Problem Solving I: Mathematical Techniques

For the basics of dimensional analysis and limiting cases, see chapter 1 of Morin or chapter 2 of Order of Magnitude Physics. Many more examples are featured in The Art of Insight; some particularly relevant sections are 2.1, 5.5, 6.3, 8.2, and 8.3. Other sections will be mentioned throughout the course. There is a total of **82** points.

1 Dimensional Analysis

Idea 1

Dimensional analysis is simply the statement that the dimensions of physical equations should match on both sides. This simple idea can sometimes solve whole problems by itself.

Dimensional analysis is also a valuable consistency check. For example, if you're trying to derive the surface area of a sphere and find $4\pi r^3$, you can instantly know you made a mistake. As another example, if a problem says the speed of an object is "small", this technically doesn't obey dimensional analysis unless we compare it to another speed. Thus, the problem might really mean you should assume the speed is small compared to the speed of light, $v \ll c$, which tells you something important.

To be precise, we should distinguish dimensions and units. The dimensions of a physical quantity determine what kind of quantity it is, while a unit is a measure of a dimension. Thus, for example, somebody's height h can be measured in units of feet or meters, but both have dimensions of length; this can be written as $[h] = [\mathrm{ft}] = [\mathrm{m}] = L$, where the brackets indicate dimensions. Another example is that angles are dimensionless, but can be measured in units of degrees or radians. These distinctions are not that important for our purposes, so we will be sloppy and conflate dimensions with units, writing the equivalent of $[h] = \mathrm{m}$.

Example 1: $F = ma$ 2018 B11

A circle of rope is spinning in outer space with an angular velocity ω_0. Transverse waves on the rope have speed v_0, as measured in a rotating reference frame where the rope is at rest. If the angular velocity of the rope is doubled, what is the new speed of transverse waves?

Solution

To solve this problem by dimensional analysis, we reason about what could possibly affect the speed of transverse waves. The result could definitely depend on the rope's length L, mass per length λ, and angular velocity ω_0. It could also depend on the tension, but since this tension balances the centrifugal force, it is determined by all of the other quantities. Thus the quantities we have are

$$[L] = \mathrm{m}, \quad [\lambda] = \mathrm{kg/m}, \quad [\omega_0] = 1/\mathrm{s}.$$

Since λ is the only thing with dimensions of mass, it can't affect the speed, because there is nothing that could cancel out the mass dimension. So the only possible answer is

$$v_0 \sim L\omega_0$$

where the $\sim$ indicates equality up to a dimensionless constant, which cannot be found by dimensional analysis alone. In practice, the constant usually won't be too big or too small, so $L\omega_0$ is a decent estimate of v_0. But even if it isn't, the dimensional analysis tells us the scaling: if ω_0 is doubled, the new speed is $2v_0$.

Example 2

Find the dimensions of the magnetic field.

Solution

To do this, we just think of some simple equation involving B, then solve for its dimensions. For example, we know that $\mathbf{F} = q(\mathbf{v} \times \mathbf{B})$, so

$$[B] = \frac{[F]}{[q][v]} = \frac{\text{kg} \cdot \text{m}}{\text{s}^2} \frac{1}{\text{C}} \frac{1}{\text{m/s}} = \frac{\text{kg}}{\text{C} \cdot \text{s}}.$$

[2] **Problem 1.** Find the dimensions of power, the gravitational constant G, the permittivity of free space ϵ_0, and the ideal gas constant R.

[1] **Problem 2.** Derive Kepler's third law for circular orbits, using only dimensional analysis. (Why do you think people didn't figure out this argument 2000 years ago?)

[2] **Problem 3.** Some questions about vibrations.

(a) The typical frequency f of a vibrating star depends only on its radius R, density ρ, and the gravitational constant G. Use dimensional analysis to find an expression for f, up to a dimensionless constant. Then estimate f for the Sun, looking up any numbers you need.

(b) The typical frequency f of a small water droplet freely vibrating in zero gravity could depends on its radius R, density ρ, surface tension γ, and the gravitational constant G. Argue that at least one of these parameters doesn't matter, and find an expression for f up to a dimensionless constant.

[3] **Problem 4.** Some questions about the speed of waves. For all estimates, you can look up any numbers you need.

(a) The speed of sound in an ideal gas depends on its pressure p and density ρ. Explain why we don't have to use the temperature T or ideal gas constant R in the dimensional analysis, and then estimate the speed of sound in air.

(b) The speed of sound in a fluid depends only on its density ρ and bulk modulus $B = -V\,dP/dV$. Estimate the speed of sound in water, which has $B = 2.1\,\text{GPa}$.

The speed of waves on top of the surface of water can depend on the water depth h, the wavelength λ, the density ρ, the surface tension γ, and the gravitational acceleration g.

(c) Find the speed of capillary waves, i.e. water waves of very short wavelength, up to a dimensionless constant.

(d) Find the speed of long-wavelength waves in very deep water, up to a dimensionless constant.

[3] **Problem 5** (Morin 1.5)**.** A particle with mass m and initial speed v is subject to a velocity-dependent damping force of the form bv^n.

(a) For $n = 0, 1, 2, \ldots$, determine how the stopping time and stopping distance depend on m, v, and b.

(b) Check that these results actually make sense as m, v, and b are changed, for a few values of n. You should find something puzzling going on. (Hint: to resolve the problem, it may be useful to find the stopping time explicitly in a few examples.)

Idea 2

Dimensional analysis applies everywhere. The argument of any function that is not a monomial, such as $\sin x$, must have no dimensions. The derivative d/dx has the opposite dimensions to x, and the dx in an integral has the same dimensions as x.

Example 3

We are given the integral
$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$
Find the value of the integral
$$\int_{-\infty}^{\infty} e^{-ax^2}\,dx.$$

Solution

In the first equation, x must be dimensionless, so both sides are dimensionless. The second equation would also be consistent if both x and a were dimensionless, but we can do better. Suppose we arbitrarily assign x dimensions of length, $[x] = \text{m}$. Then to make the argument of the exponential dimensionless, a must have dimensions $[a] = \text{m}^{-2}$. The dimensions of the left-hand side are $[dx] = \text{m}$. In order to make the dimensions work out on the right-hand side, we must have
$$\int_{-\infty}^{\infty} e^{-ax^2}\,dx \propto \frac{1}{\sqrt{a}}.$$
To find the value of the constant, treat x and a as dimensionless again. Then we know the answer must reduce to $\sqrt{\pi}$ when $a = 1$, so
$$\int_{-\infty}^{\infty} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}.$$

This procedure is completely equivalent to using u-substitution to nondimensionalize everything, but may be faster to see. In general, for all integrals except for the simplest ones, you should use either dimensional analysis or u-substitution to reduce the integral to a dimensionless one.

Remark

Consider the value of the definite integral

$$\int_{-\infty}^{y} e^{-x^2}\,dx.$$

You can try all day to compute the value of this integral, using all the integration tricks you know, but nothing will work. The function e^{-x^2} simply doesn't have an antiderivative in terms of the functions you already know, i.e. in terms of polynomials, exponents and logarithms, and trigonometric functions (for more discussion, see here).

If you ask a computer algebra system like Mathematica, it'll spit out something like $\text{erf}(x)$, which is defined by being an antiderivative of e^{-x^2}. But is this really an "analytic" solution? Isn't that just saying "the integral of e^{-x^2} is equal to the integral of e^{-x^2}"? Well, like many things in math, it depends on what the meaning of the word "is" is.

The fact is, the set of functions we regard as "elementary" is arbitrary; we just choose a set that's big enough to solve most of the problems we want, and small enough to attain fluency with. (Back in the days before calculators, it just meant all the functions whose values were tabulated in the references on hand.) If you're uncomfortable with $\text{erf}(x)$, note that a similar thing would happen if a little kid asked you what the ratio of the opposite to adjacent sides of a right triangle is. You'd say $\tan(x)$, but they could say it's tautological, because the only way to define $\tan(x)$ at their level is as the ratio of opposite to adjacent sides. Similarly, $1/x$ has no elementary antiderivative – unless you count $\log(x)$ as elementary, but ultimately $\log(x)$ is simply *defined* to be such an antiderivative. It's all tautology, but it's still useful.

[1] **Problem 6.** Evaluate the integral

$$\int \frac{dx}{x^2+1}$$

using a trigonometric substitution. Using dimensional analysis, find the integral

$$\int \frac{dx}{x^2+a^2}.$$

[2] **Problem 7.** In particle physics it is conventional to work in "natural units", where the numeric values of $\hbar$ and c are equal to 1. For example, if we take the second as the unit of time, then we can take the light-second as the unit of length, so that $c = 1$ light-second/second. This is usually sloppily written as "$\hbar = c = 1$" so that factors of $\hbar$ and c can be suppressed. However, you can always restore these factors by dimensional analysis.

According to standard references, the mass of the Higgs boson is about 125 GeV, where 1 eV is the energy gained by an electron accelerated through a voltage difference of 1 V. Fix the dimensions of this statement and find the mass of the Higgs boson in kilograms.

[3] **Problem 8.** USAPhO 2002, problem A3.

Example 4

The Schrodinger equation for an electron in the electric field of a proton is

$$-\frac{\hbar^2}{2m}\nabla^2\psi - \frac{e^2}{4\pi\epsilon_0 r}\psi = E\psi.$$

Estimate the size of the hydrogen atom.

Solution

This is yet another dimensional analysis problem: there is only one way to form a length using the quantities given above. We have

$$[m] = \text{kg}, \quad [\hbar] = \text{J}\cdot\text{s} = \text{kg}\,\text{m}^2\,\text{s}^{-1}, \quad [e^2/4\pi\epsilon_0] = \text{J}\cdot\text{m} = \text{kg}\,\text{m}^3\,\text{s}^{-2}.$$

Doing dimensional analysis, the only length scale is the Bohr radius,

$$a_0 = \frac{4\pi\hbar^2\epsilon_0}{me^2} \sim 10^{-10}\,\text{m}.$$

I've thrown in a 4π above because ϵ_0 always appears in the equations as $4\pi\epsilon_0$. The dimensional analysis would be valid without this factor, but as you'll see in problem 11, if you don't include it then annoying compensating factors of 4π will appear elsewhere.

Classically (i.e. without $\hbar$), there is no way to form a length, and hence there should be no classically stable radius for the atom. (This was one of the arguments used by Bohr to motivate quantum mechanics; it appears in the beginning of his paper introducing the Bohr model.) Once we introduce $\hbar$, there are three dimensionful parameters in the problem, as listed above. And there are exactly three fundamental dimensions. So there is only one way to create a length, which we found above, one way to create a time, one way to create an energy, and so on. This means that the solutions to the Schrodinger equation above look qualitatively the same no matter what these parameters are; all that changes are the overall length, time, and energy scales. In problem 11, you'll investigate how this conclusion changes when we add more dimensionful parameters.

Dimensional analysis is especially helpful with scaling relations. For example, a question might ask you how the radius of the hydrogen atom would change in a world where the electron mass was twice as large. You would solve this problem in the exact same way as the example above, using dimensional analysis to show that $a_0 \propto 1/m$.

[3] **Problem 9.** In this problem we'll continue the dimensional analysis of the Schrodinger equation.

(a) Estimate the typical energy scale of quantum states of the hydrogen atom, as well as the typical "velocity" of the electron, using dimensional analysis.

(b) Do the same for one-electron helium, the system consisting of a helium nucleus (containing two protons) and one electron.

(c) Estimate the electric field needed to rip the electron off the hydrogen atom.

Idea 3: Buckingham Pi Theorem

Dimensional analysis can't always pin down the form of the answer. If one has N quantities with D independent dimensions, then one can form $N - D$ independent dimensionless quantities. Dimensional analysis can't say how the answer depends on them.

A familiar but somewhat trivial example is the pendulum: its period depends on L, g, and the amplitude θ_0, three quantities which contain two dimensions (length and time). Hence we can form one dimensionless group, which is clearly just θ_0 itself. The period of a pendulum is $T = f(\theta_0)\sqrt{L/g}$.

Example 5: $F = ma$ 2014 12

A paper helicopter with rotor radius r and weight W is dropped from a height h in air with a density of ρ. Assuming the helicopter quickly reaches terminal velocity, use dimensional analysis to analyze the total flight time T.

Solution

The answer can only depend on the parameters r, W, h, and ρ. There are four quantities in total, but three dimensions (mass, length, and time), so by the Buckingham Pi theorem we can form one independent dimensionless quantity. In this case, it's clearly r/h. Continuing with routine dimensional analysis, we find

$$T = f(r/h)\, h^2 \sqrt{\frac{\rho}{W}}.$$

The form of this expression is a bit arbitrary; for instance, we could also have written $f(r/h)r^2$ in front, or even $f(r/h)r^{37}h^{-35}$. These adjustments just correspond to pulling factors of r/h out of f, not to changing the actual result.

This is as far as we can get with dimensional analysis alone, but we can go further using physical reasoning. If the helicopter quickly reaches terminal velocity, then it travels at a constant speed. So we must have $T \propto h$, which means that $f(x) \propto x$, and

$$T \propto rh\sqrt{\frac{\rho}{W}}.$$

Example 6

An hourglass is constructed with sand of density ρ and an orifice of diameter d. When the sand level above the orifice is h, what is the mass flow rate μ?

Solution

The answer can only depend on ρ, d, h, and g. The Buckingham Pi theorem gives

$$\mu = f(h/d)\rho\sqrt{gd^5}.$$

That's as far as we can get with dimensional analysis; to go further we need to know more about sand. If we were dealing with an ideal fluid, then the flow speed would be $v = \sqrt{2gh}$ by Torricelli's law, which means the flow rate has to be proportional to $\sqrt{h}$. Then $f(x) \propto \sqrt{x}$, giving the result $\mu \propto \rho d^2\sqrt{gh}$. This is a good estimate as long as the orifice isn't so small that viscosity starts to dominate.

But this isn't how sand works. Sand is a granular material, whose motion is dominated by the friction between sand grains. So the higher pressure doesn't actually propagate to the orifice, and the flow rate is independent of h, which is apparent from watching an hourglass run. Then $f(x)$ is a constant, giving $\mu \propto \rho\sqrt{gd^5}$, which has been experimentally verified.

Remark

One has to be a little careful with the Buckingham Pi theorem. For example, if all we had were 3 speeds v_i, we can form two dimensionless quantities: v_1/v_2 and v_1/v_3. (The quantity v_2/v_3 is not independent, since it is the quotient of these two.) But there are 3 quantities with 2 dimensions (length and time), so we expect only 1 dimensionless quantity.

The problem is that the two dimensions really aren't independent: for any quantity built from the v_i, a power of length always comes with an inverse power of time, so there's only one independent dimension. These considerations can be put on a more rigorous footing in linear algebra, where the Buckingham Pi theorem is merely a special case of the rank-nullity theorem. If you're ever in doubt, you can just forget about the theorem and play with the equations directly.

Remark

Dimensional analysis is an incredibly common tool in Olympiad physics because it lets you say a lot even without much advanced knowledge. If a problem ever says to find some quantity "up to a constant/dimensionless factor", or how that quantity scales as another quantity changes, or what that quantity is proportional to, it's almost certainly asking you to do dimensional analysis. Another giveaway is if the problem *looks* extremely technical and advanced, because they can't actually be.

[3] **Problem 10** (Insight)**.** In this problem we'll do one of the most famous dimensional analyses of all time: estimating the yield of the first atomic bomb blast. Such a blast will create a shockwave of air, which reaches a radius R at time t after the blast. The air density is ρ, and we want to estimate the blast energy E.

(a) Declassified photographs of the blast indicate that $R \approx 100\,\mathrm{m}$ at time $t \approx 15\,\mathrm{ms}$. The density of air is $\rho \approx 1\,\mathrm{kg/m^3}$. Estimate the blast energy E.

(b) How much mass-energy (in grams) was used up in this blast?

(c) If we measure the entire function $R(t)$, what general form would we expect it to have, if this dimensional analysis argument is correct?

Remark

The British physicist G. I. Taylor performed the dimensional analysis in problem 10 upon seeing a picture of the first atomic blast in a magazine. The result was so good that the physicists at the Manhattan project thought their security had been breached!

During World War II, the exact value of the critical mass needed to set off a nuclear explosion was important and nontrivial information. The Nazi effort to make a bomb had been stopped by Werner Heisenberg's huge overestimation of this quantity, and after the war, the specific value was kept a closely guarded secret. That is, it was until 1947, when a Chinese physicist got the answer using a rough estimate that took four lines of algebra.

[5] **Problem 11.** We now consider the Schrodinger equation for the hydrogen atom in greater depth. We begin by switching to dimensionless variables, which is useful for the same reason that writing integrals in terms of dimensionless variables is: it highlights what is independent of unit choices.

(a) Define a dimensionless length variable $\tilde{r} = r/a_0$, where a_0 is the length scale found in example 4. The ∇^2 term in the Schrodinger equation is a second derivative, the 3D generalization of d^2/dx^2. Using the chain rule, argue that

$$\tilde{\nabla}^2 = a_0^2 \nabla^2$$

where $\tilde{\nabla}$ is the gradient with respect to $\tilde{r}$.

(b) Similarly define a dimensionless energy $\tilde{E} = E/E_0$, using the energy scale E_0 found in problem 9. Show that the Schrodinger equation can be written in a form like

$$-\tilde{\nabla}^2 \psi - \frac{1}{\tilde{r}} \psi = \tilde{E} \psi$$

Here I've suppressed all dimensionless constants, like factors of 2, because they depend on how you choose to define E_0 and don't really matter at this level of precision.

The result of this part confirms what we concluded above: solutions to the Schrodinger equation don't qualitatively depend on the values of the parameters, because they all come from scaling a solution to this one dimensionless equation appropriately.

(c) This is no longer true in relativity, where the total energy is

$$E = \sqrt{p^2 c^2 + m^2 c^4}.$$

Assuming $p \ll mc$, perform a Taylor expansion to show that the next term is Ap^4, and find the coefficient A. (If you don't know how to do this, work through the next section first.)

(d) In quantum mechanics, the momentum is represented by a gradient, $p \to -i\hbar\nabla$. (We will see why in **X1**.) Show that the Schrodinger equation with the first relativistic correction is

$$-\frac{\hbar^2}{2m} \nabla^2 \psi - \frac{e^2}{4\pi\epsilon_0 r} \psi + \hbar^4 A \nabla^4 \psi = E\psi.$$

(e) Since there is now one more dimensionful quantity in the game, it is possible to combine the quantities to form a dimensionless one. Create a dimensionless quantity α that is proportional to $e^2/4\pi$, then numerically evaluate it. This is called the fine structure constant. It serves as an objective measure of the strength of the electromagnetic force, because it is dimensionless, and hence its value doesn't depend on an arbitrary unit system.

(f) As the number of protons in the nucleus increases, the relativistic correction becomes more important. Estimate the atomic number Z where the correction becomes very important.

You probably won't see any differential equations as complex as the ones in the above problem anywhere in Olympiad physics, but the key idea of using dimensionless quantities to simplify and clarify the physics can be used everywhere.

[5] **Problem 12.** IPhO 2007, problem "blue". This problem applies thermodynamics and dimensional analysis in some exotic contexts.

Example 7

Estimate the Young's modulus for a material with interatomic separation a and typical atomic bond energy E_b. Use this to estimate the spring constant of a rod of area A and length L, as well as the speed of sound, if each atom has mass m.

Solution

This example is to get you comfortable with the Young's modulus Y, which occasionally comes up. It is defined in terms of how much a material stretches as it is pulled apart,

$$Y = \frac{\text{stress}}{\text{strain}} = \frac{\text{restoring force/cross-sectional area}}{\text{change in length/length}}.$$

The Young's modulus is a useful way to characterize materials, because unlike the spring constant, it doesn't depend on the shape of the material. For example, putting two identical springs side-by-side doubles the spring constant, because they both contribute to the force. But since the stress is the force per area, it's unchanged. Similarly, putting two identical springs end-to-end halves the spring constants, because they both stretch, but since the strain is change in length per length, it's unchanged. So you would quote a material's Young's modulus instead of its spring constant, for the same reason you would quote a material's resistivity instead of its resistance.

We note that Y has the dimensions of energy per length cubed, so

$$Y \sim \frac{E_b}{a^3}$$

solely by dimensional analysis. (Of course, for this dimensional analysis to work, one has to understand why E_b and a are the only relevant quantities. It's because Y, or equivalently the spring constant k, determines the energy stored in a stretched spring. But microscopically this comes from the energy stored in interatomic bonds when they're stretched. So the relevant energy scale is the bond energy E_b, and the relevant distance scale is a, because that determines how many bonds get stretched, and by how much.)

To relate Y to the spring constant of a rod, note that

$$Y = \frac{F/A}{\Delta L/L} = \frac{L}{A}\frac{F}{\Delta L} = k\frac{L}{A}$$

for a rod, giving the estimate $k \sim AE_b/La^3$. This is correct to within an order of magnitude!

To relate Y to the speed of sound, note that the sound speed, like most wave speeds, depends on the material's inertia and its restoring force against distortions. Since the speed of sound doesn't depend on the extrinsic features of a metal object, such as a length, both of these should be measured intrinsically. The intrinsic measure of inertia is the mass density $\rho \sim m/a^3$, while the intrinsic measure of restoring force is just Y. By dimensional analysis,

$$v \sim \sqrt{\frac{Y}{\rho}} \sim \sqrt{\frac{E_b/a^3}{m/a^3}} \sim \sqrt{\frac{E_b}{m}}.$$

This is also reasonably accurate. For example, in diamond, $E_b \sim 1\,\text{eV}$ (a typical atomic energy scale), while a carbon nucleus contains 12 nucleons, so to the nearest order of magnitude, $m \sim 10m_p$, where a useful fact is $m_p \sim 1\,\text{GeV/c}^2$. Thus,

$$v \sim \sqrt{\frac{1\,\text{eV}}{10^{10}\,\text{eV}}}\,c \sim 10^{-5}c \sim 3\,\text{km/s}$$

which is roughly right. (The true answer is $12\,\text{km/s}$.)

Amazingly, we can get an even rougher estimate of v for any solid in terms of nothing besides fundamental constants. To be very rough, the binding energy is on the order of that of hydrogen. As you found in problem 9, this is, by dimensional analysis,

$$E_b \sim \frac{1}{4\pi\epsilon_0}\frac{e^2}{a_0} \sim m_e\left(\frac{e^2}{4\pi\epsilon_0\hbar}\right)^2.$$

We take the nuclear mass to be very roughly the proton mass m_p, which gives

$$\frac{v}{c} \sim \sqrt{\frac{m_e}{m_p}\left(\frac{e^2}{4\pi\epsilon_0\hbar c}\right)^2} \sim \alpha\sqrt{\frac{m_e}{m_p}}$$

where α is as found in problem 11. This expresses the speed of sound in terms of the dimensionless strength of electromagnetism α, the electron to proton mass ratio, and the speed of light. Of course, the approximations we have made here have been so rough that now the answer is off by *two* orders of magnitude, but now we know how the answer would change if the fundamental constants did.

Estimates as simple as these can be surprising to even seasoned physicists: in 2020, the simple estimate above was rediscovered and published in one of the top journals in science. If you want to learn how to do more of these estimates, this paper is a good starting point.

Remark

A warning: from these examples, you could get the idea that dimensional analysis gives you nearly godlike powers, and the ability to write down the answer to most physics problems instantly. In reality, it only works if you're pretty sure your physical system depends on only about 3 or 4 variables – and the hard part is often finding *which* variables matter. For example, as we saw above, you can't get Kepler's third law for free because that requires knowing the dimensions of G, which require knowing that gravity is an inverse square law in the first place, a luxury Kepler didn't have. And as another example, we couldn't have figured out $E = mc^2$ long before Einstein, as who would have thought that the speed of light had anything to do with the energy of a lump of matter? Without the framework of relativity, it seems as irrelevant as the speed of sound or the speed of water waves. To illustrate this point, we consider two contrasting examples below.

In reality, dimensional analysis is best for problems where it's easy to see which variables matter, problems where you're explicitly told what variables matter, and for checking your work, which you should get in the habit of doing all the time!

Example 8

Cutting-edge archeological research has found that the famed T. Rex was essentially a gigantic chicken. Suppose a T. Rex is about $N = 20$ times larger in scale than a chicken. How much larger is its weight, cross-sectional area of bone, walking speed, and maximum jump height?

Solution

These kinds of biological scaling arguments are fun to think about, though the reliability of the results is somewhat questionable – the data is extremely noisy, and if any given scaling law doesn't quite match it, you can always think a bit more, and come up with a new argument yielding a different scaling. But here are a few simple examples:

- Since the densities should match, the weight should scale with the volume, so as N^3.
- Since the maximum compressive pressure that bone can take should be the same, the bone area should scale with the weight, so also as N^3. That is, the width of the bones scales as $N^{3/2}$, while their length L scales only as N. This is the reason small animals are strong relative to their weight, while large ones need to be very bony to even stand. The largest animals today are whales, as they don't need to support their own weight.
- As a very crude model of walking, we can think of the legs as swinging like a free pendulum. The length of one step is proportional to L, while the period of the steps is proportional to $\sqrt{L}$. Thus, the walking speed scales as $\sqrt{L} \propto \sqrt{N}$.
- The energy stored in the muscle cells scales with the volume, but the mass also scales with the volume. Since the jump height satisfies $E = mgh$ where E is the energy stored by the muscles, $h \propto N^0$. So a dinosaur can't jump much higher than a human – and indeed, *we* can't jump much higher than fleas can!

There's an entire literature on these arguments. For instance, this delightful paper discusses how furry mammals shake to dry themselves off. This is an increasingly severe problem for smaller mammals, since a relatively larger amount of water will cling to them after getting wet, which can cause hypothermia. Using elementary fluid mechanics, the paper argues that the optimal frequency the mammal will shake to dry itself off scales as $f \propto m^{-3/16}$.

Example 9

A person with density ρ and total energy E stored in their muscles can jump to a height h in gravity g. How high would they be able to jump in gravity $10g$?

Solution

By dimensional analysis, the only possible answer is

$$h \propto \left(\frac{E}{\rho g}\right)^{1/4}$$

which means that in gravity $10g$, they can jump to height $h/10^{1/4}$.

But this is completely wrong! In gravity $10g$, a person wouldn't be able to jump at all; they'd be so crushed by their own weight that they wouldn't even be able to stand. The actual answer depends on details of the biomechanics of muscles and bone, which involve more dimensionful quantities than just the total energy E. So, as remarked above, you can't solve literally any problem by just listing a few relevant quantities and doing dimensional analysis – you need to make sure those are the *only* relevant quantities.

2 Approximations

Idea 4: Taylor Series

For small x, a function $f(x)$ may be approximated as

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2}f''(0) + \ldots + \frac{x^n}{n!}f^{(n)}(0) + O(x^{n+1})$$

where $O(x^{n+1})$ stands for an error term which grows at most as fast as x^{n+1}.

There are a few Taylor series that are essential to know. The most important are

$$\exp(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + O(x^4), \quad \log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - O(x^4)$$

and the small angle approximations

$$\sin x = x - \frac{x^3}{6} + O(x^5), \quad \cos x = 1 - \frac{x^2}{2} + O(x^4).$$

Another Taylor series you learned long before calculus class is

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + O(x^4).$$

Usually you'll only need the first one or two terms, but for practice we'll do examples with more. If any of these results aren't familiar, you should rederive them!

Example 10

Find the Taylor series for $\tan x$ up to, and including the fourth order term.

Solution

By the fourth order term, we mean the term proportional to x^4. (Not the fourth nonzero term, which would be $O(x^7)$.) Of course, $\tan x$ is an odd function, so the $O(x^4)$ term is zero, which means we only need to expand up to $O(x^3)$. That means we can neglect $O(x^4)$ terms and higher everywhere in the computation, subject to some caveats we'll point out later.

By definition, we have

$$\tan x = \frac{\sin x}{\cos x} = \frac{x - x^3/6 + O(x^5)}{1 - x^2/2 + O(x^4)}.$$

However, it's a little tricky because we have a Taylor series in a denominator. There are two ways to deal with this. We could multiply both sides by $\cos x$, and expand $\tan x$ in a Taylor series with unknown coefficients. Then we would get a system of equations that will allow us to solve for the coefficients recursively, a technique known as "reversion of series".

A faster method is to use the Taylor series for $1/(1-x)$. We have

$$\frac{1}{1-u} = 1 + u + O(u^2)$$

and substituting $u = x^2/2 - O(x^4)$ gives

$$\frac{1}{\cos x} = 1 + \frac{x^2}{2} + O(x^4).$$

Therefore, we conclude

$$\tan x = (x - x^3/6 + O(x^5))(1 + x^2/2 + O(x^4)) = x + x^3/3 + O(x^5).$$

Here I was fairly careful with writing out all the error terms and intermediate steps, but as you get better at this process, you'll be able to do it faster. (Of course, one could also have done this example by just directly computing the Taylor series of $\tan x$ from its derivatives. This is possible, but for more complicated situations it's generally not a good idea, because computing high derivatives of a complex expression tends to get very messy. It's better to just Taylor expand the individual pieces and combine the results, as we did here.)

Remark

Finding series up to a given order can be subtle. For example, if you want to compute an $O(x^4)$ term, it is *not* always enough to expand everything up to $O(x^4)$, because powers of x might cancel. To illustrate this, the last step here is wrong:

$$\tan x = \frac{x^3 \sin x}{x^3 \cos x} = \frac{x^4 + O(x^6)}{x^3 + O(x^5)} \neq x + O(x^5).$$

[2] **Problem 13.** Find the Taylor series for $1/\cos x$ up to and including the fourth order ($O(x^4)$) term.

[2] **Problem 14.** Extend the computation above to get the x^5 term in the Taylor series for $\tan x$.

[3] **Problem 15.** For small x, approximate the quantity

$$\frac{x^2 e^x}{(e^x - 1)^2} - 1$$

to lowest order. That is, find the first nonzero term in the Taylor series. (Hint: if you don't take enough terms in the Taylor series to begin with, you'll get an answer of zero, indicating you approximated too loosely. But if you take too many, the computation will get extremely messy.)

[3] **Problem 16.** The function $\cos^{-1}(1 - x)$ does not have a Taylor series about $x = 0$. However, it does have a series expansion about $x = 0$ in a different variable.

(a) What is this variable, and what's the first term in the series?

(b) $\star$ Can you find the next nontrivial term in the series?

Idea 5: Binomial Theorem

When the quantity xn is small, it is useful to use the binomial theorem,

$$(1 + x)^n = 1 + xn + O(x^2 n^2).$$

It applies even when n is not an integer. In particular, n can be very large, very small, or even negative. The extra terms will be small as long as xn is small. If desired, one can find higher terms using binomial coefficients,

$$(1 + x)^n = \sum_{m=0}^{\infty} \binom{n}{m} x^m$$

where the definition of the binomial coefficient is formally extended to arbitrary real n.

The binomial theorem is one of the most common approximations in physics. It's really just taking the first two terms in the Taylor series of $(1 + x)^n$, but we give it a name because it's so useful.

[1] **Problem 17.** Suppose the period of a pendulum is one second, and recall that

$$T = 2\pi \sqrt{\frac{L}{g}}.$$

If the length is increased by 3% and g is increased by 1%, use the binomial theorem to estimate how much the period changes. This kind of thinking is extremely useful when doing experimental physics, and you should be able to do it in your head.

[1] **Problem 18.** Consider an electric charge q placed at $x = 0$ and a charge $-q$ placed at $x = d$. The electric field along the x axis is then

$$E(x) = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{x^2} - \frac{1}{(x-d)^2}\right).$$

For large x, use the binomial theorem to approximate the field.

[3] **Problem 19.** Some exercises involving square roots.

(a) Manually find the Taylor series for $\sqrt{1+x}$ up to second order, and verify they agree with the binomial theorem.

(b) Approximate $\sqrt{1+2x+x^2}$ for small x using the binomial theorem. Does the result match what you expect? If not, how can you correct it?

Example 11: Birthday Paradox

If you have n people in a room, around how large does n have to be for there to be at least a 50% chance of two people sharing the same birthday?

Solution

Imagine adding people one at a time. The second person has a $1/365$ chance of sharing a birthday with the first. If they don't share a birthday, the third person has a $2/365$ chance of sharing a birthday with either, and so on. So a decent estimate for n is the n where

$$\left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)\ldots\left(1 - \frac{n}{365}\right) \approx \frac{1}{2}.$$

The surprising point of the birthday paradox is that $n \ll 365$. So we can use the binomial theorem in reverse, approximating the left-hand side as

$$\left(1 - \frac{1}{365}\right)\left(1 - \frac{1}{365}\right)^2\ldots\left(1 - \frac{1}{365}\right)^n \approx \left(1 - \frac{1}{365}\right)^{n^2/2}$$

which is valid since $n/365$ is small. It's tempting to use the binomial theorem again to write

$$\left(1 - \frac{1}{365}\right)^{n^2/2} \approx 1 - \frac{n^2}{2 \cdot 365} = \frac{1}{2}$$

which gives $n = 19$. However, this is a bad approximation, because the binomial theorem only works if $(n^2/2)(1/365)$ is very small, but here we've set it to $1/2$, which isn't particularly small. Since the series expansion variable is $1/2$, each term in the series expansion is *roughly* $1/2$ as big as the last (ignoring numerical coefficients), so we expect to be off by about $(1/2)^2 = 25\%$.

The binomial theorem is an expansion for $(1+x)^y$ which works when xy is small. Here xy isn't small, and we instead want an approximation that works when only x is small. One trick to dealing with an annoying exponent is to take the logarithm, since that just turns it into a multiplicative factor. Note that

$$\log((1+x)^y) = y\log(1+x) \approx yx$$

by Taylor series, which implies that

$$(1+x)^y \approx e^{yx}$$

when x is small, an important fact which you should remember. So we have

$$\left(1 - \frac{1}{365}\right)^{n^2/2} \approx e^{-n^2/2(365)} = \frac{1}{2}$$

and solving gives $n = 22.5$. We should round up since n is actually an integer, giving $n = 23$, which is indeed the exact answer.

Remark

Precisely how accurate is the approximation $(1+x)^y \approx e^{yx}$? Note that the only approximate step used to derive it was taking $\log(1+x) \approx x$, which means we can get the corrections by expanding to higher order. If we take the next term, $\log(1+x) \approx x - x^2/2$, then we find

$$(1+x)^y \approx e^{yx} e^{-x^2 y/2}.$$

Note that because we are approximating the logarithm of the quantity we want, the next correction is multiplicative rather than additive; we'll see a similar situation with Stirling's approximation in **T2**. Our approximation has good fractional precision as long as $x^2 y \ll 1$. In the previous example, $x^2 y/2 = (22.5/365)^2/4 = 0.1\%$, so our answer was quite accurate.

[2] **Problem 20.** Find a series approximation for x^y, given that y is small and x is neither small nor exponentially huge. (Hint: to check if you have it right, you can try concrete numbers, such as $y = 0.01$ and $x = 10$. The series expansion variable may look a bit unusual.)

Remark

If these questions seem complicated, rest assured that 90% of approximations on the USAPhO and IPhO boil down to using

$$\sin x \approx x, \quad \cos x \approx 1 - x^2/2, \quad (1+x)^n \approx 1 + xn, \quad e^x \approx 1 + x, \quad \log(1+x) \approx x.$$

I've given you a lot of subtle situations above, but it's these that you have to know by heart. Almost all situations where you will use these will look like problem 17 or problem 18.

3 Solving Equations

Idea 6

In Olympiads, you may have to find numeric solutions for equations that can't be solved analytically. A simple but reliable method is to "guess and check", starting with a reasonable first guess (e.g. derived by solving an approximated version of the equation, or sketching the graphs of both sides), plugging it into both sides, then proceeding with binary search.

[3] **Problem 21.** Sometimes, you can get an accurate numeric answer very quickly on a basic calculator by using the method of iteration, which solves equations of the form $x = f(x)$.

(a) Take a scientific calculator (in radians), put in any number, and press the "cos" button many times. Convince yourself that the final number you get is the unique solution to $x = \cos x$.

(b) What are the key features of the graphs of x and $\cos x$ that made this work? For example, why doesn't pressing $\cos^{-1}$ repeatedly give the same result? As another example, since $x = \sin x$ has a unique solution, why does repeatedly pressing sin not work so well?

(c) Find a nonzero solution for $x = \tan(x/2)$.

(d) Find a nonzero solution for $e^x - 1 = 2x$.

(e) Find a positive solution for $x^x = e$.

[2] **Problem 22.** [A] Newton's method is a more sophisticated method for solving equations, which converges substantially faster than binary search. Suppose we want to solve the equation $f(x) = 0$. Starting with a nearby guess x_0, we evaluate $f(x_0)$ and $f'(x_0)$, then find our next guess by applying the tangent line approximation at this point,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

The process repeats until we get a suitably accurate answer.

(a) Use Newton's method to solve $x = \cos x$.

(b) Newton's method converges quadratically, in the sense that for typical functions, if your current guess is ϵ away from the answer, the next guess will be $O(\epsilon^2)$ away. (This implies that the number of correct digits in the answer roughly doubles with each iteration!) Explain why, and then find an example where Newton's method *doesn't* converge this fast.

Newton's method is very important in general, but it's not that useful on Olympiads. It takes a while to set up, especially if the derivative f' is complicated, and you usually don't need that many significant figures in your answer anyway. (There are alternatives to Newton's method, such as Halley's method, that converge even faster, but the tradeoff is the same: each iteration takes more effort to calculate, as higher derivatives of f must be computed.)

Remark

You've seen several approximation methods above, and going forward, you should feel free to use whichever looks best in each situation. However, if you're solving problems using the same calculator you use for schoolwork, you should make sure to not rely on its more advanced features. In Olympiads, you're generally only allowed to use an extremely basic scientific calculator, with a tiny display and no memory except for the "Ans" key.

Example 12

In units where $c = 1$, the Lorentz factor is defined as

$$\gamma = \frac{1}{\sqrt{1 - v^2}}.$$

Suppose that a particle traveling very close to the speed of light has $\gamma = 10^8$. Find the difference Δv between its speed and the speed of light.

Solution

This problem looks easy; by some trivial algebra we find

$$\Delta v = 1 - \sqrt{1 - 1/\gamma^2}.$$

But when you plug this into a cheap scientific calculator, you get *zero*, or something that's quite far from the right result. The problem is that we are trying to find a small quantity Δv by subtracting two nearby, much larger quantities. But the calculator has limited precision, and it ends up rounding $1 - 1/\gamma^2 = 1 - 10^{-16}$ a bit, giving a completely wrong answer!

Instead, we can apply the binomial theorem to find

$$\Delta v \approx \frac{1}{2\gamma^2} + O(1/\gamma^4).$$

This is no longer the exact answer, but it's a great approximation, because the error term is around $1/\gamma^2 \sim 10^{-16}$ times as small as the answer, and it's easy for a calculator to evaluate. The lesson is that it's better to be accurate in practice than to be precise in theory.

[1] **Problem 23.** Find the solutions of the equation $x^2 - 10^{20}x + 1 = 0$ to reasonable accuracy.

[4] **Problem 24.** [A] Consider the equation $\epsilon x^3 - x^2 + 1 = 0$, where ϵ is small. Find approximate expressions for all three roots of this equation, up to and including terms of order ϵ.

4 Limiting Cases

Idea 7

Limiting cases can be used to infer how the answer to a physical problem depends on its parameters. It is primarily useful for remembering the forms of formulas, but can also be powerful enough to solve multiple choice questions by itself.

Example 13

What is the horizontal range of a rock thrown with speed v at an angle θ to the horizontal?

Solution

This result is easy to derive, but dimensional analysis and extreme cases can be used to recover the answer too. It can only depend on v, g, and θ, so by dimensional analysis it is proportional to v^2/g. This is sensible, since the range increases with v and decreases with g. Now, the range is zero in the extreme cases $\theta = 0$ and $\theta = \pi/2$, but not anywhere in between, so if we remember the range contains a simple trigonometric function, it must be $\sin(2\theta)$, so

$$R \propto \frac{v^2}{g} \sin(2\theta).$$

We can also get the prefactor by a simple limiting case, the case $\theta \ll 1$. In this case, by the small angle approximation,

$$v_x \approx v, \quad v_y \approx v\theta.$$

The time taken is $t = 2v_y/g$, so the range is

$$R \approx v_x t = \frac{2v^2}{g}\theta.$$

Thus there is no proportionality constant; the answer is

$$R = \frac{v^2}{g} \sin(2\theta).$$

In reality, it's probably faster to go through the full derivation than all of this reasoning, but if you're just not sure about whether it's a sine or a cosine, or what the prefactor is, then limiting cases can be quickly used to recover that piece. Also note that the approximations we used above are frequently useful for evaluating limiting cases.

Example 14

Consider an Atwood's machine with masses m and M, and a massless pulley. Find the tension in the string.

Solution

Since the equations involved are all linear equations, we expect the answer should also be simple. It can only depend on g, m, and M, so by dimensional analysis, it must be proportional to g. By dimensional analysis, this must be multiplied by something with one net power of mass. Since the answer remains the same if we switch the masses, it should be symmetric in m and M.

Given all of this, the simplest possible answer would be

$$T \propto g(m+M).$$

To test this, we consider some limiting cases. If $M \gg m$, the mass M is essentially in freefall, so the mass m accelerates upward with acceleration g. Then the tension is approximately $2mg$. Similarly, in the case $M \ll m$, the tension is approximately $2Mg$. These can't be satisfied by the form above.

The next simplest option is a quadratic divided by a linear expression. Both of these must be symmetric, so the most general possibility is

$$T = g\frac{A(m^2+M^2)+BmM}{m+M}.$$

Then the limiting cases can be satisfied if $A = 0$ and $B = 2$, giving

$$T = \frac{2gmM}{m+M}.$$

[1] **Problem 25.** Find the perimeter of a regular N-gon, if L is the distance from the center to any of the sides. By considering a limiting case, use this to derive the circumference of a circle.

[1] **Problem 26.** Use similar reasoning to find the acceleration of the Atwood's machine. (We will show an even easier way to do this, using "generalized coordinates", in **M4**.)

[2] **Problem 27** (Morin 1.6)**.** A person throws a ball (at an angle of her choosing, to achieve the maximum distance) with speed v from the edge of a cliff of height h. Which of the below could be an expression for the maximal range?

$$\frac{gh^2}{v^2}, \quad \frac{v^2}{g}, \quad \sqrt{\frac{v^2h}{g}}, \quad \frac{v^2}{g}\sqrt{1+\frac{2gh}{v^2}}, \quad \frac{v^2}{g}\left(1+\frac{2gh}{v^2}\right), \quad \frac{v^2/g}{1-2gh/v^2}.$$

If desired, try Morin problems 1.13, 1.14, and 1.15 for additional practice.

[2] **Problem 28.** Consider a triangle with side lengths a, b, and c. It turns out the area of its incircle can be expressed purely by multiplying and dividing combinations of these lengths. Moreover, the answer is the simplest possible one consistent with limiting cases, dimensional analysis, and symmetry. Guess it!

While we won't have more questions that are explicitly about dimensional analysis or limiting cases, these are not techniques but ways of life. For all future problems you solve, you should be constantly checking the dimensions and limiting cases to make sure everything makes sense.

5 Manipulating Differentials

You might have been taught in math class that manipulating differentials like they're just small, finite quantities, and treating derivatives like fractions is "illegal". But it's also very useful.

Idea 8

Derivatives can be treated like fractions, if all functions have a single argument.

The reason is simply the chain rule. The motion of a single particle only depends on a single parameter, so the chain rule is just the same as fraction cancellation. For example,

$$\frac{dv}{dt} = \frac{d}{dt} v(x(t)) = \frac{dv}{dx}\frac{dx}{dt}$$

which show that "canceling a dx" is valid. Similarly, you can show that

$$\frac{dy}{dx}\frac{dx}{dy} = 1$$

by considering the derivative with respect to x of the function $x(y(x)) = x$.

As a warning, for functions of multiple arguments, the idea above breaks down. For example, for a function $f(x(t), y(t))$, the chain rule says

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

where there are two terms, representing the change in f from changes only in x, and only in y. Therefore, when we start studying thermodynamics, where multivariable functions are common, we will treat differentials more carefully. But for now the basic rules will do.

Remark: Rigorous Notation

Math students tend to get very upset about the above idea: they say we shouldn't use convenient notation if it hides what's "really" going on. And they're right, if your goal is to put calculus on a rigorous footing. But in physics we have no time to luxuriate in such rigor, because we want to figure out how specific things work. The point of notation is to help us do that by suppressing mathematical clutter. A good notation suppresses *as much as possible* while still giving correct results in the context it's used.

To illustrate the point, note that elementary school arithmetic is itself an "unrigorous" notation that hides implementation details. If we wanted to be rigorous about, say, defining the number 2, we would write it as $S(1)$ where S is the successor function, obeying properties specified by the Peano axioms. And 4 is just a shorthand for $S(S(S(1)))$, so $2 + 2 = 4$ means

$$S(1) + S(1) = S(S(S(1))).$$

Even this is not "rigorous", because the Peano axioms don't specify how the numbers or the successor function are defined, just what properties they have to obey. To go deeper, we could define the integers as sets, and operations like $+$ in terms of set operations. For example, in one formulation, we start with the empty set $\varnothing$ and define

$$4 = S(S(S(1))) = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}.$$

People have seriously advocated for 1^{st} grade math to be taught this way, which has always struck me as insane. You can *always* add more arbitrary layers of structure underneath the current foundation, so such layers should only be added when absolutely necessary.

Here's another example. For uniformly accelerated motion starting from rest, $v(t) = at$, physics students would say that $v(x) = \sqrt{2ax}$ by the kinematic equations, while math students would say $v(x) = ax$ by definition. Who is correct? The point is that basic physics and math courses use equations differently. In introductory physics, we often denote several distinct mathematical functions with the same symbol, if they all represent the same physical quantity. (Otherwise, even a trivial problem would need half the alphabet.) By contrast, basic math courses carefully distinguish functions, but then denote distinct physical quantities with the same symbol; often 1 m, 1 cm, and 1 s are all replaced with the number 1.

The crucial point is that nobody is wrong. There is no One True Definition of notation, which is ultimately just squiggly marks people make by dragging graphite cylinders against sheets of wood pulp. Every community makes its own notation for its own needs. And any notation system has to forget about something, or else it would be too clunky to do anything.

Remark: Advanced Notation

As an addendum to the previous remark, it turns out that as you get deeper into math and physics, notation tends to converge. For example:

- The physicist's "wrong" use of $v(t)$ and $v(x)$ can be formalized by differential geometry: here v is a scalar field defined on the particle's path, which is a one-dimensional manifold, and $v(t)$ and $v(x)$ are parametrizations of it in different coordinate charts.
- In math classes, vectors are anything you can take linear combinations of, but in physics classes we also require that they specify a direction in physical space, which math students often criticize as wrong, or meaningless. But the physicist is actually using more advanced math, which the student doesn't know yet: the physicist's vector is a element of a vector space carrying the fundamental representation of $SO(3)$.
- Most vectors flip sign under an inversion of space, $\mathbf{r} \to -\mathbf{r}$ and $\mathbf{p} \to -\mathbf{p}$, but "axial vectors" such as $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ don't. This also strikes many math students as a blatant inconsistency, but the reality is again that an axial vector is just a more advanced mathematical object they haven't met yet, specifically a rank 2 differential form.
- More generally, the "unrigorous" manipulations of differentials above, which we showed give you the right answer anyway, gain a rigorous footing in terms of differential forms. In fact, they become the *preferred* way to denote integration on general manifolds.

Arguments about notation are mostly raised by beginning students, who see the one way they know as the only possible way. Professionals know it both ways, and adjust as needed.

Example 15

Derive the work-kinetic energy theorem, $dW = F\,dx$.

Solution

Canceling the mass from both sides, we wish to show

$$\frac{1}{2}d(v^2) = a\,dx.$$

To do this, note that

$$\frac{1}{2}d(v^2) = v\,dv = \frac{dx}{dt}\,dv = \frac{dv}{dt}\,dx = a\,dx$$

as desired. If you're not satisfied with this derivation, because of the bare differentials floating around, we can equivalently prove that $F = dW/dx$, by noting

$$\frac{dW}{dx} = mv\frac{dv}{dx} = mv\frac{dv}{dt}\frac{dt}{dx} = m\frac{dv}{dt} = F.$$

[2] **Problem 29.** Some more about power.

(a) Use similar reasoning to derive $P = Fv$.

(b) An electric train has a power line that can deliver power $P(x)$, where x is the distance along the track. If the train starts at rest at $x = 0$, find its speed at point x_0 in terms of an integral of $P(x)$. (Hint: try to get rid of the dt's to avoid having to think about the time-dependence.)

[2] **Problem 30** (Kalda)**.** The deceleration of a boat in water due to drag is given by a function $a(v)$. Given an initial velocity v_0, write the total distance the boat travels as a single integral.

[5] **Problem 31.** A particle in a potential well.

(a) Consider a particle of mass m and energy E with potential energy $V(x)$, which performs periodic motion. Write the period of the motion in terms of a single integral over x.

(b) Suppose the potential well has the form $V(x) = V_0(x/a)^n$ for $n > 0$. If the period of the motion is T_0 when it has amplitude A_0, find the period when the amplitude is A, by considering how the integral you found in part (a) scales with A.

(c) Find a special case where you can check your answer to part (b). (In fact, there are two more special cases you can check, one which requires negative n and negative V_0, and one which requires $V(x)$ to be replaced with its absolute value.)

(d) Using a similar method to part (a), write down an integral over θ giving the period of a pendulum with length L in gravity g, without the small angle approximation. Using this, compute the period of the pendulum with amplitude θ_0, up to order θ_0^2. (This result was first published by Bernoulli, in 1749.)

(e) ⋆ Part (d) is the kind of involved computation you might see in a graduate mechanics course. But if you think you're *really* tough, you can go one step further. Consider a mass m oscillating on a spring of spring constant k with amplitude A. Calculate its period of oscillation up to order A^2, accounting for special relativity. (Concretely, assume that the spring force doesn't change the rest mass m, and has a potential $U = kx^2/2$. In relativity, the force $F = -dU/dx$ still obeys $F = dp/dt$, but now $E = \gamma mc^2$ and $p = \gamma mv$, where $\gamma = 1/\sqrt{1 - v^2/c^2}$.)

6 Multiple Integrals

It's also useful to know how to set up multiple integrals. This is fairly straightforward, though technically an "advanced" topic, so we'll demonstrate it by example. For further examples, see chapter 2 of Wang and Ricardo, volume 1, or MIT OCW 18.02, lectures 16, 17, 25, and 26.

Idea 9

In most Olympiad problems, multiple integrals can be reduced to single integrals by symmetry.

Example 16

Calculate the area of a circle of radius R.

Solution

The area A is the integral of dA, i.e. the sum of the infinitesimal areas of pieces we break the circle into. As a first example, let's consider using Cartesian coordinates. Then the pieces will be the rectangular regions centered at (x, y) with sides (dx, dy), which have area $dx\,dy$. The area is thus

$$A = \int dA = \int dx \int dy.$$

The only tricky thing about setting up the integral is writing down the bounds. The inner integral is done first, so its bounds depend on the value of x. Since the boundary of the circle is $x^2 + y^2 = R^2$, the bounds are $y = \pm\sqrt{R^2 - x^2}$. Thus we have

$$A = \int_{-R}^{R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy.$$

We then just do the integrals one at a time, from the inside out, like regular integrals,

$$A = \int_{-R}^{R} 2\sqrt{R^2 - x^2}\,dx = 2R^2 \int_{-1}^{1} \sqrt{1 - u^2}\,du = 2R^2 \int_{-\pi/2}^{\pi/2} \cos^2\theta\,d\theta = \pi R^2$$

where we nondimensionalized the integral by letting $u = x/R$, and then did the trigonometric substitution $u = \sin\theta$. (To do the final integral trivially, notice that the average value of $\cos^2\theta$ along any of its periods is $1/2$.)

We can also use polar coordinates. We break the circle into regions bounded by radii r and $r + dr$, and angles θ and $\theta + d\theta$. These regions are rectangular, with side lengths of dr and

$r\,d\theta$, so the area element is $dA = r\,dr\,d\theta$. Then we have

$$A = \int_0^R r\,dr \int_0^{2\pi} d\theta = 2\pi \int_0^R r\,dr = \pi R^2$$

which is quite a bit easier. In fact, it's so much easier that we didn't even need to use double integrals at all. We could have decomposed the circle into a bunch of thin circular shells, argued that each shell contributed area $(2\pi r)\,dr$, then integrated over them,

$$A = \int_0^R 2\pi r\,dr = \pi R^2.$$

In Olympiad physics, there's usually a method like this, that allows you to get the answer without explicitly writing down any multiple integrals.

Example 17

Calculate the moment of inertia of the circle above, about the y axis, if it has total mass M and uniform density.

Solution

The moment of inertia of a small piece of the circle is

$$dI = x^2\,dm = x^2 \sigma\,dA = \frac{x^2 M}{\pi R^2}\,dA$$

where x^2 appears because x is the distance to the rotation axis, and σ is the mass density per unit area. Using Cartesian coordinates, we have

$$I = \frac{M}{\pi R^2} \int_{-R}^{R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} x^2\,dy.$$

The inner integral is still trivial; the x^2 doesn't change anything, because from the perspective of the dy integral, x is just some constant. However, the remaining integral becomes a bit nasty. In general, when this happens, we can try flipping the order of integration, giving

$$I = \frac{M}{\pi R^2} \int_{-R}^{R} dy \int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} x^2\,dx.$$

Unfortunately, this is equally difficult. Both of these integrals can be done with trigonometric substitutions, as you'll check below, but there's also a clever symmetry argument.

Notice that I is also equal to the moment of inertia about the x axis, by symmetry. So if we add them together, we get

$$2I = \int (x^2 + y^2)\,dm = \int r^2\,dm.$$

The r^2 factor has no dependence on θ at all, so the angular integral in polar coordinates is trivial. We end up with

$$2I = \frac{M}{\pi R^2} \int_0^R 2\pi r\, r^2\, dr = \frac{1}{2} MR^2$$

which gives an answer of $I = MR^2/4$, as expected.

[2] **Problem 32.** Calculate I in the previous example by explicitly performing either Cartesian integral.

[3] **Problem 33.** In this problem we'll generalize some of the ideas above to three dimensions, where we need triple integrals. Consider a ball of radius R.

(a) In Cartesian coordinates, the volume element is $dV = dx\, dy\, dz$. Set up an appropriate triple integral for the volume.

(b) The inner two integrals might look a bit nasty, but we already have essentially done them. Using the result we already know, perform the inner two integrals in a single step, and then perform the remaining integral to derive the volume of a sphere.

(c) In cylindrical coordinates, the volume element is $dV = r\, dr\, d\theta\, dz$. Set up a triple integral for the volume, and perform it. (Hint: this can either be hard, or a trivial extension of part (b), depending on what order of integration you choose.)

(d) In spherical coordinates, the volume element is $dV = r^2\, dr\, \sin\phi\, d\phi\, d\theta$. Set up a triple integral for the volume, and perform it.

(e) Let the ball have uniform density and total mass M. Compute its moment of inertia about the z-axis. (Hint: this can be reduced to a single integral if you use an appropriate trick.)

[2] **Problem 34.** Consider a spherical cap that is formed by slicing a sphere of radius R by a plane, so that the altitude from the vertex to the base is h. Find the area of its curved surface using an appropriate integral.

Remark

You might be wondering how good you have to be at integration to do Olympiad physics. The answer is: not at all! You need to understand how to set up integrals, but you almost never have to *perform* a nontrivial integral. There will almost always be a way to solve the problem without doing explicit integration at all, or an approximation you can do to render the integral trivial, or the integral will be given to you in the problem statement. The Asian Physics Olympiad takes this really far: despite having some of the hardest problems ever written, they often provide information like "$\int x^n\, dx = x^{n+1}/(n+1) + C$" as a hint! This is because physics competitions are generally written to make students think hard about physical systems, and the integrals are just viewed as baggage.

In fact, plain old AP Calculus probably has harder integrals than Olympiad physics. For example, in those classes everybody has to learn the integral

$$\int \sec x \, dx = \log |\sec x + \tan x| + C$$

which has a long history. When I was in high school, I was shocked by how the trick for doing this integral came out of nowhere; it seemed miles harder than anything else taught in the class. And it is! Historically, it arose in 1569 from Mercator's projection, where it gives the vertical distance on the map from the equator to a given latitude. For decades, cartographers simply looked up the numeric value of the integral in tables, where the Riemann sums had been done by hand. (They had no chance of solving it analytically anyway, since Napier only invented logarithms in 1614.) Gradually, tabulated values of the logarithms of trigonometric functions became available, and in 1645, Bond conjectured the correct result by noticing the close agreement of tabulated values of each side of the equation. Finally, Gregory proved the result in 1668, using what Halley called "a long train of Consequences and Complications of Proportions." So it took almost a hundred years for this integral to be sorted out! (Though to their credit, they had the handicap of not knowing about differentiation or the fundamental theorem of calculus; they were finding the area under the curve with just Euclidean geometry.)

Even though Olympiad physics tries to avoid tough integrals, doing more advanced physics tends to produce them, so physicists often get quite good at integration. By contrast, Spivak's calculus textbook for math majors only covers integration techniques in a single chapter towards the end of the book. He justifies the inclusion of this material by saying:

> Every once in a while you might actually need to evaluate an integral [...] For example, you might take a physics course [...] Even if you intend to forget how to integrate (and you probably will forget some details the first time through), you must never forget the basic methods.

That attitude is why physics students frequently win the MIT Integration Bee.