared to the other inaccuracies in an approximate computation, so this method of approximate minimization is a valuable time-saver.

Now return to the original problem: determining the Bohr radius. The approximate minimization predicts the correct value. Even if the method were not so charmed, there is no point in doing a proper, calculus minimization. The calculus method is *too accurate* given the inaccuracies in the rest of the derivation.

Engineers understand this idea of not over-engineering a system. If a bicycle most often breaks at welds in the frame, there is little point replacing the metal between the welds with expensive, high-strength aerospace materials. The new materials might last 100 years instead of 50 years, but such a replacement would be overengineering. To improve a bicycle, put effort into improving or doing without the welds.

In estimating the Bohr radius, the kinetic-energy estimate uses a crude form of the uncertainty principle, $\Delta p \Delta x \sim \hbar$, whereas the true statement is that $\Delta p \Delta x \geqslant \hbar/2$. The estimate also uses the approximation $E_{\mathrm{Kinetic}} \sim (\Delta p)^2/m$. This approximation contains m instead of $2m$ in the denominator. It also assumes that Δp can be converted into an energy as though it were a true momentum rather than merely a crude estimate for the root-mean-square momentum. The potential- and kinetic-energy estimates use a crude definition of position uncertainty Δx: that $\Delta x \sim r$. After making so many approximations, it is pointless to minimize the result using the elephant gun of differential calculus. The approximate method is as accurate as, or perhaps more accurate than the approximations in the energy.

This method of equating competing terms is **balancing**. We balanced the kinetic energy against the potential energy by assuming that they are roughly the same size. The consequence is that

$$a_0 \sim \frac{\hbar^2}{m_e (e^2/4\pi\epsilon_0)}.$$

Nature could have been unkind: The potential and kinetic energies could have differed by a factor of 10 or 100. But Nature is kind: The two energies are roughly equal, except for a constant that is nearly 1. ‘Nearly 1’ is also called **of order unity**. This rough equality occurs in many examples, and you often get a reasonable answer by pretending that two energies (or two quantities with the same units) are equal. When the quantities are potential and kinetic energy, as they often are, you get extra safety: The so-called virial theorem protects you against large errors (for more on the virial theorem, see any intermediate textbook on classical dynamics).

6.5 Bending of light by gravity