Mechanics
(draft, 2013)

by James Nearing

Physics Department

University of Miami

jnearing@miami.edu

Contents

Bibliography

.

Analytical Mechanics by Grant R. Fowles. Brooks-Cole.

Analytical Mechanics by Fowles and Cassiday. A second author was added. I prefer the original.

Mechanics by Symon. Addison-Wesley   The same subject as this text, at about the same level. It's been in print for almost 40 years, so it's got to be pretty good.

Introduction to Classical Mechanics by Arya. Allyn and Bacon   I think the recent edition is quite good.

Special Relativity by A.P. French. MIT Press   I think this remains the best introduction to the subject.

Mathematical Methods for Physics and Engineering by Riley, Hobson, and Bence. Cambridge University Press   For the quantity of well-written material here, it is surprisingly inexpensive in paperback.

Mathematical Methods in the Physical Sciences by Boas. John Wiley Pub   About the right level and with a very useful selection of topics. If you know everything in here, you'll find all your upper level courses much easier.

Street-Fighting Mathematics by Mahajan. MIT Press   "In problem solving, as in street fighting, rules are for fools: do whatever works—don't just stand there!" As practical as you can get.

Mathematical Tools for Physics by Nearing. Dover Pub or an online version   In my unbiased opinion, it's pretty good.

Schaum's Outlines by various.   There are many good and inexpensive books in this series: for example, "Complex Variables", "Advanced Calculus", "Bookkeeping and Accounting", "Advanced Mathematics for Engineers and Scientists". Amazon lists hundreds.

Linear Differential Operators by Lanczos. It exploits the close relationship between differential equations and matrices to gain deep insights. *Read the user comments on Amazon.*

A Brief on Tensor Analysis by James Simmonds. Springer   This is the only text on tensors that I will recommend. To anyone. Under any circumstances.

Linear Algebra Done Right by Axler. Springer   Don't let the title turn you away. It's pretty good.

Linear Algebra Done Wrong by Treil. an online text as a pdf.   Still another view of the subject, and maybe even better. (Views differ.)

Mathematical Prependix

.

   This preliminary chapter could as easily be an appendix to the text, but I prefer to put it here. It is a collection of topics that you will need at many places later on, but that you don't have to study in detail now. In each section I will try to indicate where the material is used, and when you get to the chapter where you need it I will indicate the reference here. You should at least skim this material now, so that you will have seen where it is. If something catches your eye and you want to study it now, don't let me stop you.

0.1 Series

This section is used in some form in every chapter in the text.

Infinite series is a tool that you see in an introductory calculus course, and you may not then have realized just how useful it is. Especially power series. There are a few series that show up so often that you need to have them instantly available. Binomial, trigonometric, exponential, geometric, occasionally the logarithm.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \cdots \qquad = \sum_0^\infty \frac{x^k}{k!} \tag{a}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} \cdots \qquad = \sum_0^\infty (-1)^k \frac{x^{2k+1}}{(2k+1)!} \tag{b}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \cdots \qquad = \sum_0^\infty (-1)^k \frac{x^{2k}}{(2k)!} \tag{c}$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \qquad = \sum_1^\infty (-1)^{k+1} \frac{x^k}{k} \qquad (|x| < 1) \tag{0.1}$$

$$(1+x)^n = 1 + nx + \frac{n(n-1)x^2}{2!} + \cdots \qquad = \sum_{k=0}^\infty \frac{n(n-1)\cdots(n-k+1)}{k!} x^k \quad (|x| < 1) \tag{e}$$

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \qquad = \sum_0^\infty \frac{x^{2k+1}}{(2k+1)!} \tag{f}$$

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \qquad = \sum_0^\infty \frac{x^{2k}}{(2k)!} \tag{g}$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots \qquad = \sum_0^\infty x^k \qquad (|x| < 1) \tag{h}$$

   The hyperbolic functions, $\sinh$ and $\cosh$ have power series similar to those for $\sin$ and $\cos$ except that the signs are all positive (f,g) instead of alternating (b,c). Differentiate the power series for sine and cosine to get the familiar differentiation formulas. Now do the same thing for the hyperbolic sine and cosine. Differentiate the series for the logarithm (d), and relate it to the last series (h) on the list,

called the "geometric series". This geometric series is a special case of the binomial series for $n = -1$.

▶ What is $\sqrt{2}$? Use the 5th series, the binomial expansion: $(1+1)^{1/2} = 1 + \frac{1}{2} = 1.5$. Not bad for two terms, even though it is at the edge of its domain of validity.

▶ What is $1/0.9$? Use the last series: $1/(1 - 0.1) = 1.1$.

▶ What is $0.99^{1/10}$? Use the binomial series again: $(1 - 0.01)^{1/10} = 1 - .001 = 0.999$.

▶ Evaluate this limit. This example pulls together several techniques with series; it is very worth your time to be able to reproduce this example on your own.

$$\lim_{x \to 0} \left[\frac{1}{1 - \sqrt{1-x}} - \frac{2}{x} \right] \tag{0.2}$$

Use the binomial expansion a few times. Work first on the complicated fraction:

$$\frac{1}{1 - \sqrt{1-x}} = \frac{1}{1 - \left[1 - \frac{1}{2}x - \frac{1}{8}x^2 - \cdots\right]} = \frac{1}{\frac{1}{2}x + \frac{1}{8}x^2 + \cdots}$$
$$= \frac{1}{\frac{1}{2}x\left[1 + \left(\frac{1}{4}x + \cdots\right)\right]} = \frac{2}{x}\left[1 - \left(\frac{1}{4}x + \cdots\right) + \left(\ \right)^2 \cdots\right]$$

Put this into the original expression

$$\left[\frac{2}{x} - \frac{1}{2} - \frac{2}{x} \right] \longrightarrow -\frac{1}{2}$$

Test this limit by a numerical experiment on a calculator. Take $x = 0.5, \ 0.1, \ 0.01$

▶ What is $(\sin 0.1)/(\sinh 0.1)$? Use (b), (f), and the geometric (h) series:
$$(0.1 - 0.001/6)/(0.1 + 0.001/6) = (1 - .01/6)/(1 + .01/6)$$
$$= (1 - .01/6)(1 - .01/6) = (1 - .02/6) = 0.996667.$$
A calculator gives 0.99667221535

▶ What is the behavior of the function $x(t) = \sqrt{c^2 t^2 + c^4/a^2} - c^2/a$ for small time? Use the binomial expansion, but first you *must* arrange the square root as $\sqrt{1 + \text{small}}$. This simply involves factoring the larger term out of the square root.

$$x(t) = \frac{c^2}{a}\sqrt{1 + \frac{a^2 t^2}{c^2}} - \frac{c^2}{a} \approx \frac{c^2}{a}\left[1 + \frac{a^2 t^2}{2c^2}\right] - \frac{c^2}{a} = \frac{1}{2}at^2 \tag{0.3}$$

This expression for $x(t)$ is the relativistic expression for motion with constant (proper) acceleration, and $at^2/2$ is the non-relativistic approximation to it. It is derived in section 9.6.

▶ Suddenly apply a force to a mass that is attached to a spring. The result for x is (see Eq. (3.33))

$$x(t) = \frac{F_0}{k}\left[1 - \cos\omega_0 t\right] \quad \text{where} \quad k \text{ is the spring constant, and} \quad \omega_0 = \sqrt{\frac{k}{m}}$$

What is the behavior of x for small time? And remember that *small* is not *zero*.

$$x(t) = \frac{F_0}{k}\left[1 - (1 - \omega_0 t^2/2 + \cdots)\right] \approx \frac{F_0}{k}\left[\omega_0^2 t^2/2\right] = \frac{F_0}{k}\frac{k}{m}\frac{t^2}{2} = \frac{F_0}{m}\frac{t^2}{2}$$

and that is $at^2/2$. At the start of the motion the spring hasn't yet stretched, so the only force is the one that you apply.

All of these expansions are special cases of the Taylor series.

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + \frac{1}{3!}(x - x_0)^3 f'''(x_0) + \cdots \qquad (0.4)$$

Where does this representation come from? If you assume that there is an expansion of the form

$$f(x) = A + B(x - x_0) + C(x - x_0)^2 + D(x - x_0)^3 + \cdots \qquad (0.5)$$

then evaluate both sides at $x = x_0$ and you immediately have $A = f(x_0)$. Now differentiate the hypothesized equation (0.5) for f

$$f'(x) = B + 2C(x - x_0) + 3D(x - x_0)^2 + \cdots$$

and again evaluate it at $x = x_0$. This gives $B = f'(x_0)$. Another derivative and an evaluation and you have $2C = f''(x_0)$. And again, and again, ...

This is where all the coefficients in Eq. (0.4) come from, and every one of the series in Eq. (0.1) can be derived this way (with $x_0 = 0$ in all cases).

The manipulations in the example of equation (0.3) are typical of the most common way that series are used in this text. When there is a complicated mathematical result for the solution to a problem, the most important step is to understand that result. Series approximations are a powerful tool to dig simple results out of complex mathematics. There are technical details to learn too, though that is not the point of this text: Under what conditions do these series converge? Under what conditions do they converge to the functions they supposedly represent?

Example
▶ What is the power series expansion of the function $f(x) = a - bx + cx^3$ about it minimum, out to second order terms? Take $a, b, c > 0$.
First where is the minimum?

$$f'(x) = -b + 2cx^2 = 0 \longrightarrow x_{\text{min}} = \pm\sqrt{b/2c}$$

Which if either of these two points *is* a minimum? For an answer you can either take a second derivative (do so) or you can sketch a graph. Close to the origin the $-bx$ term in f has it going down toward the right and up toward the left, then the cx^3 term eventually becomes very positive on the right and very negative on the left. That means the minimum is the one on the right and the one on the left is a maximum. Call it x_0.

$$x_0 = +\sqrt{b/2c}, \qquad f(x_0) = a - b\left(\frac{b}{2c}\right)^{1/2} + c\left(\frac{b}{2c}\right)^{3/2} = a - \frac{b^{3/2}}{2(2c)^{1/2}}$$

$$f'(x_0) = 0, \qquad f''(x_0) = 6cx_0 = 3\sqrt{2bc}$$

$$f(x) = f(x_0) + (x - x_0) \cdot 0 + \tfrac{1}{2}f''(x_0)(x - x_0)^2 + \cdots$$

$$= a - \frac{b^{3/2}}{2(2c)^{1/2}} + \tfrac{3}{2}\sqrt{2bc}(x - x_0)^2 + \cdots$$

0.2 Hyperbolic Functions

This section appears in chapters 3, 9, and 10.

The circular trigonometric functions such as sine and cosine are familiar, but the hyperbolic trigonometric functions may not be. These functions are defined in terms of exponentials as

$$\cosh x = \frac{e^x + e^{-x}}{2} \qquad \sinh x = \frac{e^x - e^{-x}}{2} \qquad \tanh x = \frac{\sinh x}{\cosh x} \tag{0.6}$$

Their reciprocals are sech, csch, coth in analogy with the definitions of the corresponding circular functions.

Why are these hyperbolic? First, why are the others circular? Answer: the sine and cosine satisfy a simple identity that allows them to describe a circle.

If $x = \cos\phi$ and $y = \sin\phi$,

then $x^2 + y^2 = \cos^2\phi + \sin^2\phi = 1$
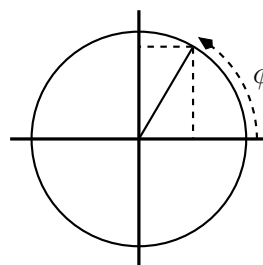
Fig. 0.1

There is a similar identity for hyperbolic functions, and its derivation involves nothing more than using the definitions.

$$\cosh^2\phi - \sinh^2\phi = \left(\frac{e^\phi + e^{-\phi}}{2}\right)^2 - \left(\frac{e^\phi - e^{-\phi}}{2}\right)^2 = \frac{e^{2\phi} + 2 + e^{-2\phi} - e^{2\phi} + 2 - e^{-2\phi}}{4} = 1$$

Divide this equation by $\cosh^2\phi$ to get $1 - \tanh^2\phi = \text{sech}^2\phi$.

If $x = \cosh\phi$ and $y = \sinh\phi$,

then $x^2 - y^2 = \cosh^2\phi - \sinh^2\phi = 1$
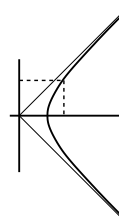
(0.7)

The coordinates x and y describe a hyperbola. In the circular case there is a simple geometric interpretation of ϕ. In the hyperbolic case there is not. It's not that there isn't an interpretation at all, just that it is not very useful.

What are the derivatives of these functions? Differentiate the equations (0.6) or the series in (0.1) for cosh and sinh, then the product (or quotient) rule for tanh.

$$\frac{d}{dx}\cosh x = \sinh x \qquad \frac{d}{dx}\sinh x = \cosh x \qquad \frac{d}{dx}\tanh x = \text{sech}^2 x \tag{0.8}$$

The hyperbolic functions involve exponentials, so it should not be too surprising that the inverse hyperbolic functions involve logarithms.

$$y = \sinh^{-1} x \qquad \text{means} \qquad x = \sinh y = \frac{1}{2}\left(e^y - e^{-y}\right) \tag{0.9}$$

Multiply by $2e^y$ and rearrange. The result is a quadratic equation.

$$e^{2y} - 2xe^y - 1 = \left(e^y\right)^2 - 2xe^y - 1 = 0 \qquad \text{which implies} \qquad e^y = x \pm \sqrt{x^2 + 1}$$

The exponential e^y is positive, so that forces the $+$ sign on the right. Now take the logarithm.

$$y = \sinh^{-1} x = \ln\left(x + \sqrt{x^2 + 1}\right)$$
$$\text{similarly,} \qquad \cosh^{-1} x = \ln\left(x \pm \sqrt{x^2 - 1}\right) \quad (x \geq 1)$$

(0.10)

Common circular trigonometric identities have their parallel here. For example

$$\cosh(x + y) = \cosh x \cosh y + \sinh x \sinh y, \qquad \sinh(x + y) = \sinh x \cosh y + \cosh x \sinh y \quad (0.11)$$

Use the series for the circular and the hyperbolic functions to see the relations between the two sets of functions. For example, what is $\cos(ix)$? Substitute ix into the third series of Eq. (0.1). Similarly, substitute ix into the series for the sine to see its relation to the hyperbolic sine.
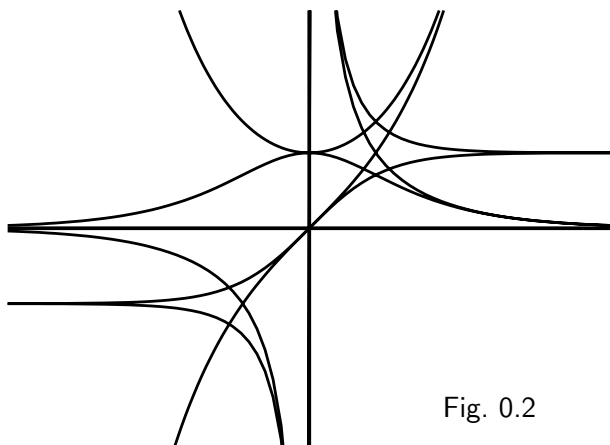
These are six graphs of the six hyperbolic functions. It's up to you to puzzle out which curves go with which functions.

To get the inverse functions, invert these graphs in the $45°$ line $x = y$.

Fig. 0.2

0.3 Coordinate Systems
This section appears in chapters 3, 4, 5, 6, 7, 8 (at least).
There are a few common coordinate systems that you will use all the time. In two dimensions there are rectangular and polar. In three dimensions the common ones are rectangular, cylindrical, and spherical. When you need parabolic coordinates or toroidal coordinates, you can look them up somewhere else, but the five drawn here are the ones you have to master.

rectangular

polar
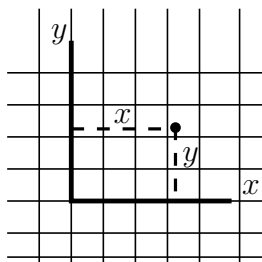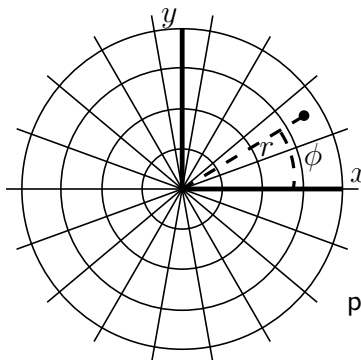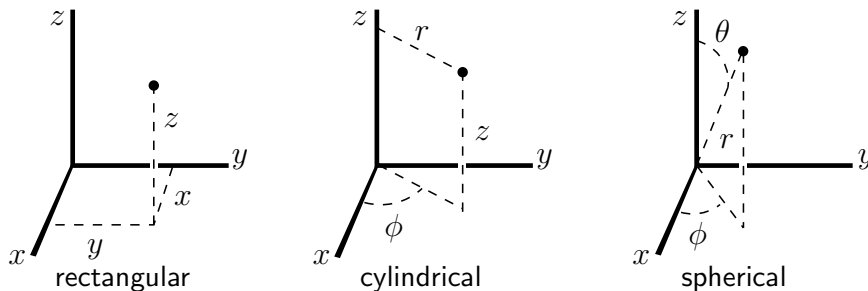
rectangular   cylindrical   spherical   Fig. 0.3

All the coordinate lines simply represent the functions that say all the *other* coordinates are constant. In the common plane rectangular coordinates the lines parallel to the x-axis are the graphs of the equations $y = 1$, or $y = -10$, *etc.* The equation $x = 5$ is a line parallel to the y-axis, and the equation for *that* line is $x = 0$.

In plane polar coordinates the coordinate lines are $r = $ constant (circles) or $\phi = $ constant (rays that start from the origin because $r \geq 0$).

The relation between these two coordinate systems is simple. When you describe a single point using (x, y) one time and (r, ϕ) another time, the equations relating these pairs are
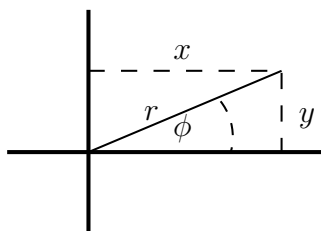
$$r = \sqrt{x^2 + y^2}, \quad \phi = \tan^{-1}(y/x)$$
$$x = r\cos\phi, \quad y = r\sin\phi$$

(0.12)

The only place to stumble in this transformation is in computing ϕ from y and x. You need both numbers y and x to specify the quadrant for ϕ because the inverse tangent is multiple valued. The signs of y and x will however tell you that $\tan^{-1}(1/1) = \pi/4$ and that $\tan^{-1}(-1/-1) = 3\pi/4$, thereby removing the ambiguity.

In three dimensions, rectangular coordinates (x, y, z) are much like those in two dimensions, except that an equation such as $y = 1$ is now a *plane* perpendicular to the y-axis as x and z vary from minus to plus infinity. To get a line requires two equations. For example the pair $\{x = 5$ and $z = 6\}$ describes a line parallel to the y-axis and puncturing the x-z plane at $(5, 6)$. The z-axis itself is described by the two equations $\{x = 0$ and $y = 0\}$.

Cylindrical coordinates (r, ϕ, z) are an extension of two-dimensional polar coordinates, simply stretched parallel to the z-axis. The z-coordinate is the same as the rectangular z-coordinate, and the (r, ϕ) are the same as the two dimensional polar coordinates. The equation $z = $ constant is a plane parallel to the x-y plane as before, and the equation $r = $ constant is now a cylinder centered along the z-axis. The third coordinate $\phi = $ constant is a half-plane with one edge along the z-axis ($0 < r < \infty$ and $-\infty < z < \infty$).
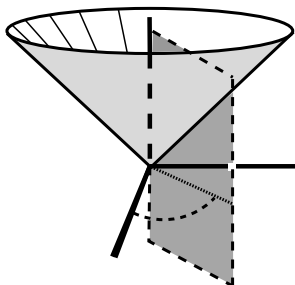
Spherical coordinates (r, θ, ϕ), use r to mean the distance to the origin, in contrast to cylindrical where it means the distance to the z-axis. The two angles are measured *from* the positive z-axis and *around* the z-axis respectively. The coordinate surfaces are now $r = $ constant (sphere), $\theta = $ constant (a cone with apex at the origin), and $\phi = $ constant (a half plane with one edge along the z-axis). In geography, the latitude and longitude define a point on the surface of the Earth. The latitude is like θ *except* that latitude is measured North and South from the equator (zero to $90°$ each), and θ is measured strictly South from the North Pole. Longitude is measured East and West from the Greenwich meridian (zero to $180°$ in each case), and ϕ is measured in one direction starting from the x-axis ($0 \le \phi \le 2\pi$). Just to keep you on your toes, some people prefer to use $-\pi \le \phi \le +\pi$.

$\rightarrow$ Watch out for varying conventions here. Commonly in math books the role of θ and ϕ in spherical coordinates are reversed, but the convention that I'm using is the standard in physics and engineering. What is not so conventional is that I choose is to make the angle ϕ the same for polar, cylindrical, and spherical coordinates. It is the angle around the z-axis and in the x-y plane. You do have to watch out for conventions used elsewhere because you will often find that θ is used for this angle in polar and cylindrical coordinates. I'm trying to be consistent here, using ϕ for the same angle in all three coordinate systems.

As with rectangular and polar coordinates, you can find the relationships between each pair of these coordinates. For rectangular and cylindrical, it is the same as with rectangular and polar, because z is the same for each. Just use the equations (0.12) again.

For the conversions between rectangular and spherical coordinates,

$$r = \sqrt{x^2 + y^2 + x^2}, \quad \theta = \cos^{-1}\frac{z}{r}, \quad \phi = \tan^{-1}(y/x)$$
$$z = r\cos\theta, \quad x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi \quad \leftarrow \tag{0.13}$$

The second line of this pair is the set of equations encountered most often.

For the conversions between cylindrical and spherical I'll leave it as an exercise, problem 0.18, but you see immediately that there is a problem in notation: what is r? I've used the same notation r in both systems, and that can be confusing. There is no standard way to do this. Some will do as I have done; others choose the Greek symbol ρ for the cylindrical radius; some will choose s; some will choose R, and there are probably other notations. I find that using r for both rarely causes confusion though, because it is usually clear from context which one is meant. On the rare occasions that I do need to make the distinction, I commonly use $r_\perp$ to indicate the perpendicular distance to the axis (the cylindrical r). That happens when both r's appear in the same equation.

In two dimensions you can draw the coordinate grid as a set of lines parallel to the x and y axes, and in three dimensions you would do the same thing, with lines parallel to the three axes. The only place to do this however is in your mind, because the drawing on paper will become so cluttered that you can't see what has happened.

In spherical coordinates the grid is formed of lines and arcs of circles: Hold the two coordinates $\theta = $ constant, $\phi = $ constant, and you have a radial line, the ray starting from the origin. For the next pair of coordinates, $r = $ constant, $\phi = $ constant is the intersection of a sphere and a half-plane that has one edge along the z-axis. It defines half of a great circle, and that is the θ coordinate curve. The third pair of coordinates are $r = $ constant, $\theta = $ constant, and they form the intersection of a sphere and a cone whose axis is along z. This defines the ϕ coordinate curve as a small circle parallel to the x-y plane.
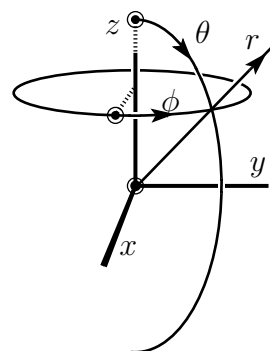
Fig. 0.4

This map of the Earth* shows the prime meridian at zero longitude, corresponding to the zero point for ϕ. Similarly, the lines of constant latitude show the relation: latitude $= |90° - \theta|$ (N or S).
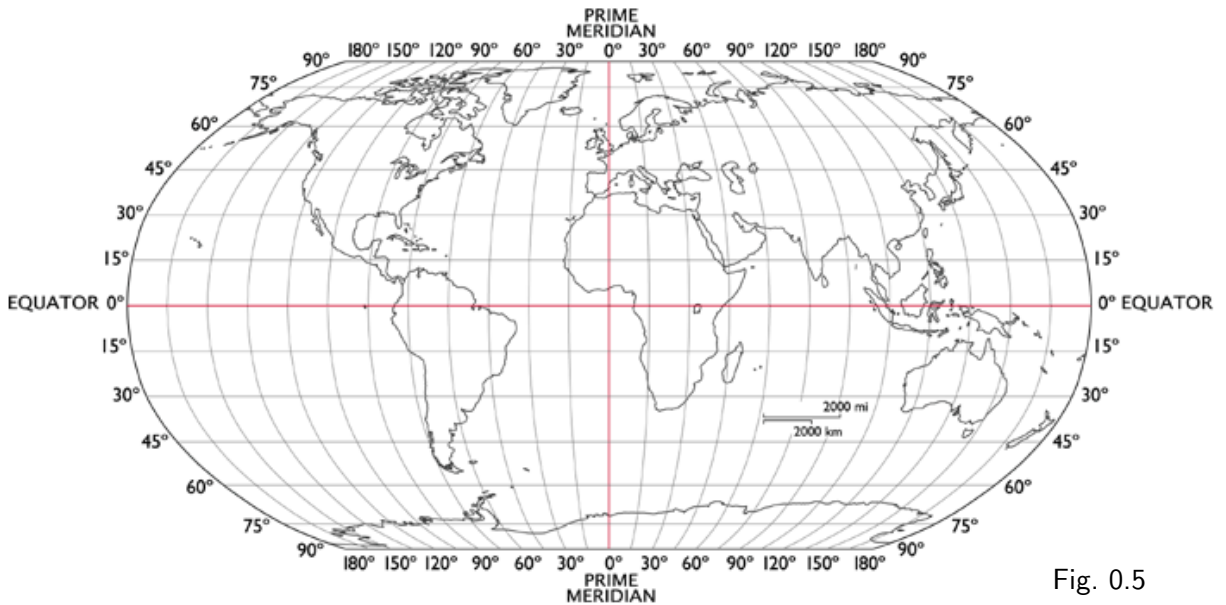
Fig. 0.5

The equations for simple curves can look very different in the various coordinate systems. In rectangular and in plane polar coordinates the equations for straight lines are respectively

$$y = mx + b \qquad \text{and} \qquad r = a\sec(\theta - \theta_0) \tag{0.14}$$

Equations for a circle can come in several forms

$$x^2 + y^2 = R^2, \qquad \text{or} \qquad r = R, \qquad \text{or} \qquad r = a\cos\phi,$$
$$\text{or} \qquad r^2 = a + br\cos(\phi - \phi_0) \tag{0.15}$$

An ellipse is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad \text{or } r = \frac{1}{A + B\cos\phi} \ \ (B < A), \quad \text{or} \ \begin{cases} x = a\cos\theta \\ y = b\sin\theta \end{cases}, \quad \text{or } r_1 + r_2 = 2a \tag{0.16}$$

Some of these are familiar. Probably some are not. The last one is in two-center bipolar coordinates — not one of your standards.

There are many examples of pictures of functions in rectangular and polar coordinates in The Famous Curves Index. In some cases, they have Java versions that allow you to play with the parameters. www-groups.dcs.st-and.ac.uk/˜history/Curves/Curves.html

Areas and Volumes

One of the traditional uses for integration is to computes areas and volumes. Even when that's not the aim, you still need to know how to set up such problems. When you compute moments of inertia or a center of mass it is a necessary tool.

For two dimensions the rectangular and polar coordinates are all you need, along with the two pictures.

---

* www.worldatlas.com
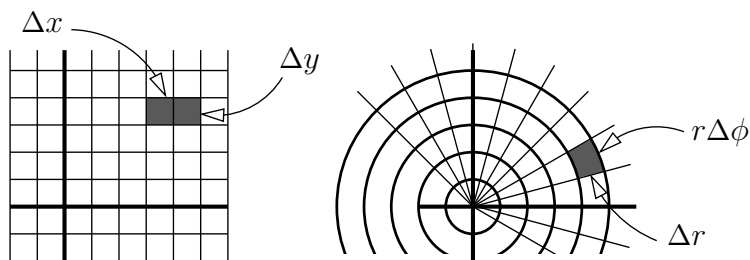
Fig. 0.6

The two shaded regions are rectangles defined respectively by the lines at x, $x + \Delta x$, y, and $y + \Delta y$ or between r and $r + \Delta r$, ϕ, $\phi + \Delta \phi$. Well, in the polar case only approximately rectangular, but in the limit that the two sides go to zero it's true. The area of a rectangle is the product of its sides, so the area elements are

$$\Delta A = \Delta x \Delta y \;\rightarrow\; dA = dx\,dy \qquad \text{and} \qquad \Delta A = \Delta r\, r \Delta \phi \;\rightarrow\; dA = r\,dr\,d\phi \qquad (0.17)$$

The area of a circle is

$$\int dA = \int_0^R r\,dr \int_0^{2\pi} d\phi = 2\pi \int_0^R r\,dr = \pi R^2$$

A comment on notation. You may be accustomed to seeing multiple integrals written differently, something like

$$\int_0^R \int_0^{2\pi} r\,d\phi\,dr$$

with the convention that you work from the inside out. There's nothing wrong with that notation, but I find it less confusing if I write the differential next to the integral sign and its associated limits. Multiplication is commutative and associative, so $r\,d\phi\,dr = d\phi\,r\,dr = dr\,r\,d\phi$. Then integrate right to left.

In three dimensions, the volume element in rectangular coordinates is just as it is in two dimensions, but with an extra factor of Δz. In cylindrical coordinates, it is the same as two-dimensional polar except for the same extra factor for the distance in the z-direction

$$dV = dx\,dy\,dz \qquad\qquad dV = r\,dr\,d\phi\,dz \qquad\qquad (0.18)$$

The spherical case requires a little more drawing.
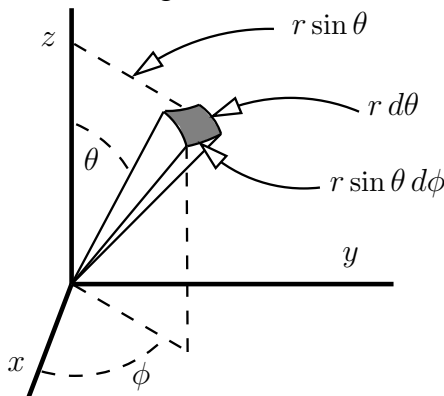
Fig. 0.7

Fix r for a moment and the piece of area on the surface of constant r is bounded by $(\theta, \theta + d\theta)$ and $(\phi, \phi + d\phi)$. Again, this forms a rectangle in the same way that dr and $r\,d\phi$ form a rectangle in plane polar coordinates even though all four sides are really arcs of circles. The sides are $r\,d\theta$ and

$r \sin \theta \, d\phi$. In the first case the radius of the circle is r, with center at the origin. In the second case the radius is $r \sin \theta$ with center *on the z-axis.*

$$dA = r^2 \sin \theta \, d\theta \, d\phi \qquad\qquad d\Omega = \sin \theta \, d\theta \, d\phi \qquad\qquad (0.19)$$

What about the volume? Just multiply this rectangle by dr to get the volume of the rectangular box with this area as a base. The combination $d\Omega$ shows up so much it has its own symbol, and Ω is called "solid angle".*

$$dV = r^2 \sin \theta \, dr \, d\theta \, d\phi$$
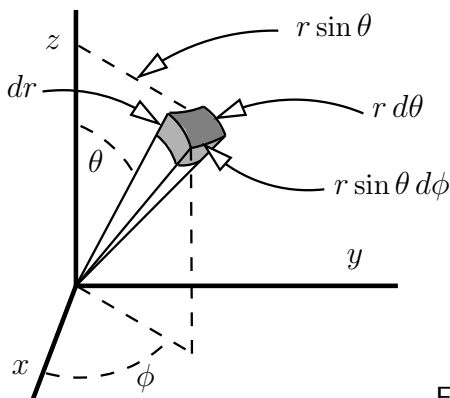
Fig. 0.8

The volume of a sphere is then

$$\int dV = \int_0^R r^2 \, dr \int_0^\pi \sin \theta \, d\theta \int_0^{2\pi} d\phi = 2\pi \int_0^R r^2 \, dr \int_0^\pi \sin \theta \, d\theta = 2\pi \int_0^R r^2 \, dr \, 2 = \frac{4}{3}\pi R^3$$

Limits of Integration

A very common problem when faced with multiple integrals is figuring out the limits of integration. What happens to those limits when you do the integrals in a different order. The key is

draw a picture of the domain of integration

and read the limits from the picture. Even in plane rectangular coordinates you will get frustrated without a picture to guide you.

The simplest example: what is the area of a right triangle, done as a double integral? Divide the area into little pieces of area, ΔA's, and add them together, finally taking a limit as each piece goes to zero. Now decide: in what systematic order are you going to add the pieces together? Certainly not randomly.
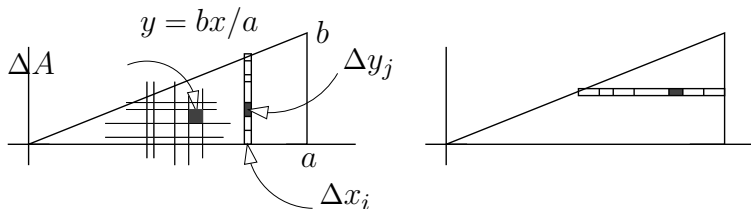
Fig. 0.9

If the typical piece of area is $\Delta A = \Delta x_i \Delta y_j$ then when you do the sum $\sum_{i,j} \Delta x_i \Delta y_j$ you must choose the order of summation. In the first picture the sum is

$$\Delta x_i \sum_j \Delta y_j \qquad \text{the } y\text{'s go from } 0 \text{ to } bx_i/a, \text{ so this is} \qquad \Delta x_i \cdot \frac{bx_i}{a}$$

---

* See Math Tools, section 8.13 for details about this concept.

The sum on Δx_i becomes an integral from 0 to a, $\int_0^a dx\, bx/a$. The sum in the other order and using the second picture starts with a sum on the Δx_i and you have to read those limits from the picture, starting at ay_j/b and ending at a.

$$A = \int_0^a dx \int_0^{bx/a} dy = \int_0^b dy \int_{ay/b}^a dx = \frac{ab}{2} \tag{0.20}$$

0.4 Vectors

This section appears in most chapters from four on.

Velocity, acceleration, momentum, angular momentum, electric field, magnetic field, gravitational field, force, torque, angular velocity. And these are just vectors that appear in *this* book.

The general definition of a vector is that it is an element of a vector space, but for here and now that is more generality than needed. The algebra of vectors in two or three dimensions starts from the pictures

$\vec{C} = \vec{A} + \vec{B}$

$\vec{D} = \vec{A} - \vec{B}$
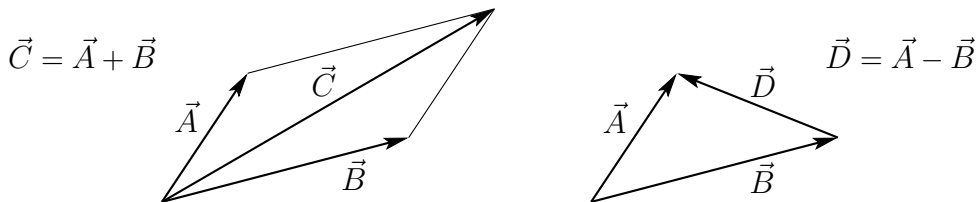
and these pictures apply whether you are talking about gravity or momentum or any of the other vectors listed above. There is a mathematical theorem guaranteeing this, saying that once you know you are working with three dimensional vectors, then they are all essentially the same. In mathematical jargon, *isomorphic*. The same for two dimensions (or seven). This means that you do not have to learn everything differently for different sorts of vectors — they all behave the same way.

Does a velocity vector look like a line with an arrowhead attached? Very few cars today have arrows sticking out of the front end, designed to skewer pedestrians,* but the import of this statement about vectors is that it doesn't matter. These pictures model velocity or magnetic fields or any other vector and you are guaranteed that they all give the same results. This is why you study the geometry of vectors as a subject of its own. There is no need to relearn it when you encounter the next sort of vector.
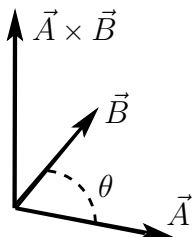
$$|\vec{A} \times \vec{B}| = AB \sin\theta$$
$$\vec{A} \cdot \vec{B} = AB \cos\theta \tag{0.21}$$

The vector product (cross product) of two vectors $\vec{A}$ and $\vec{B}$ is another vector perpendicular to these two and having magnitude $AB\sin\theta$, where θ is the indicated angle between them. The scalar product (dot product) is the scalar $AB\cos\theta$. Here I am using the convention that the magnitude of the vector $\vec{A}$ is the letter A. See the problems 0.25–0.33 for some review and practice.

The vector product as just stated is ambiguous because there are two directions that are perpendicular to the two input vectors, and they are in opposite directions. Which one to pick is chosen by convention: the right-hand rule.

---

* but find a picture of a 1950-era Oldsmobile!

Why are these the correct definitions for products of vectors? Why a cosine in one case and a sine in the other? Probably when the subject was being invented these weren't the first attempts made in defining products, but these are the ones that have good properties — especially the distributive law. $5 \cdot (3 + 4) = 5 \cdot 3 + 5 \cdot 4 = 35$, and for vectors

$$\vec{A} \cdot (\vec{B} + \vec{C}) = \vec{A} \cdot \vec{B} + \vec{A} \cdot \vec{C}, \qquad \vec{A} \times (\vec{B} + \vec{C}) = \vec{A} \times \vec{B} + \vec{A} \times \vec{C} \qquad (0.22)$$

What about the other sorts of multiplication laws: commutative and associative? $5 \cdot 4 = 4 \cdot 5$ and

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A}, \qquad \text{but} \qquad \vec{A} \times \vec{B} = -\vec{B} \times \vec{A}$$

so the commutative rule works for the dot product but not for the cross product. It anti-commutes instead. For the associative law, $5 \cdot (3 \cdot 4) = (5 \cdot 3) \cdot 4 = 60$, but

$$\vec{A} \cdot (\vec{B} \cdot \vec{C}) \quad \text{doesn't make any sense. The dot product is between } vectors.$$

For the cross product, $\vec{A} \times (\vec{B} \times \vec{C})$ makes sense, but it is not associative, and you can see that easily by writing a simple example.

$$\hat{x} \times (\hat{x} \times \hat{y}) = \hat{x} \times \hat{z} = -\hat{y}, \qquad \text{but} \qquad (\hat{x} \times \hat{x}) \times \hat{y} = 0 \times \hat{y} = 0$$

The closest that the cross product comes to associativity is the Jacobi identity,

$$\vec{A} \times (\vec{B} \times \vec{C}) + \vec{B} \times (\vec{C} \times \vec{A}) + \vec{C} \times (\vec{A} \times \vec{B}) = 0 \qquad (0.23)$$

It doesn't show up very often in ordinary vector algebra, but it appears near the end of chapter four.
A useful identity involving this triple cross product is

$$\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B}) \qquad (0.24)$$

(Do these last two equations work with the simple example $\hat{x} \times (\hat{x} \times \hat{y})$ of a few lines back?)
And the triple scalar product obeys the identities

$$\vec{A} \cdot (\vec{B} \times \vec{C}) = (\vec{A} \times \vec{B}) \cdot \vec{C} \qquad \qquad \text{interchange dot and cross}$$
$$\vec{A} \cdot (\vec{B} \times \vec{C}) = \vec{B} \cdot (\vec{C} \times \vec{A}) = \vec{C} \cdot (\vec{A} \times \vec{B}) \qquad \text{cyclic permutation of factors} \qquad (0.25)$$

Bases, Unit Vectors
The geometry of vectors can become cumbersome to manipulate, especially when there are many vectors involved. The technique of components is a way to turn much of the geometry into algebra and even arithmetic. It is to vectors what analytic geometry is to plane geometry. In rectangular coordinates the system should be familiar. In two dimensions pick two vectors, one parallel to the x-axis and one parallel to the y-axis. Call them $\hat{x}$ and $\hat{y}$, the pair forming a basis, and make them have magnitude one. Not one meter or one second, just one dimensionless.
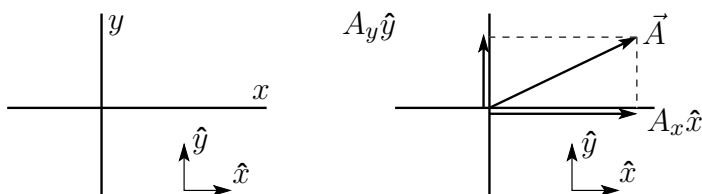
You can write any vector in the plane as the sum of two other vectors, one parallel to $\hat{x}$ and one parallel to $\hat{y}$. The former is a multiple of $\hat{x}$ (positive or negative) and the latter is a multiple of $\hat{y}$. The two components of the vector $\vec{A}$ are the two numbers A_x and A_y, with units as appropriate.

The point of writing a single vector in terms of two other vectors (three others in three dimensions) is to change geometry into algebra and arithmetic. It changes the problem of differentiating vectors into the more familiar problem of differentiating ordinary functions. The basis vectors you will encounter in this book are mostly orthogonal and normalized to one. That is a convenience not a necessity, and in other contexts you can make other choices to define a basis.

The dot product of two vectors is easy to compute in terms of these basis vectors.

$$\vec{A} \cdot \vec{B} = (A_x\hat{x} + A_y\hat{y} + A_z\hat{z}) \cdot (B_x\hat{x} + B_y\hat{y} + B_z\hat{z}) =$$
$$A_xB_x\hat{x} \cdot \hat{y} + A_xB_y\hat{x} \cdot \hat{y} + A_xB_z\hat{x} \cdot \hat{z} + A_yB_x\hat{y} \cdot \hat{y} + A_yB_y\hat{y} \cdot \hat{y}$$
$$+ A_yB_z\hat{y} \cdot \hat{z} + A_zB_x\hat{z} \cdot \hat{y} + A_zB_y\hat{z} \cdot \hat{y} + A_zB_z\hat{z} \cdot \hat{z}$$
$$= A_xB_x + A_yB_y + A_zB_z \tag{0.26}$$

The simplicity appears because the basis vectors are orthonormal. That is, each has unit magnitude and they are mutually perpendicular.

Does the product rule for derivatives apply to the product of vectors? Yes, and you can see why by differentiating the component form:

$$\frac{d}{dt}\vec{A} \cdot \vec{B} = \frac{dA_x}{dt}B_x + A_x\frac{dB_x}{dt} + \cdots = \frac{d\vec{A}}{dt} \cdot \vec{B} + \vec{A} \cdot \frac{d\vec{B}}{dt}$$

The same analysis applies to the cross product.

Index Notation

A convenient and powerful notation to manipulate the components of vectors: Denote the basis vectors by a unified notation, $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$ instead of respectively $\hat{x}$, $\hat{y}$, and $\hat{z}$. You then write the vectors as

$$\vec{A} = A_1\vec{e}_1 + A_2\vec{e}_2 + A_3\vec{e}_3 \qquad \text{and} \qquad \vec{B} = B_1\vec{e}_1 + B_2\vec{e}_2 + B_3\vec{e}_3$$

That these vectors, $\vec{e}_i$, are orthogonal and unit vectors translates to

$$\vec{e}_1 \cdot \vec{e}_2 = 0, \qquad \vec{e}_3 \cdot \vec{e}_3 = 1, \qquad \text{and all other such combinations}$$

Introduce a standard notation:

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \qquad \text{then} \qquad \vec{e}_i \cdot \vec{e}_j = \delta_{ij} \tag{0.27}$$

where the indices i and j take on any and all of the values 1, 2, 3. This is called the Kronecker delta symbol.

The dot product is exactly the same as in Eq. (0.26), only the final result just has subscripts $1, 2, 3$ instead of x, y, x. It doesn't seem to be worth all the trouble does it? It is. First a version of the calculation in this notation.

$$\vec{A} \cdot \vec{B} = \sum_i A_i\vec{e}_i \cdot \sum_j B_j\vec{e}_j = \sum_i \sum_j A_iB_j\vec{e}_i \cdot \vec{e}_j = \sum_i \sum_j A_iB_j\delta_{ij}$$

Now do the sum on j

$$\sum_j B_j \delta_{ij} = B_i \qquad \text{Remember: } \delta_{ij} = 0 \text{ for the two terms where } j \neq i$$

The rest is $\sum_i A_i B_i$, exactly as in Eq. (0.26).

   The **summation convention** was introduced by Einstein when it became clear to him that in manipulations with this notation for the components, a summation symbol $\left(e.g.\ \sum_i\right)$ *always* appeared whenever an index was repeated in a single term. So why bother to write it? The convention is that *whenever an index is repeated in a single term it is to be summed.* The immediately preceding calculation is then shortened to

$$\vec{A} \cdot \vec{B} = A_i \vec{e}_i \cdot B_j \vec{e}_j = A_i B_j\, \vec{e}_i \cdot \vec{e}_j = A_i B_j\, \delta_{ij} = A_i B_i \tag{0.28}$$

This convention also means that if you encounter an expression with *three* identical indices, such as $A_{jk} = B_i C_{ik} C_{ji}$ then go back and find your mistake. This can't happen.

   With this convention you have for example, $\delta_{ij} v_j = v_i$. (Write it out.)

   How do you handle cross products in this notation? You need to introduce another piece of notation.

$$\vec{e}_i \cdot \vec{e}_j \times \vec{e}_k = e_{ijk} \qquad \text{the alternating symbol} \tag{0.29}$$

This is ± 1 if the indices i, j, k are different from each other, *i.e.* a permutation of 123. It is $+1$ for 123 and any cyclic permutation of 123, *i.e.* 231 and 312. Interchange any two (*e.g.* 132) and you get -1. If any two indices are the same then this is zero. You can easily verify that e_{ijk} equals $\frac{1}{2}(i - j)(j - k)(k - i)$, though this identity is more of a curiosity than anything particularly useful.

$$\vec{A} \times \vec{B} = \vec{e}_i\, e_{ijk} A_j B_k, \qquad e_{ijk} e_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km} \tag{0.30}$$

These are easy to check. At worst just look at all the possible cases, and elegant derivations can come later. The second equation in Eq. (0.30) is equivalent to Eq. (0.24).

0.5 Differentiation
There's *no* chapter in this book in which this is not used.
So much of this subject depends on knowing what derivatives are and how to manipulate them that it's worth spending some space to review the subject (and maybe in the process to introduce some ideas you haven't seen).

   The standard definition of the derivative of a function of one variable is

$$\frac{df}{dx}(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{0.31}$$

The "prime" notation, $f'(x)$, for the derivative is useful too, but if that's the only notation you use it can hinder you to the point of incapacity. The Leibnitz notation of Eq. (0.31) lends itself to manipulation while the f' does not. For an immediate example, the most common method used to differentiate anything is the chain rule, and that is remarkably obscure in one notation while being quite intuitive in the other.

$$h(x) = f\big(g(x)\big) \ \longrightarrow\ h'(x) = f'(g)g' \qquad \text{or} \qquad \frac{dh}{dx} = \frac{df}{dg}\frac{dg}{dx} \tag{0.32}$$

How do you derive this? The definition of the derivative and the Leibnitz form dictate the manipulations to use. Start from the definition:

$$\frac{h(x+\Delta x)-h(x)}{\Delta x} = \frac{h(x+\Delta x)-h(x)}{g(x+\Delta x)-g(x)} \cdot \frac{g(x+\Delta x)-g(x)}{\Delta x} \qquad (0.33)$$

or in a shorter form,

$$\frac{\Delta h}{\Delta x} = \frac{\Delta h}{\Delta g} \cdot \frac{\Delta g}{\Delta x}$$

As $\Delta x \to 0$, the increment in g must approach zero also, as otherwise its derivative would not exist and the chain rule would not apply.* The second factor in Eq. (0.33) becomes, in the limit that $\Delta x \to 0$, the derivative dg/dx. Also, because in the same limit $\Delta g = g(x+\Delta x) - g(x) \to 0$, the first factor becomes df/dg, and that ends the derivation. A common way to write the composition of functions is to use, instead of $h(x) = f(g(x))$, the notation $h = f \circ g$.

A possibly unfamiliar rule occurs when you encounter the derivative of a function in which the independent variable appears in several unrelated places. For example,

$$f_1(x) = x^x, \qquad \text{or} \qquad f_2(x) = \int_0^x t\sqrt{x-t}\,dt \qquad (0.34)$$

Think of either of these as a function $f(x,x)$, where the first x is one of the two x's in x^x and the second x is the other. Now start from the definition and manipulate, adding and subtracting the same term.

$$\frac{f(x+\Delta x, x+\Delta x) - f(x,x)}{\Delta x} = \frac{f(x+\Delta x, x+\Delta x) - f(x, x+\Delta x) + f(x, x+\Delta x) - f(x,x)}{\Delta x}$$

$$= \frac{f(x+\Delta x, x+\Delta x) - f(x, x+\Delta x)}{\Delta x} + \frac{f(x, x+\Delta x) - f(x,x)}{\Delta x}$$

As $\Delta x \to 0$, the second term becomes the definition of the derivative of f with respect to the second x. In the same limit the first term becomes the derivative with respect to the first x. (This assumes that all the derivatives are continuous, so that as $\Delta x \to 0$ the fact that the first differentiation is approached from near x instead of *at* x will not matter.)

A notation common in some calculus texts can be useful here.

$$\frac{d}{dx} f(x,x) = D_1 f(x,x) + D_2 f(x,x) \qquad (0.35)$$

Here D_1 means differentiate with respect to the first argument and D_2 with respect to the second. This derivation shows that if x shows up in two different places, differentiate with respect to one of them and then with respect to the next one and add the two results. The familiar product rule is a special case of this.

$$\frac{d}{dx} f(x)g(x) = \frac{df}{dx} g + f \frac{dg}{dx} \qquad (0.36)$$

In the less familiar cases, Eq. (0.34),

$$\frac{df_1}{dx} = \frac{d}{dx} x^x = x\,x^{x-1} + \ln x\,x^x = x^x(1 + \ln x)$$

---

* What is the derivative of the function "$h(x) = x$" at the point $x = 0$? One of course. Now let $f(x) = 1/x$ and $g(x) = 1/x$, then it looks like $h = f \circ g$, but neither $f'(0)$ nor $g'(0)$ exist.

and

$$\frac{df_2}{dx} = 0 + \frac{1}{2} \int_0^x \frac{t}{\sqrt{x-t}} \, dt$$

Parametrized Differentiation
Once you see the explanation of this it seems remarkable simple, but if you encounter the phenomenon in an unfamiliar context you may not think of it.

If u is a function of time and v is a function of time, what is du/dv?

$$\text{At times } t \text{ and } t + \Delta t \quad \Delta u = u(t + \Delta t) - u(t) \quad \text{and} \quad \Delta v = v(t + \Delta t) - v(t)$$

$$\text{then} \quad \frac{\Delta u}{\Delta v} = \frac{\Delta u / \Delta t}{\Delta v / \Delta t}$$

$$\text{and the limit as } \Delta t \to 0 \text{ is} \quad \frac{du}{dv} = \frac{du/dt}{dv/dt} \qquad \text{also} \qquad = \frac{dt/dv}{dt/du} \tag{0.37}$$

Example
▶ In polar coordinates, $x = r \cos \phi$ and $y = r \sin \phi$. What is dy/dx as you go around a circle? Here r is constant, so

$$\frac{dy}{dx} = \frac{dy/d\phi}{dx/d\phi} = \frac{r \cos \phi}{-r \sin \phi} = -\cot \phi$$

It is easy to check this at angles such as $\phi = 0$ or $\pi/4$ etc. This technique will show up in a more complicated example in chapter nine, relativity, when computing relative velocity and acceleration.

0.6 Velocity, Acceleration
This section appears in most chapters from four on.
The definitions of velocity and acceleration are

$$\vec{v} = \frac{d\vec{r}}{dt} = \dot{\vec{r}} \qquad \text{and} \qquad \vec{a} = \frac{d\vec{v}}{dt} = \dot{\vec{v}}$$

This doesn't depend on your coordinate system, be it rectangular, polar, spherical, or oblate spheroidal. Computing them in different coordinate systems can get technical though. Fortunately three common cases (rectangular, polar, cylindrical) are easy. Spherical is more involved but still manageable, and the rest are rare enough that you can learn them when or if you need them. The "dot" notation over the letter is like the $'$ so common in calculus texts, but it has a particular meaning. It *always* means differentiation with respect to time.
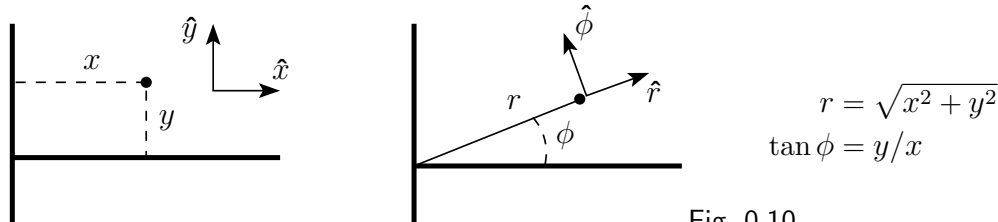

$$r = \sqrt{x^2 + y^2}$$
$$\tan \phi = y/x$$

Fig. 0.10

Rectangular is familiar

$$\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}, \qquad \text{so} \qquad \vec{v} = \dot{x}\hat{x} + \dot{y}\hat{y} + \dot{z}\hat{z} \qquad \text{and} \qquad \vec{a} = \ddot{x}\hat{x} + \ddot{y}\hat{y} + \ddot{z}\hat{z}$$

Read the components from these equations. *E.g.* $a_x = \ddot{x} = d^2x/dt^2$. For cylindrical coordinates the appropriate basis vectors conform to this system, with $\hat{r}$ pointing away from the origin and $\hat{\phi}$ perpendicular to that. $\vec{r} = z\hat{z} + r\hat{r}$, and plane polar simply omits the z. Now the unit vectors are functions of position, implying that as a particle moves these unit vectors will change, and you have to use the product rule to differentiate the terms. $\hat{z}$ is constant so it causes no trouble.

$$\vec{v} = \dot{z}\hat{z} + \dot{r}\hat{r} + r\dot{\hat{r}}$$

The one new feature is the third term, and for that you need to notice that $\hat{r}$ is a function of the coordinate ϕ, though not of z or r. To evaluate this derivative, use the chain rule.

$$\frac{d\hat{r}}{dt} = \frac{d\hat{r}}{d\phi}\frac{d\phi}{dt}$$

The first of these derivatives, $d\hat{r}/d\phi$, is now a problem in geometry, and there's a result about differentiating vectors that have a constant magnitude: the derivative is perpendicular to the original vector. To show this, let $\vec{u}$ be *any* vector of constant magnitude. That is, $\vec{u} \cdot \vec{u} = C$. Differentiate this with respect to *anything*.

$$\frac{d}{dt}\vec{u} \cdot \vec{u} = \frac{dC}{dt} = 0 = 2\vec{u} \cdot \frac{d\vec{u}}{dt} \tag{0.38}$$

That's all that's necessary, because it says that the derivative is either zero or is perpendicular to $\vec{u}$ as claimed.

$d\hat{r}/d\phi$ is perpendicular to $\hat{r}$. It is in the $\hat{\phi}$ direction. Now, what is its magnitude? A sketch answers the question. The sketch will also answer the question: "Why is it in the $+\hat{\phi}$ direction and not along $-\hat{\phi}$?"
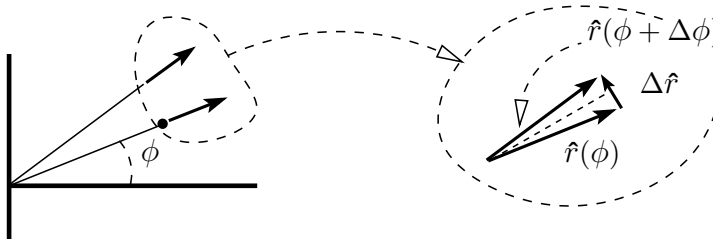
Fig. 0.11

The three vectors $\hat{r}(\phi)$, $\hat{r}(\phi + \Delta\phi)$, and $\Delta\hat{r}$ form an isosceles triangle. Construct the bisector of the vertex angle, and you immediately see that the length of $\Delta\hat{r}$ is

$$\left|\Delta\hat{r}\right| = 2 \cdot 1 \cdot \sin(\Delta\phi/2)$$

As $\Delta\phi \to 0$, the sine behaves as $\Delta\phi/2$ itself, so the quotient $\left|\Delta\hat{r}\right|/\Delta\phi \to 1$.

$$\frac{d\hat{r}}{d\phi} = \hat{\phi}, \qquad \text{and similarly} \qquad \frac{d\hat{\phi}}{d\phi} = -\hat{r} \tag{0.39}$$

Now back to velocity and acceleration.

$$\vec{v} = d\vec{r}/dt = \dot{z}\hat{z} + \dot{r}\hat{r} + r\dot{\hat{r}} = \dot{z}\hat{z} + \dot{r}\hat{r} + r\dot{\phi}\hat{\phi} \tag{0.40}$$

Another derivative:

$$\vec{a} = d\vec{v}/dt = \ddot{z}\hat{z} + \ddot{r}\hat{r} + \dot{r}\dot{\hat{r}} + \dot{r}\dot{\phi}\hat{\phi} + r\ddot{\phi}\hat{\phi} + r\dot{\phi}\dot{\hat{\phi}}$$

$$= \ddot{z}\hat{z} + \ddot{r}\hat{r} + \dot{r}\frac{d\hat{r}}{d\phi}\dot{\phi} + \dot{r}\dot{\phi}\hat{\phi} + r\ddot{\phi}\hat{\phi} + r\dot{\phi}\frac{d\hat{\phi}}{d\phi}\dot{\phi}$$

$$= \ddot{z}\hat{z} + \hat{r}(\ddot{r} - r\dot{\phi}^2) + \hat{\phi}(r\ddot{\phi} + 2\dot{r}\dot{\phi}) \tag{0.41}$$

Here you need $d\hat{\phi}/d\phi$, another derivative of a unit vector, so it is perpendicular to $\hat{\phi}$. How big is it? You can do the same sort of geometry as with $\hat{r}$, or notice that

$$\hat{r} \cdot \hat{\phi} = 0 \quad \longrightarrow \quad \frac{d}{d\phi}\hat{r} \cdot \hat{\phi} = 0 = \frac{d\hat{r}}{d\phi} \cdot \hat{\phi} + \hat{r} \cdot \frac{d\hat{\phi}}{d\phi} = 1 + \hat{r} \cdot \frac{d\hat{\phi}}{d\phi}$$

and this establishes the sign and magnitude of $d\hat{\phi}/d\phi$ as in Eq. (0.39). Increasing the angle ϕ a little bit rotates an object a little bit counterclockwise. That rotates $\hat{r}$ toward $\hat{\phi}$. It rotates $\hat{\phi}$ toward $-\hat{r}$.

That the basis vectors vary with position and are not parallel to each other as you move around the coordinate system is a familiar idea in another context: geography. $\hat{\text{Up}}$, $\hat{\text{South}}$, and $\hat{\text{East}}$ change direction as you move around the Earth, though they always remain orthogonal to each other.

For spherical coordinates the derivations can be done along the same lines as for cylindrical, but with a lot more algebra. It is not worth the trouble to go through it, and you don't need the results as often. The answer is (remember to distinguish the spherical coordinate r from the cylindrical one — they're spelled the same).

$$\begin{aligned}
\vec{r} &= r\hat{r} \\
\vec{v} &= \dot{\vec{r}} = \dot{r}\hat{r} + r\dot{\theta}\hat{\theta} + r\dot{\phi}\sin\theta\,\hat{\phi} \\
\vec{a} &= \left(\ddot{r} - r\dot{\phi}^2\sin^2\theta - r\dot{\theta}^2\right)\hat{r} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\dot{\phi}^2\sin\theta\cos\theta\right)\hat{\theta} \\
&\qquad + \left(r\ddot{\phi}\sin\theta + 2\dot{r}\dot{\phi}\sin\theta + 2r\dot{\theta}\dot{\phi}\cos\theta\right)\hat{\phi}
\end{aligned} \qquad (0.42)$$

In geographical terms,

$$\hat{r} \equiv \widehat{\text{Up}} \qquad \hat{\theta} \equiv \widehat{\text{South}} \qquad \hat{\phi} \equiv \widehat{\text{East}}$$

Example

▶ Circular motion is a familiar example from introductory courses. If $z = 0$ and $r =$ a constant, the equations (0.40) and (0.41) are

$$\vec{v} = r\dot{\phi}\hat{\phi} \qquad \text{and} \qquad \vec{a} = -\hat{r}\,r\dot{\phi}^2 + \hat{\phi}\,r\ddot{\phi} = -\hat{r}\frac{v^2}{r} + \hat{\phi}\frac{dv}{dt} \qquad (0.43)$$

The last form comes from using the magnitude of the first of these equations for $\vec{v}$, that is $v = r\dot{\phi}$, and it reproduces the familiar inward radial acceleration for circular motion ($r\dot{\phi}^2 = v^2/r$). The tangential component is $r\ddot{\phi} = d(r\dot{\phi})/dt = dv/dt$. *Now wait a minute!* If you believe this manipulation, look again more critically. Is the motion counterclockwise or clockwise and does it matter? Is $\dot{\phi}$ positive or negative? Is the magnitude of a vector positive? When you express a vector in terms of components, is the coefficient of the unit vector the magnitude of the vector or a component of the vector? Answer: the correct equation is not $v = r\dot{\phi}$, but $v_\phi = r\dot{\phi}$, stating that the phi-component of the velocity is $r\dot{\phi}$. Go back and modify these equations appropriately.

Example

▶ $x = x_0$, a constant, $y = v_0 t$ should have constant velocity and zero acceleration, but that's not so obvious if you see it in polar coordinates.

$$r = \sqrt{x^2 + y^2} = \sqrt{x_0^2 + v_0^2 t^2} \qquad \text{and} \qquad \phi = \tan^{-1}(y/x) = \tan^{-1}\frac{v_0 t}{x_0}$$
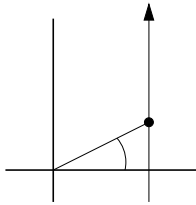
$$\vec{v} = \hat{r}\,\dot{r} + \hat{\phi}\,r\dot{\phi}$$

$$= \hat{r}\frac{v_0^2 t}{\sqrt{x_0^2 + v_0^2 t^2}} + \hat{\phi}\sqrt{x_0^2 + v_0^2 t^2}\,\frac{1}{1 + (v_0 t/x_0)^2} \cdot \frac{v_0}{x_0} \qquad (0.44)$$

$$= \hat{r}\frac{v_0^2 t}{\sqrt{x_0^2 + v_0^2 t^2}} + \hat{\phi}\frac{v_0 x_0}{\sqrt{x_0^2 + v_0^2 t^2}}$$

You can see that this has the correct behavior at $t = 0$ and as $t \to \infty$. Does it have the correct magnitude? And what is its derivative?

0.7 Complex Algebra

This section appears in chapters 3, 4, 6, and implicitly many other places.

There are some standard manipulations with complex arithmetic that take some practice. Even the basic $+$, $-$, $\times$, and $\div$ are not exactly what you learned in third grade, so I'll start with those. The standard commutative, associative, and distributive laws apply to the first three, so

$$(7 + 2i)(6 + 3i) = (6 + 3i)(7 + 2i) = 36 + 33i$$

$$(1 + 2i)\big[(3 + 4i)(5 + 6i)\big] = \big[(1 + 2i)(3 + 4i)\big](5 + 6i) = (1 + 2i)(-9 + 38i) = -85 + 20i$$

$$(1 + 2i)\big[(3 + 4i) + (5 + 6i)\big] = (1 + 2i)(3 + 4i) + (1 + 2i)(5 + 6i) =$$

$$(1 + 2i)(8 + 10i) = -12 + 26i$$

As for division, it is no more commutative here than it is for real numbers, but a simple trick allows you to simplify some expressions. The complex conjugate of a number is the number found by changing the sign of the imaginary part.

$$z = 5 + 7i \quad \longrightarrow \quad z^* = 5 - 7i$$

The * notation is a common one for this operation, though $\bar{z}$ is another notation that many prefer. What is the product of a number and its complex conjugate?

$$z = 5 + 7i, \quad z^* = 5 - 7i \quad \longrightarrow \quad z^*z = (5 - 7i)(5 + 7i) = 25 + 49 + 35i - 35i = 74$$

z^*z is always real and positive: $(a + ib)(a - ib) = a^2 + b^2$ is the square of the magnitude of the complex number, the square of $\sqrt{a^2 + b^2}$. How do you use this to manipulate division? Rationalize the denominator of a quotient.

$$\frac{1 + 2i}{3 + 4i} = \frac{(1 + 2i)(3 - 4i)}{(3 + 4i)(3 - 4i)} = \frac{11 + 2i}{25} \qquad (0.45)$$

Multiplying a number by its complex conjugate results in a real, so you can multiply the numerator and denominator of a quotient by the complex conjugate of the denominator and bring the result into a simpler form. If you ever want to make the numerator real instead, use the same idea.

Example

► A few cases of such manipulation, simplifying complex expressions:

$$\frac{3 - 4i}{2 - i} = \frac{(3 - 4i)(2 + i)}{(2 - i)(2 + i)} = \frac{10 - 5i}{5} = 2 - i.$$

$$(3i + 1)^2\left[\frac{1}{2 - i} + \frac{3i}{2 + i}\right] = (-8 + 6i)\left[\frac{(2 + i) + 3i(2 - i)}{(2 - i)(2 + i)}\right]$$

$$= (-8 + 6i)\frac{5 + 7i}{5} = \frac{2 - 26i}{5}.$$

$$\frac{i^3 + i^{10} + i}{i^2 + i^{137} + 1} = \frac{(-i) + (-1) + i}{(-1) + (i) + (1)} = \frac{-1}{i} = i.$$

What is the geometric interpretation of i? It is a factor it rotates by $90°$.
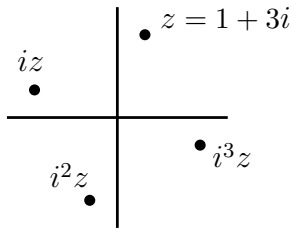
$$iz = i(1 + 3i) = -3 + i$$
$$i^2 z = i(-3 + i) = -1 - 3i$$
$$i^3 z = i(-1 - 3i) = 3 - i$$
$$i^4 z = z$$

What is i^n? Each multiplication by i rotates by $90°$ in the complex plane, so $i^4 = 1$, and $i^{217} = i^{4 \cdot 54 + 1} = i$.

Various roots of 1 or of -1 or of i appear commonly, and you need the exponential representation, Euler's formula, to find them. This is
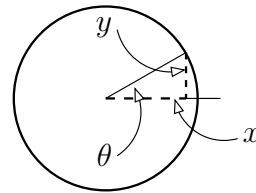
$$x + iy = r \cos\theta + ir \sin\theta = re^{i\theta} \qquad (0.46)$$

You can derive this equation from the series (0.1). Put $i\theta$ into the series for the exponential and collect the real and imaginary pieces, done in section 3.2. The result is $e^{i\theta} = \cos\theta + i\sin\theta$.

Special cases of this equation say

$$e^{2\pi i} = 1, \qquad e^{\pi i} = -1, \qquad e^{i\pi/2} = i, \qquad e^{2n\pi i} = 1$$

There are three cube roots of one, and all that you need to find them is the preceding line.

$$1^{1/3} = \left(e^{2n\pi i}\right)^{1/3}$$

Take n to be a succession of integers

$$n = 0 \longrightarrow 1^{1/3} = 1$$
$$n = 1 \longrightarrow \left(e^{2\pi i}\right)^{1/3} = e^{2\pi i/3} = \cos 2\pi/3 + i\sin 2\pi/3 = (-1 + i\sqrt{3})/2$$
$$n = 2 \longrightarrow \left(e^{4\pi i}\right)^{1/3} = e^{4\pi i/3} = \cos 4\pi/3 + i\sin 4\pi/3 = (-1 - i\sqrt{3})/2$$

If you keep going to $n = 3, 4,$ *etc.* or use negative integers, you simply repeat these three values. A picture of the roots shows them equally spaced around the unit circle, exactly as dictated by Euler's equation, and the same sort of picture appears for higher roots too.
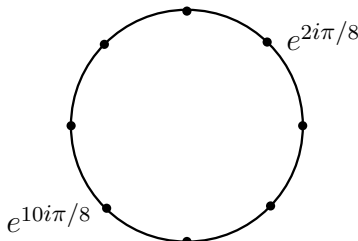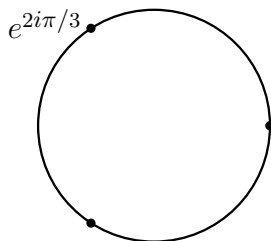
The polar form of complex numbers uses the exponential representation, and here are some examples that use this manipulation.

$$\sqrt{i} = \left(e^{i\pi/2}\right)^{1/2} = e^{i\pi/4} = \frac{1+i}{\sqrt{2}}.$$

$$\left(\frac{1-i}{1+i}\right)^3 = \left(\frac{\sqrt{2}e^{-i\pi/4}}{\sqrt{2}e^{i\pi/4}}\right)^3 = \left(e^{-i\pi/2}\right)^3 = e^{-3i\pi/2} = i.$$

$$\left(\frac{2i}{1+i\sqrt{3}}\right)^{25} = \left(\frac{2e^{i\pi/2}}{2(\frac{1}{2}+i\frac{1}{2}\sqrt{3})}\right)^{25} = \left(\frac{2e^{i\pi/2}}{2e^{i\pi/3}}\right)^{25} = \left(e^{i\pi/6}\right)^{25} = e^{i\pi(4+1/6)} = \frac{\sqrt{3}+i}{2}$$

Another application of Euler's formula is to ordinary trigonometry. What happens when you multiply two complex numbers expressed in polar form?

$$z_1 z_2 = r_1 e^{i\theta_1} r_2 e^{i\theta_2} = r_1 r_2 e^{i(\theta_1+\theta_2)} \tag{0.47}$$

Putting it into words: multiply the magnitudes and add the angles in polar form.

From this you can immediately deduce some of the common trigonometric identities. Use Euler's formula in the preceding equation and write out the two sides.

$$r_1(\cos\theta_1 + i\sin\theta_1)r_2(\cos\theta_2 + i\sin\theta_2) = r_1 r_2 \left[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right]$$

The factors r_1 and r_2 cancel. Now multiply the two binomials on the left and match the real and the imaginary parts to the corresponding terms on the right. The result is the pair of equations

$$\cos(\theta_1+\theta_2) = \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2$$
$$\sin(\theta_1+\theta_2) = \cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2 \tag{0.48}$$

and this is a much simpler derivation than usual for these common identities. You can do similar manipulations for other trigonometric identities, and in some cases you will encounter relations for which there's really no other way to get the result. That is why you will find that in physics applications where you might use sines or cosines (oscillations, waves) no one uses anything but complex exponentials. Get used to it.

The important applications of complex numbers in this text appear when you want to differentiate complex functions, especially the exponential.

$$\frac{d}{dx}e^{ix} = i\,e^{ix} = \frac{d}{dx}\left(\cos x + i\sin x\right) = -\sin x + i\cos x$$

and obviously, the second and the fourth forms agree. Do another derivative and you get

$$\frac{d^2}{dx^2}e^{ix} = i^2 e^{ix} = -e^{ix}$$

so this function e^{ix} satisfies the harmonic oscillator equation, the subject of chapter three.

There are some practice exercises on complex algebra at the end of this chapter, but for more examples see chapter three of Mathematical Tools, mentioned in the bibliography on page iv.

0.8 Separation of variables

This section appears in chapters 2, 3, 4, and in another version, in chapter 7.

The subject of differential equations is large enough to make a profession of it and still not exhaust the subject, but in this text, when you solve differential equations, there are just two methods that show up with any regularity. "Separation of variables" is one. "Linear constant coefficient equations" is the other (next section). After that there are a few equations such as Eq. (6.10) that stand on their own, and you can wait until you get there to find out about them.

A differential equation is an equation relating a function and one or more of its derivatives, and $\vec{F} = m\vec{a}$ is this semester's differential equation. The first tool in your kit is separation of variables, and it is easiest to understand by starting with an example or two. Let c be a constant.

$$\frac{dx}{dt} = c^2 + x^2 \quad \longrightarrow \quad \frac{dx}{c^2 + x^2} = dt \quad \longrightarrow \quad \int \frac{dx}{c^2 + x^2} = \int dt$$

The first of these is the differential equation to be solved. It is a first order equation, meaning that it is a relation between the function x and only the first derivative dx/dt. There are two variables here, the independent variable t, and the dependent variable x. You cannot simply integrate this with respect to t because the right side is a function of x, and that is an (unknown) function of the variable t. To separate variables put all the x's on one side of the equation and all the t's on the other. The second equation does this. It is now set up for integration.

Now do the integral, a trig substitution works: $x = c\tan\theta$.

$$dx = c\sec^2\theta\, d\theta \quad \longrightarrow \quad \int \frac{c\sec^2\theta\, d\theta}{c^2 + c^2\tan^2\theta} = \int \frac{1}{c}d\theta = \frac{1}{c}\theta = \frac{1}{c}\tan^{-1}\frac{x}{c} = t + D$$

and the solution is $x(t) = c\tan c(t + D)$. With an initial conditions such as $x(0) = x_0$ it is

$$x(0) = x_0 = c\tan(cD) \quad \longrightarrow \quad D = \frac{1}{c}\tan^{-1}(x_0/c) \quad \longrightarrow \quad x(t) = c\tan\left(ct + \tan^{-1}(x_0/c)\right)$$

Check the last expression: $x(0) = c\tan\left(\tan^{-1}(x_0/c)\right) = x_0$. *Never assume that you haven't made a mistake.* As time increases, $x(t)$ increases, so $(c^2 + x^2)$ increases, so dx/dt increases, so the slope of the curve x versus t gets bigger and bigger — that's how the tangent of t behaves.

This method looks like such a special one; the combination of factors that allow you to do this seems so improbable that it can't work very often. True. But, it happens in enough important special cases that you have to know about it and learn to recognize when it can apply.

$$1:\ dN/dt = -\lambda N, \quad 2:\ \frac{d^2x}{dt^2} = -\omega^2 x, \quad 3:\ t\frac{dx}{dt} = \alpha x + \beta, \quad 4:\ t\frac{dx}{dt} + \alpha tx = \beta x \qquad (0.49)$$

Equations 1, 3, and 4 are separable, but not 2, though in chapters 2 and 3 you will see some manipulations that will dig a separable equation out of even that one.

Wait, couldn't you manipulate the second of these to be $\frac{d^2x}{x} = -\omega^2\, dt^2$ and integrate? No! There's no such mathematics as this, so don't try.

For other examples of this method, look at Eqs. (2.13), (2.17), (2.23), (3.55).

0.9 Constant Coefficient ODEs

This sort of differential equation shows up often in this course, starting in chapter two, and commonly after that. It looks like

$$3\frac{d^2x}{dt^2} - 4\frac{dx}{dt} + 7x = 0 \qquad \text{or} \qquad \gamma\frac{d^3x}{dt^3} + \delta\frac{d^2x}{dt^2} + \epsilon\frac{dx}{dt} + \zeta x = A\cos\omega t$$

The dependent variable can have any number of derivatives, but it appears just to the first power, no x^2 or $x\frac{dx}{dt}$ or $\sin(kx)$. That makes these equation *linear*. That the coefficient of the x's are constants make these *constant coefficient* linear equations. That the first one has only terms in x or its derivatives makes it *homogeneous* and that the second one has an extra term with no x at all makes it *inhomogeneous*. The precise definition of homogeneous is that if you multiply the variable x by a constant λ, then the whole expression is multiplied by some power of λ, *i.e.* λ^n. Here $n = 1$.

The first case, the linear constant coefficient homogeneous one, has a simple solution. All you need to notice is that the derivative of an exponential is an exponential, and then try a solution $x(t) = Ae^{\alpha t}$.

$$3\frac{d^2x}{dt^2} - 4\frac{dx}{dt} + 7x = 0 \;\longrightarrow\; 3A\alpha^2 e^{\alpha t} - 4A\alpha e^{\alpha t} + 7Ae^{\alpha t} = 0$$

$$Ae^{\alpha t}[3\alpha^2 - 4\alpha + 7] = 0$$

Since neither A nor the exponential are zero, that leaves $3\alpha^2 - 4\alpha + 7 = 0$, a polynomial equation with two roots, giving two solutions to the equation. Because you are trying to undo two derivatives to get x, you will somehow get two arbitrary constants. The key property of linear homogeneous equations is that the sum of two solutions is a solution, so the full solution to this equation is

$$A_1 e^{\alpha_1 t} + A_2 e^{\alpha_2 t}, \qquad \text{where} \qquad \alpha_{1,\,2} = \left(2 \pm i\sqrt{17}\right)/3$$

How do you handle the inhomogeneous case example above? An exponential won't work here. You will not get $A\cos\omega t$ out of it in order to match the right-hand side. The sum of two solutions is no longer a solution. But, there is one simplification: If you (temporarily) throw away the inhomogeneous term ($A\cos\omega t$), you can solve the remaining homogeneous part of the equation with a simple exponential. O.k. you get a cubic equation, but it's only a polynomial equation so there are ways to handle it. This partial solution will have three arbitrary constants. Now if somehow you can find any one solution to the whole equation the trick is to add the two partial solutions.

$$\gamma\frac{d^3x_{\text{hom}}}{dt^3} + \delta\frac{d^2x_{\text{hom}}}{dt^2} + \epsilon\frac{dx_{\text{hom}}}{dt} + \zeta x_{\text{hom}} = 0, \qquad \text{with three arbitrary constants}$$

$$\gamma\frac{d^3x_{\text{inh}}}{dt^3} + \delta\frac{d^2x_{\text{inh}}}{dt^2} + \epsilon\frac{dx_{\text{inh}}}{dt} + \zeta x_{\text{inh}} = A\cos\omega t, \qquad \text{with none}$$

Then $x(t) = x_{\text{inh}}(t) + x_{\text{hom}}(t)$. How do you verify this? Plug in to the original equation and watch it work.

For the problems encountered in this book, finding the special, inhomogeneous solution will not be difficult, and later you will see some general methods for finding such solutions even when it *is* difficult.

0.10 Matrices
This section appears in chapters 4, 8, 10.

Just as you have components of vectors with respect to a basis you will have components of certain types of vector-valued functions. You have (v_x, v_y, v_z) or (v_r, v_θ, v_ϕ) with three components for a vector. An important sort of function (a linear, vector-valued function of a vector variable) appears in describing the angular momentum of a rigid body. It also appears in describing dielectric properties of a crystal. And in describing rotations of vectors. And.... Anyway, it too has components (nine this time) and these form matrices. The development of these ideas, showing the reason for the odd-looking rules that matrices obey, can wait until they're needed in section 8.2. For the moment this will be a summary of some rules, without any discussion of the reasons that they are the way they are.

For the moment then a matrix is a square array of numbers. They can be rectangular too, but not here. They can be added, multiplied, divided, even exponentiated.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix} \tag{0.50}$$

and of course subtraction just changes all the $+$ signs to $-$. What matrix plays the role of zero so that adding it changes nothing? An array of all zeroes.

I said that there are nine components and these objects have only four. If you know everything about 2×2 arrays, the extension to 3×3 is easy. Just as with with either mechanics or calculus, the step from one dimension to two is the big one. After that the step to three dimensions or even N dimensions is relatively small. Besides, it's easier to write these and they take only about $8/27$ of the arithmetic to manipulate.

Multiplication obeys

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix} \tag{0.51}$$

Run across the rows of the first matrix and down the columns of the second matrix in order to construct the entries in the product. Just as there is a zero matrix for addition, there is a unit matrix for multiplication. What is it? What entries in the first factor of (0.51) make the product equal the array of e, f, g, h, thereby reproducing the second factor? For the top left entry of the product,

$$ae + bg = e \ \text{ for all } e \text{ and for all } g \implies a = 1, \ \ b = 0$$

This makes the top right entry work too. Similarly for the bottom entries the equations are $c = 0$ and $d = 1$. That makes the identity matrix

$$(I) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The order of multiplication matters, and multiplication is not commutative. You can however check the special case showing that this identity matrix functions just as well as the right hand factor as it does on the left.

The inverse of a matrix is that matrix such that the product with the original is the identity. Set the right side of Eq. (0.51) to the identity matrix and solve the four equations in the four unknowns a, b, c, d. I'll just write the answer, but you should carry out the algebra so that the result is *yours* and not mine.

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix}^{-1} = \frac{1}{eh - fg} \begin{pmatrix} h & -f \\ -g & e \end{pmatrix} \tag{0.52}$$

Multiply this by the original matrix and verify that you get the identity. It works in either order, so check it both ways.

There is no common notation for a matrix as there is for vectors. In the latter case you see boldface type or an arrow or sometimes a squiggly underline, but for matrices there are no standards. Sometimes a boldface sans serif font is chosen for this purpose, and it serves as well as anything else so that's what I will use here.

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}, \quad \text{then} \quad \mathbf{AB} = \mathbf{C} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}$$

The inverse matrix as in Eq. (0.52) obeys

$$\mathbf{BB}^{-1} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$$

The statement that matrix multiplication is not commutative is $\mathbf{AB} \neq \mathbf{BA}$. You do have the associative law though: $\mathbf{A(BC)} = \mathbf{(AB)C}$. Also the distributive law: $\mathbf{A(B + C)} = \mathbf{AB} + \mathbf{AC}$.

Simultaneous equations

Matrices appear in many interesting and elegant contexts. They also appear in mundane settings, but these are no less important. How do you solve two linear equations in two unknowns?

$ax + by = p, \qquad cx + dy = q.$

multiply the first by d and the second by b, then subtract

$dax + dby = dp, \qquad bcx + bdy = bq \quad \longrightarrow \quad dax - bcx = dp - bq \quad \longrightarrow \quad x = \dfrac{dp - bq}{da - bc}$

multiply the first by c and the second by a, then subtract   (0.53)

$cax + cby = cp, \qquad acx + ady = aq \quad \longrightarrow \quad cby - ady = cp - aq \quad \longrightarrow \quad y = \dfrac{cp - aq}{bc - ad}$

This is matrix inversion in disguise.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} \qquad \text{or} \qquad \mathbf{Mx} = \mathbf{p}$$

Multiply both sides of this matrix equation by the inverse of $\mathbf{M}$ from Eq. (0.52).

$$\mathbf{M}^{-1}\mathbf{Mx} = \mathbf{x} = \mathbf{M}^{-1}\mathbf{p} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

This is exactly the same as the preceding explicit solution for x and y. In fact, that explicit solution is how the inverse matrix is derived, so this comparison is really circular.

 Does this always work? No. You can't divide by zero, and in Eq. (0.53) I ignored that important point.

$$dax - bcx = dp - bq \quad \longrightarrow \quad (da - bc)x = dp - bq \qquad \text{also} \qquad (da - bc)y = aq - cp \qquad (0.54)$$

What if $da - bc = 0$? then the right sides of the equations must be zero, otherwise there is no solution. There is a solution if p and q are both zero or if a particular combination of p, q, and the elements of the matrix conspire to make the right side zero.

$$ad - bc = \text{determinant of the matrix}$$

The determinant determines the nature of the solutions (deterministically of course).
 1. If the determinant is non-zero then the solution exists and is unique.
 2. If the determinant is zero and p or q is non-zero there is no solution unless special circumstances occur; then there are an infinite number of solutions.
 3. If the determinant is zero and both p and q are zero there are an infinite number of solutions.

 Case #1 is routine. You solve simultaneous equations and you expect to find a solution. The second case is exceptional, and it can be used to determine properties of the right-hand side. It will

show up in disguise in sections 7.13 and 10.7. The third case is the most common for the purposes of this book. It means that the two equations you are solving are

$$ax + by = 0, \qquad cx + dy = 0 \qquad \text{but} \qquad ad - bc = 0 \qquad (0.55)$$

If for example d and b are $\neq 0$, multiply the first of these by d: $dax + dby = 0$. Now $ad = bc$, so this equation is the same as $bcx + bdy = 0$ or $cx + dy = 0$. That means there is really only one equation for the two unknowns, not two. That in turn means that there are an infinite number of non-zero solutions x and y for the answer. Once you have found one such solution, simply multiply x and y by the same constant and you have another. You can understand this most simply by a graphical interpretation.

$$ax + by = 0 \quad \text{is a straight line through the origin.}$$

and this graph represents an infinite number of possible solutions.

And how do you write "boldface sans serif" on paper? Perhaps by using "Blackboard Bold" style: $\mathbb{ABCDEFGHIJKLMNOPQRSTUVWXYZ}$. This is a way to fake boldface type in writing.

Index Notation

A_{ij} is the set of elements of the matrix $\mathbf{A}$. The indices i and j run from one to whatever the size of the matrix is (two in these examples). The first index specifies the row and the second the column.

$$A_{\text{row, column}} = A_{ij} \longleftrightarrow \mathbf{A} \longleftrightarrow \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{matrix} \leftarrow & \text{first row} \\ \leftarrow & \text{second row} \end{matrix}$$

These are three notations for the same thing because you don't have to think of the subscripts i and j as particular values. It is like the common notation for a function, $f(x)$. You are not supposed to think of this as some particular "x" but as a placeholder for any value that argument can take on.* In this notation matrix addition and multiplication are

$$\mathbf{A} + \mathbf{B} = \mathbf{C} \longleftrightarrow A_{ij} + B_{ij} = C_{ij} \qquad \text{and} \qquad \mathbf{AB} = \mathbf{C} \longleftrightarrow \sum_{j=1}^{N} A_{ij} B_{jk} = C_{ik}$$

A column matrix has a single index

$$x_i \longleftrightarrow \mathbf{x} \longleftrightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\text{and} \qquad \mathbf{Ax} = \mathbf{y} \longleftrightarrow \sum_k A_{jk} x_k = y_j \longleftrightarrow \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

This sort of product-and-sum occurs so often that the conventional notation is again to omit the summation symbol just as in Eq. (0.28). Whenever a product appears with a repeated index *summation is implied*. The two sums just above then appear as

$$A_{ij} B_{jk} = C_{ik} \qquad \text{and} \qquad A_{jk} x_k = y_j \qquad (0.56)$$

In this kind of manipulation you will find that a repeated index always appears as a pair. If you find a combination such as $A_{ij} B_{jk} C_{jn}$ then go back and find your mistake. It shouldn't happen.

---

* Mathematicians will argue that this is bad notation, and that you should think of f at the function and $f(x)$ as the particular value of the function at the point x. They have a point. That is technically the correct thing to do, and making that distinction can help keep you out of trouble, but it is cumbersome, and this good advice is often ignored. There is a case in chapter eight however, where I will raise this issue again, and there I will side with the mathematicians.

0.11 Iterative Solutions

This material is used in chapters 4 and 5.

Sometimes a complicated equation is really a simple equation in disguise. You just have to look at it the right way. Take the quadratic equation

$$.01x^2 - x + 1 = 0, \qquad \text{and solve:} \qquad x = \left(1 \pm \sqrt{0.96}\right)/.02$$

but really, this is almost a linear equation: $-x + 1 = 0 \longrightarrow x = 1$, and that's much easier. What if you need more accuracy? Just rearrange the equation to be

$$x = 1 + .01x^2, \qquad \text{then an improved solution is} \qquad x = 1 + .01(1)^2 = 1.01$$

If that's still not good enough then you can iterate the process until you're tired of it.

$$x = 1 + .01(1.01)^2 = 1.010201 \qquad \text{and again,} \qquad x = 1 + .01(1.010201)^2 = 1.01020506060401$$

or maybe you want to do it again to get 1.0102051426447 or just 1.0102051, and you may then decide that 1.01 was probably good enough.

Why do this when you have the quadratic formula? What if the equation is $.01x^5 - x + 1 = 0$? You will have to hunt very hard to find the quintic formula. Iteration however is no harder than it was for the quadratic equation.

$$x = 1 + .01x^5 \longrightarrow x = 1.01 \longrightarrow x = 1 + .01(1.01)^5 = 1.0105101005$$
$$\longrightarrow x = 1 + .01(1.0105101005)^5 = 1.0105367 \longrightarrow x = 1.010538$$

If the equation is not algebraic, maybe you can do the same thing. If you encounter the equation $x + 1 + .01\sin x = 0$ then do a single iteration to get $x = -0.9916$. But of course all these examples were set up so that you could do such repetitive calculations easily. That doesn't always happen, but then *nothing* always works. That's why you need a large tool kit. This method will be used in sections 4.1, 4.5, and 5.3. It will greatly simplify the solution and the interpretation of the results found there.

Though you don't need it for later applications in this text, you may be wondering about the *other* solution to the quadratic equation $.01x^2 - x + 1 = 0$. (If you weren't, why not?) Sketch a graph of this function. It is $1 - x$ plus a small bit of x^2. That means that it doesn't come back up to intersect the x-axis until some large value of x. If x is large, what can balance it in this equation? Not 1 certainly. That means it must be the $.01x^2$ term that balances it—that is all there is left.

$$.01x^2 - x = 0 \longrightarrow .01x - 1 = 0 \longrightarrow x = 100$$

This is the lowest order approximation to the result. How do you improve on it? Rearrange the equation to take advantage of the fact that x is now large: Divide by it.

$$.01x^2 - x + 1 = 0 \longrightarrow .01x - 1 + \frac{1}{x} = 0 \longrightarrow x - 100 + \frac{100}{x} = 0 \longrightarrow x = 100 - \frac{100}{x}$$

Now iterate using this.

$$x = 100 - \frac{100}{100} \longrightarrow x = 99 \longrightarrow x = 100 - \frac{100}{99} = 98.9898, \quad \text{etc.}$$

Exercises

1 In Eq. (0.1) you have the series for $1/(1-x)$. Differentiate it with respect to x. Next use the binomial series for $n = -2$ and expand $1/(1-x)^2$ to see if the results match.

2 Divide one of the two equations in Eq. (0.11) by the other and manipulate to obtain an identity for $\tanh(x+y)$.

3 Verify the equations (0.11) by multiplying out the expressions on the right side using the definitions of the hyperbolic functions.

4 Verify that $x^2 + y^2 + z^2 = r^2$ in the second set of equations (0.13).

5 Either verify Eq. (0.43) or correct it if it's wrong.

6 Verify that the result in Eq. (0.44) is right. Not that the derivation is right, that the result is! Also, what is its time-derivative?

7 Use the series in Eq. (0.1) to derive Eq. (0.46).

8 Set the right side of Eq. (0.51) equal to the unit matrix. If you know the numbers e, f, g, h, then solve for a, b, c, d. For example, $ae + bg = 1$ is your first equation. Ans: Eq. (0.52)

9 Take $\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and find all the p's and q's such that the equations (0.54) do have a solution. Then how many solutions do they have?

10 Use the procedure starting at Eq. (0.5) to derive the power series for the sine and for the cosine.

11 Compute the area of the triangle with vertices $(0,0)$, $(a,0)$, (a,b). Do it twice as a double integral $dx\, dy$, interchanging limits. Now integrate the function x over the same area both ways and see if you again get the same answer both ways. Since you've done both integrals, divide the second result by the first.

12 Verify that the spherical coordinate velocity and acceleration, Eq. (0.42), reduce to the plane polar coordinate versions in the equatorial plane ($\theta = 90°$). Verify the same along constant longitude ($\phi = $ constant).

13 In what states are the most Northern, Southern, Western, and Eastern points in the United States? Ans: You're probably wrong. Look *very* closely at the map on page 8, or consult a large atlas.

14 Each day the Earth gains about $100\,000$ tons from space debris (rocks, dust, etc.) About what is the average daily change in the Earth's radius from this bombardment? Express the result in atomic diameters. If you find yourself involved in a lot of arithmetic, reread Eq. (0.31). Do *not* become involved with subtracting large, almost-equal numbers. Let the algebra do the labor for you.

Problems

0.1 Find the next order correction to the series expansion for Eq. (0.3). Check the result by a numerical comparison of the exact result versus your approximate result for a couple of modest sized values of at/c. Use $a = g$, but first convert $g = 9.81\,\mathrm{m/s^2}$ to the units light-years per year squared.

0.2 Write the power series expansion about $t = 0$ for $1/(1+t)$, then evaluate the integral $\int_0^x dt/(1+t)$ by integrating this series. Ans: See table Eq. (0.1).

0.3 Use series expansions to find the limit as $x \to 0$ of

$$\frac{1}{\sin^2 x} - \frac{1}{x^2}$$

The series for the sine and the binomial series are what you need. Test your result experimentally by putting various values of x into a hand calculator.

0.4 Improve on the preceding calculation and find the behavior of that function for small x. Find the results in a power series up through terms in x^2. Check your approximate result versus the exact one using a calculator for a couple of small x. Again you will use the sine series and the binomial expansion, but keep the next order terms. Ans: $\frac{1}{3} + \frac{1}{15}x^2$

0.5 The limit taken in Eq. (0.2) was simply the value as $x \to 0$. Improve it by keeping more terms and finding the behavior for small x instead of just for zero x. Ans: $-\frac{1}{2} - \frac{1}{8}x$

0.6 The relativistic expression for the kinetic energy of a non-zero mass particle is

$$K = mc^2 \left[\frac{1}{\sqrt{1 - v^2/c^2}} - 1 \right]$$

For small speed $(v \ll c)$ expand this to terms in v^4 at least.

0.7 The hyperbolic sine is an odd function, $\sinh(-x) = -\sinh x$, so the inverse hyperbolic sine is odd too. Equation (0.10) doesn't look odd, but prove that it is anyway.

0.8 Derive the equations (0.8).

0.9 From the definition of $\sinh$, write $\sinh(2x)$ then factor the result and derive the identity $\sinh(2x) = 2\sinh x \cosh x$. Similarly, find $\cosh(2x)$ in terms of hyperbolic functions of x.

0.10 Derive the power series expansions of $\sinh x$ and $\cosh x$ about $x = 0$. Ans: See table Eq. (0.1).

0.11 **(a)** $y = \sinh^{-1} x$ means $x = \sinh y$. Differentiate the latter equation with respect to x, solve the result for dy/dx, use a simple identity to eliminate the $\cosh$, and show that the derivative of $\sinh^{-1} x$ is $1/\sqrt{1 + x^2}$. **(b)** Repeat this calculation of the derivative, but starting from Eq. (0.10).

0.12 Just after the equation (0.10) there is a set of graphs of the various hyperbolic functions. Sort out which graph is which.

0.13 In Eq. (0.7) you see how the hyperbolic functions produce a hyperbola for a graph. Now change variables (rotate coordinates) to $x' = (x+y)/\sqrt{2}$, $y' = (x-y)/\sqrt{2}$ and draw the corresponding graph for this.

0.14 Substitute ix into the power series for $\cos x$ to get $\cos(ix)$ and show that it is $\cosh x$. Find an analogous relation for $\sin(ix)$.

0.15 From the preceding problem to get $\cos$ and $\sin$ of ix, show how to go from the known identities for $\cos(x + y)$ and $\sin(x + y)$ to the corresponding ones for $\cosh$ and $\sinh$. Ans: $\cosh(x + y) = \cosh x \cosh y + \sinh x \sinh y$

0.16 From the equation $\cos^2 x + \sin^2 x = 1$ you get a useful identity by dividing it by $\cos^2 x$. What is the analogous result starting from $\cosh^2 x - \sinh^2 x = 1$?

0.17 In problem 0.4, change the sine to a hyperbolic sine and repeat the calculation.

0.18 Similar to the relations in Eq. (0.13), find the relations between cylindrical and spherical coordinates.

0.19 Compute the area of the triangle in Eq. (0.20) using *polar* coordinates. The triangle has vertices at $(0, 0)$, $(a, 0)$, and (a, b). Carefully consider the order of integration, drawing enough pictures to allow you to make a considered choice.

0.20 Compute the volume of a sphere using (**a**) spherical coordinates, (**b**) cylindrical coordinates.

0.21 Compute the area of a sphere using (**a**) spherical coordinates, (**b**) cylindrical coordinates, (**c**) rectangular coordinates.

0.22 Compare the areas on the Earth between $10°$ and $11°$North latitude to the area between $79°$ and $80°$. Take the ratio. Can you do this by brute force, using a calculator that keeps a zillion digits? Yes, but don't. Use algebra and some thought instead before you grab the calculator. Ans: $\tan 10.5° = 0.1653$

0.23 A moment of area is $\int r^2 \, dA$. Use the area of the triangle that led to Eq. (0.20) with r measured from the origin, and evaluate this moment, doing the integral twice, once in each order shown.

0.24 Express the rectangular components x, y, z of a point in terms of its spherical coordinates. Verify $x^2 + y^2 + z^2$. Ans: *e.g.* $x = r \sin \theta \cos \phi$

0.25 (**a**) Compute the dot product of $(3\hat{x} + 4\hat{y})$ and $(4\hat{x} + 3\hat{y})$ two ways: once using components and once using the original definition of the dot product, Eq. (0.21). Equate the results and deduce the angle between the vectors. (**b**) Repeat the calculation to find the angle, but using the cross product instead — once with the definition and once with basis vectors and components. What ambiguities appear in these results? Ans: in part $16.3°$

0.26 Obtain the law of cosines in trigonometry by interpreting the product $(\vec{A} - \vec{B}) \cdot (\vec{A} - \vec{B})$. Let $\vec{C} = \vec{A} - \vec{B}$. No components please.

0.27 (**a**) For a triangle with two sides being $\vec{A}$ and $\vec{B}$, show that $\vec{A} \times \vec{B}$ has magnitude twice the area of the triangle. (**b**) Make the third side of the triangle $\vec{C}$ and (having chosen the directions appropriately) show that $\vec{A} \times \vec{B} = \vec{A} \times \vec{C}$. From there derive the law of sines. Again, no components please.

0.28 Show by drawing pictures and using the geometric definitions of the products that $\vec{A} \cdot \vec{B} \times \vec{C}$ is ($\pm$) the volume of the parallelepiped spanned by the three vectors. From this picture, show why the first of the identities Eq. (0.25) is true. The second one too.

0.29 Let $\vec{a}$ be a fixed vector. Show that the in the plane, $r^2 - 2\vec{a} \cdot \vec{r} = 0$ is the equation of a circle. Interpret this equation to show that the angle inscribed in a semicircle is a right angle.

0.30 If $\vec{A} \cdot \vec{B} = 0$ and $\vec{A}$ is not zero, show that the simultaneous equations $\vec{V} \times \vec{A} = \vec{B}$ and $\vec{V} \cdot A = p$ have the solution

$$\vec{V} = (\vec{A} \times \vec{B} + p\vec{A})/\vec{A} \cdot \vec{A}$$

0.31 If $\hat{u}$ is a unit vector in an arbitrary direction, show that for any vector $\vec{A}$ this identity holds. Also, draw a picture to show what this identity looks like.

$$\vec{A} = \hat{u}(\vec{A} \cdot \hat{u}) + \hat{u} \times (\vec{A} \times \hat{u})$$

0.32 Evaluate (remember the summation convention)

$$\delta_{ij}\delta_{jk}, \qquad \delta_{ii}\delta_{jj}, \qquad \delta_{ij}\delta_{ij}, \qquad \delta_{ii}\delta_{ii}, \qquad \delta_{ii}\delta_{jk}$$

0.33 (a) Derive the Jacobi identity, Eq. (0.23) from Eq. (0.24). (b) Derive Eq. (0.24) equation from Eq. (0.30). (c) Derive that equation too.

0.34 Carry out the calculations of the derivatives of Eq. (0.34), with answers a few lines after that. Also compute the derivatives of $f_3 = df_1/dx$ and of $f_4(x) = \int_0^{x/2} t\sqrt{x-t}\,dt$. In the case of f_4, also evaluate the integral and differentiate the result to see if the two ways to calculate the derivative give the same answer. To do the integral, notice that the factor t can be written as $(-(x-t)+x)$.

0.35 (a) Compute the derivative with respect to t of $\int_t^{2t} dx\, e^{-tx^2}$. (b) Make a change of integration variables so that the new limits are constants. Now compute the time derivative and compare the two answers.

0.36 A tricky derivative. Try it the straight-forward way first and demonstrate how it blows up on you. Next do a partial integration to put the integral into a different form and then do the derivative. $f(t)$ is an unspecified, but differentiable, function of t. Show that the result is

$$\frac{d}{dx}\int_0^x dt\,\frac{f(t)}{\sqrt{x-t}} = \frac{1}{\sqrt{x}}f(0) + \int_0^x dt\,\frac{\dot{f}(t)}{\sqrt{x-t}}$$

Test this on an explicit, non-constant, special f for which you can do the integral; then do this derivative directly and with the formula that you just derived.

0.37 The preceding problem can also be done by going back to the Δ-definition of a derivative and manipulating that. No partial integrations, just the definitions and maybe a coordinate shift.

0.38 Start from the product rule for differentiation, Eq. (0.36), integrate it and so derive the equation for partial integration, $\int u\,dv = uv - \int v\,du$.

0.39 A function of x is defined to be $\int_0^\infty t^{x-1}e^{-t}dt$. Call it $\Gamma(x)$. Change variables in the integral to $t = \alpha u$. Differentiate the result with respect to α. Of course, the answer has to be zero doesn't it? Now set $\alpha = 1$ and show that this gives $\Gamma(x+1) = x\Gamma(x)$. (b) Evaluate $\Gamma(1)$. Now what are $\Gamma(2)$, (3), (4), (5)?

0.40 Compute

$$\frac{d}{dx}\int_x^{2x} dt\, e^{xt^3}, \qquad \frac{d}{dx}\int_{-x}^x dt\, e^{xt^3}$$

without doing the integral (you won't be able to anyway). Check one point of each of your results by asking what these derivatives are at $x = 0$. [For small (not zero) x, what is the integral?] Compare your solutions to this special case. Sketching a graph may help.

0.41 (**a**) Take the square of Eq. (0.46) and deduce two common identities concerning trigonometric functions of double angles. (**b**) Take the cube of the same equation and deduce two not-so-common trigonometric identities for triple angles.

0.42 Start from the equation (0.1) for the exponential and substitute $x = i\theta$. Collect the real and imaginary parts, and use other series from (0.1) to derive Euler's formula for $e^{i\theta}$.

0.43 Express in polar form, $re^{i\phi}$

$$\frac{1+i}{1-i}, \qquad \frac{-1+i\sqrt{3}}{+1+i\sqrt{3}}, \qquad 1+i, \qquad 14-17i$$

0.44 Sketch the points in the complex plane: $|z| < 1$, $|z-2| > 1$, $z + z^* = 5$, $z - z^* = 5$.

0.45 Show that $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$ and that $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$.

0.46 Assume that the equations of problem 0.45 are valid for complex values of x, and use that to define $\cos(x + iy)$ and $\sin(x + iy)$. That is, *start* with $\cos(x + iy)$ defined by replacing $x \to x + iy$ in problem 0.45, then rearrange the result to show that (**a**) $\sin(x + iy) = \sin x \cosh y + i \cos x \sinh y$. (**b**) Find the analogous equation for $\cos(x + iy)$. (**c**) And what are these when $x = 0$? (**d**) What is $\cos^{-1} 2$? What is $\sin^{-1} 2$? Ans: in part: $1.570796 - i1.316958$

0.47 Differentiate $e^{i\theta}$ with respect to θ to derive the differentiation formulas for sine and cosine.

0.48 What is $\cosh\left(x + \frac{1}{2}i\pi\right)$?

0.49 Find the area of an ellipse. Doing the integral $\int dA = \int dx\, dy$ is easiest in rectangular coordinates, with the equation Eq. (0.16). Ans: πab

0.50 In Eq. (0.38) there's the derivative of a dot product. Can you really do this? Write $\vec{v} \cdot \vec{v}$ in rectangular components, differentiate it, and reassemble the results to show that it works.

0.51 In Eq. (0.39) it says "and similarly", leaving the rest of the derivation to you. Do it.

0.52 Fill in the steps leading to the example result, Eq. (0.44), and verify that the result is plausible. Then compute its acceleration. Pictures of course. Lots of arrows.

0.53 A car starts from rest and moves along a circular track of radius R. It has constant forward acceleration, so the distance along the track is $at^2/2$. Write its velocity and acceleration in the $\hat{r}$-$\hat{\phi}$ basis and sketch the $\vec{r}$, $\vec{v}$, and $\vec{a}$ vectors at different times.

0.54 (**a**) Compute $d\vec{a}/dt = d^3\vec{r}/dt^3$ in plane polar coordinates. (**b**) What is this for circular motion at constant speed? Ans: $\hat{r}(\dddot{r} - 3\dot{r}\dot{\phi}^2 - 3r\dot{\phi}\ddot{\phi}) + \hat{\phi}(r\dddot{\phi} + 3\dot{r}\ddot{\phi} + 3\ddot{r}\dot{\phi} - r\dot{\phi}^3)$

0.55 In Eq. (0.45) it shows how to simplify a fraction by rationalizing the denominator. What do you get if you rationalize its numerator instead?

0.56 Use the identities in Eq. (0.48) or another way if you prefer, and derive the occasionally useful identities

$$\cos x + \cos y = 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right) \quad \text{and} \quad \cos x - \cos y = 2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{y-x}{2}\right)$$

0.57 Solve by separation of variables $dx/dt = a + x$ with the initial condition that $x(0) = b$.
Ans: $x(t) = (a+b)e^t - a$

0.58 Solve the 1st, 3rd, and 4th of the equations in Eq. (0.49), finding the general solution.

0.59 Find a combination of numbers a, b, c, d, p, q in Eq. (0.54) so that one or both of p and q are not zero, the determinant *is* zero, and there is still a non-zero solution for x and y.

0.60 Sometimes a quadratic equation is a linear equation in disguise.
(**a**) Solve the equation $0.001x^2 - x + 2 = 0$ by noticing that it is almost linear, so $x = 2$ is *almost* right. (**b**) Use iteration. Get the second solution too.

0.61 How do you take the square root of a number by hand? You solve the equation $x^2 = a$ by iteration. If x_0 is an approximate answer and ϵ is the error, then

$$a = (x_0 + \epsilon)^2 = x_0^2 + 2x_0\epsilon + \epsilon^2 \approx x_0^2 + 2x_0\epsilon$$
$$\longrightarrow \quad \epsilon \approx \frac{a - x_0^2}{2x_0} \quad \longrightarrow \quad x_1 = x_0 + \epsilon = \frac{a + x_0^2}{2x_0}$$

The improved estimate of the root is this x_1. Use this to compute $\sqrt{2}$ by repeating this process with improved values at every step. Start with a stupid initial guess such as $x_0 = 1$ and iterate to see how fast it converges. This is the basis for the internal square root algorithm used inside some computers. There however, they will make a more intelligent starting choice. Ans: 1., 1.5, 1.417, 1.414216, 1.4142135623747, ...

Introduction

.

What are Newton's laws of motion? For a caricature of them (don't believe this) you have

1. An object with no external forces on it will either remain at rest or stay moving at constant velocity.
2. $\vec{F} = m\vec{a}$.
3. Action equal reaction.

Start with the last one. What in the world is "action"* let alone "reaction"? The answer is that you would never write this law in such a language today. A formulation in contemporary language is

3. If one object exerts a force on a second, then the second will exert a force on the first of the same magnitude but opposite direction. In colloquial mathematics this is

$$\vec{F}_{\text{on 1 by 2}} = -\vec{F}_{\text{on 2 by 1}}$$

Not obviously so, but another, better way to state the third law is *momentum is conserved*. See problem 4.46 for a puzzle about this.

For the first law, can't you derive it from the second? Just set the total force on an object to zero, then $\vec{a} = 0$ and the velocity is constant. Didn't Newton notice this?

As stated, the first law is obviously false. The sun goes around the Earth daily. It is moving on a circle and its velocity is certainly changing, but there's no force causing it to do this. You say that this is because of the Earth's rotation on its axis? Where, in the first law as stated above, is this precluded? The last time that I was on a merry-go-round I saw people moving up and down and going in circles around me. They weren't moving at constant velocity even though there was no force to push them into this odd motion. If you think that I'm wrong to say this, that I'm the one who's moving, then explain why I can't assume that I'm the center of the universe. What if I prefer to think that the world moves around me and that I am forever standing still?

The answer is that I *can* do this. It is a question of complexity. Newton's first law is really a definition: that of an "inertial frame", and it is only in inertial systems that the laws of nature take on an especially simple form. A statement of Newton's first law is

1. An inertial observer (or inertial frame or inertial coordinate system) is one for which IF no forces act on an object — there are no discernible interactions with any other body — THEN that object will move with constant velocity. This is the *definition* of an inertial system.

This doesn't say that there's anything wrong with a coordinate system that is not inertial. It simply says what it means to *be* inertial. The Earth is not an inertial system. The merry-go-round isn't either. Does that mean there's anything wrong with setting up a coordinate system centered on either of these? No, it simply means that you can't use Newton's laws in their simplest form. Chapter five is devoted to this question.

---

* The Oxford English Dictionary says it is either momentum times time, or kinetic energy times time, or maybe closest to our purposes, "The exertion of force by one body upon another; influence". And these don't exhaust the possibilities.

For the second law,

> 2. IF you are in an inertial frame, THEN the acceleration of a point mass is equal to the total force on it divided by its mass. $\vec{a} = \vec{F}/m$. More generally, $\vec{F} = d(m\vec{v})/dt$ if the mass that is in motion happens not to be constant.

When writing the laws this way you can't even ask that question about deriving the first law from the second. You need the definition stated in the first law even to state the second law. Do these restatements of the laws of mechanics clarify everything? No. There are still some ideas to resolve, some of them hard. What is acceleration? (easy) What is mass? (pretty easy) What is force? (tough)

Which is correct, $\vec{F} = m\vec{a}$ or $\vec{F} = d(m\vec{v})/dt$? That is an experimental question, and when the mass of an object is not constant, the two forms aren't the same. The second one is right, and that's the one that is consistent with conservation of momentum. Of course the validity of *that* law is an experimental question too. Example: a raindrop is falling and while falling it picks up more water from the surrounding moist air. Example: a rope is coiled on the table. Pick up one end, pulling it straight up at constant speed; only *some* of the mass is moving at any time. For most of the material in this text, the mass will be constant, so you can pull it out of the derivative and use the simpler form. But,

$$\text{In a non-inertial system} \quad \vec{F} \neq m\vec{a} \quad \text{and it} \quad \neq \frac{d(m\vec{v})}{dt} \quad \text{either.}$$

What do you do if you really want to work in a non-inertial system? Read chapter five.

This statement of Newton's first law defines an inertial frame, but if you're not satisfied, look up plato.stanford.edu/entries/spacetime-iframes/ for a far more detailed definition. It will tell you more than you want to know about the subject. Despite this snarky comment, it is very clear and readable. It lays out and discusses some of the subtleties that I've ignored here.

1.1 Dimensions and Units

Almost every physical quantity that you encounter will have some sort of dimensional property: mass, length, time, charge, temperature, *etc.* There is a difference between dimensions and units, though you can easily ignore this distinction without serious consequence. If mass is a dimension, then kilograms, grams, even slugs are units. For the dimension length you have units meter, decimeter, light-year, furlong, mile, and many others.

If an object has a mass of two kilograms that says that the ratio of its mass to the mass of a certain piece of platinum-iridium alloy kept in a vault near Paris is two to one. When you use a unit of measurement, that's all you are doing — stating a ratio of what you measure to some conventional standard. If it seems arbitrary that's because it is, but it is exactly what you do any time that you measure something, including your own weight (unit: pound, stone, grzywna*, or Newton).

When measuring a length, the common unit is the meter, and that is defined in terms of the properties of light in conjunction with a special atomic clock. One meter is the distance that light travels in a vacuum in $1/299\,792\,458$ seconds. It is just as arbitrary as the kilogram, but at least it uses a physical standard that can be reproduced anywhere in the world. You do not have to go to Paris for it, though perhaps that's a point against it.

What is mass? Maybe not in any very fundamental way, but what is a definition that will let me know what I am measuring? One way is to let two masses interact with each other, perhaps by putting a very light spring in between them and letting them push apart. After they have pushed apart, measure their speeds.
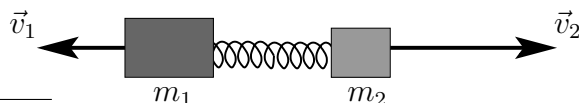

Fig. 1.1

---

* I did not make this up

Define the ratio of the two masses to be $m_1/m_2 = v_2/v_1$. That way, if m_2 is a standard kilogram then you are measuring m_1 in kilograms. What about the spring? Well, this is really some sort of limit as the spring shrinks to nothing.

Is this the only way to define such a mass ratio? No, you could measure accelerations and define this using the ratio of the measured accelerations. You can even define the ratio in another way that may make no immediate sense, but is the right definition when you encounter relativity.

$$\frac{m_1}{m_2} = \frac{v_2\sqrt{1 - v_1^2/c^2}}{v_1\sqrt{1 - v_2^2/c^2}}$$

In this equations c is the speed of light in vacuum, $c = 299\,792\,458\,\text{m/s}$, and for ordinary speeds this definition reduces to the first one. For most purposes these precise statements aren't necessary, but it is appropriate to understand that they are needed at some point in order to define the words that we use.

At a fundamental level there are just a few types of forces. They are

1. gravity
2. electromagnetism $\Big\}$ electroweak
3. weak nuclear $\Big\}$ standard model $\Big\}$?
4. strong nuclear

At one time (before Maxwell in the 1860s) electricity and magnetism were two different forces, but he found that they were just two different aspects of the same thing. More recently the electromagnetic force was united with the weak nuclear force (responsible for beta-decay). A little later these were combined with the strong nuclear force (really the interaction of quarks). Now we're stuck.

In classical mechanics there are exactly two types of forces that you encounter: gravity and electromagnetism, and in the latter case the most common manifestation is the *contact force*. Two objects touch each other and exert forces on each other. Those are in fact molecule to molecule forces and those are electrical. Fortunately you never have to examine the problem at this level; simply say that there is a force of contact between two objects and then let Newton's equations figure out how big it is. That it involves interacting molecules is someone else's business.

For this entire book the only forces are 1. gravity, 2. contact, 3. electric or magnetic fields—mostly numbers one and two. Are there exceptions? You could argue that the inertial forces in chapter five are different, but those are the result of a coordinate transformation. Friction? That is just a contact force, so it is basically electromagnetic, intermolecular forces.

1.2 Types of Mass

Mass is in some sense a resistance to acceleration. Any of the ways to define mass in the preceding section came from this idea. When two masses push each other apart the more massive one accelerates less. In the equation $\vec{a} = \vec{F}/m$, the bigger the mass the less the acceleration, other things being equal.

Mass appears in another basic physical law, gravity. The equation $\vec{F}_{\text{grav}} = m\vec{g}$ says that a gravitational field $\vec{g}$ pushes on an object in direct proportion to some special property of the object, commonly called *mass*.

Look at the last two paragraphs. The same word, "mass", appeared in two different meanings. One involved resistance to acceleration and the other involved the effect of a gravitational field. If I say that the electric field, $\vec{E}$, caused a force $m\vec{E}$ on an object, you should object. Why is it that the effect by an electric field behaves so differently from the effect by a gravitational field? There are logically

two different types of mass. One is inertial mass, describing the resistance to acceleration; the other is gravitational mass, describing the coupling to the gravitational field.

$$\vec{a} = \vec{F}_{\text{total}}/m_{\text{i}}, \quad \text{and} \quad \vec{F}_{\text{gravitational}} = m_{\text{g}}\vec{g}$$

Who says that these two masses are the same?

It is an experimental question. Are they? The question was first asked by Newton, but the real credit goes to Eötvös. In the late 1800's and continuing into the first decade of the 1900's, he carried out some miraculously precise experiments to measure the ratio of the two masses, $m_{\text{i}}/m_{\text{g}}$. Is it the same for all materials? His conclusion was yes, to a precision of about one part in 10^8. Later experiments by Dicke improved this to one part in 10^{11} and more recently Adelberger et al. added another factor of about 100. At this point it seems safe to say that the two masses are the same.

1.3 Conservation Laws

Classical mechanics is not just about forces and accelerations. Even if the problems examined are traditional ones with masses interacting — gravity, friction, magnetism, *etc.* There are other ways to approach those problems, and the conservation laws are at least as basic as $\vec{F} = m\vec{a}$. Conservation of energy, of momentum, of angular momentum can all be derived from that equation, but did you know that you can derive $\vec{F} = m\vec{a}$ from energy? If so, then which of the two is more fundamental?

Work, Energy

In one dimension with point masses, the work-energy theorem is very simple and is not at all difficult to derive from Newton's equations. There are a couple of ways to do so, one involves the full power of calculus manipulations and a few lines of algebra; the other is more intuitive, but more tedious.

Start with the first way (and if you haven't reviewed section 0.5 then go back and do it now).

$$F_x = ma_x = m\frac{dv_x}{dt} = m\frac{dv_x}{dx}\frac{dx}{dt} = mv_x\frac{dv_x}{dx} \tag{1.1}$$

This used the chain rule for differentiation and then it used the definition of velocity. Now integrate this equation with respect to x between some specified initial and final limits.

$$W = \int_{x_{\text{i}}}^{x_{\text{f}}} F_x\,dx = \int_{x_{\text{i}}}^{x_{\text{f}}} mv_x\frac{dv_x}{dx}dx = \int_{x=x_{\text{i}}}^{x=x_{\text{f}}} mv_x\,dv_x = \frac{m}{2}v_x^2\Big|_{x=x_{\text{i}}}^{x=x_{\text{f}}} = \frac{m}{2}v_{\text{f}}^2 - \frac{m}{2}v_{\text{i}}^2 = \Delta K \tag{1.2}$$

Strip the intervening material from this and it is the work-energy theorem:

$$W = \int_{x_{\text{i}}}^{x_{\text{f}}} F_x\,dx = \frac{m}{2}v_{\text{f}}^2 - \frac{m}{2}v_{\text{i}}^2 = \Delta K \tag{1.3}$$

This took two (long) lines, but you need to have a complete understanding of the chain rule for differentiation and of the methods to change variables in an integral. There are two parts to this derivation: First, knowing *how* to do the manipulations. Second, understanding what the manipulations *mean* and why they are valid.

A second way to reach this result is longer but simpler. It doesn't involve any special calculus tricks, only the definition of an integral. It starts with the special case of a constant force, then since $a_x = F_x/m$ the acceleration is a constant. That's a familiar case:

$$a_x = \text{constant} \quad \longrightarrow \quad v_x = a_x t + C, \quad \text{then} \quad x = \tfrac{1}{2}a_x t^2 + Ct + D$$

Use the initial conditions that at $t = 0$ position is $x = x_0$ and velocity is $v_x = v_0$.

$$v_0 = a_x \cdot 0 + C, \qquad x_0 = \tfrac{1}{2}a_x \cdot 0^2 + C \cdot 0 + D \qquad \text{so} \qquad v_x = a_x t + v_0, \quad \text{and} \quad x = \tfrac{1}{2}a_x t^2 + v_0 t + x_0$$

Eliminate the variable t between the last two equations

$$t = (v_x - v_0)/a_x, \qquad \text{so} \qquad x = \tfrac{1}{2}a_x (v_x - v_0)^2 / a_x^2 + v_0 (v_x - v_0)/a_x + x_0$$
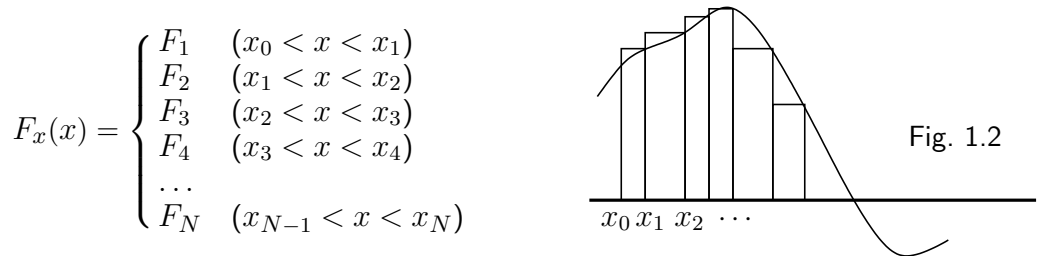
Rearrange the last result, simplifying the algebra to get

$$x = \tfrac{1}{2}(v_x^2 - v_0^2)/a_x + x_0 \qquad \text{or} \qquad ma_x(x - x_0) = \tfrac{1}{2}mv_x^2 - \tfrac{1}{2}mv_0^2$$

All of this rearrangement was just elementary algebra, and the reason for the final manipulation, multiplying by m was to get the combination ma_x in the first term. That becomes

$$F_x \cdot (x - x_0) = \tfrac{1}{2}mv_x^2 - \tfrac{1}{2}mv_0^2 \tag{1.4}$$

Now for the real case. Constant forces are textbook idealizations of the real world. At best you can approximate some force to be close enough to constant for most purposes. What about the more common case for which the force is not at all constant? Answer: use successive approximations as in calculus, and then sneak up on the result. Approximate a position-dependent force as a sequence of steps.

$$F_x(x) = \begin{cases} F_1 & (x_0 < x < x_1) \\ F_2 & (x_1 < x < x_2) \\ F_3 & (x_2 < x < x_3) \\ F_4 & (x_3 < x < x_4) \\ \dots & \\ F_N & (x_{N-1} < x < x_N) \end{cases}$$

Fig. 1.2

In each of these intervals, the equation (1.4) applies. At each point x_0, x_1,... the speed is v_0, v_1, etc.

$$F_1(x_1 - x_0) = \tfrac{1}{2}mv_1^2 - \tfrac{1}{2}mv_0^2$$
$$F_2(x_2 - x_1) = \tfrac{1}{2}mv_2^2 - \tfrac{1}{2}mv_1^2$$
$$F_3(x_3 - x_2) = \tfrac{1}{2}mv_3^2 - \tfrac{1}{2}mv_2^2$$
$$F_4(x_4 - x_3) = \tfrac{1}{2}mv_4^2 - \tfrac{1}{2}mv_3^2$$
$$F_5(x_5 - x_4) = \tfrac{1}{2}mv_5^2 - \tfrac{1}{2}mv_4^2$$
$$\dots$$
$$F_N(x_N - x_{N-1}) = \tfrac{1}{2}mv_N^2 - \tfrac{1}{2}mv_{N-1}^2$$

The sum of all these equations is

$$F_1(x_1 - x_0) + F_2(x_2 - x_1) + F_3(x_3 - x_2) + \cdots + F_N(x_N - x_{N-1}) = \tfrac{1}{2}mv_N^2 - \tfrac{1}{2}mv_0^2$$

because the terms on the right telescope: All the terms except the first and the last cancel in pairs. In conventional notation this is

$$\sum_{k=1}^{N} F_k \Delta x_k = \sum_{k=1}^{N} F_x(x_k)\, \Delta x_k = \tfrac{1}{2}mv_f^2 - \tfrac{1}{2}mv_i^2 \qquad \text{where} \qquad \begin{aligned} \Delta x_k &= x_k - x_{k-1} \\ x_N &= x_f \\ x_0 &= x_i \end{aligned}$$

The limit of this equation as $\Delta x_k \to 0$ is just the definition of the word "integral", so this reproduces Eq. (1.3).

These two derivations of the same equation, $W = \Delta K$, look completely different, but they aren't. Look closely at them and compare them step by step they are much more alike than you first think. The second derivation says that the combination of force times distance is important — especially force times a little bit of distance: $F_x \, dx$. Write down this combination and manipulate it.

$$F_x \, dx = ma_x \, dx = m\frac{dv_x}{dt} dx = m\frac{dv_x}{dt}dx = m \, dv_x \frac{dx}{dt} = mv_x \, dv_x$$

However you do it, this is the work-energy relation stating that the work on a point mass ($\int F_x \, dx$) equals the change in its kinetic energy.

What happens if this is not a point mass? What if you aren't operating in one dimension? What if the force is not a function of position alone? All good questions and all will have to be addressed in chapters 8, 4, and 2 respectively.

Mechanical Energy

For this same simple case of a point mass in one dimension, and if the force is a function of position only, you can rearrange the work-energy theorem to get a conservation law.

The work is an integral of $F_x(x) \, dx$. When you evaluate integrals, the most common way is to find an anti-derivative and then to evaluate it at the endpoints of the integration interval, the fundamental theorem of calculus: If f has an antiderivative and if f is integrable then the two are related.

$$\text{If } \ f(x) = dF(x)/dx \quad \text{then} \quad \int_a^b f(x) \, dx = F(b) - F(a) \tag{1.5}$$

In the present case the function you're integrating is $F_x(x)$ and I will denote its anti-derivative by $-U(x)$. That is,

$$F_x(x) = -\frac{dU(x)}{dx}, \qquad \text{then} \qquad \int_{x_i}^{x_f} F_x(x) \, dx = -U(x_f) + U(x_i) \tag{1.6}$$

Now apply this to the work-energy theorem, Eq. (1.3).

$$W = \Delta K \qquad \text{is} \qquad \int_{x_i}^{x_f} F_x(x) \, dx = -U(x_f) + U(x_i) = \frac{m}{2}v_f^2 - \frac{m}{2}v_i^2$$

$$\text{which is} \qquad \frac{m}{2}v_i^2 + U(x_i) = \frac{m}{2}v_f^2 + U(x_f) \tag{1.7}$$

This is the reason for the minus sign in the definition of U in Eq. (1.6). If you don't put in that minus sign you would not get a plus sign here, in this conservation of energy equation.

U is called the potential energy corresponding to the force $F_x(x)$, and this equation says that the sum of the kinetic and the potential energies has the same values throughout the motion; the sum is conserved.

Can't you *always* do this manipulation to get conservation of mechanical energy out of the work-energy theorem? Can't you always find an anti-derivative (maybe in a big table of integrals)? For example, what about common dry friction? Perhaps you remember an equation such as $F_{\text{fr}} = -\mu_k F_N$. The frictional force as an object slides over a surface is the coefficient of (kinetic) friction times the normal component of the force on the surface. It's just a constant, so you can certainly integrate a constant and get a potential energy U and then get conservation of mechanical energy.

No.

In this example of dry friction, the frictional force is velocity dependent even though it doesn't look like it—there's no velocity in the expression $-\mu_k F_N$. True, but that's because this expression is wrong. The frictional force is velocity dependent because it is always opposite the velocity. This commonly used expression for the force is simply not right. A more correct way to write it is

$$\vec{F}_{fr} = -\mu_k F_N \, \hat{v} \tag{1.8}$$

where $\hat{v} = \vec{v}/v$ is the unit vector in the direction of the velocity of the sliding mass. If it slides in the reverse direction, the force from friction is reversed and you can have two opposite values of the force at the same value of the position. It is not a function of x alone and no U exists. Why don't people (including me) write this the correct way? Answer: It's awkward and I'm lazy.

Notice that I keep using the clumsy phrase "conservation of mechanical energy" instead of "conservation of energy". Why? That's tied in to the fact that conservation of energy was one of the most difficult laws to sort out. Newton may have written his basic equations in the late 1600s, but conservation of energy did not become an accepted law of physics until the middle 1800s. These two expressions, kinetic energy and potential energy, seem easy to derive now, but historically even these were discovered through tortuous and tortious routes.* The problem is that *energy* is not just *mechanical energy*.

Energy is an abstraction, a bookkeeping device much like money. It is a prescription saying that certain mathematically described properties of a system are to be added together and if you do it right, you can come back later, redo the sum and you're going to get the same answer. The historical confusion arose in trying to figure out what terms to include in this sum. It was resolved when heat was seen to be another form of energy (a term in the sum). Then light, sound, chemical energy, and eventually mass were added to the mix. Energy is not a *thing*, even though you can get that impression by reading newspaper articles on the subject.

Did I just say that money is an abstraction? Isn't it just the bills and coins that you carry with you? No. If you buy a car, do you pay for it with hundred dollar bills? Most people don't, and bills and coins are nowhere near conserved. If you include what's in your checking account, that's not bills and coins stored in a vault. It is a set of bits stored on a computer disk. A savings account is more of the same. Then there are whatever mysterious manipulations the Federal Reserve Board can implement. The total "money" is a combination of many different things, of which just a tiny amount is in any sense concrete. If you ask five economists to define money you will likely get five answers (M_1, ..., M_5). The difference with the abstraction called energy is that we think we know what goes into it. Or do we? Cosmologists were surprised by the discovery that the expansion of the universe is accelerating, caused by something called "dark energy" for lack of a better name. No one knows what it is.

Derive $F = ma$

A few paragraphs back I said that you can derive $F = ma$ from energy. How? Take conservation of mechanical energy and differentiate it:

$$\text{``} E = \frac{1}{2}mv_x^2 + U(x) \text{ is conserved''} \qquad \text{means} \qquad \frac{dE}{dt} = 0$$

$$\frac{d}{dt}\left[\frac{1}{2}mv_x^2 + U(x)\right] = \frac{1}{2}m\,2v_x\frac{dv_x}{dt} + \frac{dU}{dx}\frac{dx}{dt} = v_x\left[m\frac{dv_x}{dt} + \frac{dU}{dx}\right] = 0$$

$$\text{or,} \quad m\frac{dv_x}{dt} = -\frac{dU}{dx}$$

---

* but not torturous

and that is $F_x = ma_x$. This is not just a pretty theorem, it is sometimes a useful way to derive the equations of motion in specific cases. Often it involves much easier manipulations to reach the answer, and I'll regularly use this method later in the text.

Momentum

There are other ways to manipulate Newton's equation. The first and least interesting way is to start from Newton's equation for a single point mass and integrate it with respect to time.

$$\vec{F} = m\vec{a} \longrightarrow \int_{t_1}^{t_2} \vec{F}\,dt = \int_{t_1}^{t_2} m\vec{a}\,dt = m\vec{v}(t_2) - m\vec{v}(t_1)$$

When there are two or more interacting masses, you get something more useful

$$\vec{F}_{\text{on 1 by 2}} = m_1\vec{a}_1, \qquad \vec{F}_{\text{on 2 by 1}} = m_2\vec{a}_2 \tag{1.9}$$

Add these, and the two forces cancel because of Newton's third law, leaving

$$0 = m_1\frac{d\vec{v}_1}{dt} + m_2\frac{d\vec{v}_2}{dt} = \frac{d}{dt}\big[m_1\vec{v}_1 + m_2\vec{v}_2\big] \tag{1.10}$$

This is conservation of momentum. It doesn't matter how complicated to two forces are as long as the third law is satisfied.

  If you have three or three million particles the result is the same, only the notation changes. Put indices i and j on the masses instead of simple 1 and 2. The total force on particle i is the sum of the forces from all the *other* particles

$$m_i\vec{a}_i = \sum_{j\neq i} \vec{F}_{\text{on i by j}}$$

Now add this equation over all values of the index i.

$$\sum_i m_i\vec{a}_i = \sum_i \sum_{j\neq i} \vec{F}_{\text{on i by j}} = 0 \qquad \text{Write this out for three masses! } \textit{Really} \tag{1.11}$$

Because acceleration is the time-derivative of velocity, this equation says that

$$\frac{d}{dt}\sum_i m_i\vec{v}_i = 0, \qquad \text{or} \qquad \sum_i m_i\vec{v}_i = \text{a constant} \tag{1.12}$$

  Can you really go from Eq. (1.9) to (1.10)? What if mass isn't constant? Then $\vec{F} \neq m\vec{a}$ anyway and you should have been using $\vec{F} = d\vec{p}/dt$. That also makes the whole process much more natural. (Go back and do it that way. It's easy.)

Angular Momentum

Another manipulation of Newton's equation uses the cross product. Pick an origin and let $\vec{r}$ be the coordinate vector of a single point mass from that origin. Notice the difference here: This result depends on having chosen an origin. It doesn't matter which origin you pick, but you have to pick one.

$$\vec{r} \times \vec{F} = \vec{r} \times \frac{d\vec{p}}{dt} = \vec{r} \times \frac{d\vec{p}}{dt} + \frac{d\vec{r}}{dt} \times \vec{p} = \frac{d}{dt}\left(\vec{r} \times \vec{p}\right) \tag{1.13}$$

What was the trick that appears after the second equals sign? The term that I added is zero because it is simply $\vec{v} \times \vec{p} = \vec{v} \times m\vec{v} = 0$. The final term is the ordinary product rule for derivatives, and the terms in this equation get names, torque and angular momentum.

$$\vec{\tau} = \frac{d\vec{L}}{dt} \qquad \text{is} \qquad \text{torque} = \text{time-derivative of angular momentum} \qquad (1.14)$$

$$\vec{L} = \vec{r}_1 \times \vec{p}_1 + \vec{r}_2 \times \vec{p}_2 + \cdots \qquad (1.15)$$

Again, with many particles, you put indices on the masses and sum over all the particles, but unlike the case of linear momentum the results are sufficiently complex that they deserve a complete chapter (eight) to themselves.

In the absence of external forces, you have the conservation law: $\vec{L}_{\text{total}} = \text{constant}$. How to derive this? The force on a single particle, the ith one, is

$$\vec{F}_i = F_{i,\,\text{external}} + \sum_{j \neq i} \vec{F}_{\text{on } i \text{ by } j}$$

This represents the total force on a single particle as caused by everything else in the universe. The index j is for all the other particles in the body. In the present case, there is no external force, and

$$\vec{\tau} = \sum_i \vec{r}_i \times \vec{F}_i = \sum_i \vec{r}_i \times \sum_{j \neq i} \vec{F}_{\text{on } i \text{ by } j}$$

Look at one particular pair of terms in this sum, the ones for which $i = 1$, $j = 2$ and $i = 2$, $j = 1$. Those are

$$\vec{r}_1 \times \vec{F}_{\text{on 1 by 2}} + \vec{r}_2 \times \vec{F}_{\text{on 2 by 1}}$$

These two forces are opposite, so this is

$$\vec{r}_1 \times \vec{F}_{\text{on 1 by 2}} - \vec{r}_2 \times \vec{F}_{\text{on 1 by 2}} = (\vec{r}_1 - \vec{r}_2) \times \vec{F}_{\text{on 1 by 2}}$$

If the forces acting between these masses act along the line between them then this cross product is zero. The same will apply to all pairs of forces. If the force is not aligned this way, then this sum is not zero, and the expression for angular momentum presented here is not conserved. See problem 4.46 for an example of this, though the resolution of the problem is left for elsewhere.

Mass
Mass is conserved. If two objects collide and they have masses m_1 and m_2, then after the collision the total mass will be the same.

$$m_1 + m_2 = m_3 + m_4$$

This doesn't mean that $m_3 = m_1$ and $m_4 = m_2$. Mass can be moved from one object to the other, perhaps by chipping off or because a loose part of one mass becomes attached to the other.

You can see why this is true if you assume that momentum conservation is valid. Measure the velocity of some mass to be $\vec{v}$ and let a friend of yours, moving by at velocity $\vec{u}$ do the same. Your friend will conclude that the velocity of the mass is $\vec{v} - \vec{u}$. If momentum conservation holds for you then it should hold for your friend. The two equations are

$$m_1\vec{v}_1 + m_2\vec{v}_2 = m_3\vec{v}_3 + m_4\vec{v}_4 \qquad \text{and} \qquad m_1(\vec{v}_1 - \vec{u}) + m_2(\vec{v}_2 - \vec{u}) = m_3(\vec{v}_3 - \vec{u}) + m_4(\vec{v}_4 - \vec{u})$$

Subtract these and the result is (for *all* $\vec{u}$)

$$m_1\vec{u} + m_2\vec{u} = m_3\vec{u} + m_4\vec{u} \quad \Longrightarrow \quad m_1 + m_2 = m_3 + m_4 \tag{1.16}$$

The idea here appears again in much more detail in chapter nine, especially section 9.11. There, in the context of special relativity you will see that mass conservation is not quite right after all.

This derivation showed that, starting from momentum conservation and looking at it from another point of view, you get a very interesting result. What if kinetic energy is conserved — a completely elastic collision — what does a different point of view say about that?

$$\frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 = \frac{1}{2}m_3v_3^2 + \frac{1}{2}m_4v_4^2$$

$$\Longrightarrow \quad \frac{1}{2}m_1(\vec{v}_1 - \vec{u})^2 + \frac{1}{2}m_2(\vec{v}_2 - \vec{u})^2 = \frac{1}{2}m_3(\vec{v}_3 - \vec{u})^2 + \frac{1}{2}m_4(\vec{v}_4 - \vec{u})^2$$

If this is true for all values of $\vec{u}$, then the coefficients of u^2 and of $\vec{u}$ itself must agree. These are

$$m_1 + m_2 = m_3 + m_4 \quad \text{and} \quad m_1\vec{v}_1 + m_2\vec{v}_2 = m_3\vec{v}_3 + m_4\vec{v}_4$$

This says that once you've assumed a completely elastic collision, you automatically get conservation of momentum and of mass. That is "automatically" if you assume that differently moving observers will all have the same basic equations for mechanics. Galileo went to great pains to argue why that is true. Look on page 298 for a quote from his text on the subject.

1.4 The Tools

Is mechanics a collection of tricks to solve variously contrived problems or is it a systematic approach to analyzing complex systems? I think it is the latter, but it often appears more like the former in introductory texts. There *are* a handful of rules that, systematically followed, will allow you to set up even very complicated problems. The resulting equations can still be hard to solve, but that's only because of the mathematics, not because of the physics.

These are mechanical systems that we're dealing with here, and that limits the sort of things you have to deal with. I will lay out a set of rules and then show by example just what they mean. The whole point is that every problem is attacked the same way. You don't learn one method now and a different method later. You will not believe me when I say this, but following these systematic rules will save you a huge amount of time.

1. Draw a sketch of the physical system. It doesn't have to be pretty, but it should convey an idea of what is happening.

2. There are typically several masses in each problem, so for each one of them list in ordinary language the things that are acting on it. A force is not a thing; an acceleration is not a thing; even an $m\vec{a}$ is not a thing. A table is a thing; a hand is a thing; a rope is a thing; the Earth is a thing. Remember: Here a thing can act on a mass only by either being in contact with the mass or by its gravitational pull.

3. For each mass, go down the list of things acting on it and draw that mass together with the vectors showing the directions of the forces exerted by those things. Also, draw an acceleration vector for each mass. Label each mass and each vector. Use symbols that convey some meaning.

4. Pick a basis in terms of which you can write the vectors. Note that you are not required to use the same basis as you turn your attention from one mass to another. You can think

of them as separate problems, and the mathematics will unite them for you. Label your coordinates.

5. For each mass write $\vec{F} = m\vec{a}$ in the basis chosen. Or, $\vec{F} = d\vec{p}/dt$ if you need the more general form.

6. Break the vector equations into equations for their various components, so that you don't have vectors to carry through the rest of the algebra.

7. Did you need to use Newton's third law? Is the magnitude of this force equal to the magnitude of that force?

8. Count the number of equations and the number of unknowns. If they match, you (may) have a winner.

9. Solve.

10. Analyze the results.

11. The most popular method to turn an easy problem into a difficult one is to skip steps.

Example

▶ Two masses are in contact and are sitting on a horizontal table. Push the left one toward the right and find their common acceleration and the forces between them. Assume zero friction.

1. Sketch:
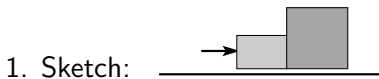
2. Acting on the left mass: 1. you, 2. table 3. Earth (gravity) 4. other mass
   Acting on the right mass: 1. table, 2. Earth 3. other mass
   Notice: *You* are not acting on the right mass.

3.
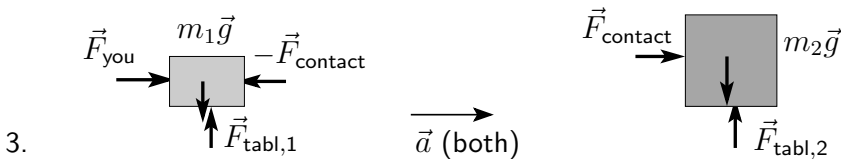
4.   The same basis for both masses

5. $\vec{F_1} = F_{\text{you}}\hat{x} - m_1 g\hat{y} + F_{\text{table,1}}\hat{y} - F_{\text{contact}}\hat{x} = m_1 a_x\hat{x}$
   $\vec{F_2} = -m_2 g\hat{y} + F_{\text{table,2}}\hat{y} + F_{\text{contact}}\hat{x} = m_2 a_x\hat{x}$

6. $F_{\text{you}} - F_{\text{contact}} = m_1 a_x$ $F_{\text{contact}} = m_2 a_x$ The y-components are of little use.

7. Newton's third law was used in applying $\pm F_{\text{contact}}$ to the two masses.

8. The unknowns are F_{contact} and a_x, and there are two equations.

9. $a_x = F_{\text{you}}/(m_1 + m_2)$ and $F_{\text{contact}} = F_{\text{you}}m_2/(m_1 + m_2)$

10. The dimensions are correct, but you do have to look.
    If $m_1 \ll m_2$ then the contact force is approximately F_{you}.
    If $m_1 \gg m_2$ then the contact force is much smaller then F_{you}.

Should you believe this? Do the experiment: Use a book and a pen or a book and a wadded-up piece of paper. Place them in contact on a table and push on one side or the other. How do you measure the contact force? Put a finger of your other hand between them. Then reverse your hands and
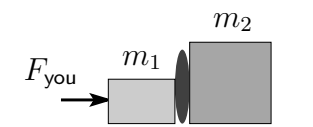
Fig. 1.3

push on the other mass to feel the contact force again.

Do you have to do this every time that you set up a problem? Yes, until the time comes that you no longer make mistakes. Then you can start skipping steps. (*I'm* still waiting.) It's also true that following all these steps is a great time-saver. You don't believe this. No one believes it at the beginning.

Example

▶ Atwood's Machine: This is a classic example of a mechanical system, a device that appears in every introductory physics book (probably by an act of Congress). You can see pictures of real versions of this apparatus at this web site of early scientific apparatus:

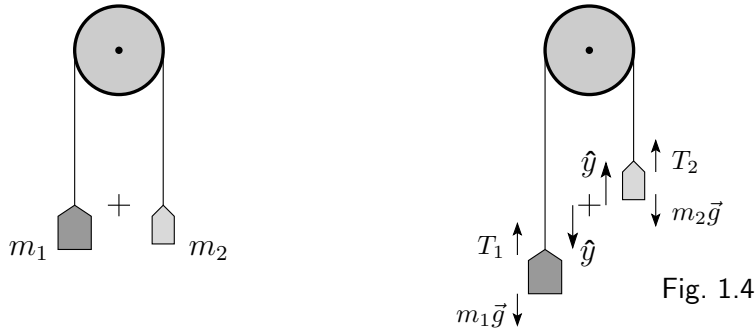physics.kenyon.edu/EarlyApparatus/Mechanics/Atwoods_Machine/Atwoods_Machine.html

Fig. 1.4

Must you execute this systematic procedure by the numbers? Maybe not. Here it is a narration, but all the steps are present:

Hang two masses on strings and run the string over a fixed pulley. The heavier mass will start to drop and the lighter one to rise, and the problem is to figure out the acceleration of either. Can I neglect the mass of the pulley? Probably not, but I will do it anyway for now. Can I neglect the mass of the string? Again, maybe not, but do it anyway this time. Later, return to the subject when you wish to add more reality to the analysis. I chose the basis with $\hat{y}$ up for one mass and with it down for the other. Is this necessary? No, but it's a convenient option because you can then easily use one coordinate y for both masses, and as long as the string behaves, a_y is the same for both.

The things that act on m_1 are gravity and the string. Acting on m_2 are again gravity and the string. The single coordinate y applies to the motion of each mass if you make the reasonable assumption that the string doesn't stretch (conservation of string). Apply Newton's law to each mass, treating the two masses as two separate problems having nothing to do with each other.

$$\vec{F}_1 = m_1\vec{g} + \vec{T}_1 = m_1\vec{a}_1, \qquad \vec{F}_2 = m_2\vec{g} + \vec{T}_2 = m_2\vec{a}_2$$
$$(1)\ F_{1y} = m_1 g - T_1 = m_1 a_y, \qquad (2)\ F_{2y} = -m_2 g + T_2 = m_2 a_y \tag{1.17}$$

These are two equations in the three unknowns (a_y, T_1, T_2), so you need another equation. That is $T_1 = T_2$. Why is this so? It comes from a torque equation. If the mass of the pulley is zero, any non-zero torque on it would give it infinite angular acceleration (Eq. (8.5)), and that can't happen. With these three equations you can solve for everything.

$$a_y = \frac{m_1 - m_2}{m_1 + m_2}g, \qquad T_1 = T_2 = \frac{2m_1 m_2}{m_1 + m_2}g \tag{1.18}$$

In analyzing this solution there are three cases that push it to its limit: $m_1 = m_2$, $m_1 \gg m_2$, and $m_1 \ll m_2$.

First: $\qquad m_1 = m_2: \qquad a_y = 0, \quad T = \frac{2mm}{m + m}g = mg$

It balances, giving zero acceleration and with just enough tension in the string to make the total force on each mass zero.

$$\text{Second:} \qquad m_1 \gg m_2: \qquad a_y \approx \frac{m_1}{m_1}g = g, \quad T \approx \frac{2m_1 m_2}{m_1}g = 2m_2 g$$

With m_1 dominant, that mass accelerates at $+g$, causing the other mass to accelerate *up* at $+g$. As m_2 experiences a downward force from gravity of $m_2 g$, this requires the string to pull up with *twice* this force, and it does.

$$\text{Third:} \qquad m_1 \ll m_2: \qquad a_y \approx -\frac{m_2}{m_2}g = -g, \quad T \approx \frac{2m_1 m_2}{m_2}g = 2m_1 g$$

Now m_2 accelerates toward $-y$ and pulls m_1 up against gravity, requiring a tension of $2m_1 g$ to do so.

Why did Atwood invent his machine and why is it important enough even to *have* a name? Perhaps he invented it as a way to harass physics students. Perhaps a better reason is as a device to verify Newton's Laws. Also as a device to measure g. In 1780, electronic timers hadn't yet been invented, and this machine slowed the motion enough to make measurements of acceleration easier. If you look at the Kenyon historical link to early scientific apparatus on page 45, you will see that Atwood's original machine does not look like the simple picture here; it has complications designed to reduce the effects of friction because good, low-friction bearings didn't exist then.

What does this have to do with energy? In chapter two you will find much more on the subject of energy, but for now I'll assume that you have seen something of the subject before and that the development of mechanical energy in the last section is familiar. You remember the gravitational potential energy to be mgh, and if you don't, then the equation (1.6) tells you that

$$F_y = -mg = -\frac{dU}{dy}, \qquad \text{so} \qquad U(y) = mgy$$

In the Atwood machine as drawn, I will assume that everything starts at rest when the masses have coordinates $y = 0$. Take the zero-point of potential energy to be zero at that point too. Later the masses will have picked up speed and their positions will have changed. Write down the total mechanical energy.

$$E = \frac{1}{2}m_1 v_y^2 + \frac{1}{2}m_2 v_y^2 + m_2 g y - m_1 g y$$

The two kinetic energies are positive and (for positive y) the m_2 gets positive potential energy and the potential energy for m_1 is negative.

Energy is conserved. That means that the derivative of the total energy with respect to time is zero.

$$E = \frac{1}{2}m_1 v_y^2 + \frac{1}{2}m_2 v_y^2 + m_2 g y - m_1 g y$$

$$\frac{dE}{dt} = 0 = \frac{d}{dt}\left[\frac{1}{2}m_1 v_y^2 + \frac{1}{2}m_2 v_y^2 + m_2 g y - m_1 g y\right]$$

$$= m_1 v_y \frac{dv_y}{dt} + m_2 v_y \frac{dv_y}{dt} + m_2 g \frac{dy}{dt} - m_1 g \frac{dy}{dt}$$

$$0 = (m_1 + m_2)\frac{dv_y}{dt} + (m_2 - m_1)g \tag{1.19}$$

The third line used the chain rule for derivatives to differentiate v_y^2 with respect to t. The fourth line used the definition of velocity, $v_y = dy/dt$, to cancel some factors. Solve the final equation for the

acceleration $a_y = dv_y/dt$ and you have the result Eq. (1.18). Or at least part of it. Using energy methods is often an easier way to get to the result you want, but it does not always give *all* the results you want. In this case it doesn't provide an equation for the tension in the string. If you need that, use $\vec{F} = m\vec{a}$ on one of the two masses and combine it with this result derived from the energy equation to find the tension.

1.5 Checking Solutions

This is not a key to all the problems in the book, but it is a key to finding out for yourself if you've made mistakes and it is a key to understanding the results that you've calculated.

If you have solved for the acceleration of the mass m_1 in the Atwood system of Figure 1.4 and you've gotten a result for the acceleration of m_1 to be $a_y = m_1 g/(m_1 - m_2)$ how can you tell if it is likely to be right?

0. Does it have the correct dimensions? mass $\cdot g$/mass is an acceleration, so this does have the right dimensions.

1. If $m_1 \gg m_2$ you expect m_1 to accelerate down, and that is positive a_y. This expression says that if m_1 is large enough $a_y \approx m_1 g/m_1 = g$, and that is right. So far, so good.

2. If $m_2 \gg m_1$ this expression goes to zero. That doesn't seem too likely.

3. If $m_1 = m_2$, this expression has zero in the denominator. It blows up. That is very unreasonable. It's wrong, so go back and find the error.

These simple procedures are the primary tools you will use to analyze your results, though other techniques will be built around them. Take another example, one that you can solve with a modest effort, but that's not what I want to do now.

Example

▶ A block of mass m is on a ramp inclined at an angle θ to the horizontal, and you are pushing it horizontally, trying to keep it from sliding. Your applied force has magnitude F_0. The box is initially moving downhill and there is a coefficient of sliding friction μ_k between the box and the ramp. What is the acceleration down the ramp?
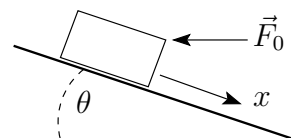

Fig. 1.5

Examine some proposed solutions and determine if any are plausible. I leave the dimension checks to you.

(1) $a_x = g[\sin\theta - \mu_\text{k}\cos\theta] - \dfrac{F_0}{m}[\sin\theta + \mu_\text{k}\cos\theta]$

    (a) If F_0 and μ_k both equal zero, this is $g\sin\theta$. That is a result you've seen before. It makes sense, going from 0 to g as θ varies from 0 to $\pi/2$, so it passes this test.

    (b) If $\theta = 0$, this is $a_x = -\mu_\text{k}g - \dfrac{F_0}{m}\mu_\text{k}$. If the friction is now zero, $\mu_\text{k} = 0$, this whole thing vanishes. This is wrong, because the force F_0 should be slowing it down. Draw a picture.

    (c) If $\theta = \pi/2$, this is $a_x = g - F_0/m$. The only thing that F_0 should be doing is increasing the normal component of the force on m, and so increasing the effect of friction. This $g - F_0/m$ is *independent* of μ_k, so that makes no sense. Again, *draw a picture.*

(2) $a_x = g[\sin\theta + \mu_\text{k}\cos\theta] + \dfrac{F_0}{m}[\cos\theta + \mu_\text{k}\sin\theta]$

    (a) This is always positive (as long as F_0 and θ are positive), but if the friction is big enough, it should decelerate the mass, making a_x negative.

    (b) Let the friction coefficient be very large. This expression for a_x says that the acceleration gets bigger (and stays positive). Friction should be slowing it down.

    (c) The bigger F_0 gets, the more positive a_x gets. That's the wrong direction.

(3) $a_x = g[\cos\theta - \mu_\text{k}\sin\theta] + \dfrac{F_0}{m}[\cos\theta - \mu_\text{k}\sin\theta]$

(a) The signs in front of the μ's are negative, tending to make the acceleration more negative. Good. That's the way it should be.

(b) If $\theta = 0$ and $F_0 = 0$, this is $a_x = g$, saying that on a level surface with no other forces, it will accelerate horizontally at a rate g. Not very likely.

(c) If $\theta = \pi/2$ and if $F_0 = 0$, this should drop with an acceleration g. Instead, this equation says $a_x = -\mu_k g$, so it is wrong.

Exercises

1 Verify the chain rule with an example constructed so that it has enough structure to show something, but is simple enough to work with easily.

$$h(t) = f\big(g(t)\big) \qquad \text{then} \qquad \frac{dh}{dt} = \frac{df}{dg}\frac{dg}{dt}$$

Take $g(t) = At^2$ and $f(x) = Bx^3$ and evaluate dh/dt two ways. First use the chain rule. Next write h explicitly in terms of t and differentiate that.

2 If your analysis of the Atwood machine had come to the conclusion for Eq. (1.18) that

$$a_y = \frac{m_1 + m_2}{m_1 - m_2}g \qquad \text{or} \qquad a_y = \frac{m_1 - m_2}{m_1 + m_2 + 1}g$$

explain why each of these is *obviously* wrong.

Problems

1.1 Derive Eq. (1.12) again, but starting from $\vec{F} = d\vec{p}/dt$, so mass may or may not be constant.

1.2 Write out Eq. (1.11) for the case of three masses.

1.3 In the deep oceans, the speed of a wave is expected to be a function of at most the gravitational field, g, the density of the water, ρ, and the wavelength of the waves, λ. Assume that the speed is given by an expression $v = g^a \rho^b \lambda^c$ where a, b, and c are unknown numbers. You know the units of all the items, v, g, ρ, λ in the equation; determine the values of a, b, and c so that the equation is dimensionally consistent.

1.4 The speed of sound in a fluid such as air may depend on the pressure, the density and the wavelength of the sound, as $v = p^a \rho^b \lambda^c$ where a, b, and c are unknown numbers. Solve for the values of a, b, and c so that the equation is dimensionally consistent. In this case there is another parameter that can enter the equation: $\gamma = c_p/c_v$ is the ratio of the specific heats, but it is dimensionless, so there's no way by looking at the dimensions to tell if or how it affects the results. In air at low frequencies it enters as $\gamma^{1/2}$, but at very high frequencies (megaHertz) it approaches one.

1.5 Solve for the acceleration for Figure 1.5, then subject your solution to the same sort of analysis shown there.

1.6 A mass m_1 hangs from a string that is wrapped around a pulley of mass M. As the mass m_1 falls with acceleration a_y, the pulley rotates. Someone claims that the acceleration of m_1 is one of the following answers. Examine each of them to determine whether any are plausible, and remember: that *does not* mean solving the problem and comparing these to your solution. The point is to explain why a proposed answer is impossible.
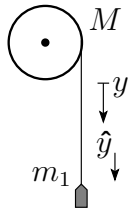
(a) $a_y = \dfrac{(M + m_1)g}{m_1}$ (b) $a_y = \dfrac{m_1 g}{m_1 - M}$ (c) $a_y = \dfrac{m_1 g}{M + 1}$ (d) $a_y = \dfrac{m_1 g}{M}$

(e) $a_y = \dfrac{Mg}{m_1 - M}$ (f) $a_y = \dfrac{Mg}{M + m_1}$ (g) $a_y = \dfrac{Mg}{m_1 + 1}$ (h) $a_y = \dfrac{Mg}{m_1}$

1.7 A mass moves in a straight line and its velocity satisfies the equation $v_x = C/x$ for some constant C. Find the force acting on this mass as a function of $x > 0$. Sketch a graph of this force. Ans: $\propto 1/x^3$

1.8 A mass m_2 is at rest. Another mass m_1, having velocity v_{1x} collides with it. Assume that all the motion is along a straight line and that both kinetic energy and momentum are conserved. Find the velocities of m_1 and m_2 after the collision. Examine special cases to see if your result is plausible. Ans: $v'_{2x} = 2m_1 v_{1x}/(m_1 + m_2)$

1.9 Two particles have the same mass m. One has zero velocity and the other has velocity $\vec{v}_0$. They collide elastically, so that kinetic energy is conserved, and they move off with velocities $\vec{v}_1$ and $\vec{v}_2$, not along a straight line this time. Write down the conservation of kinetic energy and conservation of momentum and show that the final two velocities are perpendicular to each other. No components please.

1.10 ¨(a) The same as problem 1.8, except that mass is exchanged so that the final masses are m_3 and m_4. Take $m_1 = m_2 = m$ to save on the algebra. Don't forget the conservation of mass equation. Test this in detail, looking at the effect when the final masses are the same and when

they are very different. Also, are the two equations for v_3 and for v_4 properly related to each other?
Ans: $v_3 = \frac{1}{2}v_1\left[1 \pm (m_4/m_3)^{1/2}\right]$

1.11 If two masses collide and in the process they break into three, must mass still be conserved?

1.12 Two particles have momenta $\vec{p}_1$ and $\vec{p}_2$, with coordinates $\vec{r}_1$ and $\vec{r}_2$ respectively. Write their total angular momentum, and show that if the total linear momentum is zero then the total angular momentum is independent of the origin chosen for the two $\vec{r}$'s. Recall the identity: $(\vec{A} + \vec{B}) \times \vec{C} = \vec{A} \times \vec{C} + \vec{B} \times \vec{C}$ (and draw pictures).

1.13 Read section 1.4 again. You can sometimes find an amusement park ride called a Rotor. You walk into a cylindrical room and stand against the vertical wall. The room starts to rotate, and after it has come to a sufficient rotation rate, the floor drops by a few feet. If there is enough friction with the wall, you will not slide down. What is the relation between the radius R, ω, and the coefficient of static friction μ_s ($F_{fr} < \mu_s F_N$) so that you stay up? Check the dimensions of your result and see if this minimum ω behaves in a plausible way as μ_s is changed. Recall that the magnitude of the acceleration of an object moving along a circle at constant speed is $v^2/r = r\omega^2$. Ans: $\omega^2 > g/\mu_s R$

1.14 ˙ A variation on a Rotor is a Gravitron. The difference is that the cylindrical wall is not vertical; it tilts out by an angle α (which was zero for the Rotor in the preceding problem.). What determines whether you will stay against the wall and not slide either up or down when you can no longer use the floor? ω is the angular speed of the apparatus, R is measured from the axis, and the requirement is claimed to be
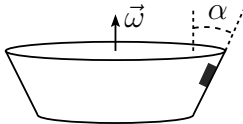
$$\frac{g}{R} \cdot \frac{1 - \mu_s \tan\alpha}{\tan\alpha + \mu_s} < \omega^2 < \frac{g}{R} \cdot \frac{1 + \mu_s \tan\alpha}{\tan\alpha - \mu_s}$$

The second inequality applies only if $\tan\alpha > \mu_s$. If that's not true then there is no upper limit on ω. Is this plausible? That is: check units; check what happens if μ_s is large or small; check what happens if α is increased or decreased; check what happens if α is negative (they don't normally do that — probably too many customer complaints); what if R is large or small; you can even vary g. See problem 4.47 for some comments on deriving these results.

1.15 Using the same stated result as in the preceding problem, what happens if you coat the wall with teflon and wear silk? *I.e.* make friction very small. How far up the wall will you come to an equilibrium? Let R_0 be the radius at the bottom and ℓ be the distance up the tilted wall. The typical Gravitron design puts this about halfway up the wall.

1.16 In the complicated-looking pulley system sketched, the masses and moments of inertia of the pulleys are negligibly small. The axle of the top pulley is attached to wall and everything is released to rotate and fall. Here are four proposed solutions for the tension in the cord that is attached to the mass m_1, that is, the magnitude of the force the cord exerts on m_1. Examine each and *explain* why it could not possibly be correct.

$$(a)\ T = \frac{m_2 m_3 g}{m_1 + m_2 + m_3}$$

$$(b)\ T = \frac{2m_1(m_2 m_3 + m_3 m_1)g}{m_1 m_2 + m_2 m_3 + m_3 m_1}$$

$$(c)\ T = \frac{8m_1 m_2 m_3 g}{m_1 m_2 + m_1 m_3 - 4m_2 m_3}$$

$$(d)\ T = \frac{m_1(2m_2 + 2m_3 - m_1)g}{m_1 + m_2 + m_3}$$
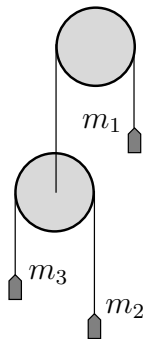
1.17 In the system as drawn, there is no friction, and the mass of the pulley is negligible. The acceleration of mass m_2 is claimed to be one of the following. They can't all be correct, so show what is wrong with most (or all) of them. That is, analyze all these proposed solutions and determine whether they are reasonable.

(a) $a_x = \dfrac{m_2 \sin \theta - m_1}{m_2 + m_1 \cos \theta} g$ (b) $a_x = \dfrac{m_2 \cos \theta + m_1}{m_2 \sin \theta + m_1} g$

(c) $a_x = \dfrac{m_2 \sin \theta - m_1 \cos \theta}{m_2 + m_1} g$
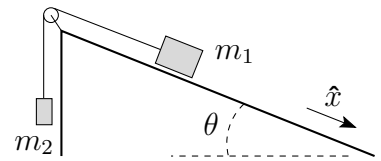
One Dimensional Motion

The simplest problems seen in a first introduction to mechanics start to become harder as soon as you add some reality. Remembering formulas for constant acceleration will get you only so far and then you come up against various forms of friction and against position- or time-dependent forces and life becomes difficult.

Toss an object straight up: The simplest model representing the force acting on it assumes that only gravity exists and that the gravitational force is a constant in time and space. In this one-dimensional case, $\vec{F} = m\vec{a}$ is (y is positive up as usual.)

$$F_y = -mg = ma_y = m\frac{d^2y}{dt^2}$$

The mass conveniently cancels and then you need to do a couple of integrals.

$$v_y(t) = \int a_y \, dt = -gt + C, \qquad y(t) = \int v_y \, dt = -\frac{1}{2}gt^2 + Ct + D$$

The two arbitrary constant C and D come from the integration and require two more equations to determine them. These are boundary or initial conditions such as

$$\left\{ \begin{array}{l} y(0) = 0 \\ v_y(0) = v_0 \end{array} \right\} \quad \text{or} \quad \left\{ \begin{array}{l} y(0) = 0 \\ y(T) = 0 \end{array} \right\} \quad \text{or} \quad \left\{ \begin{array}{l} y(0) = 0 \\ v_y(T) = v_1 \end{array} \right\} \tag{2.1}$$

In the first pair, specify the initial position and velocity. In the second pair, specify the initial position and the time you want it to return. In the third pair, specify the initial position and the velocity desired at a later time. The two unknowns C and D require two equations for you to determine them. Each of the pairs of conditions in Eq. (2.1) *is* two equations, and they give the respective results

$$\left\{ \begin{array}{l} v_y(t) = v_0 - gt \\ y(t) = v_0 t - \frac{1}{2}gt^2 \end{array} \right\} \quad \left\{ \begin{array}{l} y(t) = -\frac{1}{2}gt(t-T) \\ v_y(t) = -gt + \frac{1}{2}gT \end{array} \right\} \quad \left\{ \begin{array}{l} y(t) = -\frac{1}{2}gt^2 + (v_1 + gT)t \\ v_y(t) = -g(t-T) + v_1 \end{array} \right\}$$

You should check first that these equations are reasonable. Take some cases to see if they do what they are supposed to do. Then do problem 2.1.

What if you now remember that there's an atmosphere on Earth. What effect will that have on the motion? Usually lots. If the object is massive enough and moving slowly enough you can get away with ignoring such air resistance, but most of the time you can't. Air resistance is complicated, and you must make some simplifying model to describe it. There are a few commonly used mathematical models for air resistance, with varying simplicity and accuracy. More of the former typically means less of the latter.

Dry friction is the sort that is typical of an object sliding on a dry surface at a low speed, and a typical equation is something like $F_{\text{friction}} = \mu_k F_N$. This frictional force depends on the velocity, but only through its direction and not its magnitude. It is not even close as a way to describe air resistance.

Wet friction is a model along the lines $F_x = -bv_x$, so that it is explicitly dependent on the velocity, both in magnitude and direction. This sort of behavior is characteristic of objects sliding along lubricated surfaces or of objects moving at (very) low speed through air.

A respectable model for air resistance is $F = cv^2$. This applies to a broad range of objects as long as the speed isn't too high.

Then there are the experimentally determined functions that needed for exact work. Golf is big business, and the flight of golf balls is as well studied as the flight of any spacecraft has ever been. I would try to refer you to the experimentally determined air resistance functions for commercial golf balls, but you don't think that Titleist is going to share its data with Callaway do you?

Don't forget also that the gravitational field of the Earth decreases with altitude. Does that matter? Not for a golf ball, but does it make a difference when you fire a projectile many kilometers? Yes, and for an example, see problem 4.45.

2.1 Solving F=ma: F(t)

Time to get away from the generalities and to start solving some problems. To do that, stay in one dimension for a while so that the basic equation to solve for a single mass is

$$F_x(t, x, v_x) = m\frac{dv_x}{dt} = m\frac{d^2x}{dt^2}$$

Whether this is easy, hard, or nightmarish depends on the function F_x. Start with the simplest case, for which F_x depends on t alone, then simply do a couple of integrals.

$$F_x(t) = m\frac{d^2x}{dt^2} \implies v_x(t) = \int dt \, \frac{1}{m}F_x(t), \implies x(t) = \int dt \int dt \, \frac{1}{m}F_x(t) \qquad (2.2)$$

Each integral has a constant of integration that must be determined, so this means that you need two more equations in order to find these two constants and to finish solving the problem. Commonly these equations will come from specifying an initial position and an initial velocity, though not always, as in the examples of Eqs. (2.1).

Example
▶ Take $F_x(t) = F_0 \sin \omega t$ ($t > 0$). Assume that the mass starts from the origin and was at rest before time zero. Before doing the mathematics, what do you *expect* the mass to do? You're applying an oscillating force, as much positive as it is negative, so will m wander around the origin or will it move away? Well, which *do* you expect? Test your intuition; pre-judge the problem and say what you expect, *then* analyze the answer to check your ideas. Was your intuition correct? Of course do not discount the possibility that your intuition is right and your calculation is wrong. That can happen too.
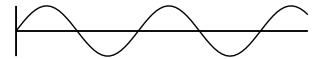
$$v_x(t) = \frac{1}{m}\int dt \, F_0 \sin \omega t = -\frac{F_0}{m\omega}\cos \omega t + C, \qquad \text{then}$$

$$x(t) = \int dt \, v_x(t) = -\frac{F_0}{m\omega^2}\sin \omega t + Ct + D$$

Apply the initial conditions.

$$v_x(0) = 0 = -\frac{F_0}{m\omega} + C, \qquad \text{so} \qquad C = \frac{F_0}{m\omega}$$

$$x(0) = 0 = -\frac{F_0}{m\omega^2}\cdot 0 + C\cdot 0 + D, \qquad \text{so} \qquad D = 0$$

Put this together and you have

$$x(t) = \frac{F_0}{m\omega}t - \frac{F_0}{m\omega^2}\sin \omega t \qquad (2.3)$$

Does this make sense? First check the dimensions. F_0 is (dimensionally) $[F_0] = \text{ML/T}^2$. This uses the fairly common notation that the square brackets denote "the dimensions of", and the individual dimensions are

$$\text{M} = \text{mass}, \qquad \text{L} = \text{length}, \qquad \text{T} = \text{time}$$

The parameter ω has to be such that ωt is an angle, so $[\omega] = 1/\text{T}$. This means that the two terms in Eq. (2.3) have the same dimensions, because dimensionally, t is the same as $1/\omega$. Now, are they the *correct* dimensions?

$$[F_0/m\omega^2] = (\text{ML/T}^2)/(\text{M/T}^2) = (\text{ML/T}^2)(\text{T}^2/\text{M}) = \text{L}$$

and the dimensions pass their test.

What is the behavior of the solution for small time? *Not zero time!* For small time use the power series expansion of the sine. Remember the equations (0.1), (b) in particular.

$$x(t) = \frac{F_0}{m\omega}t - \frac{F_0}{m\omega^2}\left[\omega t - \frac{1}{6}(\omega t)^3 + \cdots\right]$$

$$= \frac{F_0}{6m}\omega t^3 + \cdots$$

Check the dimensions again. You cannot change any dimensions by making an approximation to the result, but you can always make a mistake. This starts off as t^3. Can I understand why? For comparison I know that if it started as $a_0 t^2/2$ then this would imply that there was a force ma_0. This force function $F_0 \sin \omega t$ however starts at zero force. That implies that it has to start more slowly than t^2, and that's t^3.

What is the behavior for large time? The sine term oscillates and goes nowhere, but the t term means that it has an average drift velocity of $F_0/m\omega$. It moves away gradually. Compare problem 2.3, which replaces the sine with a cosine, and then puzzle out the differences between the two results. This graph shows Eq. (2.3); be sure to sketch a graph of the result of problem 2.3. The dashed line is the linear drift, $F_0 t/m\omega$.

x

Fig. 2.1

t
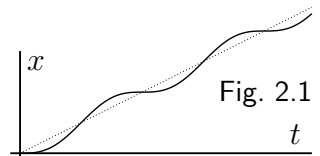
What if the time dependence of the force can't be written in a single, simple equation? What if $F_x(t)$ is a $\sin \omega t$ on Monday and a $\cos \omega t$ on Tuesday and an $e^{\omega t}$ on Wednesday? Answer: notice that the end of Monday is the beginning of Tuesday and the end of Tuesday is the beginning of Wednesday. The position and the velocity at the end of one day determine the position and the velocity at the start of the next day. You break the problem into several pieces. Can this get to be a lot of work? Of course.

Example

▶ Take a concrete example of a constant force for the time from zero to T_1 and another constant from T_1 to $T_2 > T_1$. Start from rest at the origin. I'll do this two different ways; the first one is straight-forward, but rather clumsy. The second one uses a technique that seems almost obvious after you've done it a few times, but you're not likely to invent it. It will save a lot of effort and confusion.

$$F_x(t) = \begin{cases} F_1 & (0 < t < T_1) \\ F_2 & (T_1 < t < T_2) \end{cases} \qquad\qquad (2.4)$$

F_x

x

The First Way (the clumsy one):
In the first interval

$$a_x = \frac{dv_x}{dt} = \frac{F_1}{m} \longrightarrow v_x(t) = \frac{F_1}{m}t + C \longrightarrow x(t) = \int v_x\, dt = \frac{F_1}{2m}t^2 + Ct + D \qquad (2.5)$$

Apply the initial conditions:

$$v_x(0) = 0 = C, \quad x(0) = 0 = D \qquad \text{giving} \qquad v_x(t) = \frac{F_1}{m}t \quad \text{and} \quad x(t) = \frac{F_1}{2m}t^2 \qquad (2.6)$$

all very familiar, the old $\frac{1}{2}a_x t^2 + v_0 t + x_0$.

In the second interval

$$a_x = \frac{dv_x}{dt} = \frac{F_2}{m} \longrightarrow v_x(t) = \frac{F_2}{m}t + C' \longrightarrow x(t) = \int v_x \, dt = \frac{F_2}{2m}t^2 + C't + D' \qquad (2.7)$$

Now the conditions at the start of this F_2-force are values of x and v_x at the end of the first force. The new initial conditions are the old terminal conditions.

$$v_x(T_1) = \frac{F_1}{m}T_1 \qquad \text{and} \qquad x(T_1) = \frac{F_1}{2m}T_1^2$$

Apply these equations to determine the values of C' and D'.

$$v_x(T_1) = \frac{F_2}{m}T_1 + C' = \frac{F_1}{m}T_1 \longrightarrow C' = \frac{F_1}{m}T_1 - \frac{F_2}{m}T_1$$

$$x(T_1) = \frac{F_2}{2m}T_1^2 + C'T_1 + D' = \frac{F_1}{2m}T_1^2 \longrightarrow D' = \frac{F_1}{2m}T_1^2 - \frac{F_2}{2m}T_1^2 - C'T_1$$

$$D' = \frac{F_1}{2m}T_1^2 - \frac{F_2}{2m}T_1^2 - \left(\frac{F_1}{m}T_1 - \frac{F_2}{m}T_1\right)T_1 = -\frac{F_1}{2m}T_1^2 + \frac{F_2}{2m}T_1^2$$

$$\longrightarrow \text{Surely there must be a better method!} \longleftarrow$$

The Second Way:

There is. In Eq. (2.7) you are starting a new part to the problem, but does that mean that you must keep the same coordinates as in the first part? Certainly not. Choices of coordinates are up to you and in this case the coordinate in question is t. I just automatically used the same t in the second interval as in the first interval, but instead you can start over with a new time *measured from* T_1. Call it t' and start over.

$$a_x = \frac{dv_x}{dt} = \frac{F_1}{m} \longrightarrow v_x(t) = \frac{F_1}{m}t + C \longrightarrow x(t) = \int v_x \, dt = \frac{F_1}{2m}t^2 + Ct + D \qquad \text{Eq. (2.5)}$$

The initial conditions again say

$$v_x(0) = 0 = C, \quad x(0) = 0 = D \qquad \text{giving} \qquad v_x(t) = \frac{F_1}{m}t \quad \text{and} \quad x(t) = \frac{F_1}{2m}t^2 \qquad \text{Eq. (2.6)}$$

The second part of the problem is now just as easy as the first, with $t' = t - T_1$:

$$a_x = \frac{dv_x}{dt'} \longrightarrow v_x(t') = \frac{F_2}{m}t' + C' \qquad \text{and} \qquad x(t') = \frac{F_2}{2m}t'^2 + C't' + D' \qquad (2.8)$$
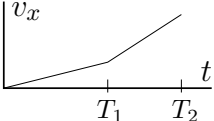
The the old terminal results are the new initial conditions:

$$v_x(t = T_1) = v_x(t' = 0) \longrightarrow \frac{F_1}{m}T_1 = \frac{F_2}{m} \cdot 0 + C'$$

$$x(t = T_1) = x(t' = 0) \longrightarrow \frac{F_1}{2m}T_1^2 = \frac{F_2}{2m} \cdot 0^2 + C' \cdot 0 + D'$$

The equations (2.8) now become

$$v_x(t') = \frac{F_2}{m}t' + \frac{F_1}{m}T_1$$
$$x(t') = \frac{F_2}{2m}t'^2 + \frac{F_1}{m}T_1t' + \frac{F_1}{2m}T_1^2$$

(2.9)

Now combine all of these results into a fairly compact set of equations. The new time coordinate is just a shifted version of the old: $t' = t - T_1$. The equations (2.6) and (2.9) become

$$v_x(t) = \frac{1}{m} \begin{cases} F_1 t & (0 < t < T_1) \\ F_2(t - T_1) + F_1 T_1 & (T_1 < t < T_2) \end{cases}$$

(2.10)

$$x(t) = \frac{1}{m} \begin{cases} \frac{1}{2}F_1 t^2 & (0 < t < T_1) \\ \frac{1}{2}F_2(t - T_1)^2 + F_1 T_1(t - T_1) + \frac{1}{2}F_1 T_1^2 & (T_1 < t < T_2) \end{cases}$$

(2.11)

> Study this version of the solution carefully. See why it is obvious. At least, see why it's obvious after you've studied it closely for a while and can explain why you should have been able to write it down directly.

When you can explain to someone else why this is evident, then you're well on your way to understanding a lot of the basic structure in the solution of complex problems.

2.2 Solving F=ma: F(v)

If the force is a function of the velocity alone,

$$F_x(v_x) = m\frac{dv_x}{dt} \qquad \text{becomes} \qquad dt = m\frac{dv_x}{F_x(v_x)}$$

This separation of variables method described in section 0.8 easily breaks the equation into two parts, each of which can be integrated.

Slide an object along a tabletop. Give it an initial speed v_0 and see how it travels. What is the frictional force? Maybe it is common dry friction, for which the magnitude of the force is approximately independent of the sliding speed. If you lubricate the surface this won't be a good approximation at all, and something such as a force of magnitude $F_{fr} = bv$ is a better approximation, with b a constant. Under other circumstances perhaps bv^2 will be a better model. How do the results of these assumptions vary, starting each case with the same velocity v_0 at position $x = 0$? There is a qualitative difference in the behaviors of these various solutions, and the only way to understand them is to grasp the result of each solution: handle it; massage it; caress it; feel how it responds. Only then can you begin to appreciate the varied behaviors that your assumptions can lead you to.

Example
▶ (a) Constant friction:

$$m\frac{d^2x}{dt^2} = -F_{fr} \longrightarrow v_x(t) = v_0 - F_{fr}t/m, \qquad x(t) = v_0 t - F_{fr}t^2/2m$$

(2.12)

This applies until the mass stops: $v_x(t) = 0$ at time mv_0/F_{fr} and for all times thereafter.

Example

▶ (b) Friction proportional to speed:

$$m\frac{dv_x}{dt} = -bv_x \longrightarrow \frac{dv_x}{v_x} = -\frac{b}{m}dt \longrightarrow \int \frac{dv_x}{v_x} = -\frac{b}{m}\int dt \longrightarrow \ln v_x = -\frac{b}{m}t + C \quad (2.13)$$

This solution used separation of variables as in section 0.8. The initial condition that $v_x(0) = v_0$ determines the constant $C = \ln v_0$, and you can combine that logarithm with the one on the left side of the equation. Next, solve for $v_x = dx/dt$ and integrate.

$$\ln \frac{v_x}{v_0} = -\frac{b}{m}t \longrightarrow v_x = v_0 e^{-bt/m} = \frac{dx}{dt} \longrightarrow x = \int dt\, v_0 e^{-bt/m} = -\frac{m}{b}v_0 e^{-bt/m} + D \quad (2.14)$$

With the same initial position as in the preceding case, $x(0) = 0$, this determines $D = mv_0/b$.

OR, you can do the same calculation more efficiently by using definite integrals to apply the initial conditions directly. When the two limits on an integral are equal, so that the integration interval is zero, then the integral is zero. That is such a simple idea that you may not immediately see what good it is. Repeat the preceding lines as definite integrals:

$$m\frac{dv_x}{dt} = -bv_x \longrightarrow \int_{v_0}^{v_x} \frac{dv_x}{v_x} = -\int_0^t \frac{b}{m}dt \longrightarrow \ln v_x - \ln v_0 = -\frac{b}{m}t \longrightarrow v_x(t) = v_0 e^{-bt/m}$$
$$(2.15)$$

When $t = 0$ the dt integral vanishes, and this in turn requires the v_x integral to vanish. That happens if the two limits there are equal, and you have then applied the initial conditions, $v_x(0) = v_0$. Next, do the integral to get x, applying the limits the same way:

$$v_x = v_0 e^{-bt/m} = \frac{dx}{dt} \longrightarrow \int_0^x dx = \int_0^t v_0 e^{-bt/m}dt \longrightarrow x(t) = -\frac{m}{b}v_0 e^{-bt/m} + \frac{mv_0}{b} \quad (2.16)$$

Example

▶ (c) Friction proportional to speed squared:

$$m\frac{dv_x}{dt} = -bv_x^2 \longrightarrow \int_{v_0}^{v_x} \frac{dv_x}{v_x^2} = -\int_0^t \frac{b}{m}dt \longrightarrow -\frac{1}{v_x} + \frac{1}{v_0} = -\frac{b}{m}t \quad (2.17)$$

Again, separation of variables handled this equation. This b is different from the one before. It doesn't even have the same dimensions. Now to get x, solve for v_x, then integrate.

$$v_x = \frac{dx}{dt} = \frac{v_0}{1 + \frac{bv_0}{m}t} \longrightarrow \int_0^x dx = \int_0^t \frac{v_0}{1 + \frac{bv_0}{m}t}dt$$
$$= \frac{m}{b}\int_0^t \frac{bv_0/m}{1 + \frac{bv_0}{m}t}dt \longrightarrow x(t) = \frac{m}{b}\ln\left(1 + \frac{bv_0}{m}t\right) \quad (2.18)$$

▶ In all three cases, start the analysis at the start. How do these solutions behave for small (not zero) time? In Eq. (2.16) for $-bv_x$ it appears that $x(t)$ blows up as $b \to 0$. Does it? The power series expansion of Eq. (0.1)(a) is what you need for this.

For $F_x = -bv_x$,

$$e^x = 1 + x + x^2/2! + x^3/3! + \cdots, \quad \text{so Eqs. (2.16) are}$$

$$v_x = v_0 e^{-bt/m} = v_0 \left[1 - \frac{bt}{m} + \frac{1}{2}\frac{b^2 t^2}{m^2} - \cdots\right] \quad \text{and}$$

$$x(t) = -\frac{m}{b}v_0 e^{-bt/m} + \frac{mv_0}{b}$$

$$= -\frac{m}{b}v_0 \left[1 - \frac{bt}{m} + \frac{1}{2}\frac{b^2 t^2}{m^2} - \cdots\right] + \frac{mv_0}{b} = v_0 t - \frac{bv_0}{2m}t^2 + \cdots$$

The t^2 term in this equation for x shows the familiar form $a_x t^2/2$, so you can immediately recognize that the initial acceleration $a_x = F_x/m = -bv_0/m$, and that says that the initial force is $-bv_0$, exactly as it should be. The apparent difficulty as $b \to 0$ in Eq. (2.16) is now gone. The series expansion not only shows that this singularity does not happen, it shows that the small time behavior (through t^2) agrees with what you would have expected if you had thought about it in advance. (Well, maybe next time...)

Equations (2.18) for the $-bv_x^2$ case call for two other series from Eqs. (0.1), the geometric series (h) and the logarithm (d):

$$1/(1+x) = 1 - x + x^2 - x^3 + \cdots \quad \text{and} \quad \ln(1+x) = x - x^2/2 + x^3/3 + \cdots$$

$$v_x = \frac{v_0}{1 + \frac{bv_0}{m}t} = v_0 \left[1 - \frac{bv_0}{m}t + \frac{b^2 v_0^2}{m^2}t^2 - \cdots\right]$$

$$x = \frac{m}{b}\ln\left(1 + \frac{bv_0}{m}t\right) = \frac{m}{b}\left[\frac{bv_0}{m}t - \frac{b^2 v_0^2}{2m^2}t^2 - \cdots\right] = v_0 t - \frac{bv_0^2}{2m}t^2 + \cdots$$

The x equation shows that the frictional force starts out as $-bv_0^2$, so the t^2 term shows the initial acceleration $a_x = F_x/m$, again because it is in the form $a_x t^2/2$. Also, x really doesn't blow up as $b \to 0$.
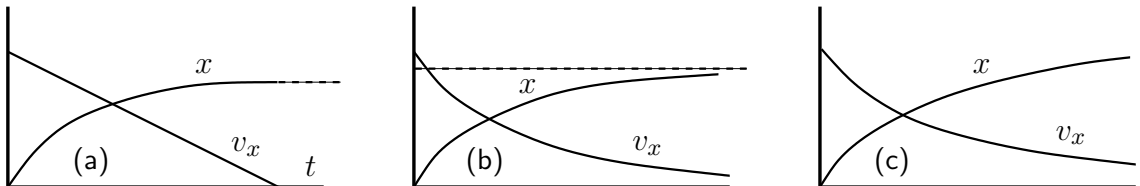

Fig. 2.2

What happens for large time?

(a) Equations (2.12) for constant friction show that the velocity reaches zero in a finite time and the position goes to a finite limit at that time.

(b) Equations (2.16) for $F_x = -bv_x$ behave differently. The velocity never reaches zero, but even as time goes to infinity the position approaches a finite value. It moves no farther than mv_0/b and that is the horizontal asymptote for x in the graph.

(c) Equations (2.18) for $-bv_x^2$ again have the property that the velocity never equals zero, but for this case the position function also goes to infinity as time increases. The x-graph keeps going up, though ever more slowly as time approaches infinity.

Why the difference? For (a), as $v_x \to 0$ the frictional force does not go to zero and it maintains its slowing effect all the way to the end. If you brake a car to a stop this way, you feel an abrupt halt, with a large jerk at the end as the acceleration drops from $-F_{\text{fr}}/m$ to zero in an instant. (Jerk is the time-derivative of acceleration, and this equation (2.12) would describe an infinite jerk.)

For (b), the friction does go to zero as the mass comes to a halt, and makes the braking more gentle. The velocity approaches zero fast enough that the integral for x converges; e^{-t} goes to zero *very* fast.

For (c), as the velocity drops toward zero the friction drops *really* fast, and as time goes on it becomes so small that it isn't even capable of bringing the mass to a halt in a finite distance. But notice that it takes a long time to get very far. Here the velocity approaches zero as $1/t$, and the corresponding integral for $x \sim \int dt/t$ is divergent.

<div align="center">

Don't skip this analysis!
Why spend all that effort to check and to analyze the solution
to a fairly simple set of problems? These are the tools that you
will use on every problem that you solve in the future. If you
ever want to publish the results of some of your work, wouldn't
it be nice if they're correct? Wouldn't you like to find your own
mistakes *before* they're in print? If you decide to become a sur-
geon, don't you want to be sure you're removing the right leg?

</div>

You're not born with intuition. That is a skill acquired only through long and careful thought — at least that's the case with *good* intuition. Most people's intuition about most things comes from casual and unconsidered experiences. A black cat passes in front of me and I then slip and skin my knee, so black cats are unlucky. I notice that when a book slides across the table it soon comes to a halt, so force is always required to sustain motion.

How do you acquire good intuition? Just solving a problem is just the start; analyzing how the result behaves is the next step. The subject of mechanics is a particularly good one on which to hone these skills because you already have ideas about how things move. You must tie those ideas to the mathematical descriptions you learn here, and then use those ties to get some intuition about the mathematics describing the real world. In the process you will probably find that some of your pictures of reality will change. Your intuition about how things move is unlikely to be 100% accurate.

When you've analyzed the results of your calculations often enough you can use the intuition that your analysis provides to solve problems without solving them! What will happen if the frictional force is

$$F_x = -F_0\left(1 - e^{-v/v_1}\right) \tag{2.19}$$

and the mass m starts with velocity v_0 as in the preceding examples? You *can* solve the problem using separation of variables, but instead, try to anticipate what the answer *will be* even before you carry out the solution.

For large v the braking force approaches the constant F_0, so in that domain the motion will look like example (a) above. For small v (what means small?) the behavior of F_x follows from a series expansion.

for small v

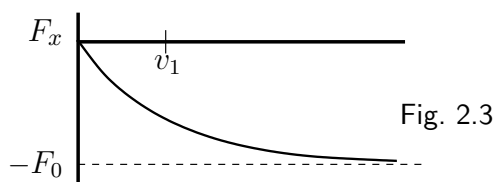$$F_x = -F_0\left(1 - [1 - \tfrac{v}{v_1} + \cdots]\right) \approx -F_0\frac{v}{v_1}$$

Fig. 2.3

and the first non-vanishing term of this looks like example (b).

Put these together and they say that if the initial velocity is well above v_1, the acceleration is constant and the velocity and position look like equations (2.12), with the velocity dropping linearly.

After the velocity has dropped to the neighborhood of v_1 or so, the braking is more like the equations (2.16), so that though the velocity never exactly goes to zero, the total distance travelled will still be finite.

2.3 Solving F=ma: F(x)

I've looked at the cases for which the force is a function of t alone and of v_x alone, now for the case that it is a function of x alone.

$$F_x(x) = m\frac{d^2x}{dt^2} = m\frac{dv_x}{dt} \tag{2.20}$$

You can't integrate this with respect to t because x isn't a constant, making it impossible to carry out the $\int F_x(x)\,dt$ until *after* you know the solution for $x(t)$. And if you already know the solution for x, why would you want to integrate it dt anyway? You can however integrate the left side of (2.20) with respect to x, and with the proper change of variables you can do the same on the right.

$$\int F_x(x)\,dx = \int m\frac{dv_x}{dt}\,dx = \int m\,dv_x\frac{dx}{dt} = \int m\,v_x\,dv_x = \frac{1}{2}mv_x^2 \tag{2.21}$$

The integral of F_x on the left defines (minus) the potential energy. Why minus? It makes later manipulation easier to interpret. Call its constant of integration E, then

$$F_x = -\frac{dU}{dx}, \qquad \text{so} \qquad -U(x) + E = \frac{1}{2}mv_x^2, \qquad \text{or} \qquad \frac{1}{2}mv_x^2 + U(x) = K + U = E \tag{2.22}$$

This constant of integration E is the total mechanical energy, kinetic plus potential, and the final equation is arranged so that the integration constant is on the right and everything else on the left. It is in the form of a conservation law. Something, $K + U$, remains constant throughout the motion. This is why there's a minus sign in defining U; the total energy is then the sum instead of the difference of two terms. Look back at section 1.3 for a more detailed presentation of this.

This is a conservation law, conservation of mechanical energy, because it says that the same mathematical expression evaluated at two different times is guaranteed to give the same value. $mv^2/2$ is the kinetic energy and U is the potential energy. This conservation of energy equation is a relation between position and velocity, and sometimes that's all you want. If you need to go all the way to the end and to find the position as a function of time, this equation is separable, as in section 0.8. Solve for $v_x = dx/dt$, then move everything involving x to one side of the equation and everything involving t to the other.

$$v_x = \frac{dx}{dt} = \sqrt{\frac{2}{m}\left(E - U(x)\right)}, \qquad \text{which rearranges to} \qquad \frac{dx}{\sqrt{2(E - U(x))/m}} = dt \tag{2.23}$$

Whether this integral is hard or easy will depend on U. At the very worst it is reduced to something that to integrate numerically, as in the example of section 3.12.

Example

▶ The absolute simplest case to start with ought to be a constant force, at least that's what you may think. Start with $y = 0$ at $t = 0$, and take the common gravitational force near the Earth's surface: $F_y = -mg = -dU/dy$. This says $U(y) = mgy$ so

$$\int_0^y \frac{dy'}{\sqrt{2(E - mgy')/m}} = \int_0^t dt' \tag{2.24}$$

Why did I start sprinkling primes around the y and t variables? Because it is poor form to use the same symbol for two different things in the same equation. y' is the variable of integration and y is the limit; t' is the integration variable and the final coordinate is the time t. If you confuse the (dummy) integration variable with one of the parameters of the problem you can cause yourself confusion and will probably make mistakes that can be very hard to find. If you look at previous pages you will find that I sometimes don't obey this rule myself even though I know better. Keep it in mind though. Perhaps you would like to go back to equations (2.15)–(2.18) and correct my notation.

The integral is easy, and if you don't show any fear, the algebra is too.

$$-\frac{1}{g}\sqrt{\frac{2(E-mgy')}{m}}\,\bigg|_0^y = -\frac{1}{g}\sqrt{\frac{2(E-mgy)}{m}} + \frac{1}{g}\sqrt{\frac{2E}{m}} = t \;\longrightarrow\; y = -\frac{1}{2}gt^2 + \sqrt{\frac{2E}{m}}\,t \quad (2.25)$$

and the last term is $v_0 t$. This is clearly the hard way to solve this problem.

Even before doing the integrals there is much more to glean from the conservation equation (2.22). You can get qualitative information even without finding a full solution, and sometimes that's more important than a detailed analytic result.

Take as an example the potential energy $U(x) = kx^2/2$, with the corresponding force $-dU/dx = -kx$ (like a simple spring). The kinetic energy as a function of x is $K(x) = E - U(x)$, and the total energy as a function of x is $E =$ constant. In this graph of U versus x you can also graph E versus x. This is the horizontal line representing the constant total energy.

This graph has several vertical lines showing the difference $K = E - U$, and they represent the kinetic energies at those values of x. As the mass moves along the x-axis its kinetic energy increases as it approaches $x = 0$, and the kinetic energy decreases after it has passed that point and moves away from the origin. As x approaches the intersection of the E and U curves, passing through the point e and headed toward f, the kinetic energy approaches zero. *At* the point of intersection f, the mass stops. Even at this stopping point there is still a force on the mass; $-dU/dx = -kx$ pushes it back toward the center and this force continues until the mass hits the other stopping point at a. From a it is pushed to the right and repeats the cycle. This example is a simple harmonic oscillator, and chapter three is devoted to it.
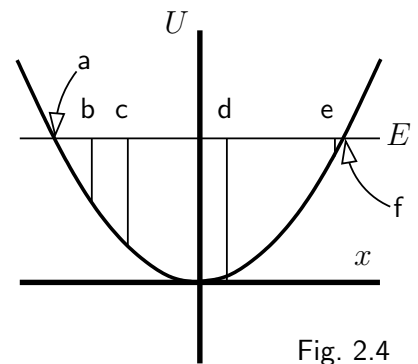

Fig. 2.4

From the graph Fig 2.4 alone, what can you say about the force function? (Yes, you already know it, but pretend you don't.) On the right, the slope dU/dx is positive, so $F_x = -dU/dx$ is negative. The same reasoning says that F_x is positive on the left. At the origin the slope is zero, so F_x is zero there. From looking at this graph of U alone can you tell that the graph of F_x is a straight line? No, just that it passes through the origin and gets more negative on the right and more positive on the left. Now remember that $F_x = -dU/dx = -kx$, so that it really is a straight line passing through the origin.

In the preceding example, $kx^2/2$, you can proceed from this qualitative analysis and find a complete solution because you can do the integral in Eq. (2.23) without too much trouble, (problem 2.15). If the force and its corresponding potential energy is more complicated this may not be practicable, but the graphical analysis is still available. Take the following potential energy (please):

U is some function of x, and however the mass moves, the combination of kinetic and potential energies stays constant. That means that just as in the simple special case above, the graph of total energy versus position is just a straight horizontal line, and $K(x) = E - U(x)$ is the vertical distance between the line E and the curve U in the graph. It can never go negative; it can however go to zero.

That's where the mass stops. Portions of the $E = $ constant graphs are solid to indicate where the mass is allowed to be ($K > 0$), and portions are dashed to indicate the forbidden regions ($K < 0$).

Fig. 2.5
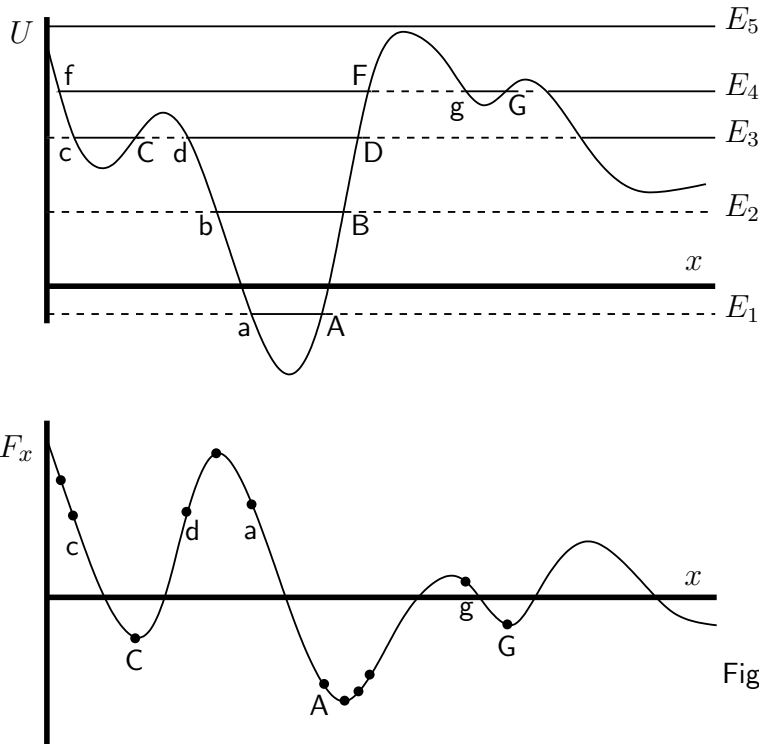
$F_x = -dU/dx$ tells you the direction of the force by reading the slope of the graph of U. This F_x-graph shows the force that you find by looking at the potential energy graph and estimating not just the sign but the size of the slope. The points at which $F_x = 0$ correspond to the minima and maxima of the potential energy, so those are easy to find. Between c and C for example there is a minimum of the potential energy; that means that the F_x curve passes through zero between those two points. Similarly, between C and d there is a maximum of U so another zero for F_x. The stopping points for the energies E_{1-4} are indicated by dots on the F_x-curve, and a few of them are labelled. For example, g and G label values of x at which $U(x) = E_4$. At that energy the mass can oscillate between these points. (At the same energy it can also oscillate between points f and F.)

$K + U = E$ shows that you can read the kinetic energy straight from the graph as the difference between the total energy E and the potential energy U, and if I know the energy then I can determine where (or whether) the mass will stop. That point is where its kinetic energy goes to zero, and there it stops. The slope of U at the same point says that it gets pushed back where it came from. For the same potential energy you can have oscillation between points a and A or between c and C or d and D or f and F or g and G, but apparently not when the energy is too big. Of course there may be more to the graph and there could be a stopping point way to the right.

In the example of E_4 between f and F, you can read the kinetic energy from the graph and tell that after it leaves f it speeds up until it passes the bottom of the curve. It then slows down but doesn't stop and then it speeds up even more before finally stopping at F. Then it returns. You can find the total period of this oscillation from f to F and back to f by doing the integral Eq. (2.23) from x_f to x_F and multiplying the result by two. The $\oint$ represents an integral over the whole cycle.

$$\text{Period } = \oint dt = \oint \frac{dx}{dx/dt} = \oint \frac{dx}{v_x} = 2 \int_{x_f}^{x_F} \frac{dx}{v} = 2 \int_{x_f}^{x_F} \frac{dx}{\sqrt{2(E_4 - U(x))/m}}$$

Example

▶ For a more interesting example, use gravity. The gravitational field of the Earth drops off with distance, the details of which are discussed in detail in chapter six. For now, say that the radial component of the force on a mass m is $F_r = -mg_0R^2/r^2$. Here the single coordinate is r, the distance from the center of the Earth, and g_0 is the gravitational field at the Earth's surface (at distance R from the center). If you fire a projectile straight up, how high will it go (ignoring air resistance)?

This is still a one-dimensional problem with coordinate r $(R < r < \infty)$, so it is the same as the equations (2.20) and (2.22) but changing the name of the variable.

$$F_r(r) = -\frac{mg_0R^2}{r^2} \qquad \text{and} \qquad F_r = -\frac{dU}{dr} \longrightarrow U(r) = -\frac{mg_0R^2}{r}$$

$$\frac{1}{2}mv_r^2 + U(r) = E \tag{2.26}$$

These equations represent firing something straight up with an initial speed v_0. They also pretend that there is no atmosphere to complicate the mathematics and that he Earth isn't rotating. How high will it go? Where is its stopping point, to use the language of this section?
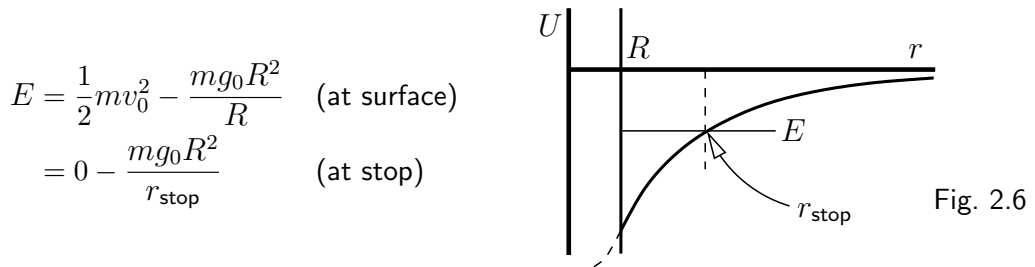
$$E = \frac{1}{2}mv_0^2 - \frac{mg_0R^2}{R} \quad \text{(at surface)}$$

$$= 0 - \frac{mg_0R^2}{r_{\text{stop}}} \quad \text{(at stop)}$$

Fig. 2.6

The graph of potential energy behaves as $1/r$ only above the Earth's surface. It has a different form inside, but you do not need that because you typically don't fire a rocket from below ground. The equations simply evaluate the energy at two points, at $r = R$ and at $r = r_{\text{stop}}$.

This conservation of energy equation gives the final height in terms of the initial speed.

$$E = \frac{1}{2}mv_0^2 - \frac{mg_0R^2}{R} = -\frac{mg_0R^2}{r_{\text{stop}}} \longrightarrow r_{\text{stop}} = \frac{R}{1 - v_0^2/2g_0R} \tag{2.27}$$

If E is large enough the horizontal line for the constant E-graph will never intersect the potential energy curve, and the projectile will never stop $(E > 0)$. Move the line of constant E upwards in this graph and watch the point of intersection with the $-1/r$ graph move to the right. Stated another way, if v_0 is increased enough the denominator in the solution for r_{stop} will go to zero and the stopping distance goes to infinity. That defines the escape speed from the planet.

STOP! Is this result for r_{stop} plausible? Does it have the correct dimensions? What is it if the initial speed is small? How big must v_0 be to escape the Earth? Compute the number, and is that number reasonable? Can you compare it to any other number you've seen? How does the number compare to the orbital speed of a satellite in near-Earth orbit? [If you don't know this number, maybe you know about how much time it takes an orbiting ship to go around the Earth.]

F(v) again

You can sometimes combine the methods of the preceding two sections by solving the problems of a velocity-dependent force as a function of x instead of t. The same trick that you saw in Eq. (2.21) to solve for forces depending on x alone also works for $F(v)$.

$$F_x(v_x) = m\frac{dv_x}{dt} = m\frac{dv_x}{dx}\frac{dx}{dt} = mv_x\frac{dv_x}{dx} \longrightarrow dx = m\frac{v_x\,dv_x}{F(v_x)} \tag{2.28}$$

Integrating this gives a relation between x and v_x. Is this more useful than getting $v_x(t)$? Usually not, but it is another tool in your utility belt, and it's sometimes just what you need. For example see problem 2.53.

2.4 Falling with resistance

A plausible example assumes that air resistance is proportional to velocity. It's not all that accurate, but it is a fairly simple place to start. A more accurate model will assume that air resistance is proportional to v^2, but this is easier.

$$ma_y = m\frac{dv_y}{dt} = m\frac{d^2y}{dt^2} = -mg - bv_y = -mg - b\frac{dy}{dt} \tag{2.29}$$

There are several ways to solve this equation. I'll use separation of variables again, moving all the v_y's to one side of the equation and all the t's to the other.

$$m\frac{dv_y}{dt} = -mg - bv_y \quad \longrightarrow \quad m\frac{dv_y}{mg + bv_y} = -dt$$

This is something that I can integrate. Use the same initial conditions as in the preceding example, $v_y(0) = v_0$, and let $u = bv_y'$

$$m\int_{v_0}^{v_y} \frac{dv_y'}{mg + bv_y'} = -\int_0^t dt = \frac{m}{b}\int_{bv_0}^{bv_y} \frac{du}{mg + u} = \frac{m}{b}\ln(mg + u)\Big|_{bv_0}^{bv_y} = \frac{m}{b}\ln\frac{mg + bv_y}{mg + bv_0} = -t$$

Solve for the velocity.

$$\frac{m}{b}\ln\frac{mg + bv_y}{mg + bv_0} = -t \quad \longrightarrow \quad v_y(t) = -mg/b + \left(v_0 + mg/b\right)e^{-bt/m} \tag{2.30}$$

Before proceeding, does this make sense? At time zero you have $v_y = v_0$ as required. (Look at both equations in Eq. (2.30) when you check.) At large time, v_y approaches $-mg/b$, and from the original differential equation of motion, Eq. (2.29), this make the acceleration zero. That is the terminal velocity.

Rewrite Eq. (2.30) directly in terms of this terminal speed.

$$v_t = \frac{mg}{b}, \qquad \text{then} \qquad v_y(t) = -v_t + (v_0 + v_t)e^{-gt/v_t} \tag{2.31}$$

To find $y(t)$ you have a simple integral with respect to time. Use the boundary condition that $y(0) = 0$ and

$$y(t) = \int_0^t dt' \left[-mg/b + (v_0 + mg/b)e^{-bt'/m} \right] = -\frac{mg}{b}t' - (bv_0 + mg)\frac{m}{b^2}e^{-bt'/m}\Big|_0^t$$

$$= -\frac{mg}{b}t + (bv_0 + mg)\frac{m}{b^2}\left[1 - e^{-bt/m}\right] = -v_t t + (v_0 + v_t)\frac{v_t}{g}\left[1 - e^{-gt/v_t}\right] \tag{2.32}$$

Do these make sense? First, check the dimensions! Do not assume that you will never make a mistake. Notice that it is easier to check dimensions now that everything is expressed in terms of v_t instead of m, b, and g. Next, see what the behavior of the solution is in some special cases. Start with the small time behavior (*not* zero time). The initial velocity of the mass should be $v_y = v_0$, and its acceleration should be $(-mg - bv_0)/m$. Why this last expression? Look at the original equation (2.29), and that

tells you the acceleration at all times, in particular the time $t = 0$. Now use a power series expansion on $y(t)$ to find the small time behavior.

$$y(t) = -\frac{mg}{b}t + (bv_0 + mg)\frac{m}{b^2}\left[1 - \left(1 - \frac{bt}{m} + \frac{(bt)^2}{2m^2} - \frac{(bt)^3}{6m^3} + \cdots\right)\right]$$

$$\approx v_0 t - (bv_0 + mg)\frac{t^2}{2m} + (bv_0 + mg)\frac{b}{6m^2}t^3$$

$$= v_0 t - \frac{1}{2m}(mg + bv_0)t^2 + (bv_0 + mg)\frac{bt^3}{6m^2}$$

$$= v_0 t - g\frac{t^2}{2} - v_0\frac{bt^2}{2m} + g\frac{bt^3}{6m} + \cdots \tag{2.33}$$

• When I multiplied out the last equation, I was careful to keep terms to a consistent order in b. That means that when I drop a b^2 term in one place I should drop it everywhere.

• Some of the terms canceled: Those that involved $1/b$. That's an important check on the algebra because if they don't cancel then everything goes to infinity as the air resistance vanishes.

• Some terms become independent of b as the viscosity approaches zero, and those are the ones that reproduce what you expect in the much simpler calculation for zero viscosity.

• The correction for small viscosity comes in the terms linear in b. These go beyond the elementary calculation, so look at the initial acceleration. That is in the t^2 terms, the ordinary $at^2/2$, so you can recognize that the acceleration has the correct initial value, though it was hidden inside a more complicated exponential expression, waiting to be dug out. The initial force is the sum of gravity (down) and the friction (also down), producing initial acceleration $a_{y0} = (-mg - bv_0)/m$.

The t^3 term is starting to get harder to interpret, but the sign is easy to understand. As the object start to rise, it starts to slow down $(-gt)$. This in turn implies that the viscous force is not quite as big as it would have been without gravity, hence this correction has a positive sign.

You're not done. What happens in the opposite case, for which the viscosity is very large? Perhaps you're firing a bullet into a barrel of honey. If b is large the exponential $e^{-bt/m}$ will go to zero very quickly. The viscous force will be much greater than the gravitational force: $bv_0 \gg mg$. In this case the solution (2.32) will be approximately

$$y(t) \approx \frac{mv_0}{b} - \frac{mgt}{b} \tag{2.34}$$

This means that it very quickly goes up a distance proportional to its initial momentum, stops at a height mv_0/b, and then slowly drifts down at its constant terminal speed mg/b. Can I see why any of this is true independent of solving the whole equation? The terminal velocity is easy to see. Just go to the original equation Eq. (2.29) and ask for the value of the velocity for which the acceleration is zero.

$$ma_y = 0 = -mg - bv_y \longrightarrow v_y = -mg/b, \quad \text{the terminal velocity}$$

For large times, the exponential $e^{-bt/m}$ dies out, and you are left with a constant velocity,

$$y(t) \approx -\frac{mg}{b}t + (v_0 + mg/b)\frac{m}{b} = -\frac{mg}{b}\left[t - \frac{m}{b}\left(1 + \frac{bv_0}{mg}\right)\right] \tag{2.35}$$

What can I say about the second term? It is positive, and that's correct, because the ball has to spend some time going up before it starts down and eventually reaches its terminal speed. The bigger the initial speed, the more time it spends going up. Did it help to rearrange the factors in the last expression and to put it in the form of something like $(t - t_0)$? Probably not, but you don't know until you try.

Still another way to squeeze information out of this result: What if the air resistance is very small? Maybe even zero? One thing that I cannot do is to let b become small in the last equations, because I was very careless in the preceding paragraph when I said "For large times..." What is large? That is meaningless because time is not dimensionless. Large compared to a femtosecond or large compared to a galactic year? What I *should* have said is that $bt/m \gg 1$, making the exponential small compared to one. Go back to the original equation (2.32), and look at that for small b. There I go again. What is "small b"? This time it is $bt/m \ll 1$, and I essentially analyzed this case in Eq. (2.33), where I said that t is small. Set $b = 0$ there, and you get $v_0 t - gt^2/2$; the correction is linear in the factor b.

Is there more that you can analyze in this simple-looking problem. Yes, see problem 2.14.

Example

▶ When solving problems, the goal is not an answer. The goal is understanding. You will not get this understanding by checking whether your answer agrees with the back of the book or if it agrees with someone else, or if a teacher says it's right. Does your solution make sense?

In section 2.2, what if (2.18) had been different, perhaps saying

$$v_x = \frac{v_0}{1 + \left(\frac{bv_0}{m}t\right)^2} \ ?$$

Would this make sense? It is dimensionally correct; at least it has the same combination of parameters $bv_0 t/m$ that appeared in Eq. (2.18), so if it was correct there then it is here too. As $t \to \infty$ this goes to zero faster than the other solution. Is that bad? Not obviously so. Before, the integral of $v_x \, dt$ went to infinity for large time because the velocity went to zero so slowly. Does that happen here? No, because

$$\int^\infty \frac{dt}{1 + t^2} \qquad \text{converges to a finite value}$$

That's different, but again, not obviously wrong. What about small time?

$$v_x \approx v_0 \left[1 - \left(\frac{bv_0}{m}t\right)^2\right] \qquad \text{for small } t. \ \text{(the geometric series, Eq. (0.1)(h))}$$

This implies that near time zero, the acceleration $\approx -2b^2 v_0^3 t/m^2 \to 0$. Even with the air resistance, $-mv^2$, assumed in deriving Eq. (2.17) this says that there is no deceleration at the beginning. That is clearly wrong because $v \neq 0$ at that time.

Example

▶ If you have a simple mechanical system with masses hanging on a string that is wrapped around a pulley, what will the acceleration be? In chapter one the equation (1.19) said

$$\frac{dv_y}{dt} = \frac{m_1 - m_2}{m_1 + m_2}g$$
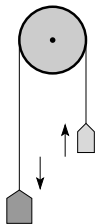
What if it had said

$$= \frac{m_1 + m_2}{m_1 - m_2}g \qquad \text{or} \qquad = \frac{m_1}{m_1 + m_2}g \ ?$$

In the first case, look at the expression for the special case that $m_1 = m_2$. The denominator vanishes, and that's almost always a bad thing. In the second case, notice that this a_y is always positive. Even though the direction of a_y should reverse if $m_2 > m_1$ changes to $m_1 > m_2$. Both of these results for dv_y/dt are wrong, and you can say this whether you know the correct solution or not.

2.5 Equilibrium

At a point of equilibrium, the total force on the mass is zero, and in one dimension that's all you need. $F_x = -dU/dx$, so a point of U at which the slope vanishes is an equilibrium point. Now, is it stable or unstable? Stable equilibrium means that if the mass is disturbed slightly from the critical point, then the force pushes it back toward the equilibrium position. If being a slight distance away from equilibrium causes it to be pushed even farther away then it is an unstable equilibrium.

                                          (2.36)
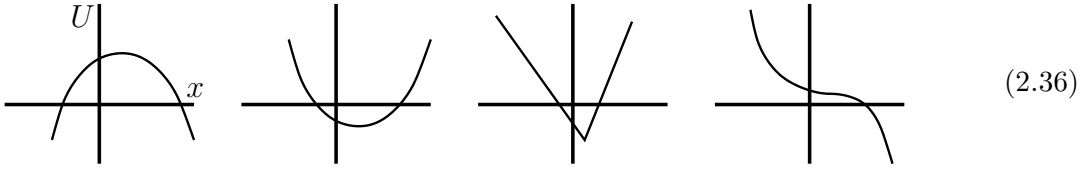
▶ The first of these curves has a potential maximum at $dU/dx = 0$, an unstable equilibrium. To the right of the maximum the force, $-dU/dx$, is positive — toward the right; to its left it's to the left.
▶ The second curve has a potential minimum as $dU/dx = 0$, a stable equilibrium because to the left the force is to the right and on the right it is to the left.
▶ The third has a potential minimum even though the derivative isn't zero there. In fact it doesn't even *exist* there. It is a stable point however. Just look at the directions of the forces on the left and the right.
▶ The fourth graph has a point with zero derivative, but it is not a stable equilibrium when you consider motions in both directions. On the left the push is back to the right, but on the right the push is *also* to the right.
▶ And the x-axis itself graphs a potential function that has zero derivative everywhere. Neutral equilibrium.
▶ See which of these curves correspond to which regions in the graph on page 61.

The graph of U on page 62 shows several equilibria. Between c and C and between a and A and between g and G there are minimum points of U. Those are stable equilibrium points. Move to the right from any of them and dU/dx becomes positive, making $F_x = -dU/dx$ negative. There are unstable equilibria between C and d and between F and g and just to the right of G. If the energy is just above that of a stable equilibrium, the motion near that point will be an oscillation around the equilibrium. The next chapter, on harmonic motion will look at that motion quantitatively.

When two atoms are bound into a diatomic molecule, the potential energy representing the force binding the atoms together is a function of their separation distance. For the example of the molecule HCl, the chlorine and hydrogen nuclei have an equilibrium separation distance of about $1.3\,\text{Å}$ ($0.13\,\text{nm}$), however there is no simple exact expression for the potential energy of the molecule as a function of the atomic separation. There is however a very good approximate function representing this energy; it was developed by Philip Morse. You can see from the graph of this function that when the interatomic distance shrinks, the force ($-dU/dr$) pushing them apart becomes very large. When the atoms are farther apart than the equilibrium position they are pulled back together *unless* they are too far apart and then the attractive force drops to zero as the slope of the graph of U becomes flat. This flattening represents the fact that the molecule can break apart if you give it too much energy — dissociation.

$$U(r) = B\left[1 - e^{-\beta(r-r_0)}\right]^2$$

Morse Potential

(2.37)
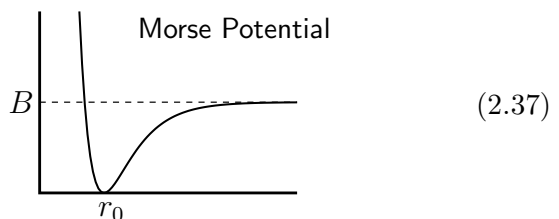
The three parameters B, β, and r_0 can be fit to the data for various diatomic molecules. For this example of HCl they are (data from T. Zielinski)

$$r_0 = 0.127\,\text{nm}, \qquad B = 7.39 \times 10^{-19}\,\text{J} = 4.61\,\text{eV}, \qquad \beta = 1.81 \times 10^{10}\,\text{m}^{-1} = 18.1\,\text{nm}^{-1} \quad (2.38)$$

Morse's function is a clever one for two reasons. First, it fits the energy data quite well with only three parameters. Second, when you get into quantum mechanics and study molecules, this choice of potential energy function leads to an equation that can be solved exactly — not easily, but it *is* possible.

2.6 Conservation of Energy

When energy is conserved it is sometimes easier to start from that conservation law than it is to start from $\vec{F} = m\vec{a}$. Differentiate Eq. (2.22) with respect to time and use $dE/dt = 0$ to state that energy is conserved. Use the chain rule a couple of times.

$$\frac{dE}{dt} = \frac{d}{dt}\left[\frac{1}{2}mv_x^2 + U(x)\right] = \frac{1}{2}m\frac{dv_x^2}{dv_x}\frac{dv_x}{dt} + \frac{dU(x)}{dx}\frac{dx}{dt} = m\,v_x\,a_x + \frac{dU(x)}{dx}v_x = 0$$

The v_x factors cancel and you have $m\,a_x = -dU/dx = F_x$, Newton's equation of motion. When problems start to get complicated, heading straight for the conservation laws is often the easiest way to get through the difficulties.

Example

▶ As an instance of this consider the Atwood machine, hang two masses over a pulley and find the acceleration of either mass. This is very easy to set up, and if you look back to chapter one in Eqs. (1.17) and (1.19), it is already there. But. Remember that the rope holding the masses has mass too. Suddenly it becomes hard just to write the equations of motion. The figure on page 45 remains the same, but that's all.

Ignore the mass of the pulley and ignore friction. The linear mass density of the rope is μ ($= dm/d\ell$). The length of the rope is L, and y is the coordinate of the top of each mass, measured up for m_2 and down for m_1. Take the potential energy to be measured from $y = 0$, then the potential energies of m_1 and of m_2 are easy. For the rope you have to divide it into three pieces. The two vertical segments and the semicircle over the top.
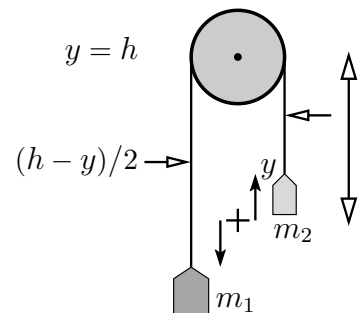
potential energy of mass m_2: $m_2 g y$

length of rope above m_2 when $y = 0$: $h = (L - \pi R)/2$

midpoint of rope above m_2 when $y \neq 0$: $(h + y)/2$

mass of rope above m_2 when $y \neq 0$: $\mu(h - y)$

potential energy rope above m_2: $\mu(h - y)\,g\,(h + y)/2$

for m_1, just change the signs of y.

$y = h$

$(h - y)/2 \rightarrow$

The total energy is

$$E = \frac{1}{2}m_1v_y^2 + \frac{1}{2}m_2v_y^2 + \frac{1}{2}\mu L v_y^2 + m_2 g y - m_1 g y$$
$$+ \mu g(h - y)(h + y)/2 + \mu g(h + y)(h - y)/2$$

The kinetic energy is clear, coming from the two masses and the rope. The potential energy for the two masses m_1 and m_2 are clear. For the rope, the length πR at the top doesn't change its position so it doesn't matter. For the two vertical segments of rope you have for m_1 and m_2 respectively

$$mgy_\text{cm} = \mu(h - y)\cdot g\cdot(h + y)/2 \qquad \text{or} \qquad \mu(h + y)\cdot g\cdot(h - y)/2$$

If you set $\mu = 0$ and differentiate this with respect to t you get the same result as Eq. (1.18).

$$\frac{dE}{dt} = (m_1 + m_2)v_y\frac{dv_y}{dt} + (m_2 - m_1)gv_y = 0, \qquad \text{or} \qquad a_y = \frac{m_1 - m_2}{m_1 + m_2}g \qquad (2.39)$$

For the more complex case in which you don't neglect the mass of the rope, do the same thing.

$$\frac{dE}{dt} = (m_1 + m_2 + \mu L)v_y\frac{dv_y}{dt} + (m_2 - m_1)gv_y - 2g\mu yv_y = 0$$

$$\text{or} \qquad (m_1 + m_2 + \mu L)\frac{d^2y}{dt^2} - 2\mu gy = (m_1 - m_2)g \qquad (2.40)$$

I'll leave the solution of this equation for chapter three, problem 3.72. If you would like to try setting up this equation starting from $\vec{F} = d\vec{p}/dt$ (not $\vec{F} = m\vec{a}$), you'll see that it is not so easy to do. Notice that even if $m_1 = m_2 = 0$, so that there are no masses hanging on the ends, you still get acceleration. Does this equation put it in the correct direction at least?

Exercises

1 What happen in Eqs. (2.10) and (2.11) if $T_1 = 0$? Is this plausible? What if $F_2 = 0$?

2 Do the dimensions in Eq. (2.14) agree? For both x and v_x.

3 Starting at Eq. (2.24) I started to be more careful to distinguish between integration variables (dummies) and parameters. What would equations (2.15) through (2.18) become with more care in writing? As it says on page 60, "go back. . . and correct my notation."

4 In Eq. (2.3) the solution $x(t)$ started off as t^3 for small time. What is the behavior of $x(t)$ near the point where $\omega t = 2\pi$ and where is this point in the graph Fig. 2.1?

5 Start from the result Eq. (2.31) and analyze the velocity for small friction. That is for large terminal speed v_t. Carry the analysis at least through terms linear in b. How does this analysis differ from analyzing the system for small times instead?

6 In Eq. (2.26) the zero point of potential energy was at infinity; the constant of integration in finding U was zero. Instead, now assume that the potential energy is zero at the Earth's surface, $U(R) = 0$. Redo the calculation leading to Eq. (2.27) and compare the answers.

7 A pendulum has a very light, rigid rod instead of a string to hold the mass at its end. (**a**) What is the potential energy of the mass as a function of the angle from the vertical? And graph it of course. (**b**) Use this to describe all the qualitatively different sorts of motion this mass can have.

8 Sketch a graph of the force function for the Morse potential energy, Eq. (2.37).

9 Write the total energy for a mass falling freely under gravity. Differentiate it with respect to t.

10 In Eq. (2.40), what is the acceleration for $\mu = 0$? More important: Analyze this answer for plausibility.

11 With the parameters stated, does the equation (2.37) really agree with the graph drawn there? Check some numbers.

Problems

2.1 For the three sets of conditions specified in Eq. (2.1), (**b**) solve for C and D and get the corresponding y and v_y, verifying (or not) the results stated in the equation following there. (**a**) *But first,* are the results stated there plausible?

2.2 ˙ The solution for a purely time-dependent force, Eq. (2.2), can be written as a single integral.

$$x(t) = x_0 + v_{x0}t + \frac{1}{m} \int_0^t dt'\,(t - t')F_x(t')$$

Differentiate this twice to verify that it works. The general framework and the reason *why* this is correct can wait until section 3.11 and problem 3.42, though deriving this particular result is problem 2.54. Read section 0.5, even if you think that you don't need to.

2.3 Repeat the example leading to Eq. (2.3) but using a force proportional to cosine instead of sine. Especially, repeat the analysis of the solution and compare the results here to those with the sine. *Explain why there is a difference from the previous result.* What is the graph that corresponds to figure 2.1?

2.4 Repeat the example leading to Eq. (2.3) but using a force proportional to sine squared instead of sine. Can you anticipate what at least some pieces of the solution should be even before you do any integrals?

2.5 A force along the x-direction is given to be F_0 for the time between zero and T. It then drops to zero in a straight line for the time from T to $2T$. (**a**) Write this function F_x using the notation of Eqs. (0.27) and (2.4). Then graph it. (**b**) A mass m starts at the origin from rest and is subject to this force. The position and velocity equations are proposed to be

$$v_x(t) = \frac{F_0}{m}t \qquad\qquad v_x = \frac{F_0}{m}T + \frac{F_0}{mT}\left[T(t - T) - \tfrac{1}{2}(t - T)^2\right]$$

$$x(t) = \frac{F_0}{2m}t^2 \qquad\qquad x = \frac{F_0}{2m}T^2 + \frac{F_0}{mT}\left[T^2(t - T) + \tfrac{1}{2}T(t - T)^2 - \tfrac{1}{6}(t - T)^3\right]$$

$$(0 < t < T) \qquad\qquad (T < t < 2T)$$

Determine if you should believe these. Try to show that they're wrong, and if you fail then maybe they're right. *First try sketching* what you expect the graph of velocity versus time should be. Then sketch the graph of this equation and compare it to yours. Then examine these to see if they behave properly. (You did check that they have the correct units didn't you.)

2.6 ˙ Derive the equations in the preceding problem, unless of course you have concluded that they are wrong. In that case, derive the correct equations. Also, assuming that the force is zero after time $2T$, what are $x(t)$ and $v_x(t)$ then?

2.7 In the example leading to the equations (2.10) and (2.11), assume that for time $t > T_2$ the applied force is zero. Compute the velocity and position for this third case and finish the graph shown there. Add a third line to each of the two equations.

2.8 For the constant force $F_y = -mg$, you already have the solution Eq. (2.25), but do the qualitative and graphical analysis as in the example following it, where $U = kx^2/2$. *Do this as if you don't already know the answer.* Use all the tools in that second analysis to verify the plausibility of this solution.

Stopping points, force, kinetic energy,... From this, find the stopping point under the initial conditions $v_y(0) = v_0$.

2.9 Instead of the specific functions of velocity as in Eq. (2.12) or (2.14), assume the friction obeys $F_x = -bv_x^\alpha$. Starting from $x(0) = 0$ and $v_x(0) = v_0$, the motion is claimed to be
(let $\beta = 1 - \alpha$, $\gamma = 2 - \alpha$)

$$v_x = \left[v_0^\beta - \beta bt/m\right]^{1/\beta} \qquad\qquad x = (m/b\gamma)\left[v_0^\gamma - \left[v_0^\beta - \beta bt/m\right]^{\gamma/\beta}\right]$$

Analyze this proposed solution to see if it is plausible. Examine dimensions, small time, large time, various ranges of α's. To analyze the special case $\alpha = 1$, recall the basic limit for the exponential function: $\lim_{n\to\infty}(1 + x/n)^n = e^x$.

2.10 Derive the result claimed in the preceding problem. (Assume $\alpha \neq 1, 2$ for reasons that should become clear in the middle of the calculation.)

2.11 A mass is initially moving at a velocity $v_x(0) = v_0 > 0$, starting from the origin. The frictional force is $-F_0 e^{kv}$ ($v > 0$). The position is claimed to be

$$x = v_0 t - \left[(1 + \gamma t)\ln(1 + \gamma t) - \gamma t\right]/k\gamma$$

where $\gamma = e^{kv_0}kF_0/m$. Analyze the results (dimensions, small time, large time, small and large k) to see if the claim is plausible.

2.12 Derive the result claimed in the preceding problem.

2.13 Start from the form of Eq. (2.26) and solve for dt. Set up the integral to find the relation between r and t, then carry out the integral in the special case that $E = 0$. The integrals aren't so bad in this case and you can spend your time analyzing the result.

2.14 You thought I spent enough time on analyzing the results of Eq. (2.30) and (2.32) but no. **(a)** Find the maximum height to which the projectile rises. **(b)** For small air resistance your result looks like it is going to infinity. Find its behavior for small b and see if it is plausible. **(c)** Approximately what is this result for large air resistance? That is, what terms dominate?
Ans: $y_{max} = -\frac{v_t^2}{g}\ln\left(1 + \frac{v_0}{v_t}\right) + \frac{v_t v_0}{g} \approx (v_0^2/2g) - (bv_0^3/3mg^2) + \cdots$

2.15 After Eq. (2.23) there is a qualitative analysis of the potential energy $U(x) = kx^2/2$. Now do the integral in Eq. (2.23) to get t in terms of x (a trig substitution) and then solve for $x(t)$. Compare your solution to the qualitative description in the text.

2.16 **(a)** Analyze the preceding quadratic potential energy qualitatively as in the text, but turn it upside down first. Now $U(x) = -kx^2/2$. **(b)** For the example in the text, I did not consider the cases for which $E < 0$. Why not? What about doing so for this new potential? For the quantitative analysis of this problem, see problem 2.56.

2.17 Answer all the questions presented immediately after Eq. (2.27).

2.18 A particle of mass m moves along the x-axis. Its potential energy is given by $U(x) = ax^3 - bx$, where $a > 0$ and $b > 0$. **(a)** What is the force F_x on this particle? **(b)** Draw U. **(c)** Find all equilibria and say if they are stable or unstable (and why). **(d)** If the mass is at the minimum of potential energy, what kinetic energy must it have to become unbound?

2.19 ˙A particle of mass m is subject to the force as specified by the potential energy $U(x) = U_0|x|/a$ where a is a (constant) length and U_0 is a positive constant having the dimensions of energy. The mass has a total energy E. What is the period of oscillation? The amplitude? How does this period vary with the amplitude of the oscillation? Sketch a graph of T versus E. Suggestion: Write $|x|$ as two cases: $x > 0$ and $x < 0$. Ans: $T = 8(a/U_0)\sqrt{mE/2}$ You *did.* draw a graph didn't you?

2.20 ˙When you brake an automobile to a halt, *none* of the velocity-dependent forms of friction in section 2.2 apply. The frictional force between the brake pads and the brake disk may be dry friction as in Eq. (2.12), but you vary the force with your foot. A smooth stop requires no big jerk at the end — that's $j_x = da_x/dt$. You are going from v_0 to a stop in a distance L. Devise a force function, $F_x(t) = ma_x$ that does this with no big jerk at the end. Also no big jerk in the middle. (It's "big" if you try to differentiate a step function.) Suggestion: Graphs, lots of graphs. Figure out what sort of graph would accomplish this and then construct a function that matches it.

2.21 A particle with mass m is subjected to a force $F_x(x) = -F_0x^3/a^3$. At time $t = 0$ the particle is at rest at position x_0 **(a)** Find $v_x(x)$, draw the potential energy graph in the process. **(b)** Set up the integral that you would have to do to go from the result in (a) to get the relation between x and t. If you find that you can do the integral easily, go back and find your mistake, starting with the units of course.

2.22 When a mass m is oscillating within a potential energy function $kx^2/2$, as on page 61, you can try assuming the simplest functional form for the motion, $x(t) = A\cos\omega t$. Do not assume that you have any knowledge of A or ω. Compute the total mechanical energy $mv^2/2 + kx^2/2$ for this assumed x. This $E(t)$ will be a function of time, but it is not supposed to be. Find the value of ω so that E is independent of t. This is easier than solving a differential equation.

2.23 If an elastic ball is bouncing on the floor, the potential energy curve is like the third one at Eq. (2.36). A straight line mgy for $y > 0$ and jumping almost straight up at $y = 0$. Assume the ball's total energy is E and use the methods of section 2.3 to find the total time between bounces. Ans: $(1/g)\sqrt{2E/m}$

2.24 Find $x(t)$ if $F_x = F_0e^{ct}$, given $x(0) = 0$ and $v_x(0) = 0$. (Is it necessary to say that you should analyze the result?) Ans: $\frac{F_0}{mc}\left[\frac{1}{c}\left(e^{ct} - 1\right) - t\right]$

2.25 Derive the equation (2.11) using the result of problem 2.2.

2.26 In the problem leading to Eq. (2.11), the force F_2 is turned off after time T_2. What are v_x and x for all times after that?

2.27 ˙For the force in Eq. (2.19), with initial velocity v_0, find the velocity and position as a function of time. Analyze the results and compare this analysis to that done in the paragraphs just after that equation.

2.28 During a bungee jump, what is the maximum force that you will feel? You will have to make (and explain) some plausible assumptions that you must make about the bungee cord. It is quite elastic, with a force when stretched that is proportional to the stretch. Is the cord tied so that it exerts a force on you from the moment you start to drop? Is it tied so that you have a distance of free fall before the cord is straightened out? Check out Wikipedia for information on bungee jumping.

2.29 ˙The assumption about air resistance made in Eq. (2.29) isn't very good, as the dependence on velocity is more nearly quadratic than linear. Take the same initial conditions as there but assume the

air resistance is $-bv_y^2$ on the way up, so $ma_y = -mg - bv_y^2$. A result claimed to be true for the velocity during the upward journey is

$$v_y(t) = v_t \frac{v_0 - v_t \tan(gt/v_t)}{v_t + v_0 \tan(gt/v_t)}$$

where v_t is the terminal speed, $mg = bv_t^2$. (a) Analyze this claimed result to determine its plausibility: small time, large time, dimensions, graphs, *etc.* (b) Show that if $v_t = 100\,\text{m/s}$ (typical for a 30 caliber bullet) then no matter how big the initial speed is the mass will stop in less than about 16 seconds. Does the same statement apply to Eq. (2.30)?

2.30 ˙ Derive the result claimed in the preceding problem.

2.31 ˙ In problem 2.29, if the mass is moving *down* then the air resistance is $F_y = +bv_y^2$, not minus. You have to handle this case separately because the form of the equation is different, and there's not a neat analytical way do both cases at once. Solve for the velocity if the mass starts from rest and drops.

2.32 A boat is slowed by a frictional force $F_x(v)$. Its speed decreases according to the equation $v(t) = c^2(t - t_1)^2$ where c is a constant and t_1 is the time at which it stops. Find the force $F_x(v)$ as a function of v. $(t < t_1)$ Ans: $-2mcv^{1/2}$

2.33 ˙ If a mass is bouncing up and down elastically on a rigid floor, the probability of seeing the mass within the interval y to $y + \Delta y$ is proportional to the amount of time it spends in that interval. (If you look at it only occasionally, the probability of seeing it somewhere is proportional to the time it spends where you look.) Stated mathematically, the probability dP to see it in the interval dy is $C\,dt$, where dt is the time spent there and C is some constant. The probability density is defined as dP/dy, so this is $C\,dt/dy = C/v$ up to y_{max}. (a) Compute this as a function of y for a bouncing ball that has total energy $E = mgy_{\text{max}}$, and the floor is at $y = 0$. The total probability must be one: $\int dP = 1$. This lets you evaluate C and then graph dP/dy versus y. Why does the graph have the shape that it does, and does it make any sense? Do the dimensions of dP/dy make sense? Examine the simulation on the right and compare it to your result. And when in doubt, remember that physics is an experimental science so you may want to try bouncing a tennis ball yourself. (b) Just as the total probability is one $(\int dP = 1)$ the mean (or average) value of the height is $\int y\,dP = \int_0^{y_{\text{max}}} y(dP/dy)\,dy$. Evaluate this. (c) Also, what is the mean value of 1? Ans: (b) $\frac{2}{3}y_{\text{max}}$ Cf. the video at `leaderboard.com/glossary_liquidmetal`

2.34 ˙ In the preceding problem, you have 100 balls all bouncing up and down randomly (but all with the same energy). (a) At any one time, if you take a photograph of all of them at once, approximately how many of the balls would be at a height from the ground up to one-half the maximum height? (b) The median height is that height so that one-half the balls are below it and one-half are above. What is that median for these bouncing balls and how does it compare to the mean value found in the preceding problem? There are 100 dots in the simulation on the right, so you may want to check your answers against that. Ans: (a) 29%, (b) $y_{\text{med}} = \frac{3}{4}y_{\text{max}}$

2.35 A force acts on m, with $F_x = Av^2/x$. Find $x(t)$ if $x(0) = x_0 > 0$ and $v_x(0) = v_0 > 0$.

2.36 A particle of mass m is initially at rest at a coordinate $x = x_0$. It is repelled from the origin by a force $F_x = A/x^3$. Solve for $x(t)$ and $v_x(t)$. Ans: $x = \left[x_0^2 + At^2/mx_0^2\right]^{1/2}$

2.37 A mass m moves in one dimension and subject to a variety of forces and initial conditions. In each case, does the mass stop in a finite time or distance as the case may be?
[a] $F_x = -\alpha/t$ for $t > 0$. At time $t = t_0$, the velocity is $v_x = v_0 > 0$; find $v_x(t)$.

[b] $F_x = -\alpha/v_x$ for $v_x > 0$. At time $t = t_0$, the velocity is $v_x = v_0 > 0$; find $v_x(t)$.
[c] $F_x = -\alpha/x$ for $x > 0$. At position $x = x_0$, the velocity is $v_x = v_0 > 0$; find $v_x(x)$.

2.38 The force on a mass m is $F_x = kv_x x$ where k is a positive constant. At time $t = 0$ the mass has $x(0) = 0$ and $v_x(0) = v_0$. Find $x(t)$.

2.39 The coefficient of friction between a mass m and the horizontal table on which it rests is μ_k. The air also provides a frictional force proportional to velocity. The mass is given an initial velocity v_0. Find $v_x(t)$. Find $x(t)$.

2.40˙ Air resistance depends on the density of the air, and there is less at greater altitude. Assume that air resistance is governed by Eq. (2.29), bv^2, and that b is proportional to air density. A reasonable model for air density is $\rho(y) = \rho_0 e^{-y/h}$, where ρ_0 is the density at sea level and h is about 8 km. When you jump from a plane in free fall you reach terminal speed fairly fast $(mg = bv_t^2)$. **(a)** What is the terminal speed as a function of y, expressed in terms of the terminal speed at sea level, v_0? **(b)** Assume that you are falling at this varying terminal speed all the way down, then what is the time to reach sea level given that you start at height H? **(c)** Someone jumps from 36 km up and doesn't open a parachute until almost at the surface, how much time does it take? Terminal speed at low altitude is about 200 km/hr if you are falling prone. **(d)** How much time would it have taken if you used the sea-level terminal speed for the entire drop? **(e)** Explain why the approximations here are unreasonable all the way to the top for this 36 km fall. **(f)** What time do you get for a 10 km fall? Ans: **(a)** $v_0 e^{y/2h}$, **(c)** $(2h/v_0)(1 - e^{-H/2h})$

2.41 A potential energy us given by $U(x) = (a/x^2) - (b/x)$, where a, b, and x are positive. Analyze the possible motions qualitatively. For what range of energy is the motion periodic? For what range is it not?

2.42˙ For the potential energy of the Earth, as described in Eqs. (2.26)-(2.27), how much time does it take to go from the surface, $r = R$, to the stopping point. As usual, what is this result in limiting cases: escape or small (not zero) v_0.

2.43 A wheel of radius R is rotating about a horizontal axis. The angle through which it has rotated is $\phi(t) = At^3$ where A is a constant and ϕ is measured in radians from the bottom. When the point that was at the bottom at time $t = 0$ has gone around N times and then comes up to the position directly to the right of the axis, a small piece of the wheel at that point breaks off and flies straight up. Use the axis of the wheel as the origin and find the displacement vector from the origin to the position of this piece at the time it reaches its maximum height.

2.44 The more general form of Newton's equation of motion is $F_x = d(mv_x)/dt$, not $= ma_x$. An open railroad car of mass M is traveling in a straight line, starting with an initial velocity v_0. It rains straight down, causing water to fill the car at a rate $dm/dt = \beta$. Find the car's velocity and the position as a function of time. Assume no friction. When you analyze the solution, one question to ask is: Why is the initial acceleration proportional to $-\dot{M}$? Ans: $x = (Mv_0/\beta)\ln(1 + \beta t/M)$

2.45¨ A uniform chain is hung vertically with its bottom point just touching a table. Release the chain, letting it drop onto the table. Compute the force the chain exerts on the table as this happens. Show that as the chain drops, this force is three times the weight of the chain that has already come to rest on the table. There are exactly two things acting on the chain: gravity and the table. One approach is to figure out what the motion will be and then take the time derivative of the chain's momentum. Another approach is to draw careful pictures at the times t and $t + \Delta t$ and to see what has happened between the two times. Remember: $\vec{F} \neq m\vec{a}$; it $= d\vec{p}/dt$.]

2.46 In chapter nine, section 9.13 you will find an analysis of a relativistic effect that causes dust particles orbiting the Sun to lose energy and to spiral gradually inward, eventually either evaporating or hitting the Sun. The equation describing the orbital radius is $r(t) = \sqrt{r_0^2 - 2\alpha t}$ where α is a parameter whose value, expressed in Astronomical Units (AU) and years is numerically $1/2800$. r_0 is the initial orbital radius. (This is for a $1\mu m$ radius particle.) An Astronomical Unit is 150×10^9 km, which is very close to the distance from the Sun to Earth. (**a**) What are the units of α? (**b**) What is the inward component of velocity as a function of t? of r? (**c**) Starting from an orbital distance of 1 AU, how much time will it take to hit the sun?

2.47 A mass m is dropped with initial velocity zero, and the forces on m are from gravity and some form of air resistance. Measure the coordinate y positive downward from the starting position. Someone got the result

$$y(t) = \frac{v_t^2}{g} \ln \left[\sinh \left(\frac{gt}{v_t} \right) \right]$$

(**a**) What is this position function for large time? Is the behavior correct? (**b**) What is this function for small time, keeping enough terms that you may be able to surmise what the assumed form for the air resistance force is? Are these results plausible?

2.48 Someone else got a different answer for the preceding problem:

$$y(t) = \frac{v_t^2}{g} \ln \left[\cosh \left(\frac{gt}{v_t} \right) \right]$$

In the same way, also analyze this solution for large and for small times, and try to determine what the air resistance force is. Determine if this solution is more or less plausible than the other solution.

2.49 For the potential energy function $U(x) = U_0 \left(x^2/a^2 - \alpha \cos(x/a) \right)$, analyze the possible motions. ($U_0 > 0$.) For example, if α is small (≈ 0.1) what motion can occur? Or if it's big (≈ 10)? Without lots of graph sketching, this is hopeless. With careful sketching, it's easy.

2.50 A molecule more complicated than HCl has a potential energy function more complicated than the Morse potential of Eq. (2.37). It is

$$U(r) = B \left[1 - e^{-\beta(r-r_0)} \right]^2 - \epsilon B e^{-2\beta^2 (r-r_1)^2}$$

where $\epsilon = 0.05$ and $r_1 = 3r_0$. Describe the behavior and possible equilibria of this molecule. For these equilibria, and using the numbers in Eq. (2.38), how much energy would it take to kick the molecule out of one equilibrium state and into the other? How much to kick it back? Look up the term "metastable". If you find this problem hard, that means that you didn't sketch a careful graph.

2.51 A simple pendulum is a point mass m on the end of a string of length ℓ. The potential energy of the mass is mgh measured from whatever height you choose. Write the total mechanical energy and compute dE/dt, setting it equal to zero to derive the differential equation of motion. Express everything in terms of the angular coordinate. And how does your result change if you change the origin from which you measure height?

2.52 Use the same frictional force as in problem 2.12, but now the mass starts from rest and drops because of gravity. Assume $F_0 < mg$.

2.53 Fire an object straight up at initial speed v_0, and assume that the air resistance is proportional to the square of the speed. Use the method of Eq. (2.28) to find (**a**) how high the object will go and (**b**) its speed when it returns to the original height. Express the results using v_t, the terminal speed, instead of the other parameters.

2.54 (a) Take the triangle in the x-y plane to the right of $x = 0$, below $y = a$, and left of $x = y$. Compute its area by a double integral $dx\,dy$. Then interchange the order of integration and compute the area again as another double integral. If you do this without thought and without carefully drawing a picture of the sums you are taking, you may by mistake get the right answer but you will be unable to do the next part of the problem.
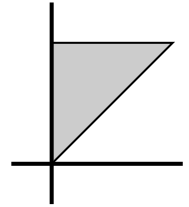
(b) If you think you did part (a) correctly then the corresponding problem for the area bounded by $y = x^2$ ($0 < x < a$) should be easy. If not, then you probably didn't do part (a) correctly and you'd better do it again.
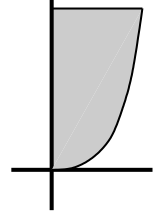
(c) Take the equation (2.2) for $x(t)$ and rewrite it as a definite integral in the spirit of Eqs. (2.15) and (2.16). (Start over from $F_x = mdv_x/dt$ to do this.) Interchange the order of integration in the double integral to derive the result of problem 2.2. Reread the comment after Eq. (2.24) concerning using the same symbol for two different things in the same equation. If you attempt to interchange the order of integration without drawing a picture of the domain of integration you will most likely not succeed.

2.55 A pencil, 20 cm long, is stood on its point, or almost so. (a) If it starts with zero velocity at the small angle θ_0 from the vertical, write the total energy as at falls over. Then separate variables to get a well-defined definite integral to determine how much time it will take to fall to the horizontal surface. (b) This is a hard integral, so to estimate it, assume that the angle θ as measured from the vertical is small all the way up to $\pi/2$, so that $\sin\theta \approx \theta$ and $\cos\theta \approx 1 - \theta^2/2$; evaluate the time to fall and evaluate it numerically for initial angles of $1°$, $10^{-3°}$, and $10^{-6°}$. Even though this approximation looks crude, it gives pretty good results. That's because for most of the duration of the fall (*i.e.* when it's just starting out), the angle *is* small.

2.56 The equation (2.23) lets you get time as a function of how far you've gone. (a) Apply this to the potential energy $U(x) = -kx^2/2$, starting at $x = x_0 > 0$ with velocity zero to get $t(x)$. (b) As $x \longrightarrow \infty$, is this finite? That is, does the integral for the time versus the distance converge? (c) What is the behavior of t versus x (or better: x versus t) for small time, and is it right? (d) More generally, for what values of the exponent in $U(x) = -kx^a/2$ ($a > 0$) will that time to reach infinity be finite and for which values is it infinite? Look to section 0.2 for an appropriate substitution to do the integral. Ans: (a) $t = \frac{1}{\omega}\cosh^{-1}(x/x_0)$, $\omega = \sqrt{k/m}$

2.57 For small speeds, constant acceleration from rest gives $x = at^2/2$. Include relativity and the result has the form $x = A + (c^2t^2 + B)^{1/2}$, where c is the speed of light. (a) Find the values of A and B so that this agrees with $x = at^2/2$ for small time. (b) Graph both x's versus t. Ans: see Eq. (9.27).

Simple Harmonic Motion

.

Why have a whole chapter devoted to the simple harmonic oscillator? Just because it is easy maybe? No, in fact some of the developments in this one subject will take some hard thought to understand. The reason for the emphasis on this one problem is that it is ubiquitous. Every time that you look at a new problem there's probably a harmonic oscillator hidden somewhere in it. Certainly any question about equilibrium will reveal an oscillator lurking in the background.

The second reason is of course: because it *is* simple. It doesn't have all the complexities that occur in other problems, but it does require using many of the tools that these other, more complex problems entail.
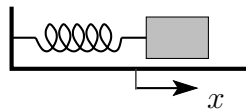
Fig. 3.1

3.1 Simplest Case

The first instance of an oscillator is a mass attached to the end of a spring. Assume that the mass is sliding on a horizontal table and that there's no friction. The force that a spring applies is, to a good approximation and for small distances, $F_x = -kx$, where the coordinate x is measured from the point of equilibrium.

$$F_x = -kx = ma_x = m\frac{d^2x}{dt^2} \tag{3.1}$$

This is a simple differential equation for $x(t)$ and there are several ways to solve it. The first and easiest is to guess the solution. You can't do this very often, so take advantage of it when you can. What functions do you know whose second derivative is proportional to themselves? A moment's thought and you can easily come up with sines, cosines, exponentials, and that's about it.

The way to find out if a guess is right is to try it. Does it satisfy the equation? $\cos t$ doesn't work. $\sin t$ doesn't work. e^t doesn't work. Try the cosine.

$$-k\cos t \overset{?}{=} m\frac{d^2\cos t}{dt^2} = -m\cos t \qquad \text{No.}$$

It has the right form of function. It has the right sign. It doesn't have the right constants. You can see even before plugging it into the differential equation that it's wrong; it doesn't even have the right *dimensions*. What is the cosine of a day? Also, the output of the cosine is dimensionless, but this $x(t)$ isn't. You need to sprinkle some constants around to take care of the units if nothing else.

Try $x(t) = A\cos\omega_0 t$ instead.

$$-kA\cos\omega_0 t \overset{?}{=} m\frac{d^2 A\cos\omega_0 t}{dt^2} = -mA\omega_0^2 \cos\omega_0 t$$

The cosines match. The signs match. All that's left are the constants, and A better not be zero or it's a trivial solution, where nothing is happening.

$$kA = mA\omega_0^2 \qquad \text{implies} \qquad \omega_0^2 = k/m \qquad \text{or} \qquad \omega_0 = \sqrt{k/m}$$

For this value of the constant ω_0 you have a solution. It says nothing about A. Will $x(t) = A\sin\omega_0 t$ work too? Yes, just try it. How about $x(t) = Ae^{\alpha t}$? Try this and you have

$$-kAe^{\alpha t} \stackrel{?}{=} m\frac{d^2 Ae^{\alpha t}}{dt^2} = mA\alpha^2 e^{\alpha t}, \qquad \text{which implies}$$

$$\alpha^2 = -k/m \qquad \text{or} \qquad \alpha = \pm\sqrt{-k/m} = \pm i\sqrt{k/m} = \pm i\,\omega_0 \tag{3.2}$$

All of these are correct, and there's more about the last one in a few pages.

For a second order differential equation there are two arbitrary constants in the complete solution. You are undoing two derivatives, and each gives a constant of integration. Also, you need two constants in order to be able to specify an initial position and an initial velocity. The solutions can come in several different forms, noting specifically that for this equation the sum of two solutions is a solution.

$$-k\big(x_1(t) + x_2(t)\big) = m\frac{d^2(x_1 + x_2)}{dt^2}$$

If x_1 and x_2 satisfy this equation separately then their sum does.

$$x(t) = A\cos\omega_0 t + B\sin\omega_0 t = Ce^{i\omega_0 t} + De^{-i\omega_0 t} \tag{3.3}$$

Other equivalent forms are

$$x(t) = F\cos(\omega_0 t + \delta_1) = G\sin(\omega_0 t + \delta_2) \tag{3.4}$$

These are all different forms of the solution and they all equal each other if the coefficients are arranged appropriately. This means that you can use whichever of these is the most useful for a particular problem: most often, one or the other of Eq. (3.3). For example, to see if the first and the fourth of these agree,

$$A\cos\omega_0 t + B\sin\omega_0 t = G\sin(\omega_0 t + \delta) = G\big(\sin\omega_0 t\cos\delta + \cos\omega_0 t\sin\delta\big) \tag{3.5}$$
$$\text{which requires} \qquad A = G\sin\delta, \qquad B = G\cos\delta$$

That last step is simply the observation that Eq. (3.5) has to be correct for *all* values of t. That's really an infinite number of equations, so the respective coefficients of $\cos\omega_0 t$ and of $\sin\omega_0 t$ must match. Now you have two equations for the two unknowns G and δ, and to solve them, take the square root of the sum of the squares of the two equations and also take the quotient of the equations.

$$G = \sqrt{A^2 + B^2} \qquad \text{and} \qquad \tan\delta = A/B \tag{3.6}$$

The relations between the other forms of solution in (3.3) and (3.4) also follow with a little algebra, problem 3.1.

Is guessing the solution somehow unsatisfactory? Maybe there's another way to get the answer, one that's more difficult and that will make you feel that you've done enough real work that to be comfortable using the result. Yes, start from the defining equation and integrate it with respect to x, getting the energy integral as in sections 1.3 and 2.3. An abbreviated form of those sections is

$$m\frac{dv_x}{dt} = -kx \implies m\int dx\,\frac{dv_x}{dt} = m\int v_x\,dv_x = \int -kx\,dx \implies \tfrac{1}{2}mv_x^2 = -\tfrac{1}{2}kx^2 + E$$

where E is an arbitrary constant of integration. This equation, which is Eq. (2.22), can be rearranged in order to apply the method of separation of variables as in section 0.8. Solve for v_x

$$v_x = \frac{dx}{dt} = \sqrt{\frac{2E}{m} - \frac{k}{m}x^2}$$

This equation separates; move all the x's to one side and all the t's to the other.

$$\frac{dx}{\sqrt{\frac{2E}{m} - \frac{k}{m}x^2}} = dt$$

and this is a standard integral, easily done by a trigonometric substitution: $x = A \sin\phi$ then $dx = A\cos\phi\, d\phi$

$$\int \frac{A\cos\phi\, d\phi}{\sqrt{\frac{2E}{m} - \frac{k}{m}A^2\sin^2\phi}} = t - t_0$$

Now choose A so that the denominator simplifies. Make the factors in the two terms match: $A^2 = 2E/k$, and the integral is now

$$\frac{\sqrt{2E/k}}{\sqrt{2E/m}} \int \frac{\cos\phi\, d\phi}{\sqrt{1 - \sin^2\phi}} = t - t_0 \quad \longrightarrow \quad \phi = \sqrt{\frac{k}{m}}(t - t_0)$$
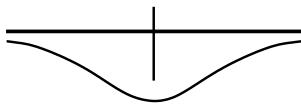
t_0 is another integration constant, and when you put ϕ back into the equation $x = A\cos\phi$ you have

$$x(t) = A\sin\phi = \sqrt{\frac{2E}{k}}\sin\omega_0(t - t_0), \qquad \text{where } E \text{ and } t_0 \text{ are arbitrary constants} \qquad (3.7)$$

This matches the fourth of the four forms in Eqs. (3.3) and (3.4).

Example

▶ A mass moves along the x-axis in a potential energy function

$$U(x) = \frac{-U_0 a^2}{a^2 + x^2} \qquad (3.8)$$

What are the equations of motion and their solution? *At least in the approximation that the motion is small enough to be harmonic.*

A couple of ways to do this: First compute the force directly from the energy function.

$$F_x = -\frac{dU}{dx} = -\frac{d}{dx}\frac{-U_0 a^2}{a^2 + x^2} = \frac{-U_0 a^2 2x}{(a^2 + x^2)^2}$$

This force is zero at $x = 0$ as you can also see from the graph of U, which has zero slope there. Near to this point the denominator is mostly $(a^2)^2 = a^4$ because x is small. In this approximation $\vec{F} = m\vec{a}$ is

$$F_x = \frac{-2U_0 x}{a^2} = m\frac{d^2 x}{dt^2}$$

and this is a harmonic oscillator as in Eq. (3.1), and whose oscillation frequency is $\omega_0 = \sqrt{2U_0/ma^2}$.

Is this the "frequency" or is this the "angular frequency"? That's a choice of units. Do you measure an angle in degrees, radians, or cycles? They're all correct, and it is a matter of convenience which you choose. 2π radians $= 1$ cycle, so 2π radians/second $= 1$ cycle/second $= 1$ Hertz. For reasons of both history and convenience it is ω if it is in radians per second and f or ν in Hertz. The important point to note however is that these manipulations with derivatives are valid *only* in radians.

You can do this calculation another way with a geometric series as in Eq. (0.1)(h). Equivalently, use the binomial series (0.1)(e) with $n = -1$.

$$U(x) = \frac{-U_0 a^2}{a^2(1 + x^2/a^2)} = -U_0\left(1 - \frac{x^2}{a^2} + \cdots\right) \quad\longrightarrow\quad -\frac{dU}{dx} = -\frac{U_0 2x}{a^2} + \cdots$$

and this again gives the harmonic oscillator differential equation $m\ddot{x} = -U_0 2x/a^2$. This series manipulation is a very common technique. Get used to it. Remember when using the binomial expansion that you need to put it in the form $(1 + \text{something small})^n$.

Whichever way you do this, you neglect a higher order term in an infinite series. This applies to both methods, except that it is explicit in the second one, so you can use that series to estimate either how good the approximation is or the range over which it makes sense. In the geometric series the next term appears as $(1 - x^2/a^2 + x^4/a^4 - \cdots)$, so you need $x^4/a^4 \ll x^2/a^2$, or $|x| \ll a$. And *where* in the graph at Eq. (3.8) are the points $|x| = a$?

Example

▶ In the first figure in this chapter, suppose now that a bullet is fired from the right, hitting m and becoming embedded in it. The bullet has mass m_0 and speed v_0. What is the motion of the mass afterwards? Use conservation of momentum to find the initial velocity of the combined mass:

$$(m + m_0)v_1 = -m_0 v_0$$

Now v_1 is the initial velocity of the combined mass, and this mass starts from the equilibrium position $x = 0$. The most convenient form for the solution is sines and cosines as in Eq. (3.3).

$$x(t) = A\cos\omega_0 t + B\sin\omega_0 t, \quad \text{then} \quad x(0) = 0 = A, \quad \text{and} \quad \dot{x}(0) = \omega_0 B = v_1$$

The solution is then

$$x(t) = -\frac{m_0 v_0}{(m + m_0)\omega_0}\sin\omega_0 t$$

First, the dimensions are right, because these are lengths. For small time, the power series expansion of the sine is $\sin\omega_0 t = \omega_0 t - \frac{1}{6}\omega_0^3 t^3 + \cdots$, so this starts as $x(t) \approx -m_0 v_0 t/(m + m_0)$. If $m_0 \gg m$ (a big bullet), this is close to $-v_0 t$. Both of these match the expected behavior of the solution.

Example

▶ And *here* is where you need to know everything in section 0.1.
What happens in a potential energy such as the one on page 62? Between points a and A or between points c and C or g and G you have oscillations. Is the motion simple harmonic? No, but as in the example Eq. (3.8) it is approximately simple harmonic for small oscillations about the equilibrium. I have no equation available to describe these graphs because they were just drawn at random, so computing with them is difficult. Instead, make a new function to illustrate the ideas.

$$U(x) = \frac{\alpha^2}{x} - \frac{\beta^2}{x - x_1} \qquad (x_1 > 0) \tag{3.9}$$

First you must sketch this function, *and I mean you. Now!* The procedure to do this sketch is:

1. Sketch U for x close to $x = 0$.
2. Sketch U for x close to $x = x_1$.
3. Sketch U for x large and positive (but *not* infinite). [Arbitrarily pick $\alpha > \beta$.]
4. Sketch U for x large and negative (but *not* infinite).
5. Now fill in the gaps.

The next step: where is the minimum energy?

$$\frac{dU}{dx} = -\frac{\alpha^2}{x^2} + \frac{\beta^2}{(x - x_1)^2} = 0$$

This is a quadratic equation in x, so it's not hard to find the roots. From the sketch that you just drew, you can easily see that one of the roots corresponds to a minimum of U, and with somewhat more effort you will be able to verify that the other root is at a maximum.

$$\alpha^2(x - x_1)^2 = \beta^2 x^2 \quad \longrightarrow \quad \alpha(x - x_1) = \pm\beta x \quad \longrightarrow \quad (\alpha \mp \beta)x = \alpha x_1$$

The root $x_0 = \alpha x_1/(\alpha + \beta)$ lies between 0 and x_1 and is the stable equilibrium point, as you see from the graph of U that you just sketched.

There are (at least) four ways to proceed from here:

1. Expand the potential energy about x_0 using the general Taylor expansion Eq. (0.4).
2. Expand the potential energy about x_0 using the binomial (or other standard) series expansion in Eq. (0.1).
3. Expand the force about x_0 using the general Taylor expansion.
4. Expand the force about x_0 using the binomial (or other) expansion.

When applicable, three and four are usually the easier methods. Start with number three:

The Taylor series is $F_x(x) = F_x(x_0) + F_x'(x_0)(x - x_0) + \cdots$

so compute $F_x'(x_0)$

$$F_x = -\frac{dU}{dx} = +\frac{\alpha^2}{x^2} - \frac{\beta^2}{(x - x_1)^2}$$

$$\frac{dF_x}{dx} = -2\frac{\alpha^2}{x^3} + 2\frac{\beta^2}{(x - x_1)^3}$$

$$\text{at } x_0 \text{ this} = -2\frac{\alpha^2}{x_0^3} + 2\frac{\beta^2}{(x_0 - x_1)^3}$$

$$= -2\frac{\alpha^2}{\big(\alpha x_1/(\alpha + \beta)\big)^3} + 2\frac{\beta^2}{\big(\alpha x_1/(\alpha + \beta) - x_1\big)^3}$$

$$= \frac{2}{x_1^3}\left[-\frac{\alpha^2(\alpha + \beta)^3}{\alpha^3} + \frac{\beta^2(\alpha + \beta)^3}{(-\beta)^3}\right]$$

$$= -\frac{2}{x_1^3}(\alpha + \beta)^3\left[\frac{1}{\alpha} + \frac{1}{\beta}\right]$$

$$= -\frac{2}{x_1^3}\frac{(\alpha + \beta)^4}{\alpha\beta} \tag{3.10}$$

The force near the equilibrium point is now a Taylor series. See also problem 3.29.

$$F_x(x) = F_x(x_0) + F_x'(x_0)(x - x_0) + \cdots = 0 - \frac{2}{x_1^3}\frac{(\alpha + \beta)^4}{\alpha\beta}z = m\frac{d^2z}{dt^2}$$

where $z = x - x_0$, and of course $d^2 z/dt^2 = d^2 x/dt^2$. This is a simple harmonic oscillator equation with solution

$$z(t) = A\cos(\omega_0 t + \delta) \qquad \text{and} \qquad \omega_0^2 = \frac{2}{x_1^3}\frac{(\alpha+\beta)^4}{m\alpha\beta} \tag{3.11}$$

$$x(t) = x_0 + A\cos(\omega_0 t + \delta)$$

What is the behavior of this solution? The first point to notice is that as $x_1 \to 0$, this frequency $\to \infty$. Why? The potential energy is singular at the origin, and when x_1 becomes smaller the point of minimum potential is squeezed between 0 and x_1. The potential energy curve that you sketched right after Eq. (3.9) then rises ever more steeply and as a result the frequency of oscillation becomes larger, exactly as Eq. (3.11) says. What happens if α or β moves toward zero?

In this example, method three is easier than four, but that's not always so. You need to know all these ways to handle problems.

3.2 Complex Exponentials

The complex exponential that appears as one of the forms of these solutions in Eq. (3.3) isn't essential for anything up to this point, but for the next steps you can't really do without it. It is time to stop and go over that bit of mathematics to be sure that you can manipulate it. First, go back and read section 0.7. That is a review of the basic arithmetic of complex numbers, and it is essential for what follows. Do all the problems 0.43 through 0.48.

For a complex number $z = x + iy$, what is $e^z = e^{x+iy}$? The first step is to use the property of exponentials that this is the same as $e^x e^{iy}$. You already know what the first factor e^x is. What then is e^{iy}? The answer was stated in Eq. (0.46), and one way to derive it is to use the infinite series expansion of the exponential.

$$e^{iy} = 1 + iy + \frac{1}{2}(iy)^2 + \frac{1}{3!}(iy)^3 + \frac{1}{4!}(iy)^4 + \cdots =$$

$$= \left[1 - \frac{1}{2!}y^2 + \frac{1}{4!}y^4 - \frac{1}{6!}y^6 + \cdots\right] + i\left[y - \frac{1}{3!}y^3 + \frac{1}{5!}y^5 - \cdots\right]$$

$$= \cos y + i\sin y$$

The last line comes from recognizing the known series expansions of the sine and cosine, Eq. (0.1). This formula, Euler's, is the single most useful equation involving complex numbers. You will see it everywhere. Complex numbers can be represented graphically as points in the x-y plane—rectangular coordinates. The polar coordinates, r and ϕ, in the same plane represent the complex exponential.
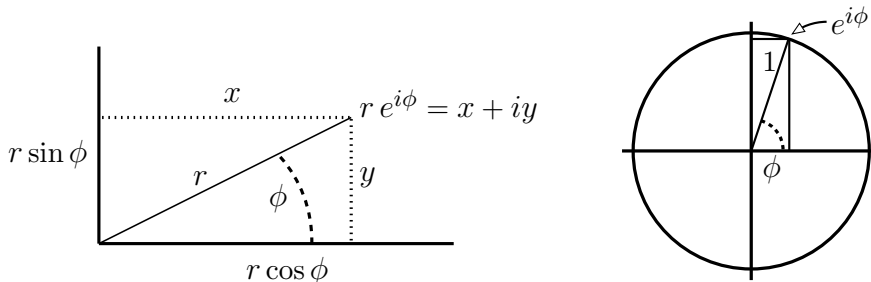
$$x + iy = r\cos\phi + ir\sin\phi = r\,e^{i\phi}$$

Fig. 3.2

How does a graph of $e^{i\phi} = \cos\phi + i\sin\phi$ appear in the complex plane? It's a circle. The radius of the circle is one and the angle ϕ is the angle between the radial line and the $+x$-axis. The magnitude

of a complex number is defined in terms of the Pythagorean Theorem as the length of this radial line, so this is

$$\left|e^{i\phi}\right| = \sqrt{\cos^2\phi + \sin^2\phi} = 1$$

The exponential is periodic,

$$e^{i(\phi+2\pi)} = e^{i\phi}e^{2i\pi} = e^{i\phi}(\cos 2\pi + i\sin 2\pi) = e^{i\phi}$$

and this is clear in the picture because as phi increases, the point representing the complex number keeps wrapping around the circle counterclockwise. The special case $\phi = \pi$ is notable: $e^{i\pi} = -1$.

Example

▶ What is the square of $e^{i\phi}$?

$$\left(e^{i\phi}\right)^2 = e^{2i\phi} = \cos 2\phi + i\sin 2\phi = (\cos\phi + i\sin\phi)^2 = \cos^2\phi - \sin^2\phi + 2i\cos\phi\sin\phi \quad (3.12)$$

Compare the third and fifth expressions and to get an instant derivation of the double angle formulas of trigonometry, because the respective real and imaginary parts here must match. If you want the triple angle formulas (not so well known), cube the exponential, problem 3.3.

When you add and subtract complex numbers, it is usually more convenient to use the rectangular form, as $x + iy$. When you multiply them the polar form has advantages, because

$$r_1 e^{i\phi_1}\, r_2 e^{i\phi_2} = r_1 r_2\, e^{i(\phi_1+\phi_2)} \quad (3.13)$$

The radii multiply as ordinary positive numbers, and the angles add. What happens if you use Euler's formula on the left side of this equation?

$$r_1(\cos\phi_1 + i\sin\phi_1)\, r_2(\cos\phi_2 + i\sin\phi_2) = r_1 r_2\, e^{i(\phi_1+\phi_2)}$$

The factors $r_1 r_2$ cancel, and you can multiply the remaining binomials. Also use Euler's formula on the right side.

$$(\cos\phi_1\cos\phi_2 - \sin\phi_1\sin\phi_2) + i(\cos\phi_1\sin\phi_2 + \sin\phi_1\cos\phi_2)$$
$$= \cos(\phi_1+\phi_2) + i\sin(\phi_1+\phi_2)$$

The real and imaginary parts must match, so this provides an immediate derivation of the sine and cosine of the sum of two angles.

When differentiating complex exponentials such as $e^{i\omega_0 t}$, the manipulations are exactly the same as with real exponentials.

$$\frac{d}{dt}e^{i\omega_0 t} = i\omega_0 e^{i\omega_0 t}$$

That the derivative results in a multiple of the original function is what makes complex exponentials so easy to use when you're trying to describe oscillations. Take another derivative of $e^{i\omega_0 t}$ to get $-\omega_0^2 e^{i\omega_0 t}$, showing again that this exponential satisfies the harmonic oscillator differential equation.

3.3 Damped Oscillators

When friction is present in a system, you still don't know what the force is. Friction is complicated and you have to make some simplifying assumptions about it in order to make any headway in solving the resulting mathematics. The sort of friction seen in introductory texts is dry friction, using an equation for the magnitude of the force that looks like $F = \mu_k F_N$, where μ_k is the coefficient of friction, F_N is the component of the force perpendicular to the sliding surfaces, and the resulting force is then independent of the magnitude of the velocity, depending only on its direction. This is a decent approximation for dry surfaces over a range of speeds and normal forces, but it isn't exact.

When the surfaces are lubricated (wet friction) the frictional force is dependent not only on the direction of the relative velocity of the surfaces, but on its magnitude. When a golf ball is flying the air resistance depends in a complicated way on its velocity, though a decent first approximation assumes that the resistance is proportional to the square of the speed.

All this leads to the fact that I have to make some assumptions, and as much as anything I'll make the assumptions for convenience, not for the best possible representation of physical reality. Having said that, I'll start by picking the frictional force to be linear in the velocity. I do this not because it is the best approximation, but because it's mathematically the simplest. In some applications, such as in electrical circuits, this approximation can be a very good one. It enters as the resistance in the circuit. I'll spend a little time on one of the other models, but any serious treatment of it will require some of the tools that won't appear until another chapter, though this linear approximation is actually a pretty good representation for an object moving slowly through a fluid.

With this *viscous damping* the equation (3.1) become

$$F_x = -kx - bv_x = -kx - b\frac{dx}{dt} = ma_x = m\frac{d^2x}{dt^2} \tag{3.14}$$

Sines and cosines are not enough here, because the first derivative changes one into the other. This differential equation falls into a class of equations that appear repeatedly. They are

linear, homogeneous, constant coefficient

differential equations of the sort described in section 0.9. That is, The dependent variable, x, or its derivatives appear just to the first power, and the coefficients are all constants. Such equations are easy to solve because of the special property of exponentials. The derivative of an exponential is itself. That doesn't mean that e^t is a solution (much less e^x). It doesn't even have the right dimensions. (For the linear, *inhomogeneous* case, see section 3.5.)

$$\text{Use}\quad x(t) = Ae^{\alpha t}, \qquad \text{then}\qquad m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = 0 \qquad \text{is}$$
$$m\alpha^2 Ae^{\alpha t} + b\alpha Ae^{\alpha t} + kAe^{\alpha t} = 0 = Ae^{\alpha t}\left[m\alpha^2 + b\alpha + k\right]$$

The exponential is not zero, and $A \neq 0$, so you're left with a quadratic equation for α.

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4km}}{2m} = -\frac{b}{2m} \pm \sqrt{\frac{b^2}{4m^2} - \frac{k}{m}} \tag{3.15}$$

The general solution to $F_x = ma_x$ is the sum,

$$x(t) = A_1 e^{\alpha_1 t} + A_2 e^{\alpha_2 t} \tag{3.16}$$

where A_1 and A_2 are arbitrary constants and the two α's are the ones that I just found.

   In the common case that the damping is not too large (the underdamped case), the argument of the square root is negative, so Eq. (3.15) becomes (introducing the parameters γ and ω' in the process)

$$\alpha = -\frac{b}{2m} \pm i\sqrt{\frac{k}{m} - \frac{b^2}{4m^2}} = -\gamma \pm i\omega', \qquad \omega' = \sqrt{\omega_0^2 - \gamma^2} \tag{3.17}$$

and this i is where the oscillations come from. Try initial conditions specifying that at time $t = 0$ the position is $x = 0$ and the velocity is $v_x = v_0$. The general solution in Eq. (3.16) produces

$$x(0) = A_1 + A_2 = 0, \qquad \text{and} \qquad \dot{x}(0) = \alpha_1 A_1 + \alpha_2 A_2 = v_0$$

Solve these for the coefficients and put them into the equation for $x(t)$.

$$x(t) = \frac{v_0}{\alpha_1 - \alpha_2}\left[e^{\alpha_1 t} - e^{\alpha_2 t}\right] \tag{3.18}$$

In the underdamped case it is easier to interpret by switching to the explicitly complex form, in which $\alpha_{1,2} = -\gamma \pm i\omega'$.

$$\begin{aligned} x(t) &= \frac{v_0}{-\gamma + i\omega' - (-\gamma - i\omega')}e^{-\gamma t}\left[e^{i\omega' t} - e^{-i\omega' t}\right] \\ &= \frac{v_0}{2i\omega'}e^{-\gamma t}\left[\cos\omega' t + i\sin\omega' t - \cos\omega' t + i\sin\omega' t\right] = \frac{v_0}{\omega'}e^{-\gamma t}\sin\omega' t \end{aligned} \tag{3.19}$$

In the end this is real, and it has to be because everything in the differential equation is real and the initial conditions are real, so if this hadn't come out real I would then have to go back and find my mistake. The oscillation frequency $\omega' = \sqrt{\omega_0^2 - \gamma^2}$ is less than the undamped frequency $\omega_0 = \sqrt{k/m}$. The damping slows the oscillations.

   Does it make sense? First check the dimensions. [Do so.] What is the behavior for small time? The power series expansion for the exponential starts off as $e^x = 1 + x + \cdots$ and for the sine as $\sin x = x - x^3/6 + \cdots$. The first terms are then

$$x(t) = \frac{v_0}{\omega'}[1 - \gamma t + \cdots][\omega' t - \omega'^3 t^3/6 + \cdots] = v_0 t - bv_0\frac{t^2}{2m} + \cdots$$

and this is easy to interpret. The first term, $v_0 t$, says that the mass starts with velocity v_0 as specified. The second terms is in the form $at^2/2$, and shows the initial acceleration to be $-bv_0/m$, which in turn says that the initial force is $-bv_0$. That is the initial viscous force because the spring hasn't yet started to act ($x = 0$). At the other extreme, for large times the motion dies out exponentially as the damping removes the energy. See Figure 3.3 for some graphs.

Overdamped Oscillations
If the damping is large, so that $b^2/4m^2 > k/m$ (or $\gamma > \omega_0$), then the α's are real and negative (the overdamped case), and equation (3.18) is all that's needed. What *sign* do these α's have?

$$\alpha_{1,2} = -\frac{b}{2m} \pm \sqrt{\frac{b^2}{4m^2} - \frac{k}{m}} \qquad \text{and} \qquad \frac{b^2}{4m^2} - \frac{k}{m} < \frac{b^2}{4m^2}$$

which makes the square root smaller in magnitude than the $-b/2m$ term. This implies that both α's are negative, and that both terms in the solution are decaying exponentials. If you put a mass on the end of a spring and then immerse everything in a vat of honey you don't get any oscillations, and in time the mass will approach $x = 0$.

If you now start a mass at $t = 0$ at the origin with velocity $v_x = v_0$, you get precisely Eq. (3.18) for the answer. You don't have to do it again. All that changes is that both α's are now real. Take α_1 as the α with the plus sign, then letting $\eta = \sqrt{\frac{b^2}{4m^2} - \frac{k}{m}}$

$$x(t) = \frac{v_0}{2\sqrt{\frac{b^2}{4m^2} - \frac{k}{m}}} e^{-bt/2m} \left[e^{\eta t} - e^{-\eta t} \right] = \frac{v_0}{\eta} e^{-bt/2m} \sinh \eta t \tag{3.20}$$

and since $\eta < b/2m$ the decaying exponential wins over the growing hyperbolic sine and the coordinate approaches zero for large time.

Critical Damping

There is one further case, though it is sort of a special instance of either of the first two. What if the discriminant of the quadratic equation inside the square root in Eq. (3.15), is zero? The two α's are then the same and you have only one solution to the differential equation. When you want to specify the two initial conditions, position and velocity, there are not enough arbitrary constants to go around and you cannot get a solution. The answer is that you already have the solution Eq. (3.18) in front of you for the case that the damping does not have this special value. If the two α's are almost, but not quite the same, maybe differing by 10^{-20} or so, then does it matter? The solution can't change much if α changes by a tiny amount. Put in mathematical language, start with the solution for the under- or overdamped case and take the limit of Eq. (3.18) as $\alpha_1 \to \alpha_2$.

$$x(t) = \lim_{\alpha_1 \to \alpha_2} \frac{v_0}{\alpha_1 - \alpha_2} \left[e^{\alpha_1 t} - e^{\alpha_2 t} \right] = v_0 \frac{d}{d\alpha} e^{\alpha t} = v_0 \, t e^{\alpha t} = v_0 \, t e^{-bt/2m} \tag{3.21}$$

This limit is nothing more than the definition of the derivative* with respect to α. One point that's easy to miss is that in evaluating this special limiting case you *must* apply the initial conditions first and only after that can you take the limit. Just try it in the other order and you will see that nothing works. Is the small t behavior of Eq. (3.21) correct?
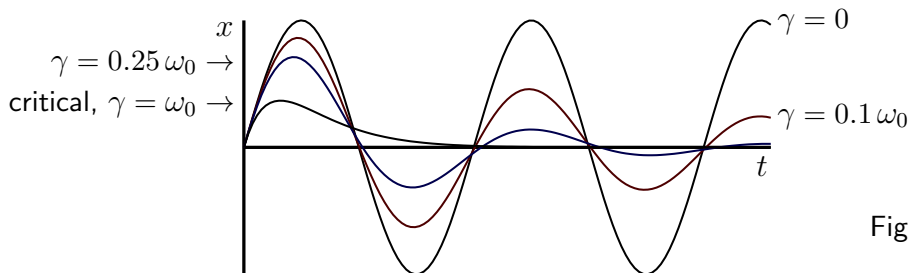

Fig. 3.3

These four graphs show the motion for values of the damping from zero through critical damping, all with the same initial velocity: Eqs. (3.19) and (3.21).

The shock absorbers in an automobile are strongly damped because you do not want the car to keep oscillating up and down after you go over a bump. (If that happens, then it is time to replace them.) If the damping is *too* large however, the ride becomes uncomfortably stiff, so the standard choice it to make the damping parameter very close to critical.

Energy

How does the energy behave in a damped oscillator? The power by the frictional force, the time

_____

* These sort of limits are not always just a derivative. Usually you have to apply a little effort, as in problem 3.12

derivative of total energy, is $dE/dt = F_{x,\text{frict}} v_x = -bv_x^2$. That is

$$\frac{dE}{dt} = \frac{d}{dt}\left(\tfrac{1}{2}mv_x^2 + \tfrac{1}{2}kx^2\right) = mv_x a_x + kxv_x = (ma_x + kx)v_x = -bv_x^2$$

$$= -2m\gamma v_x^2 = -2m\gamma^2 \frac{v_0^2}{\omega'^2} e^{-2\gamma t}\left[-\sin\omega't + \cos\omega't\right]^2 \tag{3.22}$$

$$= -2m\gamma^2 \frac{v_0^2}{\omega'^2} e^{-2\gamma t}\left[1 - \sin 2\omega't\right]$$

Caution: Do not use complex exponentials in computing energy. You must first convert everything into the real (cosine, sine) form. The square of $e^{i\omega t}$ has cross terms that do not belong. The real part of $\left[e^{i\omega t}\right]^2$ is $\cos 2\omega t$ and that is nothing like $\cos^2 \omega t$.

3.4 Other Oscillators
Any stable equilibrium will provide a harmonic oscillator for small motions. Well, almost any; see section 3.12 for one that doesn't.

Pendulum
The pendulum is one of the easiest oscillators to study and for this reason it was probably the first (Galileo). The ideal model of a pendulum has a point mass on the end of a light and inelastic string. To find its equation of motion, use plane polar coordinates with the angle ϕ measured from the vertical.pendulum

Fig. 3.4

$$\frac{d^2\phi}{dt^2} = -\frac{g}{\ell}\sin\phi \tag{3.23}$$

Derive this either by using the torque equation, $\vec{\tau} = I\vec{\alpha}$, or by using the equations (0.41) describing acceleration in polar coordinates, or by using energy conservation, $dE/dt = 0$. Take the second choice first, so look back at section 0.6. Here the radius r is a constant, ℓ, so its derivatives are zero and the acceleration is

$$\vec{a} = \hat{r}\left(\ddot{r} - r\dot{\phi}^2\right) + \hat{\phi}\left(r\ddot{\phi} + 2\dot{r}\dot{\phi}\right) = \hat{r}\left(-\ell\dot{\phi}^2\right) + \hat{\phi}\left(\ell\ddot{\phi}\right) \qquad \text{Eq. (0.41)}$$

The force on m comes from gravity and the string, taking z positive downward.

$$\vec{F} = -\hat{r}F_{\text{string}} + mg\hat{z} = -\hat{r}F_{\text{string}} + \hat{r}mg\cos\phi - \hat{\phi}mg\sin\phi$$

Combine these to get

$$-\hat{r}F_{\text{string}} + \hat{r}mg\cos\phi - \hat{\phi}mg\sin\phi = m\left(\hat{r}\left(-\ell\dot{\phi}^2\right) + \hat{\phi}\left(\ell\ddot{\phi}\right)\right)$$

$$\longrightarrow \quad -F_{\text{string}} + mg\cos\phi = -m\ell\dot{\phi}^2 \qquad \text{and} \qquad -mg\sin\phi = m\ell\ddot{\phi} \tag{3.24}$$

The last equation is the one of interest here, as it is the differential equation for ϕ versus t. The problem is that it's not a simple harmonic oscillator, so what's it doing in this chapter? It says that the second derivative of ϕ with respect to t is a sine of ϕ, not ϕ itself. The simple forms of solution such as $\cos\omega t$ or $e^{\alpha t}$ are not right. *Try it and see.* It is certainly not a linear, constant coefficient differential equation.

There is however an approximation that works surprisingly well. If the angle of oscillation is small then you can use the series expansion of the sine to get

$$\frac{d^2\phi}{dt^2} = -\frac{g}{\ell}\sin\phi = -\frac{g}{\ell}\left[\phi - \frac{1}{6}\phi^3 + \frac{1}{120}\phi^5 - \cdots\right] \approx -\frac{g}{\ell}\phi$$
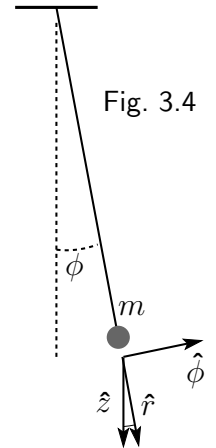
How good is this approximation? The term omitted should be a lot less than what I'm keeping.

$$\frac{1}{6}\phi^3 \ll \phi \qquad \text{so if the ratio is 1\% then} \qquad \phi^3 < .06\phi, \quad \text{or} \quad \phi < \sqrt{.06} = .24 = 14°$$

A more detailed calculation, section 4.6, shows that the error in calculating the frequency is about 0.4% when ϕ swings back and forth $\pm14°$. This is smaller than 1% because the error in estimating the sine reaches its worst value only at the ends of the swing. For most of the pendulum's motion, the error is smaller than its maximum value.

With this small angle approximation, the equation of motion is that of a simple harmonic oscillator

$$\frac{d^2\phi}{dt^2} = -\frac{g}{\ell}\phi \implies \phi(t) = A\sin(\omega_0 t + \delta), \qquad \text{where} \qquad \omega_0 = \sqrt{\frac{g}{\ell}} \tag{3.25}$$

That the frequency of oscillation is (nearly) independent of the amplitude (A) of the oscillation was a surprise to Galileo and others of his time (ca. 1600). A consequence of this observation was the development of the pendulum clock. In the next approximation, keeping the ϕ^3 term, the frequency is approximately

$$\omega_0 \approx \sqrt{\frac{g}{\ell}}\left[1 - \tfrac{1}{16}\phi_{\text{max}}^2\right] \tag{3.26}$$

See section 4.6 and problem 4.43 for the details of this.

Energy Method

Another method to derive the differential equation of motion uses conservation of energy, and this way doesn't require knowing the equation (0.41) for acceleration in polar coordinates. Write the total energy. The kinetic energy is $mv^2/2 = m\ell^2\dot\phi^2/2$, and the potential energy is the gravitational $mgh = mg\ell(1 - \cos\phi)$. Here I took the zero of potential at the bottom of the swing. Energy is conserved, so its time-derivative is zero.
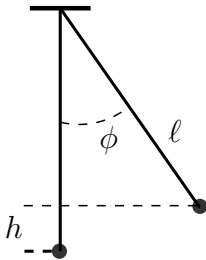

$$E = \frac{1}{2}m\ell^2\dot\phi^2 + mg\ell(1 - \cos\phi),$$

$$\text{then} \quad \frac{dE}{dt} = \frac{1}{2}m\ell^2\frac{d\dot\phi^2}{d\dot\phi}\frac{d\dot\phi}{dt} - mg\ell\frac{d\cos\phi}{d\phi}\frac{d\phi}{dt} \tag{3.27}$$

$$= m\ell^2\dot\phi\frac{d^2\phi}{dt^2} + mg\ell\sin\phi\frac{d\phi}{dt} = 0$$

The chain rule is the only mathematical tool needed to do these manipulations. The factor of $\dot\phi$ cancels, and the result is Eq. (3.23) again, but with less effort. Do you lose anything by using this simpler approach? Yes, but just the part of Eq. (3.24) for the tension in the string.

Electric Circuits

The simplest electric circuits will exhibit oscillations. All that you need is an inductor and a capacitor.
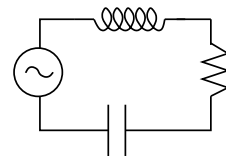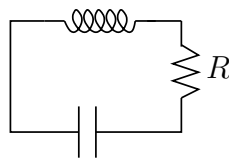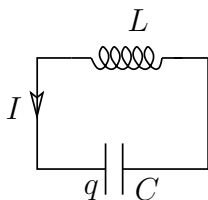
$$L\frac{dI}{dt} + \frac{q}{C} = 0, \qquad \text{and} \qquad I = \frac{dq}{dt} \implies L\frac{d^2q}{dt^2} + \frac{q}{C} = 0$$

Put in a resistor and you have

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{q}{C} = 0$$

Add an oscillating voltage source; replace the zero on the right side by $V_0 \cos \omega t$.

There is nothing new in these equations; they're the same equations as in the last several pages except that the symbols have changed. In describing mechanical friction I made a point about the sort of modelling needed to describe it, and the same question arises here. Can you really say that the electrical resistance is IR with a constant R? Sometimes and in some approximation you can, but not always. For an ordinary incandescent light bulb you probably expect this to be valid, but it too is just an approximation. The resistance in the tungsten filament depends on its temperature, and that in turn depends on the size of the current going through it. That replaces the potential difference IR with a slightly nonlinear function of I. With a fluorescent bulb, the situation is not even close to linear, and the relation between current and voltage in that case is not even single-valued. It is not a function, and the voltage-current relation even has a negative slope $(dI/dV < 0)$ in some regions. That's the reason that fluorescent fixtures use a *ballast* to keep the system stable. That ballast could be just another resistor, but in practice that would waste too much energy so other devices are used instead.

3.5 Forced Oscillations

With any sort of harmonic oscillator you can also have other forces besides the spring and the damping (or the capacitance and the resistance). If the oscillator is a child on a swing (a pendulum) you may be pushing periodically. If it's a mass on a spring you may be pulling on the end of the spring, varying the force on the mass by changing the length of the spring. In any case, the differential equation will be something like

$$m\frac{d^2x}{dt^2} = -b\frac{dx}{dt} - kx + F_{\text{external}}(t) \tag{3.28}$$

This is not a linear, homogeneous, constant coefficient differential equation, and the sum of two of its solutions is not a solution, but it comes pretty close.

Terminology: "Homogeneous" means that if you multiply the dependent variable x by a constant then every term in the equation is multiplied by that constant.* Equation (3.14) satisfies this, but (3.28) doesn't. It is an inhomogeneous equation.

"Linear" means that the dependent variable and its derivatives appear only to the first or zeroth power. Both of these equations satisfy this.

There is a plan of attack for handling inhomogeneous problems such as these.

1. Toss out the inhomogeneous term and find the general solution of this homogeneous part, complete with all the arbitrary constants.

2. Find any one solution to the full, inhomogeneous equation, and it doesn't need to have any arbitrary constants.

3. Add the results of steps 1 and 2.

This will provide a solution to the full equation complete with the arbitrary constants needed to fit the initial conditions. To show that this is true, simply plug the result into the original equation.

Example

▶ Put a mass on the end of a spring and hang the spring vertically. The equation of motion is

$$m\frac{d^2x}{dt^2} = F_x = -kx + mg \tag{3.29}$$

Step one: Toss out the mg term and you already know the solution to be
$x_{\text{hom}}(t) = A \sin \omega_0 t + B \cos \omega_0 t.$

---
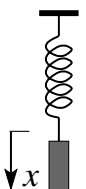
* or a power of that constant

Step two: The most efficient way to find a solution to an equation is guesswork. What function does it take so that a combination of the function and its second derivative gives a constant? A constant of course. There are more systematic ways that I'll get to in section 3.11, but this will do for the simple cases. $x_{\text{inh}} = C = mg/k$

Step three:

$$x = x_{\text{hom}} + x_{\text{inh}} = Ae^{i\omega_0 t} + Be^{-i\omega_0 t} + mg/k \qquad (3.30)$$

If the initial conditions are that I release the mass from rest at the origin,

$$x(0) = 0 = A + B + \frac{mg}{k} \qquad \text{giving} \qquad A = B = -\frac{mg}{2k}$$
$$\dot{x}(0) = 0 = Ai\omega_0 - Bi\omega_0 \qquad \qquad x(t) = \frac{mg}{k}[1 - \cos\omega_0 t] \qquad (3.31)$$

As usual, what does this say? The dimensions check easily. Also, it's real. During the derivation of the answer, all the equations were complex, so that it comes out real says that I haven't made too many mistakes. Now, what does it look like for small time? Use the series for a cosine.

$$\text{small } t \implies x \approx \frac{mg}{k}\left[1 - \left(1 - \tfrac{1}{2}\omega_0^2 t^2 + \cdots\right)\right] = \frac{1}{2}\frac{mg}{k}\omega_0^2 t^2 = \frac{1}{2}gt^2$$

The first thing it does is drop. Of course.

After that it oscillates between $x = 0$ (when $\cos\omega_0 t = 1$) and the bottom point $x = 2mg/k$ (when $\cos\omega_0 t = -1$). The equilibrium point, where the total force is zero, is halfway between these two.
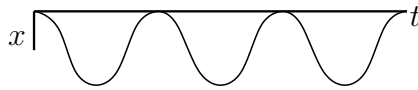
Example

▶ Apply a constant force to an oscillator, but for a finite time. Then shut it off. Assume the mass starts at the origin at rest, $m\ddot{x} = -kx + F(t)$

$$F(t) = \begin{cases} F_0 & (0 < t < T) \\ 0 & (\text{otherwise}) \end{cases} \qquad\qquad (3.32)$$

For the interval $0 < t < T$ the differential equation is $m\ddot{x} + kx = F_0$, and the solution to this is

$$x(t) = A\cos\omega_0 t + B\sin\omega_0 t + F_0/k$$

Use the initial conditions.

$$x(0) = 0 = A + F_0/k, \qquad \dot{x}(0) = 0 = B$$
$$\text{which imply} \qquad x(t) = \frac{F_0}{k}[1 - \cos\omega_0 t] \qquad (t < T) \qquad (3.33)$$

At the end of the T-interval, start again with the final conditions at the end of $0 < t < T$ as the initial conditions for $T < t < \infty$. (I will set this up the hard way first, then show an easier way.) The differential equation no longer has the F_0 term, so the solution is

$$\text{for } t > T, \qquad x(t) = C\cos\omega_0 t + D\sin\omega_0 t,$$
$$\text{from (3.33):} \quad x(T) = \frac{F_0}{k}[1 - \cos\omega_0 T], \quad \dot{x}(T) = \frac{F_0}{k}\omega_0\sin\omega_0 T$$
$$\text{match } x \text{ and } \dot{x}: \quad x(T) = C\cos\omega_0 T + D\sin\omega_0 T = \frac{F_0}{k}[1 - \cos\omega_0 T]$$
$$\dot{x}(T) = -C\omega_0\sin\omega_0 T + D\omega_0\cos\omega_0 T = \frac{F_0}{k}\omega_0\sin\omega_0 T$$

The last two lines are two equations for the two unknowns C and D. Straight-forward to solve, but this is not the easy way to do the problem. This is the same sort of difficulty that occurred in the example at Eq. (2.4), and you simplify the algebra in the same way. Look back to that equation and see what occurred there.

Instead of grinding through this algebra, jump to the easier method...

Plan ahead and choose your coordinates carefully. This includes where the clock starts. The natural origin for the second half of the problem is T, not zero. That suggests writing the solution as

$$x(t) = C \cos \omega_0(t - T) + D \sin \omega_0(t - T) \qquad (t > T)$$

$$\text{then} \quad C = x(T) = \frac{F_0}{k}[1 - \cos \omega_0 T] \quad \text{and} \quad D = \dot{x}(T)/\omega_0 = \frac{F_0}{k} \sin \omega_0 T \qquad (3.34)$$

Any time that you find yourself with an algebraic mess, look to see if you could have set it up more skillfully. If not, then show no fear and plunge ahead, but here you see what a little insight will do. Put this all together and

$$x(t) = \begin{cases} 0 & (t < 0) \\ \frac{F_0}{k}[1 - \cos \omega_0 t] & (0 < t < T) \\ \frac{F_0}{k}[(1 - \cos \omega_0 T) \cos \omega_0(t - T) + (\sin \omega_0 T) \sin \omega_0(t - T)] & (T < t) \end{cases} \qquad (3.35)$$

Examine the results.

(a) If $\omega_0 T = 2\pi$ all motion stops after $t = T$. That happens because at the time just *before* T the mass comes in to a gentle stop at the origin. If you turn off the force at that instant then it stays stopped.

(b) If the force lasts just a short time (short compared to what? $\omega_0 T \ll 1$), then the motion is approximately

$$x(t) \approx \frac{F_0}{k}\left[\tfrac{1}{2}(\omega_0 T)^2 \cos \omega_0 t + \omega_0 T \sin \omega_0 t\right]$$

The second (sine) term is the dominant one now,

$$\frac{F_0}{k}\omega_0 T \sin \omega_0 t = \frac{F_0 T}{m} \cdot \frac{m}{k}\omega_0 \sin \omega_0 t = \frac{F_0 T}{m}\frac{\omega_0}{\omega_0^2} \sin \omega_0 t = \frac{v_{\text{initial}}}{\omega_0} \sin \omega_0 t$$

and $v_{\text{initial}} = F_0 T/m$ is the velocity just after the impulse.

(c) You can find a trigonometric identity to simplify the appearance of Eq. (3.35). Does it help?

Are there still easier and more systematic ways? Yes; go to section 3.11 to see.

3.6 Harmonic Forcing

Another time-dependent force, one that appears very often and in many contexts. If I hold a spring by one end with the mass hanging from the other and I then move my hand up and down, that will provide an oscillating force on m. If a child is on a swing and I push periodically, that's a time dependent force. The simplest mathematical form for such a push is a cosine or sine. The generalization to a more complicated periodic force can wait. Call the frequency of this force Ω (capital ω), then

$$m\frac{d^2 x}{dt^2} = -kx - b\frac{dx}{dt} + F_{\text{external}}(t) \qquad \text{where} \qquad F_{\text{external}}(t) = F_0 \cos \Omega t \qquad (3.36)$$

If the damping term is absent, it is easy to guess a solution to this. You want a function such that the second derivative plus the function itself result in a cosine of Ωt. A cosine itself will do that. You have

to adjust the coefficient, but that's not hard. See problem 3.14. The damping term makes inspired guessing more difficult to do because the first derivative changes a cosine into a sine.

Here is where the complex exponential form comes into its own. Up to now there's little I've done that I couldn't do just with simple sines and cosines, but the damping term makes those much harder to deal with. The right technique is to say that the cosine is the real part of the complex exponential. Solve everything with the exponential form and then take the real part of the result.

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = F_0 e^{i\Omega t} \qquad \text{(real part understood)} \qquad (3.37)$$

A more explicit way to say this is that instead of solving one equation, solve two. Two different forcing functions (cosine and sine) give two different x's (x_r and x_i) that satisfy two different differential equations.

$$m\frac{d^2x_r}{dt^2} + b\frac{dx_r}{dt} + kx_r = F_0 \cos \Omega t \qquad \text{and} \qquad m\frac{d^2x_i}{dt^2} + b\frac{dx_i}{dt} + kx_i = F_0 \sin \Omega t$$

Now take the first of these equations and add it to i times the second.

$$m\frac{d^2(x_r + ix_i)}{dt^2} + b\frac{d(x_r + ix_i)}{dt} + k(x_r + ix_i) = F_0(\cos \Omega t + i \sin \Omega t)$$

If I let $x = x_r + ix_i$, this is Eq. (3.37).

Now, to get an exponential output on the right hand side, put an exponential in. Use a solution that's a constant times $e^{i\Omega t}$ and you have a solution.

$$x_{\text{inh}}(t) = Ce^{i\Omega t} \longrightarrow -mC\Omega^2 e^{i\Omega t} + ibC\Omega e^{i\Omega t} + kCe^{i\Omega t} = F_0 e^{i\Omega t}$$

The exponential cancels and you have the equation for C.

$$C = \frac{F_0}{-m\Omega^2 + ib\Omega + k}, \qquad \text{and} \qquad x_{\text{inh}}(t) = \frac{F_0 e^{i\Omega t}}{-m\Omega^2 + ib\Omega + k} \qquad (3.38)$$

Before worrying about adding the solution to the homogeneous part of the equation, I'll look more closely at this inhomogeneous solution. It takes a bit of manipulation to put this into a simple and useful form, and the key is to remember that while addition and subtraction of complex numbers is more easily done in terms of rectangular components, multiplication and division work more easily in polar form.

The denominator is

$$-m\Omega^2 + ib\Omega + k = A + iB = \sqrt{A^2 + B^2}\,\frac{A + iB}{\sqrt{A^2 + B^2}} \qquad (3.39)$$
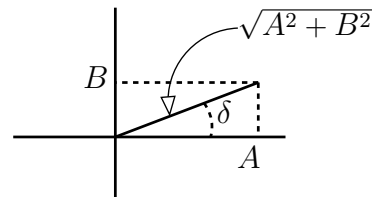
The reason for this multiplication and division by the same factor is that it makes the final fraction have magnitude one. That allows me to write it as an exponential, $e^{i\delta}$. From the picture, the cosine and the sine of the angle δ are the two terms in the fraction.

$$A + iB = \sqrt{A^2 + B^2}(\cos \delta + i \sin \delta) = \sqrt{A^2 + B^2}\,e^{i\delta} \qquad \text{and} \qquad \tan \delta = \frac{\sin \delta}{\cos \delta} = \frac{B}{A} \qquad (3.40)$$

There's a slightly tricky point about this. If $\delta = \tan^{-1}(B/A)$ then what happens when A is zero? Or if it changes sign? Do you get $\pi/2$ or $-\pi/2$? What is $\tan^{-1}(-1)$? The real way to do this is to

look back at the rectangular form, $A + iB$. Knowing the quadrants that A and B are in will specify which angle you are dealing with. The arctangent is multiple valued, and by specifying *both* rectangular components you remove the ambiguity.* The best way remains: look at the picture. Where are A and B in the complex plane? That answers the question. This way $\tan^{-1}(y/x) \to \tan^{-1}(1/-1) = 3\pi/4$ and $\tan^{-1}(-1/1) = -\pi/4$ or $+7\pi/4$.

The denominator in Eq. (3.38) is now

$$\sqrt{(-m\Omega^2 + k)^2 + b^2\Omega^2}\, e^{i\delta}, \qquad \text{where} \qquad \tan\delta = \frac{\sin\delta}{\cos\delta} = \frac{b\Omega}{k - m\Omega^2} \qquad (3.41)$$

This inhomogeneous solution is then

$$x_{\text{inh}}(t) = \frac{F_0 e^{i\Omega t}}{\sqrt{(-m\Omega^2 + k)^2 + b^2\Omega^2}\, e^{i\delta}} = \frac{F_0 e^{i(\Omega t - \delta)}}{\sqrt{(-m\Omega^2 + k)^2 + b^2\Omega^2}} \qquad (3.42)$$

For the real part of this, change $e^{i(\Omega t - \delta)} \to \cos(\Omega t - \delta)$. It is also easier to interpret the result if you divide numerator and denominator by m and write ω_0^2 instead of k/m. The angle δ is called the phase lag, appearing as it does in $(\Omega t - \delta)$.

$$x_{\text{inh}}(t) = \frac{(F_0/m)e^{i(\Omega t - \delta)}}{\sqrt{\left(-\Omega^2 + \omega_0^2\right)^2 + b^2\Omega^2/m^2}} \qquad (3.43)$$

For the real part, the numerator of $x_r(t)$ becomes $(F_0/m)\cos(\Omega t - \delta)$. This is worth more study all by itself before combining it with the homogeneous solution. Suppose that the damping coefficient b is small, and examine the denominator of the last equation. The square root factor is a function of the applied frequency Ω, and as Ω varies from zero to infinity the first term, $(\omega_0^2 - \Omega^2)^2$, varies from ω_0^4 to infinity, but notice: it vanishes in between. When $\Omega^2 = \omega_0^2 = k/m$ this parenthesis is zero and all that's left in the denominator of (3.43) is the damping factor $b\Omega/m$ and I said I'm assuming that b is small. When this denominator is small the inhomogeneous solution is large. This is the phenomenon of resonance.

In the same analysis, when the forcing frequency varies, then what happens to the angle δ in Eq. (3.41)? As $\Omega \to 0$, the angle goes to zero. As Ω increases the denominator goes to zero so $\tan\delta$ goes to infinity. Now what? The answer is really much easier to find by looking at the original complex form, Eq. (3.39).

$$-m\Omega^2 + ib\Omega + k \propto e^{i\delta}$$

In this rectangular form, small Ω makes this positive real and then $\delta \to 0$ as stated. As Ω increases you pick up a positive imaginary part, carrying δ up to $\pi/2$ when $m\Omega^2 = k$. Finally, for very large Ω this is mostly negative real $(-\Omega^2)$ plus a comparatively small amount of positive imaginary $(i\Omega)$, so $\delta \to \pi$.

The next three graphs show the amplitude and the phase δ as a function of the driving frequency Ω. There are six plots in each case, for various values of $b/m\omega_0$ from 0.5 to 0.02. The first is the amplitude (the $1/\sqrt{}$ factor in Eq. (3.42)). The second plots the phase for the same parameters. The third plots the complex number that is the whole coefficient of $F_0 e^{i\Omega t}$ in the same equation. It is worth seeing how various parts of the third graph correlate with the first two graphs.

---

* If you write a computer program that calls on the inverse tangent function, look for a library version that takes two arguments (A and B) and not just one.
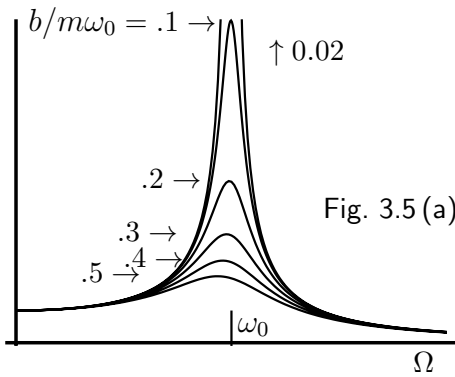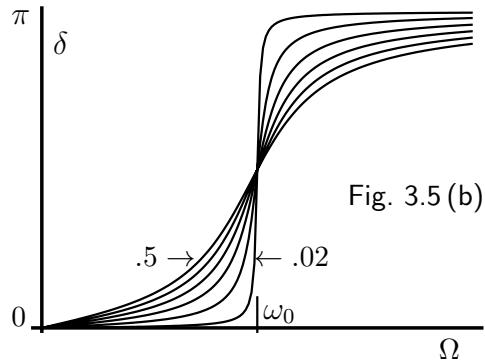
Fig. 3.5 (a)

Fig. 3.5 (b)

In the second graph you can see the phase δ as it starts from zero, meaning that when you push with a slow oscillation the mass will follow your push. If however the forcing frequency Ω is high, then $\delta \to \pi$ and the response approaches $180°$ out of phase from the applied force. These phase differences have a major effect on the structure of ocean tides, discussed in chapter five on page 174. The graph showing how the phase approaches π when the forcing frequency is well above the natural frequency helps explain why the highest high tides can occur near the time that the Moon is either rising or setting, not when it is overhead. (Did you know that?)

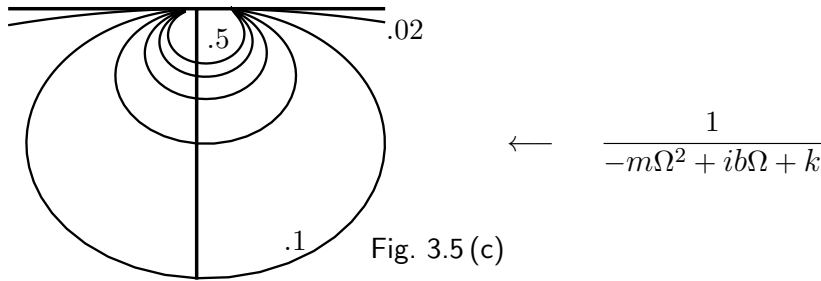$$\longleftarrow \qquad \frac{1}{-m\Omega^2 + ib\Omega + k}$$

Fig. 3.5 (c)

The next logical step would be to add the full solution of the homogeneous equation and to apply initial conditions. I won't. It is not very instructive and it is moderately complicated. I'll expect *you* to do it for the undamped case, problem 3.14, because in that special case there is enough payoff for the relatively modest effort involved.

Terminology: The inhomogeneous term that I just derived is called the *steady-state* term, and the homogeneous term computed in section 3.3 is the *transient*. That is simply because the homogeneous solution has a negative exponential in it and it eventually dies out. The steady-state term keeps going.

Is the value $\Omega^2 = k/m = \omega_0^2$ the position of the peak in the first of these graphs? That's the one that shows the amplitude of the response to the force. No it is not. The peak occurs when the denominator in Eq. (3.43) is a minimum, and to find that you have to differentiate with respect to Ω. Better yet, differentiate with respect to Ω^2.

$$\frac{d}{d\Omega^2}\left(\left(-\Omega^2 + \omega_0^2\right)^2 + b^2\Omega^2/m^2\right) = 2\left(\Omega^2 - \omega_0^2\right) + b^2/m^2 = 0$$

$$\text{then} \qquad \Omega^2 = \omega_0^2 - b^2/2m^2$$

You see from this that the resonance peak really occurs slightly below the natural frequency. For the most common interesting cases the damping is small and this distinction is not large, but it is there and you can see it in the amplitude graph, Figure 3.5 (a).

Mechanical Resonance Example

Attach a mass to the end of a spring and let it hang down from your hand. Now move your hand up and down in a low frequency oscillation. The mass will follow your hand in essentially the same motion.

Gradually increase the frequency of your hand's motion and the *amplitude* of the mass's motion will rise, eventually becoming large enough that the mass may fly off the end of the spring. This follows the left side of figures Fig. 3.5 (a) and Fig. 3.5 (b). The amplitude of m's motion increases, while it more-or-less stays in phase with the motion of your hand: $0 < \delta < \pi/2$.

Now repeat the experiment but start with a high frequency oscillation of your hand. Then gradually lower the frequency and watch the mass. Now you are following the same two figures, but coming in from the right. Again, the response of the mass will increase dramatically, but its direction of motion will be opposite to that of your hand, corresponding to the mathematical statement $\pi/2 < \delta < \pi$.

Optical Resonance Example

Light hits an atom, and as a consequence it applies an oscillating force to the atom. If the frequency of the light matches one of the natural transition frequencies in the atom, there is a large response, and the light is either strongly absorbed or strongly scattered. In the sun this phenomenon causes the Fraunhofer lines. The light from the hot sun passes through the (relatively) cooler solar atmosphere where un-ionized atoms can exist, and the specific colors of light whose frequencies match the resonant frequencies of some of the atoms is absorbed or scattered, leaving a dark line in the spectrum.
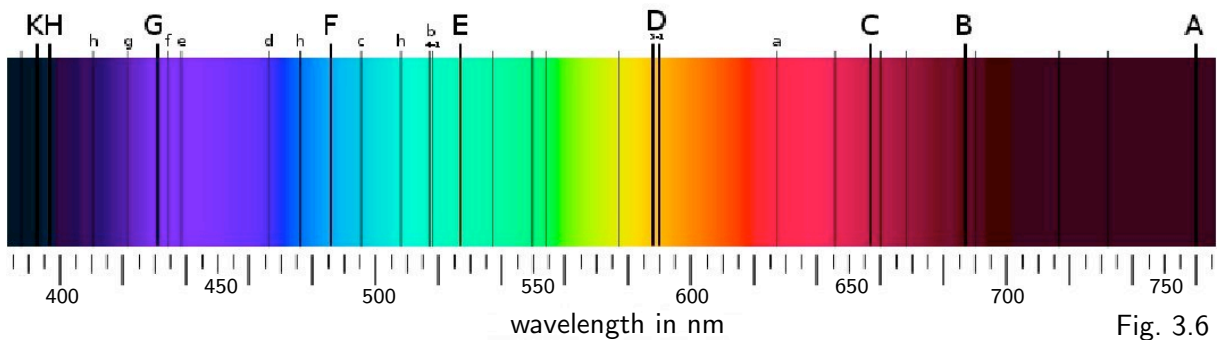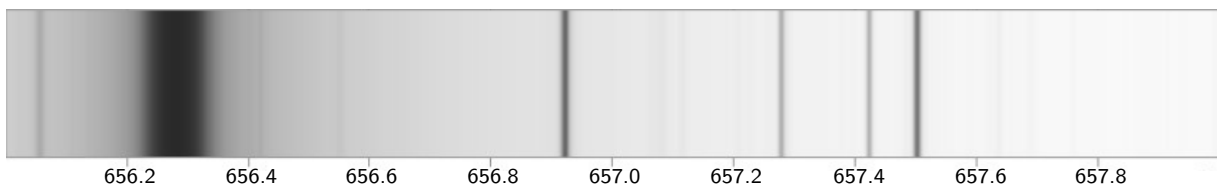

wavelength in nm

Fig. 3.6

The line labeled "C" in the first picture is the Hα line due to the presence of excited hydrogen atoms. The second picture is a greatly magnified version in which this line appears as the large blob at a wavelength just below 656.3 nm. The large width of this line is partly caused by the natural width of the resonance as in the graph on the preceding page, but the atom also suffers many collisions with its neighbors, and this accounts for most of the spreading in this case.

Detailed understanding of this phenomenon requires quantum mechanics, but it is surprising just how far the intuition that you get from the classical harmonic oscillator extends to the quantum treatment.

Example

▶ A few lines back I said that I didn't want to add the full homogeneous solution to the problem. I will make an exception here. In the undamped case, what does the solution look like with an oscillating external force?

$$F_x = -kx + F_0 \sin \Omega t = m\ddot{x} \quad \longrightarrow \quad x_{\text{hom}} = A \cos \omega_0 t + B \sin \omega_0 t \tag{3.44}$$

For the inhomogeneous part, you want a combination of x and $\ddot{x}$ to produce a $\sin \Omega t$. That's easy: a $\sin \Omega t$ does the job.

$$x_{\text{inh}}(t) = C \sin \Omega t \longrightarrow m\ddot{x}_{\text{inh}} - kx_{\text{inh}} = C\big[-m\Omega^2 + k\big]\sin \Omega t = F_0 \sin \Omega t$$

That determines the constant C and the whole solution is now

$$x(t) = A \cos \omega_0 t + B \sin \omega_0 t + \frac{F_0}{-m\Omega^2 + k}\sin \Omega t$$

Use the initial condition that the mass starts at the origin with zero velocity.

$$x(0) = 0 = A, \qquad \dot{x}(0) = 0 = B\omega_0 + \frac{F_0\Omega}{-m\Omega^2 + k}$$

Put these together and use $\omega_0^2 = k/m$ to get

$$x(t) = \frac{F_0/m}{\omega_0^2 - \Omega^2}\left[\sin \Omega t - \frac{\Omega}{\omega_0}\sin \omega_0 t\right] \tag{3.45}$$

First, in order to see if this makes sense, what *must* it look like for small time? The spring starts with no force and the added force starts at zero, so it must start with zero acceleration. Expand it:

$$x(t) = \frac{F_0/m}{\omega_0^2 - \Omega^2}\left[\Omega t - \tfrac{1}{6}\Omega^3 t^3 + \cdots - \frac{\Omega}{\omega_0}\left(\omega_0 t - \tfrac{1}{6}\omega_0^3 t^3 + \cdots\right)\right] = \frac{1}{6}\frac{F_0}{m}\Omega t^3 + \cdots$$

Is this acceleration right? Two derivatives give $\ddot{x} = (F_0/m)\Omega t + \cdots$, and that is exactly what the starting equation (3.44) says. It is, isn't it?

What does this solution look like at later times? The equation (3.45) has two oscillations of two different frequencies. Sometimes they add; sometimes they subtract. A graph will help.

Fig. 3.7

This graph uses two fairly random frequencies ($\Omega = 7.97$ and $\omega_0 = 11.3$ in some equally arbitrary units). The individual frequencies are already hard to discern in this graph.

▶ There is an interesting example of what happens to these resonances in a circumstance only slightly more complicated than these: a forced pendulum is *almost* a forced simple harmonic oscillator, but not quite. This leads to some very surprising results, and there is a qualitative description of the phenomenon in section 3.13. It's something that you're not used to in this book — pages without any equations.

3.7 Stable Motion

When do you encounter harmonic oscillators? Not just in springs and pendulums or even electric circuits. If you have a system in equilibrium and you disturb it a little, what happens? If it's a stable equilibrium then it is pushed back toward its original configuration. If it's unstable then it is pushed farther away. If you're skateboarding in a valley or on top of a hill you have different concerns. If an atom is in a molecule and it gets bumped to the side then the other atoms in the molecule will push it back into place.

Every one of these examples leads to a harmonic oscillator, and you can see why by looking at the potential energy, $E = mv^2/2 + U(x)$. The condition for equilibrium is that $F_x = -dU/dx = 0$, and whether this is stable or not depends on whether the potential energy function curves up or down.
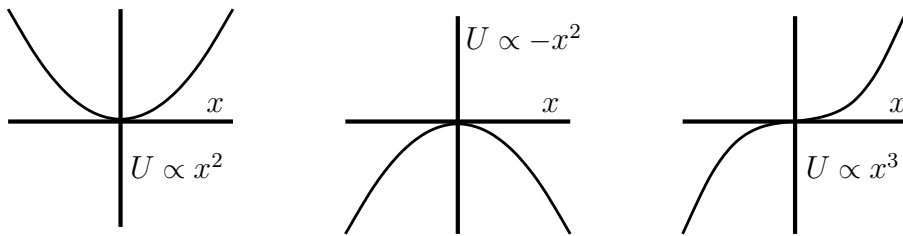

Fig. 3.8

The first graph leads to stable oscillations and the second graph provides a force that pushes the mass away from equilibrium. $F_x = -dU/dx$, so a positive slope for U implies a negative F_x and vice versa. In the first graph that force is toward the origin and in the second graph that is away from it. The third graph ($U = Ax^3$) is neither, but it never happens in practice* because even the slightest change in the system changes the graph into something quite different. If you add $\pm 10^{-20}x$ to this U then its behavior changes. See problem 3.19.

You can get the equation of motion directly from energy conservation by saying that $dE/dt = 0$.

$$\frac{d}{dt}\left[\frac{1}{2}mv^2 + U(x)\right] = 0 = mv\frac{dv}{dt} + \frac{dU}{dx}\frac{dx}{dt} = v\left[m\frac{dv}{dt} + \frac{dU}{dx}\right] \tag{3.46}$$

Cancel the v and you see that using energy methods is often a faster way to the equations of motion. What if there's friction? See problem 3.22.

The power series expansion of U allows for a quantitative analysis of the motion

$$U(x) = U(x_0) + (x - x_0)U'(x_0) + \frac{1}{2}(x - x_0)^2 U''(x_0) + \frac{1}{3!}(x - x_0)^3 U'''(x_0) + \cdots$$

If x_0 is an equilibrium point of the potential energy $U'(x_0) = 0$, then the dominant term after that is the quadratic, $U''(x_0)(x - x_0)^2/2$. If $U''(x_0) > 0$ then this looks like the first picture above and you have stable equilibrium

$$m\ddot{x} = F_x = -dU/dx = -U''(x_0)(x - x_0)$$

This is a standard harmonic oscillator with center at x_0. Take a specific example and do it several different ways.

$$U(x) = -A\operatorname{sech}(\alpha x) = -\frac{A}{\cosh \alpha x}$$
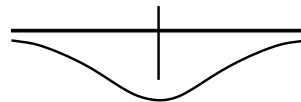

Fig. 3.9

The first method (and usually the hardest) is to plug in to the Taylor series formula to get the series representation for U out to the quadratic term. $U(0) = -A$, and

$$U'(x) = \alpha A \operatorname{sech} \alpha x \tanh \alpha x \quad \longrightarrow \quad U'(0) = 0$$
$$U''(x) = \alpha^2 A\big(-\operatorname{sech}^2 \alpha x \tanh^2 \alpha x + \operatorname{sech}^3 \alpha x\big) \quad \longrightarrow \quad U''(0) = \alpha^2 A$$
$$U(x) = -A + \tfrac{1}{2}\alpha^2 A x^2 + \cdots$$

---

* This shape is called *generically unstable*. A little change make a qualitative difference, not simply quantitative.

The equation of motion is

$$m\ddot{x} = F_x = -dU/dx = -\alpha^2 Ax \tag{3.47}$$

and this is the simple harmonic oscillator with frequency $\omega_0 = \alpha\sqrt{A/m}$. If α is small, the potential well is wide and has a weak restoring force; the corresponding frequency is small.

The second method to obtain the approximate equation of motion from the potential is to stop after doing the first derivative.

$$m\ddot{x} = F_x = -\frac{dU}{dx} = -\alpha A \operatorname{sech} \alpha x \tanh \alpha x \tag{3.48}$$

For small x, the sech is essentially 1, and $\tanh \alpha x \approx \alpha x$ so this equation reproduced Eq. (3.47) for small oscillations.

The third method is to use series expansions on the potential energy and to differentiate that. Eq. (0.1) includes the cosh and the binomial expansions.

$$U(x) = -\frac{A}{\cosh \alpha x} = -\frac{A}{1 + \frac{1}{2}\alpha^2 x^2 + \cdots} = -A\left[1 - \tfrac{1}{2}\alpha^2 x^2 + \cdots\right]$$

and the force comes from the derivative of the second term, giving the same equation as before. Once you have more practice with series, this is commonly the easiest choice.

Whichever way you arrive at it (only occasionally the first method), the equation to solve is Eq. (3.47), and that solution is the standard $x(t) = C\cos(\omega_0 t + \delta)$ with $\omega_0^2 = \alpha^2 A/m$. Is this an exact solution for Eq. (3.48)?. No, but just because the approximation that $\alpha|x| \ll 1$ is required. This is however *exactly* the same difficulty that occurred in Eq. (3.1) for a mass attached to a spring. I said that the force by the spring is represented by $F_x = -kx$, but that's an approximation too. It is valid only if $|x|$ is small enough in some unspecified way. The difference in the present example is that I can specify just what I mean by "small enough", and for the spring I couldn't.

This sort of expansion is typical of why the harmonic oscillator appears so often, even in contexts for which you don't expect it. It will even show up in the study of planetary orbits, section 6.5.

Example

▶ What is the behavior of a piece of wood floating in the water? If you leave it alone, nothing. If you disturb it then it will oscillate up and down according to $\vec{F} = m\vec{a}$. Make it a rectangular block, so the geometry is easy, and let y be the distance the bottom is below the surface (dashed line) of the water. The forces on it come from gravity and the surrounding water, and the latter is computed easily by the rule that Archimedes found for buoyant forces.

$$F_y = +mg - \rho_{\mathsf{w}}gAy = m\ddot{y}$$

Fig. 3.10
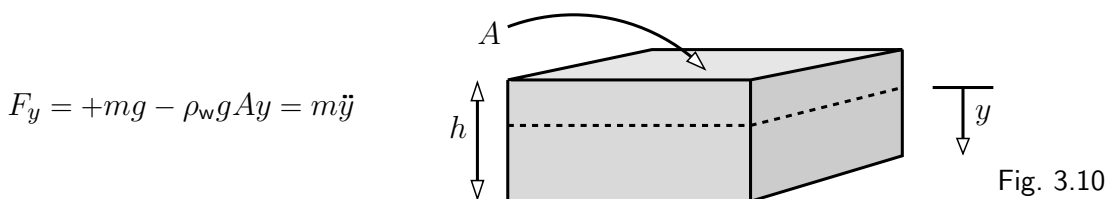
The mass density of water is ρ_{w}, and m is the mass of the box. A is the area of the top. This is an inhomogeneous, linear, constant coefficient differential equation. A solution for the inhomogeneous part is a constant, $y_{\mathsf{inh}} = m/\rho_{\mathsf{w}}A$, the equilibrium position. After that you have a harmonic oscillator with solution $y_{\mathsf{hom}}(t) = C\cos(\omega_0 t + \delta)$ and $\omega_0^2 = \rho_{\mathsf{w}}gA/m$.

In this example the differential equation is exactly a harmonic oscillator, with no approximations needed. Or is it? Water has to flow around the box as the box moves, and fluid flow is notoriously

complicated to figure out. There will at the least be fluid friction (viscosity), and probably much more. But that is another book.

If the box is not rectangular the procedure is the same, but the volume submerged will not be as simple to calculate as Ay, and the resulting differential equation will usually have to be expanded in a power series about the equilibrium point.

3.8 Unstable Motion

Balance a pencil on its point. It is easy to write the equations for its motion; just write the total energy.

$$E = \frac{1}{2}I\omega^2 + Mg\ell \cos\theta \tag{3.49}$$

I is the moment of inertia about the tip and ℓ is the distance from the tip to the center of mass. The angle θ is measured from the top. Simpler conceptually, but completely equivalent mathematically: Put a point mass M on the top of a massless rod of length ℓ and balance the rod on the other end. Then

$$E = \tfrac{1}{2}Mv^2 + Mgh = \tfrac{1}{2}M\ell^2\dot{\theta}^2 + Mg\ell \cos\theta \tag{3.50}$$

For the equation of motion, simply differentiate this with respect to time.

$$\frac{dE}{dt} = 0 = M\ell^2\dot{\theta}\ddot{\theta} - Mg\ell \sin\theta\,\dot{\theta}$$

Cancel the $\dot{\theta}$ factor and the result is

$$\ell\ddot{\theta} - g\sin\theta = 0$$

Just as with the standard pendulum, this inverted pendulum simplifies enormously with a small angle approximation. Then

$$\ell\ddot{\theta} - g\theta = 0 \tag{3.51}$$

and this linear equation is handled the same way as all the other linear equations in this chapter. Assume an exponential.

$$\theta = A\,e^{\alpha t} \;\longrightarrow\; \ell\ddot{\theta} - g\theta = \ell A\alpha^2 e^{\alpha t} - gAe^{\alpha t} = 0 \;\longrightarrow\; \ell\alpha^2 - g = 0 \;\longrightarrow\; \alpha = \pm\sqrt{g/\ell} \tag{3.52}$$

The general solution to the differential equation is now

$$\theta(t) = Ae^{\omega t} + Be^{-\omega t}, \qquad \text{with} \qquad \omega = \sqrt{g/\ell}$$

This is like using the complex exponential solutions for the simple harmonic oscillator. It is often more convenient to use sines and cosines directly in that case, and here it is often more convenient to use hyperbolic sines and cosines as in section 0.2.

$$\theta(t) = Ae^{\omega t} + Be^{-\omega t} = C\cosh\omega t + D\sinh\omega t \tag{3.53}$$

A typical initial condition is to start at rest with initial angle from the vertical θ_0, then

$$\theta(0) = C = \theta_0 \qquad \text{and} \qquad \dot{\theta}(0) = \omega D = 0$$

$$\theta(t) = \theta_0 \cosh\omega t, \quad \text{and if} \quad \ell = 20\,\text{cm} = 0.2\,\text{m} \quad \text{then} \quad \omega = \sqrt{10/0.2} = 7.07\,\text{s}^{-1} \tag{3.54}$$
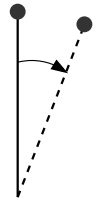
How much time would this take to hit the table starting with an initial angle of one degree?

$$1° \cosh \omega t = 90° \longrightarrow t = \omega^{-1} \cosh^{-1} 90 = 5.193/\sqrt{50} = 0.73 \text{ seconds}$$

How much more time would it take to fall this far if the initial angle is smaller? Perhaps starting at $0.1°$ or $0.01°$? Plug them in and find out just how little difference this change makes: respectively 1.09 and 1.39 seconds.

Maybe the linear approximation made in Eq. (3.51) is not good enough. How much difference would there be in the time it takes to hit the table? You have to go back to Eq. (3.50) and separate variables.

$$E = \tfrac{1}{2} M \ell^2 \dot{\theta}^2 + M g \ell \cos \theta \longrightarrow dt = \frac{d\theta}{\sqrt{2(E - M g \ell \cos \theta)/M \ell^2}} \tag{3.55}$$

Integrate this from θ_0 to $\pi/2$ to get the time to fall. What is E? Assuming that everything starts from rest that is $E = M g \ell \cos \theta_0$, making this

$$\int dt = \int_{\theta_0}^{\pi/2} \frac{d\theta}{\omega \sqrt{2(\cos \theta_0 - \cos \theta)}} \tag{3.56}$$

This is not an integral that you're likely to have seen before, and you cannot do it in terms of the familiar functions (polynomials, roots, trig, exponentials). It is an elliptic integral, and there will be a brief introduction to the subject in section 4.6. For the three angles used one paragraph back, the results using this integral are respectively 0.742 s, 1.069 s, 1.379 s. See problem 3.70.

They are remarkable close to the numbers you get by using the linear approximation in the second paragraph back, so that invites the question: Why? The linear approximation to the sine function is good even at $30°$, but at $90°$, it is terrible. The times predicted for the fall have no right to be as good as they are. Or do they? The explanation of why Eq. (3.54) gives such good results is that as the mass falls it spends most of its time moving slowly, when it is near the top. Most of its time is spent where the small angle approximation is good, and then is zips through the last part of its fall.

3.9 Coupled Oscillations
Oscillators don't always come singly. You get a single oscillator when you have a stable equilibrium and examine small oscillations about that equilibrium point. A molecule consists of several atoms, and each is in some sort of equilibrium. If they are disturbed in any way they will oscillate about that point, but now each oscillator can affect all the other oscillators. That's the subject now.

As a prototype for all oscillators, I'll use a mass on a spring again. Rather, two masses on three springs. When you understand this example you can graduate to the vibrational properties of water and then to those of hemoglobin.
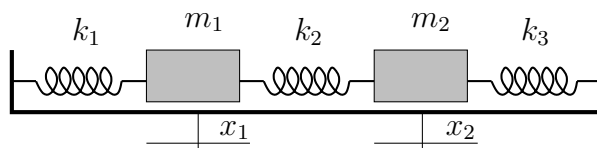

Fig. 3.11

The two coordinates, x_1 and x_2 are measured from the equilibrium position of the respective masses, so the total forces are zero at $x_1 = 0$ and $x_2 = 0$. After that, when the coordinates change then the additional forces that arise will appear because of the change in length of the attached springs. For the first mass, the force from the left spring is $-k_1 x_1$. That's easy. The force from the middle spring is proportional to the compression of that spring, that is to $(x_1 - x_2)$. Nail x_2 down and let x_1 alone vary, then it is easier to see that the force this spring applies is $-k_2(x_1 - x_2)$.

$$m_1 \frac{d^2 x_1}{dt^2} = -k_1 x_1 - k_2(x_1 - x_2) \qquad \text{and} \qquad m_2 \frac{d^2 x_2}{dt^2} = -k_3 x_2 - k_2(x_2 - x_1) \tag{3.57}$$

The reasoning for the second mass repeats that for the first. Notice that the two forces from the middle spring obey Newton's third law, and that the force *on* the middle spring is zero. That's required because of the approximation that the springs have zero mass. (Or did I forget to say that? See problem 3.44 for something about this.)

These equations are not as formidable as they may appear, because the are the familiar linear, homogeneous, constant coefficient differential equations that this whole chapter is devoted to. It's just that they are simultaneous differential equations, so you have to learn a couple of new techniques. One of the key properties of these equations, one that I've already used several times is that the sum of two solutions is a solution. That means that if the pair of functions $x_{1A}(t)$ and $x_{2A}(t)$ are a solution and if $x_{1B}(t)$ and $x_{2B}(t)$ are, then so is their sum.

$$m_1 \frac{d^2(x_{1A} + x_{1B})}{dt^2} = -k_1(x_{1A} + x_{1B}) - k_2\big((x_{1A} + x_{1B}) - (x_{2A} + x_{2B})\big)$$

$$\text{and} \quad m_1 \frac{d^2(x_{2A} + x_{2B})}{dt^2} = -k_3(x_{2A} + x_{2B}) - k_2\big((x_{2A} + x_{2B}) - (x_{1A} + x_{1B})\big)$$

Just look at the terms with subscript A and they already agree. The same with the B-terms. This is like saying that if $\cos \omega_0 t$ and $\sin \omega_0 t$ satisfy Eq. (3.1) then so does $\cos \omega_0 t + \sin \omega_0 t$.

A Digression
There is a little bit of elementary algebra that you probably already know, but I'm going to do it again anyway. *It's important.*

$$\text{Solve} \quad \begin{array}{c} ax + by = 0 \\ cx + dy = 0 \end{array} \quad \text{for } x \text{ and } y. \tag{3.58}$$

To solve for x, multiply the first equation by d and the second by b, then subtract. Now start over and eliminate x by multiplying respectively by c and a and subtracting.

$$\begin{array}{c} dax + dby = 0 \\ bcx + bdy = 0 \end{array} \Rightarrow adx - bcx = 0, \qquad \begin{array}{c} cax + cby = 0 \\ acx + ady = 0 \end{array} \Rightarrow bcy - ady = 0 \tag{3.59}$$

These are both in the form $(ad - bc) \times (x \text{ or } y) = 0$. If the product of two numbers is zero then one or both of the numbers is zero, and this implies that if $ad - bc \neq 0$ then both x and y are zero. The only way to have a non-zero solution for x or y is to have $ad - bc = 0$. The quantity $ad - bc$ is the determinant of these simultaneous equations.

A further point: If $ad = bc$ then the two equations (3.58) are the same equation. At least one of the numbers a and b is non-zero. (Otherwise why bother?) Say $a \neq 0$, then multiply the second of the pair of equations by a to get $acx + ady = 0 = acx + bcy = c(ax + by)$. The second equation is a multiple of the first. (If $c = 0$, look at that case for yourself and show that the result still holds.)

Now return to Eqs. (3.57) and solve them. These equations are linear, homogeneous, constant coefficient equations, and an exponential solution works here just as it did in the previous cases. Now however there are two functions, with two coefficients.

$$x_1(t) = Ae^{\alpha t}, \qquad x_2(t) = Be^{\alpha t} \tag{3.60}$$

Plug in.

$$m_1 \alpha^2 Ae^{\alpha t} = -k_1 Ae^{\alpha t} - k_2(Ae^{\alpha t} - Be^{\alpha t})$$
$$m_2 \alpha^2 Be^{\alpha t} = -k_3 Be^{\alpha t} - k_2(Be^{\alpha t} - Ae^{\alpha t}) \tag{3.61}$$

$e^{\alpha t}$ is a common factor; that's why this method works. Cancel it and rearrange.

$$(m_1\alpha^2 + k_1 + k_2)A - k_2 B = 0$$
$$-k_2 A + (m_2\alpha^2 + k_3 + k_2)B = 0 \tag{3.62}$$

These equations are in precisely the form of Eq. (3.58). If there is to be a non-zero solution for A and B then the determinant of the coefficients must vanish. That determines α.

Instead of grinding through the rather messy algebra of this general problem, take a special and more symmetric case. Let the two masses be equal and let the two springs on the ends be equal.

$$m_1 = m_2 = m \ \text{ and } \ k_1 = k_3 = k \implies \qquad \begin{aligned} (m\alpha^2 + k + k_2)A - k_2 B = 0 \\ -k_2 A + (m\alpha^2 + k + k_2)B = 0 \end{aligned} \tag{3.63}$$

The determinant of these equations must be zero.

$$(m\alpha^2 + k + k_2)^2 - k_2^2 = 0$$

This is the difference of squares, so it factors.

$$(m\alpha^2 + k + k_2 + k_2)(m\alpha^2 + k + k_2 - k_2) = 0 = (m\alpha^2 + k + 2k_2)(m\alpha^2 + k)$$

and the four roots for α are now

$$\pm i\sqrt{\frac{k}{m}} = \pm i\omega \qquad \text{and} \qquad \pm i\sqrt{\frac{k + 2k_2}{m}} = \pm i\omega' \tag{3.64}$$

You're not done. This tells only the frequencies of oscillation, now what are A and B—needed to get the functions x_1 and x_2.

In the equations (3.58), once you know that the determinant is zero, there is really just one equation, as each is a multiple of the other. Pick one, whichever is more convenient, and solve for the ratio of x to y. Geometrically that determines a straight line through the origin in the x-y plane.

For the present application, pick one of the alphas, say $\alpha^2 = -k/m$ and substitute it into the first of Eqs. (3.63).

$$\left[m\left(-\frac{k}{m}\right) + k + k_2 \right] A - k_2 B = 0, \qquad \text{or} \qquad A = B$$

The other choice for alpha with the same equation gives

$$\left[m\left(-\frac{k + 2k_2}{m}\right) + k + k_2 \right] A - k_2 B = 0 \qquad \text{or} \qquad A = -B$$

You now have four solutions to the original equations, and the general solution is a linear combination of all of them.

$$\omega = \sqrt{k/m}, \quad \text{then} \quad x_1(t) = A_1 e^{i\omega t} + A_2 e^{-i\omega t}, \quad \text{with} \quad x_2(t) = x_1(t)$$
$$\omega' = \sqrt{k + 2k_2/m}, \quad \text{then} \quad x_1(t) = A_3 e^{i\omega' t} + A_4 e^{-i\omega' t}, \quad \text{with} \quad x_2(t) = -x_1(t)$$

The total, general solution is now the sum of these.

$$x_1(t) = A_1 e^{i\omega t} + A_2 e^{-i\omega t} + A_3 e^{i\omega' t} + A_4 e^{-i\omega' t}$$
$$x_2(t) = A_1 e^{i\omega t} + A_2 e^{-i\omega t} - A_3 e^{i\omega' t} - A_4 e^{-i\omega' t}$$

OR, often easier to deal with,

$$x_1(t) = C_1 \cos\omega t + C_2 \sin\omega t + C_3 \cos\omega' t + C_4 \sin\omega' t$$
$$x_2(t) = C_1 \cos\omega t + C_2 \sin\omega t - C_3 \cos\omega' t - C_4 \sin\omega' t$$

(3.65)

What do these solutions look like? Pick one: $C_1 \neq 0,\ C_{2,3,4} = 0$.

$$x_1(t) = C_1 \cos\omega t, \qquad x_2(t) = C_1 \cos\omega t$$

(3.66)

The two masses are moving together, so as one moves to the right so does the other. The length of the middle spring does not change, and the $(x_1 - x_2)$ terms in Eq. (3.57) are identically zero. The solution behaves as if the middle spring is not there (left picture).

Fig. 3.12

For the other form of solution, $C_3 \neq 0,\ C_{1,2,4} = 0$.

$$x_1(t) = C_3 \cos\omega' t, \qquad x_2(t) = -C_3 \cos\omega' t$$

(3.67)

Here the two masses are oscillating in opposite directions, and the middle spring is squeezed from both sides. That's why the frequency ω' involves the combination $k + 2k_3$ (right picture). As the left mass moves by x_1 to the right, the k_3 spring is squeezed by $x_1 - x_2 = 2x_1$. This frequency is higher then the first one because there's now a larger restoring force applied to each mass. These solutions are the two modes of oscillation of the system.*

Example
▶ Take as initial conditions that you push the coordinate x_1 to the value x_0 while holding x_2 in place at zero. Then release everything from rest.

$$x_1(0) = x_0 = C_1 + C_3, \qquad x_2(0) = 0 = C_1 - C_3$$
$$\dot{x}_1(0) = 0 = \omega C_2 + \omega' C_4, \qquad \dot{x}_2(0) = 0 = \omega C_2 - \omega' C_4$$

The solutions are $C_2 = C_4 = 0$ and $C_1 = C_3 = x_0/2$.

$$x_1(t) = \frac{x_0}{2} \big[\cos\omega t + \cos\omega' t \big] \qquad x_2(t) = \frac{x_0}{2} \big[\cos\omega t - \cos\omega' t \big]$$

You will have seen equations like these before if you've studied the subject of beats in sound. These become much easier to interpret if you rewrite them using two not-so-well-known trigonometric identities for the sum and difference of two cosines. Recall problem 0.56.

$$x_1(t) = x_0 \cos\left(\frac{\omega + \omega'}{2}t\right)\cos\left(\frac{\omega - \omega'}{2}t\right), \quad x_2(t) = x_0 \sin\left(\frac{\omega + \omega'}{2}t\right)\sin\left(\frac{\omega' - \omega}{2}t\right)$$

(3.68)

---

* If you want to see this normal mode problem fully worked out *without* making assumptions of symmetry, so that $k_1 \neq k_3$ and $m_1 \neq m_2$, look up chapter four of the text Mechanics by Keith Symon. It is no different conceptually but it involves much more algebra.

ω' is larger than ω because of its extra $2k_2$, representing the extra force from the middle spring pushing the masses apart. Draw graphs of the motion for the case that k_2 is not very big, making ω' only a little larger than ω. That makes the terms involving $\omega' - \omega$ in Eq. (3.68) oscillate much more slowly than the others. A graph of $x_2(t)$ *and* the two factors that compose it is
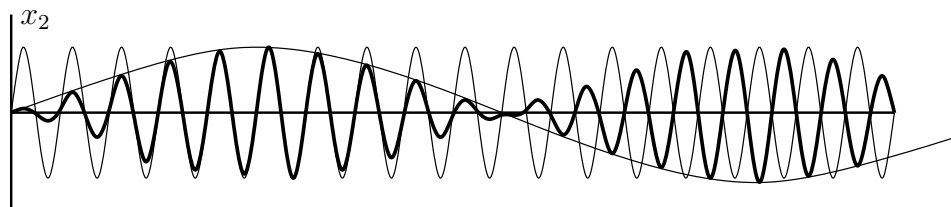

Fig. 3.13

This graph shows the functions from the second of the equations (3.68). The two sine factors that multiply to form x_2 are drawn lightly, and the product is drawn as the heavy graph. When you draw the same sort of graph for x_1 you will see that the energy that was given to mass #1 gradually shifts over to #2 and then back again to #1. This is the same sort of phenomenon you hear when you listen to two musical notes that have slightly different frequencies and that are played together. The periodic pulsations in intensity are called beats in that context, and the mathematics describing them is the same as here, with the amplitude of oscillation of mass #1 slowly going from zero to maximum and back again.

To find a life-saving application of coupled oscillators, go to section 10.5. There you can see how a tuned mass damper, which is a special kind of coupled oscillator, can be used to make tall buildings more resistant to damage from earthquakes. You may choose to skip the details there, but if nothing else, look at the interpretation accompanying Figure 10.6.

3.10 Normal Modes
The two modes of oscillation represented by the equations (3.66) and (3.67) are called *normal modes* of oscillation. The reason for this is not that the word normal means "typical" or anything like that. It is used in the sense of "perpendicular" as when two lines intersect at right angles.

You can think of the components of a mode of oscillation as the coefficients in the equation (3.66), so that the first (ω) mode is represented by the column matrix

$$P = \begin{pmatrix} C_1 \\ C_1 \end{pmatrix}$$

Similarly Eq. (3.67) states the components for the second (ω') mode as

$$Q = \begin{pmatrix} C_3 \\ -C_3 \end{pmatrix}$$

These are orthogonal to each other in the sense of the scalar product

$$\tilde{P} \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} Q = \begin{pmatrix} C_1 & C_1 \end{pmatrix} \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} C_3 \\ -C_3 \end{pmatrix} = m_1 C_1 C_3 - m_2 C_1 C_3 = 0 \qquad (3.69)$$

because $m_1 = m_2$, and where the tilde over the P means transpose. This is like the scalar product

$$\vec{A} \cdot \vec{B} = AB \cos \theta = \left(A_x \hat{x} + A_y \hat{y} \right) \cdot \left(B_x \hat{x} + B_y \hat{y} \right) = A_x B_x + A_y B_y$$

and the vectors $\vec{A}$ and $\vec{B}$ are perpendicular if $\vec{A} \cdot \vec{B} = 0$. This example could be misleading because the equal masses here make the example so special, and you may think that you could simply omit

the mass matrix in the middle of Eq. (3.69). Starting with Eq. (3.63) the masses m_1 and m_2 were set equal, and if you go back to the equation preceding it (3.62), carrying out all the calculations, you would find that the product in Eq. (3.69) would still be zero provided that you have the two masses m_1 and m_2 in the product. The modes themselves will be more complicated, but the product

$$(C_1^* \quad C_1'^*) \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} C_3 \\ C_3' \end{pmatrix} \quad \text{will still be zero.} \tag{3.70}$$

This redefinition of the scalar product may be unfamiliar, but it is part of a larger picture that is explored in chapter ten, and it is a special case of Eq. (10.22). It can wait until then. The complex conjugates weren't needed for the preceding examples, but when you write the solutions in terms of $e^{\pm i\omega t}$ instead of sines and cosines they become necessary. The coefficients will be complex too in that case. Refer to the solution in Symon's text, mentioned in the footnote on page 103 to see that more general solution, and then you can find how the masses enter into the orthogonality.

You can *use* this orthogonality relation as a guide to determine the shape of the modes in more complex cases. See for example problem 3.80.

3.11 Green's Functions

Is inspired guesswork the only way to solve for the inhomogeneous solution to (3.28)? No. There is a systematic method developed by Green, a method that turns out to be of astoundingly great utility. It is central in the subject of scattering theory in quantum mechanics and quantum field theory.*

Take the specific problem of the undamped harmonic oscillator with a forcing function.

$$m\frac{d^2 x}{dt^2} + kx = F_{\text{ext}}(t) \tag{3.71}$$

The idea behind this method is to treat the added force as a sum of impulses, and because of the atomic nature of matter, this is not so far from the truth anyway. If the mass is at rest and you kick it, you give it an initial velocity and after that it obeys the homogeneous equation $m\ddot{x} + kx = 0$. An initial velocity of v_0 at time $t = 0$ produces $x(t) = (v_0/\omega_0)\sin\omega_0 t$ for $t > 0$. If you administer the kick at time t' instead of zero, that simply shifts the origin, so

$$x(t) = \frac{v_0}{\omega_0}\sin\omega_0(t - t') \qquad (t > t') \tag{3.72}$$

I specify explicitly that the solution $= 0$ for $t < t'$. This says that it's the external force that is causing the motion, not some distant initial conditions.

How is this v_0 related to the applied force? A force F_1 applied for a (short) time interval Δt_1 changes the momentum by

$$\int dp = \int F\, dt \quad \longrightarrow \quad F_1\Delta t_1 = mv_1$$

and this v_1 is the v_0 of the preceding equation. If you later give a second kick $F_2\Delta t_2$ you don't have to know the result of the first kick to find the result of the second kick. The key fact about the equation that you're trying to solve is that it is linear.

$$m\frac{d^2 x_1}{dt^2} + kx_1 = F_1(t) \qquad \text{and} \qquad m\frac{d^2 x_2}{dt^2} + kx_2 = F_2(t)$$

$$\text{then} \qquad m\frac{d^2(x_1 + x_2)}{dt^2} + k(x_1 + x_2) = F_1(t) + F_2(t)$$

---

* George Green's life as a scientist is the stuff of fiction. He started life in the family business, as a miller, and was largely self-taught in mathematics when he published his most famous works.
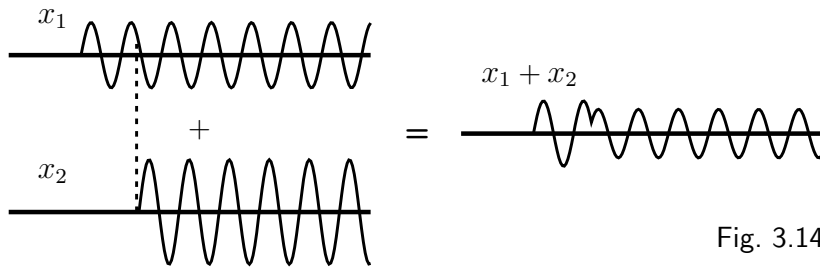
Fig. 3.14

In these graphs you apply an impulse at one time (t_1) to get the first graph. Separately apply an impulse at a later time (t_2) to get the second graph. The third graph is the sum of the first two, and it is the solution to the original differential equation when both impulses are applied.

In the example that started from Eq. (3.32), the process required solving the problem completely up to the second time (T in that case) and then using the terminal conditions before T as the initial conditions after T, leading to the solution (3.35). In the current problem, with two kicks, a similar solution would use the terminal conditions just before the second kick as the initial conditions *at* the second kick. Green's insight was that the linearity of this differential equation allows you to avoid that complexity and to handle the two impulses independently.

But what about another sort of forcing function, one that is not just the sum of a couple of impulses? Another insight: *every* function is a sum of impulses, or at least it is a limit of such a sum. Just as when you define an integral, you divide the independent variable into small pieces and approximate the function by a sequence of steps.
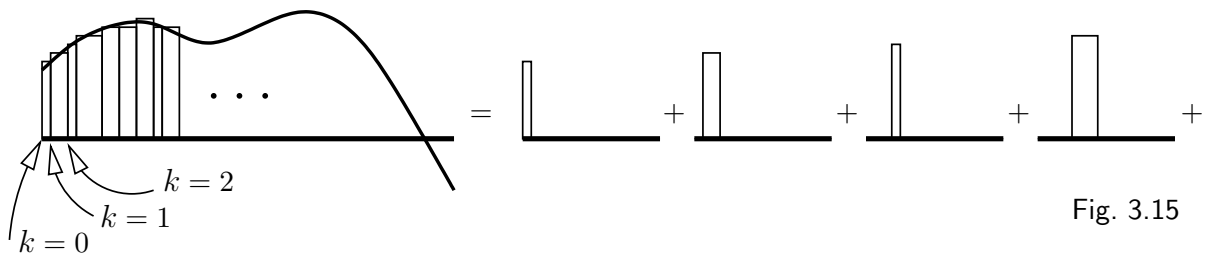

Fig. 3.15

The kth interval has width $\Delta t_k = t_k - t_{k-1}$, and the force there has height $F(t_k)$. One such box is an impulse, or at least it will be when eventually $\Delta t_k \to 0$. Take a typical one of these boxes and compute its effect on the harmonic oscillator, assuming that Δt_k is small. The equation (3.72) becomes

$$F(t_k)\Delta t_k = mv_{\text{init}} \qquad \text{so} \qquad x_k(t) = \frac{F(t_k)\Delta t_k}{m\omega_0}\sin\omega_0(t - t_k), \quad (t > t_k)$$

As when there were just two impulses, add the results of all these impulses to get the (approximate) result for the motion of m.

$$x(t) \approx \sum_{k=1}^{N} \frac{F(t_k)\Delta t_k}{m\omega_0}\begin{cases} \sin\omega_0(t - t_k) & (t > t_k) \\ 0 & (t < t_k) \end{cases}$$

As with any other integral, improve the approximation by taking smaller time intervals, and finally make it exact in the limit. That the response to the impulse is zero until the impulse happens dictates the limits of integration.

$$x(t) = \lim_{\Delta t_k \to 0} \sum_{\substack{k \\ (t_k < t)}} \frac{1}{m\omega_0} F(t_k)\,\Delta t_k\,\sin\omega_0(t - t_k) = \frac{1}{m\omega_0}\int_{-\infty}^{t} dt'\,F(t')\sin\omega_0(t - t')$$

Note that you have to be careful to use different notations for the various time variables, especially to distinguish the integration variable t' from the independent variable t.

$$G(t - t') = \frac{1}{m\omega_0} \begin{cases} \sin\omega_0(t - t') & (t > t') \\ 0 & (t < t') \end{cases} \qquad \text{so that} \qquad x(t) = \int_{-\infty}^{t} dt'\, G(t - t')F(t') \quad (3.73)$$

and this function is called Green's function for this differential equation. Because this Green's function is defined to be zero for negative arguments $t < t'$, you could write the limits on this integral as $\pm\infty$ without changing anything. I prefer the redundant form, using an explicit limit.

Does it Work?
Apply this to the example Eq. (3.32). The solution is now a plug-in:

$$x(t) = \int_{-\infty}^{t} dt'\, G(t - t')F(t') = \frac{1}{m\omega_0} \int_{-\infty}^{t} dt'\, \sin\omega_0(t - t') \cdot \begin{cases} F_0 & (0 < t' < T) \\ 0 & (\text{otherwise}) \end{cases}$$

This integral is

$$\int_{-\infty}^{t} = \int_{-\infty}^{0} + \int_{0}^{t}$$

The first integral is zero, and the next has two cases: $t < T$ and $t > T$.

$$\int_{0}^{t} dt'\, F(t')\sin\omega_0(t - t') = \begin{cases} \int_{0}^{t} dt'\, F_0 \sin\omega_0(t - t') & (0 < t < T) \\ \int_{0}^{T} dt'\, F_0 \sin\omega_0(t - t') & (T < t) \end{cases}$$

$$\text{then} \qquad x(t) = \frac{F_0}{m\omega_0^2} \begin{cases} 1 - \cos\omega_0 t & (0 < t < T) \\ \cos\omega_0(t - T) - \cos\omega_0 t & (T < t) \end{cases}$$

This is identical to the solution Eq. (3.35), though you will have to do a small amount of algebra to verify that fact. Notice that this solution is continuous at $t = T$ and it has a continuous derivative there too. This is so despite the fact that the Green's function has a discontinuity in its derivative and the force function is discontinuous at 0 and T. You are integrating over these points, and that smooths the result.

Other Green's Functions
You have seen Green's functions before, though perhaps not under that name. What is the electrostatic potential produced by a point charge? That's a Green's function. The total potential produced by a lot of charges is the sum (integral?) over the potentials produced by the individual charges. Instead of an impulse in time to produce an oscillation, this is an analogous potential function of position produced from a point in space.

The potential from a single point charge at the origin is $\phi(r) = q/4\pi\epsilon_0 r$, and the potential from several point charges is the sum over such contributions. For two charges the potential is

$$\frac{q_1}{4\pi\epsilon_0|\vec{r} - \vec{r}_1|} + \frac{q_2}{4\pi\epsilon_0|\vec{r} - \vec{r}_2|}$$

For many charges, including a continuous distribution of them you have

$$G(\vec{r} - \vec{r}') = \frac{1}{4\pi\epsilon_0|\vec{r} - \vec{r}'|} \quad \longrightarrow \quad \phi(\vec{r}) = \int dq'\, G(\vec{r} - \vec{r}') = \int d^3r'\, \rho(\vec{r}')G(\vec{r} - \vec{r}')$$

ρ is the volume charge density, and the Green's function in this case is conventionally defined without the q.

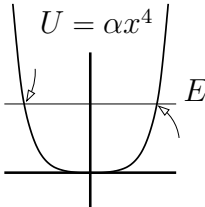The analog of Eq. (3.73) for the damped harmonic oscillator is

$$G(t - t') = \frac{1}{m\omega'} \begin{cases} e^{-\gamma(t-t')} \sin\omega'(t - t') & (t > t') \\ 0 & (t < t') \end{cases} \tag{3.74}$$

You can easily derive this yourself, problem 3.66.

3.12 Anharmonic Example˙
What happens if the approximation of a linear harmonic oscillator isn't even close to being right? There are examples in problems 2.19 and 3.32, but the latter one is sort of extreme. The first really different potential energy after x^2 that still gives oscillations is x^4. To solve that problem, write the energy equation directly and separate variables.

$$E = \frac{1}{2}m\dot{x}^2 + \alpha x^4 \longrightarrow \dot{x}^2 = \frac{2E}{m} - \frac{2\alpha}{m}x^4$$
$$\int \frac{dx}{\sqrt{1 - \alpha x^4/E}} = \int \sqrt{2E/m}\, dt \tag{3.75}$$

If you know about elliptic integrals and elliptic functions you can solve this for $x(t)$ in terms of known functions. If you don't know about them then they are unknown functions, but now they are known unknown functions and not unknown unknown functions. (See section 6.13 for more.)

In any case, it is sometimes easier to solve such problems numerically, and since there are some technical points about this example, I will proceed along this route for a while. The first thing to do is to change variables to get rid of a lot of parameters. Make all the E's, m's, and 2's disappear.

$$\text{Let}\quad \frac{\alpha}{E}x^4 = y^4 \quad \text{and} \quad \left(\frac{\alpha}{E}\right)^{1/4}\sqrt{2E/m}\, t = u \quad \text{then} \quad \int_0^y \frac{dy}{\sqrt{1 - y^4}} = u \tag{3.76}$$

Here I chose the initial condition that $x(0) = 0$, so $y = 0$ when $u = 0$. It also sets the derivative $dy/du = 1$ at $u = 0$. This integral expresses time as a function of position instead of the reverse, but you take what you can get.

If you try a brute-force numerical integration of this equation (or let a computer do the dirty work), you will run into trouble when you get near $y = 1$. That is a turning point (the arrows in the drawing), and the integrand blows up there. The integral is still finite, because the function you're integrating goes to infinity only as one over a square root, and that's integrable, but it can still give a computer hiccups.

The trick to get around this annoyance is to make a change of variables. Factor out the bad behavior near $y = 1$.

$$1 - y^4 = (1 - y)(1 + y + y^2 + y^3), \quad \text{so let} \quad \frac{dy}{\sqrt{1 - y}} = dz \longrightarrow z = 2 - 2\sqrt{1 - y}$$

with $y = 0 \longleftrightarrow z = 0$ and $y = 1 \longleftrightarrow z = 2$. You now have the integrals

$$u = \int_0^z \frac{dz}{\sqrt{1 + y + y^2 + y^3}} \quad \text{where} \quad y = z - z^2/4 \quad (0 < z < 2) \tag{3.77}$$

The point $z = 2$ corresponds to the $y = 1$ turning point of the oscillator. It comes to a stop there and reverses direction. The original integral in terms of y worked only up to that turning point, where it hit a singularity. This form in terms of z applies all the way back to when the mass returns to the origin. That is, $0 < z < 4$, though because of the periodicity it is only necessary to compute it up to $z = 2$.

The procedure to compute this is

0. Pick an integer $N \geq 1$.
1. Divide the interval $0 \leq z \leq 2$ into N pieces, $\Delta z = 2/N$.
2. Choose your favorite integration method (midpoint, trapezoid, Simpson, . . .).
3. Set $k = 1$ and $z_0 = 0$ and $z_1 = \Delta z$ and $u_0 = 0$.
4. Integrate from z_{k-1} to $z_k = z_{k-1} + \Delta z$.
5. $u_k = u_{k-1} +$ this integral.
6. Compute y_k from z_k.
7. Unless you're done $(k = N)$, increment k by 1 and go back to step 4.

This provides two arrays: the u_k (time) and the y_k (position). Plot one versus the other and then remember that the function is periodic so you can use the same results to plot the rest of the graph.
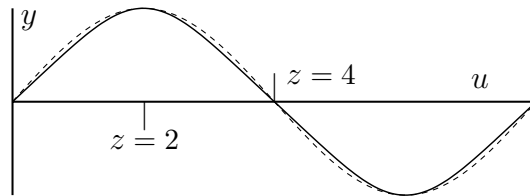

Fig. 3.16

The solid line is this quartic oscillator, and the dashed line is a pure harmonic oscillator having the same period and amplitude. They aren't as different from each other as you might have expected. The value of u when $y = 1$ (and $z = 2$) is $u = 1.311029$, and that corresponds to a quarter of a period.

3.13 Forced Pendulum, Qualitatively
The simple harmonic oscillator provides an excellent model for oscillations about equilibrium. A mass on a spring, a pendulum, a planet in an almost circular orbit, an atom vibrating in a crystal. . . . Even with the pendulum however, it is apparent that we are omitting something. In that case more than any of the others I was explicit in approximating the sine by its argument. In the series expansion $\sin\theta = \theta - \theta^3/6 + \cdots$ I dropped everything but the first term and looked back only once, in section 4.6. Is this approximation good? How good? In such an approximation are there qualitatively different phenomena that I am ignoring or are the errors only those of the next decimal place?

Regarding the last question, there are some of each. For the pendulum, as long as the oscillations don't take it near the top, there are only quantitative differences. Apply an oscillating force to it however, and there are other, radically different, things that can happen, starting with hysteresis and on up to chaos.

Before attacking this material, go back to section 3.6 on the forced harmonic oscillator, and read it again. Especially, understand the graph Figure 3.5 (a) and the material that leads up to it. What follows will then make a whole lot more sense.

Recall the phenomenon of resonance from section 3.6. Apply an oscillating force to a simple, undamped harmonic oscillator, and if the forcing frequency Ω is close to the oscillator's natural frequency, there is a large response. If the forcing frequency is *at* the natural frequency you get a response that runs away and grows without limit. Include some friction, and it tames the result so that it stays finite, as in this graph and the graph from Figure 3.5(a). Here, $b/m\omega_0 = 0.1$

Fig. 3.17
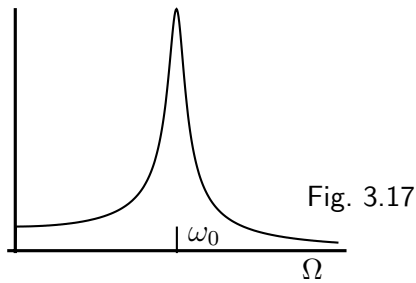
If this oscillator is a pendulum, then this analysis is good only to the extent that the small angle approximation is good, allowing the pendulum to be approximated by a simple harmonic oscillator. If this graph represents the maximum angle of oscillation for the pendulum and if the vertical scale makes the peak of the graph be $30°$ then this holds no surprises because $30° = \pi/6 = 0.5236$, and that is still very close to $\sin 30° = 0.5$. What if there is less damping, so that the peak of this resonance curve would take the pendulum to $90°$ or $150°$? Then this small angle approximation breaks down. More importantly, it breaks down not only quantitatively, but qualitatively. New phenomena appear, ones that don't show up in the small angle approximation.

The natural period of a pendulum changes as the maximum angle increases. When the maximum angle gets up to $179.9°$, it stops and waits while it considers whether to return or to go over the top. The period is then much longer at this large angle of oscillation, and it is fully solved in Eq. (4.63).

Start this analysis by pushing the pendulum at a frequency well above the natural frequency, so that you're to the right of the peak in the curve of Figure 3.17 The response is then proportionate, following the right side of a graph just like the one above. Next, gradually decrease the forcing frequency. The response gets bigger. Do it again, it's still bigger.

Fig. 3.18

Fig. 3.19
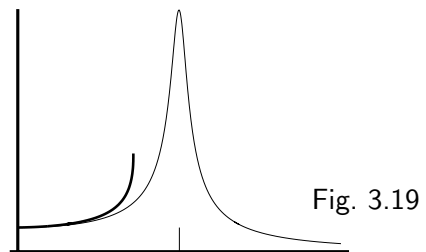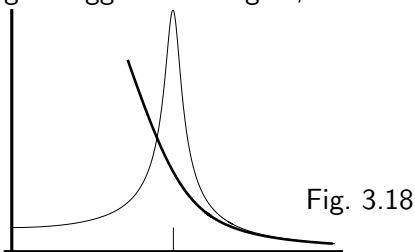
Now comes the twist: As the amplitude of the pendulum gets bigger, its natural oscillation period gets bigger, so the natural frequency gets smaller, moving to the left of the original ω_0: $\omega = 2\pi/T$. This means that as you move to the left in this graph, Figure 3.18, you are not as close to the resonance peak as you expected, and the response is not quite as large as the linear theory (and the original graph) would have said. The response follows the other, the heavier curve in that figure.

Keep going. Decrease the driving frequency Ω some more, and the amplitude gets bigger and the natural frequency again decreases and the response is still smaller than the simple previous graph above would say. As you climb the right-hand side of the same graph, it responds by leaning away toward the left. If you keep moving to the left, you can get to a frequency *below* the original ω_0 itself and the curve will still keep climbing and leaning left.

Now start over, and approach resonance from the other side. Start with a low frequency force and gradually increase the frequency, approaching ω_0 from the left. As with the analysis of the right-hand side, the response will grow as you move the forcing frequency toward the resonance, but as the amplitude of the oscillation increases, the natural frequency again decreases (that part is the same), but this time it means that the response to the oscillating force is *larger* than you would expect from the linear theory: Figure 3.19.

Continue to the right, and the response grows enough that it decreases the resonant frequency enough that the system runs away! The curve disappears from underneath you. Does it run all the way

away to infinity? No. It is not that bad. What the oscillation will do is to flip over and meet the left leaning resonance curve that you found on the way down from high frequencies — jumping to a point on the other curve.
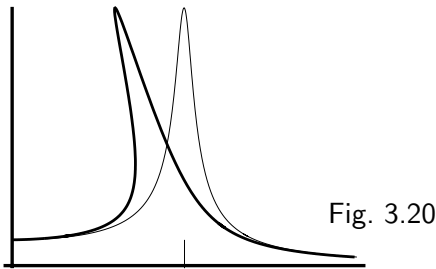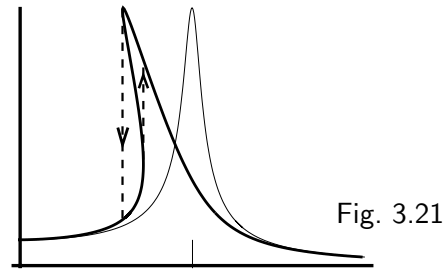

Fig. 3.20

Fig. 3.21

This picture 3.20 puts it together. The resonance curve is still continuous, and it maintains the qualitative shape that it had before. It is however *no longer a function.* It is not single valued. One value of the driving frequency can have two (or even three) possible values for the amplitude of the response. The next picture, 3.21, includes the dashed line to show what happens as you move from low frequency forcing on up toward the resonance; The response jumps to the upper curve. Similarly, if you move down from high frequencies you will reach a point at which the resonance curve disappears from beneath you and the system then jumps down to the lower amplitude motion. You can go counterclockwise around this loop as much as you want, but you will never find yourself on that middle portion of the curve in between the dashed lines. That part represents unstable motion. This multi-valued phenomenon, going around a loop, is called hysteresis.

What is the difference between these two solutions that can appear at the same forcing frequency? One will be on the lower branch and one on the upper with a larger amplitude, but how else do they differ? The answer lies in looking at the other curve that you need to describe the simple linear resonance, Figure 3.5(b). That shows the *phase* of the response. At frequencies below ω_0 the mass oscillates more-or-less in phase with the applied force. For frequencies higher than ω_0 they tend to be out of phase — approaching a $180°$ difference at high enough frequency. When the nonlinear oscillator makes a jump from the lower branch to the upper, as in Figure 3.21, the motion will change in two ways: The size of the oscillation will be bigger, but also it will switch from being pretty much in phase with the forcing function to being closer to $180°$ out of phase with it.
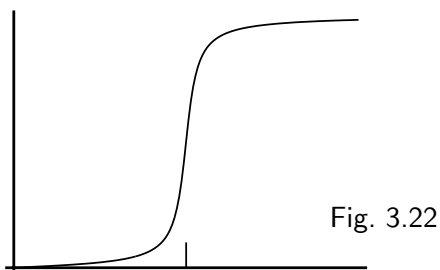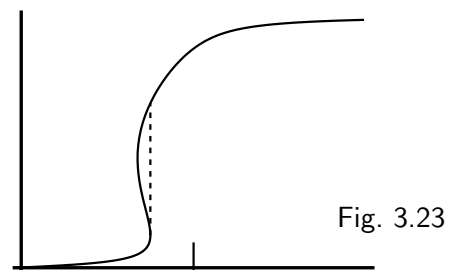

Fig. 3.22

Fig. 3.23

In Figure 3.22 you have one of the curves representing the phase lag in the simple harmonic oscillator response. It goes from 0 to π, and is one of the curves in Figure 3.5 (b), the one with $b/m\omega_0 = 0.1$ as were the curves on the preceding two pages. Figure 3.23 is approximately what the corresponding curve looks like for a pendulum, and the dashed line is the phase jump, paralleling the amplitude jump of Figure 3.21.

Complications

So far, this description resembles the simple analysis that you saw in the forced harmonic oscillator. It has some unexpected features, but nothing too bad. If only.

Continue this analysis and you will find a result that is not hard to see coming: If the initial conditions provide enough energy the pendulum can go over the top. Then you will not have oscillations

at all but circular motion instead. That's easy to predict—just put way more energy than $2mg\ell$ into the system and it is unavoidable. Now what else can happen? There is an intermediate state that occurs when the initial conditions provide an energy close to the amount needed just to carry it over the top (close to $2mg\ell$). In this circumstance the slightest change in an initial position or velocity will change the motion dramatically. It is incredibly sensitive to initial conditions, and this sort of behavior is commonly associated with chaos.

Exercises

1 What is the power series expansion of $\sec x$ about $x = 0$ out to terms in x^2? If you don't remember the derivative of a secant, just recall that $\sec x = 1/\cos x$. OR, just use $\sec x = 1/\cos x$ and $\cos x = 1 - \left(\frac{1}{2}x^2 - \frac{1}{24}x^4 + \cdots\right)$, then apply the geometric series Eq. (0.1)(h).

2 What is the power series expansion of $1/\left(a^2 + (x - b)^2\right)$ about $x = b$ out to quadratic terms What is its similar expansion about $x = 0$?

3 Where are the equilibrium points for the potential energy $U(x) = A\cos kx$ for $A > 0$. Which are stable and which are unstable?

4 Do the dimensions of Eq. (3.19) come out correctly?

5 Perhaps you're more familiar with l'Hospital's rule for evaluating certain types of limits; apply it to Eq. (3.21) to see if you get the same answer as there.

6 For the equation Eq. (3.24), it says in that text that the solution $\phi(t) = A\cos \omega t$ *won't work*. Try it and see.

7 Write in the form $re^{i\theta}$: (**a**) $1 - i$, (**b**) $(1 + i)^3$, (**c**) $1/(-1 - i)$. Pictures!

8 Write in rectangular form: (**a**) $5e^{-i\pi/2}$, (**b**) $e^{-1}e^{i\pi}$, (**c**) $e^{137i\pi/4}$.

9 In Eq. (3.53), what are C and D in terms of A and B?

10 Let α be an arbitrary complex number, $\alpha = \gamma + i\delta$, then $\frac{d}{dt}e^{\alpha t} = \alpha e^{\alpha t}$. Write this out using Euler's formula and verify that the differentiation works as promised.

11 Let α be an arbitrary complex number, then $\frac{d}{dt}e^{i\alpha t} = i\alpha e^{i\alpha t}$. Write this out using Euler's formula and verify that the differentiation works as promised.

12 Graph the exact (not approximate) solution in Eq. (3.20). Take $b/2m = 1$ and $\eta = .5$ (in some units).

13 Solve the problem Eq. (3.29) with $x(0) = 0 = \dot{x}(0)$. This time, use Green's function methods.

14 What is the general solution for $md^2x/dt^2 - kx = 0$? And sketch some of them.

15 A potential energy is given to be (in one dimension) $U(x) = \alpha x^2 + \beta x^4$, where $\alpha > 0$ and $\beta > 0$, and a mass m moves in this potential.
(**a**) Show graphically why there are oscillations.
(**b**) For small motions about equilibrium, find the motion, $x(t)$.

16 Two masses m_1 and m_2 are connected by a spring and there is no damping. When m_1 is held fixed, m_2 oscillates with a frequency ω_0. Find the frequency of oscillations when m_2 is held fixed instead.

17 Sketch the potential energy $U(x) = U_0 a^2 \cos(\alpha x)/(a^2 + x^2)$ and describe the sort of motions that occur. If there are oscillations, how do they behave?

18 What solutions do appear for $m\ddot{x} + b\dot{x} - kx = 0$ and for $m\ddot{x} - b\dot{x} + kx = 0$?

19 Just after Eq. (3.76) it states that $dy/du = 1$ at $u = 0$. Show why.

20 Use Eq. (3.73) to solve the problems presented in Eq. (3.29) with the initial condition that it is dropped from rest.

21 In going from Eq. (3.29) to Eq. (3.31), once you have the homogeneous solution in step 1, change the order, applying the initial conditions first and then adding the inhomogeneous term. What happens?

Problems

3.1 Find the relations between the other forms of the solutions given in Eq. (3.5)

3.2 Find the general solution to

$$\frac{d^2\Lambda}{d\aleph^2} = -\Xi\Lambda, \qquad (\Xi > 0)$$

3.3 From the cube of $e^{i\phi}$, derive expressions for the sine and cosine of triple angles in terms of single angles.

3.4 A particle of mass m is subject to a force computed from the potential energy: $U(x) = -U_0 e^{-\alpha x^2}$. **(a)** Set up $F_x = ma_x$ for this mass and in the approximation that it moves only a small distance from the origin find the general solution for the motion, $x(t)$. **(b)** At time $t = 0$ it is at the origin and you give it an initial velocity $v_x(0) = v_0$. Find the later motion. How big can the "small distance" be? That is, for the approximate solution to the equation to be good, how small must the oscillations be — and check the dimensions of your answer. Which of the methods on page 81 will you choose?

3.5 Express $-\cos\omega t - \sin\omega t$ as $D\sin(\omega t + \delta)$, finding D and δ. When you're done, take some special cases for t (or ωt) and check that your result matches what you started with.

3.6 Derive the formulas for the sine and cosine of sums of angles. The derivation should occupy not much more space than does Eq. (3.12).

3.7 Use Euler's formula to evaluate $\sqrt{i}$, both of them, with a sketch in the complex plane. Since you already have i expressed as a complex exponential, what is i^i?

3.8 From $e^{i\phi}$ you also have $e^{-i\phi} = \cos\phi - i\sin\phi$. Add and subtract these to express $\cos\phi$ and $\sin\phi$ in terms of complex exponentials. From these, what are $\cos(i\phi)$ and $\sin(i\phi)$? Use the standard equation for functions of sums of angles (from problem 3.6) to evaluate $\cos(x+iy)$ and $\sin(x+iy)$. Check your result numerically by trying $\cos(1+i)$ in Google. Similarly, check the results of problem 3.7.

3.9 A mass m is subject to a frictional force $F_x = -bv_x$, and $x(0) = 0$ with zero velocity. Starting at time $t = 0$ there is an additional force $F_x = ct$, where c is a constant. The motion is claimed to be

$$x = (mc/b^2)\left[-\tfrac{m}{b}\left(e^{-bt/m} - 1\right) - t\right] + \tfrac{c}{2b}t^2$$

Analyze this thoroughly to see if it is plausible.

3.10 Derive the result claimed in the preceding problem. Ans: See the preceding problem.

3.11 Derive the critically damped solution, Eq. (3.21), for the harmonic oscillator starting from the underdamped (oscillating) solution, Eq. (3.19).

3.12 In the damped harmonic oscillator specify the initial conditions that $x(0) = x_0$ and $v_x(0) = 0$. **(a)** Evaluate the solution in the general case and finally **(b)** take a limit to get the critically damped solution. Ans: (b) $x_0\left(e^{-\alpha t} + \alpha t e^{-\alpha t}\right)$, $\alpha = b/m$

3.13 If there is no damping in the harmonic oscillator and there is a forcing term $F_0 \cos\Omega t$, you can easily guess a solution for the inhomogeneous equation. For fixed ω_0, plot the amplitude of this inhomogeneous solution versus Ω. Compare the result to the graphs on page 93.

3.14 If the damping term is missing in Eq. (3.36) you easily guessed a solution for the inhomogeneous equation in the preceding problem. Add the solution of the homogeneous equation and then apply the initial conditions that $x(0) = 0$ and $\dot{x}(0) = 0$. Analyze the resulting solution in detail, small t, large t, $\Omega \ll \omega_0$, $\Omega \gg \omega_0$. Ans: $\left[F_0/m(\omega_0^2 - \Omega^2)\right]\left[\cos \Omega t - \cos \omega_0 t\right]$

3.15 In the preceding problem what happens to your solution if $\Omega = \omega_0$? Take the limit $\Omega \to \omega_0$ of your solution there, and then analyze the result. *Graphs of course!* Ans: $x = (F_0/2m\omega_0)t \sin \omega_0 t$

3.16 In the paragraph following Eq. (3.43), it says "I'm assuming that b is small." This really makes no sense, because b is not dimensionless. It has to be compared to something. What?

3.17 You can solve Eq. (3.36) without resorting to complex exponentials. You simply assume a solution of the form $x(t) = A\cos(\Omega t - \delta)$ and plug it in to the differential equation. On the left you get a combination of sines and cosines of $(\Omega t - \delta)$ and on the right you have a $\cos \Omega t$. You can expand the left side using trigonometric identities and then making the two sides match. It's a lot of algebra. OR, you can write the right-hand side as $F_0 \cos\left((\Omega t - \delta) + \delta\right)$ and expand *that*. Suddenly the mass of algebra becomes trivial. Compare the result to the solution in the text. Ans: The real part of Eq. (3.43)

3.18 (**a**) Write the potential energy for a simple pendulum as mgh and express it in terms of the angle with respect to the vertical. Expand this for small angles about the equilibrium point at the bottom to show the shape of the energy function there. Do it again for the equilibrium point at the top. Graphs?
(**b**) Express the kinetic energy in term of $\dot{\theta}$ and compute dE/dt just as in Eq. (3.46) to get the equations of motion in the two cases.
(**c**) Assume that the position is near the top, and get the differential equation for small deviations from the vertical there. What is the solution to this equation for small angles from the vertical, assuming initial conditions $\theta(0) = 0$ and $\dot{\theta}(0) = \dot{\theta}_0$. Refer to section 0.2.

3.19 The potential energy function $U(x) = Ax^3$ has $F_x = 0$ at $x = 0$, but it doesn't look like either of the two parabolic potentials that I analyzed for stability or instability. What happens to the conditions for equilibrium (stable, unstable) if I add ϵx to the potential energy, where ϵ is as small as I like, both positive *and* negative. Describe the possible motions in all cases and if there is an equilibrium, find where it is and determine its stability.

3.20 What is the behavior of the damped harmonic oscillator if it starts at $x = x_0$ and $v_x = 0$ and is *strongly* overdamped?

3.21 Using the equations (3.24), find the tension in the pendulum's cord as a function of time. Use $\phi = A\sin \omega_0 t$ as the small angle solution to the second ($\ddot{\phi}$) equation and keep terms to a consistent order in the amplitude A. Examine various special cases of your result. How do the middle and end points compare, and can you see why those results are as they are? Ans: $mg\left[1+\frac{1}{4}\phi_{\max}^2(1+3\cos 2\omega_0 t)\right]$

3.22 Derive the equation of motion for a simple harmonic oscillator directly from energy considerations as in Eq. (3.46), but now include friction. The time derivative of the total mechanical energy is the power from other forces, and that is $F_x\dot{x}$ in one dimension. Here the F_x is specifically the frictional force. In the model used in Eq. (3.14) that is $-b\dot{x}$.

3.23 The standard conservation equation for the simple harmonic oscillator is $E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2$. Consider instead the combination $\frac{1}{2}ma^2 + \frac{1}{2}kv^2$. Is this dimensionally consistent? Show that it is conserved too. Will something like this happen with any potential energy other than that of a simple harmonic oscillator?

3.24 What is the behavior of the solution Eq. (3.31) near the bottom of the oscillation. *I.e.* , do a series expansion about that point, near the time π/ω_0. Can you anticipate the answer before doing the expansion to derive it? And even if you can't, explain afterwards why you should have been able to.

3.25˙A mass m moves in one dimension in the potential energy $U(x) = ax^3 - bx + c$, where a, b and c are positive constants. Find where the mass is in stable equilibrium and find the frequency of small oscillations about there. Ans: $\omega_0 = \sqrt{2/m}\,(3ab)^{1/4}$

3.26¨An undamped harmonic oscillator is subjected to a force $F_x = \alpha + \beta t^2$ starting with the initial conditions $x(0) = 0$ and $v_x(0) = 0$. Find the subsequent motion.

3.27 A particle of mass m moves in one dimension with an applied force $F_x = -F_0 \sinh \alpha x$, where F_0 and α are positive constants. (**a**) Find the frequency of small oscillations about the equilibrium position.
(**b**) Next, find the potential energy function and sketch it. Ans: $\omega_0^2 = F_0\alpha/m$

3.28˙Another model for damping is one for which the damping force is proportional to the amplitude (not velocity) but is 90° out of phase with it. You can model this as $m\ddot{x} + hix + kx = 0$, where i is the usual imaginary unit and h is a damping parameter. It would be wrong, but you can do it. Solve this equation with the usual exponential solutions and show why it doesn't make much sense. [I did not make this up.]

3.29˙In the calculations leading to the equations (3.11) I made some algebraic mistakse, resulting first in the expression (a) and later in expression (b) for the coefficient of $z = x - x_0$.

$$(a) \quad -\frac{2(\alpha + \beta)^4}{\beta x_1^3}, \qquad (b) \quad -\frac{2(\alpha + \beta)^4}{\alpha^2 x_1^3}$$

I quickly realized that (a) was wrong and backtracked to find the source of the error. Why is (a) clearly wrong? After making another error to get (b), I realized that it too must be wrong. Why is (b) not possible as a solution? For this one, look at the behavior of $U(x)$ as the parameters α and β vary. Draw graphs!

3.30 An undamped harmonic oscillator has mass m and spring constant k. It is subject to an additional force $= ct$ for time $t > 0$, and it starts from the origin with zero velocity. (**a**) Find its position as a function of time. (**b**) What happens when $k \to 0$? Ans: $x(t) = (c/\omega_0 k)(\omega_0 t - \sin \omega_0 t)$

3.31˙Do the problem that started at Eq. (3.9), but use either method one or two instead.

3.32 A mass m is moving under the force from a potential energy that is given by

$$U(x) = U_0 \left(\frac{x}{a}\right)^{1000000}$$

Carefully sketch U and analyze the motion qualitatively first! Find the period of oscillation for m if it has a total energy E. How does this period depend on E, and compare it the corresponding dependence in the case of a simple harmonic oscillator.

3.33 A critically damped simple harmonic oscillator starts at the origin at rest, and is subject to an additional force $F_{\text{ext}} = F_0 t/t_0$. Find the subsequent motion.

3.34 A mass m is moving in one dimension, and the potential energy associated with the force on it is

$$U(x) = ax + \frac{b}{x}, \qquad (a > 0 \quad b > 0)$$

Sketch this and find the frequency for small oscillations about the equilibrium point.
Ans: $\omega_0^2 = (2/m)(a^3/b)^{1/2}$

3.35 A piece of wood is floating in the water. Unlike the example of section 3.7 however its shape is not rectangular. It is conical. The base of the cone is a circle of radius R and the height of the cone is h. Set it to float with the base down and determine the frequency for small vertical oscillations of this piece of wood. A more difficult problem is to determine the circumstances $(R,\ h)$ so that it will not tip over. I'm not asking that however. Ans: $\omega_0^2 = (3g\rho_w/\rho h)[(\rho_w - \rho)/\rho_w]^{2/3}$, where ρ_w is the density of water and ρ that of wood. What happened to R?
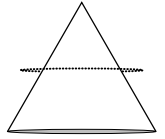
3.36 Two point masses M are fixed on the x-axis at the coordinates $x = \pm a$ and $y = 0$. A third mass m is placed on the y-axis. Find the gravitational force exerted on m by the other two masses in terms of G and the given quantities. If $|y| \ll a$, find the differential equation from $\vec{F} = m\vec{a}$ and then the general solution $y(t)$. If m starts from rest at $y = y_0 \ll a$, what is the future motion?

3.37 Same as the preceding problem except that the mass m is placed on the x-axis at $|x| \ll a$. Find the solution of the equations of motion after it is released from rest at position x_0 (small x).

3.38 The equation (3.68) has two equations. The graph of the second one appears in the text. Graph the first one and compare to two graphs to see how the energy moves from one mass to the other and back.

3.39 In (3.57), replace the spring in the middle with a viscous damper, so that the force it applies is not $\pm k_2(x_1 - x_2)$, but is proportional to the relative velocities of the masses, $\pm b(v_{x1} - v_{x2})$. (**a**) Set up the general differential equations and then solve the special symmetric case as in the text. (**b**) Without actually solving the more general case, explain why the system with this special symmetry will qualitatively behave very differently from the more general case.

3.40 The three springs for the system described by the equations (3.57) are replace by viscous dampers. Each damper produces a force proportional the relative velocity of its two ends, for example $-b_2(v_{x1} - v_{x2})$, or $-b_1 v_{x1}$. Find the general motion of such a system. Set up the equations assuming that the masses are different and the three dampers are different. To solve it, make it symmetric so that the masses are equal and the two outside dampers are equal.

3.41 What is the solution to the equation $m\ddot{x} + kx = F_0 \sin\omega_0 t$? Assume that the force starts as $t = 0$ and that the mass was not moving before then. Do this using Green's function methods. $\omega_0 = \sqrt{k/m}$. Compare this to problem 3.15. Ans: $(F_0/2m\omega_0^2)\left[\sin\omega_0 t - \omega_0 t \cos\omega_0 t\right]$

3.42 Mimic the derivation starting with Eqs. (3.71) and (3.72) and culminating in (3.73), applying it to the equation $m\, d^2x/dt^2 = F_x(t)$. This will result in a single integral that undoes two derivatives. Verify by two differentiations that your answer produces the correct result. Ans: problem 2.2.

3.43 For the potential energy of the problem 3.4, assume now that $U_0 < 0$. Start the particle at the origin with the same small initial velocity and find (approximately) the future motion. For about what time interval will this solution be good?

3.44 A mass M is on the end of a spring having spring constant k and with the spring fixed at the other end. Take the fixed end as the origin, $x = 0$. The *spring itself* has mass m. Assume that M is oscillating as $x = L + A\cos(\omega t)$, but don't assume you know what ω is. L is the unstretched (equilibrium) length of the spring. (It should cancel from the final result.) Each little piece of the spring will be moving at a speed proportional to the distance from the fixed end, and getting to its maximum value at the end where M is attached. At time t, what is the kinetic energy of each dm of the spring, and what then is the total kinetic energy in the spring? Use the usual $k(x - L)^2/2$ for the potential

energy, and add this to the kinetic energy of the spring and of the mass M. What must the value of the frequency ω be so that this is conserved? (That is, independent of time.) Ans: $\omega = \left[k/(M+\frac{1}{3}m)\right]^{1/2}$

3.45 A mass m is moving in one dimension, and the potential energy associated with the force on it is, for $x > 0$,

$$U(x) = ax^2 + b/x^2 \qquad\qquad (a > 0, \quad b > 0)$$

Sketch this function, and find the frequency and period for small oscillations about the position of equilibrium. The period of oscillation turns out to be independent of the energy, so your approximate result is in fact, exact. (Cf. problems 3.67 and 6.51.) Ans: $\sqrt{8a/m}$

3.46 The velocity of a particle varies with position as $v_x = C/x$ for $x > 0$. Find the force acting on the particle as a function of x, *i.e.* $F_x(x)$.

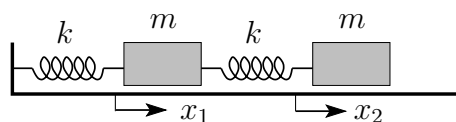**3.47** A particle of mass m moves in one dimension with a potential energy

$$U(x) = -A \cos kx \qquad\qquad (A > 0)$$

Sketch the potential energy and the corresponding force.
Write the differential equation of motion.
The particle starts at rest at the origin and, beginning at $t = 0$ is subjected to an additional force F_0. Assuming this force to be sufficiently small that the particle does not move far from equilibrium, find the approximate solution for $x(t)$.

3.48 Find the normal modes of oscillation for the system drawn here. Write the general solutions for $x_1(t)$ and $x_2(t)$. If the system was at rest for $t < 0$ and you give mass m_2 a kick to start its velocity as v_0, find the future $x_1(t)$ and $x_2(t)$.

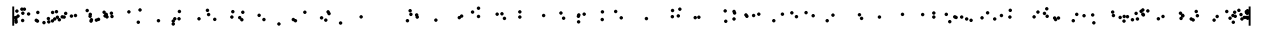**3.49** A particle of mass m moves in one dimension in a potential energy

$$U(x) = -\frac{A}{a^2 + (x-b)^2} \qquad (A > 0)$$

(a) Sketch U. (b) Write the differential equation of motion, and for the case that the particle does not move far from the position of equilibrium, find the approximate solution to the equation of motion: $x(t)$. Ans: $\omega_0^2 = 2A/ma^4$

3.50 A spring is attached at one end to a wall and at the other end to a mass m. The coordinate for m is x. Neglect the mass of the spring. Everything starts at rest and at time $t = 0$ you start to apply a force $F_x = \alpha e^{\beta t}$. (a) Find $x(t)$. Of course, you will ask how your solution behaves for small β. (b) Also, for a real spring that starts as length L, about how much time will it take before this question and answer will make no sense? Take $k = 1\,\text{N/m}$, $m = 1\,\text{kg}$, $\alpha = 1\,\text{N}$, $\beta = 1\,\text{s}$, $L = 1\,\text{m}$ or $L = 100\,\text{m}$. (Compare the latter two.)

3.51 When a mass oscillates as $x(t) = A \sin \omega_0 t$ the velocity is $dx/dt = A\omega_0 \cos \omega_0 t$. If you look at the mass at random times you will find it somewhere between $\pm A$, what is the probability that you will find it near different places between those limits? In a little interval between x and $x + dx$ the probability of seeing it is proportional to the amount of time dt that it is *in* that interval. The probability per interval is then proportional to dt/dx. Evaluate this probability density as a function of x and graph it. Is such behavior plausible? The total probability of being found somewhere is one, so the integral

of this function with respect to x must be one. Use that to evaluate the proportionality constant and then graph the function. Cf. the plot immediately below and see problem 2.33. Ans: $1/\pi\sqrt{A^2 - x^2}$

⠏⠄⠒⠶⠄⠝ ⠽⠁⠍ ⠶⠄ ⠄⠌ ⠤⠶⠁⠗ ⠄⠄⠒⠴ ⠄ ⠄⠄⠄⠄ ⠄⠄⠄⠄ ⠄⠄⠄⠄ ⠄⠄⠄⠄ ⠄⠄⠄⠄⠄ ⠄⠄⠄⠄ ⠄⠄⠄⠄

3.52 As in the preceding problem, what is the probability density for the motion in problem 3.32?

3.53 Take the preceding problem, 3.52, and add a term αx to the potential energy. Now what is the probability density? Assume that α is in some sense small. How big can it be before you have to reconsider your answer?

3.54 A simple pendulum has a spring attached to the (stiff) rod holding the mass. The other end of the spring is attached to a wall. Find the frequency of oscillation of the mass. The spring is unstretched when the pendulum is vertical.
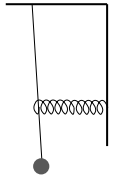
3.55 Two springs having spring constants k_1 and k_2 are attached end-to-end (in series). Derive the effective spring constant of the whole combined spring. (What is the force on the point of attachment?) Ans: $\frac{1}{k_{\text{tot}}} = \frac{1}{k_1} + \frac{1}{k_2}$

3.56 A one dimensional potential energy is $U(x) = -U_0 \frac{x}{a} e^{-x/a}$. Sketch it. What is the position of equilibrium? What is the frequency of small oscillation about equilibrium? Ans: $\omega_0^2 = U_0 e^{-1}/ma^2$ (and where is your *graph*?)

3.57 A particle of mass m moves in a straight line with a potential energy given by $U(x) = \alpha x^2 - \beta x^4$, where α and β are positive constants. (**a**) Determine the range of energies for periodic motion. (**b**) What is the period of small oscillations? Ans: (a) $|x| < \sqrt{\alpha/2\beta}$

3.58 The Morse potential, Eq. (2.37), has a stable equilibrium at $r = r_0$. For small oscillations about that point, find the frequency of oscillation of a mass m. Apply this to the HCl molecule where m is the mass of the H-atom, assuming Cl is big enough that it doesn't move much. Use the parameters stated there to compute the numerical value of this frequency. Assume that the molecule emits radiation at that frequency, what is its wavelength and where is it in the electromagnetic spectrum? See section 6.7 or problem 3.61 to see that the motion of the Cl atom modifies the answer only by changing the mass m to the "reduced mass", $m_H m_{Cl}/(m_H + m_{Cl})$. This is a change of slightly less than 3%.

3.59 Electric circuits involving inductors and capacitors will have currents that can undergo harmonic oscillations, and two such circuits can form coupled oscillators. The circuit drawn has two currents, one for each part of the circuit. The capacitances are C, the self-inductances are L, and the mutual inductance is M. The resulting equations for the two parts of the circuit are
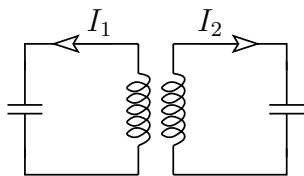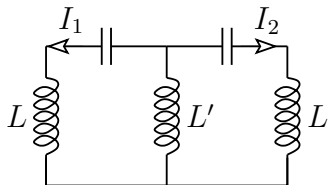

$$L\frac{d^2 I_1}{dt^2} + \frac{1}{C}I_1 + M\frac{d^2 I_2}{dt^2} = 0$$

$$L\frac{d^2 I_2}{dt^2} + \frac{1}{C}I_2 + M\frac{d^2 I_1}{dt^2} = 0$$

Find the modes of oscillation of the currents I_1 and I_2 so that you can write the general form of $I_1(t)$ and $I_2(t)$. ($L > |M|$) and why must this inequality be true? Draw pictures of how the currents move in each mode.

3.60 A circuit involving inductors and capacitors will have currents that can undergo harmonic oscillations and even coupled oscillations. The circuit drawn has two currents, one for each part of the circuit.

The two capacitors are the same, C, and the two inductances on the ends are L. The one in the middle is L'. The resulting equations for the two parts of the circuit are


$$\frac{1}{C}I_1 + L\frac{d^2 I_1}{dt^2} + L'\frac{d^2}{dt^2}(I_1 + I_2) = 0$$

$$\frac{1}{C}I_2 + L\frac{d^2 I_2}{dt^2} + L'\frac{d^2}{dt^2}(I_1 + I_2) = 0$$

Find the modes of oscillation of the currents I_1 and I_2, then write the general form of $I_1(t)$ and $I_2(t)$. Draw pictures of how the currents move in each mode.

3.61 The carbon monoxide molecule, CO, can be modeled as two masses on the ends of a spring. Solve for the oscillations of this molecule, assuming that the motion is along the single long axis between the atoms. Compare your result to what appears in section 6.7, especially Eq. (6.47), derived by different methods.

3.62 Two masses, m and m, are sitting on a table with no friction and initially at rest. They are connected by a spring. At time $t = 0$, start to apply a constant force F_0 to one of the masses, along the line connecting the masses. The motions of the two masses are (maybe)
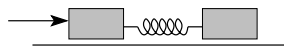
$$x_{1,2}(t) = F_0\, t^2/4m \pm (F_0/4k)\big[1 - \cos\omega_0 t\big] \qquad (\omega_0^2 = 2k/m)$$

Analyze these equations to determine if they are plausible.

3.63¨ Derive the motion, $x_1(t)$ and $x_2(t)$, in the preceding problem.

3.64 Two equal masses m are attached to three springs as in Figure 3.11. The spring constants on the ends are k and the one in the middle is k_2. Starting from equilibrium at $t = 0$, a constant horizontal force F_0 is applied to the mass on the left.
The solution proposed for the motion of the two masses is

$$x_1(t) = -\tfrac{1}{2}(a + b)\cos\omega t - \tfrac{1}{2}(a - b)\cos\omega' t + a$$
$$x_2(t) = -\tfrac{1}{2}(a + b)\cos\omega t + \tfrac{1}{2}(a - b)\cos\omega' t + b$$

$$a = \frac{k + k_2}{k(k + 2k_2)}F_0, \quad b = \frac{k}{k(k + 2k_2)}F_0 \qquad \omega = \sqrt{\frac{k}{m}}, \quad \omega' = \sqrt{\frac{k + 2k_2}{m}}$$

(**a**) Examine this solution for plausibility. (**b**) An alternative solution changes the numerator of b to k_2. Examine this in the same way.

3.65¨ For the system described in the preceding problem,
(**a**) Write all the differential equations of motion for the coordinates x_1 and x_2.
(**b**) Find an inhomogeneous solution to the equations.
(**c**) Find the general solutions for the homogeneous part of the equations.
(**d**) Write the total solution.
(**e**) At time $t = 0$ both masses are at their equilibrium position with zero velocity. Find their future positions.

3.66 For the case that the harmonic oscillator is damped (underdamped), find Green's function for the solution, deriving Eq. (3.74).

3.67 ¨The motion in the potential energy function of problem 3.45 can be solved, resulting in an anharmonic oscillator. The solution is $x(t) = (\alpha \cos^2 \gamma t + \beta \sin^2 \gamma t)^{1/2}$ for appropriate values of the constants. Verify that this is so by substituting it into the conservation of energy equation $K + U = E$ and showing that for appropriate α, β, and γ it is correct. You can solve for these in terms of E, a, and b, but it is simpler if you work backwards and let a, b, and α be the controlling parameters, expressing E, γ and β in terms of them. And of course analyze the results. *Is the frequency correct?* And draw graphs. What does the graph of $x(t)$ look like for small energy and for large energy? And can you anticipate what the graph *should* look like even before trying to use the equation to sketch it? The reason this problem can be done exactly will be clear after doing problem 6.51.
Ans: $\gamma = \sqrt{2a/m}$, $E = a\alpha + b/\alpha$, $\beta = (b/a\alpha)$, $x_{\min} = \sqrt{\alpha}$, $x_{\max} = \sqrt{\beta}$.

3.68 A potential energy is specified to be $(\alpha > 0)$

$$U(x) = \begin{cases} 0 & (x \le 0) \\ \alpha x^2 & (x > 0) \end{cases}$$

A mass m starts at a point $x = -d$ with a speed v_0 to the right. How much time does it take to return to its initial point? Also, sketch a graph of the force function that comes from this potential. I assume that you already sketched $U(x)$ without even being asked.

3.69 In Eq. (3.56) there is a special case that you can do approximately in order to check the analysis. Assume that $\theta_0 = \frac{\pi}{2} - \alpha_0$ and that α_0 is small. Expand the cosines near $\pi/2$ and do the resulting (now pretty easy) integral. This is a result that you can compare to a simple $at^2/2$ type of calculation, where you just let the mass at the end of the pendulum drop a short distance.

3.70 ¨For those who want to try a numerical integral, look at the comparison of the exact integral and the linear approximation in Eqs. (3.54) and (3.56) to check the numerical results stated there. To do this integral numerically, first recognize that the integrand is singular at $\theta = \theta_0$. If you make the substitution $\theta - \theta_0 = x^2$, then the integrand in the x variable is no longer singular, but you must evaluate it as $x \to 0$ in order to find its value *at* zero. Then integrate dx from $x = 0$ to where $\theta = \pi/2$.

3.71 ˙The Atwood machine described in section 1.4 has the masses m_1 and m_2 as indicated, but now the cord has a constant linear mass density μ. Neglect the mass of the pulley. The system starts at time zero with velocity equals zero and the initial coordinate is $y(0) = y_0$. The position y of mass m_1 is claimed to be one or another of these results. Check the dimensions of the proposed answers, then consider various special values of the parameters m_1, m_2, μ, y_0 and demonstrate that all of these proposed answers are impossible. *Do not* solve the problem and try to compare the answer to these proposed. (That comes next.)

$$y(t) = y_0 \cosh \omega t \quad \text{with} \quad \omega^2 = \frac{m_1 - m_2}{m_1 + m_2 + \mu L}$$

$$y(t) = y_0 \cosh \omega t + \frac{m_1 - m_2}{2\mu}(1 - \cosh \omega t) \quad \text{with} \quad \omega^2 = \frac{2\mu g}{m_1 + m_2 + \mu L}$$

3.72 ¨Solve the equation (2.40) for the motion of Atwood's machine. Assume It starts from rest at initial $y = y_0$. Many cases to analyze here: (**a**) $\mu = 0$, (**b**) $m_1 = m_2 = 0$, (**c**) there's an exponential in this solution. Can the acceleration ever be larger than g? The special case (**b**) will suffice for this one.

3.73 (**a**) Write the equation for conservation of energy for the potential function $U(x) = -kx^2/2$, and sketch the graph of U.
(**b**) Separate variables to get dt and write the integral that will allow you to find t in terms of x.

(**c**) Evaluate the integral for the special case that $E = 0$ (keeping the algebra down) and apply the initial condition that $x(0) = x_0$ to solve for $x(t)$. Ans: $x = x_0 e^{\omega t}$

3.74 ˙Two masses, m_1 and m_2, are sitting on a table with no friction. They are connected by a spring, and a force $F_0 \cos \omega t$ is applied to m_1 along the line connecting the masses. (**a**) Find the steady-state solution (the inhomogeneous solution) for the motion. (**b**) Graph the amplitude for the motion of each mass as a function of the applied ω.

3.75 For the simple harmonic oscillator, the period is independent of the energy. For the quartic potential, $U \propto x^4$, how does the period depend on the energy? Also compare these to the solutions of problems 2.19 and 3.32. Is there a pattern?

3.76 Two pendulums are coupled with a light spring. The two lengths of the pendulums are the same, but the masses are different. The spring is attached at a common distance ℓ from the support and the lengths of the pendulums are L. The masses are m_1 and m_2. When the masses are hanging vertically the spring is unstretched. Set up the equations of motion and find the frequencies and modes of oscillation. If you think that you're getting involved in solving a quartic equation, look more closely; it's not that hard. Verify that these solutions are orthogonal in the sense of Eq. (3.69). [This is an easy demonstration to set up and to see that the mass matrix is important in making these modes orthogonal.]
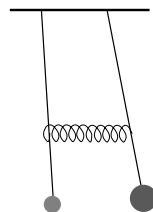
3.77 The manipulations that led from Eq. (3.75) to Eq. (3.77): do the same for the simple harmonic oscillator.

3.78 Do the numerical integral, Eq. (3.77) from 0 to 2. Use the midpoint rule with $N = 1$ and with $N = 2$ (or Simpson—look it up). The more accurate answer for this u when $z = 2$ is 1.311029.

3.79 In the same way that you plot position versus time in figure 3.16, plot the motion for problem 3.32.

3.80 In Eq. (3.70) you see the orthogonality relation for normal modes. Consider the same two-mass, three-spring example used in section 3.9, but take all the springs to be equal and take one mass to be much larger than the other, say $m_2 = 100m_1$ or so. You can go through the whole solution of a quadratic equation, the determinant of Eq. (3.62), but without the symmetry. But don't. Instead, draw pictures of what you expect the modes ought to look like and then apply the orthogonality relation or Eq. (3.69). Your first guess may not be right, but use the fact that they have to be orthogonal to adjust the factors. What then are the two frequencies, at least approximately? Justify your reasoning for the shape of the modes. In other words, solve this problem approximately without solving the differential equations.

3.81 (**a**) Show that in Eqs. (3.75) and (3.76) the relation between the dimensionless velocity and v is $dx/dt = (dy/du)\sqrt{2E/m}$. (**b**) Show that even if the energy is raised so that the amplitude of the oscillation is greater, then the shape of the curve for x versus t remains unchanged, being the same as in Figure 3.16. Only the scales change.

Three Dimensional Motion

.

The world isn't one dimensional. The added complexities that come when you leave the straight line lead to some pretty results, some surprising results, some perplexing results, and some just plain hard results.

Stay with the case of constant mass for now, so that $\vec{F} = m\vec{a}$ for a single particle. Start with the same sort of special cases as have in one dimension, for which the force is a function of time or velocity or position, and then put them together.

When $\vec{F}$ is a function of time alone, there's seems to be not much difference from the one-dimensional case except to turn everything into vectors: $d^2\vec{r}/dt^2 = \vec{F}(t)/m$. Can that be all there is to this case? Pretty nearly so. You have three times as much calculation to do, and the pictures of the results will be harder to draw, but there's no new concept here — at least in rectangular coordinates! In polar coordinates, the differential equation for r will involve the angle $\phi(t)$ and the equation for ϕ will involve r. Not necessarily simple.

4.1 Projectile Motion
A standard problem that every introductory physics text handles is that of figuring out motion in a uniform gravitational field, $\vec{F} = m\vec{g}$. There are a couple of easy integrals and the rest is interpreting the algebra. Of course you neglect air resistance in order to start with the easiest problems first. What if you don't neglect it? How do you describe air resistance mathematically? To do so fully is quite difficult and complicated, because it depends on many factors. Why does a golf ball have dimples? Because if you use a smooth ball you can't hit it nearly as far. The reasons for this are complex, involving the change in air resistance when turbulence is induced and the lift caused by the interaction of the ball's spin with the air — the Magnus force that appears in Eq. (4.51).

Air resistance with any object is strongly dependent on velocity. At low enough speeds it is typically a linear function of v, then at somewhat higher speeds it is more nearly quadratic. At still higher speeds it can even decrease with increasing speed for some ranges of v.

Assume that the resistance is linear in v, not because it is a good approximation for ordinary speeds, but because it is the only assumption allowing you to use straight-forward mathematics. Every other model is more difficult to handle. Qualitatively though, it's still pretty good.

Example
▶ In this model the equation to examine is

$$\vec{F} = m\vec{g} - b\vec{v} = m\vec{a}, \qquad \text{or} \qquad -mg\hat{y} - b\vec{v} = m\frac{d^2\vec{r}}{dt^2} \tag{4.1}$$

where I pick the y-axis to be up. Take the initial conditions to be

$$\vec{r}(0) = 0, \qquad \text{and} \qquad \vec{v}(0) = v_{x0}\,\hat{x} + v_{y0}\,\hat{y} = v_0\cos\theta\,\hat{x} + v_0\sin\theta\,\hat{y}$$

With these conditions the z-coordinate stays at zero.

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} = 0, \qquad m\frac{d^2y}{dt^2} + b\frac{dy}{dt} = -mg \tag{4.2}$$

These are linear constant coefficient differential equations, one inhomogeneous. You can solve them in several different ways, and I choose to use the method of section 0.9.

Both of Eqs. (4.2) have the same homogeneous part, so do that first. Assume an exponential.

$$x = Ae^{\alpha t} \implies mA\alpha^2 e^{\alpha t} + bA\alpha e^{\alpha t} = 0, \qquad \text{so} \qquad \alpha = 0 \qquad \text{or} \qquad \alpha = -b/m$$

The homogeneous solution is then $x(t) = A + Be^{-bt/m}$. The same for y, though with different constants.

The inhomogeneous part of the differential equation is a constant, so try a constant times time for a solution:

$$y = Ct \implies m \cdot 0 + bC = -mg, \qquad \text{or} \qquad C = -mg/b$$

Put these together and the total solution is

$$x(t) = A + Be^{-bt/m}, \qquad\qquad y(t) = A' + B'e^{-bt/m} - mgt/b$$

Apply the initial conditions to evaluate the constants.

$$x(0) = 0 = A + B, \qquad y(0) = 0 = A' + B',$$

$$\text{and} \qquad v_x(0) = v_0 \cos\theta = -\frac{b}{m}B, \qquad v_y(0) = v_0 \sin\theta = -\frac{b}{m}B' - \frac{mg}{b}$$

$$\longrightarrow \qquad \begin{aligned} x(t) &= \frac{m}{b} v_0 \cos\theta \left[1 - e^{-bt/m}\right] \\ y(t) &= \frac{m}{b}\left[v_0 \sin\theta + \frac{mg}{b}\right]\left[1 - e^{-bt/m}\right] - \frac{mgt}{b} \end{aligned} \qquad (4.3)$$

First check that the dimensions are correct. This has to be a reflex, so I'll leave it to you. [*Quickly: Look at the exponent in $e^{-bt/m}$ and determine the units for m/b. Then look at all the other places that m/b appears.*] Of course you should also go back to the very first equation where b shows up and check that the units for m/b match there.

Back in chapter 2 you looked at what linear viscosity did to the motion of a particle, but that was without gravity, and the solution appears in the equations (2.13) through (2.16) in section 2.2(b). The results here, Eqs. (4.3), should agree with those previous equations simply by turning gravity off. Check it out.

The next step is to check that these new results are plausible. If there is no air resistance after all, then you know the answer is a simple $\frac{1}{2}gt^2$ sort of result. Does this give the right answer if $b \to 0$? For $x(t)$ you get $\frac{0}{0}$. For $y(t)$ you get $\frac{1}{0} \cdot \frac{1}{0} \cdot 0 - \frac{1}{0}$, so this requires power series expansions. Start with the $e^{-bt/m}$ term; the exponential is, from Eq. (0.1)

$$e^{-bt/m} = 1 - \frac{bt}{m} + \frac{b^2 t^2}{2m^2} - \frac{b^3 t^3}{6m^3} + \cdots$$

Substitute into $x(t)$ and it is

$$\begin{aligned} x(t) &= \frac{m}{b} v_0 \cos\theta \left[1 - \left(1 - \frac{bt}{m} + \frac{b^2 t^2}{2m^2} - \frac{b^3 t^3}{6m^3} + \cdots\right)\right] \\ &= \frac{m}{b} v_0 \cos\theta \left[\frac{bt}{m} - \cdots\right] \\ &= v_0 \cos\theta\, t + \cdots \end{aligned} \qquad (4.4)$$

All the rest of the terms go to zero as $b \to 0$, and this is the correct result for zero resistance.

Next examine y:

$$y(t) = \frac{m}{b}\left[v_0 \sin\theta + \frac{mg}{b}\right]\left[1 - \left(1 - \frac{bt}{m} + \frac{b^2t^2}{2m^2} - \frac{b^3t^3}{6m^3} + \cdots\right)\right] - \frac{mgt}{b}$$

$$= \frac{m}{b}\left[v_0 \sin\theta + \frac{mg}{b}\right]\left[\frac{bt}{m} - \frac{b^2t^2}{2m^2} + \frac{b^3t^3}{6m^3} + \cdots\right] - \frac{mgt}{b} \qquad (4.5)$$

$$= v_0 \sin\theta\, t - \frac{1}{2}gt^2 + \cdots$$

Again, the rest of the terms vanish as $b \to 0$. These two limiting equations for x and y are what you *must* find or you have to go back and discover your mistake.

Go to the opposite extreme now. What happens to this trajectory after a long time? This is pretty simple: In the equations (4.3), the exponential factor $e^{-bt/m}$ will go to zero after a long enough time, and all that is left after that is

$$x(t) \approx \frac{m}{b}v_0 \cos\theta \qquad \text{and} \qquad y(t) \approx \frac{m}{b}\left[v_0 \sin\theta + \frac{mg}{b}\right] - \frac{mgt}{b} \qquad (4.6)$$

It has gone some distance horizontally and is then dropping vertically at constant velocity. The terminal speed is mg/b. It may of course hit the ground before this "long enough" time has been reached. You can get into this terminal speed zone by firing something from the edge of a cliff, allowing enough time then before hitting the ground. You also get into this domain if you fire the projectile into a large barrel of honey, increasing b sufficiently.

We're not done yet. Just as with the zero viscosity case you eliminate t between the equations (4.3) for x and y to get a non-parametric equation for the trajectory.

$$y = x\left[\tan\theta + \frac{mg}{bv_0 \cos\theta}\right] + \frac{m^2g}{b^2}\ln\left[1 - \frac{bx}{mv_0 \cos\theta}\right] \qquad (4.7)$$

That helps a lot, doesn't it?

There really are some things that to dig out of this equation: As x starts at zero, this is zero. When x increases, the argument of the logarithm drops from one to zero. The log goes to $-\infty$ there, and that point is $x = mv_0 \cos\theta/b$, the same value as computed from Eq. (4.6). What is the behavior for small x? Is it what you expect? Do a power series expansion again, remembering the series for a logarithm, $\ln(1 + x) = x + \cdots$. Then the terms in $1/\cos\theta$ cancel each other and you are left with $y = x\tan\theta$, precisely as needed.

Fig. 4.1
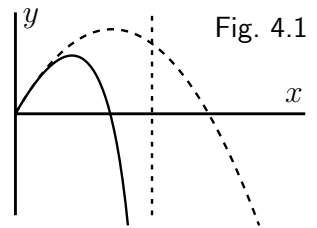
The vertical dashed line in the figure shows a vertical asymptote $x = mv_0 \cos\theta/b$, and the trajectory approaches that vertical motion as time goes to infinity (or until it hits the ground). The curved dashed line is the trajectory the of the same mass with the same initial conditions but without friction. The shape of the motion with friction is no longer the symmetric parabola that you've seen so often in elementary work, and you cannot make some of the simplifying assumptions that you may then have taken for granted. Does it take the same time to come down as to go up? No. Is the peak of the curve halfway between the left and right intercepts with the x-axis? No.

Example

▶ To get a qualitative feel for how equations such as Eq. (4.7) behave, examine it for *small* not zero values of the air resistance. Use a series expansion to see the effect of the air. The first b is in the denominator, so it looks like the right side is approaching infinity as b approaches zero, but carry on. The series for the logarithm from page 1 is all that's needed.

$$y = x\left[\tan\theta + \frac{mg}{bv_0\cos\theta}\right] + \frac{m^2g}{b^2}\ln\left[1 - \frac{bx}{mv_0\cos\theta}\right]$$

$$= x\left[\tan\theta + \frac{mg}{bv_0\cos\theta}\right] + \frac{m^2g}{b^2}\left[-\frac{bx}{mv_0\cos\theta} - \frac{1}{2}\left(\frac{bx}{mv_0\cos\theta}\right)^2 - \frac{1}{3}\left(\frac{bx}{mv_0\cos\theta}\right)^3 - \cdots\right]$$

$$= x\tan\theta - \frac{gx^2}{2v_0^2\cos^2\theta} - \frac{gbx^3}{3mv_0^3\cos^3\theta} - \frac{gb^2x^4}{4m^2v_0^4\cos^4\theta} - \cdots \tag{4.8}$$

and the $1/b$ terms have disappeared. Where does it land? Assume the simplest case, that it is fired on a level surface so that the equation for the ground is $y = 0$. You solve that as a simultaneous equation with (4.8). One root factors instantly, $x = 0$, and the rest is

$$\tan\theta - \frac{gx}{2v_0^2\cos^2\theta} - \frac{gbx^2}{3mv_0^3\cos^3\theta} - \frac{gb^2x^3}{4m^2v_0^4\cos^4\theta} - \cdots = 0 \tag{4.9}$$

You could look up the cubic formula and find the solution, but let's not. You could try examining the effect of the first new term, the one in bx^2, using the more familiar quadratic formula for that. But no, there's a better way. If the air resistance is small, then this equation is *almost* linear, consisting of the first two terms. This calls for the iterative methods described in section 0.11. *If you've skipped it, then read it now.* You will see it again several times.

Apply the iteration method to Eq. (4.9) to see the effect of the frictional factor b.

$$\tan\theta - \frac{gx}{2v_0^2\cos^2\theta} - \frac{gbx^2}{3mv_0^3\cos^3\theta} = 0 \longrightarrow x = \frac{2v_0^2}{g}\cos^2\theta\tan\theta - \frac{2b}{3mv_0\cos\theta}x^2 \tag{4.10}$$

In the lowest order approximation neglect the final term, proportional to b. This gives the zeroth order approximation

$$x_0 = \frac{2v_0^2}{g}\cos\theta\sin\theta = \frac{v_0^2}{g}\sin 2\theta$$

showing the range in a vacuum. This is the case in which the maximum range occurs for a firing angle of $45°$, where $\sin 2\theta = 1$. The correction to this as caused by air resistance is what comes next. Iterate the quadratic equation to get an improved root, x_1, by putting x_0 into the right hand side of Eq. (4.10).

$$x_1 = \frac{2v_0^2}{g}\cos\theta\sin\theta - \frac{2b}{3mv_0\cos\theta}\left[\frac{2v_0^2}{g}\cos\theta\sin\theta\right]^2$$

$$= \frac{v_0^2}{g}\sin 2\theta - \frac{8bv_0^3}{3mg^2}\cos\theta\sin^2\theta$$

$$= \frac{v_0^2}{g}\sin 2\theta - \frac{4bv_0^3}{3mg^2}\sin 2\theta\sin\theta \tag{4.11}$$

You see that the range has decreased from the simple vacuum result. (Surprise!) Look at the coefficient of the second term. It is bv_0/mg times the first term (times $\frac{4}{3}\sin\theta$). This is just the ratio of the force by the air to the force by gravity, and that is just what you expect. Or is it? Did you anticipate that it would come out this way? Maybe next time. Does the maximum range still occur for a firing angle of $45°$? No.

Why does the added factor in the second term have a $\sin\theta$ in it? Is it plausible? You see that when θ is very small, this correction is very small and when θ is up near $90°$ the correction is bigger.

That makes sense because at small angles, it is not in the air for a long time; it stays in the air far longer when fired at larger angles. The correction should then be bigger for larger θ than for smaller, and it is. Compare the size of the correction for $\theta = 1°$ and $\theta = 89°$.

Can you get a still better result by repeating this iteration, getting x_2 by putting x_1 into the right hand side of Eq. (4.10)? No! That would be wrong. Why? And what would you have to do to find the higher order correction correctly? Don't worry; we won't.

Without deriving the result, I will state that the maximum range, found by setting $dx_1/d\theta = 0$ from Eq. (4.11), occurs for an angle slightly less than $45°$:

$$\theta_{\text{max range}} \approx \frac{\pi}{4} - \frac{2bv_0}{mg\sqrt{2}}$$

So this is what you get into when you start to add the slightest bit of reality to these previously elementary problems.

4.2 General Results

If you read Newton's Principia you will not find the equation $\vec{F} = m\vec{a}$. You will not even find the equation $\vec{F} = d\vec{p}/dt$. His presentation of physics was nothing like the modern way to develop the subject, and it is close to unreadable today.* The treatment we are accustomed to was mostly developed by Leonard Euler, and if you look up Euler's collected works you will find the language and the notation remarkable modern. That is because we mostly use Euler's notation.

Starting from the basic equation describing the motion of a particle, there are some general results to derive. Start with one dimension.

$$F_x = m\frac{dv_x}{dt} \quad \longrightarrow \quad F_x v_x = mv_x\frac{dv_x}{dt} \quad \longrightarrow \quad F_x v_x = \frac{m}{2}\frac{dv_x^2}{dt}$$

Reading from right to left, this says that the power, the rate of change of energy is

$$\frac{dW}{dt} = \frac{d}{dt}\frac{1}{2}mv_x^2 = F_x v_x$$

Still another way to manipulate it is to integrate with respect to t.

$$\int_{t_i}^{t_f} F_x v_x \, dt = \int_{t_i}^{t_f} \frac{m}{2}\frac{dv_x^2}{dt}dt = \int_{t=t_i}^{t=t_f} \frac{m}{2}d(v_x^2) = \frac{1}{2}mv_f^2 - \frac{1}{2}mv_i^2 \qquad (4.12)$$

This is the simplest form of the work-energy theorem,

$$W_{\text{total}} = \int_{t_i}^{t_f} F_x v_x \, dt = \int_{x_i}^{x_f} F_x \, dx = \Delta K = K_f - K_i \qquad (4.13)$$

Does this look familiar? If so, that may be because you've studied sections 1.3 and 2.3, but I thought it worth repeating.

In one dimension, $F_x v_x \, dt = F_x \, dx$, and if the force F_x is a function of the coordinate x alone, then the integral depends on the initial and final values of x and not on how fast or how slowly the mass

---

* though there is an amazing book by S. Chandrasekhar that attempts to make it accessible: "Newton's Principia for the common reader".

went from the start to the finish. $F_x(x)$ has an anti-derivative, and the fundamental theorem of calculus says that you evaluate this anti-derivative at the endpoints and subtract. Call this anti-derivative $-U$.

$$F_x(x) = -\frac{dU}{dx}, \qquad \text{then} \qquad \int_{x_i}^{x_f} F_x(x)\,dx = -U(x)\Big|_{x_i}^{x_f} = -U(x_f) + U(x_i) \qquad (4.14)$$

Put this into the preceding equation and rearrange it.

$$-U(x_f) + U(x_i) = \frac{1}{2}mv_f^2 - \frac{1}{2}mv_i^2 \quad \longrightarrow \quad \frac{1}{2}mv_i^2 + U(x_i) = \frac{1}{2}mv_f^2 + U(x_f) \qquad (4.15)$$

This is in the form of a conservation law. Something evaluated at one time equals the same thing evaluated at a later time: Conservation of mechanical energy. This is the reason for choosing the minus sign on U. It produces a nicer result.

Staying with one dimension for a moment, what can go wrong? A simple form of force that you've seen in an introductory course is friction. Common dry friction is velocity dependent, violating the assumptions leading to this conservation law. When an object slides along a surface, dry friction is represented by $F_{fr} = \mu_k F_N$, where F_N is the perpendicular (normal) component of the object's force on the surface and μ_k is the coefficient of kinetic (sliding) friction. Assuming that this coefficient is independent of the speed of sliding, the component of force, F_x, depends on the *direction* of the motion. It may not depend on the magnitude of the velocity, but it depends on $\hat{v}$. At one value of x, if the object slides to the right then the force is to the left; if it slides left, the force is to the right. There is no $F_x(x)$ and so no $U(x)$. A more complete form for the equation of dry friction appears in Eq. (1.8).

Example

▶ For a simple example of this, look at a mass sliding up and back down a slope. It starts uphill with speed v_0. How fast is it moving when it comes back to its starting position (assuming that it does)? For this example there are two cases, corresponding to the direction of motion and correspondingly to the direction of friction.

Fig. 4.2
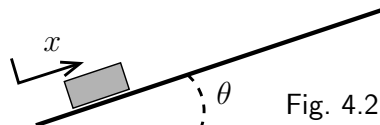
$$\text{(up)} \qquad F_x = -mg\sin\theta - \mu_k mg\cos\theta, \qquad \text{then}$$

$$W = \int_0^{x_{top}} dx\, F_x = \big(-mg\sin\theta - \mu_k mg\cos\theta\big)x_{top} = \Delta K = 0 - \frac{1}{2}mv_0^2$$

That was on the way up. Coming down it is

$$\text{(down)} \qquad F_x = -mg\sin\theta + \mu_k mg\cos\theta, \qquad \text{and}$$

$$W = \int_{x_{top}}^{0} dx\, F_x = -\big(-mg\sin\theta + \mu_k mg\cos\theta\big)x_{top} = \Delta K = \frac{1}{2}mv_f^2 - 0$$

Divide these equations to eliminate x_{top}, leaving a relation between v_f and v_0. Then simplify the result.

$$v_f^2 = v_0^2\,\frac{\sin\theta - \mu_k\cos\theta}{\sin\theta + \mu_k\cos\theta} \qquad (4.16)$$

If the friction is too large or the angle too small, of course it doesn't return. Then too there is the pesky fact that static friction is not the same as kinetic. This will affect whether it sticks at the top or not. In any case, there is no conservation of mechanical energy. There is no U.

Even in this example with friction there is a part of the force (gravity) for which potential energy exists. In the work-energy theorem it is useful to divide the force into two types: Conservative (there is a potential energy) and Nonconservative (there isn't).

$$F_x = F_{x,\,\text{cons}} + F_{x,\,\text{other}}, \qquad \text{then} \qquad F_{x,\,\text{cons}}(x) = -\frac{dU}{dx}$$

$$\text{and} \qquad W_{\text{total}} = \int_{x_i}^{x_f} F_x \, dx = -U(x) \Big|_{x_i}^{x_f} + \int_{x_i}^{x_f} F_{x,\,\text{other}} \, dx$$

Put this into (4.12) and rearrange.

$$\int_{x_i}^{x_f} F_{x,\,\text{other}}(x) \, dx = \left(\tfrac{1}{2} m v_f^2 + U(x_f)\right) - \left(\tfrac{1}{2} m v_i^2 + U(x_i)\right)$$

$$W_{\text{other}} = \Delta E \tag{4.17}$$

and this E is the total mechanical energy, $K + U$. This form of the work-energy theorem is equivalent to the others, but it includes the other forms as special cases. If there is no "other" force then this is conservation of mechanical energy. If you decide not to use the potential energy then you simply make "other" the whole thing.

Potential Energy, 3-d

Eqs. (4.12) and (4.17) applies to one dimension, and in two or three dimensions there are more complexities. Even if the force is a function of position alone, the resulting work integral may *not* be a function of position alone (*i.e.* not a function of its endpoints alone). Eq. (4.12) relates the change in kinetic energy to an integral of the force, and that integral will usually depend on the path that that goes from the initial position to the final position. It is a function, not of the endpoints of the path, but of the whole path. Is the path a straight line? Is it an arc of a circle? In Eq. (4.14), if there is a potential energy function of position the initial value of the energy has some determined value, so the final value does too. The work integral cannot then depend on the way you went from the initial to the final position and in three dimensions the same condition must hold.

There are conditions on the force needed in order that the potential energy exist. When the force is a function of the rectangular coordinates, $\vec{F}(x, y, z)$, the necessary conditions are

$$\frac{\partial F_x}{\partial y} - \frac{\partial F_y}{\partial x} = 0 \tag{4.18}$$

You also need the same equations, but with (y, z) and with (z, x) instead of (x, y), giving then a total of three equations. The partial derivative notation (∂ instead of d) means that the other two coordinates are treated as constants during the differentiation. To see why this equation is needed, look at the case where the potential energy integral goes around a loop, so that the final point is the same as the initial point. See problem 4.31 for a quick derivation.

In three dimensions the differential relation between force and energy simply extends that of Eq. (4.14) to more components.

$$F_x(x, y, z) = -\frac{\partial U}{\partial x}, \qquad F_y(x, y, z) = -\frac{\partial U}{\partial y}, \qquad F_z(x, y, z) = -\frac{\partial U}{\partial z} \tag{4.19}$$

And you see from these lines that Eq. (4.18) is the statement that you can do partial derivatives in either order:

$$\frac{\partial^2 U}{\partial x \partial y} = \frac{\partial^2 U}{\partial y \partial x}$$

All these say is that a particular component of the force is (minus) the derivative of the potential energy with respect to the distance in that direction. It is just as true of r:

$$F_r = -\frac{\partial U}{\partial r}$$

whether spherical or cylindrical. The gravitational force by the Earth on a mass m is radial and is a function of the single coordinate r (at least to a good approximation). In this approximations it has the potential energy:

$$F_r = -\frac{GMm}{r^2} = -\frac{dU}{dr} \quad \longrightarrow \quad U(r) = -\frac{GMm}{r} \tag{4.20}$$

Example
► Two dimensions are enough to see what's happening here. Take

$$U(x,y) = \tfrac{1}{2}k(x^2 + y^2), \qquad \text{then} \qquad F_x = -\frac{\partial U}{\partial x} = -kx \quad \text{and} \quad F_y = -\frac{\partial U}{\partial y} = -ky$$

Assemble the force vector from these components and

$$\vec{F}(x,y) = -kx\,\hat{x} - ky\,\hat{y} = -k\vec{r}$$
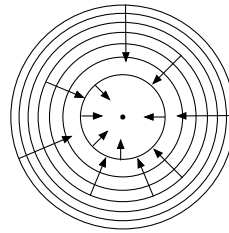
Fig. 4.3

The force is directed radially toward the origin, pointing from higher toward lower potential energies, and the force's magnitude is proportional to the distance to the origin: $F = k(x^2 + y^2)^{1/2}$. This is just a spring force. The drawing shows a set of equipotentials—curves with $U = $ a constant—for equal increments in the value of U. The arrows represent the force vectors at various starting points, with their magnitudes proportional to the distance to the origin.

   Are these equations as easy to derive in three dimensions as in one? If you're familiar enough with vector manipulations they are.

$$\vec{F} = m\vec{a} \quad \longrightarrow \quad \vec{F}\cdot\vec{v} = m\frac{d\vec{v}}{dt}\cdot\vec{v} = \frac{d(mv^2/2)}{dt}$$

$$\longrightarrow \quad \int_{t_i}^{t_f} \vec{F}\cdot\vec{v}\,dt = \int_{t_i}^{t_f} \frac{d(mv^2/2)}{dt}\,dt = \tfrac{1}{2}mv_f^2 - \tfrac{1}{2}mv_i^2 \tag{4.21}$$

Are these manipulations legal? Yes, but until you are comfortable with them you should write them out in x-y-z components to see what's happening, especially the $m\vec{a}\cdot\vec{v}$ equation.

$$\frac{d}{dt}(v_x^2 + v_y^2 + v_z^2) = 2v_x\frac{dv_x}{dt} + 2v_y\frac{dv_y}{dt} + 2v_z\frac{dv_z}{dt} = 2\vec{v}\cdot\vec{a}$$

and you see that the manipulations were valid.

   The first of the integrals in Eq. (4.21) is not necessarily well-defined without more effort. $\int_{t_i}^{t_f} \vec{F}\cdot\vec{v}\,dt$ will almost always depend on the path from the beginning to the end. Simply giving the limits t_i and t_f is not enough. In three dimensions there are many ways to get from one point to

another; in one dimension, there are not so many, but even there you should look back at the analysis at figure 4.2 and leading to Eq. (4.16) to see that the work integral from zero to zero depends on the path.

$$W = \int_0^0 F_x dx = 0 \qquad \text{if the mass doesn't move at all}$$

$$W = \int_0^{x_{\text{top}}} F_x dx + \int_{x_{\text{top}}}^0 F_x dx$$
$$= \left(-mg\sin\theta - \mu_k mg\cos\theta\right)x_{\text{top}} - \left(-mg\sin\theta + \mu_k mg\cos\theta\right)x_{\text{top}}$$
$$= -2\mu_k mg\cos\theta\, x_{\text{top}} \neq 0 \qquad \text{if it goes up and back}$$

In section 2.3 a graph of potential energy $U(x)$ became a useful tool to gain a qualitative (and quantitative) understanding of the motion of a mass. Even when you can't solve the equations explicitly you can see how the mass behaves, bouncing between stopping points. The same sort of pictures help in three dimensions too.

Just as the equation (4.13) shows that power in one dimension is $dW/dt = F_x v_x$, the equation for work in three dimensions gives

$$W = \int \vec{F} \cdot d\vec{\ell} = \int \vec{F} \cdot \vec{v}\, dt \longrightarrow \frac{dW}{dt} = \vec{F} \cdot \vec{v} \tag{4.22}$$

This is true whether the integral depends on the path or not because power is a derivative and you need to go only the short distance $\Delta\vec{r} = \vec{v}\Delta t$ to evaluate it.

$$\Delta W = \int_{\vec{r}}^{\vec{r}+\Delta\vec{r}} \vec{F} \cdot d\vec{\ell} = \vec{F}(\vec{r}) \cdot \Delta\vec{r} \longrightarrow \frac{dW}{dt} = \vec{F} \cdot \frac{d\vec{r}}{dt} = \vec{F} \cdot \vec{v}$$

If the work integral, $\int \vec{F} \cdot \vec{v}\, dt = \int \vec{F} \cdot d\vec{\ell}$, does in fact depend on only the starting and ending points, and not on how you go from one to the other, then potential energy exists, otherwise it does not. The equation (4.14) defines U as a function of the upper limit, x_f; what happens to the integral as the upper limit is changed? Rewrite that equation, taking the zero of the potential to be at x_0.

$$U(x) = -\int_{x_0}^x F_x(x')\, dx' \qquad \text{then} \qquad U(x + \Delta x) = -\int_{x_0}^{x+\Delta x} F_x(x')\, dx'$$

Subtract these, and the parts of the integrals from x_0 to x cancel.

$$U(x + \Delta x) - U(x) = -\int_x^{x+\Delta x} F_x(x')\, dx'$$

For small enough Δx the integral is just $F_x(x)\Delta x$, so you can write this as a differential

$$dU = -F_x(x)\, dx$$

What happens in three dimensions? Much the same thing. If the integral does depend on only the endpoints and not on the path from one to the other,

$$U(\vec{r}) = -\int_{\vec{r}_0}^{\vec{r}} \vec{F}(\vec{r}') \cdot d\vec{\ell}'$$

Here, $d\vec{\ell}$ is the little bit of path along which you are integrating. Make a little change in the endpoint and

$$U(\vec{r} + \Delta\vec{r}) - U(\vec{r}) = -\int_{\vec{r}_0}^{\vec{r}+\Delta\vec{r}} \vec{F}(\vec{r}') \cdot d\vec{\ell}' + \int_{\vec{r}_0}^{\vec{r}} \vec{F}(\vec{r}') \cdot d\vec{\ell}' = -\int_{\vec{r}}^{\vec{r}+\Delta\vec{r}} \vec{F}(\vec{r}') \cdot d\vec{\ell}'$$

and just as in one dimension, when the increment $\Delta\vec{r}$ is small enough the integral is just a product, and this is

$$dU = -\vec{F} \cdot d\vec{r} \tag{4.23}$$

How to picture this? As you move around, the potential energy will change. If you move in a direction perpendicular to the force, the dot product $\vec{F} \cdot d\vec{r}$ is zero and the potential energy does not change in that direction. Move in the direction of $\vec{F}$ and U decreases. Move in against $\vec{F}$ and the potential energy increases. For the gravitational field of the Earth, the force is close to radial (in the pretty good approximation that the Earth is spherical). That implies that the surfaces of constant U are spheres. The equation for U is Eq. (4.20), $U = -GMm/r$, and if this is a constant then r is constant. The *equipotentials* are spheres.
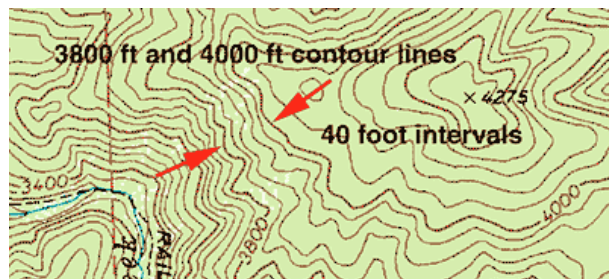

Fig. 4.4

This is a map of the Earth's gravitational potential energy mgh over the surface of the Earth in a mountainous region. The factor mg is omitted from the contours, so that they represent the height above mean sea level as you wander over the ground. The component of force you must contend with on your hike is downhill, and Eq. (4.23) says that if you move ($d\vec{\ell}$) perpendicular to the downhill direction then U doesn't change. You are walking along a contour. Compare the simpler set of contours in the figure 4.3.

Example

▶ If in two dimensions $U = \alpha y - \beta x^2$ where α and β are positive constants, the equipotentials are

$$\alpha y - \beta x^2 = C \tag{4.24}$$

and Each C defines a different parabola as an equipotential. The force corresponding to this potential is everywhere perpendicular to these parabolas. What *is* the force? Equation (4.23) determines the answer. If $d\vec{r}$ is in the direction such that y is a constant, then

$$d\vec{\ell} = \hat{x}dx, \qquad \text{so} \qquad dU = -\vec{F} \cdot d\vec{\ell} = -\vec{F} \cdot \hat{x}dx = -F_x\,dx$$

That just gives the first of the equations (4.19), $F_x = -\partial U/\partial x = 2\beta x$. For the other component, just move in the y-direction to get $F_y = -\alpha$. Draw some equipotentials and draw some force vectors.

4.3 E and B fields

How does a charge behave in an electromagnetic field? The force law is

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) = m\frac{d^2\vec{r}}{dt^2} \tag{4.25}$$

The values of the electric and magnetic fields have to be computed somewhere else, because that's another subject. Suppose that someone else derives them or measures them, the problem now is to determine their effect on charges. If it sounds straight-forward, that's probably because you've looked at the cases of uniform fields in a beginning course on the subject. Even there, you probably saw only the two fields separately. Put them together and you already get some interesting and surprising results. When the fields are functions of position the mathematics escalates dramatically, but there's one important simplification that applies to all magnetic fields, uniform or not: a magnetic field does no work. Recall the expression (4.22) for power: $\vec{F} \cdot \vec{v}$. If the force is magnetic then

$$\text{power} = \vec{F} \cdot \vec{v} = q\vec{v} \times \vec{B} \cdot \vec{v} = 0 \tag{4.26}$$

It's the electric field alone that does work.*

Start with the case of uniform fields $\vec{E} = 0$ and $\vec{B} = B_0\hat{z}$. You've seen the result. The charge will go in a circle. Well no, a helix is more likely. The equation (4.26) says that the kinetic energy and hence the speed of the charge will be constant. Recall from a previous course that the expression for the radial acceleration of a mass moving in a circle is v^2/r, so use that here. The magnetic force is toward the center of the circle, so

$$F = qvB_0 = mv^2/r \quad \longrightarrow \quad v/r = qB_0/m = \omega$$

The period of the circular orbit is the circumference over the speed: $2\pi r/v = 2\pi/\omega$, and this ω is the (constant) angular frequency of the orbit, (radians/second). Where does a helix enter? Don't forget that there is no force along the z-axis, so the component of velocity in that direction is a constant. Combine that with the circular motion around the z-axis and you have a helix.
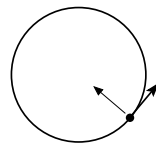
Here I will get the same result using methods that are not so elementary, but that allow for later generalization.

$$m\frac{d^2\vec{r}}{dt^2} = m\frac{d\vec{v}}{dt} = q\vec{v} \times B_0\hat{z} = q[\hat{x}\dot{x} + \hat{y}\dot{y} + \hat{z}\dot{z}] \times B_0\hat{z} = qB_0[\dot{y}\hat{x} - \dot{x}\hat{y}] \tag{4.27}$$

Let $qB_0/m = \omega$, then this breaks into components

$$\frac{d\dot{x}}{dt} = \omega\dot{y}, \qquad \frac{d\dot{y}}{dt} = -\omega\dot{x}, \qquad \frac{d\dot{z}}{dt} = 0 \tag{4.28}$$

The third of these equations stands by itself and says that the z-component of velocity is a constant. As for the other two, they are simultaneous equations for $\dot{x}$ and $\dot{y}$. How do you solve them? Several ways. (1) Just as with algebraic equations: eliminate one of the variables between them and get a single equation in one dependent variable. (2) Follow the methods of section 3.9 and assume an exponential solution. (3) Use matrix methods. (4) Use operator methods. (5) Use complex algebra.

---

* Then how does it appear that in an electric motor the magnetic field is pushing on the currents and doing work? This is a subtle problem involving the Hall effect, and I won't try to resolve it here.

The first method is not the most general, but it is a perfectly good way to get to the solution. More powerful methods are not needed, so they will wait for a page or two. To eliminate one of the coordinates between the equations (4.28), take $\dot{y}$ from the first equation and substitute it into the second equation.

$$\dot{y} = \frac{1}{\omega}\frac{d\dot{x}}{dt} \longrightarrow \frac{d\dot{y}}{dt} = \frac{1}{\omega}\frac{d^2\dot{x}}{dt^2} = -\omega\dot{x} \longrightarrow \frac{d^2\dot{x}}{dt^2} = -\omega^2\dot{x}$$

This is a simple harmonic oscillator equation for $\dot{x}$, so the solution is

$$\dot{x} = A\cos(\omega t + \delta), \quad \text{then the first equation of (4.28) says} \quad \frac{d\dot{x}}{dt} = -A\omega\sin(\omega t + \delta) = \omega\dot{y} \quad (4.29)$$

These are the solutions for $\dot{x}$ and for $\dot{y}$. One integral each and you have the functions $x(t)$ and $y(t)$ themselves (plus z).
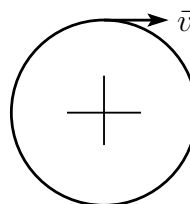
$$x(t) = \frac{1}{\omega}A\sin(\omega t + \delta) + C, \qquad y(t) = \frac{1}{\omega}A\cos(\omega t + \delta) + D, \qquad z(t) = z_0 + v_z t \qquad (4.30)$$

The constant A is arbitrary, so A/ω is too. Call the combination another arbitrary constant, and I will pick the letter R.

$$x(t) = R\sin(\omega t + \delta) + C, \qquad y(t) = R\cos(\omega t + \delta) + D \qquad (4.31)$$

Verify that these really do satisfy the equations (4.28). They describe a circle of radius R and with a center at the x-y coordinates C and D.

$$(x - C)^2 + (y - D)^2 = R^2$$

Is it going in the correct direction? Set $\delta = 0$, $C = 0$, $D = 0$, and take $q > 0$ to check.

$$x(t) = R\sin(\omega t), \qquad y(t) = R\cos(\omega t) \qquad (4.32)$$

Equation (4.32) starts at $(x, y) = (0, R)$, and a little time later the x-coordinate is a little bit positive. This says $\vec{v} \propto \hat{x}$, with $\vec{B} \propto \hat{z}$. The acceleration is along $\vec{v} \times \vec{B}$, and this is proportional to $\hat{x} \times \hat{z} = -\hat{y}$. This is exactly what the picture shows—acceleration toward the center of the circle. (Where did I use assumption that $q > 0$? And what if it isn't?)

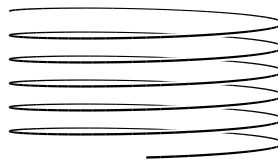With $v_z < 0$ in the picture, you have

Fig. 4.5

The quantity $\omega = qB/m$ is called the cyclotron frequency. The name refers to its application in the first particle accelerator to achieve high energy at low voltages. It was invented in 1932, and despite its antiquity, this accelerator still has some applications, mostly in medicine.

Operator Solution
This is an excuse to show a totally different way to attack problems. It will look like nothing you've seen before and you may well ask if it is legitimate. It is. The equation of motion for a charge in a magnetic field is

$$m\frac{d\vec{v}}{dt} = q\vec{v} \times \vec{B}, \qquad \text{or} \qquad \frac{d\vec{v}}{dt} = -\frac{q}{m}\vec{B} \times \vec{v}$$

This has the same form as the simple equation $dx/dt = \alpha x$, and the solution to that is an exponential. Starting from the initial position x_0 the result is

$$\frac{dx}{dt} = \alpha x \quad \longrightarrow \quad x = e^{\alpha t} x_0$$

Then why can't I solve the previous line as

$$\frac{d\vec{v}}{dt} = -\frac{q}{m}\vec{B} \times \vec{v} \quad \longrightarrow \quad \vec{v} = e^{-\frac{qt}{m}\vec{B}\times}\vec{v}_0 \tag{4.33}$$

Now if I can just figure out what the exponential of "$\vec{B}\times$" is.

I will define it by its power series representation. To keep the algebraic factors from accumulating, let $\vec{\omega} = -q\vec{B}/m$, then

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \qquad \text{with } x = t\,\vec{\omega}\times \qquad \text{gives}$$

$$\vec{v}(t) = e^{t\,\vec{\omega}\times}\,\vec{v}_0 = \vec{v}_0 + t\vec{\omega}\times\vec{v}_0 + \frac{t^2}{2!}\vec{\omega}\times(\vec{\omega}\times\vec{v}_0) + \frac{t^3}{3!}\vec{\omega}\times(\vec{\omega}\times(\vec{\omega}\times\vec{v}_0)) + \cdots \tag{4.34}$$

There is still some manipulation to put this series into a manageable form (less than you may think though). Each term in the series is a well-defined product of vectors, so I can at least ask if it satisfies the differential equation. You do that the usual way: plug in. Differentiate the series (4.34) term by term.

$$\frac{d}{dt}\vec{v}(t) = \vec{\omega}\times\vec{v}_0 + \frac{2t}{2!}\vec{\omega}\times(\vec{\omega}\times\vec{v}_0) + \frac{3t^2}{3!}\vec{\omega}\times(\vec{\omega}\times(\vec{\omega}\times\vec{v}_0)) + \cdots$$

and this is $\vec{\omega}\times$(4.34) itself. It works!

Now for some vector manipulation. The initial vector $\vec{v}_0$ can be in any direction, but in the series for $\vec{v}$ all the terms after the first are perpendicular to the direction of $\vec{\omega}$ (or $\vec{B}$). Call $\vec{B}$ the $\hat{z}$-direction, then v_z stays constant and only the other components change. Just set $v_z = 0$ for now, and add it back later.

$$\vec{\omega}\times\vec{v}_0 \qquad \text{rotates } \vec{v}_0 \text{ by } 90° \text{ about } \vec{\omega}$$

and each successive factor of $\vec{\omega}\times$ effects a rotation by another $90°$ about the $+\vec{\omega}$ direction. Looking down along the z-axis ($\vec{B}$ (oût) and $\vec{\omega} = -q\vec{B}/m$ (în)), the first few are (and yes, the directions are right; look again)
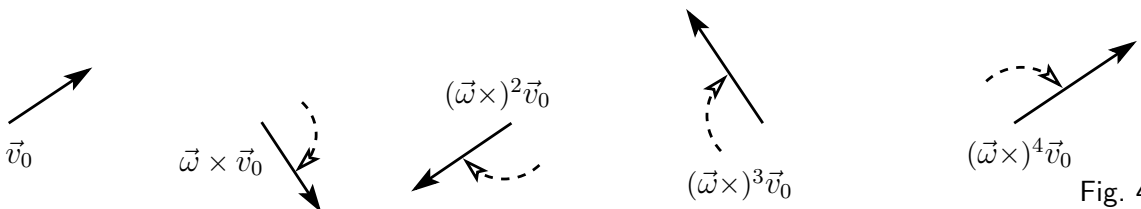

Fig. 4.6

Now the series of terms in Eq. (4.34) are easy to sum. The 1$^{\text{st}}$, 3$^{\text{rd}}$, 5$^{\text{th}}$, *etc.* are in ($\pm$) one direction and the 2$^{\text{nd}}$, 4$^{\text{th}}$, *etc.* are in ($\pm$) the perpendicular direction. In drawing the picture the lengths of the vectors change because of the factor ω. If you draw the picture using the unit vector $\hat{\omega}\times$ as the factor, then all the lengths would be the same. The cross product by a unit vector $\hat{z}$ applied to the $\hat{x}$-$\hat{y}$ plane is purely a $90°$ rotation about the z-axis.

Write the initial velocity as two terms. One along the z-axis and the other perpendicular to it: $\vec{v}_0 = \hat{z}v_{0z} + \vec{v}_1$, and only the $\vec{v}_1$ is rotated. Equation (4.34) is

$$
\begin{aligned}
\vec{v}(t) &= \hat{z}v_{0z} + \vec{v}_1 + \omega t \hat{\omega}\times\vec{v}_1 + \frac{\omega^2 t^2}{2!}\left(-\vec{v}_1\right) + \frac{\omega^3 t^3}{3!}\left(-\hat{\omega}\times\vec{v}_1\right) + \frac{\omega^4 t^4}{4!}\left(\vec{v}_1\right) + \ldots \\
&= \hat{z}v_{0z} + \vec{v}_1\left[1 - \frac{\omega^2 t^2}{2!} + \frac{\omega^4 t^4}{4!} - \cdots\right] + \hat{\omega}\times\vec{v}_1\left[\omega t - \frac{\omega^3 t^3}{3!} + \frac{\omega^5 t^5}{5!} - \cdots\right] \\
&= \hat{z}v_{0z} + \vec{v}_1\cos\omega t + \hat{\omega}\times\vec{v}_1\sin\omega t \tag{4.35}
\end{aligned}
$$

The position as a function of time is now

$$
\vec{r}_0 + \int_0^t dt'\,\vec{v}(t') = \vec{r}_0 + \hat{z}v_{0z}t + \frac{1}{\omega}\hat{\omega}\times\vec{v}_1 + \frac{1}{\omega}\left(\vec{v}_1\sin\omega t - \hat{\omega}\times\vec{v}_1\cos\omega t\right) \tag{4.36}
$$

Does this agree with Eq. (4.31)? It represents motion along a helix: a constant velocity in one direction and uniform circular motion about that direction as an axis. See problem 4.21.

Both E and B

With both fields present (but still uniform) the equations are

$$
\vec{F} = q(\vec{E} + \vec{v}\times\vec{B}) = m\frac{d\vec{v}}{dt} \tag{4.37}
$$

This is a linear, inhomogeneous differential equation for $\vec{v}$. As such, divide the problem in two, finding the general solution to the homogeneous part and finding any one solution to the inhomogeneous part. The homogeneous part is Eq. (4.31), so that's done. For a solution to the inhomogeneous part, notice that $\vec{E}$ and $\vec{B}$ are constants, so a constant $\vec{v}$ is a plausible guess. It isn't right because $\vec{v}\times\vec{B}$ is perpendicular to $\vec{B}$, and that is not necessarily in the direction of $\vec{E}$. A function that is linear in t works.

Instead of pursuing a general solution to these equations, take a special case and solve it by the simplest possible methods. After that go for the generalities. Take $\vec{B}$ along the $\hat{z}$-direction and $\vec{E}$ along the $\hat{y}$-direction, both uniform. ($q > 0$.) Start the charge at the origin with zero velocity. The first thing to happen is that the charge accelerates along $\hat{y}$ because of the electric field, and the magnetic field does nothing (initially) because the velocity is zero. In a little time the velocity has a $\hat{y}$-component and now the magnetic field will act. The magnetic force $q\vec{v}\times\vec{B}$ is in the $\hat{y}\times\hat{z} = \hat{x}$-direction, so it pushes the charge along the x-axis, perpendicular to both $\vec{E}$ and $\vec{B}$.
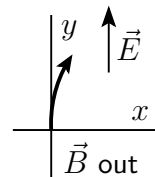
$$
m\frac{d\vec{v}}{dt} = q(E\hat{y} + \vec{v}\times B\hat{z}) \quad\longrightarrow\quad
\begin{aligned}
m\dot{v}_x &= qBv_y \\
m\dot{v}_y &= qE - qBv_x \\
m\dot{v}_z &= 0
\end{aligned}
\tag{4.38}
$$

The third equation says that v_z started at zero and so it stays at zero. Let $\omega = qB/m$, then without the E-term these are the equations (4.28) with solutions $v_x(t)$ and $v_y(t)$ in (4.29), so you know the

solutions to the homogeneous part of these equations. Now try guessing a particular solution for the inhomogeneous part, and start with the simplest guesses. Try a constant, and if that doesn't work, add a term proportional to t. If v_y is a constant, the first implies that $v_x \propto t$, then the second equation can't be right. If v_x is a constant, the first implies that v_y is zero, and the second does work for the right value of the constant v_x.

$$v_{x,\text{ inh}} = \frac{E}{B}, \quad v_{y,\text{ inh}} = 0, \quad v_{x,\text{ hom}} = A\cos(\omega t + \delta), \quad v_{y,\text{ hom}} = -A\sin(\omega t + \delta)$$

The general solution is the sum of these.

$$v_x(t) = \frac{E}{B} + A\cos(\omega t + \delta), \qquad v_y(t) = -A\sin(\omega t + \delta) \tag{4.39}$$

The velocity at time zero is zero, so this determines $\delta = 0$ and $A = -E/B$. Now integrate these, starting at the origin, $x(0) = 0 = y(0)$, and getting

$$x(t) = \frac{E}{B}t - \frac{E}{B\omega}\sin(\omega t), \qquad y(t) = \frac{E}{B\omega}\big[1 - \cos(\omega t)\big] \tag{4.40}$$
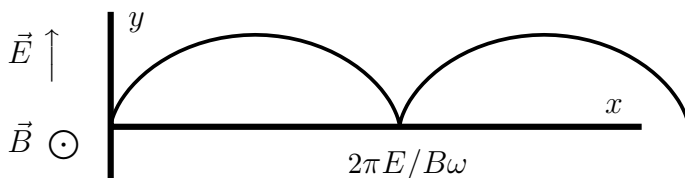

Fig. 4.7

This curve is called a cycloid.* If a pebble is caught in the tread of a car's tire then as the car moves along the road, the pebble will trace this cycloidal curve.

What is the initial behavior of this motion? Expand the solution for small time.

$$x(t) \approx \frac{E}{B}t - \frac{E}{B\omega}\big[\omega t - \omega^3 t^3/6\big] = \frac{E\omega^2}{6B}t^3 \quad \text{and} \quad y(t) \approx \frac{E}{B\omega}\big[1 - 1 + \omega^2 t^2/2\big] = \frac{E\omega}{2B}t^2 = \frac{qE}{2m}t^2$$

For these small times, the t^2 term is much bigger than t^3, so the motion starts along the y-direction as expected. The shape of the curve near the origin behaves as $y \propto x^{2/3}$, consistent with the graph just before Eq. (4.38).

The average velocity comes from Eq. (4.40) and is $\hat{x}E/B$, perpendicular to both fields, and you can write it as the product $\vec{E} \times \vec{B}/B^2$. Is it possible for this result to be bigger than the speed of light? Of course, just make the magnetic field small enough, but then you must use the relativistic version of these equations. The result when with these more complicated equations and small enough B is that there is no oscillation, and the mass will continue to accelerate, approaching the speed of light as time goes on. No more cycloid. The condition for oscillatory motion is simply $E < cB$, though well below this point the equations will no longer be exactly those of a cycloid.

In solving the equations for crossed electric and magnetic fields, all the equations used components. I didn't try to show manipulations using the whole vector formalism with $\vec{E}$'s and $\vec{B}$'s. Can you do it that way? Yes, but it is a lot of effort without enough benefit.

Other Initial Conditions

The equations (4.40) assume that the charge started from rest. A more general initial condition is easy enough to implement simply by integrating the equations (4.39).

$$x(t) = \alpha + \frac{E}{B}t + \beta\sin(\omega t + \delta), \qquad y(t) = \gamma - \beta\cos(\omega t + \delta) \tag{4.41}$$

---

* see the Famous Curves Index mentioned at the end of section 0.3.

Depending on the initial conditions or equivalently, on your choices of the four arbitrary constants in these equations, you will get other shapes for the curves. They look like

Fig. 4.8

The curves are called trochoids instead of cycloids, and these two examples simply have respectively positive and negative initial components of velocity in the x-direction.

Non-uniform B-fields

You may think that a non-uniform magnetic field will be a little more difficult than a uniform one. You would be wrong; even without $\vec{E}$ in the mix it's a *lot* more difficult. The simplest such B-field has space dependence linear in x, y, and z, and it is something like

$$\vec{B} = B_0\hat{z} + \alpha\left[z\hat{z} - \tfrac{1}{2}x\hat{x} - \tfrac{1}{2}y\hat{y}\right]$$

The factors are necessary in order to satisfy Maxwell's equations. The analog of the equations (4.27) and (4.28) involve products of functions, such as $x\,dy/dt$ and $y\,dz/dt$. This makes the differential equations non-linear and far more difficult to solve. In fact they *cannot* be solved exactly in terms of standard functions. They require either finding purely numerical solutions or using far more sophisticated approximation methods.

4.4 Magnetic Mirrors

There is one case however on which I want to spend some time. That's not because it is any easier, but because the result is both interesting and important, so that it is worth the time to look at it. The Earth's magnetic field is not uniform, being larger near the magnetic poles and weaker in between. The solar wind consists of charged particles, electrons and ions, sent at high speed from the sun, and in the absence of the Earth's magnetic field its abrasive effect would gradually strip the Earth of its atmosphere. This may be what has happened on Mars because that planet's magnetic field is far too weak to deflect these particles.

When charges reach the Earth, many are deflected, but some are trapped in the Earth's field and form belts of charged particles that bounce around in *magnetic mirrors*. These are the Van Allen belts. In a uniform magnetic field, a charge moves according to Eq. (4.30) — a helix. The Earth's field is not uniform, but to a good approximation an electron will spiral around a B-field line, moving in a sort of distorted helix, and the axis of this helical form bends around to follow the field. Furthermore, as the field gets stronger the radius of the electron's orbit will get smaller and its motion along the field line will slow. Eventually the electron's orbit will stop and reverse, sending it back toward the other magnetic pole. The mirror effect is what traps the charges in the Van Allen belts.

The equations describing the Earth's magnetic field are sufficiently complicated that you don't want to start with them. Instead, take a simplified model of the field that requires less geometry but that still captures the essence of the problem.

$+z$ $\vec{B} = B_0\hat{z} + B_0\left[z^2\hat{z} - z(x\hat{x} + y\hat{y})\right]/\ell^2$ (4.42)

B-field lines

$z = 0$
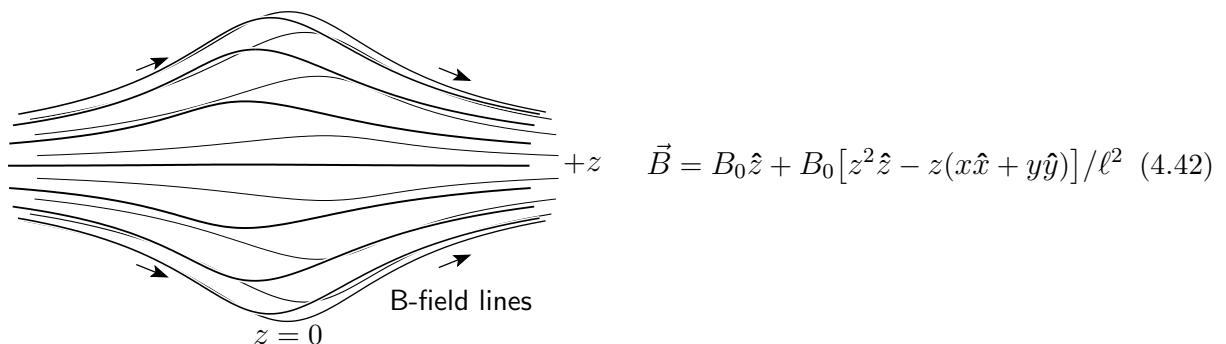
The z-axis is left and right, and the x-axis is up. Look at how the $\vec{B}$-field varies along the z-axis. When $x = y = 0$ this equation is $\vec{B} = B_0(1 + z^2/\ell^2)\hat{z}$, so that the field gets stronger near the two ends. Off the axis, and in moving toward $\pm z$ this equation picks up a component of $\vec{B}$ that points *toward* the z-axis on the right and away from the axis on the left. That's the minus sign in front of the last term. This then resembles the Earth's field, which is largest near the North and South magnetic poles, it differs however in that it doesn't have to curve around the Earth. This mathematical model simply straightens out the field lines while still maintaining the essential structure of the Earth's field. The parameter ℓ is the scale over which the field changes a lot. On the axis, going from $z = 0$ to $z = \pm\ell$ the field strength doubles and the field lines come closer to each other. Without deriving it I will simply state that the equation for the magnetic field lines is $r_\perp = \sqrt{x^2 + y^2} = \alpha/\sqrt{\ell^2 + z^2}$, where α is a parameter that specifies which field line you are referring to, and the line lies in a plane with the z-axis, which is left-to-right. (See problem 4.36 for the derivation. With the right trick it's really easy.) The picture shows many field lines, with the magnetic field strength largest on the left and right.

The equations describing the motion of an electron in a magnetic field are

$$\frac{d\vec{r}}{dt} = \vec{v}, \qquad \frac{d\vec{v}}{dt} = \frac{q}{m}\vec{v} \times \vec{B}$$

With initial data $\vec{r}(0) = \vec{r}_0$ and $\vec{v}(0) = \vec{v}_0$ you can, using a lot of arithmetic, march the electron a little bit at a time through its motion. You replace the derivative $d\vec{r}/dt$ with a difference quotient, $\Delta\vec{r}/\Delta t$. Fix the small but non-zero time interval: Δt. The step from time zero to time Δt is approximately

$$\vec{v}(\Delta t) = \vec{v}(0) + (q/m)\vec{v}(0) \times \vec{B}(\vec{r}(0))\Delta t, \qquad \vec{r}(\Delta t) = \vec{r}(0) + \vec{v}(\Delta t)\Delta t \qquad (4.43)$$

After the first step comes the second step, from Δt to $2\Delta t$. It's the same as the first step except that the initial conditions are now the values of $\vec{r}(\Delta t)$ and $\vec{v}(\Delta t)$ that you just computed. And keep doing this one Δt at a time. (Computers are good at blindly repetitive arithmetic.)* Repeat this for $n = 0$ on up to $n =$ the end of your patience,

$$\vec{v}((n+1)\Delta t) = \vec{v}(n\Delta t) + (q/m)\vec{v}(n\Delta t) \times \vec{B}(\vec{r}(n\Delta t))\Delta t$$
$$\vec{r}((n+1)\Delta t) = \vec{r}(n\Delta t) + \vec{v}((n+1)\Delta t)\Delta t$$

Is this result accurate enough? If not, use a smaller Δt. If *still* not, then you should learn one of the more sophisticated methods to do this sort of computation. Runge-Kutta or Runge-Kutta-Fehlberg are popular techniques. Look up chapter 11 of "Mathematical Tools for Physics" for an introduction to the subject. The primitive form that I wrote above is rarely used in practice because it is not very accurate, and the better methods are not much harder to implement.

In making this plot, I used a fourth-order Runge-Kutta method to step through the motion of the charge as it traverses the field.
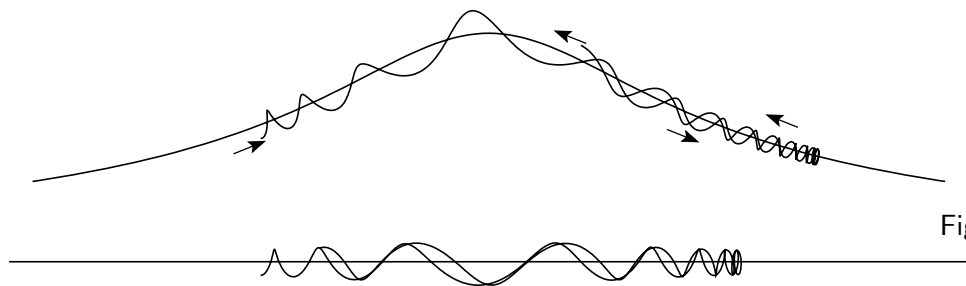

Fig. 4.9

---

* Why isn't the last term in Eq. (4.43) $\vec{v}(0)\Delta t$? Both forms are correct, but the one I gave is slightly more stable.

This picture shows two of the magnetic field lines as in the preceding drawing, one of them along the axis. For both field lines the charge starts on the left side with initial velocity at some angle to the field, and it spirals around the direction of $\vec{B}$ as represented by the field line. As the field becomes weaker toward the midpoint the orbital radius increases, the transverse component of the velocity decreases, and the longitudinal component of velocity increases. Toward the right side, the longitudinal component of velocity decreases to zero and the charge is reflected back along its path. The center of the orbit follows the field line to good accuracy.

How is it possible to understand this motion without recourse to all this numerical computation? There are two key facts that provide insight, allowing you to understand most of these results without detailed calculations. Deriving them gets into the subject of adiabatic invariants, which is well beyond anything in this book, but I will simply state the results and use them: As the charge moves through the field it circles around a field line, and the center of the circle follows that field line. Also, the product of the local magnetic field and the area of the circular orbit stays constant (the adiabatic invariant). Both of these statements are approximate but useful. If you accept them, they explain the results of the preceding paragraph. One assumption needed to make this valid is that the movement along the field is slow compared to the rotation around it. $v_\parallel \ll v_\perp$. The second assumption is that the field does not change much over the diameter of the circle.
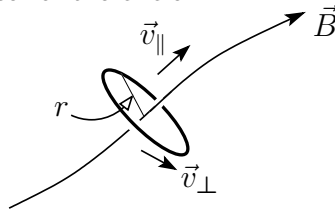

Fig. 4.10

For a little circle around the field line, take B as the value at the center of the circle. The charge has speed v with components parallel to the axis and perpendicular to the axis: call them $v_\parallel$ and $v_\perp$. The speed is constant $v^2 = v_\parallel^2 + v_\perp^2$. Why constant? Magnetic forces do no work, Eq. (4.26), so $mv^2/2$ doesn't change. For this circle, the radial component of $\vec{F} = m\vec{a}$ is

$$m\frac{v_\perp^2}{r} = qv_\perp B \longrightarrow \frac{v_\perp}{r} = \omega = \frac{qB}{m}$$

I claimed that $\pi r^2 B$ is an adiabatic invariant, which means that it's nearly constant. Combine that equation with this equation for ω.

$$\pi r^2 B = \Phi \qquad \text{and} \qquad \frac{v_\perp^2}{r^2} = \frac{q^2 B^2}{m^2} \qquad \text{imply} \qquad \frac{v_\perp^2}{\Phi/\pi B} = \frac{q^2 B^2}{m^2}$$

Rearrange this to get

$$v_\perp^2 = \frac{q^2 \Phi}{\pi m^2} B \tag{4.44}$$

The factors are less important than the relation between $v_\perp$ and B. It says that as the magnetic field gets stronger, the component $v_\perp$ gets bigger. *But*, the speed stays constant, so that says that the other component, $v_\parallel$, must get smaller. When B gets big enough, that parallel component goes to zero and the motion along the field line stops. *That's* the magnetic mirror.

You can get a surprising result out of this by writing the equation for the kinetic energy of the charge. (A constant, remember.)

$$K = \frac{1}{2}mv^2 = \frac{1}{2}m\left(v_\parallel^2 + v_\perp^2\right) = \frac{1}{2}m\left(v_\parallel^2 + \frac{q^2\Phi}{\pi m^2}B\right) \tag{4.45}$$

Look at the motion along the axis, where the field has only the z-component. From Eq. (4.42),

$$B = B_0 + B_0 z^2/\ell^2 \quad \longrightarrow \quad K = \frac{1}{2}m\left(v_\parallel^2 + \frac{q^2\Phi}{\pi m^2}(B_0 + B_0 z^2/\ell^2)\right)$$

Compare this to the energy equation for a simple oscillator, $E = mv^2/2 + kx^2/2$.

$$\frac{1}{2}mv_\parallel^2 + \frac{1}{2}\frac{q^2\Phi}{\pi m}B_0\frac{z^2}{\ell^2} = \text{a constant} \tag{4.46}$$

That is a harmonic oscillator. This result means that not only does the motion along the field line stop, it reverses and keeps bouncing back and forth in simple harmonic motion until it hits something.

The Earth's magnetic field is, to a decent approximation, that of a magnetic dipole. Do the same energy analysis that led to Eq. (4.46), and you will find that the motion of the electron matches that of a pendulum, the subject of sections 3.4 and 4.6.

To get an idea of the frequency of oscillation of these electrons in the Van Allen Belt,

$$\omega_0^2 = \frac{k}{m} = \frac{1}{m}\cdot\frac{q^2\Phi B_0}{\pi m\ell^2}$$

The factor $\Phi = \pi r^2 B = \pi r_0^2 B_0$, where r_0 is radius of the electron's circular motion at the middle of the oscillation; $B = B_0$ there. The electron charge is $q = -e$.

$$\omega_0^2 = \frac{e^2 r_0^2 B_0^2}{m^2\ell^2} = \left(\frac{eB_0}{m}\frac{r_0}{\ell}\right)^2$$

The quantities e and m are known. B_0 for Earth is about $1/2$ Gauss. ℓ is some plausible part of the Earth's radius, perhaps 4000 km. Finally, r_0 is the least well-known, and will vary greatly from one electron to the next. Make a guess of one millimeter and you at least have a place to start:

$$\omega_0 = \frac{eB_0}{m}\frac{r_0}{\ell} = \frac{1.6 \times 10^{-19} \cdot 0.5 \times 10^{-4}}{9.1 \times 10^{-31}} \cdot \frac{1.0 \times 10^{-3}}{4 \times 10^6} = 0.002\,\text{s}^{-1}$$

so this period, $2\pi/\omega_0$, is about an hour. If you now change r_0 from $1\,\text{mm}$ to $1\,\text{m}$, this period changes from about an hour to about 3 seconds.

4.5 Approximate Solutions
Sometimes "... a sleazy approximation that provides good physical insight into what's going on in some system is far more useful than an unintelligible exact result." Here I present a particular method by showing some examples.

> If you haven't read section 0.11 and solved prob-
> lem 0.60(**a**), then now is a good time for it. The
> rest of this section will make much more sense
> after that. Eq. (4.11) also used this method.

Example
▶ To illustrate a particular method, iteration, look again at Eqs. (4.1) and (4.2), but this time see what happens if by assuming from the beginning that the air resistance is small, or equivalently that the mass is large. To a first approximation then, Eq. (4.1) is

$$m\vec{g} = -mg\hat{y} = m\frac{d^2\vec{r}_0}{dt^2}$$

denoting the approximate solution by the subscript zero. With the same initial conditions as before ($\vec{r}(0) = 0$ and $\dot{\vec{r}}(0) = \vec{v}_0$), this is

$$\frac{d^2\vec{r}_0}{dt^2} = \vec{g} \quad\longrightarrow\quad \ddot{r}_0 = \vec{A} + \vec{B}t + \tfrac{1}{2}\vec{g}t^2 \quad\longrightarrow\quad \vec{r}_0(t) = \vec{v}_0 t + \tfrac{1}{2}\vec{g}t^2 \tag{4.47}$$

So far this is all old material. Now use this approximate result to construct an improved result. If the $b\vec{v}$ term is small compared to the other terms in the equation (4.2) for $\vec{r}$, then write

$$\frac{d^2\vec{r}}{dt^2} - \vec{g} = -\frac{b}{m}\frac{d\vec{r}}{dt} \quad\longrightarrow\quad \frac{d^2\vec{r}_1}{dt^2} - \vec{g} = -\frac{b}{m}\frac{d\vec{r}_0}{dt} = -\frac{b}{m}(\vec{v}_0 + \vec{g}t) \tag{4.48}$$

This uses the approximate solution $\vec{r}_0$ in the small term $-b\vec{v}/m$, and because this term is small anyway, using a slightly wrong value for $\vec{r}(t)$ will introduce only a small error. The function $\vec{r}_1(t)$ is the improved approximation for the solution. This can be integrated just as easily as the original equation that didn't include air resistance:

$$\vec{r}_1 = \vec{A} + \vec{B}t + \tfrac{1}{2}\vec{g}t^2 - \frac{b}{m}\left(\tfrac{1}{2}\vec{v}_0 t^2 + \tfrac{1}{6}\vec{g}t^3\right) \tag{4.49}$$

Again, use the initial conditions $\dot{\vec{r}}_1(0) = \vec{v}_0$ and $\vec{r}_1(0) = 0$, to get

$$\vec{r}_1(t) = \vec{v}_0 t + \tfrac{1}{2}\vec{g}t^2 - \frac{b}{m}\left(\tfrac{1}{2}\vec{v}_0 t^2 + \tfrac{1}{6}\vec{g}t^3\right) \tag{4.50}$$

   Just how good is this approximation, and does it really give information about the exact result? There's no doubt that it's simpler, but is simpler good enough? Fortunately, in this case you have an exact solution, so this example is a laboratory in which to test the validity of the iterative solution. Go back to the equations (4.3) – (4.5). Those equations include the exact solution and then they expand it in a series, keeping only the lowest order term — the equivalent of $\vec{r}_0(t)$ in the language of this section. All that you need to do is to keep the next order terms in those equations, one more power of $\frac{b}{m}$, and you will see that they match Eq. (4.50) perfectly. Notice also that deriving Eq. (4.50) is far quicker and simpler than all the labor in section 4.1. If you *know* that you are aiming for an approximate solution to a problem, you are generally better off making the approximations at the beginning, not at the end of the calculation.

   In the last equation, how do the two new terms compare? Factor out $\tfrac{1}{2}t$ from each and you have $\frac{b}{m}\frac{t}{2}\left(\vec{v}_0 t + \tfrac{1}{3}\vec{g}t^2\right)$. The first term in the parentheses is the distance the object would have gone in time t and the second term is $\tfrac{2}{3}$ the distance it would have dropped. If you're dealing with an object that is moving rather fast, then the second term can be much smaller than the first and the expression for $\vec{r}_1(t)$ would still be nearly a parabola. *But notice:* the tip of this parabola is not at the top. The parabola is tilted, as if there is an effective gravitational field given by $\vec{g}' = \vec{g} - \frac{b}{m}\vec{v}_0$. Compare this to the drawing of the exact solution in Figure 4.1.

Example

▶ What about a problem that can't be solved exactly? (That is, almost all other problems.) Air resistance is more accurately described as quadratic in v, and in chapter two, where the motion was along a straight line, calculating this was left to a couple of homework problems: 2.29 – 2.31. The results there are complicated but manageable. If you try to do the same thing in three dimensions, you get an equation that cannot be solved at all without resorting either to numerical methods or as here, to an iterative technique.

   The differential equation for the motion in this case, written as vectors, is

$$m\frac{d^2\vec{r}}{dt^2} = m\vec{g} - bv^2\hat{v} = m\vec{g} - bv\vec{v}$$

This "b" is of course different from the previous one, and the beginning of the problem is the same:

$$\frac{d^2\vec{r}_0}{dt^2} = m\vec{g} \quad \longrightarrow \quad \vec{r}_0 = \vec{A} + \vec{B}t + \tfrac{1}{2}\vec{g}t^2 \quad \longrightarrow \quad \vec{r}_0(t) = \vec{v}_0 t + \tfrac{1}{2}\vec{g}t^2$$

The next step is the same too, at least in principle.

$$\frac{d^2\vec{r}_1}{dt^2} - \vec{g} = -\frac{b}{m}|\dot{\vec{r}}_0|\dot{\vec{r}}_0$$

where $\quad \dot{\vec{r}}_0 = \vec{v}_0 + \vec{g}t \quad$ and $\quad |\dot{\vec{r}}_0| = \left[v_0^2 + 2\vec{v}_0 \cdot \vec{g}t + g^2 t^2\right]^{1/2}$

Can you integrate this? Yes, although you probably don't want to. The result is a complicated combination of algebraic and inverse hyperbolic functions. It is not enlightening, and spending time here to carry out the algebra just isn't worth it. I have included this example only to show that this iterative method can be used even in complicated cases, and if you really *need* the results, you can get them.

Curve Balls
A more interesting example: Why does a curve ball curve? In baseball the pitcher stands 60 ft (18.3 m) from the batter and throws the ball at a speed that would violate traffic laws anywhere but on the Autobahn. A skilled pitcher can throw the ball so that it curves in various directions by a scarily large amount (scarily if you're the batter). If you're more familiar with soccer, or international football (the spherical kind) then you know that a properly kicked ball can move on a surprisingly bent path. This is the same phenomenon as the curved baseball path, so I'll describe everything in one language: baseball.
　　The spinning ball is moving through the air, and because the air has some viscosity some of the air just next to the surface of the ball is dragged along with that surface. The details are complicated, but when the dust settles the additional force due to the interaction of the ball's spin with the air can be reasonable well described by a simple expression called the Magnus force.

$$\vec{F}_{\mathsf{M}} = C\vec{\omega} \times \vec{v} \tag{4.51}$$

What is C? Wind tunnel tests verify that this force is proportional to ωv, so C can't involve either of those factors. Dimensional analysis can provide a lot of information. What can C possibly depend on?

$\rho,$ the density of air, dimensions M/L^3
$\eta,$ the viscosity of air, dimensions M/LT
$R,$ the radius of the ball, dimensions L

C itself must have dimensions force over acceleration $=\mathsf{M}$. What combination of factors will produce the correct dimensions for C?

$$[C] = \mathsf{M} = [\rho]^a \, [\eta]^b \, [R]^c = \left(\frac{\mathsf{M}}{\mathsf{L}^3}\right)^a \left(\frac{\mathsf{M}}{\mathsf{LT}}\right)^b (\mathsf{L})^c$$

$$\longrightarrow \quad a+b=1, \qquad -3a-b+c=0, \qquad -b=0$$

$$\longrightarrow \quad a=1, \quad b=0, \quad c=3$$

The expression for C is then some dimensionless factor times ρR^3. A value of $\tfrac{1}{2}\rho R^3$ is a good place to start for this factor, but it's never quite that simple.

On a baseball the size of this force is small compared to gravity, so this iterative method should work well here. Follow the pattern of the example Eqs. (4.47)–(4.49).

$$m\frac{d^2\vec{r}}{dt^2} = m\vec{g} + \vec{F}_{\text{air resistance}} + C\vec{\omega} \times \vec{v} \tag{4.52}$$

The air resistance can be either the simple linear one, or quadratic, or something worse, but for the present purposes it doesn't matter. The iterative method allows you to solve that part separately from the Magnus force and then, for this first order iteration, to add the two results.

Do the Magnus force alone then.

$$m\frac{d^2\vec{r}}{dt^2} = m\vec{g} \longrightarrow \vec{r}_0(t) = \vec{v}_0 t + \tfrac{1}{2}\vec{g}t^2$$

$$m\frac{d^2\vec{r}_1}{dt^2} = m\vec{g} + C\vec{\omega} \times \dot{\vec{r}}_0(t) = C\vec{\omega} \times (\vec{v}_0 + \vec{g}t)$$

$$\vec{r}_1(t) = \vec{A} + \vec{B}t + \tfrac{1}{2}\vec{g}t^2 + (C/m)\vec{\omega} \times \left(\tfrac{1}{2}\vec{v}_0 t^2 + \tfrac{1}{6}\vec{g}t^3\right)$$

$$= \vec{v}_0 t + \tfrac{1}{2}\vec{g}t^2 + (C/m)\,\vec{\omega} \times \left(\tfrac{1}{2}\vec{v}_0 t^2 + \tfrac{1}{6}\vec{g}t^3\right) \tag{4.53}$$

Now for the fun part: Where is the curve ball in this equation? Start by looking at the relative sizes of the last two terms. Factor a $\tfrac{1}{2}t$ from $\left(\tfrac{1}{2}\vec{v}_0 t^2 + \tfrac{1}{6}\vec{g}t^3\right)$ to get two terms, $\vec{v}_0 t$ and $\tfrac{1}{3}\vec{g}t^2$. When the ball reaches the plate, the first of these is just the distance $\vec{v}_0 t$ from the pitcher to the batter. The second is $\tfrac{2}{3} \cdot \tfrac{1}{2}\vec{g}t^2$, and that is $\tfrac{2}{3}$ the amount that the ball will naturally drop by the time it reaches the plate. In this case, the second term is really going to be much less than the first. Concentrate then on

$$\vec{v}_0 t + \tfrac{1}{2}t^2\left[\vec{g} + (C/m)\vec{\omega} \times \vec{v}_0\right] \tag{4.54}$$

and picture it as the batter will see it. The ball is coming at high speed almost straight at you. At least you *hope* that it is "almost". Assume that the ball is thrown horizontally, simply because that makes it easier to picture what will happen.

$$\vec{v}_0 \text{ is horizontal.} \qquad \vec{g} \text{ is down.} \qquad \vec{\omega} \times \vec{v}_0 \text{ is perpendicular to } \vec{v}_0$$

This means that the ball's acceleration is not necessarily *down*. It can be in any direction that is perpendicular to the ball's initial velocity. If the Magnus term is large enough, it could in principal be up, though probably no human pitcher is able do this.

What is the ratio of the second term in the brackets of Eq. (4.54) to the first? Take $C = \tfrac{1}{2}\rho R^3$, then the ratio is

$$\frac{\rho R^3 \omega v_0}{2mg} = \frac{(1.27\,\text{kg/m}^3)\cdot(3.7\,\text{cm})^3}{2\cdot 145\,\text{gm}\cdot 9.8\,\text{m/s}^2}\,\omega v_0 = 2.3 \times 10^{-5}\,\frac{\text{s}^2}{\text{m}}\,\omega v_0$$

If v_0 is 85 mph, or 38 m/s, and ω is 20 rev/s, or 126 rad/s, then this ratio is about 0.1, and that's a big effect. It is especially big when you remember that the batter has only a time $d/v_0 = 18\,\text{m}/(38\,\text{m/s}) = 0.5\,\text{s}$ from when the pitcher throws the ball to when the batter is expected to hit it.

Take left and right from the pitcher's vantage to describe this.

If $\vec{\omega} \times \vec{v}_0$ is $\widehat{\text{down}}$ ($\vec{\omega}$ left) then the added force is down; the ball sinks faster than $\vec{g}$ alone would cause.

If $\vec{\omega} \times \vec{v}_0$ is $\widehat{\text{up}}$ ($\vec{\omega}$ right) then it will sink less than expected.

If $\vec{\omega} \times \vec{v}_0$ is $\widehat{\text{left}}$ ($\vec{\omega}$ up) then it is a left curving ball. And of course $\vec{\omega}$ down implies a right curve.

What is the *shape* of this curve in the last two cases? That's easy. It just looks like the usual parabolic trajectory of a ball moving in a gravitational field, but that field isn't pointing exactly down; it has a component either left or right. Instead of $\vec{g}$, you have $\vec{g}\,'$. Equation (4.54) is

$$\vec{r}(t) = \vec{v}_0 t + \tfrac{1}{2}\vec{g}\,' t^2 \qquad \text{with} \qquad \vec{g}\,' = \vec{g} + (C/m)\vec{\omega} \times \vec{v}_0 \qquad (4.55)$$

Will this be a tilted parabola as in Figure 4.1 or Eq. (4.50)? No, because $\vec{\omega} \times \vec{v}_0$ is always perpendicular to $\vec{v}_0$. Instead, this parabola is rotated around the $\vec{v}_0$ direction as an axis.

Why do batters think that the ball "breaks"? Why should it appear that the ball is coming pretty nearly straight at you for a while and then it takes a sudden detour? The pitcher will throw the ball nearly horizontally, maybe a little up. During the first half of the ball's trip to the plate it will drop at most one-fourth of its total drop, and less than that if it starts out with a slight upward component of velocity. On average the distance from the batter to the ball is three times as large during the first half of the ball's trip as in the second half, so the batter sees it move sideways through an *angle* that is about $3 \times 3 = 9$ times greater during the last half of the trip as during the first half. Hence the illusion of a breaking ball.

Compare the effects for a baseball with mass 145 gm, radius 3.7 cm and the Association football (soccer ball) with mass 425 gm, radius 11 cm. The coefficient C/m in the latter case is larger by a factor $(11/3.7)^3/(425/145) = 9$. Of course you still have to get a large enough ω. That's the art and why Beckham earned the big bucks.

4.6 Pendulum, large angles˙

The pendulum was an example of a simple harmonic oscillator in the last chapter, but it wasn't exactly harmonic, any more than any other oscillation about an equilibrium will be. In Eq. (3.23) you had something that was not a simple harmonic oscillator, and it became simple only in the small angle approximation. What if the angle isn't small?

Solving for the position as a function of time is difficult, but if all you want is the *period* of oscillation then that's not so hard — all that is involved is doing a difficult integral. In the equation (3.27) the energy is
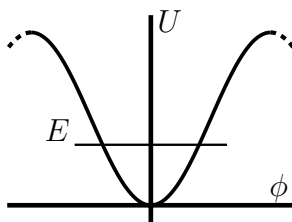
$$E = \frac{1}{2}m\ell^2\dot{\phi}^2 + mg\ell(1 - \cos\phi) \qquad (4.56)$$

In that chapter, I differentiated the energy, $dE/dt = 0$ and got the differential equations of motion. Instead, solve this energy equation for dt directly.

$$\frac{2}{m\ell^2}\left[E - mg\ell(1 - \cos\phi)\right] = \left(\frac{d\phi}{dt}\right)^2 \quad \longrightarrow \quad dt = \frac{d\phi}{\sqrt{(2E/m\ell^2) - (2g/\ell)(1 - \cos\phi)}}$$

This separates the equation, and all you have to do now is to integrate. It is a tough integral but not impossible, and it helps to do a little manipulation first. Remember the trigonometric identity for half-angles?

$$\sin^2\frac{\phi}{2} = \frac{1}{2}(1 - \cos\phi)$$

$$dt = \frac{d\phi}{\sqrt{(2E/m\ell^2) - (4g/\ell)\sin^2(\phi/2)}} = \sqrt{\frac{\ell}{4g}} \frac{d\phi}{\sqrt{(E/2mg\ell) - \sin^2(\phi/2)}} \tag{4.57}$$

Why manipulate the expression in this way? It is not obvious, and I'm sure that the first person to solve this problem did not find the right path through the maze the first time.

The motion of the pendulum is back and forth between the stopping points as in section 2.3. The stopping points occur when the kinetic energy drops to zero — when the line of constant E in the graph, Eq. (4.56), crosses the curve $U(\phi)$. That is when $\dot{\phi} = 0$,

$$\dot{\phi} = 0 \implies E = mg\ell(1 - \cos\phi) = 2mg\ell\sin^2(\phi/2), \qquad \phi = \phi_{\text{max}}$$

and that is the point where the denominator under $d\phi$ vanishes in Eq. (4.57).

Now try to do the integral of dt to get the period of the pendulum, and up to this point all the manipulations have been the same kind that you see in a "techniques of integration" chapter in your introductory calculus text. But now comes a clever trick. Again, it is nothing beyond what is done in beginning calculus (a change of variables), but it is *not* an obvious change. The period of the pendulum is four times the time it takes to go from the center to the stopping point

$$T = 4\int dt = 4\int_0^{\phi_{\text{max}}} \sqrt{\frac{\ell}{4g}} \frac{d\phi}{\sqrt{(E/2mg\ell) - \sin^2(\phi/2)}} \qquad \text{where} \qquad \phi_{\text{max}} = 2\sin^{-1}\sqrt{\frac{E}{2mg\ell}}$$

The integral goes from zero to a limit that depends on the energy. It would be much easier to handle if the limit is fixed. The trick: let

$$\sqrt{\frac{E}{2mg\ell}}\,\sin\theta = \sin\frac{\phi}{2} \tag{4.58}$$

This is designed so that as θ goes from zero to $\pi/2$, the angle ϕ goes from zero to ϕ_{max}. Why should this help? If it did not I wouldn't be spending time on it. Execute the change of variables.

$$\sqrt{\frac{E}{2mg\ell}}\,\cos\theta\,d\theta = \frac{1}{2}\cos\frac{\phi}{2}\,d\phi$$

$$\frac{d\phi}{\sqrt{(E/2mg\ell) - \sin^2(\phi/2)}} = \sqrt{\frac{E}{2mg\ell}}\,\frac{\cos\theta\,d\theta}{\frac{1}{2}\cos\frac{\phi}{2}\sqrt{(E/2mg\ell) - \sin^2(\phi/2)}}$$

$$= 2\frac{\cos\theta\,d\theta}{\sqrt{1 - \sin^2(\phi/2)}\sqrt{1 - \sin^2\theta}} \tag{4.59}$$

$$= 2\frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}} \qquad \text{where} \quad k^2 = \frac{E}{2mg\ell} = \sin^2\frac{\phi_{\text{max}}}{2}$$

Assemble everything to get the period of oscillation

$$T = 4\int dt = 4\sqrt{\frac{\ell}{4g}}\int_0^{\pi/2} d\theta\,\frac{2}{\sqrt{1 - k^2\sin^2\theta}} = 4\sqrt{\frac{\ell}{g}}\int_0^{\pi/2} d\theta\,\frac{1}{\sqrt{1 - k^2\sin^2\theta}} \tag{4.60}$$

There are a lot of missing steps in this derivation, but not big ones, and you should carry out these details for yourself. Why k^2? That's simply the most common convention in the treatment of these

Elliptic Integrals though not the only one. In Abramowitz and Stegun they use the letter m instead of k^2. The standard notation for this "elliptic integral of the first kind" is

$$K = \int_0^{\pi/2} d\theta \, \frac{1}{\sqrt{1 - k^2 \sin^2 \theta}}, \qquad \text{so} \qquad T = 4K \sqrt{\frac{\ell}{g}} \qquad (4.61)$$

In section 6.13 you will find a brief introduction to these functions and some of their properties.

Should I believe the answer? If the oscillations are small enough this must reduce to the solution for T with oscillations having small enough energy to allow small angle approximations. Small enough energy means that the parameter $k^2 = \sqrt{E/mg\ell} \ll 1$. Start by making it zero.

$$T = 4\sqrt{\frac{\ell}{g}} \int_0^{\pi/2} d\theta \, \frac{1}{1} = 4\sqrt{\frac{\ell}{g}} \frac{\pi}{2} = 2\pi \sqrt{\frac{\ell}{g}}$$

and that is $2\pi/\omega_0$ from Eq. (3.25), so it works.

Now what if the energy is not so small? k^2 can still be small but not negligible. Do a power series expansion using the binomial series of Eq. (0.1).

$$(1 + x)^n = 1 + nx + \frac{n(n-1)x^2}{2!} + \cdots \qquad \text{and for } n = -1/2 \text{ this is}$$

$$1 - \frac{1}{2}x + \frac{1 \cdot 3}{2^2 \cdot 2!}x^2 - \frac{1 \cdot 3 \cdot 5}{2^3 \cdot 3!}x^3 + \cdots$$

$$T = 4\sqrt{\frac{\ell}{g}} \int_0^{\pi/2} d\theta \, \left[1 + \frac{1}{2}k^2 \sin^2 \theta + \frac{1 \cdot 3}{2^2 \cdot 2!}k^4 \sin^4 \theta + \cdots \right] \qquad (4.62)$$

$$= 2\pi \sqrt{\frac{\ell}{g}} \left[1 + \left(\frac{1}{2}\right)^2 \sin^2 \frac{\phi_{\text{max}}}{2} + \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 \sin^4 \frac{\phi_{\text{max}}}{2} + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 \sin^6 \frac{\phi_{\text{max}}}{2} + \cdots \right]$$

$$(4.63)$$

You can easily do the second term, the one in $\sin^2 \theta$, yourself, but the rest of them require another integral, that of $\sin^{2n} \theta \, d\theta$, and you can either learn how to do it (probably with complex variables) or have a big enough table of integrals such as that by Gradshteyn and Ryzhik. The integral here is equation number 3.621.3 in their book:

$$\int_0^{\pi/2} d\theta \, \sin^{2n} \theta = \frac{(2n-1)!!}{(2n)!!} \frac{\pi}{2}$$

The double factorial notation means the product of every other integer up to the given one. For example $5!! = 15$ and $6!! = 48$.

For an angle of $\phi_{\text{max}} = 30°$ this series is

$$2\pi \sqrt{\frac{\ell}{g}} \left[1 + 0.01674 + 0.000005 + \cdots \right], \qquad [\,] = 1.017 \qquad (4.64)$$

For 90° it is

$$2\pi \sqrt{\frac{\ell}{g}} \left[1 + 0.1250 + 0.0352 + 0.0122 + 0.0003 + \cdots \right], \qquad [\,] = 1.173$$

Even at $90°$ this series converges rapidly and changes the small angle result by about 17%. If you want to try $179°$ you will quickly conclude that there's got to be a better way, because this time the sine of $\phi_{max}/2$ is close to one and the series converges very slowly. At this point, look up a reference on elliptic integrals to find that the quantity in brackets is approximately 3.90, evaluated using Eq. (4.65) in the next paragraph.

$$k^2 = \sin^2(179°/2) = 0.9999238 \longrightarrow K \approx 6.128 \longrightarrow [\,] = \tfrac{2}{\pi}K = 3.90$$

The defining equation for this elliptic integral can be evaluated by the series shown for not too big values of k, and the equation following is an approximate expression that is useful for larger k (near one). It is from Abramowitz and Stegun, equation (17.3.26).

$$K = \int_0^{\pi/2} d\theta \, \frac{1}{\sqrt{1 - k^2 \sin^2 \theta}} = \frac{\pi}{2}\left[1 + \left(\frac{1}{2}\right)^2 k^2 + \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 k^4 + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 k^6 + \cdots\right]$$

$$\sim \frac{1}{2}\ln\left(\frac{16}{1 - k^2}\right) \quad \text{as } k \to 1 \tag{4.65}$$

Though this function approaches infinity as $k \to 1$, it does so slowly. The logarithm has a very mild sort of infinity.

Exercises

1 Do problem 0.60.

2 On page 126 it says that "getting x_2 by putting x_1 into Eq. (4.10)" would be wrong. Why?

3 Using Eq. (4.11), compare the size of the correction term for $\theta = 1°$ and for $\theta = 89°$.

4 Apply Eq. (4.18) to Eq. (4.19).

5 What happens to the equation (4.1) if there is a wind, $\vec{v}_w$? This wind can be in any direction.

6 Draw several equipotentials for the potential energy in Eq. (4.24). Draw some force vectors in the same picture.

7 Check the radial force function $\vec{F}(x, y) = (k/r^2)\hat{r} = (k/r^3)\vec{r}$ against Eq. (4.18).

8 Using only the elementary methods (before you ever heard of a differential equation), Find the relations between speed, radius, angular frequency, and magnetic field strength for a charged particle moving in a uniform magnetic field.

9 Rearrange Eq. (4.46) in terms of the cyclotron frequency at $z = 0$, qB_0/m, and the radius of the orbit at $z = 0$, r_0.

10 Can you use the same method as in Eq. (4.34) to define $e^{\vec{B}\cdot}$?

11 In the paragraph after Eq. (4.55), it says "an angle that is about $3 \times 3 = 9$ times greater". Where do the two factors of three come from?

12 If a pendulum clock had an error equivalent to the correction in Eq. (4.64), how much would it gain or lose in one day (and which)? And is the number in this equation correct?

Problems

4.1 Fill in the missing steps in taking the limits, Eqs. (4.4) and (4.5). Also, find the next order terms in b for these equations — the first non-vanishing correction.

4.2 Derive Eq. (4.7).

4.3 What is the maximum height as found from either Eq. (4.3) or (4.7).

4.4 In the equations (4.3), if you turn off gravity the results should reduce to those of Eq. (2.16). Do they?

4.5 Expand Eq. (4.7) for small b, keeping terms up to those linear in b. What is the range of the projectile, assuming a level surface, and getting your answer correct to terms linear in b. Is your answer plausible?

4.6 A mass drops from rest without air resistance, starting from coordinates $x = x_0$, $y = y_0$, $z = 0$. Compute its angular momentum (Eq. (1.13)) about the origin at time t; compare it to the torque on the mass. Repeat this, but pick your origin for the angular momentum and torque about the point $(x, y, z) = (2x_0, 0, 0)$.

4.7 The equation (4.7) will tell you where the projectile hits the ground: the range. The equation for the ground is $y = 0$ if you assume that the ground is level. First take the case $b = 0$ and solve for the range; for that you have the simple equations (4.4) and (4.5). For small air resistance (small b) the range is claimed to be approximately $(2v_0^2/g)[\sin\theta\cos\theta - \frac{4}{3}(bv_0/mg)\sin^2\theta\cos\theta]$. Analyze this claim for plausibility or lack thereof.

4.8̈ Derive the expression in the preceding problem. Expand Eq. (4.7) to the first order in b and find the first order correction to the range. If you haven't done problem 0.60, now would be a good time; the quadratic equation that you get here is best solved by iteration. If you find something not quite right about the answer stated in the preceding problem, find a better answer and then check *it* for plausibility.

4.9 If you fire a projectile up a hill and want to find its range, the equation for the ground is no longer $y = 0$. Assume that the hill is represented by a straight line at an angle α from the horizontal and write its equation. Find the range of a projectile fired up such a hill. For this problem, assume zero air resistance. Do you measure the position where it hits horizontally or along the hill? I chose along the hill. (And is the firing angle ϕ measured from the horizontal or from the hill itself? Your choice.) Do you choose your coordinate system x-y as horizontal and vertical or do you choose it parallel and perpendicular to the hill? Again, your choice. Instead of solving the problem, the answer for the range is claimed to be $(2v_0^2/g)\left[\sin\phi\cos(\phi+\alpha)/\cos^2\alpha\right]$. Is this plausible? Analyze it completely. And what is ϕ? You figure out whether I measured ϕ from the hill or the horizontal.

4.10˙ Solve the the preceding problem for the range. At what angle do you fire it to get the maximum range? Show that it is halfway between the hill and the vertical.

4.11 The projectile analyzed in section 4.1 now has a headwind; the air has a horizontal velocity $-v_{\text{air}}\hat{x}$. Solve for the trajectory now and analyze the result. What wind will it take to bring the mass back to its starting point?

4.12 Following the methods of section 3.11, find the Green's function solution for equations of the form Eq. (4.2), but for an arbitrary force on the right: $m\frac{d^2y}{dt^2} + b\frac{dy}{dt} = F(t)$.

Ans: $\frac{1}{b}\int_{-\infty}^{t} dt' \, F(t')\left(1 - e^{-b(t-t')/m}\right)$

4.13 (a) Now use the formula from the preceding problem and apply it to Eq. (4.2).
(b) The result doesn't agree with the y part of Eq. (4.3) because the initial condition $v_y(0) = v_0 \sin \theta$ is not included. You can handle that if you remember that $-mg$ is *not* the only force that acts on the projectile. You had to apply a force to give it the initial velocity; that is an impulse, a large force acting for a very short time, and it is $F_y(0)\,dt = mv_0 \sin \theta$. If you include this in the integral, so that the very first time interval contributes a large chunk to the whole integral, you add this to the result from part (a) to get the full solution for $y(t)$ in agreement with Eq. (4.3).

4.14 Start from $\vec{r}(t) = \hat{x}\,r(t) \cos\left(\phi(t)\right) + \hat{y}\,r(t)\sin\left(\phi(t)\right)$ and differentiate it once and twice. Rearrange the components and group the terms in order to derive the expression for acceleration found in Eq. (0.41), expressed in polar coordinates. Here you are getting the same result by a different method, and you will need this result a lot in chapter six.

4.15 Draw pictures of the equipotentials and of $\vec{F}$ for the example in Eq. (4.24).

4.16 A mass m is attached to a very light string, and the other end of the string is attached to the ceiling. Set the mass moving so that it goes in a circle parallel to the floor, a conical pendulum. Find the relations among the length ℓ of the string, the angle α that the string makes with the vertical, the radius of the circle, the speed of the mass, the tension in the cord, and any other parameters that may enter the problem. How does the speed vary with the angle from the vertical? Same for the rotational frequency. In the end, express your answer in terms of constants and the angle α. And of course your answer makes sense doesn't it?

4.17 In the preceding problem of a conical pendulum, add some reality to it by assuming now that there is a small air resistance affecting the motion of the mass. Take this force to be $-b\vec{v}$. As energy it lost, the angle α will gradually drop, and it is claimed that the rate at which this angle changes is

$$\frac{d\alpha}{dt} = -2\frac{b}{m}\frac{\sin\alpha \cos\alpha}{1 + 3\cos^2\alpha}$$

Is this plausible? Especially, look at both small and large angles.

4.18 Derive the result claimed in the preceding problem. What is the kinetic energy and the potential energy of the mass? Express the total mechanical energy in term of the angle α the string makes with the vertical, eliminating any other parameters that depend on time. Assume that the air resistance is $-b\vec{v}$, and determine the rate of loss of energy due to this resistance, dE/dt. (a) What then is $d\alpha/dt$? Recall: $d\alpha/dt = (dE/dt)/(dE/d\alpha)$. Ans: see the preceding problem. (b) Since you've done all this work, what is the result if the air resistance is $-bv^2$, which is a more realistic value anyway. You need not do any more than about two additional lines of algebra because you have already done the hard part. There are some significant differences in the way these two solutions behave. What are they?
Ans: (b) $d\alpha/dt = -\left[2b\ell^{1/2}g^{1/2}/m\right]\left[\sin^2\alpha \cos^{1/2}\alpha/(1 + 3\cos^2\alpha)\right]$

4.19 Eq. (4.28) can be solved several different ways, as mentioned in the paragraph right after it. Do it using method (2), assuming a solution $\dot{x} = A_1 e^{\alpha t}$ and $\dot{y} = A_2 e^{\alpha t}$ just as in Eq. (3.60). Follow that method through to get a non-zero solution for A_1 and A_2 and then on to the complete solution.

4.20 As stated in the second paragraph of section 4.3, you can solve for the motion of a charge in a uniform magnetic field by elementary means, using such equations as $a = v^2/r$. Do so. Is the result compatible with the expression $\omega = qB_0/m$ as found in the text?

4.21 Show that the solutions in Eqs. (4.36) and (4.31) agree.

4.22 For motion in uniform electric and magnetic fields, start a charge at the origin with zero velocity and take $\vec{E} = E_0\hat{x}$, with $\vec{B} = B_0\hat{x}$. Find the motion of the charge.

4.23 There is a lot of algebra between equations (4.56) and (4.63). (**a**) Fill in the missing steps. (**b**) Then evaluate the integrals for the first two terms in the series Eqs. (4.62) and (4.63), through $\sin^2$.

4.24˙There is one case for which the integral in Eq. (4.57) is possible in terms of elementary functions: $E = 2mg\ell$. Start the pendulum at the bottom with exactly this kinetic energy and you want the angle as a function of time. The answer is proposed to be

$$\phi(t) = 2\tan^{-1}\sinh\omega t \qquad \text{where} \qquad \omega = \sqrt{g/\ell}$$

(**a**) First show that this is the same as

$$2\cos^{-1}\text{sech}\,\omega t \qquad \text{and} \qquad 2\sin^{-1}\tanh\omega t$$

You never know which form of the result will be easier to interpret. (**b**) Also show that this result is, for large values of time, approximately $\phi \approx \pi - 4e^{-\omega t}$. If you can't find the trigonometric identities to tell you this, then at least check it numerically for a couple of values of ωt. (**c**) Are these equations for ϕ plausible? Is the energy even correct? Include a graph of $\phi(t)$.

4.25¨Derive the result claimed in the preceding problem. This is a tough integral, so instead of beating your head against it (or looking up the subject of Gudermannians), verify that this function as written in any one of these three forms, satisfies the conservation of energy equation (4.56).

4.26 A particle of mass m can slide freely along a wire. The wire is straight and horizontal, and is made to rotate in a horizontal plane about an axis perpendicular to one end at a fixed angular speed ω_0. At time $t = 0$ the mass is at radius r_0 and has zero radial velocity. Find the future motion of m. Assume zero friction. What is the force that the wire exerts on m? Have you done problem 4.14 yet? Sketch the motion. Ans: $r = r_0\cosh\omega_0 t$, $F_\phi = 2m\omega_0^2 r_0\sinh\omega_0 t$

4.27 A particle of mass m is displaced slightly from its equilibrium position at the top of a smooth fixed sphere of radius R, and it slides because of gravity. Through what angle does the particle move before it leaves the sphere?

4.28 In Eq. (2.27) you determine the escape speed from the Earth. That assumed that the planet is not rotating. At the equator you have the speed of the ground added to the total velocity vector. Take that into account and determine the minimum speed that you have to give the rocket relative to the ground so that it escapes. Put in the numbers and compare the two speeds.

4.29 A gravitational force is $\vec{F} = -GMm\hat{r}/r^2$. Show that $F_x = -GMmx/r^3$ and that this force satisfies Eq. (4.18).

4.30 In Eq. (4.59) and those preceding it, what happens if $E > 2mg\ell$? And draw some graphs.

4.31 If the work integral in three dimensions is independent of path, then in particular, the work going around a closed loop is zero. (One possible path from a point to itself just sits still, going nowhere.) Apply this to a very tiny rectangular loop in the x-y plane, from y to $y + \Delta y$ and from x to $x + \Delta x$. Since this rectangle is so small, you can approximate the value of the integral along each side by assuming that $\vec{F}$ has the constant value it attains at the midpoint of each side. For example, on the right-hand side $\vec{F} = \vec{F}(x + \Delta x, y + \frac{1}{2}\Delta y)$. For that side, $d\vec{\ell} \to \hat{y}\Delta y$. The integral in this approximation is just the sum of four dot products. Divide by the area of the rectangle and take the limit as Δx and $\Delta y \to 0$. Show that the result is Eq. (4.18).
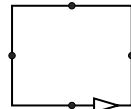
4.32 Using crossed electric and magnetic fields as in Eq. (4.38), find the initial conditions on the velocity so that the charge is unaccelerated, and determine conditions on the magnitudes of $\vec{E}$ and $\vec{B}$ so that this speed is less than that of light.

4.33 To obtain Eq. (4.40) the charge started at rest. What is the shape of the path of the charge if the initial velocity is in the $-\hat{x}$ direction? In the $+\hat{x}$ direction?

4.34 ˙A charged particle has an initial velocity perpendicular to a uniform magnetic field (no $\vec{E}$). Now there is friction with the air, assumed to be a force $-b\vec{v}$. Find (and of course sketch) the charge's subsequent motion. Note: This assumption $-b\vec{v}$ is not a good one for this process. A constant magnitude force (like dry friction) is better, but the mathematics is much harder in that case. Suggestion: You've seen in chapter three what a damped oscillator looks like. This is more of the same, so a substitution $\vec{v}(t) = \vec{v}_1(t)e^{\alpha t}$ will help. You will have to choose α. Using the complex combination $\dot{z} = \dot{x} + i\dot{y}$ is useful too.. Ans: $x(t) = R\cos(\omega t + \phi)e^{-bt/m}$, $\quad y(t) = -R\sin(\omega t + \phi)e^{-bt/m}$ (depending on choice of origin)

4.35 In the magnetic mirror of section 4.4, and for motion along the central axis, where is the stopping point? Did you do the exercise #8 to express this in terms of the cyclotron frequency?

4.36 On page 138 you find the equation to describe all the magnetic field lines for the mirror stated to be $\sqrt{x^2 + y^2} = r_\perp = \alpha/\sqrt{\ell^2 + z^2}$. To derive this, use Gauss's law for magnetism, $\oint \vec{B} \cdot d\vec{A} = 0$, applying it to the closed surface that consists of two disks, centered on the z-axis and parallel to the x-y plane at $z = z_1$ and $z = z_2$. The rest of the closed surface follows the $\vec{B}$-field lines to connect the edges of the two disks. (That means that given z_1 and z_2, the radius of one disk determines the radius of the other, and no flux escapes the side.)

4.37 In the same spirit as the operator solution in section 4.3, starting at Eq. (4.33), interpret the following operator acting on functions of the real variable x.

$$e^{h\frac{d}{dx}} f(x) \qquad (h \text{ is a constant})$$

4.38 ˙Going back to section 4.1 again, there is an equation, (4.11), that gives an approximate result for the firing range with air resistance. At what angle is the range maximum? It is no longer $45°$. Differentiate with respect to θ and remember that the second term is small and *also* that $\cos 2\theta$ is small near the root. This will let you use an iterative method again to get a simple solution. $\theta = \frac{\pi}{4} + \delta$. Ans: $\delta = bv_0/(3mg\sqrt{2})$

4.39 An Atwood machine has two masses hung over a pulley, and idealized versions use massless strings and pulleys. In this otherwise ideal Atwood machine, submerge one of the masses in water so that it feels a force $-b\vec{v}$. Now solve for the motion assuming that the system starts from rest. Neglect buoyancy effects.

4.40 ˙There's more to say about the coupled oscillators of section 3.9. Apply an oscillating force to one of the masses: $F_0 \cos \Omega t$. Write the equations of motion as done there, and now examine the *inhomogeneous* (steady-state) solution to the equations in the symmetric case $m_1 = m_2$ and $k_1 = k_3$. It is just like the forced simple harmonic oscillator except that the two functions $x_1(t)$ and $x_2(t)$ will have different amplitudes. Compare it to problem 3.13. Instead of solving the problem, a proposed solution is

$$x_1(t) = \frac{F_0}{m} \frac{\frac{1}{2}(\omega_1^2 + \omega_2^2) - \Omega^2}{(\omega_1^2 - \Omega^2)(\omega_2^2 - \Omega^2)} \cos \Omega t \qquad\qquad x_2(t) = \frac{F_0}{m} \frac{\frac{1}{2}(\omega_1^2 - \omega_2^2)}{(\omega_1^2 - \Omega^2)(\omega_2^2 - \Omega^2)} \cos \Omega t$$

$$\omega_1^2 = \frac{k + 2k_2}{m} \qquad\qquad \omega_2^2 = \frac{k}{m}$$

Examine this to see if it is plausible. *Most important:* draw graphs of the amplitudes of each mass's oscillation. One graph for each mass, showing the amplitude of its oscillation versus the applied frequency Ω (or versus Ω^2 will be easier). I suggest you start by examining the neighborhood of the singularities. Why do the graphs behave this way?

4.41 ˙Solve the differential equations of motion that you wrote in the preceding problem. Follow the same procedure as in problem 3.13, except that you have simultaneous equations now, and derive the solution stated in the preceding problem.

4.42 ˙What if air resistance is modeled the same way as common dry friction? That is $\vec{F}_{\text{air resistance}} = -b\hat{v} = -b\vec{v}/v$. Use the iteration method of section 4.5 with the same initial conditions used in the first example there. That is, look at the trajectory that starts with Eq. (4.47) and use this form of friction to find the velocity as a function of time. This form is a good approximation for very energetic charged particles moving through air.

4.43 In chapter three, Eq. (3.26), there is an approximate expression for the frequency of a pendulum. Derive it from Eq. (4.63).

4.44 In problem 3.39 one spring was replaced by a damper. What if all three springs are replaced by dampers? Again, make the problem symmetric so that the two on the end are the same.

4.45 ˙The Earth's gravity drops off with height. $g(y) = -g_0 R^2/r^2$ where R is the Earth's radius; $r = R + y$ is the distance from its center; g_0 is the gravitational field at the surface. Expand this to the first order in y, getting the variation of g with altitude near the surface. Fire a projectile from the surface at initial speed v_0 and angle from the horizontal θ. Set up and solve $\vec{F} = m\vec{a}$ for this case, neglecting air resistance this time. Still assuming a flat Earth, find where it hits the ground and then the difference between this answer and the traditional one that assumes g has the constant value it had at the surface. How large is this change if $\theta = 45°$ and v_0 is large enough to get an uncorrected distance of 30 km? How does the answer scale with the uncorrected distance so that without further effort you can answer the same question changing 30 to 3 or to 60. Depending on how you solved this problem you may need to do a series expansion to get a simple result. Ans: about 100 m.

4.46 ˙The electromagnetic force that one charged particle exerts on another starts with the Coulomb law. The next approximation is that a moving charge creates a magnetic field that will affect the other charge if it too is moving. The total of these is

$$\vec{F}_{\text{on 2 by 1}} = \frac{q_1 q_2}{4\pi\epsilon_0 r^2} \left[\hat{r} + \frac{1}{c^2}\vec{v}_2 \times (\vec{v}_1 \times \hat{r}) \right]$$

where $\hat{r}$ is the unit vector pointing from charge q_1 toward charge q_2. What is the total force on the two interacting particles, $\vec{F}_{\text{on 1 by 2}} + \vec{F}_{\text{on 2 by 1}}$? To simplify the result you will have to hunt up the *Jacobi Identity* involving cross products. Be sure to check some special cases to verify your result. What happened to Newton's third law? And if you aren't able to answer this last question then ask someone about the momentum carried by radiation. Ans: $\left(q_1 q_2/4\pi\epsilon_0 c^2 r^2\right)\hat{r} \times \left(\vec{v}_1 \times \vec{v}_2\right)$

4.47 ˙Derive the result stated in the Gravitron problem 1.14. Recall that the frictional component of the force is related to the normal component by $-\mu_s F_N < F_{\text{fr}} < +\mu_s F_N$ If you follow the rules stated in section 1.4, not skipping any steps, then this is a straight-forward problem. The only twist is that because of the nature of dry friction you have to write down two inequalities instead of one equation. Pick your basis with some thought to it. Don't forget that when you multiply or divide an inequality by a negative number you have to reverse the direction of the inequality, so you have to watch for various cases.

4.48 In the discussion after the approximate solution found in Eq. (4.50), it says that if the projectile is moving rather fast, the solution still looks like a parabola, but a tilted one. Analyze this statement to see of it makes sense. Draw enough pictures to explain what is happening.

4.49 ¨This is in the spirit of problem 2.33. Place a gun at the origin and fire a bullet with speed v_0 at an angle θ with respect to the horizontal ground. Where does it land? (Ignore air resistance.) give that distance the coordinate x as measured from the origin. Now fire many bullets at random angles uniformly spread from 0 to $\pi/2$, all with the same speed and all in the same plane. (**a**) What is the probability density for a bullet's landing near the point x. Stated another way, the fraction of the bullets that are fired between angles θ and $\theta + \Delta\theta$ is $\Delta\theta/(\pi/2)$. These land between x and $x + \Delta x$. What is the fraction per Δx? In the limit as $\Delta x \to 0$ this gives dP/dx as the probability density for the bullets to land. Graph dP/dx versus x. The integral $\int dP$ should be one. Is it? *Cautions:* A particular value of x can come from two different θ's. You must add these fractions. *Also:* Sometimes a positive $d\theta$ corresponds to a positive dx; sometimes not. (**b**) The mean value of x is $\langle x \rangle = \int x \, dP = \int x \frac{dP}{dx} dx$. What is it? Also show its position on your graph of dP/dx. (**c**) Where is the median x—half above, half below.
Ans: $2/\left(\pi\sqrt{x_{\text{max}}^2 - x^2}\right)$ $(0 < x < x_{\text{max}})$, $\quad \langle x \rangle = 0.637\, x_{\text{max}}, \quad x_{\text{med}} = 0.707\, x_{\text{max}}$

Non-Inertial Systems

The first of Newton's laws, chapter one, is the definition of an inertial system. The second law says that *if* you're in an inertial system *then* $\vec{F} = m\vec{a}$ (or $d\vec{p}/dt$). The Earth is not an inertial system; it is rotating. That means that you can't use this equation on the Earth. Well no, not really. The Earth isn't rotating all that fast, so it's probably o.k. to ignore the error. Is it? How big is the error and are there any circumstances where it matters? Look at page 34 again.

For something small, the size of a house or a bathtub or even a golf course, the effect of the Earth's rotation is not large enough to detect without delicate instruments and careful controls. If you want to understand the Gulf Stream or weather systems or how to fire long-range artillery, then the effect is very important.

To use the basic equation, Newton's second law, requires being in an inertial system. That means that to describe something on the Earth you should use a sun-centered coordinate system. Trying to understand the weather is complicated enough without adding that sort of obstacle, so the better way to do it is to work out the mathematics of finding what changes to make in $\vec{F} = m\vec{a}$ in order to apply it to a non-inertial system.

5.1 Galilean Transformation

Before jumping to the rotating case, work out the simpler example of transforming to a coordinate system that is accelerating in a straight line, a car perhaps. And before that I'll look at the transformation to a coordinate system moving at constant velocity, the Galilean transformation. My coordinate system is (x, y, z, t) and your (moving) coordinate system is (x', y', z', t'). You are in motion along the my x-axis at a velocity v_0, so the change of coordinates is

$$x' = x - v_0 t, \qquad y' = y, \qquad z' = z, \qquad t' = t \tag{5.1}$$

To check that the signs are right, your moving coordinate's origin is at $x' = 0$ and that's $x - v_0 t = 0$. This is exactly what it's supposed to be because it is what I will write down in *my* coordinate system for your position. Now what are the velocity and acceleration in your system?

$x' = x - v_0 t$

Fig. 5.1

$$\frac{dx'}{dt'} = \frac{d(x - v_0 t)}{dt} = \frac{dx}{dt} - v_0$$

$$\frac{d^2 x'}{dt^2} = \frac{d}{dt}\left(\frac{dx}{dt} - v_0\right) = \frac{d^2 x}{dt^2} \tag{5.2}$$
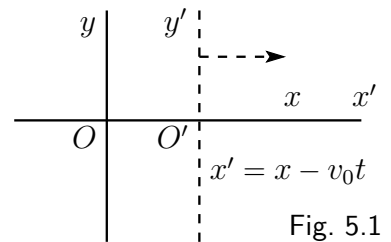
Time is the same for each system, so it doesn't matter whether you use t or t'. (That will change when in special relativity, section 9.6.) Here the acceleration is the same for both, so $\vec{F} = m\vec{a}$ is the same for both. Or is it? Suppose that the force is velocity dependent; frictional forces commonly depend on velocity. Does that mean that the force will be different in the two systems? No. All frictional forces depend on the *relative* velocity of two things. A book sliding on a table, an airplane going through the air, a swimmer in water. When you transform the velocity of one object you have to do it for all, and that means that the difference of two velocities will not change and that the forces remain unaffected.

Accelerated System

Now suppose that your moving coordinate system is accelerated at a constant rate.

$$x' = x - v_0 t - \frac{1}{2} a_0 t^2, \qquad y' = y, \qquad z' = z, \qquad t' = t \tag{5.3}$$

The position $x' = 0$ again defines the position of the moving observer. Repeat the previous calculation:

$$\frac{dx'}{dt'} = \frac{d(x - v_0 t - a_0 t^2/2)}{dt} = \frac{dx}{dt} - v_0 - a_0 t$$

$$\frac{d^2 x'}{dt^2} = \frac{d}{dt}\left(\frac{dx}{dt} - v_0 - a_0 t\right) = \frac{d^2 x}{dt^2} - a_0 \tag{5.4}$$

In the primed coordinate system Newton's second law now becomes

$$F_x = m\frac{d^2 x}{dt^2} = m\left(\frac{d^2 x'}{dt^2} + a_0\right)$$

Rearrange this and make is look like the old $F = ma$.

$$F_x - ma_0 = m\frac{d^2 x'}{dt^2} \qquad \text{The } y \text{ and } z \text{ part are unchanged} \tag{5.5}$$

If you are in the accelerated system this looks like the normal $\vec{F} = m\vec{a}$, but with an extra term that looks and behaves just like an extra force. It is not one of the basic forces of nature (gravity, electromagnetism, *etc.*) but mathematically you can handle it as if it is. It is variously called an "inertial force" or a "fictitious force" and it appears simply because of the transformation to a non-inertial system of coordinates.

Could you do something even more complicated than constant acceleration? Sure. Let the position of the moving observer be $f(t)$ where f is anything you want. Mimic the derivation that starts with Eq. (5.3) and everything works.

$$x' = x - f(t) \;\longrightarrow\; \frac{dx'}{dt} = \frac{dx}{dt} - \dot{f}(t) \;\longrightarrow\; \frac{d^2 x'}{dt^2} = \frac{d^2 x}{dt^2} - \ddot{f}(t) \;\longrightarrow\; F_x - m\ddot{f} = m\frac{d^2 x'}{dt^2}$$

Example

▶ If you are in an elevator that is accelerating upward, take the y-axis up and let the y' coordinate be fixed inside the elevator. In the elevator's coordinate system the forces on you are (1) gravity, (2) floor, and (3) inertial force. Here you are analyzing the motion in your own (self-centered) coordinate system, so you must include the extra inertial force. Your coordinate y' is constant, $d^2 y'/dt^2 = 0$, and the equation of motion (5.5) in this system is

$$m\frac{d^2 y'}{dt^2} = 0 = F_{\text{floor}} - mg - ma_0, \qquad \text{implying} \qquad F_{\text{floor}} = m(g + a_0)$$
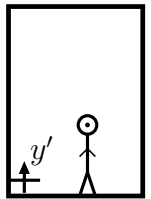
Fig. 5.2

The first two terms, from the floor and from gravity, are the two "real" forces, but the non-inertial system adds the extra term. The force that you feel from the floor makes it feel that your weight has increased, because from your perspective you are standing still and not accelerating.

What would this system look like from the inertial frame? The final answer must be the same, but now there are just the two real forces.

$$m\frac{d^2y}{dt^2} = ma_0 = F_{\text{floor}} - mg, \qquad \text{implying} \qquad F_{\text{floor}} = m(g + a_0)$$

You never feel a gravitational force! You don't feel your weight because every atom in your body is pulled in the same way. What you perceive as weight is the force by the floor on your feet, and that's really molecule-to-molecule contact—electromagnetic forces in disguise. If you jump off a diving board, that force from the floor is removed and you temporarily feel weightless. That doesn't mean that gravity has stopped, just that the force that you are able to feel (the contact force) has stopped. If you don't believe this, the next time you jump off a high-diving board carry a bathroom scale and weigh yourself on the way down. For an example to (sort of) contradict this claim, see Exercise 9 on page 179.

Example
▶ When you are a passenger in a car that is turning the corner to the left, accelerating left, you are sitting still (with respect to yourself). It's the rest of the universe that is accelerating. From your self-centered viewpoint your acceleration is zero and there are two horizontal forces on you: the seat is pushing you left and some other force is pushing you right. That's $-ma_0$. The force from the seat and the inertial force combined to keep you in equilibrium. From the vantage of an inertial pedestrian there's only one horizontal force pushing on you and that's the seat pushing left. You respond to that push by accelerating left—moving the ma_0 term from one side of the equation to the other.
▶ This idea of an accelerated coordinate system will be essential in understanding the daily tides in the ocean. See section 5.6.

5.2 Rotating System
To work out the transformation to a rotating system, it's the same idea as in the preceding work, but with vectors. The key calculation is pretty much the same as for straight-line motion: what is the time derivative of a vector that is itself expressed in the transformed system? The result has two terms, one from the change of the vector within the transformed system and the other from the change of the transformed system itself. Sort of like Eq. (0.35). For an arbitrary time-dependent vector $\vec{Q}$, the result will be

$$\frac{d\vec{Q}}{dt} = \frac{d\vec{Q}}{dt}\bigg|' + \vec{\omega} \times \vec{Q} \tag{5.6}$$

To derive this, look at it in terms of components and it is nothing more than the product rule. The vector $\vec{Q}$ is the vector $\vec{Q}$ in either coordinate system.

$$\vec{Q} = Q_x\hat{x} + Q_y\hat{y} + Q_z\hat{z} = Q'_x\hat{x}' + Q'_y\hat{y}' + Q'_z\hat{z}'$$

The first set of components, Q_x, Q_y, Q_z, is expressed with respect to a stationary set of basis vectors, so to differentiate $\vec{Q}$ all that's needed is to differentiate its components. The second (primed) set is in terms of the rotating basis vectors. For the time-derivative of $\vec{Q}$ in *this* basis, use the product rule for differentiation.

$$\frac{d\vec{Q}}{dt} = \frac{dQ'_x}{dt}\hat{x}' + Q'_x\frac{d\hat{x}'}{dt} + \cdots \tag{5.7}$$

The first term, such as $(dQ'_x/dt)\hat{x}'$, of each pair is nothing more than the time derivative that the rotating observer sees. That person says that $\hat{x}'$ is fixed, but that the component Q'_x may be changing.

To complete the equation, all that's left is to do the derivatives of the rotating unit vectors with respect to the fixed (inertial) system. That is, to compute $d\hat{x}'/dt$. The first observation to make about such a derivative is that rotations leave magnitudes unchanged. For any vector, unit vector or not, if the time derivative of u^2 is zero, then

$$\frac{d}{dt}(\vec{u} \cdot \vec{u}) = 0 = 2\vec{u} \cdot \frac{d\vec{u}}{dt}$$

That means that if the vector $\vec{u}$ has constant length then the time derivative of $\vec{u}$ is perpendicular to $\vec{u}$ itself. To get the details of this derivative, draw a picture of a vector rotating about an axis with angular speed ω.

Fig. 5.3

In time dt the tip of the vector $\vec{u}$ rotates by an angle $d\phi = \omega dt$ about the $\vec{\omega}$-axis, and along a circle of radius $u \sin \theta$. The length of the arc along this circle is then $rd\phi = u \sin \theta \, \omega \, dt$, and the length of the derivative is

$$\frac{|d\vec{u}|}{dt} = \frac{u \sin \theta \, \omega dt}{dt} = u \sin \theta \, \omega$$

The direction of $d\vec{u}/dt$ is perpendicular to $\vec{u}$ and from the picture it is also perpendicular to the vector $\vec{\omega}$. Put those two facts together with this magnitude, $u \, \omega \sin \theta$, and that is the definition of the cross product, where $\vec{\omega}$ is the angular velocity vector.

$$\frac{d\vec{u}}{dt} = \vec{\omega} \times \vec{u} \tag{5.8}$$

Now go back to Eq. (5.7) and apply this result to the derivatives of the rotating unit vectors.

$$\begin{aligned}
\frac{d\vec{Q}}{dt} &= \frac{dQ'_x}{dt}\hat{x}' + Q'_x \frac{d\hat{x}'}{dt} + \cdots \\
&= \frac{dQ'_x}{dt}\hat{x}' + Q'_x \vec{\omega} \times \hat{x}' + \frac{dQ'_y}{dt}\hat{y}' + Q'_y \vec{\omega} \times \hat{y}' + \frac{dQ'_z}{dt}\hat{z}' + Q'_z \vec{\omega} \times \hat{z}' \\
&= \frac{dQ'_x}{dt}\hat{x}' + \frac{dQ'_y}{dt}\hat{y}' + \frac{dQ'_z}{dt}\hat{z}' + \vec{\omega} \times \vec{Q}
\end{aligned} \tag{5.9}$$

The first three terms form the time-derivative of $\vec{Q}$ as seen in the rotating system. That's the meaning of the first term in Eq. (5.6).

Now back to the question of transforming $\vec{F} = m\vec{a}$ into a rotating coordinate system. In the preceding equation $\vec{Q}$ can be anything. I'll first compute it for the vector $\vec{r}$ and then for the velocity vector.

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{d\vec{r}}{dt}\bigg|' + \vec{\omega} \times \vec{r}$$

In order not to get too bogged down in notation, call the first term on the right $\dot{\vec{r}}$, and **establish, for this chapter* only, the convention** that d/dt means the time derivative in the inertial system and the dot means the time derivative in the rotating system.

$$\vec{v} = \frac{d\vec{r}}{dt} = \dot{\vec{r}} + \vec{\omega} \times \vec{r} \tag{5.10}$$

---

* and section 8.6
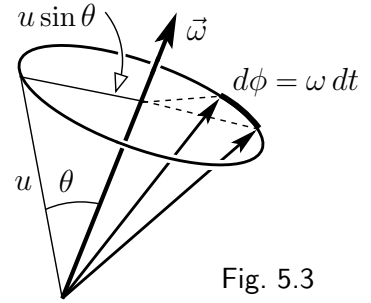
If $\omega = 0$ this equation says that there's no difference between the inertial system and the non-rotating system. Duh. If $\dot{\vec{r}} = 0$ then an object's velocity comes solely from the fact that it is held fixed in the rotating system. *E.g.* $d\hat{x}'/dt = \vec{\omega} \times \hat{x}'$.

   Now repeat the process with $\vec{Q} = \vec{v}$ and compute the acceleration

$$\vec{a} = \frac{d\vec{v}}{dt} = \frac{d}{dt}\big[\dot{\vec{r}} + \vec{\omega} \times \vec{r}\big]$$
$$= \big[\dot{\vec{r}} + \vec{\omega} \times \vec{r}\big]^{\cdot} + \vec{\omega} \times \big[\dot{\vec{r}} + \vec{\omega} \times \vec{r}\big]$$
$$= \ddot{\vec{r}} + \dot{\vec{\omega}} \times \vec{r} + \vec{\omega} \times \dot{\vec{r}} + \vec{\omega} \times \big[\dot{\vec{r}} + \vec{\omega} \times \vec{r}\big]$$
$$= \ddot{\vec{r}} + \dot{\vec{\omega}} \times \vec{r} + 2\vec{\omega} \times \dot{\vec{r}} + \vec{\omega} \times (\vec{\omega} \times \vec{r}) \tag{5.11}$$

Put this into $\vec{F} = m\vec{a}$ and move everything but the $\ddot{\vec{r}}$ term to the other side.

$$\vec{F} - m\dot{\vec{\omega}} \times \vec{r} - 2m\vec{\omega} \times \dot{\vec{r}} - m\vec{\omega} \times (\vec{\omega} \times \vec{r}) = m\ddot{\vec{r}} \tag{5.12}$$

This is the modified form of Newton's second law needed in a rotating coordinate system. Everything in this equation is now expressed in the rotating system, so all that you need to have given is $\vec{\omega}$ and $\dot{\vec{\omega}}$. That last vector, $\dot{\vec{\omega}}$, has a simplification that you can easily overlook. Apply the same equation (5.9) to the vector $\vec{\omega}$ itself. How does $\dot{\vec{\omega}}$ compare to $d\vec{\omega}/dt$?

$$\frac{d\vec{\omega}}{dt} = \dot{\vec{\omega}} + \vec{\omega} \times \vec{\omega} = \dot{\vec{\omega}} \tag{5.13}$$

This time derivative of $\vec{\omega}$ is the same in both systems. This means for example that if the rotation axis is fixed in the stationary system, then $\vec{\omega}$ is fixed in the rotating system. Look at the figure on page 284 for a not-so-obvious example of this. In fact the $\dot{\vec{\omega}}$ term in Eq. (5.12) seldom comes up; it's the other two terms that are important. They're important enough to have names:

$$-2m\vec{\omega} \times \dot{\vec{r}} \quad \text{``Coriolis force''} \qquad\qquad -m\vec{\omega} \times (\vec{\omega} \times \vec{r}) \quad \text{``Centrifugal Force''}$$

Example
▶ You are rotating at constant angular velocity and according to you a mass is standing still at a distance r away. What does this equation say? You say that the mass is not moving, so your coordinates have $\dot{\vec{r}}$ and $\ddot{\vec{r}}$ both zero. Also $\dot{\vec{\omega}} = 0$. Take your angular velocity to be $\hat{z}'\,\omega$, and put the mass along the x'-axis, at $\hat{x}'\,r$. Plug in to Eq. (5.12).
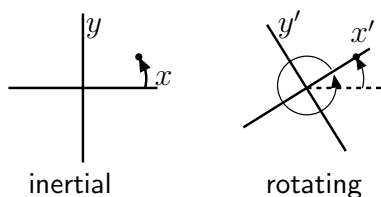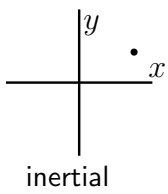

$$\vec{F} - 0 - 0 - m\hat{z}'\,\omega \times (\hat{z}'\,\omega \times \vec{r}) = 0$$
$$\vec{F} - m\hat{z}'\,\omega \times \hat{y}'\,\omega r = \vec{F} + m\hat{x}'\,r\omega^2 = 0 \tag{5.14}$$

This says that there must be some force $\vec{F}$ pulling the mass toward you and holding it in place. Perhaps it is gravity, perhaps it is your hand, but it is some real force on m, and the equation dictates just what value that force must have. For the moment I am preserving the notation that x' is the coordinate in the rotating frame. (Of course $\hat{z}' = \hat{z}$.)

Example

▶ What does this equation say about a mass that is just sitting at rest in empty space, but where you are still rotating at constant $\vec{\omega} = \omega\hat{z}$ and looking at it move around you. Make it a distant star, and you are on the rotating Earth. Just to keep the picture simple, examine the specific time that the mass crosses your x'-axis. That way according to you the star's velocity will by in the $-\hat{y}'$ direction and its acceleration will (according to you) be the usual $v^2/r = r\omega^2$ toward the center of the circle. Use the same equation (5.12) as before, then

$$\vec{F} = 0, \qquad \dot{\vec{\omega}} = 0, \qquad \dot{\vec{r}} = -\hat{y}' \, r\omega, \qquad \ddot{\vec{r}} = -\hat{x}' \, r\omega^2$$
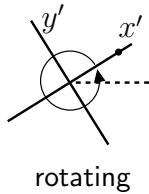
$$0 - 0 - 2m(\omega\hat{z}') \times (-\hat{y}' \, r\omega)$$
$$-m\omega\hat{z}' \times (\omega\hat{z}' \times \hat{x}' \, r) = m(-\hat{x}' \, r\omega^2) \qquad (5.15)$$
$$-2m\hat{x}' \, r\omega^2 + m\hat{x}' \, r\omega^2 = -m\hat{x}' r\omega^2$$

inertial   rotating

This equation says that it is a combination of the Coriolis force and the centrifugal force that keeps the star rotating about you (*according to you*) The first of this pair of examples, Eq. (5.14), should look very familiar. This second one, Eq. (5.15), is different from what you are used to, and it will take some time to get a feeling for the Coriolis force.

5.3 Coriolis Force

Picturing the effects of this term $(-2m\vec{\omega} \times \dot{\vec{r}})$ takes some effort because it doesn't behave the way that you're accustomed to, and it takes practice with a variety of examples to become used to it. The most interesting applications occur when you try to figure out what effects the Earth's rotation has on the motion of the atmosphere and the oceans, but that can wait.
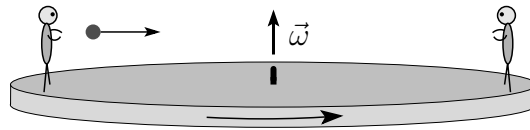
Fig. 5.4

Example

▶ When you want to throw a ball to someone else, you expect to throw it in the direction toward the person trying to catch it. What if however, the two of you are in a playground, standing on opposite sides of a spinning platform? If you throw it straight at your friend on the other side of the turntable, what direction will it go? What path will it take? The answer will depend very much on who's watching the ball, and someone standing on the ground will give a description very different from yours. If you are in the Rotor of problem 1.13 or in a rotating space station, you can ask the same question.
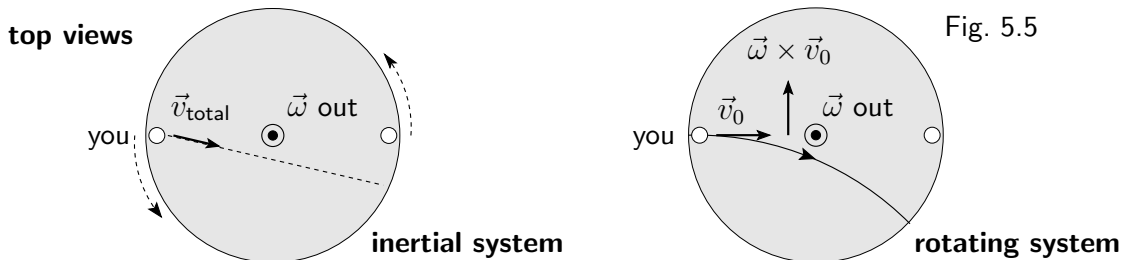
Fig. 5.5

The left picture is a top view as seen by someone standing still and watching the platform rotate — the inertial system. (Forget about gravity for this example; it just causes the ball to drop.) You throw the ball straight toward your friend, but a bystander will say that you are moving and that you gave the ball an initial sideways component of velocity $R\omega$. In that inertial system as pictured on

the left, the equations to describe the motion are simple: the ball moves in a straight line. This is true even if you choose not to throw the ball directly across, but instead give it your own extra sideways component v_1. The solution is trivial:

$$x = -R + v_0 t \qquad\qquad y = (v_1 - R\omega)t \qquad\qquad (5.16)$$

If $v_1 = 0$, the ball starts with a negative y-component of velocity because of the turntable's rotation. If $v_1 = R\omega$, the ball will (in this inertial system) go straight across the turntable, passing through its center. Does that mean that the ball will hit your friend? No. Remember that in this system that person is moving too, rotating around by an angle ωt in the time $t = 2R/v_0$. *Draw the picture.*

The right hand picture is a top view *in* the rotating system, so that the platform appears stationary as the world turns around you. The initial Coriolis force is $-2m$ times the vector $\vec{\omega} \times \vec{v}_0$ drawn in the second figure, and that pushes the ball to the thrower's right. It gives a trajectory something like the one drawn. This is a circumstance for which the equations just aren't enough to let you easily visualize what happens in this rotating system. Fortunately, there are videos readily available demonstrating exactly what happens. Here is a good example on YouTube, and a search for the word Coriolis on that site will provide others.

<div align="center">www.youtube.com/watch?v=LAX3ALdienQ</div>

How to compute the details of this trajectory? There are two ways. You can write $\vec{F} = m\vec{a}$ in the rotating system, using Eq. (5.12) and solving it. Or, you can solve it in the non-rotating coordinate system for which all those extra inertial forces are absent. Then do a coordinate transformation at the end to get the equations in the rotating system. The first way involves more algebra, but it does not involve any math different from what you've done in previous chapters. It is however, tedious. The second way is easier, but it involves methods that are less familiar. I choose the latter.

The equations that relate the coordinates of an object as seen in two coordinate systems, rotated by the angle θ with respect to each other are

$$\begin{aligned} x_R &= x_I \cos\theta + y_I \sin\theta \\ y_R &= -x_I \sin\theta + y_I \cos\theta \end{aligned}$$
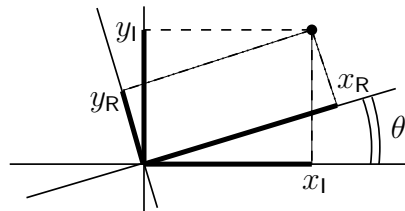
Fig. 5.6

(Check the special cases: $\theta = 0$ and $\theta = 90°$. At least.) The derivation of these equations is at Eq. (9.4), but you can probably do it yourself by drawing one line: the radial line from the origin to the dot. Designate ϕ as the angle it makes with the x_I axis. Then you need a couple of common trig identities, the sine and cosine of the difference of angles.

The equations for the motion in the inertial system are already here, in Eq. (5.16): (x_I and y_I). The rotating system makes an angle $\theta = \omega t$ with respect to the inertial system, so the coordinates in *that* system are

$$\begin{aligned} x_R &= (-R + v_0 t)\cos\omega t + (v_1 - R\omega)t\sin\omega t \\ y_R &= -(-R + v_0 t)\sin\omega t + (v_1 - R\omega)t\cos\omega t \end{aligned} \qquad (5.17)$$

This starts the ball moving at the left edge of the disk with initial velocity in the rotating system $v_0\hat{x} + v_1\hat{y}$. The picture on the right in Figure 5.5 has $v_1 = 0$, so the thrower there was aiming directly across. Near $t = 0$, do a series expansion, keeping only a couple of terms and assuming that the thrower aims straight across the turntable: $v_1 = 0$.

$$\begin{aligned} x &= -R\big(1 - \tfrac{1}{2}\omega^2 t^2\big) + v_0 t\big(1 - \tfrac{1}{2}\omega^2 t^2\big) - R\omega t\big(\omega t - \tfrac{1}{6}\omega^3 t^3\big) \\ y &= R\big(\omega t - \tfrac{1}{6}\omega^3 t^3\big) - v_0 t\big(\omega t - \tfrac{1}{6}\omega^3 t^3\big) - R\omega t\big(1 - \tfrac{1}{2}\omega^2 t^2\big) \end{aligned}$$

The constant terms: $x = -R$, $y = 0$. Yes, it starts at the left edge.

The "t" terms: $x = -R + v_0 t$, $y = 0$. No initial motion along y, but just a couple of paragraphs back didn't I say that it should have a sideways component of magnitude ωR? What's wrong? Go back and read that paragraph again.

The "t^2" terms: $x = -R + v_0 t - \frac{1}{2} R\omega^2 t^2$, $y = -v_0 \omega t^2$. It has an acceleration left and down in this picture. I can interpret the $R\omega^2$ term as coming from the centrifugal force; it is outward and has the right magnitude. The $v_0\omega$ term is from the Coriolis force; it is sideways (the cross product) and the size of this acceleration is $2v_0\omega$.

Here are two pictures of the equations Eq. (5.17). In the first picture, the lowest trajectory has $v_1 = 0$ so that it is thrown straight across (in the rotating frame). Each successive path has larger v_1 so that the ball is thrown laterally harder and harder. The 4th one comes close enough to the catcher that he should easily be able to field it.

In the second picture both v_0 and v_1 are varied, decreasing v_0 and increasing v_1 so that successive curves represent throws that come closer to being a loop. For the last throw the ball returns very near to the person who threw it, and without a boomerang.
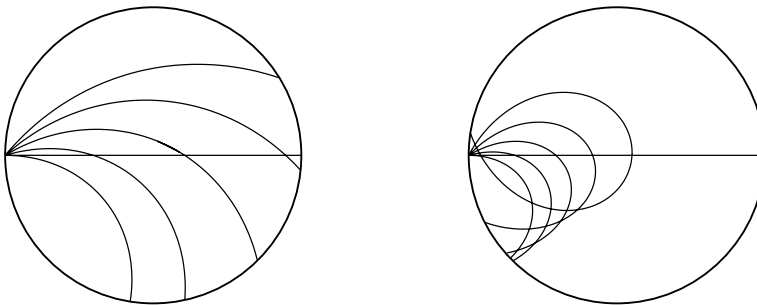
Fig. 5.7

Coriolis Force on Earth
On the Earth, $\vec{\omega}$ is pointing out of the North Pole. (The sun rises in the East.) If you are standing at the North Pole the vector $\vec{r}$ from the Earth's center to you is parallel to $\vec{\omega}$, and the term representing the centrifugal force on you has a factor $(\vec{\omega} \times \vec{r})$, so it is zero. If you now throw a rock horizontally, what does the Coriolis term do? $\vec{\omega}$ is up and $\dot{\vec{r}}$ is horizontal in front of you.

Fig. 5.8

rotating system   inertial system
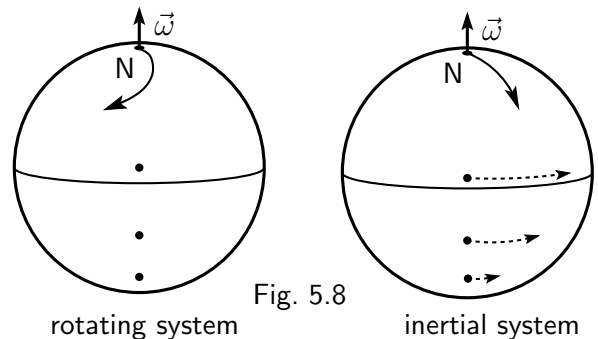
$$-2m\vec{\omega} \times \dot{\vec{r}} \qquad \text{is to your right.}$$

The rock that you threw will (in this rotating system) experience a Coriolis force to the right and that is the direction its trajectory will curve, as in the first picture. This is easy to understand if you just step off the Earth for a moment. No more Coriolis force, but the Earth is rotating counterclockwise under you. The rock will go straight, but the Earth turns left underneath it as in the second picture. It's the same thing.

This Coriolis force resembles the magnetic force on a moving charge, $q\vec{v} \times \vec{B}$. You shouldn't then be surprised when it twists motion around at right angles just as the magnetic field does for the motion of a charge.

From the previous paragraph you may get the impression that the Coriolis force is easy to understand intuitively. To reassure you that it is not, consider the case that you are standing on the equator and that you throw the same rock straight up. Now see what $-2m\vec{\omega} \times \dot{\vec{r}}$ becomes.

$$-2m\vec{\omega} \times \dot{\vec{r}} \propto -\hat{\text{north}} \times \hat{\text{up}} = +\hat{\text{west}}$$

and the rock will experience a Coriolis force that pushes it a little west of where
you threw it. This is not so obvious, even if you stand off the Earth to look at it.
It then gets even more complicated because when the rock comes back down, its
velocity is reversed — coming down instead of up. The Coriolis force is reversed
as it drops, pushing it east. Which one wins? Wait until Eqs. (5.22) and (5.26) to see.
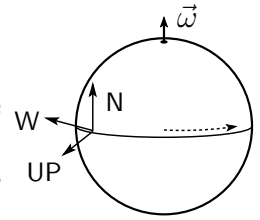
    If you play golf must you consider the Coriolis force? What if you're handling a large gun on
a battleship, aiming at something five or ten kilometers away? For both of these cases ignore the
curvature of the Earth; that isn't important in the first case and approximating the Earth as flat doesn't
even have a huge effect in the second case. Does it make sense to have a rotating, flat Earth? Yes. All
it means is that you're still dealing with distances that don't take you over the horizon. For now I will
ignore the centrifugal force term and air resistance, what's left is

$$m\ddot{\vec{r}} = m\vec{g} - 2m\vec{\omega} \times \dot{\vec{r}} \tag{5.18}$$

To see how big this new term is, plug in some numbers. Fire a rifle bullet at 300 m/s and compare the
Coriolis term to the gravitational term.

$$\frac{2m\omega v}{mg} = \frac{2 \cdot (2\pi/\text{day}) \cdot (300 \text{ m/s})}{10 \text{ m/s}^2} \cdot \frac{1 \text{ day}}{86\,400 \text{ s}} = \frac{4\pi \cdot 300}{864\,000} \approx 4 \times 10^{-3}$$

This correction is small even at this speed, and at low speeds such as walking, or at the speed of water
draining from a sink, it is more like 10^{-5} so the Coriolis effect at these small speeds is imperceptible,
despite all contrary urban legends and stage magic.

    Equation (5.18) is a linear, inhomogeneous, constant coefficient differential equation for $\vec{r}$, and
as such you can solve the homogeneous part using an exponential solution, just as in section 3.9. In
components this would look like

$$x = Ae^{\alpha t}, \qquad y = Be^{\alpha t}, \qquad z = Ce^{\alpha t} \qquad \text{or more simply,} \qquad \vec{r} = \vec{r}_0 e^{\alpha t}$$

I'm not going to do it this way. It is not that it's wrong. It is rather that it is complicated and hard to
interpret the results. I can take advantage of the fact that the Earth is rotating slowly, so the corrections
due to the ω terms are small. That suggests an iterative attack, for which I ignore the Coriolis term at
first and then go back and treat it as a correction. This is the method described in section 0.11 and
used both at the end of section 4.1 and in section 4.5.

    In the equation (5.18), the lowest order approximation to the equation is

$$\ddot{\vec{r}} = \vec{g} - 2\vec{\omega} \times \dot{\vec{r}} \longrightarrow \ddot{\vec{r}}_0 = \vec{g} \tag{5.19}$$

This is pretty easy. Take initial conditions to start from the origin with velocity $\vec{v}_0$, then

$$\dot{\vec{r}}_0 = \vec{g}t + \vec{C} = \vec{g}t + \vec{v}_0,$$
$$\text{and} \qquad \vec{r}_0 = \tfrac{1}{2}\vec{g}t^2 + \vec{C}t + \vec{D} = \tfrac{1}{2}\vec{g}t^2 + \vec{v}_0 t \tag{5.20}$$

The subscript zero on r indicates that this is the lowest order approximation to the final answer, ignoring
Coriolis effects. Use these equations as input to the right-hand side of Eq. (5.18), canceling the m's.
The improved approximation is $\vec{r}_1$, which now satisfies

$$\ddot{\vec{r}}_1 = \vec{g} - 2\vec{\omega} \times \dot{\vec{r}}_0 = \vec{g} - 2\vec{\omega} \times (\vec{g}t + \vec{v}_0) \tag{5.21}$$

Integrate this equation to get $\vec{r}_1$, including a couple of arbitrary constants, then reapply the initial conditions to evaluate these constants. $\vec{r}_1(0) = 0$ and $\dot{\vec{r}}_1(0) = \vec{v}_0$.

$$\vec{r}_1 = \tfrac{1}{2}\vec{g}t^2 - 2\vec{\omega} \times \left(\tfrac{1}{6}\vec{g}t^3 + \tfrac{1}{2}\vec{v}_0t^2\right) + \vec{E}t + \vec{F} = \vec{v}_0t + \tfrac{1}{2}\vec{g}t^2 - 2\vec{\omega} \times \left(\tfrac{1}{6}\vec{g}t^3 + \tfrac{1}{2}\vec{v}_0t^2\right) \tag{5.22}$$

If you need still higher order accuracy, put *this* into the right side of Eq. (5.18) and repeat the process to get a still better result: $\ddot{\vec{r}}_2 = \vec{g} - 2\vec{\omega} \times \dot{\vec{r}}_1$. If this looks familiar then you've probably studied section 4.5 about curve balls. If not, then you may want to look back at it.

To understand this result, try lots of special cases:

1. At the North Pole and firing directly up, the $\vec{g}$ and $\vec{v}_0$ terms in Eq. (5.22) are along the vector $\vec{\omega}$, so the cross products vanish and the Coriolis force contributes nothing. What goes straight up comes straight down.

2. At the equator aiming due North the $\vec{\omega} \times \vec{v}_0$ term is zero, but the term in $-\vec{\omega} \times \vec{g}$ points East.

$$\vec{r}_1 = \vec{v}_0t + \tfrac{1}{2}\vec{g}t^2 - \tfrac{1}{3}\vec{\omega} \times \vec{g}t^3$$

These three terms are respectively North, Down, and East. Why East? Step off the Earth for a moment and watch what happens. The projectile moves North and starts to drop. As it drops, it is moving closer to the axis of rotation, *but* it keeps the larger Eastern component of velocity that it had when it started at the larger distance from the axis. That means that it is getting ahead of its surroundings and that means that it drifts East.

3. At the equator aiming due East the $-\vec{\omega} \times \vec{v}_0$ term is Up and the $-\vec{\omega} \times \vec{g}$ term still points East.

Eq. (5.22) is $\vec{r}_1 = \vec{v}_0t + \tfrac{1}{2}\vec{g}t^2 - \tfrac{1}{3}\vec{\omega} \times \vec{g}t^3 - \vec{\omega} \times \vec{v}_0t^2$

These terms are respectively East, Down, East, and Up. The fourth one is new, and it is not easy to track down a simple interpretation for it. It does come out of the equations though.

4. With the same equation but at the equator firing straight up $\vec{\omega} \times \vec{g}$ is West, and $\vec{\omega} \times \vec{v}_0$ is East. To find where the projectile hits the ground, note that neither of these two terms affect the amount of time that the projectile stays in the air, because East and West are not up and down. The time aloft is governed solely by the old terms in $\vec{r}_1(t)$: $\tfrac{1}{2}\vec{g}t^2 + \vec{v}_0t = 0$. At the value of t that this determines, compute the position of the projectile, $\vec{r}_1$, from Eq. (5.22).

$$-\frac{1}{2}gt^2 + v_0t = 0 \implies t_{\mathrm{f}} = \frac{2v_0}{g} \qquad \text{then the corrected term } \vec{r}_1 \text{ back at ground level is}$$

$$\vec{r}_1(t_{\mathrm{f}}) = -\vec{\omega} \times \left(\tfrac{1}{3}\vec{g}\frac{2v_0}{g} + \vec{v}_0\right)\left(\frac{2v_0}{g}\right)^2$$

$$= -\left(-\tfrac{2}{3} + 1\right)\vec{\omega} \times \vec{v}_0\left(\frac{2v_0}{g}\right)^2 = \frac{4\omega v_0^3}{3g^2} \text{ West} \tag{5.23}$$

For a speed of 300 m/s this is

$$\frac{4 \cdot (2\pi/\text{day}) \cdot (300 \text{ m/s})^3}{3 \cdot (10 \text{ m/s}^2)^2} \cdot \frac{1 \text{ day}}{86\,400 \text{ s}} = 26 \text{ m} \tag{5.24}$$

At least this answers the question raised a few paragraphs back: Which part of the Coriolis force wins? The one while the rock is going up or the one when it's going down? Well, which is it?

Does this result of $26\,\mathrm{m}$ seem to be much larger than you would expect? To see how reasonable this is, ask how much time the bullet is it in the air and how high does it go? Answer:

$$t_f = \frac{2 \cdot v_0}{g} = \frac{2 \cdot 300}{10} = 60\,\text{seconds}, \qquad \text{max height} = \frac{v_0^2}{2g} = \frac{2 \cdot 300^2}{2 \cdot 10} = 4.5\,\text{km} \tag{5.25}$$

Will a bullet go that high? No, not even close. Air resistance is *very* important here, and it cuts these results by a big factor. Still, the Coriolis effect is significant.

To see how much air resistance can affect these results, go back to problem 2.30. For a 30 caliber bullet (not tumbling) the terminal speed is about $100\,\mathrm{m/s}$, and if it is fired up at 300 or $400\,\mathrm{m/s}$ it comes to the maximum height in under 15 seconds. You can make a very crude estimate of the Coriolis effect in this case by assuming the extreme case that it reaches terminal speed very fast. Then replace the $300\,\mathrm{m/s}$ in Eq. (5.24) by $100\,\mathrm{m/s}$, changing the result from $26\,\mathrm{m}$ to $1\,\mathrm{m}$. Reality is closer to the latter number than the former, but working out the details requires a little effort. A simpler version of this calculation appears in problem 5.13.

What is the trajectory of the bullet you just fired? It is no longer a straight line up and down, but something described by Eq. (5.22). For this, use a coordinate system with x East, y North, and z Up—like the coordinate system sketched below, but on the equator. The motion described by this equation is then all in the x-z plane.

$$\vec{\omega} = \omega\hat{y}, \qquad \vec{g} = -g\hat{z}, \qquad \vec{v}_0 = v_0\hat{z}$$
$$\vec{r}_1 = x_1\hat{x} + z_1\hat{z} = -\tfrac{1}{2}g\hat{z}t^2 - \omega\hat{y} \times \left(-\tfrac{1}{3}g\hat{z}t^3 + v_0\hat{z}t^2\right) + v_0\hat{z}t$$
$$x_1 = \tfrac{1}{3}\omega g t^3 - \omega v_0 t^2, \qquad z_1 = -\tfrac{1}{2}g t^2 + v_0 t \tag{5.26}$$

These equations, $x_1(t)$ and $z_1(t)$ describe the x-z coordinates in terms of the time parameter, and this graph shows the path of the bullet. The horizontal scale on this graph is greatly exaggerated so that you can see the effect.

5. For another example, fire a projectile and see where it lands, but this time start from an arbitrary latitude. Again use coordinates x-East, y-North, $\hat{z} = \hat{x} \times \hat{y}$. The latitude is λ, so
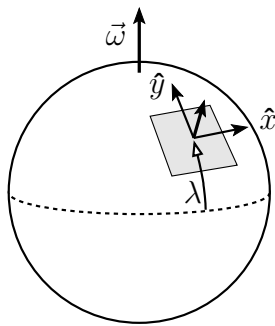
$$\vec{\omega} = \omega(\hat{y}\cos\lambda + \hat{z}\sin\lambda) \tag{5.27}$$

The initial conditions are $\dot{x}_0 = 0$ and the other initial velocity components non-zero. Eq. (5.22) is

$$\vec{r}_1 = \tfrac{1}{2}\vec{g}t^2 + \vec{v}_0 t - 2\omega(\hat{y}\cos\lambda + \hat{z}\sin\lambda) \times \left(\tfrac{1}{6}\vec{g}t^3 + \tfrac{1}{2}\vec{v}_0 t^2\right) \tag{5.28}$$

$\vec{g} = -g\hat{z}$, and $\vec{v}_0 = \dot{y}_0\hat{y} + \dot{z}_0\hat{z}$, so this is

$$\vec{r}_1 = \tfrac{1}{2}\vec{g}t^2 + \vec{v}_0 t + \tfrac{1}{3}\omega g t^3 \hat{x}\cos\lambda + \omega t^2\hat{x}(\dot{y}_0\sin\lambda - \dot{z}_0\cos\lambda)$$

which says that the y and z-coordinates are not affected by ω (to this approximation), only x is. Where will it land? That is determined by z.

$$z_1(t) = z_0(t) = 0 = \dot{z}_0 t - g t^2/2 \implies t_f = 2\dot{z}_0/g \qquad \text{then} \qquad y_1 = \dot{y}_0 t_f = 2\dot{z}_0\dot{y}_0/g \tag{5.29}$$

$$\text{and} \qquad x_1 = \frac{1}{3}g\omega \cos \lambda \, (2\dot{z}_0/g)^3 + \omega \big(\dot{y}_0 \sin \lambda - \dot{z}_0 \cos \lambda \big)(2\dot{z}_0/g)^2$$

Simplify this. Group terms and express the initial velocity in terms of speed v_0 and the angle α above the horizontal.

$$x_1 = -\frac{4}{3}\omega \frac{\dot{z}_0^3}{g^2} \cos \lambda + 4\omega \sin \lambda \frac{\dot{y}_0 \dot{z}_0^2}{g^2} = \omega \frac{v_0^3}{g^2}\left[-\frac{4}{3}\sin^3 \alpha \cos \lambda + 4\sin^2 \alpha \cos \alpha \sin \lambda \right] \qquad (5.30)$$

How big is this for a golf ball? Choose the latitude of Scotland, $\lambda = 55°$, with spherical coordinate $\theta = 90° - \lambda$, and hit the golf ball $\alpha = 45°$ above the horizontal. The average world champion golfer can hit the ball about 300 meters, so from Eq. (5.29)

$$y_1 = \frac{2\dot{z}_0 \dot{y}_0}{g} = \frac{2v_0^2 \cos \alpha \sin \alpha}{g} = \frac{v_0^2 \sin 2\alpha}{g} = 300\,\text{m} \implies v_0 = \sqrt{3000} = 55\,\text{m/s}$$

Then the deflection is

$$x_1 = \frac{2\pi}{\text{day}} \frac{(55\,\text{m/s})^3}{(10\,\text{m/s}^2)^2}\left[-\frac{4}{3}\cdot\frac{1}{2^{3/2}}\cdot 0.57 + 4\cdot\frac{1}{2^{3/2}}\cdot 0.82 \right]\left(\frac{1\,\text{day}}{86\,400\,\text{s}} \right) = 0.11\,\text{m} = 11\,\text{cm}$$

Compared to all the other variables, especially wind, this is not much, and I suppose that an expert golfer will have even this effect trained into the muscles. What happens when the tournament is in Southern Australia?

Take the same equations and fire a large artillery shell in the same way, but fast enough to go $30\,\text{km}$ instead of $300\,\text{m}$. Everything else is the same, so all I have to do is see how the result scales with the range. The distance is 100 times greater, so Eq. (5.29) for y_1 says the speed is larger by $100^{1/2}$. Then Eq. (5.30) says that the deviation, the value of x, varies as speed cubed, so $100^{3/2} = 1000$ and the deviation is 110 meters, which is not so small any more.

The result for x_1 is positive, meaning that the deflection is toward the East. Stand off the Earth a moment and try to visualize the motion in an inertial coordinate system. Is that the direction it should move?

Example

▶ For us, the most important application of the Coriolis force is to weather, and a large part about understanding weather comes not from $\vec{F} = m\vec{a}$, but from $\vec{F} = 0$. The accelerations are commonly very small, and the forces within the atmosphere almost cancel out. Are there exceptions? Yes! Start with tornadoes and work down. Still, this is a place to start.

If there is a low pressure or high pressure region somewhere, and there always is, why does it rotate and why does it rotate in the direction that it does? Look at the forces involved in the case of a Low. A surrounding higher pressure region will exert a larger force on the adjacent air than does the low pressure region — no surprise there. You would expect then that air would be pushed by the surrounding high pressure toward the inner region of lower pressure and that everything would quickly reach an equilibrium. In the absence of the Earth's rotation, that is exactly what would happen,

Fig. 5.9

But then there's the Coriolis force, $-2m\vec{\omega}\times\vec{v}$. In the Northern hemisphere, the vertical component of $\vec{\omega}$ is up. If the surrounding air pressure accelerates some air in toward the center of the low pressure region, then $-\vec{\omega}\times\vec{v}$ is in the direction $-\textbf{up}\times\textbf{in}$, and that is to the right. The air will start to circulate counter-clockwise around the center of the low pressure region. *Now* there is a Coriolis force from the velocity of this counter-clockwise circulation. The $-\vec{\omega}\times\vec{v}$ term from *this*
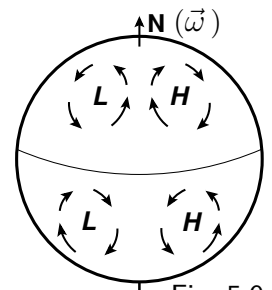
is in the direction $-\textbf{up} \times \textbf{right}$, and that is outward, away from the center of the low pressure region. When the circulation is big enough then the force from the pressure gradient and the Coriolis force from the wind around the center will tend to cancel, leaving a steady counter-clockwise circulation around the low, called "cyclonic" flow. With a high pressure region, the circulation reverses, and the result is "anti-cyclonic" flow — clockwise in the Northern hemisphere. This balancing of the Coriolis force and the pressure gradient forms a pardigm in the study of the Earth's atmosphere and oceans. Why does the Gulf Stream exist? Why are there hurricanes? This balance of forces is the first place to look.

The preceding paragraph is a bare sketch of the very complex reality. The atmosphere is not in equilibrium. There's friction. Everything depends on altitude. And probably a dozen other complications. It's a full profession to sort it all out: "Geophysical Fluid Dynamics".

A Useful Analogy
There is a familiar analog to the Coriolis Force: Magnetism.

$$\text{Coriolis force:} \quad -2m\vec{\omega} \times \vec{v}, \qquad\qquad \text{Magnetic force:} \quad q\vec{v} \times \vec{B}$$

Except for an exchange of symbols between $\vec{B}$ and $\vec{\omega}$ and an extra factor of $q/2m$, they are the same. This means that whatever intuition that you have about the motion of charges in a magnetic field can now be carried over directly to intuition about the motion of of masses in a rotating system, with $\vec{\omega}$ replacing $\vec{B}$. Go back to the several examples in this section and translate them into this language. Does it help? Only you can answer that.

5.4 Centrifugal Force
This is easier to visualize than the Coriolis force; It behaves more like things you're accustomed to. Return to the general equation (5.12) and look closely at the other inertial force term, $-m\vec{\omega} \times (\vec{\omega} \times \vec{r})$, and now ignore the Coriolis term. Use $\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})$

$$-m\vec{\omega} \times (\vec{\omega} \times \vec{r}) = -m\vec{\omega}(\omega \cdot \vec{r}) + m\vec{r}\omega^2$$

Choose the coordinate system so that $\hat{z}$ is along $\vec{\omega}$, and this is

$$-\hat{z}m\omega(\omega z) + m(x\hat{x} + y\hat{y} + z\hat{z})\omega^2 = m\omega^2(x\hat{x} + y\hat{y})$$

It points away from the axis of rotation. In this rotating system an extra force then appears, pushing away from the axis.

There is a problem in notation here. The symbol $\vec{r}$ is overused. It is the vector from the origin in spherical coordinates. It is the vector from the axis in cylindrical coordinates. This doesn't usually cause confusion, but here I need both in the same discussion. I'll resolve the dispute by using a special notation for the vector perpendicular to the axis as in cylindrical coordinates:

$$\vec{r}_{\perp} = x\hat{x} + y\hat{y} \qquad \text{then} \qquad \vec{r} = z\hat{z} + \vec{r}_{\perp} \tag{5.31}$$

$\vec{r}_{\perp}$ is perpendicular to the z-axis, letting $\vec{r}$ remain as the vector from the origin. The centrifugal force term is then $m\omega^2\vec{r}_{\perp}$. This is not a commonly used notation, so don't fret if no one else uses it.

When in a car turning a corner you can take the perfectly defensible point of view that you are the center of the universe and that you are not moving. This is certainly not an inertial system, so you must include the centrifugal force term to make sense of the world. $\ddot{\vec{r}} = 0$, so

$$\vec{F} - m\vec{\omega} \times (\vec{\omega} \times \vec{r}) = m\ddot{\vec{r}} = 0, \qquad \text{or} \qquad \vec{F} + m\omega^2\vec{r}_{\perp} = 0$$

The seat cushion friction pushes you in toward the center of rotation and the centrifugal force pushes you out, giving a total force of zero. You then have zero acceleration with respect to yourself.

When standing still on the surface of the Earth, what effect does the centrifugal force have? At the North and South Poles nothing, because $\vec{r}_\perp = 0$ there. At the equator, $\omega^2 \vec{r}_\perp$ is away from the Earth and of magnitude $\omega^2 R = (2\pi/1\,\text{day})^2 (6400\,\text{km}) = 0.034\,\text{m/s}^2$. This gives you an apparent decrease in weight of about 0.3%.

It is not quite as simple as the preceding paragraph implies, because the Earth isn't exactly spherical. It has an equatorial bulge with about $41\,\text{km}$ greater diameter at the equator than through the poles. This changes the gravitational field of the Earth from the simple $1/r^2$ of a spherical Earth to a more complicated form, and that fact provides a significant change—comparable to the 0.3% mentioned above. You haven't computed your apparent weight without including both effects.

Example

▶ Put water into a bucket, suspend it from a rope, and spin the rope about its vertical axis so that the bucket spins in the same way. The surface of the water will come to an equilibrium in a concave shape; what is that shape?
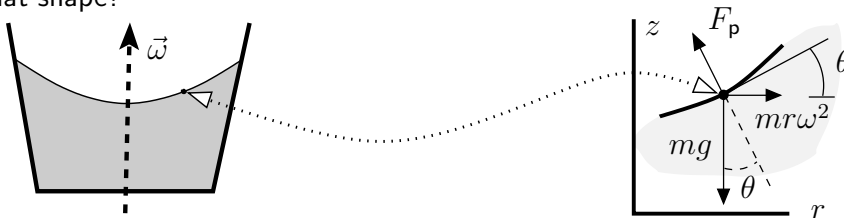
In the rotating coordinate system, the forces on a molecule at the surface are

1. centrifugal,   2. gravity,   and   3. the surrounding molecules,

and their sum is zero. These forces have respective magnitudes $mr\omega^2$, mg, and an unknown F_p. All that you know about $\vec{F}_\text{p}$ is that its direction is perpendicular to the surface, but that is enough. Temporarily and to save clutter, this r is the cylindrical coordinate. The z-coordinate of the surface is a function of this radial coordinate r, and from the picture its slope is

$$\tan\theta = \frac{dz}{dr} \quad \text{with} \quad \tan\theta = \frac{mr\omega^2}{mg}$$

Combine these to get

$$\frac{dz}{dr} = \frac{r\omega^2}{g} \longrightarrow z = \frac{r^2\omega^2}{2g} + C \tag{5.32}$$

This is a paraboloid. The bigger the spin rate, the steeper the surface is. This result is the foundation of spin casting, a method for constructing very large mirrors for astronomical telescopes (among other things). In that application, the content of the bucket becomes some high quality molten glass. The resulting parabolic shape forms an ideal telescope mirror.

5.5 Shape of the Earth

In an Earth-centered coordinate system the effective force is the gravitational force plus the centrifugal force. What does this do to the shape of the Earth? Think of the oceans first, because it is easier to see that the ocean surface should be an equipotential. If it isn't then the higher parts would slide over to the lower parts at a lower potential energy and bring the surface back to an equipotential. What about the tides? That brings in the Moon, and one thing at a time if you don't mind.

What is the potential energy of a mass m in this system? The force on a mass m at the distance r from the center of a spherical Earth is

$$\vec{F} = F_r \hat{r} \qquad \text{where} \qquad F_r(r) = -GMm/r^2$$

M is the mass of the Earth. The equation (2.22) relates force to the potential energy; this is

$$F_r = -dU/dr \qquad \text{which implies} \qquad U(r) = -GMm/r$$

The potential energy associated with the centrifugal force follows from the same defining equation, but here the only component is the cylindrical radius $r_\perp$, not the spherical r.

$$\vec{F} = +m\omega^2 \vec{r}_\perp, \quad \text{or} \quad F_{r_\perp} = +m\omega^2 r_\perp = -\frac{d}{dr_\perp} U(r_\perp) \longrightarrow U = -m\omega^2 r_\perp^2/2$$

The total potential energy for a mass m near the Earth's surface is the sum of these two:

$$U = -\frac{GMm}{r} - \frac{1}{2} m\omega^2 r_\perp^2 \qquad (5.33)$$

But wait, the whole point of this section is the shape of the Earth, so how can I assume that the Earth is spherical in getting that first term? You can't really, but as a first approximation to figuring it out I'll try the simplifying assumption that it is almost spherical and then return to see if that is good enough. (It is pretty good first try, but *not* good enough; the result will be off by a factor of about two.)

The centrifugal potential energy is like a harmonic oscillator energy turned upside down, and this centrifugal term is much smaller than the gravitational energy, so the Earth is nearly spherical. Its shape is $r = R + \epsilon$, where R is some average radius that can be specified later. Using spherical coordinates, as in Figure 0.3, ϵ is a function of the angle θ alone, measured from the North Pole. The distance to the axis is $r_\perp = r \sin\theta$.

$$U = -\frac{GMm}{R+\epsilon} - \frac{1}{2} m\omega^2 (R+\epsilon)^2 \sin^2\theta$$
$$= -\frac{GMm}{R}\left(1 - \frac{\epsilon}{R}\right) - \frac{1}{2} m\omega^2 (R^2 + 2R\epsilon) \sin^2\theta$$
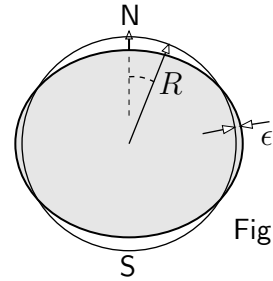
Fig. 5.10

This is a binomial expansion on each term, keeping only the first order in ϵ. For an equipotential, set this to a constant.

$$U_0 = -\frac{GMm}{R} - \frac{1}{2} m\omega^2 R^2 \sin^2\theta + \frac{GMm}{R}\frac{\epsilon}{R} - m\omega^2 R\epsilon \sin^2\theta \qquad (5.34)$$

The fourth term is much less than the third term. To see this, simply interpret the meaning of the two coefficients of ϵ:

$$\frac{GMm}{R^2} \quad \text{is the gravitational force on } m, \text{ and } \quad mR\omega^2 \quad \text{is the centrifugal force.}$$

The latter is much less than the former ($\approx 0.3\%$), so I can neglect the $m\omega^2 R\epsilon \sin^2\theta$ term in Eq. (5.34). This leaves

$$U_0 = -\frac{GMm}{R} - \frac{1}{2} m\omega^2 R^2 \sin^2\theta + \frac{GMm}{R}\frac{\epsilon}{R}$$
$$\text{or} \quad \epsilon = \frac{R^2}{GMm}\left[U_0 + \frac{GMm}{R} + \frac{1}{2} m\omega^2 R^2 \sin^2\theta\right]$$

The last term is the only θ-dependent part of ϵ, so pull it out for examination.

$$\epsilon = \cdots + \frac{R^2}{GMm}\frac{1}{2}m\omega^2 R^2 \sin^2\theta = \cdots + \frac{1}{2}\frac{m\omega^2 R^2}{GMm/R^2}\sin^2\theta$$

Pull out a factor of R from the top in order to make the final quotient dimensionless. Also cancel the m's:

$$\epsilon = \cdots + \frac{R}{2}\frac{\omega^2 R}{GM/R^2}\sin^2\theta \tag{5.35}$$

The coefficient of $\sin^2\theta$ is $R/2$ times the ratio of the centrifugal acceleration at the equator to g.

Compute the difference of the equatorial and polar diameters (2ϵ). It is

$$\frac{R^2\omega^2}{g} = \frac{(6400\,\text{km})^2(2\pi/\text{day})^2}{10\,\text{m/s}^2}\cdot\left(\frac{1\,\text{day}}{86\,400\text{s}}\right)^2\cdot\left(\frac{1000\,\text{m}}{1\,\text{km}}\right) = 22\,\text{km}$$

How accurate it this $22\,\text{km}$ result? Not very, as the measured value of the difference in the diameters is closer to twice this. The major error in this calculation is in the assumption that the gravitational field is spherically symmetric. It is a plausible assumption, but it turns out to introduce an error that is about as big as the effect you're calculating, so it is not good enough. Working out the more correct version would take too much time and effort for this single result, so I'll leave it to other sources.

Is it necessary to evaluate the constant C? No. There is however one manipulation that relates this expression to something that you will encounter often in electromagnetism and even elsewhere in this chapter. It's simply a different way to write ϵ, and though not really needed here, I'll describe it anyway. The equatorial bulge is $\epsilon(\theta) \propto \sin^2\theta$. This measures the deviation from a sphere by taking it to be zero at the poles. That's arbitrary, but convenient. Another way to set the zero point is to make the average value over the surface of the Earth zero. That will make the redefined ϵ negative at the two poles and positive at the equator. The result is

$$\epsilon'(\theta) = -\frac{R}{3}\frac{\omega^2 R}{GM/R^2}\left(\tfrac{3}{2}\cos^2\theta - \tfrac{1}{2}\right) \tag{5.36}$$

The last expression that appears in parentheses is a combination that you will soon come to know and love (or not). It is a Legendre polynomial, and these sort of polynomials appear often whenever you are looking at any problems involving potentials. And lots of other places too. You can at least see that because $\cos^2\theta = 1 - \sin^2\theta$, this expression gives the same coefficient as in Eq. (5.35). Use spherical coordinates as in Figure 0.7, and integrate this over the whole Earth and the result is zero.

5.6 Tides

Why are there tides? The Moon's pull on the side of the Earth nearer the Moon is stronger than its pull on the farther side of Earth. Now, why are there *two* tides? Air is much easier to push around than water, so are there air tides? (Yes.) Are there Earth tides? (Also yes, but smaller.) For now, look at the ocean.

There are several steps in understanding this. The first is to ask what determines the shape of the ocean's surface. That is the same question that I asked when determining the shape of the rotating Earth. The surface of the ocean is an equipotential. This time however the combining forces are not gravity and centrifugal forces, but the gravity of the Earth and of the Moon and the inertial force from the acceleration of the Earth. Take into account the presence of the Moon but ignore the distortion of

the Earth caused by the Earth's rotation. The potential energies from the Earth and from the Moon are the first term in Eq. (5.33).

$$-\frac{GMm}{r} \quad \text{and} \quad -\frac{GM'm}{r'}$$

Fig. 5.11
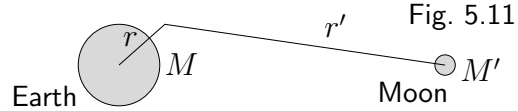
Why am I carrying along the factor m for the mass of a drop of ocean water? No good reason. The gravitational potential, as distinct from the gravitational potential *energy* is the potential energy *per mass*. It's easier to deal with $V = -GM/r$ and $-GM'/r'$. Near the Earth's surface, where you can use $U = mgh$, the gravitational potential is $V = gh$. And why is all of this in a chapter on non-inertial systems? That will be the final step, yet to come. Remember that the Moon pulls on the Earth, accelerating it. What will *that* do?

Find the total gravitational potential near the Earth's surface by adding the potentials from both Earth and Moon. Use R for the distance between the centers of the two bodies.

$$V = -\frac{GM}{r} - \frac{GM'}{r'}$$
$$= -\frac{GM}{r} - \frac{GM'}{\sqrt{r^2 + R^2 - 2rR\cos\theta}}$$

Fig. 5.12
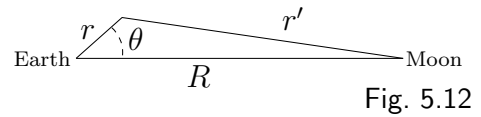
The last line used the law of cosines (problem 0.26) in order to express all the information in terms of coordinates centered on the Earth.

Now recall the geometry of the system. $r = 6400$ km and $R = 380\,000$ km. This suggests a series expansion because $r \ll R$.

$$V = -\frac{GM}{r} - \frac{GM'}{R\sqrt{1 + (r^2/R^2) - 2(r/R)\cos\theta}}$$

Use the binomial expansion with $n = -1/2$ (and you *will* need terms out to x^2).

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \cdots = 1 - \frac{1}{2}x + \frac{3}{8}x^2 + \cdots$$

$$V = -\frac{GM}{r} - \frac{GM'}{R}\left[1 - \frac{1}{2}\left(\frac{r^2}{R^2} - 2\frac{r}{R}\cos\theta\right) + \frac{3}{8}\left(\frac{r^2}{R^2} - 2\frac{r}{R}\cos\theta\right)^2 + \cdots\right]$$

$$= -\frac{GM}{r} - \frac{GM'}{R}\left[1 + \frac{r}{R}\cos\theta + \frac{r^2}{R^2}\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) + \cdots\right] \qquad (5.37)$$

In the last step I collected all the terms out to the order r^2/R^2, dropping the higher powers of r/R. Again, Legendre polynomials show up.

If you drop *all* the terms in r/R because it is small, the equipotential is

$$V = -\frac{GM}{r} - \frac{GM'}{R} = \text{a constant}$$

and that says that r is constant, a sphere, and to the lowest order that is correct. Call it R_E. The ocean surface is only approximately a sphere, so try including the next term.

$$-\frac{GM}{r} - \frac{GM'}{R}\left[1 + \frac{r}{R}\cos\theta\right] = \text{a constant} \qquad (5.38)$$

The radius r is approximately R_E, but it's the correction that is important here. Let δ be the tide's height above (or below) mean sea level, then $r = R_E + \delta$ and

(a) $\qquad -\dfrac{GM}{R_E + \delta} - \dfrac{GM'}{R}\left[1 + \dfrac{R_E + \delta}{R}\cos\theta\right] = C$

(b) $\qquad -\dfrac{GM}{R_E}\left[1 - \dfrac{\delta}{R_E}\right] - \dfrac{GM'}{R}\left[1 + \dfrac{R_E + \delta}{R}\cos\theta\right] = C$

(c) $\qquad \dfrac{GM}{R_E^2}\delta = C + \dfrac{GM}{R_E} + \dfrac{GM'}{R} + \dfrac{GM'R_E}{R^2}\cos\theta$

(d) $\qquad \delta = \dfrac{R_E^2}{GM}\cdot\dfrac{GM'R_E}{R^2}\cos\theta = \dfrac{M'}{M}\dfrac{R_E^3}{R^2}\cos\theta$ $\hspace{3cm}$ (5.39)

$\qquad\qquad \delta = \dfrac{7.35\times 10^{22}\ \text{kg}}{5.98\times 10^{24}\ \text{kg}}\cdot\dfrac{\left(6.37\times 10^6\ \text{m}\right)^3}{\left(3.84\times 10^8\ \text{m}\right)^2}\cos\theta = 21.5\ \text{m}\ \cos\theta$

(a) is Eq. (5.38) written in terms of δ.
(b) used the binomial expansion on the first term.
(c) solved for δ while dropping $\delta/R^2 \ll \delta/R_E^2$.
(d) that the mean sea level is at $\delta = 0$ determines C.*

The equation (5.39) is most positive at $\theta = 0$ and most negative at $\theta = \pi$. That is one (very large) high tide underneath the Moon and one very low tide on the opposite side of the Earth. The range from low tide to high tide is $2\times 21.5 = 43$ m. There are two difficulties with this approximation. First it has only one high tide per day, and second the height of that tide is about 22 meters, making life along the ocean coasts very wet. About 3/4 of Florida would be underwater at high tide, as would the entirety of some nations.
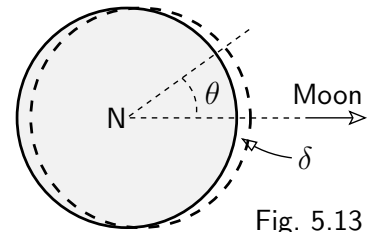
Fig. 5.13

What is missing? It is the fact that the Earth is accelerating toward the Moon, and that it is not an inertial system (even ignoring its daily rotation). This figure looks down from above the North Pole with the Moon on the right.

What is the Earth's acceleration as caused by the Moon?

$$a = F/M = GMM'/MR^2 = GM'/R^2 \approx 3\times 10^{-5}\ \text{m/s}^2$$

It doesn't sound like much, but it is enough to change everything. Call this acceleration a_0 and apply Eq. (5.5), to transform to the accelerated (Earth) coordinate system.

$$F_x - ma_0 = m\frac{d^2 x'}{dt^2}$$

This says that a mass m sitting on the Earth or in its oceans feels an inertial force ma_0 down, and here "down" means away from the Moon. In this system a_0 behaves like an extra, uniform gravitational field. What is the potential energy for that? mgh, or in this case $ma_0 h$. The "height" h is the distance toward the Moon from your origin, the center of the Earth. The "inertial energy" is then

$$U_{\text{inertial}} = ma_0 h = ma_0 r\cos\theta = m\frac{GM'}{R^2}r\cos\theta, \quad \text{so} \quad V_{\text{inertial}} = \frac{U}{m} = \frac{GM'}{R^2}r\cos\theta$$

* *How so?* What is the average value of $\cos\theta$ over the whole sphere? Use Eq. (0.18).

Add this to Eq. (5.37) and that troublesome $\cos\theta$ term precisely cancels, leaving

$$V_{\text{total}} = -\frac{GM}{r} - \frac{GM'}{R}\left[1 + \frac{r^2}{R^2}\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) + \cdots\right]$$

Already you can see that the problem of the single high tide is gone, because where before there was a $\cos\theta$ there is now a $\cos^2\theta$ and that is positive on both sides of the Earth. Now solve the size of the tides in exactly the same way as led to Eq. (5.39). (That's up to you to do.)
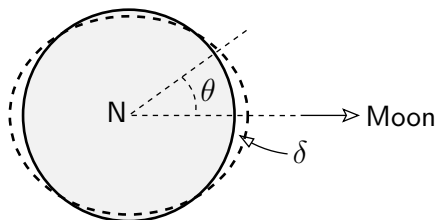
$$\delta = \frac{M'}{M}\frac{R_E^4}{R^3}\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) \qquad (5.40)$$

This is positive at $\theta = 0$ and $\theta = \pi$—toward and away from the Moon. It is negative halfway between, the low tides. How big is it? As compared to the previous result, Eq. (5.39), it is smaller by a factor $R_E/R = 6400/380000 = 0.017$, bringing it down to a size comparable to what you see at the shore, though at $21.5 \times 0.017 = 0.36\,\text{m}$ it seems a bit *too* small. (But keep going. There's a lot more to come.) The tidal range (low to high tide) is $\frac{3}{2} \times 0.36 = 0.54\,\text{m}$. Why $\frac{3}{2}$? Evaluate the quantity in parentheses in Eq. (5.40) at $\theta = 0$ and at $\theta = \pi/2$. The difference is $1 - (-\frac{1}{2}) = \frac{3}{2}$.

Notice that the tidal effect varies as the inverse cube of the distance to the Moon ($\propto 1/R^3$). That's because the gravitational force from the Moon varies as $1/r^2$, so the *change* in the force over the diameter of the Earth varies as the derivative of this with respect to distance; hence $1/r^3$. The Moon is slowly moving away from the Earth, and some billion or more years ago it was only half its current distance. This $1/r^3$ observation implies that the tides wer eight times higher then.

A simple point: The Moon goes around the Earth in a month. Why are there daily tides? The Earth rotates, and if the bulge stays aligned with the Moon, the Earth's surface rotates underneath the bulge to give the two daily tides.

Does the sun have an effect? It is much farther away, but it is also much more massive. Let $M_\odot$ be the sun's mass and $R_\odot$ the Earth-sun distance, then the ratio of the tidal effects is, from Eq. (5.40)

$$\frac{\text{solar tide}}{\text{lunar tide}} = \frac{(M_\odot/M)(R_E^4/R_\odot^3)}{(M'/M)(R_E^4/R^3)} = \frac{M_\odot}{M'}\frac{R^3}{R_\odot^3} = \frac{2.0 \times 10^{30}\,\text{kg}}{7.4 \times 10^{22}\,\text{kg}}\left(\frac{380\,000\,\text{km}}{1.5 \times 10^8\,\text{km}}\right)^3 = 0.44$$

The solar tide is almost half as large as the lunar tide. This means that when the Sun and Moon are roughly lined up near new Moon and full Moon the effects add, and when they are at right angles as seen from Earth, they try to cancel. The trigonometric factor in Eq. (5.40) is $+1$ at $\theta = 0$ and $-1/2$ at $\theta = 90°$. The high tide when they are adding is then 1.44 times the Moon's effect alone. The high tide when the are subtracting is a factor of $1-0.22=0.78$. The ratio of the highest high tide to lowest high tide is then about $1.44/0.78 = 1.85$.

Is this the whole story of tides? Far from it. There's the Earth's rotation. Then the Moon's orbit is not above the equator nor is its orbital distance constant. Nor is the Earth's orbital distance from the Sun constant and the Earth's rotational axis is tilted with respect to the plane of its orbit. Then friction. And don't forget continents. And maybe resonant interactions with natural ocean sloshing (e.g. The Bay of Fundy, which has a tidal range of 15 to 20 meters). We have barely started.

The most important of these is the fact that the oceans are *not* in static equilibrium. It is a dynamical system in which the periodic tidal force from the Moon (and Sun) are acting on a system that is already a sort of harmonic oscillator. To see why, ignore the continents for a start and think of a wave moving West across the ocean. Why West? We're in an Earth-centered system, in which the Moon rises in the East, so if it starts a wave moving in the ocean, that's the direction it will start it moving. This wave would try to go around the Earth with its own natural period, and depending on latitude, the natural period of such a long wave trying to

Fig. 5.14

go around the Earth can be more or less than a day. How does the natural period of this circumpolar wave vary with latitude? Simple: the distance around is $\ell = 2\pi R \sin\theta = 2\pi \cos\lambda$, where θ is the usual polar coordinate and λ is the latitude. The natural period of the wave around the Earth is this distance over its speed, ℓ/v.
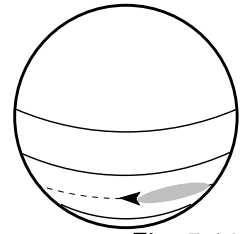
The Moon's apparent daily motion around the Earth (still viewed in our system) applies a force to this wave. How does the frequency of this force compare to the natural frequency with which ocean waves will naturally move around the Earth?

Look back at the material on forced harmonic oscillators in chapter three, and the graphs, Figure 3.5. Notice especially the graph of δ, showing the phase difference between the forcing function and the response function. When the forcing function has a frequency higher than the natural frequency of the oscillator, the response of the oscillator is between $90°$ and $180°$ out of phase with the force. That's what happens with the tides. The period of the Moon's tidal force is just over 12 hours (two per day remember), but the natural period for the tidal sloshing is substantially longer. This forced oscillator (the ocean) will have a steady-state response that follows the frequency of the Moon's tidal force, and the phase difference between force and response puts them as much as $180°$ out of phase. This implies that the high tides will *not* occur when the Moon is overhead (or even on the opposite side of the Earth). At the extreme case, for which there is a $180°$ phase shift, high tide would occur when the Moon is on the horizon. That is a six hour difference.* In the more realistic cases, the tidal bulge is ahead the point under the Moon's position, but by less than $90°$ of longitude. Watch out for a point of confusion here, as degrees of longitude and degrees of phase shift differ from each other by a factor of about two because there are almost two high tides per day. See section 7.11 for a more quantitative analysis of this subject.

Sample Tide Table† for the Eastern Atlantic (St. Augustine, FL) June 1, 2011 (New Moon)

high /low	tide time	height feet	sunrise sunset	moonrise moonset
Low	3:03 AM	0.5		
High	8:52 AM	4.3	6:25 AM	6:03 AM
Low	2:57 PM	0.3		
High	9:14 PM	5.4	8:21 PM	8:26 PM

In this sample from a large tide table, notice that the high tides are not exactly at moon- or sunrise, but they *are* very far from having the Moon or Sun at the Zenith. In the morning they're about midway between high and low tides, while in the evening the moon/sun-rise is only an hour away from the time of high tides.

The Moon does not orbit above the equator, so one lunar tidal bulge can be north of the equator and one south of it. As your point on the Earth moves around, you may move through the middle of a bulge at one time and then about 12 hours later move through the edge of the bulge. You experience

---

* Recall: $180°$ phase difference is about $90°$ of longitude.

a higher high tide and then a lower high tide. They aren't the same, and if you are far enough north you may experience only one high tide in a day.

Qualitatively, what does friction do to the water? A simplistic model would say that the Earth is rotating underneath the tidal bulge, so it pushes the water ahead. This means that as the Moon orbits the Earth, the Earth's rotation pushes the bulge a few degrees ahead of the Moon's position. As you are carried around by the Earth's surface, you will pass under the Moon's position and a little later you will pass through the center of the tidal bulge. This too simple model is overwhelmed by the dynamic processes described in the last few paragraphs, but one aspect of it does apply: The tidal friction affects the Earth. It acts as a brake, gradually increasing the length of a day. The tidal bulge also applies a *torque* to the Moon: The leading bulge applies a component of force forward, parallel to the Moon's velocity vector, and the trailing bulge applies a smaller component backwards — the trailing bulge is farther away. The actual computation of this effect can't be done separately from the resonance phenomenon discussed two paragraphs back.
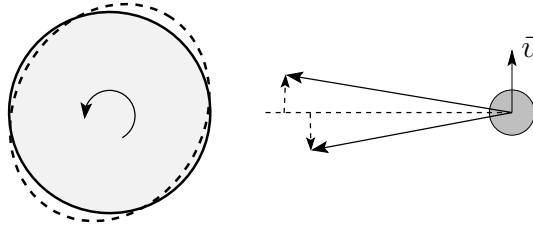
Fig. 5.15

This torque on the Moon tries to accelerate the Moon, and that puts it into a higher orbit. The tides on the Earth cause the Moon to recede gradually from the Earth. You can see this another way: The tidal friction slows the Earth's rotation, causing it to lose some of its rotational kinetic energy $(I\omega^2/2)$. Where does that energy go? Part of it goes into heating, the way friction usually does, but a significant part of it puts the Moon into a gradually higher energy orbit (farther away). This recession rate is a few centimeters per year. Notice that this forward force on the Moon causes the Moon to *slow*.

There is more discussion of the tides in section 7.11, part of chapter seven on waves. There you will find a more detailed analysis about the relationship between the position of the Moon and the position of high tide. In particular, how can low tide happen when the Moon is more-or-less overhead?

5.7 Foucault Pendulum˙

A classic experiment to give a decisive answer to the question: Does the Earth rotate? After all, just because such authorities as Galileo and Newton said it does, is that enough? As late as the middle 1800's this was an important question because people wanted to have some simple, easy-to-see evidence of the rotation, evidence not depending on the motion of the stars. If a pendulum is set up at the South* Pole, what will the Earth's rotation do to it? Stand off the Earth again and the pendulum will be swinging back and forth in a plane while the Earth rotates underneath it. To someone standing on the surface it will appear that the plane of the pendulum's swing will rotate counterclockwise once every 23 hours and 56 minutes, the time it takes the Earth to rotate once on its axis. Setting up a pendulum at the South Pole is a bit far to go for most people, but if you're at the University of New South Wales, you may think that is close enough.
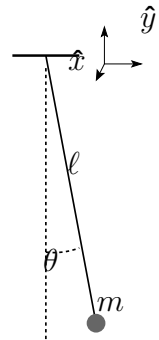

Fig. 5.16

www.phys.unsw.edu.au/PHYSICS_!/FOUCAULT_PENDULUM/foucault_pendulum

To figure out what happens at any point on the Earth other than the poles takes some effort and the apparatus of rotating coordinate systems developed in this chapter. To describe the mechanics, use the same coordinate system as in Eq. (5.27), except that now there's another force on the mass, that

---

* Why not the North Pole? There's an ocean there.

of the cord that suspends the mass from the ceiling. These equations become

$$m\ddot{\vec{r}} = -mg\hat{z} - F_{\text{cord}}\hat{r} - 2m\vec{\omega} \times \dot{\vec{r}}$$
$$m\ddot{\vec{r}} = -mg\hat{z} + F_{\text{cord}}(\hat{z}\cos\theta - \hat{x}\sin\theta\cos\phi - \hat{y}\sin\theta\sin\phi) - 2m\vec{\omega} \times \dot{\vec{r}} \tag{5.41}$$

θ and ϕ are the spherical coordinates from the point of suspension on the ceiling as in section 0.3, *except* that contrary to tradition the angle θ is measured from the $-z$ axis. The pendulum becomes a simple problem only in the approximation that it is oscillating through a small angle. That means that the cord is almost vertical. θ is small and the z-component of acceleration is small, so the z-component of the force is small too.

$$\cos\theta \approx 1 \qquad \text{and} \qquad F_{\text{cord}} \approx mg \tag{5.42}$$

The x and y-coordinates are related as in Eq. (0.13) to the spherical coordinates:

$$x = \ell\sin\theta\cos\phi, \qquad \text{and} \qquad y = \ell\sin\theta\sin\phi$$

Put these into the equation of motion and approximately

$$m\ddot{\vec{r}} = \frac{mg}{\ell}(-\hat{x}x - \hat{y}y) - 2m\vec{\omega} \times \dot{\vec{r}} = \frac{mg}{\ell}(-\hat{x}x - \hat{y}y) - 2m\omega(\hat{y}\cos\lambda + \hat{z}\sin\lambda) \times \dot{\vec{r}} \tag{5.43}$$

Look at the components in the x-y plane, the plane parallel to the Earth's surface.

$$\ddot{x} = -\frac{g}{\ell}x + 2\omega\sin\lambda\,\dot{y} \qquad\qquad \ddot{y} = -\frac{g}{\ell}y - 2\omega\sin\lambda\,\dot{x} \tag{5.44}$$

To solve these, note that these are simultaneous, linear, constant coefficient equations and perhaps use an exponential solution. This is a perfectly correct method and it works, but I'll show two other methods to solve these equations more easily.

The use of complex algebra leads to much simpler equations. Take the first of the equations (5.44) and add i times the second equation.

$$\ddot{x} + i\ddot{y} = -\frac{g}{\ell}(x + iy) + 2\omega\sin\lambda(\dot{y} - i\dot{x}) = -\frac{g}{\ell}(x + iy) - 2i\omega\sin\lambda(\dot{x} + i\dot{y}) \tag{5.45}$$

Let $q = x + iy$ and this is

$$\ddot{q} = -\frac{g}{\ell}q - 2i\omega\sin\lambda\dot{q}$$

This is an ordinary constant coefficient equation with a solution

$$q(t) = Ae^{\alpha t} \qquad \text{with} \qquad \alpha^2 + 2i\omega\sin\lambda\,\alpha + \frac{g}{\ell} = 0$$

$$\alpha = -i\omega\sin\lambda \pm \sqrt{-\omega^2\sin^2\lambda - g/\ell} = i\left[-\omega\sin\lambda \pm \sqrt{\omega^2\sin^2\lambda + g/\ell}\right]$$

To interpret this, recall that the Earth is rotating slowly so that $\omega^2 \ll g/\ell$. Use this observation to get a simpler approximate value of α by a series expansion. Let $\omega_0^2 = g/\ell$, then $\omega^2/\omega_0^2 \ll 1$.

$$\alpha = i\left[-\omega\sin\lambda \pm \omega_0\sqrt{1 + \omega^2\sin^2\lambda/\omega_0^2}\right] = i\left[-\omega\sin\lambda \pm \omega_0(1 + \omega^2\sin^2\lambda/2\omega_0^2)\right] = \pm i\omega_0 - i\omega\sin\lambda$$

More directly, note that you are going to keep just the lowest order terms in ω, dropping ω^2. This means that you can drop it while it is still inside the square root, getting the answer more quickly.
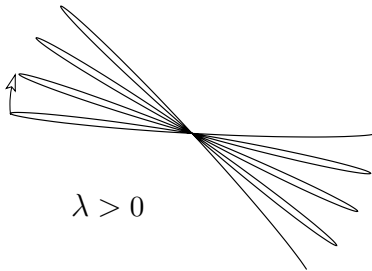
Start the pendulum by pulling it to one side and releasing it gently.

$$x(0) = R, \quad \dot{x}(0) = 0, \quad y(0) = 0, \quad \dot{y}(0) = 0 \quad \longrightarrow \quad q(0) = R, \quad \dot{q}(0) = 0$$
$$q(t) = Ae^{i(\omega_0 - \omega \sin \lambda)t} + Be^{i(-\omega_0 - \omega \sin \lambda)t}$$
$$q(0) = A + B = R, \qquad \dot{q}(0) = Ai(\omega_0 - \omega \sin \lambda) + Bi(-\omega_0 - \omega \sin \lambda) = 0$$

In these equations for the coefficients A and B, do I need to include the terms in $\omega \sin \lambda$? No. All the action occurs in the exponents controlling the phase of the exponentials.

$$A = B = R/2, \qquad q(t) = \frac{R}{2} \left[e^{i(\omega_0 - \omega \sin \lambda)t} + e^{i(-\omega_0 - \omega \sin \lambda)t} \right]$$
$$= R \cos \omega_0 t \, e^{-i\omega \sin \lambda t}$$

The real and imaginary parts of this are

$\lambda > 0$

$$x(t) = R \cos \omega_0 t \, \cos(\omega \sin \lambda t)$$
$$y(t) = -R \cos \omega_0 t \, \sin(\omega \sin \lambda t) \tag{5.46}$$

and this is an oscillation in the x-y plane with frequency ω_0, but the plane of the oscillation is rotating with the slow rotation rate $\omega \sin \lambda$. In this picture the rotation of this plane is greatly exaggerated, and during one swing the pendulum appears to execute a figure eight. If the oscillation period of this pendulum is two seconds, the period of rotation of the Earth would have to be less than one minute for this picture. For the real case the figure-eight motion will squeeze into something so close to a line you can't tell the difference, and the pendulum will appear to move back and forth in a plane that rotates once per day if you're at the Earth's pole or once every two days at latitude 30°.

Sometimes you will see an explanation of the Foucault pendulum claiming that the plane of the pendulum's oscillation stays fixed with respect to the stars. That's true if you are *at* the North or South Pole, but nowhere else. The equator is an extreme case that clearly shows the error of this statement. Simply start the pendulum moving in a north-south plane and it stays there, having no rotation with respect to the Earth, but rotating once per day with respect to the stars.

The Slick Way
And then there's another, more clever way to solve this. Go back to the equations (5.44). Forget about the fact that these refer to a rotating coordinate system and pretend that someone handed them to you in an inertial system. These are now (except for m's) $\vec{F} = m\vec{a}$ in a (pretend) inertial system. Transform these x-y coordinates into a new rotating system. If I do it right I can cancel the ω terms.

The new rotating system will have (unknown) angular velocity about the z-axis, $\vec{\omega}' = \hat{z}\omega'$. Equation (5.12) tells you how to modify the equations of motion by adding the inertial forces. As before, don't bother with the centrifugal force term and simply add the Coriolis term to the forces, $-2\vec{\omega}' \times \dot{\vec{r}}$.

$$\ddot{x} = -\frac{g}{\ell}x + 2\omega \sin \lambda \, \dot{y} \qquad \qquad \ddot{x} = -\frac{g}{\ell}x + 2\omega \sin \lambda \, \dot{y} + 2\omega'\dot{y}$$
$$\ddot{y} = -\frac{g}{\ell}y - 2\omega \sin \lambda \, \dot{x} \qquad \longrightarrow \qquad \ddot{y} = -\frac{g}{\ell}y - 2\omega \sin \lambda \, \dot{x} - 2\omega'\dot{x}$$

If I choose $\omega' = -\omega \sin \lambda$ the equations become two separate harmonic oscillators and in that system the pendulum does not rotate. Go back to the system fixed in the Earth and it says that the plane of oscillation of the pendulum will rotate at this rate ω' and it even gets the direction right. If you're at the North Pole it gives $\omega' = -\omega$, or $\vec{\omega}' = -\hat{z}\omega$ so that it is rotating clockwise. Again step off the Earth for a moment to verify that this is correct. (I had to use dots for time derivatives in both systems. What else can I do?)

In the next chapter, in section 6.6 you will see an analysis of the precession of a pendulum even *without* a rotating Earth. The case examined there is quite a long way from the one here because it concerns a pendulum going around almost in a circle, but the basic ideas will still apply. If the pendulum is swinging back and forth but started with even a tiny amount of angular momentum about the vertical axis, so that the original motion is not a straight line but a very narrow ellipse, then there is a precession that can easily be much larger than that caused by the Earth's rotation. It will completely mask the Foucault effect. This is one reason it was so difficult to build a pendulum to measure the Earth's rotation. Foucault's achievement was not trivial!*

The Really Slick Way
Go back to Eq. (5.41) and leave everything in vector form. After all isn't the use of vectors supposed to simplify the mathematics? Take the cross product of each side with $\vec{r}$.

$$m\ddot{\vec{r}} = -mg\hat{z} - F_{\text{cord}}\hat{r} - 2m\vec{\omega} \times \dot{\vec{r}} \qquad \text{Eq. (5.41)}$$

$$m\vec{r} \times \ddot{\vec{r}} = -mg\vec{r} \times \hat{z} - \vec{r} \times F_{\text{cord}}\hat{r} - 2m\vec{r} \times (\vec{\omega} \times \dot{\vec{r}}) \qquad (5.47)$$

$$m(\vec{r} \times \dot{\vec{r}})^{\cdot} = -mg\vec{r} \times \hat{z} - 0 - 2m[\vec{\omega}(\vec{r} \cdot \dot{\vec{r}}) - \dot{\vec{r}}(\vec{\omega} \cdot \vec{r})]$$

$$\dot{\vec{L}} = -mg\vec{r} \times \hat{z} + 2m\dot{\vec{r}}(\vec{\omega} \cdot \vec{r})$$

$$= -mg\vec{r} \times \hat{z} - 2m\dot{\vec{r}}\omega\ell \sin \lambda \qquad (5.48)$$

This is how you derive the angular momentum equation from $\vec{F} = m\vec{a}$, as in section 1.3. The right side is the torque: The first term is the vector version of $mg\ell \sin \theta$, and the last term is the one that causes all the interesting problems.

As was done a couple of paragraphs back, pretend this is a stationary system and go to a new system rotating with respect to it.

$$\frac{d\vec{L}}{dt} = -mg\vec{r} \times \hat{z} - 2m\dot{\vec{r}}\omega\ell \sin \lambda = \dot{\vec{L}} + \vec{\omega}' \times \vec{L} \qquad (5.49)$$

where this new unknown $\vec{\omega}'$ is the rotation with respect to the old coordinate system.

$$\dot{\vec{L}} + \vec{\omega}' \times \vec{L} = \dot{\vec{L}} + m\vec{\omega}' \times (\vec{r} \times \dot{\vec{r}}) = \dot{\vec{L}} + m[\vec{r}(\vec{\omega}' \cdot \dot{\vec{r}}) - \dot{\vec{r}}(\vec{\omega}' \cdot \vec{r})] \qquad (5.50)$$

If the rotation of this system is about the vertical axis, the vector $\vec{\omega}'$ is perpendicular to the velocity vector $\dot{\vec{r}}$. (Remember: small angle approximations here.) The equation is now

$$\dot{\vec{L}} - m\dot{\vec{r}}(\vec{\omega}' \cdot \vec{r}) = -mg\vec{r} \times \hat{z} - 2m\dot{\vec{r}}\omega\ell \sin \lambda \qquad (5.51)$$

Now let $\vec{\omega}' = -2\omega \sin \lambda \hat{z}$ and the second terms on the two sides then cancel, leaving only the $mg\ell \sin \theta$ term that causes the pendulum to swing back and forth. There's no rotation left, so this defines the coordinate system in which the pendulum does not precess. Notice: slicker does not mean easier.

---

* The book "Pendulum: Leon Foucault and the Triumph of Science" by Aczel is a good history of the subject.

Exercises

1 When a car turns a corner with a radial acceleration of magnitude no more than g, how fast can it be moving to go from the right lane into the right lane if the lanes are 12 feet wide?

2 A mass is at rest in a coordinate system that is also at rest. Now the mass remains at rest with respect to this coordinate system, but the system starts to rotate at constant angular acceleration, so $\omega = \alpha t$. Just as the rotation starts, what does Eq. (5.11) say about the mass, and what force does it take to hold it at rest in this system?

3 If a vector is stationary in an inertial system, what is its time derivative in a rotating system?

4 Find a root of $0.001x^3 + x + 1 = 0$ to at least four significant figures.

5 Find a root of $x - 2 + .01 \sin x = 0$ to five or more significant figures.

6 In Figure 5.4, if the person throwing the ball throws it directly to his right, what direction will the Coriolis force on it be? And if he throws it to his left?

7 The same as the immediately preceding exercise, only now describe what an aerial view by a stationary observer will be.

8 Fill in the algebra to get Eq. (5.40).

9 If the Sun is compressed to a point, a black hole, and you are falling into it feet first, how far from the origin would you have to be so that the gravitational field on your head and on your feet differ by $10 \, \text{m/s}^2$ (the onset of spaghettification)? This is just an extreme version of a tide. [Ans: about 3700 km] See also problem 6.14 for another aspect of this question. You can do this the hard way or you can compute dg/dr. *This* is a unique case in which you can actually feel the gravitational force. (Briefly)

10 In the preceding exercise on spaghettification, at what speed would you be moving if you are falling in from very far away?

11 On page 175 it says that the Moon is receding from the Earth at a few centimeters per year. How far away would this say that it will be in one billion years? Why is this estimate a great exaggeration?

Problems

5.1 Generalize the equations (5.3) to arbitrary straight-line motion: $x' = x - f(t)$ where f is any function. How are velocity and acceleration transformed?

5.2 If you are in a car that's accelerating at a_0, use Eqs. (5.5) to describe the total forces that you perceive, assuming that you use the accelerated coordinate system in which you are at rest.
A cord is hanging down from the roof inside the car. What angle does it make with the vertical in this case? Ans: $\tan^{-1}(a_0/g)$

5.3 As in the preceding problem the cord is hanging down from the roof inside the car, but this time you are standing by the side of the road in an inertial coordinate system. Now figure out all the forces and deduce what angle the cord makes with the vertical. Ans: Still $\tan^{-1}(a_0/g)$

5.4 A mass is hanging at the end of a one-meter length of light string, and it is suspended from the roof of the elevator that you are taking. You notice that the period of the pendulum is 3 seconds. Are you in trouble? If so, how much?

5.5 Derive Eqs. (5.17) from Eqs. (5.16) using Eqs. (9.5) on page 305. Here, the angle $\alpha = \omega t$.

5.6¨Derive the equations (5.17) by solving the equations (5.12) directly. Pick $\vec{\omega} = \omega \hat{z}$, and there's a trick that will save you a *lot* of work: Instead of the basis $\hat{x}$ and $\hat{y}$, use $(\hat{x} + i\hat{y})$ and $(\hat{x} - i\hat{y})$.

5.7 Fill in the missing steps in Eqs. (5.18)–(5.30).

5.8˙In two dimensions, you have basis vectors $\hat{x}$ and $\hat{y}$. Now take two other unit vectors that are rotating so that $\phi = \omega t$. Write $\hat{x}'$ and $\hat{y}'$ in terms of $\hat{x}$ and $\hat{y}$ as in Figure 5.6. A vector $\vec{r}$ is expressed in the rotating system as $x'(t)\hat{x}' + y'(t)\hat{y}'$. Express the rotating unit vectors in terms of the stationary $\hat{x}$ and $\hat{y}$ and from this compute $d\vec{r}/dt$. Compare this to the general result in the text. Apply this to the particular example for which $x'(t) = x_0$ and $y'(t) = v_0 t$, drawing enough pictures to be sure that you understand what is happening.
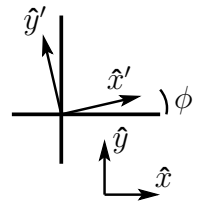
5.9 For the golf ball and the artillery shell from Eq. (5.30) what result do you get if you aim south instead of north? What happens in the southern hemisphere at $55°$ South latitude?

5.10˙Stand in the center of a rotating platform and fire a bullet horizontally. ($\vec{\omega}$ is up.) Turn off gravity for a moment to make this a two-dimensional problem. (**a**) What is the trajectory of the bullet in your coordinate system? That is, what are $r(t)$ and $\phi(t)$ as seen by you. For this problem, you see that it is much easier to ask someone who is in an inertial system to solve it, then translate what this means into your $r(t)$ and your $\phi(t)$. (**b**) Having done this, determine if your results satisfy the equations of motion in your rotating system, *i.e.* Eq. (5.12). Also, draw pictures of what the trajectory looks like to you and what it looks like to your inertial friend. You will also need Eqs. (0.40) and (0.41). (**c**) Also express your results in polar coordinates as an equation for r in terms of ϕ so you can more easily see how the shape of the trajectory will appear to you. Ans: (c) $r = -\frac{v}{\omega}\phi$

5.11 Analyze Eq. (5.30) for various special cases, *e.g.* equator, North or South pole, fired vertically, fired almost horizontally—all combinations. When is it zero?

5.12¨There is a force in the x-y plane $\vec{F_0} = k\vec{r}_\perp = k(x\hat{x} + y\hat{y})$ where $k > 0$. A mass m under this force will move rapidly away from the neighborhood of the origin. (Show this.) The magnetic force on a charged particle is $\vec{F_1} = q\vec{v} \times \vec{B}$. If $\vec{B}$ is along the z-axis, show that a sufficiently strong

magnetic field will stabilize the motion of the (charged) mass, and that the motion of this mass will then be confined to some region around the origin. For this problem you're better off in rectangular coordinates. The Coriolis force has the same structure as this magnetic force, and in analyzing the motion of the Trojan asteroids the problem is quite similar to this one (but with a lot more algebra). To see that analogous problem of the Trojan asteroids worked out, look up chapter 7 of the mechanics text by Symon, mentioned in the bibliography, page iv. Ans: Stable if $qB/m > 2k/m$

5.13 ¨You can include air resistance in the calculations that started at Eq. (5.19): add a term $-b\vec{v}$. The solution for $\vec{r}_0$ is already done starting at Eq. (2.29), and you can use that result here in order to improve the calculation leading to Eqs. (5.23) and (5.24). Assume that the projectile is fired straight up with the same initial speed 300 m/s and that the terminal speed is 100 m/s. Now where will it land? A still better calculation would be to use the added force $\propto v^2$ as in problem 2.30, but that would be cruel.

5.14 The masses shown are connected by a string of negligible mass over a pulley, also of negligible mass. The pulley accelerates up at a_0 because of an external force. Find the accelerations of the two masses in the inertial system. Take $m_1 \geq m_2$. If $a_0 = 0$ or if $g = 0$, what results do you get?
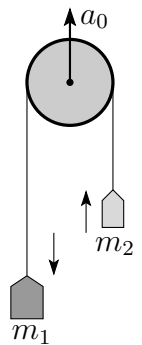
5.15 In the example starting at Eq. (5.27), derive algebraically how the size of the deflection depends on the initial speed of the golf ball. That is, how does the deflection scale as you multiply the speed by some factor?

5.16 ˙Assume that Coriolis is the only force present. No gravity, no centrifugal, with $\vec{\omega}$ straight up and a flat Earth. Fire a rifle horizontally and get the trajectory. Use rectangular coordinates to solve this problem, and don't throw away any roots just because you expect them to be unimportant. Describe the shape of the trajectory, including a sketch of course. Ans: $\vec{r} = \frac{1}{2\omega}\vec{v}_0 \sin 2\omega t + \frac{1}{2\omega}\hat{\omega} \times \vec{v}_0 (\cos 2\omega t - 1)$

5.17 ˙Add centrifugal force to the preceding problem and now find the trajectory.

5.18 (a) In the analysis of the trajectory of an object thrown straight up from the equator, leading to Eq. (5.26), what is the horizontal velocity of the projectile when it returns to the ground? (b) In trying to understand why the effect of the up and down motion gives a net result to the West of the starting point, how does the answer to part (a) help? Note: After working out the solution to this problem I realized that the original picture on page 165 was wrong due to a programming error. The original drawing is here; what's wrong with it?
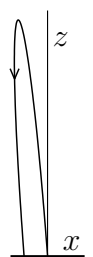
5.19 If you didn't do problem 5.7, then at least fill in the steps leading to Eq. (5.22). Then continue the process to find the next iteration.

5.20 Drop a mass from a height h, initially at rest. Where does it land? You are at latitude λ. Neglect air resistance, just the opposite of the next problem.

5.21 Jump from a great height and you quickly reach terminal speed because of air resistance. Assume that you Jump from height h and quickly reach terminal velocity v_t straight down (and there's no wind), where do you land? Also assume that the terminal speed is a constant, even though it really depends on the air density, which varies with height. You are at latitude λ. Take terminal speed to be 60 m/s, $\lambda = 30°$, and $h = 2$ km. Reaching terminal speed quickly—take it to mean that all the forces except Coriolis balance. For air resistance, use the simplest form, $-b\vec{v}$ and evaluate the numbers on the equator. Ans: about $1\frac{1}{2}$ m East, a negligible effect.

5.22 An insect is trying to walk on a rotating turntable, and with respect to the turntable its motion is a constant speed v_0 directly away from the center. (**a**) Find the force (magnitude and direction) that the turntable must exert on the insect. The coefficient of static friction is μ_s. (**b**) At what point will it start to slip? (**c**) Under what conditions will it *always* slip? Ans: (**a**) $\vec{F} = \hat{\phi}\, 2m\omega vt - \hat{r}\, m\omega^2 vt$, (**b**) Slip when $\omega v \left[4 + \omega^2 t^2\right]^{1/2} > \mu_s g$

5.23 Under what circumstances will the deviation of a golf ball as found in Eq. (5.30) be zero? For a reasonable choice of α, where is this?

5.24 If a bird is flying in level flight at speed v and desires to fly in a straight line despite a horizontal component of Coriolis force, how much force must the bird exert to do so? Do this for a flight north and for flight west. Pick your favorite bird and latitude and evaluate the size of this force.

5.25 Repeat problem 4.26 using the methods of this chapter.

5.26 The differential equation for a charged particle in a magnetic field and an electric field is

$$m\frac{d^2\vec{r}}{dt^2} = q\left[\vec{E} + \frac{d\vec{r}}{dt} \times \vec{B}\right]$$

as in section 4.3. Write this equation in a coordinate system rotating with constant angular velocity $\vec{\omega}$. Show that if $\vec{B}$ is a constant you can pick $\vec{\omega}$ so as to eliminate $\vec{B}$ if you neglect terms proportional to B^2. What is the differential equation to solve now? You may use components to solve this problem only if you want to do a lot of extra algebra and then get results that are difficult to interpret. (Larmor's Theorem)

5.27˙ Find the differential equation of motion for a pendulum in a coordinate system where the pendulum is stationary. Note that you will need *all* the terms in Eq. (5.12), including the one that hasn't been used in this chapter at all. You don't have to include the Earth's rotation for this one. Ans: Eq. (3.23).

5.28˙ Fill in the missing steps (and explain the steps that *are* there) for Eqs. (5.47) through (5.51). Also draw some vectors to explain the signs.

5.29 For the problem 4.26, a mass on a wire, assume now that there is dry friction, $F_{fr} = \mu_k F_N$.

5.30 Tie a mass to the end of a string and whirl it around your head in a circle.
(**a**) Ignore gravity and find the relation between the radius of the circle, the speed of the mass and the tension in the string, the magnitude of the force that it applies. Does it matter if the motion is clockwise or counterclockwise?
(**b**) Now you are standing at the center of a platform rotating with angular speed ω. Repeat the problem of part (a). Solve this in the rotating system, but verify that it agrees with what you would expect in the stationary system.

5.31˙ An insect is trying to walk on a rotating turntable, and with respect to the turntable its motion is a constant speed v_0 around a circle of radius R and centered at the origin. (**a**) Find the force (magnitude and direction) that the turntable exerts on the insect. The coefficient of static friction is μ_s. (**b**) For what speed v_0 will it slip? Assume that the insect is moving around the circle in the same direction that the turntable is rotating. Be sure to simplify the result before you start to interpret it. Ans: (**b**) $v_0 = \mu_s\sqrt{gR} - R\omega$

5.32 A mass m is free to slide on a circular wire of radius R. The wire is rotating about a vertical axis with fixed angular velocity ω. Find the equation of motion for m then find all points of equilibrium for M and determine their stability for various ω. Ans: $\phi = 0,\ \pm\cos^{-1}(g/R\omega^2)$ (but only for some ω).
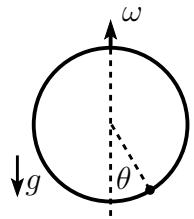
5.33 For the stable equilibria in the preceding problem, find the frequency of oscillation. Ans: $\sqrt{g/R - \omega^2}$ and $\omega\sqrt{1 - (g/R\omega^2)^2}$.

5.34 A particle with no forces on it will move in a straight line at constant velocity — in an inertial coordinate system. Now analyze this problem in a coordinate system that is not only rotating, but is doing so with non-constant ω. Let $\omega = \alpha t$ about the z-axis so that the angular acceleration is constant, and start with the rotated coordinates coinciding with the stationary ones at $t = 0$. Now the modified $F = ma$ has all the extra terms in it *including* the $\dot\omega$ term that we don't normally need. It does not have the "real" force term. Don't try to solve the equations (or if you succeed, please tell me how). Instead write the known solution in the inertial frame, $x = x_0$, $y = v_0 t$, and transform it to the rotating system. Then show that this satisfies the equations. It will be a sufficient amount of algebra to verify one of the two components.

5.35 In a carnival ride you go inside a cylindrical room and stand against the wall. The room spins up and then the floor drops out from under you while you remain against the vertical wall. Find the relation between your coefficient of friction with the wall and the rotation rate and anything else you need in order that you stay pinned against the wall and off the floor. Ans: $\omega^2 > g/r\mu_s$

5.36 Estimate the equatorial bulge that the planet Jupiter should have. Ans: NASA fact sheet: nssdc.gsfc.nasa.gov/planetary/factsheet/jupiterfact.html

5.37 With the result of Eq. (5.40), is the mean sea level really at $\delta = 0$? What is the average of this equation over the surface of the Earth? $\oint \delta\, dA = R_E^2 \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\phi\, \delta$

5.38 (**a**) Considering just the lunar tides, what is the numerical value the equation (5.40) predicts for the daily range from high tide to low tide? (**b**) Now add the solar tide and get the range from highest high tide to lowest low tide.

5.39 Use the data from the tide table on page 174 to sketch the positions of the Earth and Moon and the shape of the ocean surface at moonrise and at moonset. Of course extrapolation from one point to the whole ocean is dicey, but do what you can.

5.40 At the start of section 5.7 it says that a rotation of the Earth occurs every 23 hours, 56 minutes (and 4.1 seconds). Why not 24 hours? And are these numbers close to right?

5.41 (**a**) Derive the result for a spinning bucket containing water as in Eq. (5.32), but using the inertial system — no centrifugal or Coriolis force. (**b**) Also derive the same result using the concept of centrifugal potential energy as in section 5.5.

5.42 In the equation (5.43) there is a z-component in the last term and it was very conveniently ignored. What effect does it have?

5.43 Leave the results for the Foucault pendulum, Eqs. (5.46), in terms of $q = x + iy$ as in the equation preceding it. What is the phase of q and what does that have to do with interpreting the results?

5.44 In some oversimplified explanations of the Foucault pendulum you will read the statement that with respect to the inertial system the plane in which the pendulum oscillates does not rotate. Show

why this is false (with two exceptions). With respect to the inertial frame, what is the rotation rate of the plane of oscillation?

5.45 In section 4.3 you saw the solution of the equations (4.28) by a couple of methods, but not by using complex algebra as in Eq. (5.45). Return to the magnetic field problem and solve it using complex algebra this time.

5.46 You are to design a space station in the form of a cylinder that spins along its axis to provide an artificial gravity for people standing on the inside surface. The centrifugal force there is to be mg and the maximum Coriolis force is to be less than $0.05\,mg$ for someone moving at one meter per second in any direction. What radius and angular speed must the cylinder have (inequalities)? And would you like to be a juggler on such a space station?

5.47 ˙ The torque on the Earth shown in Figure 5.15 causes the Moon to increase its distance from the Earth. Apply conservation of angular momentum, $L_z = \frac{3}{5}M_\mathsf{E}R_\mathsf{E}^2\omega_\mathsf{E} + m_\mathrm{moon}r^2\omega$, to find the change in the Moon's orbital distance from Earth as caused by an increase in our day by one second. Make the simplifying approximation that the Moon is going around the Earth, not that they are going around each other.
Ans: $dr = \frac{3}{10}\left(M_\mathsf{E}/m_\mathsf{M}\right)\left(R_\mathsf{E}^2 T_\mathsf{M}/r_\mathsf{M}T_\mathsf{E}^2\right)dT = 0.8$ m, where r and m are the distance to and the mass of the Moon.

5.48 You are in a space ship accelerating at g so that it feels like home. A light beam enters horizontally through a window on one side of the ship and hits the far wall. In the inertial system it travels in a straight line, but in your system it doesn't. (**a**) What curve does it follow according to you? (**b**) At what angle will it be aimed by the time it hits the wall? The width of the ship is L. Ans: $\Delta\alpha = gL/c^2$

5.49 ˙ Same as the preceding problem, only now the light doesn't enter horizontally, but at some angle to the horizontal. Now compute the deflection of the angle when it hits the far wall, showing the answer is the same as for the case that it entered horizontally.

5.50 ¨ Assume that light will behave the same way in a gravitational field that it does in an accelerated system, then you can deduce the angle by which light is deflected as it passes by the sun. To a first approximation light travels on a straight line. At each point it is in a gravitational field GM/r^2, so as it moves a distance Δx you can think of it as passing through a little space ship that is accelerating at GM/r^2 in a radial direction. Then figure out the width of the ship in terms of the Δx and the angle it enters the ship. Then find the corresponding deflection angle. Then get an integral to find the total deflection angle of light passing the sun. Note: the real answer is twice as big as this one because the correct calculation uses general relativity theory. Ans: $2GM/c^2R \approx 0.87''$
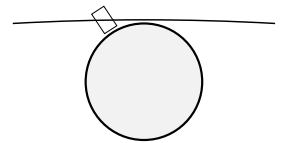
5.51 In problem 5.50, get an approximate value for the result by assuming that the Sun's gravitational field is constant and equal to its value at the surface. Then picture the ship as one large box of width $2R$. What is the deflection of light in this crude approximation? Ans: Same as the exact answer to problem 5.50. You get lucky.

5.52 ¨ Hang a very long cable straight down from a point in a high orbit so that the cable's bottom end touches the surface of the Earth and so that its far end is a distance L straight up. If L is long enough, the whole system will orbit in equilibrium with the cable remaining straight and with zero tension at both the top and bottom ends. This must of course be done at the equator. This *skyhook* has been proposed as an economical way to get cargo into space. Really. The Earth's radius is R, its angular speed is ω, and g_0 is the gravitational field strength at its surface ($g = g_0R^2/r^2$). How long must L

be so that this cable will stay up? Assume that its linear mass density is constant. As a check on your algebra, if $L \to 0$, what must ω do? To solve the (cubic) equation approximately, assume that $L \gg R$. Ans: $L \approx 150\,000$ km, but an iteration on the equation will make it closer to $130\,000$ km.

5.53 ¨In the preceding problem concerning the skyhook, consider the total force on the mass between r and $r + \Delta r$, and find the tension in the cable as a function of height. Where is it maximum? And how large is that maximum for a cable of linear mass density λ? Look up the breaking stress [tension over cross-sectional area] for some standard cable and see if the cable will snap or not, first expressing the maximum stress you just calculated in terms of volume mass density ρ. Now look up the breaking stress for carbon fiber.   Ans: $20\,000$ km. $T_{\text{max}}/\text{area} = \text{stress} = \rho g_0 R\left[-\frac{3R}{2r} + 1 + \frac{1}{2}R\omega^2/g_0\right] \approx \rho g_0 R\left(1 - \frac{3}{2} \cdot 6400/20000\right) = 0.52\rho g_0 R$

5.54 The next time that you get on a merry-go-round, you may decide to take along a bucket that is half-full of water. Put it on down on the platform and see what it does after you've come up to speed. What angle does the surface of the water make with the ground?

Orbits

.

When two objects are orbiting each other, for example two stars in a binary system, or the Earth and Moon, or even the Earth and the Sun, you have to take into account that fact that each is pulled by the other and that both are accelerating. I'm going to pretend that's not happening for the moment. For the Earth going around the Sun, this is a very good approximation, and you can comfortably assume that only the Earth is moving and that the Sun is fixed. Section 6.7 shows a fairly easy way to correct this simplification, finding along the way that there *are* cases where the correction really is important.

6.1 Harmonic Oscillator

Most of the problems involving orbits will involve stars and planets or sometimes colliding nuclei, but there's one example that doesn't, and it is unusual in the way you solve it. Most of these problems are best attacked in spherical or cylindrical coordinates, but the three-dimensional harmonic oscillator is better done in a rectangular system.

$$\vec{F} = -k\vec{r}, \qquad \text{or with less symmetry}$$
$$\vec{F} = -k_1 x \hat{x} - k_2 y \hat{y} - k_3 z \hat{z}$$
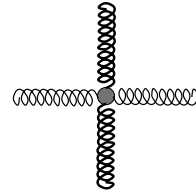
Fig. 6.1

In the second form the constants can be different in three different direction, and in the first form everything is symmetric about the origin. Picture this as a mass held in place by a set of springs, or as an atom sitting in a crystal lattice and held in place by neighboring atoms. Both represent oscillations about a stable equilibrium so both are, in the small motion approximation, harmonic oscillators.

Simply apply $\vec{F} = m\vec{a}$ to get

$$\frac{d^2 x}{dt^2} = -\frac{k_1}{m} x, \qquad \frac{d^2 y}{dt^2} = -\frac{k_2}{m} y, \qquad \frac{d^2 z}{dt^2} = -\frac{k_3}{m} z \tag{6.1}$$

and you know the solutions to these.

$$x(t) = A_1 \cos(\omega_1 t + \delta_1), \qquad y(t) = A_2 \cos(\omega_2 t + \delta_2), \qquad z(t) = A_3 \cos(\omega_3 t + \delta_3) \tag{6.2}$$

or any of the other convenient forms, such as complex exponentials. If this is the symmetric case, then the frequencies are all the same.

It is easier to draw pictures in two dimensions, so stay with that case for now, setting $z = 0$. Also, start with the symmetric case so there is only one k. I can pick the starting time so that $\delta_1 = 0$, and that changes nothing. Now the equations are

$$x(t) = A_1 \cos(\omega t), \qquad y(t) = A_2 \cos(\omega t + \delta) \tag{6.3}$$

A couple of special cases first. Take $\delta = 0$, then y is a constant times x, so this is a straight line, but just over the values specified by the cosine. It is a line segment with slope $y/x = A_2/A_1$ and length $2\sqrt{A_1^2 + A_2^2}$.

If $\delta = \pi/2$ the equation for y is a (negative) sine.

$$x(t) = A_1 \cos(\omega t), \qquad y(t) = -A_2 \sin(\omega t) \tag{6.4}$$

Eliminate the time parameter with the identity $\cos^2 \omega t + \sin^2 \omega t = 1$.

$$\frac{x^2}{A_1^2} + \frac{y^2}{A_2^2} = 1$$
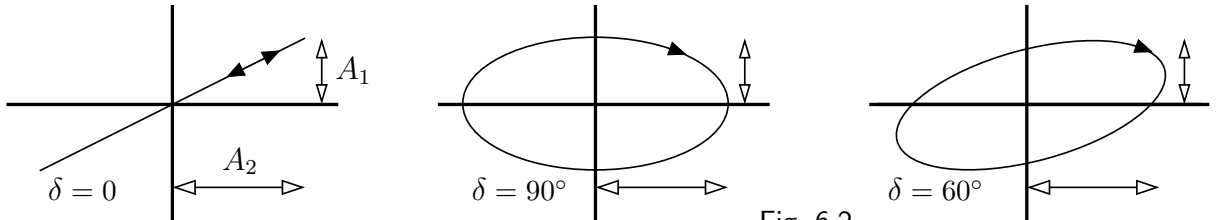
and this is the equation for an ellipse.

$\delta = 0$ $\delta = 90°$ $\delta = 60°$

Fig. 6.2

Any other value of δ will produce an ellipse oriented at some angle to the axes, as in the third figure. If $\delta < 0$ the direction of the orbit reverses. See problem 6.3 for the details. Even the third, tilted, picture above is an ellipse, so that means that by using the right coordinate system, an x'-y' set that is itself tilted, you can put the general form of Eq. (6.3) into the simpler looking form of Eq. (6.4). Proving this and constructing the algebra is not all that easy and not very interesting.

Lissajous Figures
What if the k's aren't equal? Eq. (6.2) will then have $\omega_1 \neq \omega_2$. With this simple-looking change, the orbits get complicated. Take $\omega_2 = 2\omega_1$ as an example.

$$x(t) = A_1 \cos \omega_1 t, \qquad y(t) = A_2 \cos(2\omega_1 t + \delta)$$

If $\delta = 0$ this is

$$x = A_1 \cos \omega t, \qquad y = A_2(2 \cos^2 \omega t - 1),$$
$$y = A_2(2x^2/A_1^2 - 1)$$
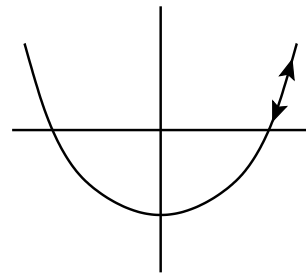
It is an arc of a parabola, but *only* in the finite domain $-A_1 \le x \le A_1$.
If $\delta = \pi/2$ the picture is

$$x = A_1 \cos \omega t, \qquad y = A_2 \sin 2\omega t \tag{6.5}$$
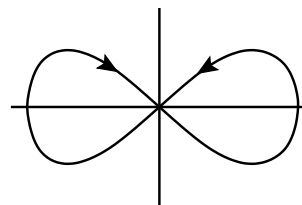
It is a figure eight, being a type of curve called a lemniscate. A change to $\delta = \pi/4$ results in a curve halfway between these two, and you will have to spend a little time to figure out what that means. Very quickly you find that a computer graphics program is a useful adjunct to your tool kit.

These general curves are Lissajous figures, and they get complicated rapidly. Just let $\omega_2 = \omega_1\sqrt{2}$ for the example in figure 6.3, and for more details, including a Java Applet, see Wikipedia and
www-groups.dcs.st-and.ac.uk/~history/Curves/Curves.html

Do such complicated curves show up in real situations? Yes. A variation on these equations appears in the full three-dimensional form much like Eq. (6.2) when you examine one aspect of the gravitational three body problem. The Trojan asteroids are in the orbit of Jupiter but both leading and trailing the planet by $60°$. They move according to equations much like these.
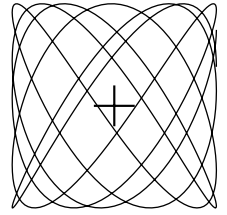
Fig. 6.3

For a more mundane example, there is a device called a "sand pendulum". The support of the pendulum lets it swing in two perpendicular directions and with two different frequencies. The bottom end of the pendulum executes a Lissajous figure and in the process the tip of the pendulum draws that figure in sand. Look for images of these on YouTube.

6.2 Planetary Orbits
In the early 1600's Kepler wrote a series of books on the structure of the solar system. Most of his ideas are forgotten by all but specialists in the history of science. There are three exceptions, now simply called Kepler's laws. They were right and they were important.

1. Planets move along orbits that are ellipses, and the Sun is at one focus of the ellipse.
2. The line from the Sun to a planet sweeps out equal areas in equal times.
3. The periods of the planets are related as $T = ka^{3/2}$, where T is the period, a is then the semi-major axis of the ellipse, and k is an empirical constant.

For the next major step, can you deduce all of Kepler's empirical laws from one basic idea? If you're Newton, yes. He published his results in 1687, in the process inventing what we now call mathematical physics.

Newton's problem is one mass m in motion near a center of force such that $\vec{F} = f(r)\hat{r}$. For Newton's theory of gravity that is $f(r) = -GMm/r^2$, but until I need to I will simply leave this as $f(r)$, and the setup will then apply to any central force.

Starting from $\vec{F} = f(r)\hat{r} = m\vec{a}$, you can show that the trajectory of the planet lies in a single plane. There are a couple of ways to see this, and the simplest is to think of the problem in rectangular coordinates. Start the planet at time zero in the x-y plane with a velocity that is also in that plane. $d\vec{v} = \vec{a}\,dt$ has no z-component, so in each increment of time the new velocity vector will remain with no z-component. That's all it means to say that the motion remains in the x-y plane.
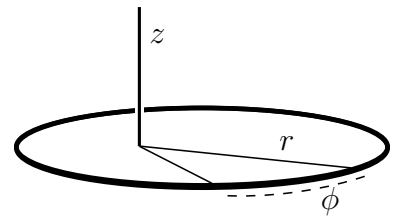
Is there a more elegant way to show this? Always. The angular momentum obeys equation (1.14), $\vec{\tau} = \vec{r} \times \vec{F} = d\vec{L}/dt$. The torque is $\vec{r} \times f(r)\hat{r} = 0$, so the angular momentum $\vec{L} = \vec{r} \times \vec{p} = \vec{r} \times m\vec{v}$ remains constant. This says that the velocity is always perpendicular to the constant angular momentum vector, and that defines the plane in which the planet moves. $\vec{L} = L_z\hat{z}$ in the coordinate system of the preceding paragraph.

Use cylindrical coordinates to describe the motion, as in section 0.6, with the z-axis perpendicular to the plane of the orbit. What is this angular momentum in polar coordinates? Eq. (0.40) implies

$$\vec{L} = \vec{r} \times m\vec{v} = m\,\vec{r} \times \left(\dot{r}\,\hat{r} + r\dot{\phi}\,\hat{\phi}\right) = mr\hat{r} \times \hat{\phi}\,r\dot{\phi} = mr^2\dot{\phi}\,\hat{z} \tag{6.6}$$

The $\vec{r} \times \hat{r}$ term in the torque is zero, so this calculation implies that $r^2\dot{\phi}$ is a constant.

In this system the equation of motion is, using the acceleration from Eq. (0.41)

$$f(r)\hat{r} = m\ddot{\vec{r}} = m\left[\hat{r}(\ddot{r} - r\dot{\phi}^2) + \hat{\phi}(r\ddot{\phi} + 2\dot{r}\dot{\phi})\right] \tag{6.7}$$

The components provide two differential equations, one of which does not even depend on the nature of the force, except for that fact that it is purely radial.

$$m(\ddot{r} - r\dot{\phi}^2) = f(r) \qquad \text{and} \qquad r\ddot{\phi} + 2\dot{r}\dot{\phi} = 0 \tag{6.8}$$

You can solve the second of these independently of the first, and there are a couple of ways to do so. If you're up on your differential equations you can find an integrating factor. That's something to multiply the equation by so that the result is a derivative of something else—then integrate it.
OR, look back at the preceding paragraph on angular momentum. To say that something stays constant is to say that its time-derivative is zero. That is,

$$\frac{d\vec{L}}{dt} = 0 = \frac{d}{dt}mr^2\dot{\phi}\,\hat{z}$$

Since m and $\hat{z}$ are constant, this is

$$\frac{d}{dt}r^2\dot{\phi} = r^2\ddot{\phi} + 2r\dot{r}\dot{\phi} = 0$$

Divide this equation by r, and suddenly the second of the equations (6.8) is solved. It is

$$r^2\dot{\phi} = \text{a constant. Call it } \ell \tag{6.9}$$

This $\ell = r^2\dot{\phi}$ is the angular momentum per mass, as in Eq. (6.6).
    Now use the expression $\dot{\phi} = \ell/r^2$ from Eq. (6.9) in the first of the differential equations (6.8).

$$\ddot{r} - r\left(\frac{\ell}{r^2}\right)^2 = \frac{1}{m}f \qquad \text{or} \qquad \ddot{r} = \frac{1}{m}f(r) + \frac{\ell^2}{r^3} \tag{6.10}$$

This looks like (and is) a one-dimensional problem with a force that is a combination of the real force f and something extra that came from the $r\dot{\phi}^2$ term in the original equations. This combination is referred to as an *effective force*. That it is one dimensional doesn't necessarily make it easy. For example if the force f is zero, you know that the motion will be a straight line, but it's hard to see a straight line coming out of Eq. (6.10).

Example
▶ What does this orbital equation say about simple circular orbits?

$$\text{Here,} \quad \ddot{r} = \frac{1}{m}f(r) + \frac{\ell^2}{r^3} = 0, \qquad \text{so} \qquad f(r) = -\frac{m\ell^2}{r^3} = -\frac{m(r^2\dot{\phi})^2}{r^3} = -mr\dot{\phi}^2$$

This is the familiar elementary result that you get using $a = r\omega^2 = v^2/r$, because ω *is* $\dot{\phi}$.

Example
▶ Halley's Comet has a period of 75.32 years. It goes a long way out and then falls in in a narrow orbit toward the Sun, making a tight swing around it. From this information alone, and using Kepler's 3rd law, estimate the maximum distance from the Sun that this comet reaches.

    Kepler's law is $T = ka^{3/2}$; apply it both to the Earth and to this comet.

$$T_{\mathsf{E}} = ka_{\mathsf{E}}^{3/2}, \qquad \text{and} \qquad T_{\mathsf{H}} = ka_{\mathsf{H}}^{3/2}$$

Divide these.

$$\frac{T_H}{T_E} = \left(\frac{a_H}{a_E}\right)^{3/2} \quad \longrightarrow \quad \frac{a_H}{a_E} = \left(\frac{T_H}{T_E}\right)^{2/3}$$

The longest distance the comet could go occurs if its elliptical orbit is a straight line — an extreme case. That would make the maximum distance from the sun twice its semi-major axis.

$$2a_H = 2a_E \left(\frac{T_H}{T_E}\right)^{2/3} = 2a_E \cdot 75.32^{2/3} = 35.67 \, a_E \tag{6.11}$$

The true distance will be a little less than this, and is measured to be 35.1 AU, where one astronomical unit (AU) equals 150×10^9 m $= 150$ Gm, which is almost exactly the semi-major axis of Earth's orbit. This puts the greatest radial distance of Halley's comet out beyond Neptune's orbit.

6.3 Kepler Problem

Now pick a specific f that varies as $1/r^2$. A straightforward attack on equation (6.10) is difficult, but there's a clue leading to a method that does work. Kepler's analysis of the orbits of planets shows that there are some simple-looking results, one of which is that the orbits are shaped as ellipses. Knowing the shape of the orbit doesn't tell you anything about how the planet moves *as a function of time.* It provides no clue about solving (6.10) directly. The clue is the simple result found if you eliminate the time variable in favor of something else. This works, and the other independent variable to use is ϕ. If I can solve for the function $r(\phi)$ perhaps I'll recognize it as the equation for an ellipse. Now how many people know enough analytic geometry to write the polar equation for an ellipse? Rectangular coordinates maybe, but probably not polar.

The equation in question is

$$r(\phi) = \frac{A}{B + C \cos \phi} \qquad (|C| < |B|) \tag{6.12}$$

A more familiar rectangular equation for an ellipse is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{centered at the origin, or more generally,} \quad \frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1 \tag{6.13}$$

Write Eq. (6.12) as $Br + Cr \cos \phi = A$; write r and $r \cos \phi$ in terms of x and y, and you can finish it easily in problem 6.4, showing that the polar form and the rectangular form describe the same curve. Another way to define an ellipse parallels the definition of a circle. A circle is the set of points all at a constant distance from one fixed point. An ellipse take *two* points and looks for the set of points so that the sum of the distances to the two points is a constant, as in the picture at Eq. (6.17).* The two points are the foci of the ellipse.

This polar equation for an ellipse doesn't have any obvious properties that I might recognize in a differential equation, but what about $1/r$? That's proportional to $B + C \cos \phi$ and we spent a whole chapter on functions such as that — harmonic oscillators. Working backwards from the experimental results then suggests two changes of variables, both the independent and the dependent ones.

$$t \to \phi \qquad \text{and} \qquad r \to u = 1/r$$

---

* What about difference, product, or quotient of distances? Also the sum or difference of squares of distances? They're fun too.

Now it's a lot of application of the chain rule. You can do the changes one at a time in either order, but once you've done this you will ask if there's an easier way. There is: Do both at once.

$$\text{Recall:}\quad \ell = r^2\dot{\phi},\quad \text{so}\quad \ell u^2 = \dot{\phi}$$

$$\frac{dr}{dt} = \frac{dr}{du}\frac{du}{d\phi}\frac{d\phi}{dt} = -\frac{1}{u^2}\frac{du}{d\phi}\ell u^2 = -\ell\frac{du}{d\phi}$$

$$\frac{d^2r}{dt^2} = -\ell\frac{d^2u}{d\phi^2}\frac{d\phi}{dt} = -\ell^2 u^2\frac{d^2u}{d\phi^2}$$

Put this change into Eq. (6.10) to get

$$-\ell^2 u^2\frac{d^2u}{d\phi^2} = \frac{1}{m}f\left(\frac{1}{u}\right) + \ell^2 u^3, \qquad \text{or} \qquad \frac{d^2u}{d\phi^2} + u = -\frac{1}{m\ell^2 u^2}f\left(\frac{1}{u}\right) \qquad (6.14)$$

This is starting to look like a harmonic oscillator. Can you solve it? That depends on f, and for which functions f is it *easy*? If the right side is a constant or if it is proportional to u, then you can do it without difficulty, otherwise it's hard. What happens with the Kepler problem?

$$f(r) = -\frac{GMm}{r^2} \implies \frac{d^2u}{d\phi^2} + u = -\frac{1}{m\ell^2 u^2}\left[-GMmu^2\right] = +\frac{GM}{\ell^2} \qquad (6.15)$$

Leave the other easy case for later, section 6.9, because though those results are amusing, they're not very important. (What f would that be?) Most of the rest of this chapter will concern the exact solutions to the Kepler problem and approximate solutions to all the others.

Equation (6.15) is easy.

$$u = \frac{GM}{\ell^2} + C\cos(\phi - \phi_0), \qquad \text{or} \qquad \frac{1}{u} = r = \frac{\ell^2/GM}{1 + \epsilon\cos(\phi - \phi_0)} \qquad (6.16)$$

C was an arbitrary constant so $\epsilon = C\ell^2/GM$ is too, and it's dimensionless. This is the ellipse as stated in Kepler's laws as long as the parameter ϵ has magnitude less than one. (More on that later, section 6.10.) The numerator of (6.16) is a length, and ϕ_0 is an angle. I may as well assume that $\epsilon \geq 0$ because if not, then I can redefine the parameter ϕ_0 by adding π to it. That changes the sign in front of the cosine back to positive again. With this convention on ϵ, the angle ϕ_0 is the direction in which the denominator is largest. It points toward the smallest r in the orbit.

When talking about planets going around the Sun, the orbital point nearest the Sun is the perihelion, and the one farthest away is the aphelion. For objects orbiting the Earth, the corresponding names are perigee and apogee, and the generic terms are periapsis and apoapsis.
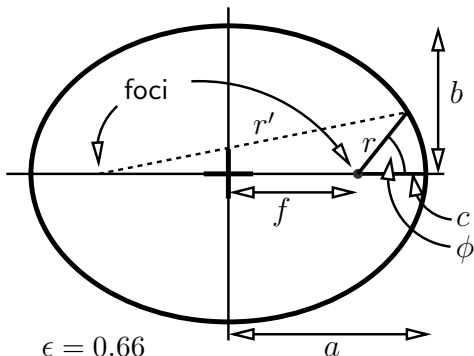
$$r = \frac{a(1 - \epsilon^2)}{1 + \epsilon\cos\phi}$$
$$f = a\epsilon$$
$$b = a\sqrt{1 - \epsilon^2}$$
$$a^2 - b^2 = f^2 \qquad (6.17)$$
$$r + r' = 2a$$
$$A = \pi ab$$
$$c = a - f = a(1 - \epsilon)$$

This figure summarizes many properties of the ellipse. The two governing parameters are a, the semi-major axis, and ϵ, the eccentricity. Other values that can be derived in terms of these start with the semi-minor axis b, the distance from the center to a focus f, and the perihelion distance c. The origin of the polar coordinate system is at the right hand focus. That's where the Sun is. The other focus is symmetrically placed on the other side of the ellipse, and the sum of the distances from the two foci $(r + r')$ is a constant. The ellipse in the drawing has a fairly large eccentricity, $\epsilon = 0.66$, which is far larger than the eccentricity of any planet. Earth's is 0.017, and even Mercury has an eccentricity of only 0.21, the largest of all the planets.* On the scale of this drawing, Earth's orbit is so non-eccentric that the Sun's position would be almost in the center of the ellipse—just on the edge of the darkened part of the vertical axis in the picture. Only Venus has a lower eccentricity than Earth's: 0.0068. The error in drawing Earth's orbit as a perfect circle is about the thickness of the line representing the orbit. Despite this, Tycho's observational data, taken before 1600 and before the invention of the telescope, were so accurate that Kepler was able to use them to discover his laws. For the derivations of these properties of the ellipse, see the exercises on page 221

Example
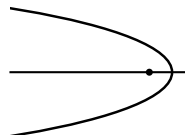
▶ What happens to this ellipse if it is stretched out to infinity along the major axis? That depends on how it's done, because you could stretch it while keeping b constant. That will result in two lines parallel to the horizontal axis and a distance $2b$ apart. A more interesting limit appears if you let $a \to \infty$ while keeping c constant. c is the distance of closest approach on the right end.

$$c = a(1 - \epsilon), \quad \text{so this requires} \quad \epsilon \to 1 \quad \text{as} \quad a \to \infty$$

Go to the first equation in Eq. (6.17), the one for r.

$$r = \frac{a(1 - \epsilon^2)}{1 + \epsilon \cos \phi} = \frac{c(1 + \epsilon)}{1 + \epsilon \cos \phi} \quad \longrightarrow \quad r = \frac{2c}{1 + \cos \phi} \quad \text{as} \quad \epsilon \to 1$$

To see what this represents, put it in rectangular coordinates, with the origin at the focus.

$$r + r \cos \phi = 2c \quad \longrightarrow \quad \sqrt{x^2 + y^2} + x = 2c$$
$$\longrightarrow \quad \sqrt{x^2 + y^2} = 2c - x \quad \longrightarrow \quad x^2 + y^2 = 4c^2 - 4cx + x^2 \tag{6.18}$$
$$\longrightarrow \quad y^2 = 4c^2 - 4cx$$

This is a parabola opening to the left $(x \propto -y^2)$. It has its vertex at $y = 0$, $x = c$.

How much time does it take a planet to orbit the Sun? There are enough relations here to figure that out. Start with $\ell = r^2 \dot\phi$. Solve it for

$$dt = \frac{1}{\ell} r^2 \, d\phi$$
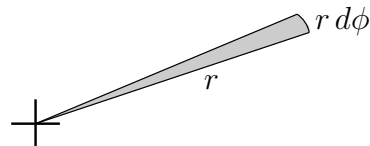
At this point you can substitute Eq. (6.16) and if you know something about complex variables and contour integration you can do the integral fairly easily. (If you don't you should learn, but that's another story.) There's an easier trick here anyway. The factor $r^2 d\phi$ is an area. It is twice the area of a triangle with vertex at the origin.

$$dA = \frac{1}{2} r \cdot r \, d\phi \quad \text{so} \quad dt = \frac{2}{\ell} dA \tag{6.19}$$

---

* Now that Pluto (0.25) has been dethroned.

Notice that this equation *is* Kepler's second law as in section 6.2! The area swept out is proportional to the time elapsed. Integrate dt over the whole ellipse, $0 < \phi < 2\pi$, and you have the period $T = 2A/\ell$. Combine this with the known area of an ellipse, πab, Eq. (6.16), and the other properties in Eq. (6.17).

$$T = \frac{2\pi ab}{\ell} = \frac{2\pi a^2 \sqrt{1 - \epsilon^2}}{\ell}. \qquad \text{Now use} \quad \ell^2 = GMa(1 - \epsilon^2) \text{ from Eqs. (6.16) and (6.17)}$$

Solve the last expression for $\sqrt{1 - \epsilon^2}$ to get

$$\sqrt{1 - \epsilon^2} = \frac{\ell}{\sqrt{GMa}} \qquad \text{then} \qquad T = \frac{2\pi a^2}{\ell} \cdot \frac{\ell}{\sqrt{GMa}} = \frac{2\pi}{\sqrt{GM}} a^{3/2} \qquad (6.20)$$

and this is the third of Kepler's laws. When the eccentricity is zero the semi-major axis is the radius of the circle, $T \propto r^{3/2}$, and you can derive that special case by elementary methods.

Energy
Angular momentum conservation has played an important role in this analysis; count the number of times that $\ell = r^2\dot{\phi}$ has been used so far. What about energy? That should have some significance too.

$$E = \frac{1}{2}mv^2 + U(r), \qquad \text{where} \qquad -\frac{dU}{dr} = f(r)$$

In the same cylindrical coordinates, this is

$$E = \frac{1}{2}m\big[\dot{r}\hat{r} + r\dot{\phi}\hat{\phi}\big]^2 + U = \frac{1}{2}m\big[\dot{r}^2 + r^2\dot{\phi}^2\big] + U \qquad (6.21)$$

Combine this equation with angular momentum, $\ell = r^2\dot{\phi}$, to eliminate the ϕ variable.

$$E = \frac{1}{2}m\left[\dot{r}^2 + \frac{\ell^2}{r^2}\right] + U = \frac{1}{2}m\dot{r}^2 + \left[\frac{m\ell^2}{2r^2} + U(r)\right] \qquad (6.22)$$

This looks like a one-dimensional problem again, with the x-coordinate replaced by r. The role of potential energy is played by the final combination in the last equation, not U alone, but combined with another term called the "centrifugal potential" energy. The combination is called the "effective potential energy".

$$U_{\text{eff}} = U(r) + \frac{m\ell^2}{2r^2} \qquad (6.23)$$

For the Kepler problem Eq. (6.23) is

$$U_{\text{eff}}(r) = -\frac{GMm}{r} + \frac{m\ell^2}{2r^2} \qquad (6.24)$$

The same sort of analysis done in section 2.3 applies here too. In this effectively one-dimensional problem the total energy determines the qualitative behavior of the solution. If $E = E_2$, the allowed

motion is back and forth between the two stopping points where the line E_2 intersects the effective potential energy curve. This is just the elliptical orbit described in the preceding section, and the stopping points are the perihelion and aphelion. You cannot however tell from just this simple energy analysis that the orbit is an ellipse. The value of the energy in this case is

focus

$\epsilon = 0, 0.33, 0.66, 0.99$
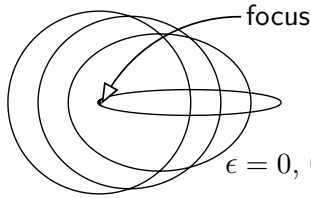
$$E_2 = -\frac{m\ell^2}{2a^2(1 - \epsilon^2)} = -\frac{GMm}{2a} \tag{6.25}$$

You can derive this by finding the stopping point for E_2 and then expressing it in terms of the minimum (or maximum) value of r from Eq. (6.17). See problem 6.32. The energy depends *solely* on the semi-major axis of the ellipse, a. The set of ellipses drawn here all have the same energy and a common focus. That focus is at the center of the leftmost ellipse (the circle) and that dot is less than the line thickness from the leftmost point in the narrowest ellipse. It's hard to see.

If $E = E_1$, the single allowed energy lies at the bottom of the curve and that says that r is a constant in time—a circle. If $E = E_3$ the motion is unbounded. The mass comes in from far away, stops at the intersection of E_3 with U_{eff} and returns to infinity. That is orbit is a hyperbola, discussed in section 6.10.

The boundary between the elliptic and the hyperbolic cases is the line $E = 0$. This orbit is unbounded, like the hyperbola, but its shape is a parabola instead, as in Eq. (6.18)

There are more relationships among these parameters, and some of them are even useful. They are all easy to derive from the other equations in the last couple of pages, and in section 6.11 there is a lot of manipulation that will require several of them. Let c denote the orbital distance at perihelion, then with $\epsilon < 1$

$$
\begin{aligned}
c = a - f = a(1 - \epsilon) \qquad && f = \frac{c\epsilon}{1 - \epsilon} \qquad && b = c\sqrt{\frac{1 + \epsilon}{1 - \epsilon}} \\
E = -\frac{m\ell^2}{2c^2}\frac{1 - \epsilon}{1 + \epsilon} \qquad && \ell^2 = GMc(1 + \epsilon) \qquad && r = \frac{c(1 + \epsilon)}{1 + \epsilon \cos \phi}
\end{aligned}
\tag{6.26}
$$

6.4 Insolation

Notice the spelling; this refers to the power received from the Sun, not what you put in your attic. The power radiated by the Sun is $P_\odot = 2 \times 10^{30}$ Watts. It spreads out radially and at a distance r is spread over a sphere of area $4\pi r^2$. The irradiance*, power per area, at that distance is $P_\odot/4\pi r^2$. How does the power hitting the Earth vary with the season? From the first of Eqs. (6.17),

$$\pi R^2 \cdot \frac{P_\odot}{4\pi r^2} = \frac{P_\odot R^2}{4}\frac{(1 + \epsilon \cos \phi)^2}{a^2(1 - \epsilon^2)^2}$$

where R is the Earth's radius. The ratio at perihelion ($\phi = 0$) to that at aphelion is, with $\epsilon = 0.017$,

$$\frac{(1 + \epsilon)^2}{(1 - \epsilon)^2} = 1.07$$

The perihelion is in January and the aphelion is in July. "$\odot$" is a commonly used symbol for the Sun.

---

* The term intensity is sometimes used here, but that word is also used for power per area per solid angle. (That's radiance.) It's better to use the technical term and avoid confusion.

The annual energy received from the Sun is the integral

$$P_\odot \pi R^2 \int dt \, \frac{1}{4\pi r^2} = P_\odot \pi R^2 \oint d\phi \, \frac{dt}{d\phi} \frac{1}{4\pi r^2} = P_\odot \pi R^2 \oint d\phi \frac{r^2}{\ell} \frac{1}{4\pi r^2} = P_\odot \pi R^2 \frac{2\pi}{\ell 4\pi} = P_\odot \frac{\pi R^2}{2\ell} \quad (6.27)$$

and in this equation the eccentricity cancels, leaving only the dependence on angular momentum. As a check, what should this be for a circular orbit, where $\ell = vr$, and does it agree with this result?

It is worth looking at the result expressed in terms of other variables. Combine pieces of Eqs. (6.26) to get

$$\ell = \sqrt{GMa(1 - \epsilon^2)}, \qquad \text{so} \qquad P_\odot \frac{\pi R^2}{2\ell} = P_\odot \frac{\pi R^2}{2\sqrt{GMa(1 - \epsilon^2)}} \quad (6.28)$$

Without further information about the parameters, it is not clear that there are any implications to be drawn from these equations. A detailed analysis of the interactions of the planets does lead to important consequences however. The solar system consists of the Sun and Jupiter plus detritus. Maybe you can include Saturn too, but that's about it. The question is: what is the effect of Jupiter on Earth's orbit. That is far beyond this text, but I can quote some results.

The shape of Earth's orbit changes over time so that the eccentricity will vary slightly, as will the direction of the perihelion with respect to the distant stars, and the time scale for these changes is the order of $100\,000$ years. One thing that does not change on this time scale is a, the semi-major axis. The immediate consequence of this is that the annual energy reaching the Earth will vary over time as ϵ varies. This is apparent in the last form in Eq. (6.28). This is a sufficiently complicated problem that it is still not fully understood, and you can search for information relating to it under the heading of "Milankovitch cycles". There you will see that I've barely started on the subject.

What is the *quantitative* implication of the preceding paragraph? That depends on how much ϵ varies. If it varies from 0.01 to 0.06 (not out of line with expected values), the ratio of the mean insolations will be $[(1 - .01^2)/(1 - .06^2)]^{1/2} = 1.0018$. A naïve calculation would imply that this will provide a temperature change of about $0.13°\,$C. Life is never that simple of course. There are probably interactions with the precession of the Earth itself, and it may matter whether perihelion occurs during the Northern winter as it does now, or during the Northern summer as it will in about a dozen millenia.

6.5 Approximate Solutions
There are few orbit problems that can be solved exactly in terms of familiar functions. And if there are others that can be solved in terms of unfamiliar functions, then which gives more insight, an exact complicated solution or an approximate simple solution? Usually the latter.

Some examples of force laws that can really happen:
(1) Is our sun exactly spherical? If it is slightly oblate the way that the Earth is, then its gravitational field isn't exactly the same as a point mass at its center. It will have a small part that drops off as $1/r^4$ added to the usual $1/r^2$ force.
(2) If the solar system is embedded in a background of a uniformly dense material (dark matter? dust?) how does that affect orbits?
(3) There are relativistic corrections to the equations, and they will affect the orbit.
(4) The gravitational pull by the other planets, especially Jupiter, contribute. This is by far the largest of all these effects, but deriving it can wait until the end of the chapter, section 6.12, as it is somewhat involved.

There are two sorts of questions I want to examine about orbits: What is their shape? What is the time dependence of the object in orbit? There are two starting points for this analysis, Eqs. (6.10) and (6.14). The second is concerned with *only* the shape of the orbit and the first can do either, but not as easily.

Example

► A reminder of the methods: If you apply $F = ma$ to the case $F_x(x) = m\ddot{x}$, what happens at or near a point of equilibrium? Use series expansions. For example, $F_x(x) = \alpha x - \beta x^3$ and equilibrium occurs when this is zero.

$$F_x(x_0) = 0 \qquad\qquad \longrightarrow \qquad \alpha x_0 - \beta x_0^3 = 0$$
$$x_0 = (\alpha/\beta)^{1/2}$$
$$F_x(x) = F_x(x_0) + (x - x_0)F_x'(x_0) + \cdots \qquad\qquad F_x' = \alpha - 3\beta x^2$$
$$F_x'(x_0) = -2\alpha$$
$$F_x(x) = 0 + -2\alpha(x - x_0)$$
$$m\ddot{x} = F'(x_0)(x - x_0) \qquad\qquad m\ddot{x} = -2\alpha(x - x_0)$$
$$\text{Let } z = x - x_0$$
$$m\ddot{z} - F'(x_0)\,z = 0 \qquad\qquad m\ddot{z} + 2\alpha z = 0$$

And the last equation is a harmonic oscillator, assuming that α is positive of course.

First Method

Start with Eq. (6.10). First write the equation defining a circular orbit, $r = r_0$. This is for an arbitrary radial force at this point.

$$\ddot{r} = \frac{1}{m}f(r) + \frac{\ell^2}{r^3}, \qquad \text{then} \qquad 0 = \frac{1}{m}f(r_0) + \frac{\ell^2}{r_0^3} \tag{6.29}$$

For a *nearly* circular orbit, r will be nearly r_0. That means that I can write $r = r_0 + x$ and assume that x is small compared to r_0. Next do a series expansion to determine the properties of x. As with the sort of expansions so familiar from chapter three, this will be a harmonic oscillator.

$$\ddot{r} = \ddot{x} = \frac{1}{m}f(r_0 + x) + \frac{\ell^2}{(r_0 + x)^3} = \frac{1}{m}\left[f(r_0) + xf'(r_0) + \cdots\right] + \frac{\ell^2}{r_0^3(1 + x/r_0)^3}$$
$$= \frac{1}{m}\left[f(r_0) + xf'(r_0) + \cdots\right] + \frac{\ell^2}{r_0^3}\left(1 - 3\frac{x}{r_0}\right)$$
$$\ddot{x} = \left[\frac{1}{m}f'(r_0) - 3\frac{\ell^2}{r_0^4}\right]x = \frac{1}{m}\left[f'(r_0) + \frac{3}{r_0}f(r_0)\right]x \tag{6.30}$$

The leading terms, the ones without the x, add to zero. I used Eq. (6.29) twice in the last line, first to eliminate the constant term, then to rearrange the coefficient of x. Is this a harmonic oscillator? That depends on the sign of the last factor: Is it negative? Try the Kepler case; you already know the exact answer there, so it's a good laboratory.

A graph of the effective radial force versus r:

$$f(r) = -\frac{GMm}{r^2},$$

$$f'(r_0) + \frac{3}{r_0}f(r_0) = 2\frac{GMm}{r_0^3} - \frac{3}{r_0}\frac{GMm}{r_0^2} = -\frac{GMm}{r_0^3}$$

$$f_{\text{eff}}(r) = -\frac{GMm}{r^2} + \frac{\ell^2}{r^3}$$
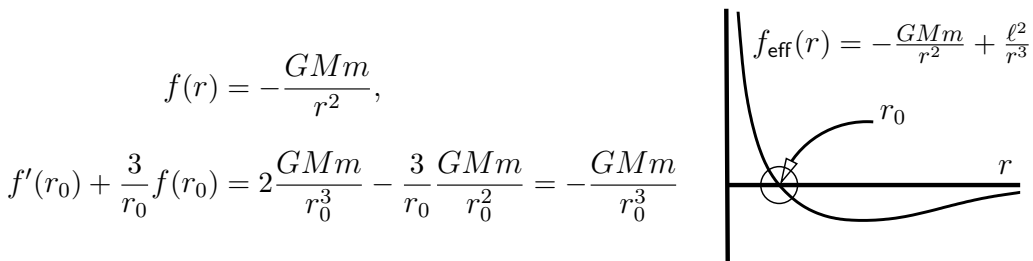
The differential equation (6.30) for x is then a harmonic oscillator equation.

$$\ddot{x} = -\frac{GM}{r_0^3}x \implies r(t) = r_0 + x = r_0 + x_0\cos(\omega_0 t + \delta), \quad \text{where} \quad \omega_0^2 = \frac{GM}{r_0^3} = \frac{\ell^2}{r_0^4}$$

Where did that last equation for ω_0^2 come from? That is Eq. (6.29) for the case that $f(r) = -GMm/r^2$. This equation tells only the way that r oscillates. It doesn't by itself describe the orbit because you need to know the angular variable $\phi(t)$ for that. This angular coordinate will come from Eq. (6.9) for angular momentum per mass.

$$\dot{\phi} = \frac{\ell}{r^2} = \frac{\ell}{(r_0 + x)^2} \longrightarrow \frac{\ell}{r_0^2}\left(1 - \frac{2x}{r_0}\right) = \omega_0\left(1 - \frac{2x}{r_0}\right)$$

$$\text{If} \quad x = x_0\cos\omega_0 t, \quad \text{then} \quad \phi(t) = \omega_0 t - \frac{2x_0}{r_0}\sin\omega_0 t \quad (6.31)$$

If all you want is the shape of the orbit, the second (small) term in $\phi(t)$ isn't important, so I will start by ignoring it. At the end, come back and see what it does.

$$\phi(t) = \omega_0 t, \quad r(t) = r_0 + x_0\cos\omega_0 t \quad (6.32)$$

As t increases from 0 to $2\pi/\omega_0$, the angle ϕ goes from zero to 2π. The planet orbits once. In that same period, r starts at $r_0 + x_0$, decreases to $r_0 - x_0$, and then returns to its starting point. It is a closed orbit.
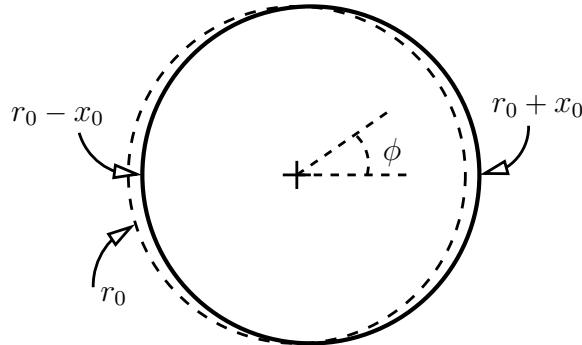
Fig. 6.4

Compare this to the exact result from Eq. (6.16), and then look at the first term in its series expansion.

$$r = \frac{\ell^2/GM}{1 + \epsilon\cos(\phi - \phi_0)} \approx r_0\big(1 - \epsilon\cos(\phi - \phi_0)\big)$$

Is ℓ^2/GM equal to r_0? Yes, go back and check it out. How about the sign? Take $\phi_0 = \pi$ and it's the same. $x_0 = \epsilon r_0$.

The shape of the approximate orbit matches the shape of the exact orbit. Now, what about that extra term in Eq. (6.31)? $\phi = \omega_0 t - (2x_0/r_0)\sin\omega_0 t$. At time zero, the positions match, but $\dot{\phi}$ is a little smaller than the ω_0 that it had without this extra term. Farther away from the sun, the angular speed is less, and that is nothing more than conservation of angular momentum, $mr^2\dot{\phi}$. Now what about the linear speed?

$$v = (r_0 + x_0)\dot{\phi} = (r_0 + x_0)\omega_0(1 - 2x_0/r_0) = r_0\omega_0(1 - x_0/r_0)$$

It is a little less than the speed $r_0\omega_0$ for the original circular orbit. That makes sense; the planet is farther away, so its kinetic energy is a little less. It has climbed up the potential well. On the other side of the orbit the speed is correspondingly larger than $r\omega_0$ so it catches up.

To see an example of what another shape of orbit could be, suppose that the rate of radial oscillation is exactly five times the revolution rate.

$$\phi = \omega_0 t, \qquad r = r_0 + \epsilon \cos(5\omega_0 t)$$
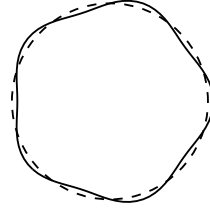
Second Method

Start with Eq. (6.14) to examine the orbit alone, without following the time dependence. The aim again will be to examine some *almost* circular orbits. With $u = 1/r$,

$$\frac{d^2 u}{d\phi^2} + u = -\frac{1}{m\ell^2 u^2} f\left(\frac{1}{u}\right) \qquad\qquad \text{Eq. (6.14)} \quad (6.33)$$

A circular orbit is $r(\phi) = r_0 = \text{constant}$. For this case the term $d^2 u/d\phi^2$ is zero, so $u_0 = 1/r_0$ satisfies

$$u_0 = -\frac{1}{m\ell^2 u_0^2} f\left(\frac{1}{u_0}\right) \qquad\qquad (6.34)$$

This may be a hard equation to solve or it may be easy, but leave it alone for now because it isn't yet clear just how much of the equation *needs* to be solved. If and when I have to come back and solve it I will. To determine the shape of an orbit that is almost circular, assume that u is almost constant, so look for a solution in the form

$$u(\phi) = u_0 + x(\phi), \qquad \text{so that} \qquad \frac{d^2 x}{d\phi^2} + u_0 + x = -\frac{1}{m\ell^2(u_0 + x)^2} f\left(\frac{1}{u_0 + x}\right)$$

When $x \equiv 0$ this of course satisfies the immediately preceding equation for the circle. If x is small ($\ll u_0$) do a series expansion on everything

$$\frac{d^2 x}{d\phi^2} + u_0 + x = -\left[\frac{1}{m\ell^2 u_0^2(1 + x/u_0)^2}\right] f\left(\frac{1}{u_0(1 + x/u_0)}\right)$$

$$= -\frac{1}{m\ell^2 u_0^2}(1 - 2x/u_0) f\left(\frac{1}{u_0}(1 - x/u_0)\right)$$

$$= -\frac{1}{m\ell^2 u_0^2}(1 - 2x/u_0)\left[f\left(\frac{1}{u_0}\right) - f'\left(\frac{1}{u_0}\right)\frac{x}{u_0^2}\right]$$

$$= -\frac{1}{m\ell^2 u_0^2} f\left(\frac{1}{u_0}\right) + \frac{1}{m\ell^2 u_0^2}\left[f\left(\frac{1}{u_0}\right)\frac{2x}{u_0} + \frac{x}{u_0^2} f'\left(\frac{1}{u_0}\right)\right]$$

$$\frac{d^2 x}{d\phi^2} + x = +\frac{1}{m\ell^2 u_0^2}\left[f\left(\frac{1}{u_0}\right)\frac{2}{u_0} + \frac{1}{u_0^2} f'\left(\frac{1}{u_0}\right)\right] x \qquad\qquad (6.35)$$

The last simplification used the equation for a circular orbit, (6.34). This is now a simple equation for $x(\phi)$, because all the coefficients are constants. Does this provide a solution for $r(\phi)$? Just one more step.

$$r = \frac{1}{u} = \frac{1}{r_0 + x} = \frac{1}{u_0}\left(1 - \frac{x}{u_0}\right) = r_0 - r_0^2 x \qquad\qquad (6.36)$$

You already know the exact solution for the Kepler problem, where the force varies as $1/r^2$, so the first question to ask should be whether this approximate equation predicts correct results when you *know* the answer. I'll leave this to the problem (6.11), but with the caution that it has a minor pitfall near the start of the problem, so if you get a result that fails to agree with the exact one, go way back.

Example

▶ Suppose the radial force is constant, $f = -f_0$. The equation (6.35) is then

$$\frac{d^2x}{d\phi^2} + x = \frac{1}{m\ell^2 u_0^2} f \frac{2}{u_0} x \qquad \text{then} \qquad \frac{d^2x}{d\phi^2} + x\left[1 + 2f_0/m\ell^2 u_0^3\right] = 0$$

I still need the equation for r_0.

$$u_0 = +\frac{1}{m\ell^2 u_0^2} f_0, \qquad \text{or} \qquad u_0^3 = f_0/m\ell^2$$

The equation for the perturbation x is now

$$\frac{d^2x}{d\phi^2} + x\left[1 + 2\right] = 0, \qquad x(\phi) = A\cos(\sqrt{3}\phi + \delta)$$

The orbit is $r(\phi) = r_0 + r_1\cos(\sqrt{3}\phi + \delta)$, where $r_1 = -Ar_0^2$ is another form for the arbitrary constant representing the small deviation from a circular orbit.

When ϕ increases by 2π, the oscillation in radius will have gone through a phase of $2\pi\sqrt{3}$, a little less that $1\,3/4$ cycles. In the Kepler problem, when the planet has gone around once, $\phi = 2\pi$, the planet is back where it started and with the same velocity. The orbit is closed. In contrast, here no no matter how many times the mass orbits, it never comes back to exactly the same position and velocity. $\sqrt{3}$ is not a rational number; it can't be expressed as the quotient of two integers, so no integer multiple of it will be an integer. Can this happen in the real world?
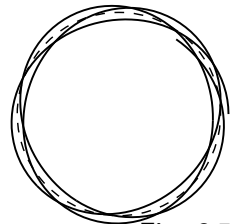
Fig. 6.5

Yes, remember that calculations for the Kepler problem assumed that the single thing affecting the motions of a planet is the sun. What about the pull by the other planets? When you take them into account the problem is of course far more difficult, but one consequence is that the elliptical orbits are only a (very good) first approximation and the orbits don't *quite* close. The perihelion of Mercury precesses around the Sun at the rate of about 575 seconds of arc per Earth century. That is $0.16°$ per century. It doesn't sound like much, but it was measured by about 1800 and (most but not all of) the explanation was worked out as due to the pulls by the other planets.

Example

▶ If the solar system is filed with a uniform dust cloud of density ρ, it will provide an added force on the planets: $F_r(r) = -Gm\rho r/3$. Evaluate the effect on the planet's orbit caused by this dust. It will be enough to find the orbit without finding the time dependence.

The total force on the planet is $-GMm/r^2 - Gm\rho r/3$, and using the same change of variables as in Eq. (6.14), $u = 1/r$, the equation for the shape of the orbit is

$$\frac{d^2u}{d\phi^2} + u = -\frac{1}{m\ell^2 u^2} f\left(\frac{1}{u}\right) = -\frac{1}{m\ell^2 u^2}\left[-GMmu^2 - Gm\rho/3u\right] = \frac{GM}{\ell^2} + \frac{G\rho}{3\ell^2 u^3}$$

For small oscillations about the constant radius, $u = u_0 + x$, and this becomes

$$\frac{d^2x}{d\phi^2} + u_0 + x = \frac{GM}{\ell^2} + \frac{G\rho}{3\ell^2(u_0 + x)^3} \approx \frac{GM}{\ell^2} + \frac{G\rho}{3\ell^2 u_0^3}(1 - 3x/u_0)$$

The equation for the circular orbit is just $x \equiv 0$, which determines u_0. The equation for x itself is then

$$\frac{d^2 x}{d\phi^2} + x = -\frac{G\rho}{3\ell^2 u_0^3}\frac{3x}{u_0}$$

This is familiar: $x(\phi) = x_0 \cos(\omega\phi)$ where

$$\omega = \left(1 + G\rho/\ell^2 u_0^4\right)^{1/2} \approx \left(1 + G\rho/2\ell^2 u_0^4\right) \approx \left(1 + G r_0^4 \rho/2\ell^2\right)$$

(Check the units.) This is a little larger than one, so when ϕ goes around by 2π, the product $\omega\phi$ is slightly large than 2π. This means that the perihelion will have been reached slightly before the full revolution about the sun. The planet precesses backwards (retrograde). By how much? Just set $\omega\phi = 2\pi$ to get $\phi = 2\pi/\omega$. This in turn gives

$$\frac{2\pi}{\omega} - 2\pi = 2\pi\left(\frac{1}{\omega} - 1\right) = 2\pi\left(\frac{1}{1 + G r_0^4 \rho/2\ell^2} - 1\right) = -\frac{\pi G r_0^4 \rho}{\ell^2} = \Delta\phi$$

How does this vary? The first thing to ask then is what is ℓ? Look back and find that $\ell^2 = GMa$ for a circular orbit (with no perturbation). That implies that this $\Delta\phi \approx -\pi a^3 \rho/M$. That's $\frac{3}{4}$ the ratio of the mass of dust in the orbital sphere to the solar mass, a very plausible result.

  Does friction ever play a role in the motion of objects going around the Sun? Surprisingly, yes. There is an effect discussed in section 9.13 that is unimportant for describing the motion of planets, but is very important for the orbital motion of dust in the solar system. There is no dust you say? Actually there is some, but not much, and the reason there's not much is a relativistic effect described in that section, showing how sunlight exerts a drag on the dust.

  For a major application of these methods see sections 6.12 and 6.13. They describe a calculation of the precession of the orbit of Mercury, and this was the basis of an important test of Einstein's theory of gravity (General Relativity).

6.6 Spherical Pendulum

This is an example of a problem that involves all the apparatus of perturbed orbits, but that doesn't require you to leave Earth to check its validity. Hang a heavy mass from the ceiling and start it moving around a circle parallel to the floor. Finding the period of this motion is an elementary exercise, giving $T = 2\pi\sqrt{L\cos\theta/g}$, where L is the length of the cord.

  That is a conical pendulum, but what if the orbit of the pendulum is not exactly circular? The coordinate system to describe this is of course spherical. You can use $\vec{F} = m\vec{a}$ directly, using the result stated in Eq. (0.42), and this is a perfectly workable method. This problem is still sufficiently simple however, that you can use conservation laws to find the equations of motion. The velocity in spherical coordinates is easy:
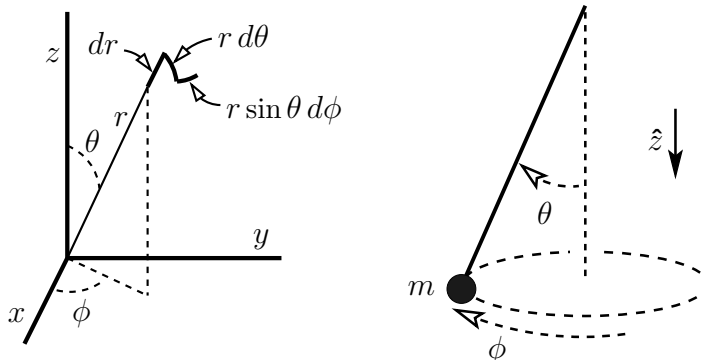

Fig. 6.6

Change the r-coordinate by dr and that is the displacement in the $\hat{r}$-direction. Change the θ-coordinate by $d\theta$ and the displacement in the $\hat{\theta}$-direction is $r\,d\theta$. Change the ϕ-coordinate by $d\phi$ and the displacement in the $\hat{\phi}$-direction is $r\sin\theta\,d\theta$ — note that as ϕ changes the point moves around a small circle of radius $r\sin\theta$. Put these together, divide by dt, and

$$\vec{v} = \frac{d\vec{r}}{dt} = \dot{r}\,\hat{r} + r\dot{\theta}\,\hat{\theta} + r\sin\theta\,\dot{\phi}\,\hat{\phi}$$

Set the z-coordinate to be positive downward and the energy is

$$E = \frac{1}{2}mv^2 + U = \frac{m}{2}\left[\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\right] - mgr\cos\theta \tag{6.37}$$

The z-component of angular momentum is also conserved because there is no torque about that axis. Compute it

$$\vec{r}\times m\vec{v} = mr\hat{r}\times\left[\dot{r}\,\hat{r} + r\dot{\theta}\,\hat{\theta} + r\sin\theta\dot{\phi}\,\hat{\phi}\right] = mr\hat{r}\times\left[r\dot{\theta}\,\hat{\theta} + r\sin\theta\dot{\phi}\,\hat{\phi}\right]$$

The product $\hat{r}\times\hat{\theta}$ is in the $\hat{\phi}$-direction, and that is perpendicular to the z-axis so it doesn't contribute. The other term has $\hat{r}\times\hat{\phi} = -\hat{\theta}$, which does have a z-component. (Remember the definitions of the unit vectors: $\hat{r}$ is in the direction of increasing r; $\hat{\phi}$ is in the direction of increasing ϕ; $\hat{\theta}$ is in the direction of increasing θ.)

$$\hat{z}\cdot\vec{L} = \hat{z}\cdot mr\hat{r}\times r\sin\theta\,\dot{\phi}\,\hat{\phi} = \hat{z}\cdot mrr\sin\theta\,\dot{\phi}\,(-\hat{\theta}) = mr^2\sin^2\theta\,\dot{\phi} \tag{6.38}$$

That is $L_z = mr^2\sin^2\theta\dot{\phi}$. Simplify the expression for the energy by eliminating $\dot{\phi}$ from E.

$$E = \frac{m}{2}\left[\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\frac{L_z^2}{m^2r^4\sin^4\theta}\right] - mgr\cos\theta$$

For the spherical pendulum the length of the cord is constant so $\dot{r} = 0$ and the equation for conservation of energy is $dE/dt = 0$

$$\frac{dE}{dt} = \frac{d}{dt}\left[\frac{m}{2}\left[r^2\dot{\theta}^2 + \frac{L_z^2}{m^2r^2\sin^2\theta}\right] - mgr\cos\theta\right] = 0$$

$$= mr^2\dot{\theta}\ddot{\theta} + \frac{L_z^2}{mr^2}\frac{-\cos\theta}{\sin^3\theta}\dot{\theta} + mgr\sin\theta\,\dot{\theta}$$

Cancel the common factor $\dot{\theta}$ and rearrange the terms to get

$$\ddot{\theta} + \frac{g}{r}\sin\theta - \frac{L_z^2}{m^2r^4}\frac{\cos\theta}{\sin^3\theta} = 0 \qquad \text{with} \qquad \dot{\phi} = \frac{L_z}{mr^2\sin^2\theta} \tag{6.39}$$

These have the same structure as the equations (6.9) and (6.10), which were analyzed starting in Eq. (6.29).

But first, there's a simple special case to check. What if the angular momentum is zero? That is, the pendulum is going back and forth in a single plane. The equation is then

$$\ddot{\theta} + \frac{g}{r}\sin\theta = 0$$

and this is Eq. (3.23), precisely as required.

If the pendulum is swinging in an almost circular orbit that is almost parallel to the floor, use a series expansion about this orbit. Start with the exact, simple solution, θ a constant, so that the orbit is a circle,

$$\frac{g}{r}\sin\theta_0 - \frac{L_z^2}{m^2r^4}\frac{\cos\theta_0}{\sin^3\theta_0} = 0 \qquad \text{then} \qquad \sin^4\theta_0 = \frac{r}{g}\frac{L_z^2}{m^2r^4}\cos\theta_0 \tag{6.40}$$

Eliminate L_z in favor of $\dot\phi_0$, combining (6.40) with the second of the equations (6.39). $\dot\phi_0$ is the angular speed for the circular orbit. It yields

$$\dot\phi_0 = \sqrt{g/r\cos\theta_0} \tag{6.41}$$

This special case of a conical pendulum gives an equation that you can also (and should) derive by elementary methods, problem 4.16.

Now expand the differential equation (6.39) about θ_0. Let $\theta = \theta_0 + \delta$, then

$$\ddot\delta + \frac{g}{r}\sin\left(\theta_0 + \delta\right) - \frac{L_z^2}{m^2r^4}\frac{\cos\left(\theta_0 + \delta\right)}{\sin^3\left(\theta + \delta\right)} = 0$$

Expand in powers of δ up to the first.

$$\ddot\delta + \frac{g}{r}\left[\sin\theta_0 + \delta\cos\theta_0\right] - \frac{L_z^2}{m^2r^4}\frac{\cos\theta_0 - \delta\sin\theta_0}{\sin^3\theta_0 + 3\delta\sin^2\theta_0\cos\theta_0} = 0$$

$$\ddot\delta + \frac{g}{r}\left[\sin\theta_0 + \delta\cos\theta_0\right] - \frac{g\sin^4\theta_0}{r\cos\theta_0}\left[\frac{\cos\theta_0}{\sin^3\theta_0} - \frac{\sin\theta_0}{\sin^3\theta_0}\delta - \frac{\cos\theta_0}{\sin^3\theta_0}3\frac{\cos\theta_0}{\sin\theta_0}\delta\right] = 0$$

$$\ddot\delta + \delta\frac{g}{r}\left[\cos\theta_0 + \frac{\sin^2\theta_0}{\cos\theta_0} + 3\cos\theta_0\right] = 0$$

$$\ddot\delta + \frac{g}{r}\frac{1 + 3\cos^2\theta_0}{\cos\theta_0}\delta = 0$$

This manipulation used the general form of Taylor's series, Eq. (0.4). Then the relation for L_z in Eq. (6.40). Then the binomial expansion. Then some routine manipulation.

Now everything is set up to solve and analyze.

$$\ddot\delta + \frac{g}{r}\frac{1 + 3\cos^2\theta_0}{\cos\theta_0}\delta = 0 \qquad \text{and} \qquad \dot\phi_0 = \sqrt{\frac{g}{r\cos\theta_0}} \tag{6.42}$$

The simplest case first, always. Suppose that the angle θ_0 is small, so that the pendulum is hanging almost vertically, then

$$\ddot\delta + \frac{4g}{r}\delta = 0 \qquad \text{and} \qquad \dot\phi_0 = \sqrt{\frac{g}{r}}$$

The rate of oscillation about the circular orbit is governed by δ:
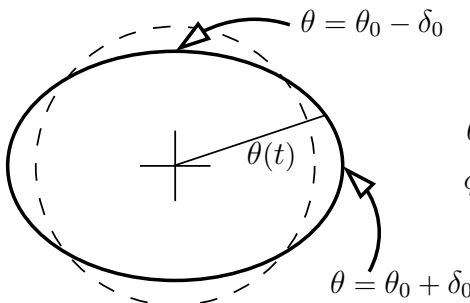
$$\theta(t) = \theta_0 + \delta(t) = \theta_0 + \delta_0\cos\omega t$$
$$\phi(t) = \dot\phi_0 t \qquad \text{with} \qquad \omega = 2\dot\phi_0 \tag{6.43}$$

This is a closed orbit, and is approximately an ellipse. When the angle ϕ goes from 0 to π, the argument of the cosine goes from 0 to π and all the way to 2π, giving two points of maximum excursion. Does it go clockwise or counterclockwise? That depends on how it starts out.

Is there an easy way to see that this is the form the solution must take? Yes, solve it in rectangular coordinates and go back to the middle of section 6.1, Figure 6.2.

What if the mean angle θ_0 is small but not negligible? The key parameter that determines the result is the ratio of the two angular frequencies in Eq. (6.42).

$$\frac{\omega^2}{\dot{\phi}_0^2} = \frac{(1 + 3\cos^2\theta_0)/\cos\theta_0}{1/\cos\theta_0} = 1 + 3\cos^2\theta_0 = 1 + 3\left(1 - \theta_0^2/2 + \cdots\right)^2 = 4 - 3\theta_0^2$$

The ratio of the frequencies is $\omega/\dot{\phi}_0 = 2\sqrt{1 - 3\theta_0^2/4} = 2\left(1 - 3\theta_0^2/8\right)$. This is not a simple ratio of integers, so the orbit doesn't close — it precesses. What happens to the position of the farthest point on the orbit as m goes around once and returns to the farthest point? The radial oscillation is governed by ωt, and when that has oscillated twice what had ϕ done?

When $\omega t = 4\pi$, then angle ϕ changes by

$$\dot{\phi}_0 t = \frac{\dot{\phi}_0}{\omega}\omega t = \frac{1}{2\left(1 - 3\theta_0^2/8\right)} \cdot 4\pi = 2\pi\left(1 + 3\theta_0^2/8\right)$$

The apex of the ellipse has precessed forward by an angle $3\pi\theta_0^2/4$ in one orbit, and this is something that you can check experimentally by hanging a cord from the ceiling. Attach a mass to it and set it going in an *almost* circular path. You can put paper on the floor to mark the changing apex and see how far it goes in ten revolutions or so. Because of friction, the angle θ_0 will decrease slightly during the experiment. It will be a reasonable approximation to use the average of its initial and final values of θ_0 in computing $3\theta_0^2/8$. This forward precession is called "prograde". The opposite is "retrograde".

6.7 Center of Mass Transformation
When there are two bodies interacting, such as the Sun and Jupiter or Earth and the Moon, there are six coordinates to contend with. This can quickly become unmanageable, so it's fortunate that there's a way to transform these six coordinates to two simpler and independent sets of three each. Assume that you have two masses that act on each other; also allow for some external forces.

$$\vec{F}_{\text{on }1} = \vec{F}_{\text{on }1\text{ by }2} + \vec{F}_{\text{on }1,\text{ external}} = m_1 d^2\vec{r}_1/dt^2$$
$$\vec{F}_{\text{on }2} = \vec{F}_{\text{on }2\text{ by }1} + \vec{F}_{\text{on }2,\text{ external}} = m_2 d^2\vec{r}_2/dt^2$$

$$(6.44)$$

Add these equations and use the property that the two mutual forces add to zero by Newton's third law: $\vec{F}_{\text{on }1\text{ by }2} = -\vec{F}_{\text{on }2\text{ by }1}$. This gives

$$\vec{F}_{\text{on }1,\text{ ext}} + \vec{F}_{\text{on }2,\text{ ext}} = m_1\frac{d^2\vec{r}_1}{dt^2} + m_2\frac{d^2\vec{r}_2}{dt^2} = (m_1 + m_2)\frac{d^2}{dt^2}\left(\frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}\right)$$
$$\vec{F}_{\text{total}} = m_{\text{total}}\frac{d^2\vec{r}_{\text{cm}}}{dt^2}$$

$$(6.45)$$

The last parenthesis presents the definition of the center of mass of two objects — of N objects by extending the sums in the numerator and the denominator.

$\vec{r}_{\text{cm}}$ represents three of the six new coordinates. For the other three, that will depend on the nature of the external forces. To make things clearer, start by assuming *no* external forces, $\vec{F}_{\text{on }1,2\text{ ext}} = 0$. Now

in the equations (6.44) divide by m_1 and m_2 respectively, then subtract 1 from 2 and use Newton's third law again.

$$\frac{1}{m_2}\vec{F}_{\text{on 2 by 1}} - \frac{1}{m_1}\vec{F}_{\text{on 1 by 2}} = d^2\vec{r}_2/dt^2 - d^2\vec{r}_1/dt^2$$

$$\left(\frac{1}{m_2} + \frac{1}{m_1}\right)\vec{F}_{\text{on 2 by 1}} = \frac{d^2}{dt^2}(\vec{r}_2 - \vec{r}_1)$$

(6.46)

Rearrange the last equation: multiply by $m_1 m_2$; divide by $(m_1 + m_2)$.

$$\vec{F}_{\text{on 2 by 1}} = \frac{m_1 m_2}{m_1 + m_2}\frac{d^2\vec{r}}{dt^2} \qquad \text{where} \qquad \vec{r} = \vec{r}_2 - \vec{r}_1$$

(6.47)

This is an equation for the other three new coordinates, $\vec{r}$, the relative coordinates.

That this switch from $(\vec{r}_1, \vec{r}_2)$ to $(\vec{r}_{\text{cm}}, \vec{r})$ is just a change of coordinates, and can be reversed if needed, is easy to see with a little algebra; just solve the first line to get the second

$$\vec{r}_{\text{cm}} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2} \qquad\qquad \vec{r} = \vec{r}_2 - \vec{r}_1$$

$$\vec{r}_2 = \vec{r}_{\text{cm}} + \frac{m_1}{m_1 + m_2}\vec{r} \qquad\qquad \vec{r}_1 = \vec{r}_{\text{cm}} - \frac{m_2}{m_1 + m_2}\vec{r}$$

(6.48)
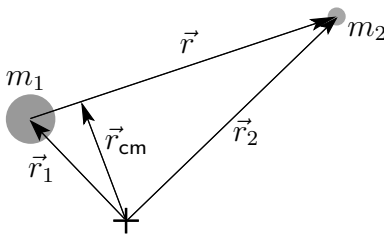
What happens to these equations when $m_1 \gg m_2$? when $m_2 \gg m_1$?

How often do you encounter cases in which the external forces are zero (or at least negligible)? Atoms or molecules in a gas will feel only small average forces from their neighbors, so this change of coordinates will work in analyzing a diatomic molecule or a hydrogen atom. Even when there are external forces, you can sometimes use this transformation anyway. For the case of gravity, the forces are proportional to the masses of the object being affected, and if the external gravitational field is fairly uniform, then

$$\vec{F}_{\text{on 1, ext}} = m_1\vec{g}_{\text{ext}}, \qquad \vec{F}_{\text{on 2, ext}} = m_2\vec{g}_{\text{ext}}$$

(6.49)

with nearly the same external $\vec{g}$ for both. The Sun acts on the Earth and the Moon, and at this distance the Sun's gravitational field at the Earth's position is nearly the same as that at the Moon's position.

Use Eq. (6.49) in Eqs. (6.44), and divide the equations by m_1 and m_2 respectively. Now subtract, and the external force terms will cancel.

$$\frac{1}{m_2}\vec{F}_{\text{on 2 by 1}} + \frac{1}{m_2}m_2\vec{g}_{\text{ext}} - \frac{1}{m_1}\vec{F}_{\text{on 1 by 2}} - \frac{1}{m_1}m_1\vec{g}_{\text{ext}} = d^2\vec{r}_2/dt^2 - d^2\vec{r}_1/dt^2$$

The external forces are gone, and the internal forces are related because of Newton's third law just as in Eqs. (6.46) and (6.47).

$$\vec{F}_{\text{on 2 by 1}} = \frac{m_1 m_2}{m_1 + m_2}\frac{d^2\vec{r}}{dt^2} \qquad \text{where} \qquad \vec{r} = \vec{r}_2 - \vec{r}_1$$

(6.50)

This process took the six coordinates and separated them into three plus three. The center of mass $\vec{r}_{cm}$ obeys equation (6.45) involving the total external force. The relative coordinate $\vec{r}$ obeys the equation (6.50), and it is affected by the internal force alone. Each of these has the appearance of a single mass under a single force, though the masses in the two cases are neither m_1 nor m_2, but

$$m_{total} = m_1 + m_2 \quad \text{and} \quad m_{reduced} = \mu = \frac{m_1 m_2}{m_1 + m_2} \tag{6.51}$$

The *reduced mass* is the combination that takes the place of the ordinary mass in the transformed equation. These two equations are

$$\vec{F}_{tot,ext} = M_{tot}\frac{d^2\vec{r}_{cm}}{dt^2} \quad \text{and} \quad \vec{F}_{on\ 2\ by\ 1} = \mu\frac{d^2\vec{r}}{dt^2} \tag{6.52}$$

where μ is a common notation for the reduced mass, though some prefer to use m_r.

If the mass m_1 is much larger than m_2, this reduced mass is close to m_2 and the correction from treating m_1 as fixed is small—the Sun (m_1) and Earth (m_2) for example, or even the Sun and Jupiter. All the mathematics done in solving the gravitational orbit problem is now the same except for a change in one symbol: the m in "$\vec{F} = m\vec{a}$" becomes μ, but watch out, because the m in GMm/r^2 remains m. Eq. (6.20) for the orbital period becomes (check it out). Look back at section 1.2 and figure out which *types* of mass apply to these two m's.

$$T = \frac{2\pi}{\sqrt{G(M + m)}}a^{3/2} \tag{6.53}$$

The equation for the acceleration of the center of mass is independent of anything else, and the interesting action is in the other, the relative coordinate. When you study some quantum theory, enough to be able to compute the structure of the hydrogen atom, the reduced mass of the nucleus and the electron is what appears in the energy calculations. It is a small correction because the proton's mass is 1836 times that of the electron. Replace the proton by a deuteron (twice the mass), and the energies will change only slightly, though still enough to be detected.

For the Earth-Moon system, $m_{Earth} = 81m_{Moon}$, so the reduced mass $\mu = (81/82)m_{Moon}$. The distance of the center of mass from the origin is $|\vec{r}_{cm} - \vec{r}_1| = m_2 r/(m_1 + m_2)$, and this is about $4700\,\text{km}$, placing it a little more than $1000\,\text{km}$ below the Earth's surface. In a binary star system, two stars orbit each other. If their masses are the same, then the reduced mass is $\mu = m^2/2m = m/2$.

This center of mass change of variables is of great general utility. It appears in the study of atomic and molecular structure. It greatly simplifies calculations in scattering theory. It is necessary in doing orbital calculations of binary stars. The next section uses it in trying to understand how other planetary systems were first discovered.

Is it really correct to apply this transformation to the Earth-Moon system, and to ignore the fact that the Sun's gravitational field is not really uniform? To a good approximation yes, but not for precise results. For those it is impossible to ignore the variation in the Sun's gravity with distance, and the problem is far more involved. It can't fully be reduced to something elementary by a simple transformation such as this one, and you have to go much farther into advanced mechanics to calculate the details of the Moon's orbit. Newton himself made great progress in analyzing the problem, but the whole theory took a couple more centuries to figure out. The name Delaunay is not so well known, but in the middle 1800's his pioneering work established the foundations for modern developments in the subject.

Example
▶ Two separate masses have a combined mass M. They are released at zero velocity and attract each other, finally colliding. For what ratio of the masses will the time until collision be the shortest?

The relative distance, r obeys the equation (6.50) or equivalently, (6.52), so starting from rest and then moving along a straight line, the equation becomes

$$\mu \frac{d^2 r}{dt^2} = -\frac{Gm_1 m_2}{r^2} \longrightarrow \frac{d^2 r}{dt^2} = -\frac{G(m_1 + m_2)}{r^2} = -\frac{GM}{r^2}$$

This acceleration depends on the sum of the masses only, not their separate values. It doesn't make any difference what the mass ratio is; the time is the same.

6.8 Extrasolar Planets
For centuries, since the realization that Earth is a planet, people have speculated about the existence of planets around other stars. Now there is in orbit a telescope with the primary function of scanning for planets around other stars. It has found so many it's hard to keep up.

Transits
To see a planet of another star directly requires luck, and the year 2008 was the first time that someone observed a planet* passing in front of its star, causing the star's light to be dimmed. It is a tiny effect, but now it is being used even by (some very good) amateur astronomers to find new planets. Of course for this to be possible, the plane of the planet's orbit has to allow the planet to pass directly between us and the star, and that is a rare occurrence. Or is it?

If a star has radius R_{st} and a planet of radius r_{pl} is orbiting the star at a distance R_{orbit}, what is the probability that the planet will pass in front of the star? To be specific, ask for the likelihood that the center of the planet passes in front of the star.
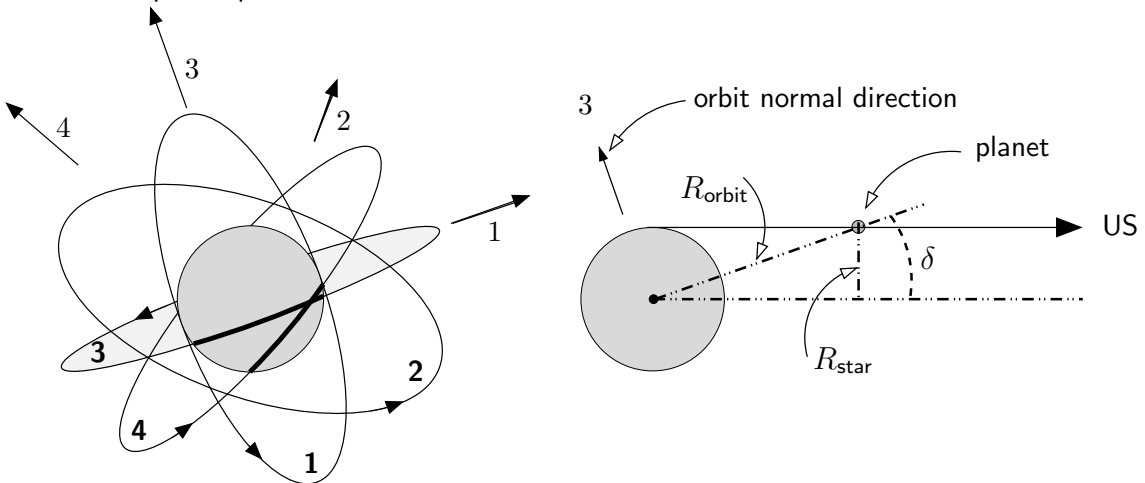
In the left sketch, it shows the star with several possible (circular) planetary orbits. They are all the same distance from the star, but they don't all pass in front of it; only the number 3 and 4 do, with 1 just grazing the edge. The vectors around the outside represent directions *normal* to the respectively numbered orbital planes, and those are the precisely the vectors that will determine just how likely we are to see a transit. From the center of the star, all directions are equally likely, so you expect that these normal directions are randomly and uniformly distributed around the sphere. If such a normal points directly at us, that means that the planetary orbit lies in the plane perpendicular to our line of sight and will appear as a circle, definitely not giving a transit. Only those normals that lie close enough to the plane *perpendicular* to our line of sight will produce transits.

The right hand sketch places us now as observers on the far right side, and only orbits whose normals are within an angle δ, with $\sin\delta = R_{star}/R_{orbit}$, of the perpendicular plane will have their

---

* other than Mercury or Venus

center pass in front of the star. The orbit sketched has its orbital plane seen edge on, so that its normal points toward the upper left and so that the center of the planet just grazes the edge of the star as seen by US. On a sphere, what is the area within an angle δ of the equator as compared to the area of the whole sphere?

$$\int r^2 d\Omega = \int_0^{2\pi} d\phi \int_{-\delta+\pi/2}^{+\delta+\pi/2} d\theta \sin\theta r^2 = 2\pi r^2 \Big[-\cos\theta\Big]_{-\delta+\pi/2}^{+\delta+\pi/2} = 2\pi r^2 \cdot 2\sin\delta$$

The whole area is $4\pi r^2$, so the ratio of these two numbers is the probability of a transit, and that is simply $\sin\delta = R_{\text{star}}/R_{\text{orbit}}$. This is the case in which you ask about the center of the planet passing in front of the star. If you ask about *any* part of the star passing in front, the probability becomes slightly larger.

How big is this number for the Earth? About $700\,000\,\text{km}/150\,000\,000\,\text{km} = 0.0047$, which means that if we use this method and find one planet in an Earth-sized orbit around a Sun-sized star, the probability is that there are more than $1/0.0047 \approx 200$ other such planets that we can't detect this way. The surprise has been just how many planets have been found at all by this method. The easiest planet to observe would be one in a small orbit, so that the probability of its passing in front of the star is much larger. Also, if it does pass in front of the star it will do it frequently — the orbital period is small. The next thing to make it easy to see is if it is a large planet, more like Jupiter than like Earth. Then the drop in the light from the star will be much more, being proportional to the planet's area. The Kepler satellite, now in orbit but out of commission, was specially designed to search for such planets, and it has been a roaring success. Many such planets have now been found, far more than previous theories of planetary formation had predicted. Back to the drawing boards...

The detailed measurement of these transits gives still more information about the orbits, and you cannot assume that they are circles as I have done in this introduction. If there are variations in the times of the transits, it even provides a sensitive method to detect the presence of *other* planets that do not pass in front of the star, but whose gravitational pull affects the ones that you do detect. ("Transit Timing Variation")

Wobbling
The first time that a planet was found around another star was 1995, and it involved a method that has the center of mass transformation at its core. It did *not* involve the stellar dimming caused by a planet passing in front of the star. That transit method came later, and each has its uses.

From Eqs. (6.52), if there is no net external force, then the center of mass will move with constant velocity, and using Eqs. (6.48) you see that the star ($\vec{r}_1$) will wobble because of the pull of the planet. As the relative coordinate $\vec{r}$ represents the planet's orbital motion, the star's motion will wobble by the amount $-m_2\vec{r}/(m_1+m_2)$. Because the planet's mass m_2 will be a lot less than the star's, m_1, the star's motion will be small, but it *can* be detected.

Solve for $\vec{r}_{\text{st}}$ and $\vec{r}_{\text{pl}}$.

$$\vec{r}_{\text{cm}} = \frac{m_{\text{st}}\vec{r}_{\text{st}} + m_{\text{pl}}\vec{r}_{\text{pl}}}{m_{\text{st}} + m_{\text{pl}}} \qquad \text{and} \qquad \vec{r} = \vec{r}_{\text{pl}} - \vec{r}_{\text{st}}$$

$$\implies \vec{r}_{\text{st}} = \vec{r}_{\text{cm}} - \frac{m_{\text{pl}}}{m_{\text{st}} + m_{\text{pl}}}\vec{r} \qquad \text{and} \qquad \vec{r}_{\text{pl}} = \vec{r}_{\text{cm}} + \frac{m_{\text{st}}}{m_{\text{st}} + m_{\text{pl}}}\vec{r}$$

When $m_{\text{st}} \gg m_{\text{pl}}$, the small mass m_{pl} moves on a big ellipse. The star in turn moves on an orbit scaled down by a factor $m_{\text{pl}}/m_{\text{st}}$ from the planet's orbit, and even for a planet as large as Jupiter this is $1/1050$.
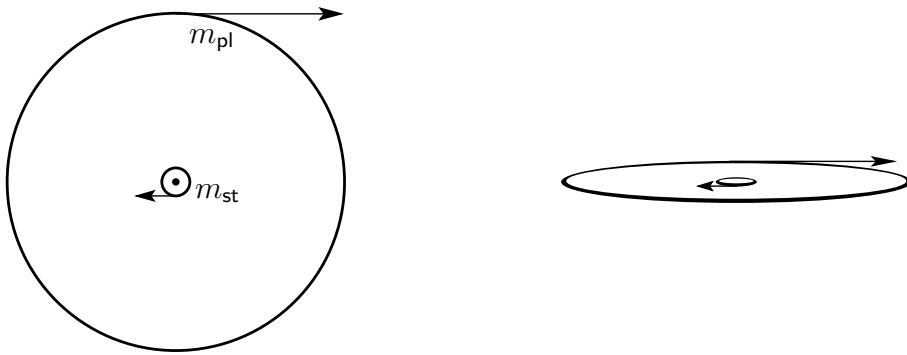
How do you detect the motion of the star? The Doppler effect. As the planet orbits its star, the star's velocity is $\dot{\vec{r}}_{st} = -\dot{\vec{r}} m_{pl}/(m_{st} + m_{pl})$. As a first attack on the problem, start with the simplifying assumption that the orbit is circular. The equation for the motion is then

$$Gm_{st}m_{pl}/r^2 = \mu r\omega^2 \quad \longrightarrow \quad G(m_{st} + m_{pl}) = r^3\omega^2 \tag{6.54}$$

The speed of the star in its own circular orbit is then

$$v_{st} = |\dot{\vec{r}}_{st}| = \frac{m_{pl}}{m_{st} + m_{pl}} r\omega = \frac{m_{pl}}{m_{st} + m_{pl}} \sqrt{\frac{G(m_{st} + m_{pl})}{r}} = m_{pl} \sqrt{\frac{G}{r(m_{st} + m_{pl})}} \tag{6.55}$$

The component of the star's velocity in the direction of the Earth is $v_{st} \cos\omega t \sin\alpha$, where α is the angle the normal to the orbit makes with the direction to Earth. The Doppler equation then says that the wavelength of the light received is related to that emitted as

$$\frac{\Delta\lambda}{\lambda} = \frac{v_{st}}{c} \cos\omega t \sin\alpha \tag{6.56}$$

If you can measure this shift over time, what information can you dig out of it? The orbital frequency ω of course and the product $v_{st} \sin\alpha$. Next, when examining a particular star you already know its spectral classification and so you have a decent estimate of its mass m_{st}, at least to within $5-10\%$. This mass is much larger than a planetary mass, Eq. (6.54) provides the distance r between the star and the planet. Equation (6.55) isn't enough to determine the mass m_{pl} of the planet, because you know only $v_{st} \sin\alpha$, not v_{st}. Multiply this equation by $\sin\alpha$ you can determine $m_{pl} \sin\alpha$, which determines a lower bound on the planetary mass.

What can go wrong? $\Delta\lambda/\lambda$ is the tough part. It is small and difficult to measure, and the techniques to measure it were developed in the 1980's, but not with enough precision to believe the results. Only as recently as 1995 did anyone develop sufficiently reliable techniques to produce measurements that are widely believed. How big is v/c? Calculate it first for our sun and Jupiter:

$$m_{st} = 1.99 \times 10^{30}\,\text{kg}, \quad m_{pl} = 1.9 \times 10^{27}\,\text{kg}, \quad r = 7.78 \times 10^8\,\text{km}$$

$$\longrightarrow \quad \frac{v_{st}}{c} = \frac{1.9 \times 10^{27}}{3 \times 10^8} \sqrt{\frac{6.67 \times 10^{-11}}{7.78 \times 10^8 \cdot 2.0 \times 10^{30}}} = 1.3 \times 10^{-6}$$

$\Delta\lambda/\lambda$ is still smaller by a factor $\sin\alpha$, so this is an impressively difficult quantity to measure, and it is not surprising that the first planets found were larger than Jupiter. Their masses were several times more and their orbital radius was much less than that of Jupiter, a combination that very much surprised astronomers.

What if the orbit isn't circular? It's an ellipse, the planet has varying speed so that the Doppler shift is not simply a cosine as in Eq. (6.56). There is enough information in this time dependence to determine the eccentricity of the orbit, but I will leave that to other references.

6.9 Another Orbit

In section 6.3, Eq. (6.14) you find the equation for the shape of an orbit in a general central force. The Kepler problem has the force $f(r) = -GMm/r^2$, leading to an easily solved equation. There's another case for which the orbital equation is just as easy, though it's more of a curiosity than anything else. The inverse cube force make this differential equation simple, though interpreting the solutions is more effort. Equation (6.14) is

$$ u = \frac{1}{r}, \qquad \ell = r^2\dot{\phi}, \qquad \frac{d^2u}{d\phi^2} + u = -\frac{1}{m\ell^2u^2}f\left(\frac{1}{u}\right) $$

Take the attractive force $f(r) = -m\alpha/r^3$. This is not a physically relevant example, but the solution is interesting anyway.

$$ \frac{d^2u}{d\phi^2} + u = +\frac{1}{m\ell^2u^2}\alpha m u^3 = \frac{\alpha}{\ell^2}u \quad \longrightarrow \quad \frac{d^2u}{d\phi^2} + \left(1 - \frac{\alpha}{\ell^2}\right)u = 0 \qquad (6.57) $$

This is sometimes another harmonic oscillator, but with various forms of solution, depending on the parameters.

$$ \frac{\alpha}{\ell^2} < 1 \qquad \frac{1}{u} = r(\phi) = \frac{r_0}{\cos\left(\beta(\phi - \phi_0)\right)} \qquad \beta = \sqrt{1 - \alpha/\ell^2} $$

$$ \frac{\alpha}{\ell^2} = 1 \qquad \frac{1}{u} = r(\phi) = \frac{r_0}{\phi - \phi_0} $$

$$ \frac{\alpha}{\ell^2} > 1 \qquad \frac{1}{u} = r(\phi) = \frac{r_0}{\cosh\left(\gamma(\phi - \phi_0)\right)} \qquad \gamma = \sqrt{-1 + \alpha/\ell^2} $$

$$ r(\phi) = \frac{r_0}{\sinh\left(\gamma(\phi - \phi_0)\right)} $$

$$ r(\phi) = r_0 e^{\pm\gamma(\phi - \phi_0)} $$

The arbitrary constants can be written in several ways, but these forms make the solutions easier to interpret.

Why no 1/sine? That's the same as 1/cosine with a different phase angle ϕ_0. There's no similar (real) transformation for the hyperbolic functions, but see problem 0.48. Here are a set of plots of these orbits. Sometimes the value of one of the parameters β or γ makes a significant change in the shape, so there are more plots than there are separate functions. It is up to you to figure out which plot goes with which function and very roughly what parameter values are involved. The parameter $\phi_0 = 0$ in all cases, and the scales are not necessarily the same from one sketch to the next, so you can say nothing about r_0.

1.   2.   3.
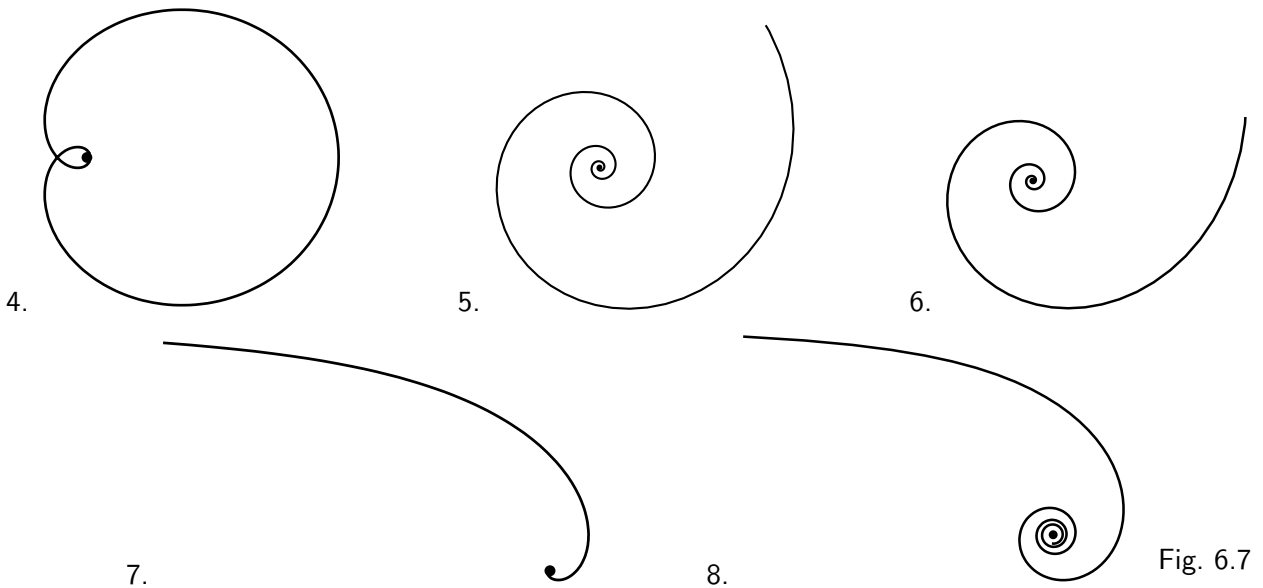
4.   5.   6.
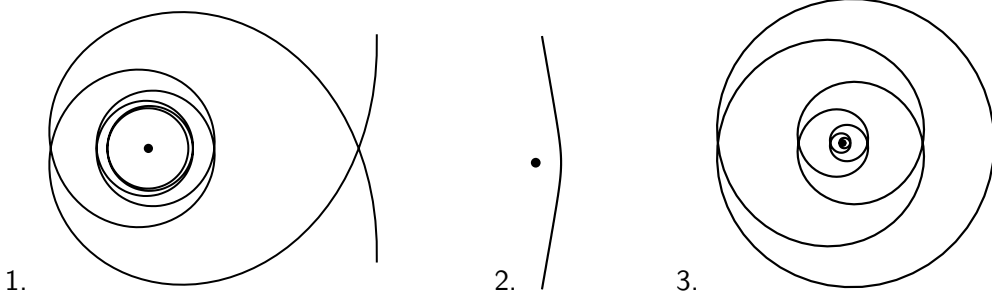
7.   8.   Fig. 6.7

You can see from these pictures that there is no stable orbit. A circular orbit is just zero distance away from an orbit that either crashes into the center or flies away. The inverse cube orbit is what to expect in a universe with four dimensions of space. For a contrast, look at the effective potential energy graph for $F \propto -1/r$ and you will see that *nothing* can escape anything else and all stars would coalesce. This is the potential you get in two dimensions, so three is the magic number. We got lucky. Graph $U_{\text{effective}}$ for the inverse cube force. That will explain a lot.

6.10 Hyperbolic Orbits

In the analysis of the Kepler orbits, equations (6.15) and (6.16), the eccentricity ϵ was less than one. What if it isn't? The polar equation for the ellipse from Eq. (6.17) becomes the equation for a hyperbola, though it needs a sign change when $\epsilon > 1$.

$$ r = \frac{a(\epsilon^2 - 1)}{1 + \epsilon \cos \phi} $$

Now when $\epsilon > 1$ the denominator can vanish. That is for $\phi = \cos^{-1}(-1/\epsilon)$, $\pm$. These angles have $\phi > \pi/2$ and $\phi < -\pi/2$ at the symmetric positions. To verify that this is a hyperbola, put it in the more familiar rectangular form.

$$ r + \epsilon r \cos \phi = a(\epsilon^2 - 1) \longrightarrow \sqrt{x^2 + y^2} + \epsilon x = a(\epsilon^2 - 1) $$
$$ x^2 + y^2 = \left[a(\epsilon^2 - 1) - \epsilon x \right]^2 = a^2(\epsilon^2 - 1)^2 - 2a\epsilon(\epsilon^2 - 1)x + \epsilon^2 x^2 $$
$$ x^2(1 - \epsilon^2) - 2a\epsilon(1 - \epsilon^2)x + (1 - \epsilon^2)a^2\epsilon^2 - (1 - \epsilon^2)a^2\epsilon^2 + y^2 = a^2(\epsilon^2 - 1)^2 $$
$$ (1 - \epsilon^2)(x - a\epsilon)^2 + y^2 = a^2(\epsilon^2 - 1)^2 + a^2(1 - \epsilon^2)\epsilon^2 = -a^2(\epsilon^2 - 1) $$
$$ \frac{(x - \epsilon a)^2}{a^2} - \frac{y^2}{a^2(\epsilon^2 - 1)} = 1 $$
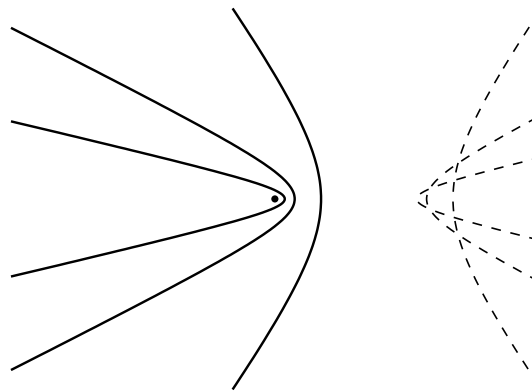
The hyperbola has two branches, but only one applies to the orbit. The other branch is the dashed curve, and it will not be needed until problem 6.56. In order to draw the solid curves easily, parametrize them by

$$x = \epsilon a - a \cosh \alpha, \qquad y = a\sqrt{\epsilon^2 - 1}\,\sinh \alpha \qquad (6.58)$$

For the other (dashed) branch, change the sign in front of the $\cosh \alpha$ in this expression for x. The orbits drawn here have eccentricities 1.15, 1.3, 1.7—which is which? The orbits of comets can be ellipses with very large eccentricity, so that they will return every 76 or 10000 years as the case may be. Cometary orbits can also be hyperbolas such as these, and such a comet will then appear only once, then leave forever. These hyperbolic orbits correspond to the energy E_3 at Eq. (6.24).

6.11 Time Dependence˙
All the calculations concerning the Kepler problem resulted in the shape of the orbit, but they didn't give the position as a function of *time.* The reason I've put that question off is that there is no neat, explicit solution to the equations in terms of time. There is however a solution, just not an explicit one. The calculation is a little intricate, but no single step is too bad.

This procedure will *not* produce equations for r and ϕ as functions of t. Instead you will get three equations that give you r, ϕ, and t as functions of a fourth parameter.

The typical way to solve for $r(t)$ from Eq. (6.10) is to get the energy integral and then to use separation of variables. The procedure here is close to that, but with a twist. The energy integral is Eq. (6.22), and after solving for $\dot{r}$ it is separable, just as in Eq. (2.23). Proceed along this route, and the integral to do for $t(r)$ would be

$$E = \frac{1}{2}m\dot{r}^2 + \frac{m\ell^2}{2r^2} - \frac{GMm}{r} \;\longrightarrow\; dt = \sqrt{\frac{m}{2}}\,dr \Big/ \sqrt{E - \frac{m\ell^2}{2r^2} + \frac{GMm}{r}} \qquad (6.59)$$

To do the integral, multiply the numerator and denominator of the dr integral by r, making the big square root a quadratic polynomial. Now complete the square on that polynomial and integrate. That will yield $t(r)$, not $r(t)$, and you can't invert the resulting expression to get r explicitly. The clever thing to do is to redefine the independent variable t. I find the simplest way to present it follows Sundman* and changes variables in the energy equation.

---

* It is sometimes said that the general solution for the time dependence of three mutually orbiting masses is impossible—the three body problem. Poincaré did prove that it is impossible using the particular methods that he applied, but Sundman showed that there are other ways, demonstrating how to find a series solution in powers of $t^{1/3}$, albeit a very slowly converging one. The n-body solution wasn't found until the 1990's.

Doing this, it is easier to work with the energy equation from the first half of Eq. (6.59) and not the manipulated and separated form that is the second half. Change variables from t to s, where

$$\frac{ds}{dt} = \frac{\gamma}{r} \qquad \text{then} \qquad \frac{dr}{dt} = \frac{dr}{ds}\frac{ds}{dt} = \frac{\gamma}{r}\frac{dr}{ds}. \qquad \text{Let} \quad r' = \frac{dr}{ds}$$

$$E = \frac{m}{2}\frac{\gamma^2}{r^2}(r')^2 - \frac{GMm}{r} + \frac{m\ell^2}{2r^2} \quad \longrightarrow \quad \frac{1}{2}m(r')^2 - \frac{GMm}{\gamma^2}r + \frac{m\ell^2}{2\gamma^2} = \frac{E}{\gamma^2}r^2 \qquad (6.60)$$

This is a harmonic oscillator if $E < 0$. The parameter γ is still arbitrary and I can choose it for convenience. Two plausible choices are $\gamma = 1$ and $\gamma = \sqrt{GM}$, but I will let the choice hang for now because it is just as easy to carry the γ along until you are forced to decide. Is this really a harmonic oscillator? Yes, just complete the square (remember, $E < 0$).

$$\frac{1}{2}m(r')^2 - \frac{E}{\gamma^2}\left(r + \frac{GMm}{2E}\right)^2 = -\frac{m\ell^2}{2\gamma^2} - \frac{G^2M^2m^2}{4E\gamma^2} \qquad (6.61)$$

This is in the same form as the energy conservation equation that for a mass on a spring:

$$\tfrac{1}{2}mv^2 + \tfrac{1}{2}k(x - x_0)^2 = E_0 \implies x(t) = x_0 + A\cos(\omega t + \phi), \quad \omega^2 = k/m, \quad kA^2/2 = E_0$$

With this you can read the frequency ($\sqrt{k/m}$), the center (x_0) and the amplitude of oscillation ($A = \sqrt{2E_0/k}$). The translation table for the present case is then

$$x(t) \to r(s), \quad m \to m, \quad k \to -2E/\gamma^2, \quad x_0 \to -GMm/2E, \quad E_0 \to -\frac{m\ell^2}{2\gamma^2} - \frac{G^2M^2m^2}{4E\gamma^2}$$

Look back again at the picture of an ellipse at Eq. (6.17). The value of r varies from a maximum of $a + f$ down to a minimum of $a - f$. Translate this into the language of a harmonic oscillator and that says $x_0 \to a$ and $A \to f$. The equation for $r(s)$ is then

• r
$$r(s) = a - f\cos\omega(s - s_0) \qquad (6.62)$$

Translate further:

$$\omega^2 = \frac{k}{m} \to -\frac{2E}{m\gamma^2}, \qquad x_0 \to a = -\frac{GMm}{2E},$$

$$A^2 = \frac{2E_0}{k} \to f^2 = \left[-2\frac{m\ell^2}{2\gamma^2} - 2\frac{G^2M^2m^2}{4E\gamma^2}\right]\Big/\left(-2E/\gamma^2\right) = \frac{m\ell^2}{2E} + \frac{(GMm)^2}{4E^2} \qquad (6.63)$$

What about t? Take $s = 0$ when $t = 0$, and for convenience choose $s_0 = 0$.

• t
$$\frac{ds}{dt} = \frac{\gamma}{r} \quad \longrightarrow \quad \gamma dt = r\,ds = (a - f\cos\omega s)\,ds \quad \longrightarrow \quad t(s) = \frac{1}{\gamma}\left[as - \frac{f}{\omega}\sin\omega s\right] \qquad (6.64)$$

$$r^2\dot{\phi} = \ell, \qquad \text{so} \qquad \frac{d\phi}{dt} = \frac{d\phi}{ds}\frac{ds}{dt} = \frac{d\phi}{ds}\frac{\gamma}{r} = \frac{\ell}{r^2} \quad \longrightarrow \quad \phi' = \frac{\ell}{\gamma r}$$

• ϕ
$$\phi(s) = \frac{\ell}{\gamma}\int_0^s ds\,\frac{1}{a - f\cos\omega s} = \frac{2\ell}{\gamma\omega b}\tan^{-1}\left(\frac{b}{a(1-\epsilon)}\tan(\omega s/2)\right) \qquad (6.65)$$

$$= \frac{2\ell}{\gamma\omega b}\tan^{-1}\left(\sqrt{\frac{1+\epsilon}{1-\epsilon}}\tan(\omega s/2)\right) = \frac{2\ell}{\gamma\omega b}\tan^{-1}\left(\cot(\alpha)\tan(\omega s/2)\right)$$

$$= \frac{2\ell}{\gamma\omega b}\sin^{-1}\left(\frac{\cos(\alpha)\sin(\omega s/2)}{\sqrt{\sin^2(\alpha)\cos^2(\omega s/2) + \cos^2(\alpha)\sin^2(\omega s/2)}}\right)$$

Here b is the usual semi-minor axis of the ellipse, $b^2 = a^2 - f^2$ and $f = a\epsilon$ and $\cos 2\alpha = \epsilon$. (Gradshteyn and Ryzhik 2.553.3)
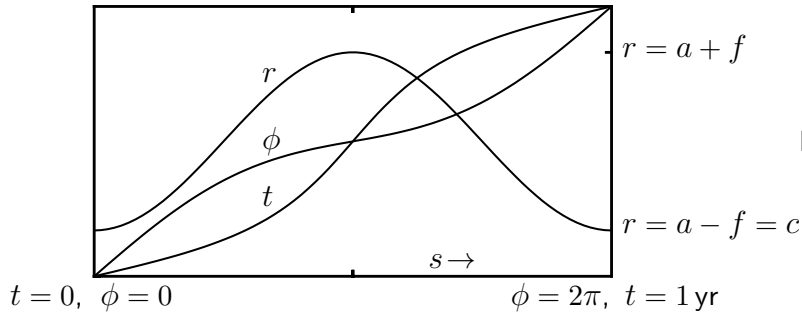

Fig. 6.9

These three graphs show r, ϕ, and t as functions of s for an orbit with eccentricity 0.66 as in the picture at Eq. (6.17). To read and interpret these graphs, take as an example a specific part of the orbit: the perihelion. For minimum $r = a - f$, s is zero—the left end of the graph. What is $d\phi/dt$ at that point?

$$\frac{d\phi}{dt} = \frac{d\phi}{ds}\frac{ds}{dt} = \frac{d\phi}{ds} \Big/ \frac{dt}{ds}$$

At the left (or right) ends of the graph the slope of the t-curve is at its smallest and the slope of the ϕ-curve is at its largest. This gives $\dot{\phi}$ its largest value there, at perihelion, just as Kepler's law says. The planet doesn't spend much time near the perihelion. At the other extreme, the aphelion in the middle, look at the slopes and verify that $\dot{\phi}$ is smallest there.

Why are there so many different versions of the result for $\phi(s)$ in Eq. (6.65)? The answer is that when solving a complicated problem you don't immediately know which form of the result will be the most useful. It is prudent to write it in several different ways until you get a better idea of the route to take. For example, there's a minor technical difficulty when plotting the graph of ϕ if you use the first form. The arctangent is multiple valued, so you have to use some care in evaluating it. If you use any of the first three versions you may need to write the tangent as (sine over cosine) and then to treat the arctangent as a function of two variables. That will let you keep track of which quadrant you are in. The fourth version doesn't have this difficulty. The third version however will have advantages when you are start to look at unbound ($E > 0$) orbits.

Does this complicated looking set of equations reduce to the correct results when the orbit is a circle? $\epsilon = 0$, $b = a$, $f = 0$, so

$$r(s) = -\frac{GMm}{2E} = a, \qquad t(s) = \frac{1}{\gamma}as, \qquad \phi(s) = \frac{\ell}{\gamma}\frac{2}{\omega a}\tan^{-1}\tan(\omega s/2) = \frac{\ell}{\gamma a}s$$

The statement that $A = 0$ (no variation in r) determines ℓ.

$$A^2 = \frac{m\ell^2}{2E} + \frac{(GMm)^2}{4E^2} = 0 \;\longrightarrow\; \ell^2 = -\frac{G^2M^2m}{2E} \;\longrightarrow\; \ell = \frac{GM}{\sqrt{-2E/m}}$$

$$\frac{d\phi}{dt} = \frac{(\ell/\gamma a)ds}{(a/\gamma)ds} = \frac{\ell}{a^2}$$

and that is $(r^2\dot{\phi}/r^2)$ as it should be.

6.12 Perihelion of Mercury[.]

This section is more demanding than most of this text, but it shows the power of the tools that you've studied. The solar system is a whole lot more complicated than anything done here so far, and this section develops one aspect of this complexity.

1. The gravitational forces are not just from the sun.
2. The planets move only *approximately* in closed orbits along ellipses.
3. It *is* possible to figure out the effects that the planets have on each other, at least to a very good approximation. (We can't do everything here.)

The development of perturbed orbits a few sections back has an important application that at first may seem impossible to do: Precession of the perihelion of Mercury. Each planet is pulled by the sun, but it is also pulled by all the other planets, and that will affect the shapes of their orbits. Mercury's orbit in particular was found to have measurable deviations from a simple ellipse. Its orbit precessed by an amount that was large enough to be observed as far back as 1800, and this measurement is associated with a significant milestone in the history of physics because the calculations of the precession and the measurements of the precession did *not* agree.* Accurate measurements of the perihelion's motion were made in the 1850's, and precise methods to calculate the predicted motions had been developed by then. For more than half a century the disagreement remained a vexing problem. Only when Einstein published his theory of gravity (general relativity) in 1916 was the problem resolved. The drawing is greatly exaggerated, both in the eccentricity of Mercury's orbit and in its rate of precession. The dots are near where the perihelion is on these orbits. The angle of precession in one Earth century is
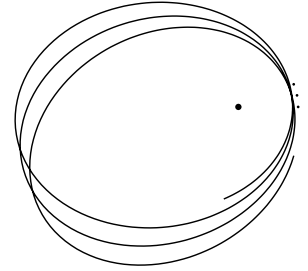
$$532''(\text{Newtonian}) + 43''(\text{general relativity}) = 575''(\text{measured})$$

How did they compute this precession? That's what I want to do here—at least in part. The person who found the discrepancy between theory and observation was Urbain Le Verrier, and you can read something about him in Wikipedia. (He also was one of the two people who independently predicted not only the existence, but the position of Neptune.)

The perturbation methods described so far in this chapter involved making small changes in the potential energy and in the corresponding radial force, $F_r = -dU(r)/dr$. In all cases the small change involved adding something that was spherically symmetric. All the changes to $U(r)$ added another small term $U_1(r)$, and whether it was a cloud of dust in the solar system or a bulge at the Sun's equator it was symmetric. The pull on Mercury by the planets Venus or Jupiter is not at all symmetric; their attractions are not toward the Sun and they are very much time-dependent. So now what?

There is a clever trick due to Gauss. He noted that planets orbit the Sun in periods that are measured in years— $1/4$ year in the case of Mercury, 165 years for Neptune. The measured precession rate of Mercury is only about $0.16°$ in one Earth century, and this is an extremely slow motion compared to that of any planet, so why not smear out and *average* the positions of the planets? They're whizzing around so (comparatively) fast that their detailed position as a function of time shouldn't matter, only their orbit. The average in this case means that you replace the orbit of a given planet by a ring of uniform linear mass density, then the effect of such a ring on the motion of Mercury will be symmetric. Here I will make the pretty good simplifying approximation that all the planetary orbits (except Mercury) are circles and all of them orbit in exactly the same plane as Mercury. Refinements beyond *this* are more difficult and I'll save mention of them until the end. It is quite amazing however just how close this simplified model comes (about 4%).

---

* Close, but no cigar.

This is now, at least in principle, the same sort of perturbation problem developed earlier in this chapter. The only initial difficulty is in figuring out the field of a ring of mass. The answer for the potential energy is

$$U_1(r) = -\frac{GMm}{2\pi} \int_0^{2\pi} d\phi \frac{1}{\sqrt{R^2 + r^2 - 2Rr\cos\phi}}, \qquad V_1 = \frac{U_1}{m} \qquad (6.66)$$

M is the mass of the outer planet; R is its orbital radius; m is Mercury's mass, U is the potential energy, and V is the gravitational potential. This r is in the plane of the ring, and the denominator in the integral comes from using the law of cosines. Don't be intimidated. This sort of calculation becomes fairly routine once you've taken one E&M course beyond the first. ($dM = M\,d\phi/2\pi$)

As a check on this expression for the potential, what is it at $r = 0$?*

$$V_1(0) = -\frac{GM}{2\pi} \int_0^{2\pi} d\phi \frac{1}{\sqrt{R^2}} = -\frac{GM}{2\pi}\frac{2\pi}{R} = -\frac{GM}{R}$$

and since the entire mass M is at the constant distance R, this is correct. Rather than spending time discussing how to evaluate this integral, I will state the result for its lowest order power series representation. At the end of this section there is a more complete result. Out to terms in r^2 this potential is (see problem 6.72 for how to do this.)

$$V_1(r) \approx -\frac{GM}{R}\left(1 + \frac{r^2}{4R^2}\right), \qquad U_1(r) = mV_1(r), \quad F_r(r) = -\frac{dU_1}{dr} \approx +\frac{GMm}{2R^3}r \qquad (6.67)$$

The next pictures shows a view from above the plane of the planet's orbit on the right and a side view of the orbit on the left. It also shows a graph of $V(r)$, and you can see both its curved (r^2) shape near $r = 0$ and its behavior near $r = R$. I chose to label it as the orbit of Venus, but it could as easily be any of the other planets farther out.

Fig. 6.10
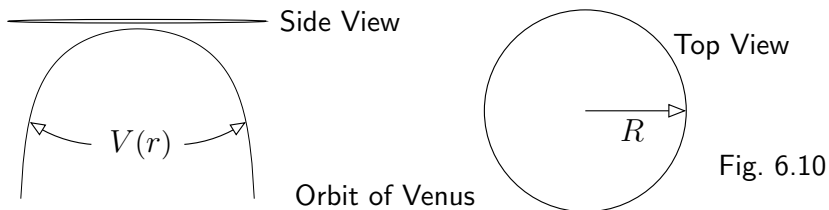
The radius of Mercury's orbit is about one-half of Venus's, so you can see roughly where in this graph to put Mercury. It also means that the behavior of V near $r = R$ will not be relevant to the problem.

Now for the effect of this extra force on the planet Mercury, and this takes us back to section 6.5. The core equation is Eq. (6.14), where $u = 1/r$ and $M_\odot$ is the solar mass. Also, $f = -dU/dr$ and leads to Eq. (6.35).

$$f(r) = -\frac{GM_\odot m}{r^2} + \frac{GMm}{2R^3}r = -GM_\odot mu^2 + \frac{GMm}{2R^3 u}$$

$$\frac{d^2u}{d\phi^2} + u = -\frac{1}{m\ell^2 u^2}f\left(\frac{1}{u}\right) = \frac{GM_\odot}{\ell^2} - \frac{GM}{2\ell^2 R^3 u^3}$$

* and what is it for $r \gg R$?

Follow the same procedure as there, looking for oscillations about a circular orbit: $u(\phi) = u_0 + x(\phi)$, with $x \ll u_0 = 1/r_0$, and keeping terms only to first order in x.

$$\frac{d^2(u_0 + x)}{d\phi^2} + u_0 + x = \frac{GM_\odot}{\ell^2} - \frac{GM}{2R^3(u_0 + x)^3}$$

$$\frac{d^2x}{d\phi^2} + u_0 + x = \frac{GM_\odot}{\ell^2} - \frac{GM}{2\ell^2 R^3 u_0^3}\left(1 - 3\frac{x}{u_0}\right)$$

$$\longrightarrow \quad u_0 = \frac{GM_\odot}{\ell^2} - \frac{GM}{2\ell^2 R^3 u_0^3} \quad \text{and} \quad \frac{d^2x}{d\phi^2} + x = +\frac{GM}{2\ell^2 R^3 u_0^3}\frac{3x}{u_0}$$

The solution for x is then

$$x(\phi) = x_0 \cos\omega\phi, \qquad \text{where} \qquad \omega^2 = 1 - \frac{3GM}{2\ell^2 R^3 u_0^4}$$

Remember that this is a first order expansion, so in putting u_0 into this ω to evaluate it, keep terms to a consistent order. The equation for u_0 has the GM term as a small correction to the first term, so when you use this to evaluate ω, it would be inconsistent to keep it. Use $\ell^2 = GM_\odot r_0$ to simplify the equations.

$$\frac{1}{r_0} = u_0 \approx \frac{GM_\odot}{\ell^2} \quad \longrightarrow \quad \omega^2 = 1 - \frac{3}{2}\frac{GM}{u_0}\frac{1}{\ell^2 R^3 u_0^3} \approx 1 - \frac{3}{2}\frac{M}{M_\odot}\frac{r_0^3}{R^3}$$

$$r = \frac{1}{u} = \frac{1}{u_0 + x} = \frac{1}{u_0}(1 - \frac{x}{u_0}) = r_0 - r_0^2 x = r_0 - r_0^2 x_0 \cos\omega\phi \tag{6.68}$$

$$\text{with} \qquad \omega = 1 - \frac{3}{4}\frac{M}{M_\odot}\frac{r_0^3}{R^3} \tag{6.69}$$

This is what we now have to analyze, putting in the numbers and seeing what it predicts.

First: $\omega < 1$, so in order that $\omega\phi = 2\pi$, the angle ϕ must be greater than 2π. Mercury doesn't get back to the perihelion until it has gone slightly more than once around the Sun. This is called *prograde* precession, the opposite of *retrograde*.

Second: How much more than once around? $\omega \cdot (2\pi + \delta) = 2\pi$, or to this order

$$\left(1 - \frac{3}{4}\frac{M}{M_\odot}\frac{r_0^3}{R^3}\right)(2\pi + \delta) = 2\pi \quad \longrightarrow \quad \delta = +\frac{3}{4}\frac{M}{M_\odot}\frac{r_0^3}{R^3}\cdot 2\pi \tag{6.70}$$

Third: What is r_0? Mercury's path is really an ellipse, recall. From Eq. (6.68) its perihelion is at a radius $r_0 - r_0^2 x_0$ and its aphelion (at $\phi = \pi$) is at $r_0 + r_0^2 x_0$. That means that $2r_0$ is the whole major axis and that r_0 is then the semi-major axis, a.

Each of the outer planets will contribute to the yearly advance of the perihelion of Mercury an amount given by Eq. (6.70).

$$\delta_{\text{total}} = \sum_{k=\text{Venus}}^{k=\text{Neptune}} \frac{3}{4}\frac{M_k}{M_\odot}\frac{a^3}{R_k^3}\cdot 2\pi \tag{6.71}$$

	R/R_E	M/M_E	
Mercury	0.38710	0.05527	
Venus	0.72334	0.8150	1.767×10^{-6}
Earth	1.0	$1.0 = 5.9736 \times 10^{24}$kg	8.3095×10^{-7}
Moon		0.0123	
Mars	1.52371	0.10745	2.4933×10^{-8}
Jupiter	5.2029	317.83	1.85246×10^{-6}
Saturn	9.537	95.159	9.005×10^{-8}
Uranus	19.189	14.500	1.685×10^{-9}
Neptune	30.061	17.204	5.204×10^{-10}
Pluto	39.482	0.0022	
Sun		$333\,000. = 1.989 \times 10^{30}$kg $= M_\odot$	

Should some of the larger moons be included? Ganymede, a moon of Jupiter, is 0.025 the mass of Earth, but our own moon is half the mass of Ganymede and it's five times closer to the Sun, so it will count for much more, and it is still a small correction.

This table is expressed as ratios to the orbital distance and the mass of Earth, and because Eq. (6.71) involves only ratios, that is all that's needed.

$$
\frac{3\pi}{2} \frac{a^3}{M_\odot} \sum_{k=\text{Venus}}^{k=\text{Neptune}} \frac{M_k}{R_k^3} = \frac{3\pi}{2} \frac{0.3871^3}{333000} \left[\frac{0.815}{0.72334^3} + \frac{1.0123}{1^3} + \frac{0.1074}{1.5237^3} + \frac{317.83}{5.2029^3} + \cdots \right]
$$

$$
= 4.5686 \times 10^{-6} \text{ radians per orbit of Mercury}
$$

$$
= 0.9423'' \text{ per orbit} \times 365.24 \times 100/87.97
$$

$$
= 391'' \text{ per Earth century} = 0.11° \text{ per Earth century}
$$

The right hand column in the table above is made up of the successive terms in this series. You can see that the order of influence is Jupiter, Venus, Earth, with the other planets well behind. Go to higher orders in the calculation of $U_1(r)$, keeping terms in r^4/R^4, and Venus will take and keep the lead. This estimate is not a bad start, but $0.11°$ as compared to $0.16°$ is not yet good enough.

It *is* possible to calculate this precession without making the approximation in Eq. (6.67), that the potential is proportional to $(1+r^2/4R^2)$. The methods described in the next section allow manipulation of the potential in Eq. (6.66) exactly, or if you prefer, there is a more familiar type of method available: Use a series representation from Eqs. (6.73) and (6.74). That is

$$
V_1(r) = \frac{U_1(r)}{m} = -\frac{GM}{R} \left[1 + \left(\frac{1}{2} \right)^2 \left(\frac{r}{R} \right)^2 + \left(\frac{1 \cdot 3}{2 \cdot 4} \right)^2 \left(\frac{r}{R} \right)^4 + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \right)^2 \left(\frac{r}{R} \right)^6 + \cdots \right]
$$

Use each successive term in this series to compute its effect on Mercury, following exactly the same procedure as for the r^2 term. You will need a lot of terms to get a really good result, which is why, in the next section, I want to show a less familiar method that avoids this series.

6.13 Elliptic Integrals˙
I'm including this section hoping to persuade you that if you encounter an integral or a function that you've never heard of before, you shouldn't be intimidated. If you see a "sine integral" when studying optics then you should be ready to look up what it is. If you study electromagnetism or any of dozens of other topics then don't be surprised by Bessel functions, there are books, tables, and computer programs to deal with them. Gamma functions too are ubiquitous, and they are easy. Legendre polynomials are easy too, and you've already run into them here in Eq. (5.36). Legendre *functions* can be a nuisance,

but there are tools readily available to handle them. If however you should encounter a Mathieu function in a dark alley, you will have my sympathy.

Here are two commonly used functions, K and E: the "complete elliptic integrals of the first and second kinds". Along with their definitions, here are some useful formulas straight out of standard references.

$$K(k) = \int_0^{\pi/2} d\theta \frac{1}{\sqrt{1 - k^2 \sin^2 \theta}} = \int_0^1 dx \frac{1}{\sqrt{(1 - x^2)(1 - k^2 x^2)}} = \frac{1}{2} \int_0^\pi d\phi \frac{1}{\sqrt{1 + k^2 - 2k \cos \phi}}$$

$$E(k) = \int_0^{\pi/2} d\theta \sqrt{1 - k^2 \sin^2 \theta} = \int_0^1 dx \sqrt{\frac{1 - k^2 x^2}{1 - x^2}} \tag{6.72}$$
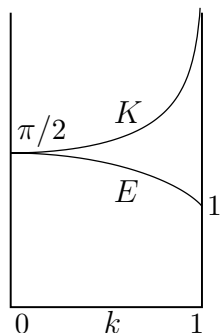
$$\frac{dK}{dk} = \frac{E(k)}{k(1 - k^2)} - \frac{K(k)}{k} \qquad \frac{dE}{dk} = \frac{1}{k}(E(k) - K(k))$$

$$K(k) = \frac{\pi}{2} \left[1 + \left(\frac{1}{2}\right)^2 k^2 + \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 k^4 + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 k^6 + \cdots \right] \tag{6.73}$$

$$E(k) = \frac{\pi}{2} \left[1 - \left(\frac{1}{2}\right)^2 k^2 - \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 \frac{k^4}{3} - \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 \frac{k^6}{5} - \cdots \right]$$

From the two equations for dK/dk and dE/dk you see that higher derivatives of K and E can also be expressed in terms of the original K and E functions simply by differentiating these derivatives repeatedly — sort of like the sine and cosine. Equation (6.66) calls for the first of these functions K, but the perturbation equations that follow it will require the first and second derivatives of K.

$$U_1(r) = -\frac{GMm}{2\pi} \int_0^{2\pi} d\phi \frac{1}{\sqrt{R^2 + r^2 - 2Rr \cos \phi}}$$

$$= -\frac{GMm}{2\pi R} \int_0^{2\pi} d\phi \frac{1}{\sqrt{1 + (r^2/R^2) - 2(r/R) \cos \phi}}$$

$$= -\frac{2GMm}{\pi R} K(r/R) \tag{6.74}$$

A reason not to shy from using these integrals is that there are some very fast and efficient ways to evaluate them, and those methods are much faster then a straight numerical integration or a series expansion would be. Also, what we need now is to use Eq. (6.14), which requires knowing only f and f', and those come from dU/dr and d^2U/dr^2, and those in turn require evaluating K and E only at a/R_k for several planets. This is really not too much effort for a very important result. And why are these functions called "elliptic"? See problem 6.73 to find out.

The graph above shows the behavior of K and E in the domain $0 \leq k < 1$ and they are obviously even functions. The left end of the graph at $k = 0$ is easy to evaluate from the equations (6.72) using the $d\theta$ integrals. At the other end, at $k = 1$, the dx form make the calculation of E easy, but if you look for K at the same point you will see that the integral diverges. It's a mild divergence, being only logarithmic, and the behavior of K near $k = 1$ is $K(k) \sim \frac{1}{2} \ln (8/(1 - k))$. The second derivatives of K and E come by differentiating the expressions at Eq. (6.73) that give their first derivatives. It is tedious but straight forward, and the results are

$$\frac{d^2 E}{dk^2} = \frac{-E + (1 - k^2)K}{k^2(1 - k^2)}, \qquad \frac{d^2 K}{dk^2} = \frac{(-1 + 3k^2)E + (1 - 2k^2)(1 - k^2)K}{k^2(1 - k^2)^2}$$

Now return to the perturbed orbital equations. $u = 1/r$, and M and R are the other planet's mass and orbital radius. Use the method that starts with Eq. (6.33):

$$\frac{d^2u}{d\phi^2} + u = -\frac{1}{m\ell^2 u^2}f\left(\frac{1}{u}\right)$$

$$\text{with} \quad f(r) = -\frac{d}{dr}U(r) = -\frac{d}{dr}\left[\frac{-GM_\odot m}{r} - \frac{2GM\,m}{\pi R}K(r/R)\right]$$

$$= \frac{-GM_\odot m}{r^2} + \frac{2GM\,m}{\pi R^2}K'(r/R)$$

Using the same $u = u_0 + x$ with $x(\phi)$ being the perturbation from a circular orbit,

$$\frac{d^2u}{d\phi^2} + u = \frac{GM_\odot}{\ell^2} - \frac{2GM}{\pi R^2\ell^2 u^2}K'(1/Ru)$$

$$\frac{d^2(u_0 + x)}{d\phi^2} + u_0 + x = \frac{GM_\odot}{\ell^2} - \frac{2GM}{\pi R^2\ell^2(u_0+x)^2}K'\left(\frac{1}{R(u_0+x)}\right)$$

$$\frac{d^2x}{d\phi^2} + x = -u_0 + \frac{GM_\odot}{\ell^2} - \frac{2GM}{\pi R^2\ell^2 u_0^2}\left(1 - \frac{2x}{u_0}\right)K'\left(\frac{1}{Ru_0}\left(1 - \frac{x}{u_0}\right)\right)$$

$$= -u_0 + \frac{GM_\odot}{\ell^2} - \frac{2GM}{\pi R^2\ell^2 u_0^2}\left(1 - \frac{2x}{u_0}\right)\left[K'\left(\frac{1}{Ru_0}\right) - K''\left(\frac{1}{Ru_0}\right)\frac{x}{Ru_0^2}\right]$$

The terms without "x" on the right of the final equation define the circular orbit u_0. The terms *with* an x are the ones that contribute to the precession.

$$\frac{d^2x}{d\phi^2} + x\left[1 - \frac{2GM}{\pi R^2\ell^2 u_0^2}\left[K'\left(\frac{1}{Ru_0}\right)\frac{2}{u_0} + K''\left(\frac{1}{Ru_0}\right)\frac{1}{Ru_0^2}\right]\right] = 0$$

$$\omega^2 = 1 - \frac{2GMr_0^2}{\pi R^2\ell^2}\left[2r_0 K'\left(\frac{r_0}{R}\right) + \frac{r_0^2}{R}K''\left(\frac{r_0}{R}\right)\right]$$

Simplify this using $r_0 = a$ and $\ell^2 = GM_\odot a$ to this order, then

$$\omega = 1 - \frac{Ma^2}{\pi M_\odot R^2}\left[2K'\left(\frac{a}{R}\right) + \frac{a}{R}K''\left(\frac{a}{R}\right)\right]$$

The first check: Does this agree with the first order calculation of Eq. (6.69)? I think that it does, but you'd better look, using the power series for K. As in the previous, lowest order calculation the precession comes from $\omega\cdot(2\pi + \delta) = 2\pi$. Then

$$\delta = +\frac{2Ma^2}{M_\odot R^2}\left[2K'\left(\frac{a}{R}\right) + \frac{a}{R}K''\left(\frac{a}{R}\right)\right] \tag{6.75}$$

To check the sign, look at the graph of K on the preceding page. What are the signs of K' and of K'', and does that match the sign for δ in Eq. (6.70)? To include the effects from several planets, use the same notation as in Eq. (6.71):

$$\delta = \sum_{k=\text{Venus}}^{k=\text{Neptune}} \frac{2M_k a^2}{M_\odot R_k^2}\left[2K'\left(\frac{a}{R_k}\right) + \frac{a}{R_k}K''\left(\frac{a}{R_k}\right)\right] \tag{6.76}$$

$$= 553.4'' \text{ per Earth century}$$

This is about 4% larger than the result usually quoted for this effect $(531.6'')$. One error in the present approximation is that the plane of Mercury's orbit is $7°$ off from the plane of most other planets. Another is that its orbit is more eccentric than the other planets, $\epsilon = 0.21$, so that assuming a small deviation from a circle may not be good enough. Also, the orbits of the other planets are not circles, and the limitations of replacing the moving planet by a *uniform* ring of mass must be reexamined because planetary speeds vary over an orbit, making the effective mass density bigger where it moves more slowly. All of these contribute to the effect and account for the change from 553 to 531 seconds. You will have to find those calculations elsewhere, as they go well beyond what you see here.

Computing elliptic integrals

This uses the "arithmetic-geometric mean" method, and if you think that it's not obvious why it works, you're right. It involves making some incredibly clever changes of variable in the original integrals.*

$n = 0$

0. $a_0 = 1$ $\qquad\qquad b_0 = \left(1 - k^2\right)^{1/2}$ $\qquad c_0 = k$ $\qquad\qquad d = c_0^2$

1. $a_{n+1} = \frac{1}{2}(a_n + b_n)$ $\quad b_{n+1} = \left(a_n b_n\right)^{1/2}$ $\quad c_{n+1} = \frac{1}{2}(a_n - b_n)$ $\quad d = d + c_{n+1}^2 \cdot 2^{n+1}$

2. if c_{n+1} is small enough then stop, else increment n and repeat step #1

$$K(k) = \pi/(2a_{n+1}), \qquad E(k) = K(k)\cdot(1 - d/2)$$

This process converges very fast. For example, if $k = 1/2$, $b_0 = \sqrt{3/4}$ and

0. $a_0 = 1$ $\qquad\qquad b_0 = 0.86602540$ $\qquad c_0 = 0.5$ $\qquad\qquad d = 0.25$
1. $a_1 = 0.933012702$ $\quad b_1 = 0.930604859$ $\quad c_1 = 0.066987298$ $\quad d = 0.258974600$
2. $a_2 = 0.931808780$ $\quad b_2 = 0.931808003$ $\quad c_2 = 0.001203921$ $\quad d = 0.258980394$
3. $a_3 = 0.931808392$ $\quad b_3 = 0.931808392$ $\quad c_3 = 0.00000039$ $\quad d = 0.25898039$

$$K(0.5) = \frac{\pi}{2a_3} = 1.685750354812596 \qquad E(0.5) = 1.467462209339427$$

These numbers were really computed with almost twice the number of digits shown, and it took no more steps. If all that you want is K, then you don't need either c_n or d. Here, however, computing K' and K'' involves E, so you need both.

Exercises

1 Sketch a Lissajous figure for the case $\omega_2 = 10\omega_1$.

2 For the two-dimensional harmonic oscillator of Eqs. (6.1)–(6.3), what is the angular momentum? Is it constant? *I.e.* take $z = 0$ here.

3 From the information given at and just after Eq. (6.17), what is the ratio of the semi-minor axis to the semi-major axis for Earth and for Mercury? What is the ratio of the perihelion distance to the aphelion distance in the same two cases?

4 What figure results if the sum of the squares of the distances from two points add to a constant?

5 For the thinnest ellipse at Eq. (6.25), what are b and c?

---

* See the Wikipedia entry for arithmetic-geometric mean.

6 Communications satellites are placed in orbits above the Earth's equator so that relative to the Earth, they stay in a fixed position. How far are they from the Earth's center and from the Earth's surface?

7 A mass is in orbit about a fixed center, and the attractive force law is such that it can orbit in a circle at radius r_0 from that center. The angular speed is $\dot\phi_0 = \omega_0$. The orbit is slightly perturbed, so that the mass orbits in a nearly circular orbit and the radius oscillates about the circle with frequency $3\omega_0/2$. Draw the resulting orbit and explain the reasons that your drawing appears as it does.

8 What is the ratio of the solar power reaching Earth at its perihelion (January) to that at its aphelion (July)? Now what about Mercury and Mars?

9 Check the dimensions of the equations (6.27) and (6.28).

10 In the set of ellipses at Eq. (6.25), if the circle represents Earth's orbit, radius 150×10^6 km, what is the perihelion distance for the most eccentric of these orbits? And what is its semi-minor axis? In the same picture, where is the orbit of Mars?

11 In looking for planetary transits of another star as on page 206, what do the pictures look like in the extreme cases that $\sin\delta \ll 1$ and the case that $\sin\delta$ is almost equal to one.

12 The expression $L_z = mr^2 \sin^2\theta\dot\phi$ follows Eq. (6.38). Show how to derive it in one or two lines.

13 Jupiter's orbital period is about 12 years. Orbiting at the same distance, what would Jupiter's period be if the Jupiter had the same mass as the Sun? Or if the Sun had the same mass as Jupiter?

14 Graph $U_{\text{effective}}(r)$ for forces proportional to $-1/r$, $-1/r^2$, and $-1/r^3$.

15 User the first terms of the series expansion for K to check that Eq. (6.76) produces Eq. (6.71).

not **All About Ellipses**

16 If you remember that the polar equation for an ellipse is something like $r = K/(1 + \epsilon\cos\phi)$, but you don't remember what K is, evaluate this expression at the two ends of the major axis, at $\phi = 0, \pi$, and add the results to solve for K. For the answer, see the first equation in Eq. (6.17).

17 Where does the second of the equations in Eq. (6.17) come from? Imagine two equal point masses at the two ends of the major axis, and compute their center of mass, measured from the origin. Use the equation from the immediately preceding exercise.

18 The third equation in Eq. (6.17) tells you b. How so? The length of the semi-minor axis is the maximum y-coordinate on the ellipse, so maximize $y = r\sin\phi$ by taking $dy/d\phi = 0$.

19 The fourth equation in Eq. (6.17) follows from the two preceding exercises.

20 The fifth equation in Eq. (6.17) is just a little more involved. Use the same figure as with the four preceding exercises, but now express everything in terms of vectors. $\vec{r}' = \vec{r} + \vec{F}$, and $\vec{F} = 2\vec{f}$. To relate the lengths, use $\vec{r}' \cdot \vec{r}' = (\vec{F} - \vec{r}) \cdot (\vec{F} - \vec{r})$ and eliminate $r\cos\phi$ by using the preceding exercises.

21 The sixth equation is problem 0.49. If you haven't done it, now would be a good time.
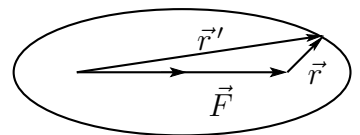
Problems

6.1 In Eq. (6.10), if the force f is zero, the motion should somehow represent a straight line. Take the case as represented in rectangular coordinates by $x = x_0$ and $y = v_0t$; translate it into an expression for $r(t)$, and show that it works.

6.2 The Lissajous curve in Eq. (6.5) can be expressed as a polynomial equation in x and y. Do so.

6.3˙ Referring to the equations (6.3) and to the picture, Figure 6.2, show that the angle by which the third ellipse is rotated is ϕ, where $\tan 2\phi = A_2^2 \sin 2\delta / (A_1^2 - A_2^2 \cos 2\delta)$. If your answer comes out in a different form, then check a few numerical values of the parameters to see if the two expressions agree.

6.4 Transform Eq. (6.12) into the form Eq. (6.13) to show that this is the polar form of an ellipse. Not really. Find the conditions on A, B, and C so that it *is* an ellipse. What shapes do you get if these conditions are violated? Catalog and sketch the possibilities, including $B = C$.

6.5 (**a**) For a planet in circular orbit around the sun, what is the planet's speed as a function of distance from the sun? (**b**) A year on Mars is very close to two Earth years. One astronomical unit is (very close to) Earth's mean distance from the sun; How many AU is Mars from the sun?

6.6 In Eq. (6.14), let the force be zero. Solve, and show what the shape of the orbit is.

6.7 If r_1 and r_2 are the distances of perihelion and aphelion for a planet, what are their arithmetic and their geometric means? Ans: a and b

▶ If you didn't do problems 4.16 and 4.17, now would be a good time. Also problem 4.14.

6.8 The Earth-Sun distance is 150 Gm. Start from elementary principles, commonly known data, and the value of G to deduce the mass of the Sun.

6.9 A particle in a central force moves in a spiral orbit given by $r = a\phi$. Find the force and find the time dependence of ϕ, *i.e.* , $\phi(t)$.

6.10 Improve on the result calculated in Eq. (6.11) by using the fact that the orbit of Halley's comet has eccentricity $\epsilon = 0.96714$.

6.11 Start from Eq. (6.35) and examine the case of the Kepler problem. Plot the polar coordinate orbit $r(\phi)$ and compare the result to the exact solution.

6.12 The force on an orbiting planet of mass m is given as the sum of two terms: One is a central force; the other is a velocity dependent friction term.

$$\vec{F}_1 = -F_0 \operatorname{sech}^3(\alpha r) K_3(\beta r)\hat{r}, \qquad \vec{F}_2 = -\gamma \vec{v}$$

where sech is the hyperbolic secant and K_3 is the Bessel function of the third kind with imaginary argument. The angular momentum about the origin is given at time $t = 0$ to be $\vec{L}_0$. Find the angular momentum at later times.

6.13 The force on a mass m is given to be $\vec{F} = k\hat{r}/r^3$. The mass is given an initial speed v_0 and is aimed so that in the absence of forces, it will miss the center of the repulsive force by a distance b (the impact parameter). Find the distance of closest approach to the origin and what is its speed when it reaches that point? (Don't solve for the orbit unless you need the extra work.) Ans: $\sqrt{b^2 + k/mv_0^2}$

6.14 An object starts at a negligibly small speed from a long distance away from the sun, falling straight into the sun. (**a**) As it passes distance to the sun equal to the radius of the Earth's orbit, how much time does it have before it hits the sun? Express the answer as a ratio to the period of Earth's orbit, and then express it in weeks or days. (**b**) And since you've done the labor already, go back to the exercise #4 on page 179 to find how much time it would take from the onset of spaghettification to hit the center of attraction in that case, and with the same assumption of falling from a great distance. Ans: (**a**) almost a month and (**b**) about 0.3 second.

6.15 If the Earth stopped still with respect to the sun, how much time would it take to fall into the sun? Set up the equation of motion for straight line radial motion, get the conservation of energy equation and rearrange to get dt in terms of dr. Set up a well-defined definite integral for the final answer, then evaluate it. This is an integral you may have to look up. Express the answer as a multiple of the orbital period of the planet. How many weeks or days is that for the Earth? Ans: $1/4\sqrt{2}$ years—about 2 months

6.16 A central attractive force is given by the equation $\vec{F} = -\alpha r^{\beta}\hat{r}$. Here $\alpha > 0$, but β can have either sign. (**a**) Assume that the orbit is almost circular and find the values of β for which the orbit is stable. (**b**) Find the values of β for which the orbit is closed—at least in the perturbation approximation used in this chapter. Ans: **a**: $\beta > -3$

6.17 A particle in a central force moves in a spiral orbit given by $r = a/\phi$. Find the force that does this. And sketch this orbit.

6.18 Derive Eq. (6.53).

6.19 A mass m is moving in the central force $\vec{F} = -k\vec{r}$. For a circular orbit, what is the frequency (or period) of the orbit as a function of the orbit's radius?

6.20 For the same force as the preceding problem, find the shape of an almost circular orbit and explain why it comes out as it does.

6.21 Two masses m_1 and m_2 are attracted to each other by a force directly proportional to their distance apart. Transform to the center of mass coordinate system and solve the problem of their motion as a function of time.

6.22 For the same force as in problems 6.19 and 6.20, transform to a system rotating with the frequency that you found in the first of these problems. One solution is of course that the mass is at rest at its initial radius and angle, say $r = r_0$ and $\phi = 0$. Expand the equations about this to first order and get linear differential equations in $r - r_0 = \delta$ and in ϕ. For small δ and ϕ plot the orbit. Do you believe it?

6.23 Assume for the moment that the Earth and Moon are held stationary. A small mass is placed at a point between the Earth and Moon where the total gravitational force on it is zero. Find this point and then (**a**) The mass is displaced a small amount along the line joining the Earth and Moon and released from rest. Find the force on it for small distances from that equilibrium position and then solve the equations of motion for its position. (**b**) Like (a) but it is displaced along a line perpendicular to the line joining the Earth and Moon.

6.24 In the figure at Eq. (6.17), if the orbit represents an asteroid that comes in just close enough to the sun to graze Earth's orbit, to $1\,\text{AU}$, then how far out will it go? What planet has an orbit of about that radius? Ans: $4.88\,\text{AU}$

6.25 Two particles of mass m_1 and m_2 are repelled from each other by a force of magnitude k/r^2. They start at rest; find the speed of each as a function of the distance travelled.

6.26 Imagine a spherical, non-rotating planet of mass M and radius R. A satellite is sent up from the surface of the planet with a speed v_0 at angle $30°$ from the local vertical. In its subsequent path it reaches a maximum distance $5R/2$ from the center of the planet. (**a**) Find v_0. Ignore the atmosphere. (**b**) Now assume that the planet is rotating with angular speed ω and the satellite is fired from the equator. What is v_0 now? Two cases: fired east, fired west. What is this effect for the Earth, and in particular what is the ratio of the energies required to launch the satellite? Also compare the smaller of these to the energy required to launch from the North Pole. Ans: **a**: $\sqrt{\frac{6}{5}gR}/\cos 30°$

6.27 An asteroid is headed in the general direction of Earth, and its speed when far away is v_0 relative to the Earth. It is aimed so that in the absence of gravity it would pass the Earth at a distance b from Earth's center. (**a**) In the presence of gravity how far from the center will it pass? You can do this simply using conservation laws, or the hard way using all the apparatus of this chapter. (**b**) If it just barely scrapes the planet, what would b have to be? Express this last result in terms of v_0 and the escape speed from the planet. Ans: (**b**) $b > r\left[1 + v_{\text{esc}}^2/v_0^2\right]^{1/2}$ and of course, is this answer plausible?

6.28 For the spherical pendulum with an almost circular orbit, what is the shape of the orbit if you have set it moving so fast that the cord holding the mass is almost, but not quite, horizontal?

6.29 In a noted experiment by Dicke, he concluded that the sun is not exactly spherical but has a slight bulge toward the equator. This would cause an alteration in the gravitational potential energy to

$$U(r) = -\frac{GMm}{r} - \eta\frac{GMmR^2}{r^3}$$

where $\eta \ll 1$. (This form for the correction applies only for the case that the planet is in the plane of the sun's equator.) For a planet in an almost circular orbit, this added term will cause a small precession of the planet's orbit. Find this precession by doing a perturbation analysis about a circular orbit.

6.30 If the gravitational field has a small inverse cube part in addition to the usual one, you can solve for the orbit without approximation.

$$F_r(r) = -\frac{GMm}{r^2} - \delta\frac{GMmR}{r^3}$$

Go back to the orbital differential equation for a general force and apply it to this. Just get the relation for $r(\phi)$ instead of the coordinates as functions of time. Plot the orbit for both positive and negative small δ.

6.31 "As the Earth and Moon move around the Sun together, the Moon's orbit is always concave toward the Sun." Is this true? If so, does it mean that the Moon is orbiting the Sun and not the Earth?

6.32 Derive Eq. (6.25). Also, what is the radius for the orbit with energy E_1 in the same picture?

6.33 What is the minimum energy E_1 from Eq. (6.24) and interpret this result in terms of Eq. (6.25).

6.34 Find the force law for a central force that allows a mass to move in a spiral orbit $r = k\phi^2$ where k is a constant.

6.35 Toss an object straight up at the surface of the Earth, and you usually assume that g is constant. Of course, it isn't. For a coordinate y measured from the Earth's surface,

$$g = \frac{GM}{r^2} = \frac{GM}{(R + y)^2} = \frac{GM}{R^2(1 + y/R)^2}$$

Expand this to first order in y/R and use the resulting force to find the motion of a mass thrown straight up at initial speed v_0, neglecting air resistance and the rotation of the Earth. What is the maximum height in this case? Compare the result to the simple case that assumes g remains at the same value it had at the surface. How big is the effect if the simpler version gave a result of 10 km? 100 km?

6.36 ˙Space debris is a problem in near-Earth orbit. If you are in a ship in such an orbit, the fact that you're moving at several kilometers per second is not the problem, because most debris will be moving at about the same speed in about the same direction. The key word is "about". Assume that your orbit is circular with radius r_1 and that someone once lost a ball-peen hammer that wound up in a orbit with perigee r_1 and apogee $r_2 = r_1 + \delta$. If it hits your ship, how fast will it be moving relative to you? Do this for small δ, and express the result in terms of r_1, g, and δ. Here g is the Earth's gravitational field at this height. For low orbit, if $\delta = 100$ km, what is this speed? This speed corresponds to a fall from what height at the Earth's surface? Ans: $\delta\sqrt{g/r_1}\,/4$, almost 50 meters
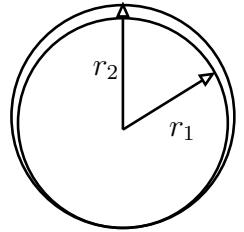
6.37 A mass m is orbiting in a field with a central force $F_r = -Ar^4$. It has angular momentum L. $A > 0$, $L > 0$. (**a**) What is the potential energy for this force? Take it to be zero at $r = 0$. (**b**) Find the energy for which this orbit will be circular, and what is its radius? Ans: (**b**) $r = (L^2/mA)^{1/7}$

6.38 ˙A mass is moving in a central potential in a circular orbit. This circular orbit passes *through* the center of the force. That is, the center of the field is on the circumference of the circle. This type of motion is possible for exactly one form of central force. Find that form. (It is something like r^{-n} for some n.) First show that the orbit can be described by the equation $r = R_{max}\cos\phi$

6.39 The Hohmann transfer orbit between planets (assume circular planetary orbits) puts the spacecraft in an elliptical orbit that is tangent to the orbit of the starting planet and tangent to the orbit of the destination planet. How much time would this orbit take getting from Earth to Mars? This neglects the acceleration time at the start and the acceleration time at the end. What directions are these two accelerations? Ans: 8.5 months

6.40 ¨A small moon of mass m and radius a orbits a planet of mass M while keeping the same face toward the planet. Show that if this moon is too close to the planet, then loose rocks lying on the surface of the moon will be lifted off. What is this distance, and what is it for the Earth-moon system? The moon's mass is 1/81 of the Earth's; its radius is 1740 km. Ans: r (center-to-center) $\approx 11\,000$ km

6.41 ˙A planet is orbiting a star, *but* all that you know about the force of attraction is that its potential energy is $U(r)$ and that the planet does have an orbit. What is the angle through which the planet will orbit as it goes from minimum r to maximum r? The result will be an integral in which $U(r)$ appears. Use conservation of energy and angular momentum. Eliminate dt between them.
Ans: $\int_{r_1}^{r_2} (\ell/r^2)\,dr/\sqrt{2(E - U(r) - m\ell^2/2r^2)/m}$

6.42 ˙Test your result in the preceding problem for the special case that $U(r) = -GMm/r$ to see if it gives the correct answer there.

6.43 In section 6.9 there are plots of many orbits. Figure out which ones go with which equations. There is one pair that is too close to call, but in that case you can still determine which pair of equations are the candidates. The parameters β and γ can make a big change in the shapes, and in turn the angular momentum per mass, ℓ, controls these — how?

6.44 Return to the analysis of section 6.9, but now make the force repulsive instead of attractive and analyze the solutions. These orbits are not nearly as wild.

6.45 A mass is sliding around the inside of a cone that has vertex angle 2α. Assume no friction and that the mass is sliding in a horizontal circle. Find the relation between its angular speed and the height of the circle above the vertex of the cone. Ans: $\dot\phi_0 = \sqrt{g/z_0}\cot\alpha$
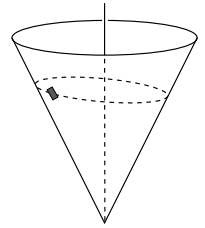
6.46 `Now take the physical system of the preceding problem and find the shape of the orbit, assuming that it is only *almost* horizontal. Use cylindrical coordinates, r-ϕ-z, and write down the total energy of the mass. Write the z-component of angular momentum, which is conserved. Because r and z are not independent variables in this situation, eliminate z. Eliminate ϕ and solve for the almost circular radial motion as a function of time. Combine this with $\dot\phi$ to determine the shape of the orbit. Ans: $\omega_{\text{radial}}/\dot\phi_0 = \sqrt{3}\sin\alpha$

6.47``For the elliptical orbits of Eq. (6.3), what is the semi-major axis for arbitrary δ?

6.48 A planet is orbiting a star in an almost circular orbit of mean radius r_0. If there is, in addition to the usual GMm/r^2 gravitational force, a small harmonic oscillator force $\vec F_1 = -k\vec r$, find the effect on the orbit in the approximation that the orbit differs little from a circle.

6.49 The problem 6.15 can be done with no integration, using three or four lines of algebra and Kepler's laws. Do so.

6.50 (**a**) For two masses m_1 and m_2, what is the total momentum in terms of coordinates $\vec r$ and $\vec r_{\text{cm}}$ (and their derivatives) as in Eq. (6.48).
(**b**) What is their total kinetic energy in terms of CM variables.
(**c**) What is their total angular momentum in terms of CM variables.
In all three cases you should have no cross terms between the relative and the CM coordinates. Do the answers depend on which one is mass #1 and which is #2? Ans: $m_{\text{tot}}\vec v_{\text{cm}}$, $\frac{1}{2}m_{\text{tot}}v_{\text{cm}}^2 + \frac{1}{2}\mu v_{\text{rel}}^2$, $m_{\text{tot}}\vec r_{\text{cm}} \times \vec v_{\text{cm}} + \mu\vec r_{\text{rel}} \times \vec v_{\text{rel}}$

6.51 `Set up the two-dimensional isotropic harmonic oscillator in polar coordinates as in sections 6.2 and 6.3. (**a**) Write the differential equation for $r(t)$. (**b**) What is the effective potential energy for this system? (**c**) You have the solution to this problem in rectangular coordinates from section 6.1, so from that you can construct the polar solution for $r(t)$. (**d**) Why is the solution expressed in problem 3.67 correct and do the results there agree with what you get from the calculation here?

6.52``Two masses are tied to the two ends of a light string, and one of them is suspended through a hole in a table. Assume that there is no friction anywhere and that the mass m_1 is moving on a circle, held at a constant radius by the tension in the string. (**a**) Find v, r, ω for this motion. (**b**) For the general case in which m_1 has arbitrary motion in r and ϕ, write the total energy and the angular momentum and use them to find the orbit of m_1 if it is *almost* circular. Ans: $\dot\phi_0^2/\omega_0^2 = (m_1 + m_2)/3m_1$.
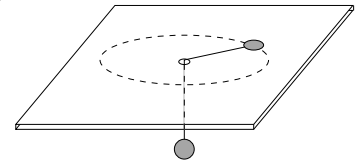
6.53 `A satellite is in a circular orbit of radius r_0 above the Earth. (**a**) What is its total (kinetic plus potential) energy? If the orbit is low enough (say 200-300 km or so) there will be some very small friction with the atmosphere, applying force of magnitude F_{fr} to the satellite. (**b**) In one orbit, what is the change in energy of the satellite? (**c**) In that orbit, what will be the change in r? What then is the change in r per time for the satellite, dr/dt, the decay rate? (**d**) Make the implausible assumption that the frictional force is a constant all the way down to the Earth's surface, how much time will it be before the satellite hits the ground? (**e**) How does the speed of the satellite vary over time because of this frictional force?

6.54 Show that you can quickly and easily go from the beginning of section 6.6 all the way to the results implied by Eq. (6.43) if you start with small angles, use rectangular coordinates, and follow the methods of section 6.1.

6.55 In section 6.8 the equation $\sin \delta = R_{star}/R_{orbit}$ appears. Fill in the missing steps in its derivation.

6.56 The hyperbolic orbits in section 6.10 led to a picture in which only one half the hyperbola was the orbit. The other half (the dashed lines) didn't seem to correspond to anything. Show that they represent the solution if the inverse square force is *repulsive* instead of attractive. For an example of this, consider two positively charged particles, *e.g.* a proton and an atomic nucleus.

6.57¨For those who know the mathematical technique, start from $dt = r^2 \, d\phi/\ell$ and (6.16), using contour integration to derive Kepler's law Eq. (6.20).

6.58˙For the orbital equation (6.14), what is the equation if the force has two terms, the usual Newtonian gravity plus a $1/r^4$ term.

$$f(r) = -\frac{GMm}{r^2} - 3\frac{GMm\ell^2}{c^2 r^4}$$

Here, c is the speed of light. Does this even have correct dimensions? This is the orbital equation derived from Einstein's theory of gravity, the General Theory of relativity. Use the methods starting with equation (6.34) to find the angle by which the elliptical orbit described by the first term alone will precess in each orbit. Look up the orbital parameters for the planet Mercury to get the numerical value for this result and then convert it into the conventionally reported value of the precession per Earth century ($43''$ per century). Cf. section 6.12, which addresses the (larger) effect of the gravitational pull by the other planets.

6.59¨Add a uniform magnetic field to the spring force of Eq. (6.1), assuming that all the spring constants are equal, and that the mass has a charge q. Choose the z-axis along the direction of the magnetic field. Solve for the motion of the mass and sketch some solutions, especially for weak B-fields.

6.60 The carbon monoxide molecule, CO, can be modeled as two masses on the ends of a spring having unstretched length ℓ. Solve for the oscillations of this molecule, the normal modes, assuming that the motion is along the single long axis between the atoms. Compare your result to what appears in section 6.7, especially Eq. (6.47).

6.61 Estimate the period for the close orbit of a pebble around a rock. More than one asteroid has been observed with such a satellite. Assume the density of the rock is comparable to the average density of Earth. Look up Ida and Dactyl, pictured on the right.

6.62 In section 6.9 you see a variety of complicated orbits. Now examine the case that the inverse cube force is repulsive, and solve for those orbits in this case.

6.63 Verify that the equations (6.58) describe the hyperbola as stated there.

6.64 Take all the complicated results in section 6.11 and apply them to the case of a circular orbit to see if they give the correct results. (This is how I found some missing factors.)

6.65˙For the special case that an asteroid starts from an initial distance R from the sun at zero velocity, what do the equations for $r(s)$, $t(s)$, $\phi(s)$ in section 6.11 become? And analyze the results. Check out the various versions of Eq. (6.65) to compare their utility. How much time does it take to hit the sun?

6.66 For the Kepler orbit, use the equation for r in terms of ϕ, the first of Eqs. (6.17) and compute the velocity of the planet, $\vec{v} = \dot{r}\hat{r} + r\dot{\phi}\hat{\phi}$. Use ℓ to express this in terms of ϕ alone. Show that the vector coefficient of ϵ is the constant unit vector $\hat{y} = \hat{r}\sin\phi + \hat{\phi}\cos\phi$. Finally use this to show that the graph of $\vec{v}(t)$ is a circle centered at $\ell\epsilon\hat{y}/a(1-\epsilon^2)$. This is called a hodograph.

6.67 Easier than the preceding problem for the Kepler orbit, what is the locus of $\vec{v}(t)$ for the anisotropic two-dimensional harmonic oscillator?

6.68 In section 6.3 you saw the change of variables from $r(t)$ and $\phi(t)$ to $u = 1/r$ with the independent variable ϕ. Try another approach: Again use ϕ as the independent variable, but use v_ϕ, the angular component of the velocity as the dependent variable. What differential equation do you get for this? Solve it. If you can't, then go back and find your mistake.

6.69 Show that the ratio of the solar tidal effect at perihelion is $(1 + 6\epsilon) \approx 1.1$ times the solar tide at aphelion.

6.70 A variation on the problem 6.41 is actually useful. Again for an arbitrary $U(r)$, an object starts from very far away and is headed toward the general direction of the center of the force, though aimed so as to miss. Far away it has a speed v_0 and its line of motion is aimed to pass at a distance b from the origin, the impact parameter. Follow the procedure described in problem 6.41 to eliminate dt between the two conservation laws, and find the angle (the "scattering angle") at which it will leave the system after having come to its distance of closest approach. Since the first half of the orbit is a mirror image of the second half, find the direction change between the distance of closest approach and when it's infinitely far away. Then double the result. The equation for $r_{\min}$ comes from conservation of energy and of angular momentum. Check: what if $U \equiv 0$? What if $v_0 \to \infty$? In this check, you may need the equation $\int_0^\infty d\theta \,\text{sech}\,\theta = \pi/2$, depending on what change of variables you make for the integral.
Ans: $-\pi + 2\int_{r_{\min}}^\infty (\ell/r^2)\,dr/\sqrt{2(E - U(r) - m\ell^2/2r^2)/m}$, where $\ell = v_0 b$

6.71 Use the result of the preceding problem to find the scattering angle for an α-particle aimed at a gold nucleus. Ignore the motion of the nucleus, as it is quite massive. This result is important in understanding Rutherford's experimental discovery of the *existence* of the nucleus. Here the potential energy is kq_1q_2/r, where $q_1 = 2e$, $q_2 = Ze$, and $Z = 79$ for gold.

6.72 Derive $V_1(r)$ in Eq. (6.67) from Eq. (6.66). Factor R from the denominator and use the binomial expansion for $(1 + x)^{-1/2}$ out to second order in x.

6.73 An ellipse can be written as two parametric equations:

$$x(\phi) = a\cos\phi \qquad \text{and} \qquad y(\phi) = b\sin\phi$$

Use these to find the circumference of an ellipse: $ds = \sqrt{dx^2 + dy^2}$. Express the result in terms of an elliptic integral as in Eq. (6.72). Is the result correct in the two cases for which you can write down the answer easily? If you study more orbit theory you will encounter the term "eccentric anomaly". That's what this ϕ is.
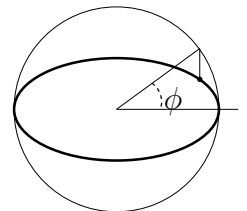
6.74 (Irrelevant to this material, but fun) An ellipse is a conic section. That means that the intersection of a plane and a cone is an ellipse (or a hyperbola or a parabola if the angles are different). You can inscribe two spheres inside the cone and tangent to the intersecting plane at two points. What are these two points? (What else could they *be?*)
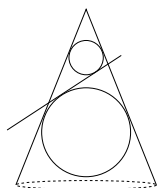
6.75 Show that $E(ik) = \sqrt{1 + k^2}E\left(k/\sqrt{1 + k^2}\right)$.

Also, $K(ik) = \frac{1}{\sqrt{1+k^2}}K\left(k/\sqrt{1 + k^2}\right)$.

This needs only the simplest trigonometric identities, so it's not necessary for you to panic.

6.76 Derive Eq. (6.66). $dM = M\, d\phi/2\pi$.

6.77¨ The expression for potential energy of a ring, Eq. (6.66), is exactly the same if $r > R$, and you can express it in terms of an elliptic integral by factoring the larger radius r out of the denominator instead of R. (**a**) Show that you get $U_1(r) = -(2GMm/\pi r)K(R/r)$ this time. Also, check that it behaves properly for large r and for r near to R. (**b**) The largest perturbation on the orbit of Saturn comes from Jupiter; just look at the numbers following Eq. (6.71) to see why. Compute the precession of the perihelion of Saturn because of the presence of Jupiter. (**c**) Add the effects from Uranus and Neptune, as they use the same forms of the equations that appear in the text, applying to Mercury $(r < R)$.
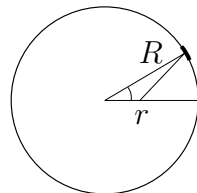
Waves

.

The general subject of waves is so large that it would take several books just to get some idea of the extent of the subject. I won't do that. Instead, this chapter contains a few of the ideas and some of the simplest applications, chosen because they appear in so many contexts.

7.1 A String

The prototype for this subject is a wave on a string. It can be a guitar string or a bullwhip, but the basic analysis is the same and it starts from $\vec{F} = m\vec{a}$.

Assume that the string is stretched straight along the x-axis, and I'll assume that its oscillations occur in a single plane so that the y-coordinate says how far the string is from this equilibrium position. The shape of the string is then defined by the equation $y = f(x, t)$, representing the displacement of the string from the axis at point x and time t.

The key trick in setting up the equations of motion is to imagine slicing the string apart and then examining the segment between x and $x + \Delta x$. This segment will then have forces acting on it caused by the part of the string $(< x)$ pulling in one direction and part of the string $(> x + \Delta x)$ pulling in another. There will also be the forces by gravity and by the surrounding air. If this sort of division, using an imaginary cut in the string, looks fairly plausible and maybe even commonplace, it wasn't when the subject was invented. That you can imagine slicing the string (without really doing it) was the subject of great debate and worry by those who did it the first time.
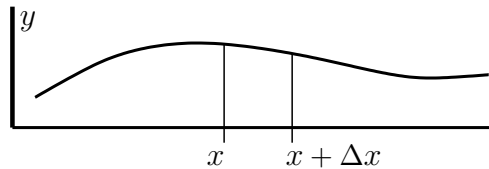
As with most other problems involving oscillations, this becomes simple only when the displacements are small. Without that assumption the resulting differential equations will be extraordinarily difficult. If you've ever looked at a guitar or any other stringed instrument being played, you know that the motion of the string is very tiny and so this small motion approximation will be a very good one.

The linear mass density is $\mu = dm/dx$ and the tension in the string is $T(x)$. This tension, the force that one part of the string exerts on the part it is connected to, need not be a constant, and the same is true of the mass density. The interval Δx is small (eventually $\rightarrow 0$), so the mass in the interval is $\Delta m = \mu(x)\Delta x$.
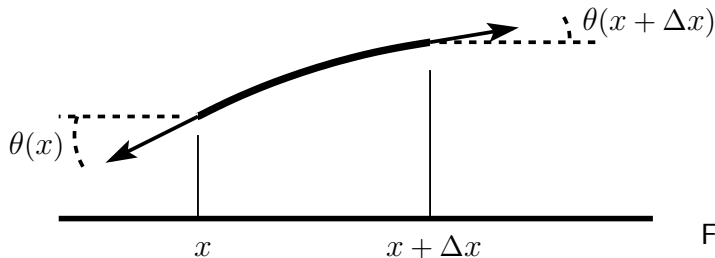
Fig. 7.1

The forces on this Δm come from the two adjacent parts of the string, from gravity, and from the surrounding air.

$$F_y = -T(x)\sin\theta(x) + T(x + \Delta x)\sin\left(\theta(x + \Delta x)\right) - \Delta m\, g - b\Delta x\, v_y \tag{7.1}$$

Here I'm taking the simplest model for the air resistance, linear viscosity, with a proportionality factor b. This string is horizontal, so that determines the gravitational term. The angle that the string makes with the x-axis depends on x, so to get the y-component of the force I needed to indicate that explicitly.

Up to here the presentation is straight-forward, I've just been very careful in expressing the forces. Now for a trick. The displacement of the string is small and the angle that it makes with the axis is small. That means that I can use a small angle approximation for the sine. Not the usual one though.

$$\text{small angle} \implies \sin\theta \approx \theta \approx \tan\theta \qquad [\cos\theta \approx 1]$$

Instead of approximating the sine by theta, approximate it by the tangent of theta. The reason for wanting to do this is that the tangent is simply the derivative of y with respect to x. Equation (7.1) now becomes

$$F_y = -T(x)\frac{\partial f}{\partial x}(x,t) + T(x+\Delta x)\frac{\partial f}{\partial x}(x+\Delta x, t) - \mu(x)\Delta x\, g - b\Delta x\frac{\partial f}{\partial t}(x+\Delta x/2, t) \qquad (7.2)$$

I took the last term at the center point of the interval, but it won't matter in the end. I switched to partial derivative notation because there are two independent variables, and when doing the x-derivative time is fixed. For the t-derivative, position is fixed. For the ma side of the equation,

$$F_y = \Delta m\, a_y = \mu(x)\Delta x\frac{\partial^2 f}{\partial t^2}(x,t)$$

If you think that this ought to be evaluated at $x+\Delta x/2$ instead, try it, and verify that it won't matter. If I now take the limit as $\Delta x \to 0$, I get the most uninformative equation, $0 = 0$. Not very helpful. Instead, first divide by Δx and *then* take the limit.

$$\frac{1}{\Delta x}\left[-T(x)\frac{\partial f}{\partial x}(x,t) + T(x+\Delta x)\frac{\partial f}{\partial x}(x+\Delta x, t)\right] - \mu(x)g - b\frac{\partial f}{\partial t}(x+\Delta x/2, t)$$

$$= \mu(x)\frac{\partial^2 f}{\partial t^2}(x,t)$$

Now as $\Delta x \to 0$, most of the terms the equation simply become the function at x. The first term [] is different however. It goes to $0/0$, and it is nothing more than the definition of a derivative. [something$(x+\Delta x)$ − something(x) all divided by Δx]. The limit is

$$\frac{\partial}{\partial x}\left(T\frac{\partial f}{\partial x}\right) - \mu g - b\frac{\partial f}{\partial t} = \mu\frac{\partial^2 f}{\partial t^2} \qquad (7.3)$$

Have I left out any physics? Yes. A thin string is easy to bend, but a thicker one resists deformation. It is stiff and I've ignored this property. Is that important? It is if you are a professional piano tuner and have a demanding client. One aspect of this problem will appear in section 7.14. For now, think of this as a thin *chain* for which it takes no effort to rotate one link with respect to the next one.

What about F_x on the mass Δm? Go back to the equation (7.1) and the picture that accompanies it.

$$F_x = -T(x)\cos\theta(x) + T(x+\Delta x)\cos\left(\theta(x+\Delta x)\right) \qquad (7.4)$$

In this example, because the string is horizontal there is no x-component from gravity. This implies that when the string is horizontal, then for small angles and for motion in the y-direction, $F_x = \Delta m\, a_x$ is zero and the tension is constant. I will leave the T factor as is anyway, allowing that it may vary

with x, not because it's needed for this problem with the string but because it's harmless, not causing any extra work. Also there are plenty of wave problems where something analogous to this factor *will* show up and you may as well get used to it here. In most cases here it will be constant. Of course if the string is vertical instead of horizontal there is a term in $\Delta m\,g$ in this equation but not in the F_y equation. In that case, put the extra term $(\pm)\Delta m\,g$ into Eq. (7.4) and it will determine how the tension varies with x.

7.2 Static case

Start with the simplest circumstance. Nothing is moving. A string is stretched taut between two posts and it sags. Let the tension and the mass density be constant here.

The equation (7.1) for the shape of the string is now

$$T\frac{\partial^2 f}{\partial x^2} - \mu g = 0 \tag{7.5}$$

This is easy to integrate. (Drop the t from $f(x,t)$.)

$$f(x) = \mu g x^2/2T + Ax + B$$

Apply the boundary conditions that the string is tied at the two ends

$$f(0) = f(L) = 0 \implies B = 0 \quad \text{and} \quad \mu g L^2/2T + AL = 0$$

The curve is then

$$f(x) = -\frac{\mu g}{2T}x(L - x) \tag{7.6}$$

The shape of the wires hung between telephone poles in this small angle approximation* is a parabola. The bigger the tension, the less it sags. The heavier it is, the more it sags. Also look at problem 7.33 for a different way to look at this question.

7.3 The Wave Equation

In Eq. (7.3), the simplest case that exhibits waves comes by assuming that the mass density and the tension are constant and that gravity and air resistance are negligible.

$$T\frac{\partial f}{\partial x^2} = \mu\frac{\partial^2 f}{\partial t^2} \quad \text{or} \quad \frac{\partial^2 f}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = 0 \quad \text{where} \quad v^2 = \frac{T}{\mu} \tag{7.7}$$

This is called "the" wave equation, as if there were no others. It *is* the most important wave equation, and it fully deserves the definite article. The parameter v has dimensions of speed, and that's easy to see because f is the same in the numerator in both terms of the equation, so the denominators, x^2 and $v^2 t^2$ must have the same dimensions. v is the propagation speed of the wave.

How do you solve the wave equation? There are many ways, starting with inspired guesswork and carrying on through separation of variables and on into the world of characteristics. I'll mostly stay with the first method, though one direct approach to the solution appears in problem 7.21. For now, try a cosine, $f(x,t) = A\cos(kx - \omega t + \delta)$ and see what happens.

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = -Ak^2\cos(kx - \omega t + \delta) + A\omega^2\frac{1}{v^2}\cos(kx - \omega t + \delta)$$
$$= -A\cos(kx - \omega t + \delta)\left[k^2 - \omega^2/v^2\right] = 0 \tag{7.8}$$

---

* The static problem can be solved without the small angle approximation, and the result is a catenary — a hyperbolic cosine

The cosine isn't identically zero and A better not be or there's nothing there. What's left determines the relation between k and ω.

$$k^2 - \omega^2/v^2 = 0, \qquad \text{or} \qquad v^2 = \frac{\omega^2}{k^2} = \frac{T}{\mu} \tag{7.9}$$

This is already an infinite number of independent solutions because I can make ω anything that I want, and as long as k has the corresponding value $\pm\omega/v$ this works. Compare the harmonic oscillator, where there are exactly *two* independent solutions and one frequency for the system, not an infinite number. This richness of solutions here is typical of partial differential equations such as (7.7).
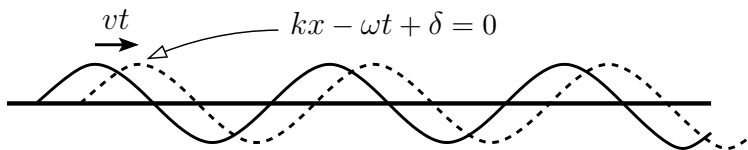
Fig. 7.2

Is v really a speed? Draw a picture of the function y versus x, and where do you find its peak? A peak will occur wherever the argument of the cosine is zero or a multiple of 2π. That define the equation for the position of a peak. Pick one:

$$kx - \omega t + \delta = 0 \qquad \text{implies} \qquad x = \frac{\omega}{k}t + \frac{\delta}{k} = \pm vt + \text{a constant} \tag{7.10}$$

The position of the maximum moves at precisely the value of v, either to the right or left depending on the sign of ω/k.* Most commonly the frequency, omega, is taken as positive and k can have either sign, thereby determining the direction of the wave's travel as right or left as k is positive or negative. k is called the wave number, and has dimensions of angle per distance, just as ω is angle per time. The distance in which the wave repeats is the wavelength, λ. This implies that $|k|\lambda = 2\pi$.

The equation (7.7) is linear. That means that the sum of two solutions is a solution, which implies that you can create arbitrarily complicated waves by adding waves of different frequencies. All that's needed is that $\omega_1^2/k_1^2 = \omega_2^2/k_2^2 = \ldots = v^2$.

$$A_1 \cos(k_1 x - \omega_1 t + \delta_1) + A_2 \cos(k_2 x - \omega_2 t + \delta_1) + A_3 \cos(k_3 x - \omega_3 t + \delta_3) + \cdots$$
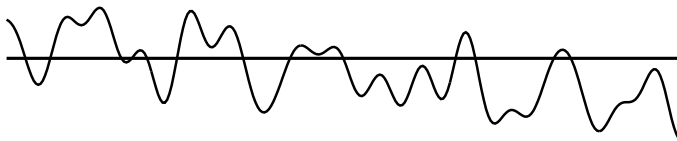
This is a picture of such a combination formed by picking a set of random values of the A's, k's, and δ's (at $t = 0$). Sound obeys the same wave equation as Eq. (7.7), and you know that sound waves can carry enormously complex information, whether in music or in words.

Are there any other solutions besides the cosines (or sines for that matter)? Yes, just as the ordinary differential equation $md^2x/dt^2 = -kx$ has two arbitrary constants in its solution, $A\cos\omega t + B\sin\omega t$, the wave equation has two arbitrary functions!

$$y = f(x, t) = f_1(x - vt) + f_2(x + vt) \tag{7.11}$$

---

* This is called the *phase velocity* to distinguish it from other definitions of velocity that will show up later, in section 7.10. The name comes from the fact that you are following a point of constant phase, the argument of the cosine: $(kx - \omega t + \delta)$ in the wave.

The demonstration of this fact is no more than plugging in to the equation and using the chain rule.

$$\frac{\partial f_1}{\partial x} = f_1'(x - vt), \quad \frac{\partial^2 f_1}{\partial x^2} = f_1''(x - vt), \quad \frac{\partial f_1}{\partial t} = -vf_1'(x - vt), \quad \frac{\partial^2 f_1}{\partial t^2} = (-v)^2 f_1''(x - vt)$$

Substitute these into Eq. (7.7) and it works. Note: the "prime" notation is the derivative of f with respect to its argument, not the derivative with respect to x, so that $\sin' = \cos$. The same calculation works for $f_2(x + vt)$. This means that *any* wave form will move along the string. All that you need do is to get it started. The cosine solution that I started with is simply a special case of this general solution. And yes, Eq. (7.11) *is* the most general solution for the wave equation. The proof is tedious but not difficult, involving only a change of variables. See problem 7.21.

Any function (well, any function with two derivatives) will provide a solution, and that is why waves can carry so much information. To find the behavior of a single mass, you need to specify its initial position and velocity, and essentially the same thing happens here. Specify the initial position and initial velocity at each point on the string. Assume

$$f(x, 0) = F(x), \qquad \frac{\partial f}{\partial t}(x, 0) = G(x) \tag{7.12}$$

Apply Eq. (7.11) for the general solution, then these become

$$f_1(x) + f_2(x) = F(x) \qquad \text{and} \qquad -vf_1'(x) + vf_2'(x) = G(x)$$

Differentiate the first of these to get $f_1'(x) + f_2'(x) = F'(x)$ and you now have two algebraic equations for the two unknowns f_1' and f_2'. Solve:

$$f_1'(x) = \tfrac{1}{2}F'(x) - \tfrac{1}{2v}G(x), \qquad f_2'(x) = \tfrac{1}{2}F'(x) + \tfrac{1}{2v}G(x)$$

And now do two integrals. These are all functions of a single variable, so nothing unusual happens. You have constants of integration, but they're easy to handle.

$$f_1(x) = \frac{1}{2}F(x) - \frac{1}{2v}\int_0^x G(x')\,dx' \qquad f_2(x) = \frac{1}{2}F(x) + \frac{1}{2v}\int_0^x G(x')\,dx' \tag{7.13}$$

$$y = f(x, t) = \frac{1}{2}F(x - vt) - \frac{1}{2v}\int_0^{x-vt} G(x')\,dx' + \frac{1}{2}F(x + vt) + \frac{1}{2v}\int_0^{x+vt} G(x')\,dx'$$

$$= \frac{1}{2}\big[F(x - vt) + F(x + vt)\big] + \frac{1}{2v}\int_{x-vt}^{x+vt} G(x')\,dx' \tag{7.14}$$

You should of course directly verify that this expression satisfies the initial conditions. See problem 7.8. And is there anything special about taking the lower limits in Eq. (7.13) as zero? (No.)

Example
► Take $F(x) = A(L^2 - x^2)$ for $-L < x < +L$ and zero elsewhere. Take the initial velocity to be zero: $G \equiv 0$, then

$$f(x, t) = \frac{1}{2}\big[F(x - vt) + F(x + vt)\big] \tag{7.15}$$

This is a plucked string, and the disturbance goes equally in both directions. For a struck string, see problem 7.10.

7.4 Energy and Power

Waves can transport energy. Certainly ocean waves do, and sound waves that are loud enough can carry enough energy to damage your ears. On a string, the analysis is pretty straight-forward, leading to the ideas of kinetic energy density and potential energy density.

Start with the kinetic energy; it's easier. Do the same thing that I did to get the wave equation in the first place: look at a small piece of the string in the interval from x to $x + \Delta x$. It has a mass $\Delta m \approx \mu(x)\Delta x$. Its kinetic energy is then given by the usual $mv^2/2$ formula.

Watch out! v is not v.

The v that shows up in the wave equation is the speed with which the disturbance moves along the string, $\sqrt{T/\mu}$. The v that appears in the expression for kinetic energy is the speed of the mass m. They are not at all the same thing. The former is the speed with which a shape moves, not the speed with which a mass moves. That latter one occurs at a fixed value of x so it is a partial derivative, $\partial f/\partial t$.

$$\Delta K \approx \frac{1}{2}\mu\Delta x \left(\frac{\partial f}{\partial t}\right)^2 \qquad \text{so} \qquad \frac{\Delta K}{\Delta x} \to \frac{dK}{dx} = \frac{1}{2}\mu\left(\frac{\partial f}{\partial t}\right)^2 \tag{7.16}$$

This is the kinetic energy density (Joules per meter). There's nothing in this derivation that requires μ to be a constant. It applies just as well to the general equation, Eq. (7.3).

For the first solution, the cosine, what is this? $f(x,t) = A\cos(kx - \omega t)$, so

$$\frac{dK}{dx} = \frac{1}{2}\mu\big(-A\omega\sin(kx - \omega t)\big)^2 = \frac{1}{2}\mu A^2 \sin^2(kx - \omega t)$$
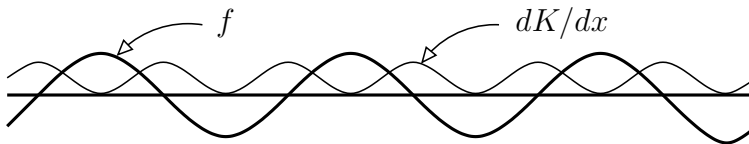
Fig. 7.3

This shows the two functions f and dK/dx on the same graph, and you can see that the kinetic energy density is zero when the magnitude of the wave is at its maximum and vice versa. When the string is at its peak, the molecules at that peak have stopped and have not yet turned around — that's zero speed.

The potential energy is a little trickier. When you use the equation $\Delta m = \mu\Delta x$, the mass is what occurred along the axis in the interval Δx. The length of the string when it is deformed can't stay *exactly* the same though, because the y-coordinates at its ends aren't the same as when it was at rest. Its length is

$$\Delta\ell = \sqrt{(\Delta x)^2 + (\Delta y)^2}$$

It takes some work to stretch this string. All that I have to do is to compute this work and that will give me the potential energy in this piece of stretched string. The tension in the string is T, or $T(x)$, and the work that I do in changing the length of this string segment is this force times the stretch.

$$\Delta W \approx T\big[\Delta\ell - \Delta x\big]$$

To evaluate this, use the binomial series.

$$\Delta\ell = \Delta x\sqrt{1 + \frac{(\Delta y)^2}{(\Delta x)^2}} = \Delta x\left[1 + \frac{1}{2}\frac{(\Delta y)^2}{(\Delta x)^2} + \cdots\right]$$

(The slope is small.) Subtract the original length Δx and

$$
\Delta W \approx T\Delta x \frac{1}{2}\frac{(\Delta y)^2}{(\Delta x)^2} \qquad \text{so} \qquad \frac{\Delta W}{\Delta x} \to \frac{dW}{dx} = \frac{1}{2}T\left(\frac{\partial y}{\partial x}\right)^2 \tag{7.17}
$$

This is the potential energy density. As before, this derivation applies to the general equation, for which T can be a function of position. The total energy density is the sum of these.

$$
\frac{dE}{dx} = \frac{1}{2}\mu\left(\frac{\partial f}{\partial t}\right)^2 + \frac{1}{2}T\left(\frac{\partial f}{\partial x}\right)^2 \tag{7.18}
$$

How does energy flow down the string? One part of the string pulls on its neighbor and does work on it. What is the power transmitted? There are a couple of ways to do this and I'll do one here, leaving the other (much easier) method to you in problem 7.2.* The total energy within the interval $a < x < b$ is the integral over the energy density.

$$
E(t) = \int_a^b dx\,\frac{dE}{dx} = \int_a^b dx\left[\frac{1}{2}\mu\left(\frac{\partial f}{\partial t}\right)^2 + \frac{1}{2}T\left(\frac{\partial f}{\partial x}\right)^2\right] \tag{7.19}
$$

Note that this is a function of time as the wave moves past. The time variable does not appear in the limits of the integral; they're fixed. It remains in the wave function f itself, because it is just the x-variable that you are integrating. The t-variable is just sitting inside the integral waiting to be differentiated. (Look ahead to Eq. (7.25) to see a concrete example of this.) From conservation of energy, the total power flowing into this region is the time-derivative of E.

$$
\frac{dE}{dt} = \frac{d}{dt}\int_a^b dx\,\frac{dE}{dx}
$$

The limits a and b are not functions of time, so I can simply move the derivative under the integral, make it partial, and do the differentiation on the two terms.

$$
\frac{dE}{dt} = \int_a^b dx\left[\mu\frac{\partial f}{\partial t}\frac{\partial^2 f}{\partial t^2} + T\frac{\partial f}{\partial x}\frac{\partial^2 f}{\partial x\partial t}\right] \tag{7.20}
$$

Now what? I haven't used that fact that $y = f(x,t)$ satisfies the wave equation, and I'll have to do that or I can't get anywhere. The only obvious place that I can apply the equation is on the first term. It becomes

$$
\int_a^b dx\left[\mu\frac{\partial f}{\partial t}v^2\frac{\partial^2 f}{\partial x^2} + T\frac{\partial f}{\partial x}\frac{\partial^2 f}{\partial x\partial t}\right] \tag{7.21}
$$

Now integrate by parts. You can do it on either term; I'll pick the first. I'm assuming that T is independent of x, though the more general case is no different: problem 7.15

$$
\int_a^b dx\,\mu\frac{\partial f}{\partial t}v^2\frac{\partial^2 f}{\partial x^2} = \mu\frac{\partial f}{\partial t}v^2\frac{\partial f}{\partial x}\Big|_a^b - \int_a^b dx\,\mu\frac{\partial^2 f}{\partial t\partial x}v^2\frac{\partial f}{\partial x} \tag{7.22}
$$

Put this back into the equation (7.21) and the two integral terms cancel. Do they? You'd better check. All that's left is

$$
\frac{dE}{dt} = \mu\frac{\partial f}{\partial t}v^2\frac{\partial f}{\partial x}\Big|_a^b = T\frac{\partial f}{\partial t}\frac{\partial f}{\partial x}\Big|_a^b = T(b)\frac{\partial f}{\partial t}(b,t)\frac{\partial f}{\partial x}(b,t) - T(a)\frac{\partial f}{\partial t}(a,t)\frac{\partial f}{\partial x}(a,t) \tag{7.23}
$$

---

* Maybe you will want to start with that problem and then come back here.

This says that the rate of change of the total energy in this interval comes from two terms: one at the right end and one at the left end. This is what I ought to have expected. The energy within this interval can come only from something happening at the ends; it can't jump over and appear at some distant point. I can then interpret the expression

$$P = -T\frac{\partial f}{\partial t}\frac{\partial f}{\partial x} \tag{7.24}$$

as the power moving (to the right) past the point x. Why the minus sign? Look at the term in Eq. (7.23) at the point a. If that term (with its minus sign) is positive, it represents energy coming into the interval from the left. Having gone through this, you can better appreciate the simplicity of the alternate derivation in problem 7.2. Then why did I bother with this more complicated approach? The *methods*, such as differentiating the integral, are important and common and you have likely not seen them applied before.

Example
▶ Take the cosine wave solution from Eq. (7.8), setting $\delta = 0$ for convenience ($f = A\cos(kx - \omega t)$), and see what the total energy in an interval is.

$$\frac{dE}{dx} = \frac{1}{2}\mu\left(\frac{\partial f}{\partial t}\right)^2 + \frac{1}{2}T\left(\frac{\partial f}{\partial x}\right)^2, \quad \text{and} \quad f(x,t) = A\cos(kx - \omega t)$$

$$E = \int_a^b dx\,\frac{1}{2}A^2\left[\mu\omega^2\sin^2() + Tk^2\sin^2()\right] = A^2Tk^2\int_a^b dx\,\sin^2(kx - \omega t)$$

$$= A^2Tk^2\int_a^b dx\left[1 - \cos\left(2(kx - \omega t)\right)\right]$$

$$= A^2Tk^2\left[(b-a) - \frac{1}{2k}\left(\sin(2(kb - \omega t)) - \sin(2(ka - \omega t))\right)\right] \tag{7.25}$$

(Remember Eq. (7.9).) The first term in the result represents the average energy in the interval, and the other terms represent the fluctuations in total energy as the wave enters at a and leaves at b.

The power itself in this example is

$$P = -T\frac{\partial f}{\partial t}\frac{\partial f}{\partial x} = -T\left(+ A\omega\sin(kx - \omega t)\right)\left(- Ak\sin(kx - \omega t)\right)$$

$$= +A^2Tk\omega\sin^2(kx - \omega t) = A^2\mu\omega^2 v\sin^2(kx - \omega t) \tag{7.26}$$

Notice that the power is zero where the wave is a maximum and vice versa. Or is it? And what is the time derivative of the energy computed in Eq. (7.25) and do you believe it?

7.5 Complex Form
The equation (7.11) showed $f(x \pm vt)$ as the most general solutions of the wave equation. Nothing restricts f to being real, and using complex notation provides a great simplification in many calculations about waves. This is the same idea as in using $e^{i\omega t}$ instead of cosines and sines for the simple harmonic oscillator. In the easy problems it doesn't matter, but when the going gets tough. . . .

$$\text{Instead of} \quad \cos(kx - \omega t), \quad \text{use} \quad e^{ikx - i\omega t} \quad \text{(real part understood)}$$

It will take some practice to get used to using this complex form, but eventually you will wonder why someone would do anything else. If you doubt this, try solving problems 7.45 or 7.46 using sines and cosines, then solve them with complex exponentials. The length of one method is measured in pages; the other way is measured in lines. It is up to you to decide which you prefer. Note: When you compute energy or power you cannot use complex amplitudes. It involves squares of the amplitude, and the real part of $(e^{ix})^2$ is not the same as $(\cos x)^2$

7.6 Three Dimensional Waves

Waves on a string are deceptively simple because they can move only in a single dimension. Sound waves, light waves, and seismic waves are examples of more complex waves that are not confined to a line. Sound in air is the first step in adding this complexity because the dependent variable is the scalar pressure variation from the mean air pressure, and it satisfies the same equation as a wave on a string but with more variables.

$$\frac{\partial^2 f}{\partial x^2} - \frac{\mu}{T}\frac{\partial^2 f}{\partial t^2} = 0 \qquad \text{becomes} \qquad \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} - \frac{1}{v^2}\frac{\partial^2 p}{\partial t^2} = 0 \qquad (7.27)$$

Here $v^2 = \gamma p_0/\rho_0$ for ordinary frequencies. p_0 and ρ_0 are the average background air pressure and density, while $p(x,y,z,t)$ is the deviation from this background. The factor γ is the ratio of specific heats that appears so often in the subject of thermodynamics. For frequencies of about 10^8 Hz and more, the factor deviates from γ to something that eventually approaches one as ω increases.

Without going into the derivation of this equation, what do its solutions look like? The answer is not much different from the wave on a string, only with more possibilities. A trigonometric function or a complex exponential will still represent a wave of a particular frequency. You can write

$$p = A\cos\left(k_x x + k_y y + k_z z - \omega t\right) = A\cos\left(\vec{k}\cdot\vec{r} - \omega t\right) \quad \text{or} \quad Ae^{i(\vec{k}\cdot\vec{r}-\omega t)} \qquad (7.28)$$

Put this into the equation to see if it works.

$$\frac{\partial^2 p}{\partial x^2} = -k_x^2 A\cos\left(k_x x + k_y y + k_z z - \omega t\right)$$

with similar results for the derivatives with respect to y and to z. Assemble these to get

$$-A(k_x^2 + k_y^2 + k_z^2)\cos\left(k_x x + k_y y + k_z z - \omega t\right) - \frac{1}{v^2}(-A\omega^2)\cos\left(k_x x + k_y y + k_z z - \omega t\right) \stackrel{?}{=} 0$$

The requirement for this to work is that the coefficient of the cosine be zero.

$$-(k_x^2 + k_y^2 + k_z^2) + \frac{\omega^2}{v^2} = 0 \longrightarrow \frac{\omega^2}{k^2} = v^2, \qquad \text{where} \qquad k^2 = k_x^2 + k_y^2 + k_z^2$$

It doesn't matter if you use a cosine or a sine or a complex exponential: $e^{i(\vec{k}\cdot\vec{r}-\omega t)}$ is just as much a solution as a cosine. $\vec{k}$ is the wave number vector, with units of radians per meter. It is the spacial analog of frequency. Do this calculation yourself for the complex exponential.

One aspect of this is actually easier if you use the three dimensional vector notation: Which direction is the wave moving? In one dimension, if the wave is moving right you had $\cos(kx - \omega t)$ or $\sin(kx - \omega t)$ or $e^{i(kx-\omega t)}$. If the wave is moving left you can change the sign in front of the ω or you can change the sign in front of the k. For the cosine it doesn't matter, but for the sine it does, turning the function upside down. In three dimensions, when you change directions there's really one thing you can do. Change the direction of the vector $\vec{k}$. There is then no ambiguity in how to proceed.

7.7 Reflections

Waves often meet barriers. Light waves hit glass. Ocean waves hit a reef. Sound waves hit a wall. The analysis of all these is similar, so spending some time looking at reflections on a string will not be wasted.

Construct an example that typifies the process by tying two strings together, one thicker and heavier than the other. Light is partially reflected when it hits glass (just look in a window), so it is not

surprising when the same thing happens here. Let the mass density be μ_1 to the left of the origin and μ_2 to the right. A wave comes in from the left and some of it gets through the boundary, but not all. Some is reflected. What equation do you have to solve? The same equation that started this chapter, Eq. (7.3). After dropping the μg term, I assumed that T and μ were constant in order to produce "the" wave equation. Now μ isn't constant, but T will be for the first example.

$$T\frac{\partial^2 f}{\partial x^2} = \mu(x)\frac{\partial^2 f}{\partial t^2} \quad\longrightarrow\quad \frac{\partial^2 f}{\partial x^2} = \frac{1}{v^2(x)}\frac{\partial^2 f}{\partial t^2} \tag{7.29}$$

The mass density is constant in two domains, $\mu = \mu_1$ for $x < 0$ and μ_2 for $x > 0$. In each region there are the corresponding (constant) wave speeds.

$$v(x) = v_1 = \sqrt{T/\mu_1}, \quad (x < 0) \qquad \text{and} \qquad = v_2 = \sqrt{T/\mu_2}, \quad (x > 0)$$

This equation is exactly the same equation as when light (coming along the normal) hits a sheet of glass. The speed ω/k of the light wave in air decreases in glass by a factor of n, the index of refraction. The function f is an electric field instead of the lateral displacement of a string. The mathematics is the same.

If you look only at the left part of the string or only at the right part, you already know the solutions. They are waves such as $A\cos(kx - \omega t)$ or $B\cos(-kx - \omega t)$, and all you need is that the functions satisfy the wave equation, which means $\omega/k = v_1$ or $= v_2$ in the two halves. If they have the same frequency then*

$$\begin{aligned} k_1 &= \omega/v_1 \\ k_2 &= \omega/v_2 \end{aligned} \qquad \begin{aligned} f(x,t) &= \begin{cases} A\cos(k_1 x - \omega t) + B\cos(-k_1 x - \omega t) & (x < 0) \\ C\cos(k_2 x - \omega t) + D\cos(-k_2 x - \omega t) & (x > 0) \end{cases} \\ \text{or it} &= \begin{cases} Ae^{i(k_1 x - \omega t)} + Be^{i(-k_1 x - \omega t)} & (x < 0) \\ Ce^{i(k_2 x - \omega t)} + De^{i(-k_2 x - \omega t)} & (x > 0) \end{cases} \end{aligned} \tag{7.30}$$

There's something missing. This works for $x < 0$ and for $x > 0$, but what about $x = 0$? It does *not* work there, at least not automatically, and this leads to the subject of boundary conditions. You do not have a solution to the differential equation if you don't have it everywhere. If it misses one point, then it's not a solution.

Boundary Conditions
For the particular problem of a string with two mass densities, the boundary conditions are

$$f \text{ is continuous}, \qquad \partial f/\partial x \text{ is continuous} \tag{7.31}$$

and where do these come from? I will show two different ways to get the answer. One by physical reasoning, the other mathematical. Fortunately they give the same answer.

The string hasn't broken. That is the first condition, the continuity of f.

Next, look at the point of attachment; did you tie a knot to attach the strings together? If so that knot has a mass and you must apply $\vec{F} = m\vec{a}$ to it. In specifying the mass density $\mu(x)$ I tacitly assumed that there is no knot at $x = 0$, just some sort of massless attachment. The force F_y on that zero-mass connection must be zero: $a_y = F_y/m = F_y/0$. Let the point $-\epsilon$ be a little left of zero and $+\epsilon$ a little to the right of zero. The equation (7.2) then says

$$\begin{aligned} F_y(\text{at zero}) &= -T(-\epsilon)\frac{\partial f}{\partial x}(-\epsilon, t) + T(+\epsilon)\frac{\partial f}{\partial x}(+\epsilon, t) \quad \text{(in the limit)} \\ &\longrightarrow -T\frac{\partial f}{\partial x}(0-, t) + T\frac{\partial f}{\partial x}(0+, t) = 0 \end{aligned} \tag{7.32}$$

---

* Do they have the same frequency? This is addressed at the end of this section.

This says that the derivative is continuous. The slope just to the right of the origin is the same as the slope just to the left. What if the knot tying the two parts of the string together isn't massless? Then there is a different condition on the slope, a constraint on the discontinuity itself because the right side of equation (7.32) isn't zero in those circumstances. See problem 7.45. What about F_x? If the knot tying the strings together is massless, this better be zero; that's why T is continuous.

For a straight mathematical treatment of the boundary conditions, remember that you are trying to solve the differential equation (7.29) *everywhere,* not just to the left and to the right of the origin, but also *at* the origin. The method to find the boundary conditions is a typical math trick: assume the opposite and show that you get a contradiction. Assume that f *does* have a discontinuity at $x = 0$. Then differentiate it with respect to x in order to plug it into the differential equation. You can't easily differentiate a discontinuous function, so draw a graph assuming that it rises in a very short interval. Draw graphs of f, $\partial f/\partial x$, and $\partial^2 f/\partial x^2$. The first graph has a step (almost), so the second has a spike. The third (f'') has a positive then a negative spike. The density $\mu(x)$ has a step too (almost).
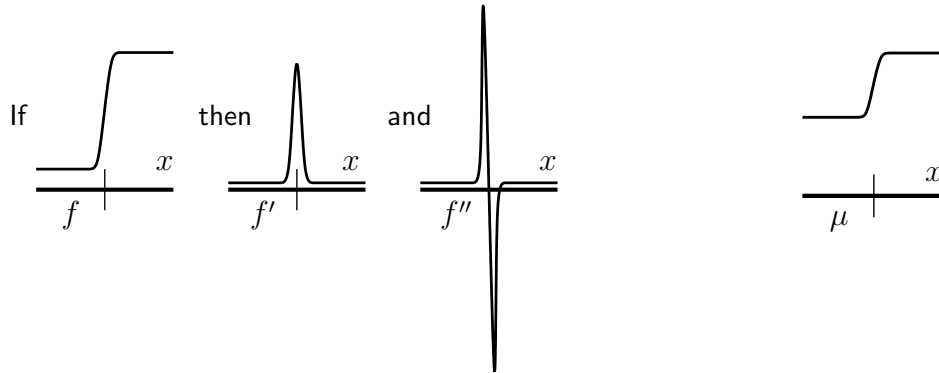
In Eq. (7.29) the time derivatives do not change the shape of the curve; only the x-derivatives and the factor $\mu(x)$ can do that. What is the graph of the terms in the differential equation?

$$T\frac{\partial^2 f}{\partial x^2} \overset{?}{=} \mu(x)\frac{\partial^2 f}{\partial t^2} \qquad \text{is graphically}$$
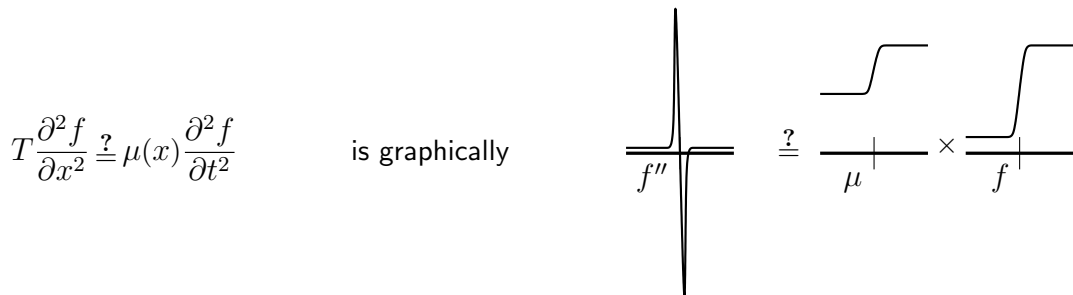
There is no way that the product of two steps can equal the double spike on the left of the equation. This is a contradiction, so f must be continuous.

Can it have a discontinuous derivative? No, and the reasoning is the same. Assume the slope of f changes abruptly at $x = 0$, and the graphs of the derivatives of f are
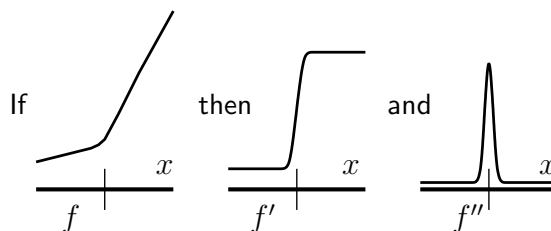
Again, try to insert these functions into the differential equation to see if it works.

$$T\frac{\partial^2 f}{\partial x^2} \stackrel{?}{=} \mu(x)\frac{\partial^2 f}{\partial t^2} \qquad \text{has the graph}$$
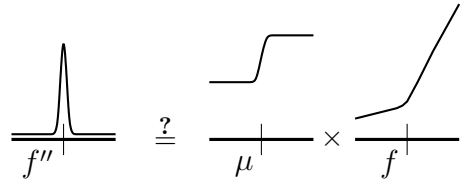
As before, this does not match. The right side has a step, but no spike. This assumed discontinuous slope doesn't work, so the slope must be continuous.

Why go through all this extra manipulation when the physical derivation was so simple and clear? The reason is that this sort of analysis appears in many other places where the physical intuition is not clear. For example, it lets you find the boundary conditions that occur when you solve Maxwell's equations. I've even used it in the somewhat esoteric subject of singular Sturm-Liouville equations, and I don't know *anyone* who can do that intuitively.

Back to the Problem
This started with a wave coming along a cord on which there is a discontinuous mass density. Now the boundary conditions are well-settled, it is time to attack the original problem. When a wave comes in from the left and hits the discontinuity, some part of the wave will continue forward and some part will be reflected. If the incoming wave has a frequency ω then the transmitted and the reflected waves will have that frequency too.

$$f(x,t) = \begin{cases} A\cos(kx - \omega t) + B\cos(kx + \omega t) & (x < 0) \\ C\cos(k'x - \omega t) & (x > 0) \end{cases} \qquad (7.33)$$

The boundary conditions are

$$\begin{aligned}
f(0-, t) = f(0+, t) &\implies A\cos(-\omega t) + B\cos(\omega t) = C\cos(-\omega t) \\
&\implies A + B = C \\
f'(0-, t) = f'(0+, t) &\implies -kA\sin(-\omega t) - kB\sin\omega t = -k'C\sin(-\omega t) \\
&\implies kA - kB = k'C
\end{aligned}$$

$\mu_1 \qquad\qquad \mu_2 \qquad\qquad\qquad\qquad \lambda_1 \qquad\qquad \lambda_2 \qquad\qquad$ Fig. 7.4
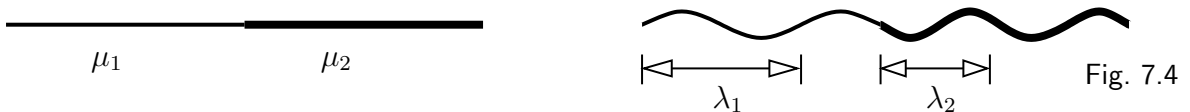
Solve the equations for B and C in terms of A because you can control the incoming wave amplitude (A) and you want to know where the wave goes.

$$A + B = C, \qquad kA - kB = k'C \longrightarrow B = \frac{k - k'}{k + k'}A \qquad \text{and} \qquad C = \frac{2k}{k + k'}A \qquad (7.34)$$

What happens to the energy in the wave? The incoming wave (A) carries energy and the two outgoing waves $(B$ and $C)$ carry energy. You may then expect conservation of energy to say that $A = B + C$, or maybe $A^2 = B^2 + C^2$. Neither is correct; compute them and find out. Why not? Simply that these coefficients don't represent the power. To get that you must use Eq. (7.24) applied to the wave Eq. (7.33). On the left, $x < 0$ this is

$$\begin{aligned}
P &= -T\frac{\partial f}{\partial t}\frac{\partial f}{\partial x} \\
&= -T\big[A\omega\sin(kx - \omega t) - B\omega\sin(kx + \omega t)\big]\big[-Ak\sin(kx - \omega t) - Bk\sin(kx + \omega t)\big] \\
&= T\omega k A^2\Big[\sin^2(kx - \omega t) - \frac{B^2}{A^2}\sin^2(kx + \omega t)\Big] \qquad (x < 0)
\end{aligned}$$

Notice how nicely this exhibits two terms that you can easily interpret as flow right minus flow left. Next, on the right the power flow is

$$P = -T\left[C\omega\sin(k'x - \omega t)\right]\left[-Ck'\sin(k'x - \omega t)\right] = TC^2\omega k'\sin^2(k'x - \omega t)$$

$$= T\omega k A^2 \frac{k'}{k}\frac{C^2}{A^2}\sin^2(k'x - \omega t) \qquad (x > 0)$$

The last manipulation makes it easier to compare the various expressions. Now all of them have the factor $T\omega k A^2$. Except for this factor and the $\sin^2$ factors, the power flow is

$$\text{1 (incoming from the left),} \qquad \frac{B^2}{A^2} = \left(\frac{k - k'}{k + k'}\right)^2 \text{ (outgoing to the left),}$$

$$\frac{k'}{k}\frac{C^2}{A^2} = \frac{k'}{k}\left(\frac{2k}{k + k'}\right)^2 \text{ (outgoing to the right)}$$

(7.35)

The two outgoing power factors add to

$$\left(\frac{k - k'}{k + k'}\right)^2 + \frac{k'}{k}\left(\frac{2k}{k + k'}\right)^2 = \frac{k^2 - 2kk' + k'^2 + 4kk'}{(k + k')^2} = 1$$

and energy is conserved.

In trying to understand why light is reflected at a window, you are solving Maxwell's equations involving electric and magnetic fields. The details of the solution are different, but for the case that the light is coming in perpendicular to the surface, the resulting equations (7.35) are identical. If light hits the glass at an angle other than along the normal, then the equations are more complicated and depend on the polarization of the light. (Look up the Fresnel equations.) For light you usually see the equations (7.35) expressed in terms of the index of refraction, so that the reflected and transmitted intensities for light going from vacuum or air to glass with index $n = c/v = (\omega/k)/(\omega/k') = k'/k$ are

$$\text{1 (incoming to the surface),} \qquad \frac{B^2}{A^2} = \left(\frac{1 - n}{1 + n}\right)^2 \text{ (reflected from the glass),}$$

$$\frac{k'}{k}\frac{C^2}{A^2} = n\left(\frac{2}{1 + n}\right)^2 \text{ (transmitted into the glass)}$$

(7.36)

Example

▶ Do this calculation using complex exponentials instead. The setup is already there in the Eqs. (7.30). For the same wave entering from the left you again have $D = 0$, so

$$f(x, t) = \begin{cases} Ae^{i(kx - \omega t)} + Be^{i(-kx - \omega t)} & (x < 0) \\ Ce^{i(k'x - \omega t)} & (x > 0) \end{cases}$$

The $e^{-i\omega t}$ is a common factor, so when you apply the continuity conditions, it will cancel. At $x = 0$ then

$$f(0-, t) = f(0+, t) \longrightarrow A + B = C \qquad \text{and}$$

$$\frac{\partial f}{\partial x}(0-, t) = \frac{\partial f}{\partial x}(0+, t) \longrightarrow ikA - ikB = ik'C$$

Multiply the first of these by ik and add to get C. Multiply the first by ik' and subtract to get B. The results are the same as before.

Conservation of Frequency

In the equations (7.30) I assumed that the frequency is the same on both sides of the boundary. It's easy to believe, but shouldn't it be possible to prove it? Assume that they are *not* the same and discover the consequences:

$$k_1 = \omega_1/v_1 \qquad f(x,t) = \begin{cases} A\cos(k_1 x - \omega_1 t) + B\cos(-k_1 x - \omega_1 t) & (x < 0) \\ C\cos(k_2 x - \omega_2 t) + D\cos(-k_2 x - \omega_2 t) & (x > 0) \end{cases}$$
$$k_2 = \omega_2/v_2$$

What happens at the boundary? The wave is supposed to be continuous, so this is

$$f(0-,t) = f(0+,t) \longrightarrow$$
$$A\cos(-\omega_1 t) + B\cos(\omega_1 t) = C\cos(-\omega_2 t) + D\cos(\omega_2 t)$$
$$(A+B)\cos(\omega_1 t) = (C+D)\cos(\omega_2 t)$$

If either one of $(A+B)$ or $(C+D)$ is non-zero, this is a contradiction. The two cosines are linearly independent. Specifically, take the case $t = 0$ then take $t = \pi/(2\omega_2)$ for which the right side is zero but not the left. What else is needed? Don't forget that the slope is continuous too. That gives still another equation involving $k_1(A-B)$ and $k_2(C-D)$. They have to vanish also and that makes the whole solution identically zero. This nails down the intuitively plausible statement that the transmitted frequency is the same as for incoming frequency. If the power company is supposed to delivery electricity at $60\,\text{Hz}$ and you discover that it is only $59.9\,\text{Hz}$, would you believe them if they ascribe it to frequency loss in the lines?

7.8 Standing Waves
A guitar string doesn't have a wave moving toward plus or minus infinity. The wave stops at the end of the string, so I need different kinds of solutions for this circumstance. I will consider the case for which T and μ are constants, Eq. (7.7), but I'll use a method for solving it that is common in these contexts: separation of variables. I could simply guess the solution, as I did in Eq. (7.8), but this method is almost as easy and it is of great general utility. It is in an extension of the ideas in section 0.8 about the separation of ordinary differential equations.

Assume a solution in the form of a product of functions, one of x and one of t.

$$f(x,t) = F(x)G(t)$$

Not all solutions will be like this, but the important result is that all solutions are *sums* of solutions like this. That means that if I can find all the solutions that are products then I can add them and get all the other solutions that aren't.* To see if it works, plug in.

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = 0 = \frac{d^2 F(x)}{dx^2}G(t) - \frac{1}{v^2}F(x)\frac{d^2 G(t)}{dt^2}$$

Now move one term to the other side of the equation and divide by the product FG.

$$\frac{1}{F(x)}\frac{d^2 F(x)}{dx^2} = \frac{1}{v^2 G(t)}\frac{d^2 G(t)}{dt^2}$$

The left side depends solely on the independent variable x. The right side depends solely on the independent variable t. I can change x any way that I want without changing t, and in doing so the right side stays constant. That means that the left side must be constant. A similar argument holds for the right side, allowing t to vary. And the constants have to be the same.**

$$\frac{1}{F(x)}\frac{d^2 F(x)}{dx^2} = C, \qquad \frac{1}{G(t)}\frac{d^2 G(t)}{dt^2} = Cv^2 \tag{7.37}$$

---

* *Proving* this is hard work. I won't.

** If T or μ are functions of x then v is too, and you'd better move the v^2 to the left side to make this work. But then you must go back to the form where T is inside the derivative on the left anyway.

These are ordinary differential equations, and they're the same equation that I spent a whole chapter on: the harmonic oscillator. Or are they? Harmonic motion requires that C is negative and how do I know that? At this point I don't. C can be anything, even complex, and it provides a valid solution to the wave equation. The equation doesn't know that the string is tied down at its ends—that fact produces a set of *boundary conditions* that will determine the details of the solution, including the nature of C.

In order to represent a system that's tied at the two ends I have to specify the function at those points. Pick the coordinate x to be zero at the left and L at the right. The y-value is zero where the string is at rest.

$$y = f(0, t) = 0 \qquad \text{and} \qquad y = f(L, t) = 0$$

Apply this to the separated solution $F(x)G(t)$ and it implies $F(0) = 0$ and $F(L) = 0$. The equation that F satisfies is $F'' = CF$, and if C is a positive number its solutions are real exponentials. A hyperbolic sine vanishes at one point, but *only* at one point, so you can't make it match these boundary conditions at both ends. If C is negative you have sinusoidal solutions, and these vanish at many places, allowing the possibility of matching the boundary conditions at both places. (What if $C = 0$?)

If C is negative, denote it by $C = -k^2$, then

$$\frac{d^2 F(x)}{dx^2} = -k^2 F \implies F(x) = A \sin(kx + \delta)$$

You must use the same constant for the G equation,

$$\ddot{G} = Cv^2 G = -k^2 v^2 G \implies G(t) = B \sin(\omega t + \delta') \qquad \text{where} \qquad \omega^2 = k^2 v^2 \qquad (7.38)$$

Now apply the boundary conditions.

$$f(0, t) = F(0)G(t) = 0, \qquad \text{so} \qquad F(0) = A \sin \delta = 0 \qquad \text{and} \qquad \delta = 0$$

You can pick another δ, such as π or -9π, but that just changes everything by an overall minus sign and just redefines A.

Next the other end of the string:

$$f(L, t) = F(L)G(t) = 0 \implies F(L) = A \sin kL = 0 \qquad \text{so} \qquad kL = n\pi$$

for integer n. Now put them together. $\omega = kv$, so

$$f(x, t) = A \sin\left(\frac{n\pi x}{L}\right) \sin(\omega_n t + \delta_n) \qquad \text{where} \qquad \omega_n = \frac{n\pi v}{L}, \qquad n = 1, 2, 3, \dots \qquad (7.39)$$

Negative n are unnecessary here because they would just reproduce the same functions as do the positive n. The δ_n's are arbitrary.

Some pictures of these waves at maximum displacement, *i.e.* when $\sin(\omega_n t + \delta_n) = 1$:

$$n = 1 \qquad n = 2 \qquad n = 3 \qquad n = 4$$

Fig. 7.5

Stroboscopic pictures of the motion
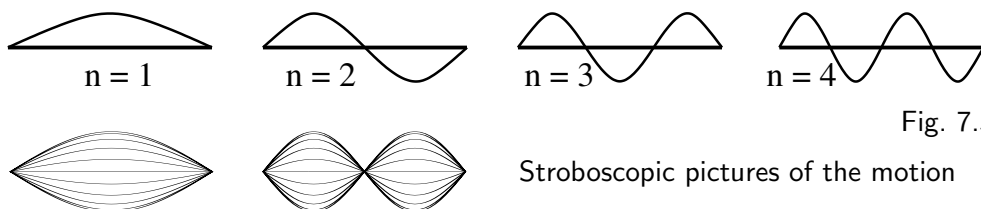
The first set of graphs show the waves at one instant, the peak displacement for each mode. The second set shows the waves at eleven instants, equally spaced in time from peak positive to peak negative amplitude, *i.e.* over half a period of oscillation.

The respective frequencies for various n are $\omega_1 = \pi v/L$, $\omega_2 = 2\pi v/L$, *etc.*, and they are the sequence of harmonics that heard when you pluck the string on a musical instrument.

The lowest frequency, the fundamental, is $\omega_1 = \pi\sqrt{T/\mu}\,/L$. It says that the greater the tension in the string, the higher the frequency. The same happens if the mass density of the string is less. Tuning such a musical instrument means adjusting the tension. Could you change the mass density instead? Actually yes. Look at the strings of a guitar or a piano and compare the strings for low frequencies and those for high frequencies. The former have higher linear mass density.

You can hear the separate harmonics from a guitar string if you pluck it properly. Touch the tip of your finger to the midpoint of the string and pluck the string with your thumb, immediately releasing contact from your finger. You will hear the upper harmonic ($n = 2$) because your finger damped out the $n = 1$ motion without affecting the $n = 2$ component. Similarly place your finger at the 1/3 or 1/4 point to hear still higher harmonics. An expert musician can use this technique to good effect.

Orthogonality

These standing wave solutions have a special property that won't appear particularly important now, but it will later. The idea of orthogonal vectors is one that you're familiar with: $\vec{A} \cdot \vec{B} = 0$. There is a similar concept for waves. For each positive integer n, the equation (7.39) provides a mode of oscillation shaped as

$$u_n(x) = \sin\left(n\pi x/L\right), \qquad n = 1,\ 2,\ 3,\ \dots$$

Take two different such modes and evaluate the integral

$$I = \int_0^L dx\ \sin\left(\frac{n\pi x}{L}\right)\sin\left(\frac{m\pi x}{L}\right) \qquad \text{for } n \neq m \tag{7.40}$$

This simply requires the right trigonometric identity, and that is

$$\sin x \sin y = \tfrac{1}{2}\left[\cos(x+y) - \cos(x-y)\right]$$

$$I = \int_0^L dx\ \frac{1}{2}\left[\cos\left(\frac{(n+m)\pi x}{L}\right) - \cos\left(\frac{(n-m)\pi x}{L}\right)\right]$$

$$= \frac{1}{2}\left[\frac{L}{(n+m)\pi}\sin\left(\frac{(n+m)\pi x}{L}\right) - \frac{L}{(n-m)\pi}\sin\left(\frac{(n-m)\pi x}{L}\right)\right]_0^L$$

$$= 0$$

These functions u_n and u_m are said to be orthogonal to each other because of this integration. For an analogy, you can compute the common three-dimensional scalar product as $A_x B_x + A_y B_y + A_z B_z$, a sum over three indices running over x, y, and z. The integral Eq. (7.40) is analogous to a sum over a continuous index that runs from zero to L.

In what will probably seem like very different sorts of problems, essentially the same idea will appear in the equations (8.42) and (10.21).

7.9 Other Waves

The structure of waves on a string with constant tension and constant mass density is so simple that is is misleading. Studying this is like studying electromagnetic waves in a vacuum. They are important, but as soon as you encounter matter they become complicated. The speed of light is a function of the wavelength (rainbows); the speed of light depends on the polarization of the wave (birefringence,

optical activity); the speed of light *by the definition of speed that you've seen in this chapter so far* can be greater than c, its speed in vacuum. In other words it is a lot more complicated than you may think.

What about the most commonly observed wave in nature, surface waves on the ocean. People saw these long before anyone knew that sound is a wave, much less that light is. Oddly enough, these ocean waves are some of the more complicated sorts of waves to analyze mathematically. For example, the wave speed is $v = \sqrt{g\lambda/2\pi}$ for deep water. This means that long wavelength waves travel much faster than short wavelength waves. A little ripple in a pond can travel at walking speed or less. A tsunami can travel at hundreds of kilometers per hour. The equations that these waves obey is *not* the wave equation of Eq. (7.7), the one that this chapter is devoted to. It is very different, and its wave solutions appear because of some very special boundary conditions at the ocean surface.

That you can listen to music at all without its sounding garbled tells you that for sound in air and within the range of human hearing the dependence of the sound speed on wavelength must be very slight. The speed does change at ultrasonic frequencies because at those very short wavelengths heat can flow quickly from a compression region to a rarefaction region and make the process nearly isothermal. At low frequencies and long wavelengths the distance is too great for this heat transfer to occur in the short time between wave crests, and in the low frequency domain that we are accustomed to hearing, the compression is adiabatic. That makes it harder to compress the air than it is in the isothermal case, and the speed for low frequencies is larger by the factor $\sqrt{\gamma = 7/5} \approx 1.18$. γ is the ratio of specific heats discussed so extensively in thermodynamics, and the transition between the two speeds occurs in the neighborhood of 100 MHz for air, so on this scale humans can hear only very low frequencies.

Earthquake waves have speeds that depend on whether they are longitudinally or transversely polarized. And then there are torsional waves in which the rocks are twisted instead of being pushed forward and backward or side to side. This difference in the speed of different types of earthquake waves allows geologists to get information about what is under the ground. It also allows them to determine the distance to an earthquake that occurs a thousand kilometers away — without receiving a phone call from someone in the middle of the quake itself.

7.10 Other Velocities

I have alluded to that fact that the definition of the velocity of a wave is more complicated than what is in Eq. (7.10), $v = \omega/k$. That expression is called the phase velocity because it describes the motion of a point having constant phase, $kx - \omega t =$ constant, but it is not the whole story. There are four different definitions for the velocity of a wave:

phase velocity,   group velocity,   information (or signal) velocity,   energy velocity

The first two are all that I will describe here, as the other two get very technical very fast. They were described by Sommerfeld and Brillouin.*

If a wave has the property that its phase velocity ω/k is itself independent of frequency, that's all there is. It's the end of the subject because all these four sorts of wave velocity are then identical. A light wave in a vacuum is like this. The more interesting problem is the one for which the phase velocity is *not* independent of frequency. The most visible example is that of light when it is passing through a medium such as water or glass or even air, because in those cases the index of refraction, $n = c/v_{\text{phase}}$, depends on the frequency of the light, and that is the reason there are rainbows. This is *dispersion*.

Any real wave has a beginning and an end, and contains a mixture of frequencies. If these frequencies travel at different speeds, the wave cannot maintain the exact same shape that it did at

---

* Leon Brillouin: *Wave Propagation and Group Velocity*, Academic Press 1960, still a standard reference.

the beginning. It will deform, typically spreading its width. It can in some circumstances have truly odd behavior.

Analyzing this in detail takes Fourier analysis, but I will take an easier route and examine the simplest possible case to show how the concept of group velocity arises. Any real wave is a mixture of many frequencies, and the simplest circumstance to deal with is two frequencies. What happens when there are two waves moving in the same direction, and one has a frequency slightly different from the other? Assume that they have the same amplitude and have frequencies ω and ω'. The two waves will then have (phase) velocities $v = \omega/k$ and $v' = \omega'/k'$, but you *cannot* simply say that "the velocity" of the wave is the average of these two numbers. Notice that the wave numbers k and k' aren't the same either.

The total wave is

$$f(x,t) = A\cos(kx - \omega t) + A\cos(k'x - \omega' t) \tag{7.41}$$

You need a trig identity here, problem 0.56, in order to add these two cosines.

$$\cos x + \cos y = 2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)$$

$$f(x,t) = 2A\cos\left(\frac{k+k'}{2}x - \frac{\omega+\omega'}{2}t\right)\cos\left(\frac{k-k'}{2}x - \frac{\omega-\omega'}{2}t\right) \tag{7.42}$$

Picture this wave at time zero, and assume that the frequencies are close, so the wave numbers k and k' are close also. That means that the $(k - k')$ in the second factor is much smaller than the $(k+k')$ in the first, so the first cosine factor oscillates in space much more rapidly than the second. The whole function at this instant is the product of two spatial oscillations: one with a short wavelength and one with a long wavelength: $f(x,0) = 2A\cos\left((k + k')x/2\right)\cos\left((k - k')x/2\right)$.
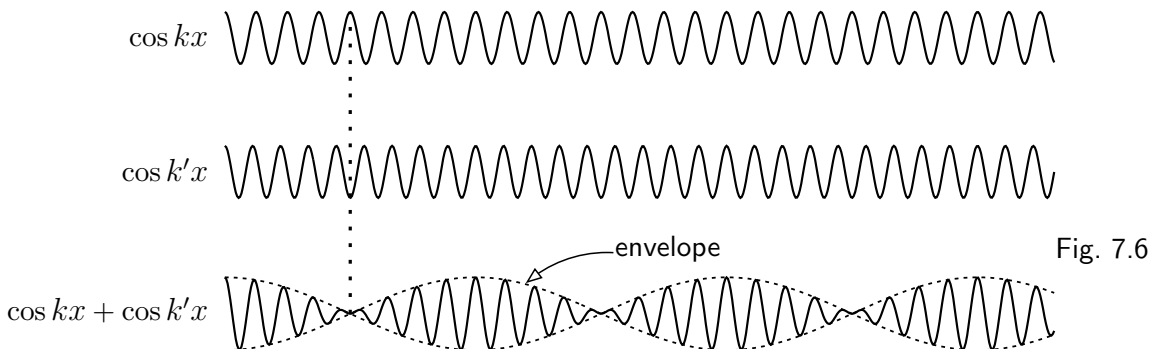
Fig. 7.6

The outline of the sum, the *envelope*, follows the $\pm\cos\left((k - k')x/2\right)$ factor, and that is shown with the slowly-varying curves. When this envelope drops to zero, you can see that the two component waves in the upper two graphs are $180°$ out of phase. Follow the vertical dotted line.

The velocity of one of these waves alone is $\omega/k = v_k$, and that of the other wave is $\omega'/k' = v_{k'}$. The combined wave does not have a single velocity, so examine the two factors in Eq. (7.42). Each factor has its own phase velocity.

$$
\begin{aligned}
&\cos\left(\frac{k+k'}{2}x - \frac{\omega+\omega'}{2}t\right) &&\text{has phase velocity} &&\frac{\omega+\omega'}{k+k'}\\
&\cos\left(\frac{k-k'}{2}x - \frac{\omega-\omega'}{2}t\right) &&\text{has phase velocity} &&\frac{\omega-\omega'}{k-k'}
\end{aligned} \tag{7.43}
$$

Unless the ratios ω/k and ω'/k' are the same, these two velocities can be quite different *even if the two frequencies are very close!* If the two frequencies are close, the first of these velocities, $(\omega+\omega')/(k+k')$,

is close to the average of v_k and $v_{k'}$; all three are about the same, the phase velocity. The second equation however has a velocity $\Delta\omega/\Delta k$ that is almost a derivative and there's no reason to expect this to be close to any of the other velocities. That would be like saying that the slope of a function, $f'(x)$, is the same as $f(x)/x$.

$$\frac{\omega - \omega'}{k - k'} \longrightarrow \frac{d\omega}{dk} = v_{\text{group}} \tag{7.44}$$

This is called the group velocity. The shorter wavelength waves within the envelope move at the phase velocity, while the envelope of the wave moves at the group velocity. The still image on the previous page is animated at the link

www.physics.miami.edu/nearing/mathmethods/groupvelocity.html and in more detail at
www.falstad.com/dispersion/groupneg.html

How different can these velocities be? In quantum mechanics you will see a relation between ω and k that follows $\omega = \hbar k^2/2m$, so

$$v_{\text{phase}} = \frac{\omega}{k} = \frac{\hbar k}{2m}, \qquad \text{while} \qquad v_{\text{group}} = \frac{d\omega}{dk} = \frac{\hbar k}{m}$$

This group velocity is twice the phase velocity. For ocean waves, see problem 7.27. For an example in which the group velocity can even be in the opposite direction from the phase velocity, see Eq. (10.42) with the graph, Figure 10.12. For light travelling in a medium, both the phase and group velocities can, under some circumstances, exceed c, but the signal and energy velocities (the ones that I'm not doing here) will always stay below c.

7.11 Waves and Tides

Near the end of chapter five, at page 173, there was a discussion of ocean tides and some of the probably confusing and certainly counter-intuitive aspects of the subject. In particular, how is it that a high tide can occur when the Moon is near the horizon? Shouldn't the high tide be more-or-less underneath the Moon (or at the antipode) as in the picture at Eq. (5.40)? In describing the phenomena involved I did a lot of hand-waving and made many qualitative statements. Here I propose to present some quantitative analysis to back up the previous work.

A simplified model that shows the phenomena uses a wave on a string. This time however, wrap the string around in a circle and set it spinning the way some people can spin a lasso. The rope forms a circle and is rotating about the central axis perpendicular to the plane of the circle. To simplify the problem, let it be rotating in empty space, and the only tension in the rope is there *because* it is spinning. The rope is still a rope, so the basic equation is $\vec{F} = m\vec{a}$, and that in turn implies the wave equation Eq. (7.3) but without the μg or $b\,\partial f/\partial t$ terms.

The linear mass density is no different from before, but what is the tension? Look at the arc of the rope having length $R\Delta\phi$. At the two ends the rest of the rope pull on it with the tension T, and the only component of that force that survives is the one pointing toward the center of the circle. The vector tangent to the circle rotates by $\pm\Delta\phi/2$ as it goes from the midpoint of the arc to one or the other of its ends, so its component toward the center is $T\sin(\Delta\phi/2) \approx \frac{1}{2}T\Delta\phi$ and the total force is on the arc is $2T\sin(\Delta\phi/2) \approx T\Delta\phi$.
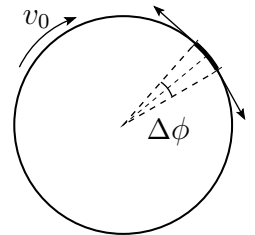
The mass of this segment is $\mu R\Delta\phi$ and its acceleration is toward the center with magnitude $v_0^2/R = R\omega_0^2$. Here v_0 is the speed at which the rope is moving around. Combine these:

$$T\Delta\phi = \mu R\Delta\phi \cdot v_0^2/R \longrightarrow \frac{T}{\mu} = v_0^2$$

From Eq. (7.7), this is the familiar relation for the speed at which a wave will move along the string. It says that the wave speed is the same as the speed at which the rope itself is moving: v_0. The wave equation will then govern the motion of waves around the loop in either direction.

Not so fast. In which coordinate system is the speed of the wave $\sqrt{T/\mu}$? *Not* the one where you are standing on the side watching the rope spinning around. The coordinate system in which the wave equation first appeared, back in section 7.1, was the one in which you were standing next to the rope and the rope wasn't going anywhere. Here, that is the system rotating with angular speed $\omega_0 = v_0/R$. In that system the waves will move clockwise and counterclockwise with speed v_0. What do these waves look like to the stationary bystander? In one direction the wave will move at speed $2v_0$ and in the other direction it is a standing wave. To see in more persuasive detail what happens when you try to run after a wave, do problem 7.34, or at least study the answer stated there.

Now to get back to tides: The ocean tides on Earth will correspond to the waves moving around this circular string. To make the parallel more realistic, look at the rope in the rotating system, so that the waves will go in either direction at speed v_0. Now what will correspond to the tidal force from the Moon? That force, because of the Earth's acceleration toward the Moon, behaves as if there are two pulls that tug on opposite sides of the Earth and its oceans, as in section 5.6. In that system, the average total force (vector) over the whole surface is zero as in problem 5.37. In this simplified model with a rope, the function that does this is an added radial force that varies as $\cos 2\phi$, or because the Moon is orbiting the Earth, as $\cos\left(2(\phi - \Omega t)\right)$. For the Earth-Moon system Ω is a little less than $2\pi/$day. Remember: We are now in the Earth-centered system, in which the Moon takes slightly more than one day to go around us because the Moon's orbit is in the same direction as the Earth's rotation on its axis, though about 28 times slower.

Translate the wave equation (7.3) into angular variables and add the external tidal force. $x \to R\phi$

$$\frac{\partial}{\partial x}\left(T\frac{\partial f}{\partial x}\right) - \mu g - b\frac{\partial f}{\partial t} = \mu\frac{\partial^2 f}{\partial t^2} \longrightarrow T\frac{\partial^2 f}{\partial x^2} = \mu\frac{\partial^2 f}{\partial t^2}$$

$$\longrightarrow \quad \frac{T}{R^2}\frac{\partial^2 f}{\partial \phi^2} + \mathcal{F}\cos\left(2(\phi - \Omega t)\right) = \mu\frac{\partial^2 f}{\partial t^2}$$

Here, $\mathcal{F}$ is the peak force density (force per length) of the mock tidal force. This is an inhomogeneous partial differential equation, solved the same way that you solve inhomogeneous ordinary differential equations. First, find some one solution to the inhomogeneous part, then find the general solution to the homogeneous part, then add them.

Two derivatives of a cosine come back to a cosine, so try that for the inhomogeneous term.

$$f \stackrel{?}{=} A\cos\left(2(\phi - \Omega t)\right)$$

$$\longrightarrow \frac{T}{R^2}(-4A)\cos\left(2(\phi - \Omega t)\right) + \mathcal{F}\cos\left(2(\phi - \Omega t)\right) \stackrel{?}{=} \mu\left(-4A\Omega^2\right)\cos\left(2(\phi - \Omega t)\right)$$

$$\mathcal{F} = \mu\left(-4A\Omega^2\right) - \frac{T}{R^2}(-4A)$$

$$\longrightarrow A = \frac{\mathcal{F}/\mu}{4\left((T/\mu R^2) - \Omega^2\right)} = \frac{\mathcal{F}/\mu}{4\left(\omega_0^2 - \Omega^2\right)} \qquad\qquad \frac{T}{\mu R^2} = \frac{v_0^2}{R^2} = \omega_0^2 \qquad (7.45)$$

$\Omega < \omega_0$ $\qquad\qquad$ $\Omega > \omega_0$
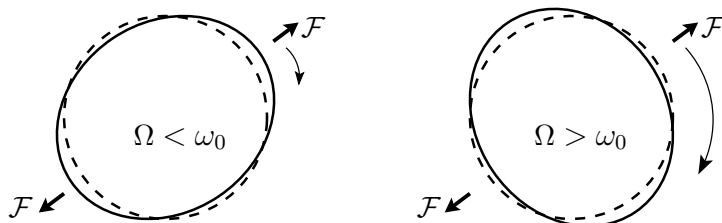
Fig. 7.7

The dashed circles represents the mean sea level, a constant height. Or, it represents the constant circle that the rope maintains when it is in equilibrium. The solid lines represent the tides going around the Earth daily. Or, they represent the disturbance in the rope as $\mathcal{F}$ moves around. The natural frequency of the wave around the rope is $\omega_0 = v_0/R$, and when the forcing frequency Ω is less than this, the factor in Eq. (7.45) is positive. When it is greater, the factor is negative. This is just like the results in Eq. (3.43) only without the complication of friction. Of course if $\Omega = \omega_0$ this solution doesn't work, but that is familiar from the previous study of resonance in chapter three. You need the damping effects to make it finite. That adds reality and complexity to the problem. In the case of the Earth-Moon system, there is friction and there are continents and there are enough other interesting features to make a profession of the subject.

For the Earth-Moon system, just how do ω_0 and Ω compare? As noted above, Ω is just a little less than $2\pi/\text{day}$. For ω_0 you need some information about the speed of *very* long wavelength ocean waves. This will depend strongly on ocean depth, but a number such as $200\,\text{m/s}$ is a plausible place to start. That corresponds to an ocean depth of 4 km. For ω_0, divide this speed by the radius, *but what radius?* Not necessarily the Earth's. For simplicity, assume that the Moon orbits above the equator. For a wave going around the Earth at the equator, $\omega_0 = (200 \text{ m/s})/R$ is right, but at other latitudes the radius of the circle around the Earth is $R\sin\theta$, and that is the relevant radius here.

At the equator

$$\omega_0 = \frac{200 \text{ m/s}}{6400 \text{ km}} = 4.33 \times 10^{-5}/\text{s} = 2.7/\text{day} < 2\pi/\text{day} \approx \Omega$$

In the Antarctic Ocean, $\theta \approx 30°$, so $\omega_0 \approx 2.7/\sin 30° = 5.4/\text{day}$. That is much closer to the forcing frequency Ω, so you expect (and get) larger tides.

7.12 An Algebraic Aside

How do you solve a cubic equation? (And what does this have to do with waves? Patience.) That depends on the cubic. Try to solve

$$x - 1 - 0.01x^3 = 0$$

If you have read section 0.11 or solved problem 0.60 or worked through Eqs. (4.9)-(4.11), you may think that you have done this before and you have, but not this way. What follows is a different approach to the same sort of problem, one that closely parallels what I will do for the wave equation. Just notice that $0.01 \ll 1$. There is a solution to the equation near $x = 1$, and I can use the equation itself to find a series representation for this root. The coefficient "0.01" is the expansion parameter. The idea is to assume the form of the series even though you do not know any of its coefficients. Then use the equation you're solving to *determine* those coefficients. Let $\epsilon = 0.01$ be the parameter.

$$x - 1 - \epsilon x^3 = 0, \qquad \text{and} \qquad x = x_0 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \cdots \tag{7.46}$$

You do not know any of the coefficients x_k, but plug the assumed solution into the equation anyway. In cubing the series, treat x_0 as one term and everything else as a second term; that way you are cubing a binomial.

$$x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots - 1 - \epsilon\left(x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots\right)^3 = 0$$

$$x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots - 1$$
$$-\epsilon\left(x_0^3 + 3x_0^2\left(\epsilon x_1 + \epsilon^2 x_2 + \cdots\right) + 3x_0\left(\epsilon x_1 + \cdots\right)^2 + \left(\epsilon x_1 + \cdots\right)^3\right) = 0$$

Expand everything and collect terms in powers of ϵ, multiplying out the square and cube of the parenthesized expressions and carefully getting all the terms with like powers of ϵ. This can get tedious, but carry on. This is very different from the iteration methods used earlier.

$$x_0 + \epsilon x_1 + \epsilon^2 x_2 + \cdots - 1 - \epsilon\left(x_0^3 + 3\epsilon x_0^2 x_1 + 3\epsilon^2 x_0^2 x_2 + 3\epsilon^2 x_0 x_1^2 + \cdots\right) = 0$$

For this to be a valid series expansion in ϵ, all the like coefficients of powers of ϵ must match.

$$\begin{aligned}
\epsilon^0 &: & x_0 - 1 &= 0 \\
\epsilon^1 &: & x_1 - x_0^3 &= 0 \\
\epsilon^2 &: & x_2 - 3x_0^2 x_1 &= 0 \\
\epsilon^3 &: & x_3 - 3x_0^2 x_2 - 3x_0 x_1^2 &= 0
\end{aligned} \tag{7.47}$$

Solve these equations in sequence to get

$$x_0 = 1, \qquad x_1 = 1, \qquad x_2 = 3, \qquad x_3 = 12, \cdots$$

The value of the root is $x = 1 + .01 + .01^2 \cdot 3 + .01^3 \cdot 12 + \cdots = 1.010312$ Put this into the original cubic and the value is $x - 1 - .01x^3 = -5.61 \times 10^{-7}$. For comparison, put $x = 1$ into the equation and the error is 10^{-2}. The next term in the series for x is $\epsilon^4 x_4 = 10^{-8} x_4$. This example shows the essential idea of perturbation theory. Continuing this to higher orders in ϵ produces 1.01031257881011 as a better approximation to the exact solution.

The spirit of this approximation is that you start with an easy solution to a simplified form of the problem and then improve the solution by using the complicated part to generate the correction. Are there ways other then series expansions to do this? Yes, the subject of numerical analysis and of perturbation theory is well-developed and there are many variations on this theme, which brings me to the reason for this diversion.

7.13 Perturbation Theory˙
If the mass density of the string isn't constant, Eq. (7.3) is more complicated than the simple wave equation. The same is true if the tension isn't constant throughout. Can this happen? Yes, easily. For a non-uniform density place a wad of chewing gum on the string. Or maybe use a bullwhip for the string. To get non-constant tension, simply hold the string vertically. The part of the string at the top has to support everything below, so the tension at the top will be greater than that at the bottom. Will this be a big effect in a musical instrument? No. The tension in that case is always far greater than the weight of the string, so this effect is not important, but if the string is not under great tension then this will alter the harmonic structure substantially. There *are* other examples of physical systems where the same equation appears and for which the effects of variable T (or its analog) are important.

There are a few special problems for which you can solve the equations exactly when T or μ are not constant, but instead of spending time on those I'll show a general procedure that's valid for any T or μ as long as they are *almost* constant.

The technique, *perturbation theory*, parallels the series solution for the cubic equation in the preceding section, and it starts from what you already know, the solution for the uniform string. Here the solution for the non-uniform case should be very close to the known, simple solution.

Take T to be constant and μ to be almost constant.

$$\mu(x) = \mu_0 + \mu_1(x), \qquad |\mu_1| \ll \mu_0$$

If μ_1 is absent, I know how to solve the problem for which the string is tied down at two ends; that's section 7.8 and Eq. (7.39). I said that the idea is to expand everything in power series, but what is the variable to use? I'll make one up. Change the preceding equation to

$$\mu(x) = \mu_0 + \epsilon\mu_1(x)$$

and use ϵ as a dimensionless parameter for the series. This allows me to keep track of the terms easily.

To be certain that I'm not missing anything, I will return to the original wave equation (7.3), but with $T = $ a constant and no gravity and no friction.

$$T\frac{\partial^2 f}{\partial x^2} = \mu(x)\frac{\partial^2 f}{\partial t^2}$$

The technique of separation of variables still works, even when μ (or even T) depends on x.

$$f(x,t) = F(x)G(t), \qquad \text{then} \qquad T\frac{d^2 F(x)}{dx^2}G(t) = \mu(x)F(x)\frac{d^2 G(t)}{dt^2}$$

To separate variables this time, divide by $FG\mu$.

$$\frac{T}{\mu F}\frac{d^2 F}{dx^2} = \frac{1}{G}\frac{d^2 G(t)}{dt^2}$$

and again the equation is separated, with all the x's on one side and all the t's on the other. They must therefore be (the same) constant. Back in Eq. (7.37) I called it C, but in the end it turned out to be proportional to ω^2, so I'll use that notation from the outset.

$$\frac{1}{G}\frac{d^2 G(t)}{dt^2} = -\omega^2, \qquad \frac{T}{\mu F}\frac{d^2 F}{dx^2} = -\omega^2 \tag{7.48}$$

The time equation is unchanged, so the solution is the familiar sine, cosine, or exponential: $e^{i\omega t}$ or $\sin(\omega t + \delta)$. The difference is in the equation for F, and that's where the action occurs. I take the same boundary conditions as before, $F(0) = F(L) = 0$, so that the string is tied down.

$$T\frac{d^2 F}{dx^2} = -\mu\omega^2 F \tag{7.49}$$

Now however μ, F, and ω^2 are different, though they should be *close* to the unperturbed case. Write all of them in terms of a power series in ϵ, just as in Eq. (7.46). The density μ of course has only two terms, so it hardly warrants the name power series.

$$\begin{aligned}
\mu &= \mu_0 + \epsilon\mu_1 \\
F &= F_0 + \epsilon F_1 + \epsilon^2 F_2 + \epsilon^3 F_3 + \cdots \\
\omega^2 &= \omega_0^2 + \epsilon\omega_1^2 + \epsilon^2\omega_2^2 + \cdots
\end{aligned} \tag{7.50}$$

The coefficients F_0, F_1, ω_2^2, *etc.* are unknowns to be determined. The expansion for ω may look a little odd, but it is ω^2 that shows up as the natural parameter, not ω itself, so this is the way to write it. It would probably be more consistent notation to call the coefficients $(\omega^2)_0$, $(\omega^2)_1$, ..., but that is being unnecessarily fussy and it is too clumsy, so I won't. Now plug in.

$$T\big[F_0'' + \epsilon F_1'' + \epsilon^2 F_2'' + \cdots\big] = -\big[\mu_0 + \epsilon\mu_1\big]\big[\omega_0^2 + \epsilon\omega_1^2 + \epsilon^2\omega_2^2 + \cdots\big]\big[F_0 + \epsilon F_1 + \epsilon^2 F_2 + \cdots\big]$$

This looks like I'm taking a difficult problem and turning it into an impossible one. Now I have an infinite number of unknowns. *But,*

The left side is a power series in ϵ. The right side is a power series in ϵ. They're supposed to agree no matter what ϵ is. For that to happen all the coefficients of like powers have to match. The coefficient of ϵ^0, the coefficient of ϵ^1, *etc.*

$$\begin{aligned}
TF_0'' &= -\mu_0\omega_0^2 F_0 \\
TF_1'' &= -\mu_0\omega_0^2 F_1 - \mu_1\omega_0^2 F_0 - \mu_0\omega_1^2 F_0 \\
TF_2'' &= -\mu_0\omega_0^2 F_2 - \mu_2\omega_0^2 F_0 - \mu_0\omega_2^2 F_0 - \mu_1\omega_1^2 F_0 - \mu_1\omega_0^2 F_1 - \mu_0\omega_1^2 F_1
\end{aligned} \tag{7.51}$$

This is now an infinite number of equations for an infinite number of unknowns. It doesn't look promising. First, some reassurance: I'm going to go just as far as the first two of these equations. F_2 will have to find itself.

Rearrange them.

$$TF_0''' + \mu_0\omega_0^2 F_0 = 0 \tag{a}$$
$$TF_1'' + \mu_0\omega_0^2 F_1 = -\mu_1\omega_0^2 F_0 - \mu_0\omega_1^2 F_0 \tag{b} \quad (7.52)$$
$$TF_2'' + \mu_0\omega_0^2 F_2 = -\mu_2\omega_0^2 F_0 - \mu_0\omega_2^2 F_0 - \mu_1\omega_1^2 F_0 - \mu_1\omega_0^2 F_1 - \mu_0\omega_1^2 F_1 \tag{c}$$

The systematic procedure to solve these:
1. Solve the 1st, easy equation.
2. Use the F_0 from step one to find the frequency correction ω_1^2 from the 2nd equation.
3. Solve the 2nd equation for F_1.
4. Use the F_1 from step two to find the frequency correction ω_2^2 from the 3nd equation.
5. Solve for F_2 *etc.*(I'm not going past step three anyway.)

I haven't yet chosen a particular problem that I want to solve, trying to leave the setup general. Now to be specific, I have to set down the boundary conditions, and the ones I choose are the same ones that I used before. Tie the string down at two ends. This says that $f(0,t) = f(L,t) = 0$ as before. Eq. (7.52)(a) is the same as Eq. (7.38), so it has the same solution, Eq. (7.39).

$$TF_0'' + \mu_0\omega_0^2 F_0 = 0 \longrightarrow F_0(x) = A\sin\left(\frac{n\pi x}{L}\right), \quad \text{with} \quad T\frac{n^2\pi^2}{L^2} = \mu_0\omega_0^2$$

Now comes the key observation that breaks open this complicated looking set of equations.

Take the unperturbed solution, $F_0(x)$, multiply the left-hand side of the equation (7.52)(b) by F_0, and integrate from 0 to L with respect to x, you get zero. Stated in the language of vectors,

$$F_0 \quad \text{is orthogonal to} \quad TF_1'' + \mu_0\omega_0^2 F_1 \tag{7.53}$$

whatever F_1 is (as long as it satisfies the same boundary conditions and vanishes at 0 and L).

The proof of this involves nothing more than a couple of partial integrations. I have to compute the following integral, showing that it is zero.

$$\int_0^L dx\, F_0(x)\left[TF_1'' + \mu_0\omega_0^2 F_1\right] \overset{?}{=} 0 \tag{7.54}$$

Work on the first term by itself, recalling that T and μ_0 are constants.

$$\int_0^L dx\, F_0(x)\, TF_1'' = TF_0(x)F_1'(x)\Big|_0^L - T\int_0^L dx\, F_0'(x)F_1'(x)$$

$$= TF_0(x)F_1'(x)\Big|_0^L - TF_0'(x)F_1(x)\Big|_0^L + T\int_0^L dx\, F_0''(x)F_1(x)$$

Put this into the integral (7.54).

$$TF_0(x)F_1(x)\Big|_0^L - TF_0'(x)F_1(x)\Big|_0^L + \int_0^L dx\,\left[TF_0''(x) + F_0(x)\mu_0\omega_0^2\right]F_1(x) = 0 \tag{7.55}$$

This last integral is zero because by assumption F_0 satisfies the original, unperturbed equation (7.52)(a), and that is precisely what appears in brackets as the coefficient of F_1 inside the integral. The other

terms are evaluated at the boundaries, 0 and L, and the boundary conditions that I placed on the problem, that $F(0) = 0 = F(L)$, make this zero because F_0 and F_1 will vanish at these points. This result parallels those you will see in seemingly very different contexts in Eqs. (8.40) and (10.19).

Now multiply Eq. (7.52)(b) by F_0 and integrate the whole equation from 0 to L. The left side is now zero.

$$\int_0^L dx\, F_0(x)\left[TF_1'' + \mu_0\omega_0^2 F_1\right] = 0 = \int_0^L dx\left[-\mu_1\omega_0^2 F_0^2 - \mu_0\omega_1^2 F_0^2\right] \tag{7.56}$$

The single unknown here is ω_1^2, the first correction for the frequency-squared. Solve for it.

$$\omega_1^2 = -\,\omega_0^2\int_0^L dx\,\mu_1(x)F_0(x)^2 \,\bigg/\, \mu_0\int_0^L dx\,F_0(x)^2 \tag{7.57}$$

The denominator is easy, and the numerator's complexity depends on μ_1. I didn't have to solve any complicated differential equations to get this result, I have only to do two integrals.

Example

▶ What if the string is tapered slightly, so that the mass density varies linearly from one end to the other? Take $\mu_1(x) = \alpha x$, making the string slightly thicker toward $x = L$.

$$F_0(x) = \sin(n\pi x/L), \qquad \text{and} \qquad \omega_0^2 = n^2\pi^2 T/\mu_0 L^2$$

The integrals are

$$\int_0^L dx\,\sin^2(n\pi x/L) = L/2, \qquad \text{and}$$

$$\int_0^L dx\,x\sin^2(n\pi x/L) = \int_0^L dx\left[(x - L/2) + L/2\right]\sin^2(n\pi x/L) = L^2/4 \tag{7.58}$$

The correction to the frequency is then to this order

$$\omega_1^2 = -\frac{\omega_0^2\alpha L^2/4}{\mu_0 L/2} = -\omega_0^2\frac{\alpha L}{2\mu_0} \qquad \text{so that} \qquad \omega^2 = \omega_0^2\left[1 - \frac{\alpha L}{2\mu_0}\right] \tag{7.59}$$

Is this plausible? If α is positive this is a decrease in the frequency, and that's just what you should expect because it represents an increase in the linear mass density over the μ_0 that I started with, and the density appears in the denominator of $\sqrt{T/\mu}$. And what about the frequency itself, not just the squared frequency? See problem 7.16. Are the dimensions of (7.59) right?

Example

▶ How is the frequency altered if a small mass is added at one point along the string? Perhaps a tenacious fly lands on it. In this case the perturbing mass density μ_1 is zero every place except very near the one point. Specifically a small mass m is added at the point $x = d$. In the numerator of Eq. (7.57) the integrand is then zero everywhere except very close to the point d.

$$\int_0^L dx\,\mu_1(x)F_0(x)^2 = \int_{d-\epsilon}^{d+\epsilon} dx\,\mu_1(x)F_0^2(x)$$

where 2ϵ is the tiny width of the mass. As F_0 is a smooth function it won't change significantly over this small range of x and I can approximate the integral by evaluating F_0 at $x = d$.

$$\approx \int_{d-\epsilon}^{d+\epsilon} dx\, \mu_1(x) F_0^2(d) = m\, F_0^2(d) = m \sin^2(n\pi d/L)$$

The denominator in the frequency correction is $\mu_0 L/2$, so the whole correction is

$$\omega_1^2 = -\omega_0^2 \frac{m \sin^2(n\pi d/L)}{\mu_0 L/2}$$

Notice how this varies with d. Take $n = 2$ for example. There are two humps ($\pm$) with a node at the center, $L/2$. If the mass is placed exactly at the node, then $\sin^2(2\pi L/2L) = 0$ and there is no shift in the frequency. This makes sense because the center point on the string isn't moving anyway, so it doesn't matter if there is a extra mass there or not. In the same $n = 2$ mode, if you put the extra mass at $L/4$ or $3L/4$ then the effect is maximum. It is negative; it has the correct dimensions.

Finding F_1

All the preceding calculation did was to get the lowest order correction for the frequency, and that's often all that you want. But what about the wave itself? Sometimes you need more than the frequency. That's Eq. (7.52)(b).

$$T F_1'' + \mu_0 \omega_0^2 F_1 = -\mu_1 \omega_0^2 F_0 - \mu_0 \omega_1^2 F_0$$

Now you know the right side of the equation because you've found ω_1^2 and you already knew ω_0^2 and F_0. This is an inhomogeneous linear differential equation for F_1, and you have seen examples of exactly this equation in chapter three. The sections 3.5 and 3.11 were devoted to it. The independent variable was t instead of x, and the dependent variable was x instead of F_1, but that makes no difference because the equations are the same. The only difference between this calculation and those is the extra use of boundary conditions.

      Why would you need F_1? One reason is that the equation for the *second* order correction to ω^2 uses it. That is in the third of the equations (7.52), and you need F_1 to determine the right side of that equation. Do the same thing as in equations (7.52)(b) and (7.56) and suddenly you have ω_2^2. For here and now, it is not worth pursuing this matter, but precisely this sort of calculation appears in quantum mechanics and the higher order corrections can be important in that context.

7.14 Stiffness

In the first section of this chapter I mentioned that I had left out some physics, the stiffness of the string. If you look at a guitar or inside a piano you will see that the low pitched notes have heavy, strings and they are stiff and harder to bend. How does this affect the frequencies you hear when the instrument is played? The tools in the preceding section on perturbation theory apply here too.

      The first question: how to describe stiffness mathematically? For a wire of circular cross section the extra force comes from a combination of Young's modulus (Y), the cross-sectional area (S), and a dimensionless factor ($\alpha = 1/4\pi$).

$$\mu \frac{\partial^2 f}{\partial t^2} = T \frac{\partial^2 f}{\partial x^2} - \alpha Y S^2 \frac{\partial^4 f}{\partial x^4} \tag{7.60}$$

I won't attempt to derive this here, I will simply use it. Does it at least have the correct dimensions?

      What does this do to the frequency of standing waves? You can solve this particular problem exactly, but treating the added term as a perturbation is enough for many circumstances. Start from

the same solution as Eq. (7.39) and introduce a dimensionless expansion parameter in front of the new term: ϵ. Use the same sort of expansion as in Eq. (7.50) though now there is no change in the mass density, no μ_1. There is no change in the tension either, so that doesn't need an expansion of its own. The x and t variables separate just as in Eq. (7.48), so that the equation (7.49) is

$$T\frac{d^2 F}{dx^2} - \epsilon \alpha Y S^2 \frac{d^4 F}{dx^4} = -\mu \omega^2 F \qquad (7.61)$$

Use Eqs. (7.50) for F and ω^2, and collect the terms in ϵ^0 and ϵ^1. I won't go beyond this.

$$T\frac{d^2 F_0}{dx^2} = -\mu \omega_0^2 F_0$$

$$T\frac{d^2 F_1}{dx^2} - \alpha Y S^2 \frac{d^4 F_0}{dx^4} = -\mu \omega_0^2 F_1 - \mu \omega_1^2 F_0$$

Rearrange this to put it in the form of Eqs. (7.51).

$$T\frac{d^2 F_0}{dx^2} + \mu \omega_0^2 F_0 = 0$$

$$T\frac{d^2 F_1}{dx^2} + \mu \omega_0^2 F_1 = \alpha Y S^2 \frac{d^4 F_0}{dx^4} - \mu \omega_1^2 F_0 \qquad (7.62)$$

The first of these gives the same standing wave as before,

$$F_0(x) = A \sin\left(\frac{n\pi x}{L}\right), \qquad \text{with} \qquad T\frac{n^2 \pi^2}{L^2} = \mu \omega_0^2$$

Now use the same equation (7.54) as before. Multiply the second of these equations (7.62) by F_0 and integrate from 0 to L.

$$\int_0^L dx\, F_0(x)\left[T\frac{d^2 F_1}{dx^2} + \mu \omega_0^2 F_1\right] = 0 = \int_0^L dx\, F_0(x)\left[\alpha Y S^2 \frac{d^4 F_0}{dx^4} - \mu \omega_1^2 F_0\right]$$

This determines the correction ω_1^2 for the frequency.

$$\omega_1^2 = \frac{\alpha Y S^2}{\mu} \frac{\int_0^L dx\, F_0(x) F_0''''(x)}{\int_0^L dx\, F_0(x)^2} = \frac{\alpha Y S^2}{\mu} \frac{n^4 \pi^4}{L^4}$$

The total frequency squared to this order is then

$$\omega^2 = \frac{T}{\mu} \frac{n^2 \pi^2}{L^2} + \frac{\alpha Y S^2}{\mu} \frac{n^4 \pi^4}{L^4} \qquad (7.63)$$

This says that the overtones of the string are not quite in a simple harmonic sequence, ω is approximately proportional to n, but the overtones are stretched to slightly higher values. Can you hear this? I can't, but a good piano tuner can.

Exercises

1 Do the algebra to get the equations Eq. (7.34).

2 Solve $x - 1 + .01x^4 = 0$ to $\pm 10^{-4}$.

3 Apply the equations for reflection (7.36) and determine the fraction of the light power reflected from ordinary window glass. While you're at it, what is the sum of the two expressions there?

4 What is the time derivative of the energy computed in Eq. (7.25) and do you believe it?

5 Is there any circumstance for which the energy calculated in Eq. (7.26) is a constant?

6 You are supposed to get electric power from the local utility as an alternating voltage at 60 cycles per second. You discover that it is coming in at only 59 cycles per second. The power company's response to your complaint is that this is due to frequency loss in the lines. Your response?

7 In the development of section 7.7, at the discontinuity the frequency ω stayed the same, but the wave number k changed. Why can't it be the reverse, with ω changing and k staying the same?

8 In the equation (7.29) does the original derivation of the wave equation in section 7.1 have to be modified? Perhaps adding a term in $d\mu(x)/dx$?

9 For light travelling through a dilute gas such as air, and in the visible part of the spectrum, the relation between ω and k is, to a good approximation $\omega = \alpha k + \beta k^2$. Here the constant α is close to c and the other term is small over the visible range. The index of refraction for blue light is larger than that for red light. Is β positive or negative?

10 On page 246 it says that the frequency at which the speed of sound in air changes is about 100 MHz. About what wavelength is this, and how does it compare to intermolecular distances in air?

11 For what index of refraction is the transmitted and the reflected power the same?

12 In Eq. (7.58) I used the manipulation $x = [(x - L/2) + L/2]$. Why does this help you to do the problem in your head? *Sketch some graphs.*

13 What is f if $f'(x) = f(x)/x$? And what does this have to do with group velocity?

14 In deriving Eq. (7.31) by physical reasoning, I tacitly assumed that the tension, T, is continuous. Why is that correct?

Problems

7.1 For the wave on a uniform string, $A\cos(kx - \omega t)$, compute the kinetic and potential energy densities and show that they are the same. Does this work for the more general function $f(x - vt)$?

7.2 You can derive Eq. (7.24) directly from the expression $\vec{F} \cdot \vec{v}$ for power in a simple mechanical system. Do so. In the usual basis just write the vectors $\vec{F}$ and $\vec{v}$ as in Figure 7.1. (Recall: $\vec{F}$ *on what by what?*)

7.3 A string can be tapered, as with a bullwhip. Stretch it between two posts and determine how much it sags, assuming the static case. What is the shape of the string now? Assume that the linear mass density is $\mu(x) = A + Bx$ for $0 < x < L$. Does it matter if $A = 0$? And does your result reduce to the original result for which the density is constant? Where is the lowest point and where is the center of mass of the string? For computing the center of mass, assume that the string is almost straight. Ans: If $A = 0$, the minimum is at $L/\sqrt{3}$, $< x_{cm}$.

7.4 Start from Eq. (7.21) again and do the partial integration on the second term instead, showing that you get the same result Eq. (7.24) this way too.

7.5 Show that the two equations (7.16) and (7.17) remain valid even if T and μ are functions of x. Then show that the equations for power, (7.23) and (7.24) are also valid here.

7.6 The wave speed for water waves in very deep water depends at most on the density, ρ, the wavelength, λ, and g. (**a**) Use dimensional analysis to determine the only possible form for the wave speed in terms of these. (**b**) For shallow water or long wavelength the speed depends on ρ, g, and the depth d. What is the form for the speed now? (**c**) A general expression for wave speed is $v = \left[(g\lambda/2\pi)\tanh(2\pi d/\lambda)\right]^{1/2}$. Show that this reproduces the first two results.

7.7 A cord is hanging vertically. (**a**) Find the tension in the cord as a function of the distance from the bottom. (**b**) What is the speed of a wave as a function of position on the cord? (**c**) Compute the acceleration, dv/dt, of a wave pulse moving up the cord. Ans: $g/2$

7.8 (**a**) Verify that the solution presented in Eq. (7.14) does satisfy both of the initial conditions on f as given by Eq. (7.12). (**b**) What happens if the two limits of integration were not taken to be 0 and 0, but two other numbers: either the same as each other or different?

7.9 What initial conditions on the wave function will guarantee that the wave is moving to the right only. I suggest that you don't start with the final formula to answer this. For the simple initial pulse in the example of Eq. (7.15), sketch and interpret the F and G that do this.

7.10 What would the resulting wave be if in the example at Eq. (7.15) you have $F \equiv 0$ and $G(x) = a^2v_0/(a^2 + x^2)$ This means that you struck the string instead of plucking it. Sketches please. This function is chosen so that it should make the integrals easy and the graph sketching and interpretation not too bad. Also, sketch the shape of the initial velocity curve for small and large values of a. As $t \to \infty$, you should find that $y(0, t) \to \pi a v_0/2v$.

7.11 (**a**) For the plucked string at Eq. (7.15), what is the total energy before the string is released? (**b**) Compute its total energy after it has been released and the two pulses have separated?

7.12 For the wave $f = A\cos(kx - \omega t)$, explicitly evaluate the integral Eq. (7.19) and then differentiate the result with respect to time, interpreting the result.

7.13 (**a**) Find the kinetic energy density and the potential energy density in a standing wave between two fixed points. (**b**) What is the total kinetic and total potential energy and is the sum constant? Does it have to be?

7.14 Find the ϵ^4 term for the solution of the cubic equation (7.46). Put the resulting number into the original equation and examine the accuracy of the total.

7.15 In Eq. (7.22) the factor v^2 was treated as a constant, as if T and μ are independent of x. Go back to the calculation and repeat it, but allowing T and μ to depend on x. You should get the same answer.

7.16 In the perturbation calculation for standing waves, it was not the frequency that was easy to calculate but the square of the frequency. Show that to first order in ϵ, the correction to the frequency produces

$$
\omega = \sqrt{\omega_0^2} \left[1 + \frac{\epsilon}{2} \frac{\omega_1^2}{\omega_0^2} + \cdots \right]
$$

7.17 (**a**) Follow the procedure to go from (7.30) to Eq. (7.34), but use a sine instead of a cosine. Compare the results. (**b**) Perhaps you remember from an introduction to optics and the discussion there of the $180°$ phase change on reflection at a surface. Show that using a sine instead of a cosine gives the *opposite* answer for whether there is a phase change going from low index to high or high to low.

7.18 Starting from the solution to the wave equation, Eq. (7.14), what relation is there between F and G so that the result is a wave moving only to the right?

7.19 Following the development of section 7.7, show what happens to the reflected and to the transmitted waves if $\mu_2 \gg \mu_1$. Same question if $\mu_2 \ll \mu_1$.

7.20 Starting from the modes of oscillation of a string of uniform tension and mass density held at points $x = 0$ and $x = L$, change the mass density by the small amount $\mu_1(x) = \mu_A \sin(N\pi x/L)$, where N is a given integer. Find the effect on the frequencies of the oscillation modes of the string assuming that $\mu_A \ll \mu_0$. How does the effect of this perturbation depend on which mode of oscillation you are dealing with? Sketch a plot of the size of the correction versus n, the mode number.

7.21 ˙In the wave equation (7.7) you can solve for the most general solution directly by making the right change of variables. Let $z = x - vt$ and $w = x + vt$. Use the multivariable chain rule four times to evaluate the derivatives $\partial^2 f/\partial x^2$ and $\partial^2 f/\partial t^2$ in terms of derivatives with respect to z and w. For example,

$$
\frac{\partial f}{\partial x}\bigg|_t = \frac{\partial f}{\partial z}\bigg|_w \frac{\partial z}{\partial x}\bigg|_t + \frac{\partial f}{\partial w}\bigg|_z \frac{\partial w}{\partial x}\bigg|_t
$$

where the sub-w or -t *etc.* indicate which variable is held constant. Show that the wave equation in the variables z-w is

$$
4\frac{\partial^2 f}{\partial z \partial w} = 0 \qquad \text{or with more explicit notation,} \qquad 4\frac{\partial}{\partial z}\bigg|_w \frac{\partial}{\partial w}\bigg|_z f = 0
$$

Now do two integrals to get the answer for f in terms of z and w. When you integrate a partial derivative, note that you don't get an arbitrary *constant,* but an arbitrary function of the other variable. When you are done, replace z and w by their values in terms of x and t, and you will have actually *solved* the wave equation by straight integration instead of just guessing the answer.

7.22 In the calculation leading up to Eq. (7.59), in which the string is slightly by $\mu_1 = \alpha x$, there is another way to arrange the calculation. Let the original μ_0 be replaced by $\mu_0 + \alpha L/2$ with $\mu_1 = \alpha(x - L/2)$. Recompute ω_0^2 and ω_1^2 using these definitions. Show that the final answer agrees with the old one, at least to first order in α. That is, you may have to do a power series expansion to see if they agree.

7.23 Sketch the wave Eq. (7.15) when the two parts have moved only partway from each other. $vt = L/4$ perhaps.

7.24 ˙A uniform string is stretched between two walls and allowed to oscillate. Now attach a set of small, evenly spaced masses to the string and to lowest order, find the effect on the frequency of oscillation. The string has length L. The $N - 1$ masses m are placed a distance a apart so that $Na = L$. Under which circumstances is the effect biggest and when is it smallest?

7.25 Three strings are attached to walls as shown and they vibrate only perpendicular to the plane of equilibrium (in and out in the drawing). Find the lowest frequency normal mode. The strings are initially horizontal and are tied at the middle with a massless knot. They are at $120°$ from each other. Ans: $\omega = \pi v/2L$
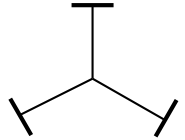
7.26 In the equations (7.41) and (7.42), suppose that they describe sound within the frequency range of human hearing. In this problem the two velocities w/k and w'/k' are the same. In this simpler case, determine what these equations and the following graphs say about what you will hear.

7.27 The phase velocity of deep water ocean waves is $v_{\text{phase}} = \sqrt{g\lambda/2\pi}$. Express this as a relation between ω and k, then compute the group velocity and compare it to the phase velocity. How large is this phase velocity for wavelengths of $1\,\text{cm}$ and $1\,\text{km}$? Same for the group velocities.

7.28 A string is tied down at one end, and the other end is attached to a very light ring that is free to move without friction on a rod. Take the mass of the ring to be zero and find the boundary conditions that must be satisfied at the two ends of the string. Find the modes and frequencies of oscillation of the string. Ans: $\omega_n = (2n + 1)\pi v/L$, with $v = (T/\mu)^{1/2}$ as usual. $n = 0, 1, \ldots$
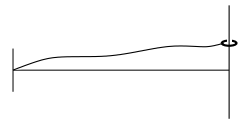
7.29 In Eq. (7.32) I tacitly assumed that the tension T itself is continuous. Prove that it is. (Look at F_x.)

7.30 ˙Take the system defined in problem 7.28, but instead of letting the ring slide up and down freely, grab hold of it and force it to move as $y = A\cos\omega t$. After the string has settled into its steady-state motion so that it too is oscillating with a frequency ω, what is the motion of the string? Explain the singularities you (should) find.

7.31 ¨Return to problem 7.25 and find the next lowest frequency mode of oscillation. Again, look at vibrations perpendicular to the plane of equilibrium.

7.32 ¨A string of length L is made of two parts: mass density μ_1 for $0 < x < L/2$ and density μ_2 for $L/2 < x < L$. The string is fixed at $x = 0$ and at $x = L$. Show that a mode of oscillation, $y = f(x, t) = F(x)G(t)$, will have the form

$$F(x) = A\sin k_1 x \quad (0 < x < L/2), \qquad F(x) = B\sin k_2(x - L) \quad (L/2 < x < L)$$

where k_1 and k_2 can, when needed, be expressed in terms of the tension, the frequency, and the respective mass densities. **(a)** Use the fact that F and dF/dx are continuous to get equations determining

the mode, and find the condition for a non-trivial solution to exist (a mode of oscillation). Don't try to solve the equation explicitly, but check it by showing that it gives the right result when the two mass densities happen to be equal. **(b)** Also, what happens to the modes if the mass density $\mu_2 \to \infty$? **(c)** If $\mu_2 \to 0$ be extra careful in your analysis. Make it *small*, not zero in order to see what happens. Your sketches should properly describe what is happening.

Show that one of your limiting cases looks like
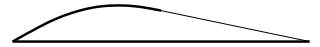
7.33 Return to the problem of the static, sagging string as in section 7.2. Assume as a guess that the shape of the sagging string is approximately a quadratic function of x. This quadratic satisfies the boundary conditions $f(0) = f(L) = 0$, but *don't* assume that it satisfies the wave equation. That leaves one free parameter in the function. Instead, compute its energy. This will be the sum of two terms, gravitational potential energy and the stretching of the string as see in Eq. (7.17) for the elastic potential energy. Use this quadratic to compute the total potential energy, $\int dm\, gy$ for gravity and $\int dx\, (dW/dx)$ for the elastic energy. Minimize this total energy as a function of the remaining free parameter in your $f(x)$. How does the shape at this minimum energy solution compare to the solution in Eq. (7.6)? This method is useful in finding approximate solutions to complicated problems, though you're rarely as lucky as in this example.

7.34 ¨**(a)** For those who want to demonstrate their mastery of the multivariable chain rule, what is the wave equation as seen by someone who is moving parallel to the x-axis with velocity u? That is, show that the change from variables (x, t) to variables (x', t') where $x' = x - ut$ and $t' = t$ changes the wave equation from

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = 0 \qquad \text{to} \qquad \frac{\partial^2 f}{\partial x'^2} - \frac{1}{v^2}\left[\frac{\partial^2 f}{\partial t'^2} - 2u\frac{\partial^2 f}{\partial x'\partial t'} + u^2\frac{\partial^2 f}{\partial x'^2}\right] = 0$$

(b) Show that the general solution to the original equation, Eq. (7.11), is now the sum of the two terms

$$y(x', t') = f(x' - ct') \qquad \text{where} \qquad c = -u \pm v$$

There is an application of this is section 7.11.

7.35 (b′) Someone else got a different answer for the preceding equation, getting the final solution to be

$$y(x', t') = f(x' - ct') \qquad \text{where} \qquad c = u \pm v$$

Analyze the proposed solutions to explain what they look like to the moving observer and decide which, if either, may be right. **(c′)** Describe the two proposed solutions in the special case that $u = v$.

7.36 Using the solution to problem 7.34, see what happens if the person chasing the wave is moving at the same speed as the wave. **(a)** Show that there is a static solution now, and find what form it can take. **(b)** Show that the other, non-static, solution satisfies a first order partial differential equation.

7.37 ˙Solve for the modes of a string governed by Eq. (7.61) and fixed at the ends, but now do it exactly. Expand your answer for small corrections and compare it to the approximate result in the text.

7.38 In the first of the equations (7.43), the velocity is *almost* the same as either wave alone. Show that the arithmetic mean of the two separate velocities differs from the original by terms of second order. *I.e.*

$$\frac{\omega + \omega'}{k + k'} = \frac{1}{2}\left(\frac{\omega}{k} + \frac{\omega'}{k'}\right) + \text{terms of order } (\Delta\omega)^2 \text{ and } (\Delta k)^2$$

7.39 In chapter ten, Eq. (10.41), you find the relation $\omega = 2\sqrt{T/\mu a^2}\,\sin(ka/2)$. Compute the phase and group velocities in this case. Graph and compare them.

7.40 The equation (7.60), which includes the stiffness of the string, has wave solutions: $e^{i(kx-\omega t)}$. Put this into the equation and find the relation between ω and k. What are the group and phase velocities? Sketch graphs of both as a function of either ω or k (your choice). As a start, assume that the stiffness is small, but then ask if there are ranges of frequency for which this added term has a noticable effect on the speeds even if it is small.

7.41 Use the same picture as in problem 7.28. Now however the ring at the right end has a mass M, though it is still free to slide up and down without friction. (**a**) Show that the boundary condition at the right end is

$$-T\frac{\partial f}{\partial x}(L,t) = M\frac{\partial^2 f}{\partial t^2}(L,t)$$

and find the modes of oscillation. You will not get explicit algebraic solutions for the values of k or ω, but draw graphs of the equation that has to be solved for k and indicate there what the solutions will be. (**b**) What does the graphical solution approach in the limit that $M \to 0$? (**c**) What does it approach in the limit that $M \to \infty$? In this second case, show that there is a solution that is not the same as when the string is tied down at the end, and interpret it.

7.42 Find an approximate solution to problem 7.32, assuming that the mass density of the string is μ_1 everywhere and that you add a small perturbing mass density $\delta\mu = \mu_2 - \mu_1 \ll \mu_1$ to the right half. Use perturbation theory to find the approximate change in the frequency(squared) for all the modes of oscillation.

7.43 Do the preceding problem 7.42 another way: Change the original, unperturbed mass density to the constant value $(\mu_1 + \mu_2)/2$. (**a**) Then use as a perturbing mass density the function

$$\delta\mu(x) = \begin{cases} -(\mu_2 - \mu_1)/2 & (0 < x < L/2) \\ +(\mu_2 - \mu_1)/2 & (L/2 < x < L) \end{cases}$$

and use perturbation theory to find the correction to the normal mode values of ω^2. (**b**) Show that to lowest order in $(\mu_2 - \mu_1)$ the results of the two ways to solve this problem agree. Also note that one method is easier.

7.44 For the example of a stiff string, Eq. (7.60), the first order correction to the frequency (squared) appears in Eq. (7.63). Now go back to the equation for the correction to the wave function itself and solve that, Eq. (7.62), for F_1. (*Surprise?*) Perhaps you would like to explain what happened.

7.45 Take a long, uniform string and attach a point mass to it in the middle. Stretch the string out with tension T and send a wave toward this mass, coming from the left side. What are the amplitudes of the transmitted and the reflected waves? The boundary condition at the mass $(x = 0)$ is like that of problem 7.41 except that the force comes from both sides so there are two terms on the left (from $\frac{\partial f}{\partial x}(0-)$ and $\frac{\partial f}{\partial x}(0+)$. Naturally you will examine how the results depend on M and on any of the parameters for the incoming wave. Is energy conserved? Are the phases of the reflected and transmitted waves fixed or do they depend on any aspects of the incoming wave? This is a problem in which *you really do* want to use the complex forms for the waves as in section 7.5. Ans: Let $\gamma = Mk/2\mu = M\pi/\lambda\mu$. Reflection amplitude $= i\gamma/(1-i\gamma)$. Transmission amplitude $= 1/(1-i\gamma)$. Fraction of power reflected $= \gamma^2/(1+\gamma^2)$.

7.46 Use the string plus mass system of the preceding problem, but rotate everything so that the string is vertical. Notice that the tension in the string is not the same on the two sides of the central mass. Now compute the reflection and transmission amplitudes when a wave comes in from one or the other end of the string. Also check conservation of energy. Assume that the string is light enough that the tension within each side of the string is constant.

Rigid Body Motion

There's no such thing as a rigid body of course, because everything is at least slightly flexible. No matter. To a good approximation a lot of objects are reasonably rigid, meaning that in the course of their motions the distances between all the points in the body are nearly constant. $d|\vec{r}_i - \vec{r}_j|/dt = 0$, where the $\vec{r}$'s are coordinates of different atoms.

Unless you live on the Pacific Rim, it is a useful approximation to say that the Earth is rigid. The tire of a car flexes under the car's weight, but you can still treat it as rigid for many purposes. A satellite in orbit around the Earth and having only electronic components is a rigid body — or is it? One early Earth-orbital satellite was lost by making this approximation; it wasn't quite rigid enough. See the end of section 8.6, Figure 8.14.

When describing rotations, there are a couple of ideas that may seem fairly intuitive, but I want to state them explicitly.

1. Any motion (except pure translation) of a rigid body is a pure rotation about some axis.
2. Any motion of a rigid body is a translation of the center of mass plus rotation about the center of mass.

These are basically geometric theorems and their proofs aren't very enlightening, so I won't go through them.

8.1 Center of Mass

Newton's equation $\vec{F} = m\vec{a}$ applies to a point mass. Except maybe for an electron, there are no point masses, so how can you get away with treating things as if they are? Is the Earth a point mass? It is if the subject is stellar astronomy; it isn't if you're a meteorologist.

Any real object is made of many others — atoms at least, and the total force on it is the sum of all the forces on the individual pieces. In this calculation, I will assume that each mass remains constant, so I can move masses inside and outside of derivatives at will. The idea of the center of mass came up in section 6.7, but in that context I was trying to introduce a change of variables to simplify an orbit problem. The purpose now is different, so I'll repeat the calculation in a form that is more useful here.

The point masses are $\{m_k, \ k = 1, 2, \ldots\}$, and the force on one such mass is $\vec{F}_k$.

$$\vec{F}_k = m_k \frac{d\vec{v}_k}{dt} \implies \sum_k \vec{F}_k = \sum_k m_k \frac{d\vec{v}_k}{dt}$$

Each $\vec{F}_k$ has two parts, one from interactions with the other parts of the body and one from outside. Handle these separately

$$\vec{F}_k = \vec{F}_{k,\text{ external}} + \sum_{k' \neq k} \vec{F}_{k \text{ by } k'} \tag{8.1}$$

If the Earth the Moon are a system of two particles, the force on the Moon is the sum of the forces from Earth and the Sun. Similarly on the Earth from the Moon and the Sun, treating the Sun as an external force on the Earth-Moon system. Add all the forces in Eq. (8.1).

$$\sum_k \vec{F}_k = \sum_k \vec{F}_{k,\text{ external}} + \sum_k \sum_{k' \neq k} \vec{F}_{k \text{ by } k'}$$

If these are the Earth and the Moon, with the Sun providing an external force, this is

$$\vec{F}_{\text{on E}} + \vec{F}_{\text{on M}} = \vec{F}_{\text{on E by S}} + \vec{F}_{\text{on M by S}} + \underbrace{\vec{F}_{\text{on E by M}} + \vec{F}_{\text{on M by E}}}_{= 0}$$

$$= \vec{F}_{\text{on E by S}} + \vec{F}_{\text{on M by S}}$$

Newton's third law says that the force on k by k' is minus the force on k' by k, so in this sum over all the possible pairs of masses k and k', every term is canceled by another. $\vec{F}_{k \text{ by } k'} = -\vec{F}_{k' \text{ by } k}$ This implies

$$\sum_k \vec{F}_k = \sum_k \vec{F}_{k, \text{ external}} = \sum_k m_k \frac{d\vec{v}_k}{dt} = \sum_k m_k \frac{d^2 \vec{r}_k}{dt^2} = \frac{d^2}{dt^2} \sum_k m_k \vec{r}_k$$

Now manipulate the last equation. Multiply and divide by the same factor, the total mass.

$$M_{\text{total}} = \sum_k m_k, \qquad \text{then} \qquad \frac{d^2}{dt^2} \sum_k m_k \vec{r}_k = M_{\text{total}} \frac{d^2}{dt^2} \left(\frac{\sum_k m_k \vec{r}_k}{M_{\text{total}}} \right) \qquad (8.2)$$

The final parenthesized expression defines the center of mass, and I'll switch to the integral notation for the sum as it is more convenient to manipulate.

$$\vec{r}_{\text{cm}} = \frac{1}{M_{\text{tot}}} \int dm \, \vec{r} \qquad \text{and} \qquad \vec{F}_{\text{tot, ext}} = M_{\text{tot}} \frac{d^2 \vec{r}_{\text{cm}}}{dt^2} \qquad (8.3)$$

This is $\vec{F} = m\vec{a}$ for real masses, not fictional point masses. What it says is that all the calculations done while pretending that you're dealing with a point mass are correct. You just have to interpret them as describing the motion of the center of mass of your object. There is nothing in this calculation that requires the masses to form a rigid body (*e.g.* the solar system, acted on by the rest of the galaxy), but rigid bodies will be the case for most of this chapter.

Example

▶ Where is the center of mass of the Earth-Moon system? Treat each as a point mass and use a single coordinate to describe their position.

$$x_{\text{cm}} = \frac{m_E x_1 + m_M x_2}{m_E + m_M} = \frac{x_2}{1 + (m_E / m_M)}$$

$$= \frac{380\,000\,\text{km}}{1 + 80} = 4700\,\text{km}$$

Fig. 8.1

The Earth's radius is 6400 km, so this is 1700 km below the surface. This is certainly not a rigid body.

Example

▶ Where is the center of mass of a uniform wire bent into a semicircle?

$$M\vec{r}_{\text{cm}} = \int dm \, (\hat{x}\,x + \hat{y}\,y) = \int dm \, (\hat{x}R\cos\phi + \hat{y}R\sin\phi)$$

$$= \int_0^\pi M \frac{d\phi}{\pi} (\hat{x}R\cos\phi + \hat{y}R\sin\phi)$$

$$\vec{r}_{\text{cm}} = \hat{y} \frac{2R}{\pi} = \hat{y}\,0.64R$$
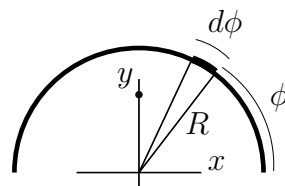
Example

► Where is the center of mass of a uniform, solid hemisphere? One way to do this is to use spherical coordinates, as in Figure 0.8.

$$dm = \rho\, dV = \rho\, r^2 \sin\theta\, dr\, d\theta\, d\phi \quad \text{with} \quad \rho = M/\tfrac{2}{3}\pi R^3 \quad \text{and} \quad z = r\cos\theta$$

The center of mass will be along the central axis somewhere, so $\int z\, dm$ is all that's needed.

$$\int z\, dm = \int_0^R dr \int_0^{\pi/2} d\theta \int_0^{2\pi} d\phi\, r^2 \sin\theta\, \rho\, r\cos\theta$$

$$= \rho \cdot \frac{R^4}{4} \cdot \frac{1}{2} \cdot 2\pi = \frac{M}{\tfrac{2}{3}\pi R^3} \frac{\pi R^4}{4} = \frac{3}{8} MR$$
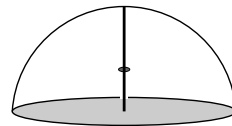
so $z_{\text{cm}} = \tfrac{3}{8}R$. This is less than halfway from the center to the top, and that is plausible, in that there is more mass toward the bottom.

In the first of these examples, I treated the Earth and Moon as point masses. The last I saw, they aren't. Is it correct to do this? Can you compute the center of mass of an object by treating portions of it as if they are themselves point masses? Yes, but it requires a little proof.

$$M_{\text{tot}}\vec{r}_{\text{cm}} = \int dm\,\vec{r} = \int_{\text{first mass}} dm\,\vec{r} + \int_{\text{second mass}} dm\,\vec{r} = m_1\vec{r}_{\text{cm}\,1} + m_2\vec{r}_{\text{cm}\,2} \qquad (8.4)$$

Now divide by the total mass and that's all there is to it.

What does that mean?

In Eq. (8.3) you see the notation $\int \vec{r}\, dm$. How do you integrate position with respect to mass? The answer involves looking back to where that expression came from and then combining that with the concept of an integral. Eq. (8.2) doesn't involve an integral but a sum, $\sum_k m_k \vec{r}_k$. The concept of an integral is that you divide a complicated thing into pieces in order to get an approximate result for a sum, then improve the approximation by refining the divisions.

Here you can divide the region into small volumes and find how much mass is in each volume. Call the volumes ΔV_ℓ and the masses ΔM_ℓ. The coordinate vector to the center of each volume is $\vec{r}_\ell$, so an approximate value for the sum is

$$\sum_\ell \vec{r}_\ell \Delta M_\ell$$

That doesn't look any different from the original sum but it *is* different, because each of these smaller number of terms represents a group of point masses. The idea of an integral is that you improve the approximation by taking smaller volumes and in the ideal case take a limit.

Who says that a particular mass ΔM_6 will go to zero? It may not. For the regions where the limit $\Delta M_\ell/\Delta V_\ell$ approaches a limit, call it the mass density ρ and let $dM = \rho dV$. If it doesn't approach a limit because there is a big point mass there, you are simply left with a discrete, separate term in the sum, $M\vec{r}$. The final limit is then an ordinary volume integral plus some separate terms that come from point masses. This notation, $\int \vec{r}\, dm$, is simply a shorthand for what you would probably do anyway if you're confronted with a mixture of continuous and discrete distributions of mass.*

---

* For a full-blown mathematical treatment of this way to look at an integral, look up Riemann-Stieltjes integrals. It's a useful concept.

8.2 Angular Momentum

The first step in describing the dynamics of a rigid body is to start with a point mass and apply a force to it.

$$\vec{F} = d\vec{p}/dt$$

Pick an origin and take the cross product of this equation with $\vec{r}$, the coordinate vector for the point mass.

$$\vec{r} \times \vec{F} = \vec{r} \times \frac{d\vec{p}}{dt}$$

$$= \vec{r} \times \frac{d\vec{p}}{dt} + \frac{d\vec{r}}{dt} \times \vec{p}$$

$$= \frac{d}{dt}(\vec{r} \times \vec{p})$$

The second line adds zero because $\vec{v} \times m\vec{v} = 0$. The next step is the product rule for derivatives, combining the terms. This is

$$\vec{\tau} = \vec{r} \times \vec{F} = \frac{d}{dt}(\vec{r} \times \vec{p}) = \frac{d\vec{L}}{dt} \tag{8.5}$$

Torque is the time derivative of angular momentum (for a point mass).

For a rigid body, the only change in Eq. (8.5) is to add an index to the factors. The index says that now there are many masses and you simply sum over those masses to get the total torque and total angular momentum.

$$\vec{\tau} = \sum_k \vec{r}_k \times \vec{F}_k = \frac{d}{dt}\left(\sum_k \vec{r}_k \times \vec{p}_k\right) = \frac{d\vec{L}}{dt} \tag{8.6}$$

The next, and most important, step is to organize the bookkeeping for this sum describing the angular momentum.

$$\vec{L} = \sum_k \vec{r}_k \times \vec{p}_k$$

To describe rotation about an axis, use the equation (5.8), $\vec{v} = \vec{\omega} \times \vec{r}$, giving

$$\vec{L} = \sum_k \vec{r}_k \times m_k \vec{v}_k = \sum_k \vec{r}_k \times m_k (\vec{\omega} \times \vec{r}_k)$$

This shows how the angular momentum depends on the angular velocity and it is the defining equation for the function called the tensor of inertia. What's a "tensor"? Wait a few pages for a full explanation.

$$\vec{L} = I(\vec{\omega}) = \sum_k \vec{r}_k \times m_k (\vec{\omega} \times \vec{r}_k)$$

The function I is a vector-valued function of a vector variable, and that is the central defining property to keep your eye on. This is where you start to learn about tensors, and it is at the core of the rest of this chapter.

When you deal with a rigid body, most often it is a continuous distribution of mass and not some discrete masses. A more appropriate notation for both cases is then
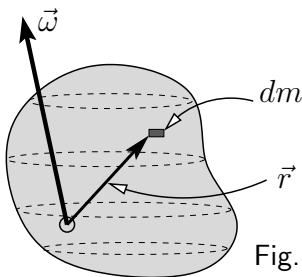
Fig. 8.2

$$\vec{L} = I(\vec{\omega}) = \int dm\,\vec{r} \times (\vec{\omega} \times \vec{r})$$

$$= \int dm\,[r^2\vec{\omega} - \vec{r}(\vec{\omega} \cdot \vec{r})] \tag{8.7}$$

$$[\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})] \quad (0.24)$$

Either way that you write it, you are computing the properties of the rigid body with respect to a point, the origin — *not* an axis as you may have been expecting. This manipulation used the vector identity (0.24) for the triple cross product,

Example

▶ A straight, uniform rigid rod rotating about an axis through one end. Mass M, length a. Place it along the x-axis with one end at the origin, and make the axis of rotation the z-axis. The element of mass is $dm = M\,dx/a$ between x and $x + dx$, then

$$\vec{\omega} \times \vec{r} = \omega \hat{z} \times x\,\hat{x} = \omega x\,\hat{y}$$

$$\vec{r} \times (\text{this}) = x\,\hat{x} \times \omega x\,\hat{y} = \omega x^2 \hat{z}$$

$$\int dm\,(\text{that}) = \int_0^a dx\,\frac{M}{a}\omega x^2 \hat{z} = \frac{M}{a}\omega \hat{z} \frac{x^3}{3}\Big|_0^a = \frac{1}{3}Ma^2\,\omega \hat{z}$$

Fig. 8.3
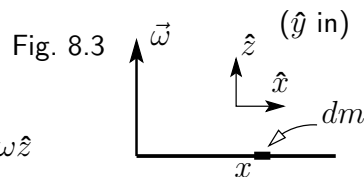
In this example the angular momentum is parallel to the angular velocity, and the proportionality factor is called the moment of inertia: $\vec{L} = I_0\vec{\omega}$, and $I_0 = Ma^2/3$ is the moment of inertia of the rod about this axis through its end. As you will quickly find, life is not always so simple. Here I_0 is a number multiplying $\vec{\omega}$. On the preceding page and in most of this chapter, I is a function relating $\vec{\omega}$ and $\vec{L}$. You will see where they fit together shortly.

Example

▶ A dumbbell is spinning about an axis as shown. Treat the two masses as points and compute the sum for $\vec{L}$.
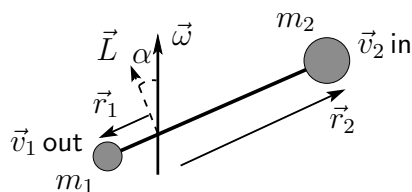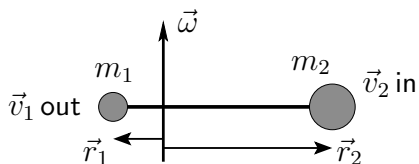
Fig. 8.4

In the first picture, $\vec{r}_1 \times \vec{v}_1$ is up along the same direction as $\vec{\omega}$. The same is true for the second mass, and you can write $\vec{L}$ as a constant times $\vec{\omega}$. In the *second* picture however, $\vec{r}_1 \times \vec{v}_1$ points to the upper left, perpendicular to the line between the masses. The same for $\vec{r}_2 \times \vec{v}_2$. Neither is parallel to $\vec{\omega}$, so in this second simple-looking example a relation such as $\vec{L} = I_0\vec{\omega}$ is just can't be true. It's still easy to figure out $\vec{L}$ though. The angle between $\vec{\omega}$ and $\vec{r}_1$ is $90° + \alpha$, and between $\vec{\omega}$ and $\vec{v}_2$ it is $90° - \alpha$.

$$\vec{v}_1 = \omega \times \vec{r}_1 \quad \rightarrow \quad v_1 = \omega r_1 \cos \alpha \qquad \text{then} \qquad \vec{r}_1 \times \vec{v}_1 = \omega r_1^2 \cos \alpha \ \text{ to the upper left}$$

Similarly $|\vec{r}_2 \times \vec{v}_2| = \omega r_2^2 \cos \alpha$ in the same direction, and the total angular momentum is the sum of two terms,

$$\vec{L} = \left(m_1 r_1^2 + m_2 r_2^2\right)\omega \cos \alpha \ \text{ to the upper left} \qquad \textit{[check } \alpha = 0 \textit{ and } \alpha = 90°\textit{]}$$

Example

▶ A square plate is rotating about an axis that is at an angle α from the axis of symmetry of the plate (the z-axis). In the coordinate system drawn, $\vec{\omega}$ is in the y-z plane, and the plate (in the $z = 0$ plane) has x and y coordinates that go from $-a/2$ to $+a/2$. Use the area mass density, $\sigma = m/a^2$. In the equation (8.7),

$$\vec{L} = \int dm\left[r^2\vec{\omega} - \vec{r}(\vec{\omega} \cdot \vec{r})\right]$$

$$\vec{r} = x\hat{x} + y\hat{y}$$

$$\vec{\omega} = \hat{z}\,\omega \cos \alpha + \hat{y}\,\omega \sin \alpha$$

$$dm = \sigma\,dx\,dy, \qquad \text{where} \qquad \sigma = m/a^2$$
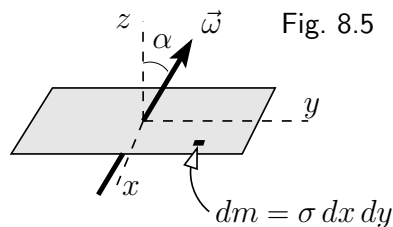
Fig. 8.5

$$\vec{L} = \int_{-a/2}^{a/2} dx \int_{-a/2}^{a/2} dy\, \sigma \left[(x^2 + y^2)(\hat{z}\omega\cos\alpha + \hat{y}\omega\sin\alpha) - (x\hat{x} + y\hat{y})\omega\, y\sin\alpha \right]$$

There's a little simplification before plunging ahead: the $y^2\omega\sin\alpha$ terms cancel; the term with an xy factor integrates to zero because the limits on x (or y) go from $-a/2$ to $+a/2$. What is left is

$$\begin{aligned}
\vec{L} &= \sigma \int_{-a/2}^{a/2} dx \int_{-a/2}^{a/2} dy \left[(x^2 + y^2)\hat{z}\omega\cos\alpha + x^2\hat{y}\omega\sin\alpha \right] \\
&= \sigma \int_{-a/2}^{a/2} dx \left[(x^2 a + \tfrac{1}{12}a^3)\hat{z}\omega\cos\alpha + x^2 a\hat{y}\omega\sin\alpha \right] \\
&= \sigma \left[(\tfrac{1}{12}a^4 + \tfrac{1}{12}a^4)\hat{z}\omega\cos\alpha + \tfrac{1}{12}a^4\hat{y}\omega\sin\alpha \right] \\
&= \tfrac{1}{12}ma^2 \left[2\hat{z}\cos\alpha + \hat{y}\sin\alpha \right]\omega
\end{aligned} \tag{8.8}$$

Compare this to the direction of $\vec{\omega} = \hat{z}\,\omega\cos\alpha + \hat{y}\,\omega\sin\alpha$ and you see that $\vec{L}$ and $\vec{\omega}$ don't align. The angular momentum points closer to the z-axis than does $\vec{\omega}$. Draw $\vec{L}$.
What happens if α equals 0 or $90°$?

Definition

The key defining property of a tensor:

A linear, vector-valued function of a vector variable is a tensor.

That a function f is linear means

$$\begin{aligned}
f(\vec{v}_1 + \vec{v}_2) &= f(\vec{v}_1) + f(\vec{v}_2) \qquad \text{and} \qquad f(c\vec{v}) = c\, f(\vec{v}) \\
&\text{or in a single line:} \quad f(a\vec{v}_1 + b\vec{v}_2) = a\, f(\vec{v}_1) + b\, f(\vec{v}_2)
\end{aligned} \tag{8.9}$$

and a, b, c are scalars. Another name for this type of function is "linear transformation" or "linear operator". This doesn't completely define the subject of tensors, but it covers the case at hand. For a general definition to cover the other cases, see chapter 12 of Mathematical Tools.

A common example of a tensor is a rotation. Rotate all vectors by $30°$ about the z-axis, and this defines a function. You have a vector in and a vector out. If you add two vectors and rotate the result or if you rotate the two vectors and then add them, you get the same answer. $\vec{A} + \vec{B} = \vec{C}$ both before and after rotation. If you double the length of a vector before or after rotating it, the answer is the same. Rotation defines a linear function.†
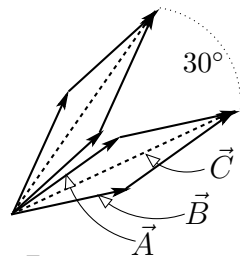
The function I in the equation $\vec{L} = I(\vec{\omega})$ satisfies the definition (8.9). It has a vector input and a vector output. The cross product and the integral in Eq. (8.7) are linear, so the combination is linear too.

Fig. 8.6

$$\begin{aligned}
I(\vec{\omega}_1 + \vec{\omega}_2) &= \int dm\, \vec{r} \times \left((\vec{\omega}_1 + \vec{\omega}_2) \times \vec{r} \right) = \int dm\, \vec{r} \times (\vec{\omega}_1 \times \vec{r} + \vec{\omega}_2 \times \vec{r}) \\
&= \int dm\, \vec{r} \times (\vec{\omega}_1 \times \vec{r}) + \int dm\, \vec{r} \times (\vec{\omega}_2 \times \vec{r}) \\
&= I(\vec{\omega}_1) + I(\vec{\omega}_2)
\end{aligned} \tag{8.10}$$

And $I(c\vec{\omega}) = c\, I(\vec{\omega})$ even more easily.

---

† This is a more restrictive definition of "linear" than used with differential equations. It would be "linear, homogeneous" in that context, but here that's all there is.

Notice: I did not refer to "the function $I(\vec{\omega})$". This object, $I(\vec{\omega}) = \vec{L}$, is not a function but a vector. It is the output of the function I when you feed the vector $\vec{\omega}$ to it, and *that* output is $\vec{L}$, the angular momentum vector. This distinction between a function and the value of the function for some particular argument sounds trivial and pedantic, and don't you commonly refer to $f(x)$ as a function? It is really not; it is the value of the function at the point x. Does this matter? When you're talking about the function f and the number $f(x)$ for common real-valued functions of real variables, it probably makes no difference and it seldom causes confusion. Here however if you don't pay attention to this distinction, it will cause needless difficulties, so watch the use of language in this chapter. At the end of section 0.10 I briefly raised this issue in a footnote but it was not important then. In the picture just above, "f" is the whole graph and "$f(x)$" is the value of f at the point x.
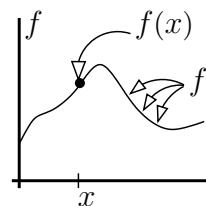
Example

▶ Suspend a mass with springs and let it come to equilibrium. Now apply a small force to the mass and look at its displacement. The displacement is a function of the applied force, and it is a linear function of that applied force. This defines a tensor f, where $\vec{d} = f(\vec{F})$. If the springs are all the same, then $\vec{d}$ will be in the same direction as $\vec{F}$ and it is as simple as $\vec{d} = \vec{F}/k$, but if the springs on the left and right are weaker than the others then this displacement won't even be in the same direction as $\vec{F}$. Push in a direction between the springs and the displacement will be more toward the weaker springs then the stronger.
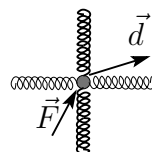
Fig. 8.7

Parallel Axis Theorem

There are a couple of useful relations for computing these functions, the parallel axis theorem and the perpendicular axis theorem. Calling them "theorems" is stretching the word a bit, but they are useful. The first relates a tensor for a body to its tensor when the center of mass is the origin. It says that the tensor of inertia about a point P equals the tensor of inertia about the center of mass plus another term that is the tensor of inertia of a point mass place *at* the center of mass and using the original P as the origin. The proof is easy: add and subtract a term, $\vec{r}_{cm}$.

$$\int dm\, \vec{r} \times (\vec{\omega} \times \vec{r}) = \int dm \left((\vec{r} - \vec{r}_{cm}) + \vec{r}_{cm}\right) \times \left(\vec{\omega} \times \left((\vec{r} - \vec{r}_{cm}) + \vec{r}_{cm}\right)\right)$$

$$= \int dm\, (\vec{r} - \vec{r}_{cm}) \times (\vec{\omega} \times (\vec{r} - \vec{r}_{cm})) + \int dm\, \vec{r}_{cm} \times (\vec{\omega} \times \vec{r}_{cm}) \quad (8.11)$$

What happened to the cross terms? They're zero.

$$\int dm\, (\vec{r} - \vec{r}_{cm}) \times (\vec{\omega} \times \vec{r}_{cm}) = \left[\int dm\, (\vec{r} - \vec{r}_{cm})\right] \times (\vec{\omega} \times \vec{r}_{cm})$$

The last parentheses formed a single constant inside the integral, so you can pull it outside the integration (on the right). The definition of the center of mass is
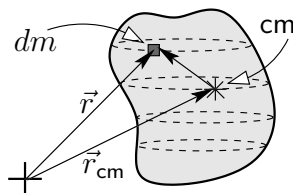
Fig. 8.8

$$\vec{r}_{cm} = \frac{1}{M_{total}} \int \vec{r}\, dm \quad (8.12)$$

so that makes the integral in the cross term vanish. The two terms in Eq. (8.11) are now precisely the statement of the parallel axis theorem because in the first integral, $(\vec{r} - \vec{r}_{cm})$ is the coordinate of the mass dm with respect to the center of mass, and in the second integral the entire integrand is a constant.

$$I(\vec{\omega}) = I_{cm}(\vec{\omega}) + M_{total}\, \vec{r}_{cm} \times (\vec{\omega} \times \vec{r}_{cm}) \quad (8.13)$$

The perpendicular axis theorem is even easier, but it refers to the components of the tensor, not to the tensor itself. It can wait until Eq. (8.30).

Example

▶ What is the angular momentum of a Hula Hoop when you are spinning it about your waist? The hoop is a circle, but it is not spinning it about its center but about its edge. If you are working really hard at it, then to a good approximation a point on the edge is nearly standing still. What is the hoop's angular momentum? Use the parallel axis theorem and start by finding the angular momentum assuming that the hoop is spinning about an axis through the center of the circle. $\vec{\omega}$ is along this axis.

In the integral about the center (in the second picture), $\vec{L} = \int dm\,\vec{r} \times (\vec{\omega} \times \vec{r})$, and all the radii have the same magnitude; all the $\vec{\omega} \times \vec{r}$ are tangent to the circle with magnitude ωr. In turn, $\vec{r}\times$ this vector is along the axis with magnitude ωr^2. That makes $\vec{L} = Mr^2\vec{\omega}$ along the axis (out of the page and parallel to the angular velocity).

Now add the second term of Eq. (8.13), where $\vec{r}_{cm}$ points from the right edge to the center of the circle. Do the two cross products and this term adds an equal amount $Mr^2\vec{\omega}$ to the total. $\vec{L} = 2Mr^2\vec{\omega}$.

$\vec{\omega} \times \vec{r}$
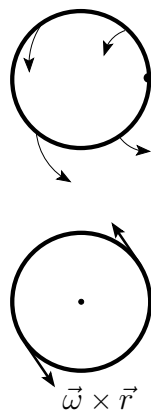
8.3 Tensor Components

Computing with tensors is like computing with vectors — the geometry gets out of hand quickly. For vectors you can use the tool of components to do calculations, and the same thing is true here. Tensors have components, but instead of the usual three components for vectors, there are nine* for tensors.

Go directly for the components of the inertia tensor. Other tensors are conceptually no different. Start with the vectors $\vec{\omega}$ and $\vec{L}$. These have components with respect to whatever basis you've chosen:

$$\vec{\omega} = \omega_x \hat{x} + \omega_y \hat{y} + \omega_z \hat{z} \qquad \text{and} \qquad \vec{L} = L_x \hat{x} + L_y \hat{y} + L_z \hat{z}$$

Now relate these vectors with the function I.

$$\vec{L} = I(\vec{\omega}) \qquad \text{is} \qquad L_x \hat{x} + L_y \hat{y} + L_z \hat{z} = I(\omega_x \hat{x} + \omega_y \hat{y} + \omega_z \hat{z}) \tag{8.14}$$

Use the linearity property of Eq. (8.9), which is the defining property of a tensor.

$$I(\omega_x \hat{x} + \omega_y \hat{y} + \omega_z \hat{z}) = \omega_x I(\hat{x}) + \omega_y I(\hat{y}) + \omega_z I(\hat{z}) \tag{8.15}$$

The expression $I(\hat{x})$ is a vector. As such it has three components. Denote these as

$$I(\hat{x}) = I_{xx}\hat{x} + I_{yx}\hat{y} + I_{zx}\hat{z} \tag{8.16}$$

In the same way, express the other terms in components:

$$I(\hat{y}) = I_{xy}\hat{x} + I_{yy}\hat{y} + I_{zy}\hat{z} \qquad \text{and} \qquad I(\hat{z}) = I_{xz}\hat{x} + I_{yz}\hat{y} + I_{zz}\hat{z} \tag{8.17}$$

The indices denote which output basis vector and which input basis vector are referred to, and the order of the indices is carefully chosen for later convenience. For example, I_{xy} is the x-component of the vector $I(\hat{y})$. Insert equations (8.15), (8.16), and (8.17) into equation (8.14) to get

$$\begin{aligned} L_x \hat{x} + L_y \hat{y} + L_z \hat{z} = \ & \omega_x \big[I_{xx}\hat{x} + I_{yx}\hat{y} + I_{zx}\hat{z} \big] \\ & +\omega_y \big[I_{xy}\hat{x} + I_{yy}\hat{y} + I_{zy}\hat{z} \big] \\ & +\omega_z \big[I_{xz}\hat{x} + I_{yz}\hat{y} + I_{zz}\hat{z} \big] \end{aligned}$$

---

* There are generalizations of this statement for other kinds of tensors, but for today this is good enough.

For these vectors to be equal, their respective components must match. Equate the coefficient of $\hat{x}$ on the left to the coefficient of $\hat{x}$ on the right. Now repeat the process for $\hat{y}$ and $\hat{z}$. This expresses the components of $\vec{L}$ in terms of the components of $\vec{\omega}$.

$$
\begin{aligned}
L_x &= I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z \\
L_y &= I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z \\
L_z &= I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z
\end{aligned}
\tag{8.18}
$$

In matrix notation this is

$$
\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}
\tag{8.19}
$$

and this is the reason for the peculiar-looking way that you are told to multiply matrices, across the row of the first factor and down the columns of the second. It all comes from the equations (8.14)-(8.17), defining the components of a tensor. The meaning of the expression (8.19) *is* the set of equations (8.18). The matrix indices are arranged as $I_{(\text{which row}), (\text{which column})}$. Compare the placement of the indices in Eq. (8.18) and in Eq. (8.17). The notation is designed to come out in the way conventionally used with matrices.

An abbreviated form for these equations is easier to write by using a unified notation for the basis vectors. $\vec{e}_i$, with the index i running over the set x, y, z (or more often 1, 2, 3). $\vec{e}_i$ replaces the three vectors $\hat{x}$, $\hat{y}$, $\hat{z}$, and the equations (8.16) and (8.17) become one equation.

$$
I(\vec{e}_i) = \sum_j I_{ji}\vec{e}_j
\tag{8.20}
$$

Recall the discussion at Eq. (0.27).

The manipulations that take you from equation (8.14) through (8.19) become far more compact using the notation of Eq. (8.20).

$$
\vec{L} = \sum_i L_i\vec{e}_i = I(\vec{\omega}) = I\Big(\sum_j \omega_j\vec{e}_j\Big) = \sum_j \omega_j I(\vec{e}_j) = \sum_j \omega_j \sum_i I_{ij}\vec{e}_i
\tag{8.21}
$$

The basis vectors $\vec{e}_i$ are independent, so the coefficients of $\vec{e}_i$ must agree on the two sides of the equation. That is,

$$
L_i = \sum_j I_{ij}\omega_j
\tag{8.22}
$$

These two lines are so much more compact that you should compare their content line-by-line with the preceding equations to verify that they are what they claim to be.

To compute these components, go back to the definition of the inertia tensor, Eq. (8.7). You need a vector identity for the triple cross product to manipulate this, Eq. (0.24).

$$
\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B}) \implies
$$
$$
I(\vec{\omega}) = \int dm\, \vec{r} \times (\vec{\omega} \times \vec{r}) = \int dm\, [r^2\vec{\omega} - \vec{r}(\vec{\omega} \cdot \vec{r})]
\tag{8.23}
$$

Now compute $I(\hat{x})$.

$$
\begin{aligned}
I(\hat{x}) = \int dm\, [r^2\hat{x} - \vec{r}(\hat{x} \cdot \vec{r})] &= \int dm\, [(x^2 + y^2 + z^2)\hat{x} - (x\hat{x} + y\hat{y} + z\hat{z})x] \\
&= \int dm\, [(y^2 + z^2)\hat{x} - yx\hat{y} - zx\hat{z}] \\
&= I_{xx}\hat{x} + I_{yx}\hat{y} + I_{zx}\hat{z}
\end{aligned}
$$

The last line is the defining equation (8.16). Now all that you need do is to read off the coefficients of $\hat{x}$, $\hat{y}$, and $\hat{z}$, getting the first column of the matrix of components, and the calculations for the other six components are identical, giving the second and third *columns* of the matrix.

$$I_{xx} = \int dm\,(y^2 + z^2) \qquad I_{xy} = -\int dm\,xy \qquad I_{xz} = -\int dm\,xz$$

$$I_{yx} = -\int dm\,yx \qquad I_{yy} = \int dm\,(x^2 + z^2) \qquad I_{yz} = -\int dm\,yz$$

$$I_{zx} = -\int dm\,zx \qquad I_{zy} = -\int dm\,zy \qquad I_{zz} = \int dm\,(x^2 + y^2)$$

The three diagonal components, I_{xx}, I_{yy}, and I_{zz}, are the moments of inertia about the three axes, and the other components are called the products of inertia. There's some symmetry among the nine components, as $I_{xy} = I_{yx}$ etc, making the matrix in Eq. (8.19) symmetric. This is a special property of the inertia tensor, and it is not true for all tensors. You can write this matrix as

$$(I) = \int dm \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{pmatrix} \tag{8.24}$$

In the spirit of Eq. (8.21) can these calculations of the components of the inertia tensor be more compact? Yes. Use x_1, x_2, and x_3 for the coordinates instead of x, y, and z, then apply the summation convention as in Eq. (0.28), then Eq. (8.23) becomes

$$L_i \vec{e}_i = I(\vec{\omega}) = \int dm \big[x_j x_j \omega_i \vec{e}_i - x_i \vec{e}_i \omega_j x_j \big]$$

$$L_i = \int dm \big[x_j x_j \omega_i - x_i \omega_j x_j \big] = \int dm \big[x_j x_j \delta_{ik} - x_i x_k \big] \omega_k$$

$$I_{ik} = \int dm \big[x_j x_j \delta_{ik} - x_i x_k \big] \qquad L_i = I_{ik} \omega_k \tag{8.25}$$

Recall that now a repeated index within a single term is automatically summed. Also, such a summed index is necessarily a dummy variable, so that you can change its name at will, a fact that was used to advantage in the second line of this little calculation (look for it). This same summation convention also simplifies the appearance of Eq. (8.21) by removing all the explicit summation symbols. They are unnecessary now.

Example

▶ A thin uniform rectangular plate is placed with one corner at the origin and with the sides along the x and y axes: $(0 < x < a)$, $(0 < y < b)$. What are the components of the tensor of inertia?

Look at Eq. (8.24): A lot of terms are zero — everywhere there's a z. This means that the only integrals to do are

$$\int dm \quad \text{of} \quad x^2, \quad y^2, \quad xy$$

The area mass density is $\sigma = m/ab$, and

$$\int_0^a dx \int_0^b dy\, x^2 = \frac{ba^3}{3}, \qquad \int_0^a dx \int_0^b dy\, y^2 = \frac{ab^3}{3}, \qquad \int_0^a dx \int_0^b dy\, xy = \frac{a^2 b^2}{4}$$

Multiply each of the these expressions by the area mass density and put them into their appropriate slots in the matrix for the tensor components.

$$(I) = \frac{m}{ab} \begin{pmatrix} \frac{1}{3}ab^3 & -\frac{1}{4}a^2b^2 & 0 \\ -\frac{1}{4}a^2b^2 & \frac{1}{3}ba^3 & 0 \\ 0 & 0 & \frac{1}{3}(ab^3 + ba^3) \end{pmatrix} = m \begin{pmatrix} \frac{1}{3}b^2 & -\frac{1}{4}ab & 0 \\ -\frac{1}{4}ab & \frac{1}{3}a^2 & 0 \\ 0 & 0 & \frac{1}{3}(a^2 + b^2) \end{pmatrix} \tag{8.26}$$

What happens to this result if a or b equals zero? Does that result make sense? Perhaps relate it to some equation appearing earlier in this chapter.

In a previous study of rotations, you probably encountered the moment of inertia as a subject in its own right, leading to the equation $\vec{L} = I\vec{\omega}$. You can now see that this was only a first approximation to the subject. But then what *is* that old I as expressed in the new language? The answer is

$$\int r_\perp^2 \, dm \qquad \text{or} \qquad \vec{\omega} \cdot I(\vec{\omega})/\omega^2 \tag{8.27}$$

where $r_\perp$ is the perpendicular distance to your axis. The first expression is what appears in an introductory text, though perhaps not in this notation. If you use a coordinate system where this axis is the z-axis, then $r_\perp = \sqrt{x^2 + y^2}$ and this is I_{zz}, the bottom right element of the above matrix. The other diagonal elements of the matrix are then the moments of inertia about the x- and y-axes. The second, more complicated looking expression in Eq. (8.27) is a more general way to relate the tensor to the moment. You can derive it in problem 8.11. Notice how easy it is to abuse the notation, using I for the moment of inertia and I for the tensor of inertia. Stay alert.

Example

► Repeat a previous example, Figure 8.4, only now express it in a new language. Two point masses are at the ends of a light rod lying in the y-z plane as shown on the right. The integral is just a sum over two terms this time. For both masses the value of x is zero, and when you evaluate this sum, the $y^2 + z^2$ factor is just r_1^2 or r_2^2 depending on which mass you're dealing with.

Fig. 8.9 $\hat{x}$ out
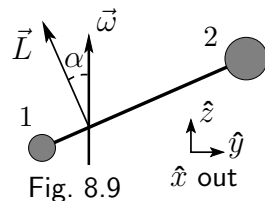
$$I_{xx} = m_1 r_1^2 + m_2 r_2^2$$
$$I_{yy} = (m_1 r_1^2 + m_2 r_2^2) \sin^2 \alpha, \qquad I_{zz} = (m_1 r_1^2 + m_2 r_2^2) \cos^2 \alpha$$
$$I_{xy} = I_{xz} = 0, \qquad I_{yz} = -(m_1 r_1^2 + m_2 r_2^2) \sin \alpha \cos \alpha$$

As a matrix this is
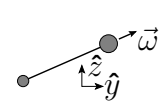
$$(m_1 r_1^2 + m_2 r_2^2) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin^2 \alpha & -\cos \alpha \sin \alpha \\ 0 & -\cos \alpha \sin \alpha & \cos^2 \alpha \end{pmatrix} \tag{8.28}$$

If the angular velocity is along the z-axis, the angular momentum components are

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = (m_1 r_1^2 + m_2 r_2^2) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin^2 \alpha & -\cos \alpha \sin \alpha \\ 0 & -\cos \alpha \sin \alpha & \cos^2 \alpha \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \omega \end{pmatrix}$$

$$= (m_1 r_1^2 + m_2 r_2^2)\omega \begin{pmatrix} 0 \\ -\cos \alpha \sin \alpha \\ \cos^2 \alpha \end{pmatrix}$$

$$\implies \vec{L} = (m_1 r_1^2 + m_2 r_2^2)\, \omega \cos \alpha (-\hat{y} \sin \alpha + \hat{z} \cos \alpha) \tag{8.29}$$

The angular momentum is pointing toward the upper left, perpendicular to the axis between the masses, and not in the direction of $\vec{\omega}$. This behavior is typical, because it is only in the case of a symmetrical rotating body or in cases of a special $\vec{\omega}$ that these two vectors will line up. The impression that you may have gotten in your first introduction to this subject was no doubt restricted to such select examples.

If the angular velocity is along the axis of the rod, $\vec{\omega} = \hat{y}\sin\alpha + \hat{z}\cos\alpha$, then compute $\vec{L}$ as

$$(m_1 r_1^2 + m_2 r_2^2)\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin^2\alpha & -\cos\alpha\sin\alpha \\ 0 & -\cos\alpha\sin\alpha & \cos^2\alpha \end{pmatrix}\begin{pmatrix} 0 \\ \omega\cos\alpha \\ \omega\sin\alpha \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Around this line of rotation, all the mass is *on* the axis, so there's no angular momentum in this approximation that these are *point* masses.

If the axis of rotation is in the direction $(-\hat{y}\sin\alpha + \hat{z}\cos\alpha)$, the direction of the result in Eq. (8.29), and the direction perpendicular to the axis between the masses, compute the angular momentum:

$$(m_1 r_1^2 + m_2 r_2^2)\begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin^2\alpha & -\cos\alpha\sin\alpha \\ 0 & -\cos\alpha\sin\alpha & \cos^2\alpha \end{pmatrix}\begin{pmatrix} 0 \\ -\omega\sin\alpha \\ \omega\cos\alpha \end{pmatrix} = (m_1 r_1^2 + m_2 r_2^2)\begin{pmatrix} 0 \\ -\omega\sin\alpha \\ \omega\cos\alpha \end{pmatrix}$$

This $\vec{L}$ is in the same direction as $\vec{\omega}$, and if you start with $\vec{\omega} = \omega\hat{x}$ (out of the page), you get a similar result, with the same proportionality factor, $(m_1 r_1^2 + m_2 r_2^2)$. Along *these* directions the vectors $\vec{\omega}$ and $\vec{L}$ are aligned.

Perpendicular Axis Theorem

This is a special result about the components of the inertia tensor, and it is occasionally useful.

It applies when a mass is distributed in a plane, so that it is essentially two-dimensional. Make that the x-y plane, then $I_{zz} = I_{xx} + I_{yy}$. The proof involves nothing more than writing the values of the components from Eq. (8.24):

$$I_{zz} = \int dm\,(x^2 + y^2),$$

$$\text{but} \qquad I_{xx} = \int dm\,(y^2 + z^2) = \int dm\,y^2 \qquad\qquad (8.30)$$

$$\text{and} \qquad I_{yy} = \int dm\,(x^2 + z^2) = \int dm\,x^2$$

because $z = 0$ for all dm, and that's all there is to it.

Example

▶ A thin disk of mass M and radius R has its center at the origin and has $z = 0$. Compute the inertia components:

$$I_{zz} = \int dm\,(x^2 + y^2) = \int_0^R 2\pi r\,dr\,\frac{M}{\pi R^2}r^2 = \frac{MR^2}{2}$$

Use the perpendicular axis theorem, and because the symmetry of the disk implies that $I_{xx} = I_{yy} = I_{zz}/2$, you see that both of these are $MR^2/4$. The
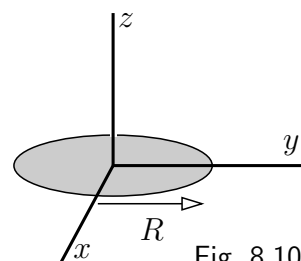
Fig. 8.10

products of inertia are all zero because they all involve integrating an odd function over a symmetric domain: $\int_{-a}^{a} x \, dx = 0$.

$$(I) = \frac{MR^2}{4} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{8.31}$$

Suppose that at some instant the angular velocity is between the $\hat{z}$ and $\hat{y}$ directions: $\vec{\omega} = \omega(\hat{z} \cos \alpha + \hat{y} \sin \alpha)$. Then $\vec{L} = MR^2 \omega(2\hat{z} \cos \alpha + \hat{y} \sin \alpha)/4$. It points in a direction closer to $\hat{z}$ than does $\vec{\omega}$. If I keep turning the disk about this $\vec{\omega}$ direction, the angular momentum will rotate around it. All the components of the inertia tensor are now time dependent, but don't worry about it because I've already found the relation between $\vec{L}$ and $\vec{\omega}$, and that's all I need. $\vec{L}$ will trace out a cone around $\vec{\omega}$, and that means that there must be a time-varying torque to make this happen: $\vec{\tau} = d\vec{L}/dt$.

The angle between the two vectors is (problem 8.27).

$$\sin \beta = \cos \alpha \sin \alpha / \sqrt{3 \cos^2 \alpha + 1} \tag{8.32}$$
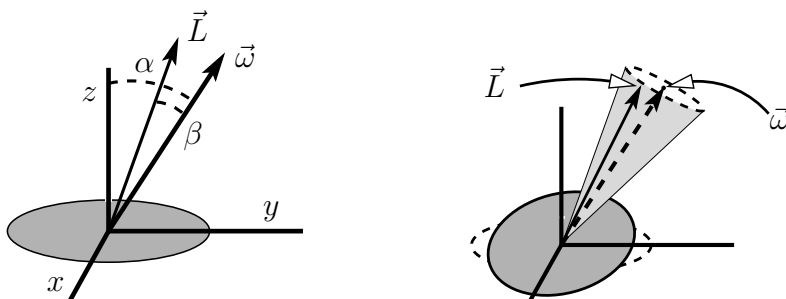
Fig. 8.11

How much torque does it take to keep this disk rotating about the line $\vec{\omega}$? Forget the coordinate system; that will just get in the way now. The vector $\vec{L}$ is spinning about a fixed line defined by $\vec{\omega}$. The length of $\vec{L}$ is constant, so its time derivative is perpendicular to itself. It is also perpendicular to $\vec{\omega}$, and its value is then

$$d\vec{L}/dt = \vec{\omega} \times \vec{L} = \vec{\tau}$$

The magnitude of this torque is $\omega L \sin \beta$, and you can get the magnitude L from the vector $\vec{L}$ in the preceding paragraph.

$$\omega L \sin \beta = \tfrac{1}{4}\omega MR^2 \omega \left(4 \cos^2 \alpha + \sin^2 \alpha\right)^{1/2} \cos \alpha \sin \alpha / \sqrt{3 \cos^2 \alpha + 1} = \frac{1}{4} M^2 \omega^2 \cos \alpha \sin \alpha$$

There was so much cancellation of the complicated factors in this calculation that you should suspect that there's an easier way. There is. The magnitudes and angles are constant, so why not evaluate the magnitude of the product at the initial time?

$$\vec{\omega} \times \vec{L} = \omega(\hat{z} \cos \alpha + \hat{y} \sin \alpha) \times MR^2 \omega(2\hat{z} \cos \alpha + \hat{y} \sin \alpha)/4 = \frac{1}{4} MR^2 \omega^2 \hat{x} \cos \alpha \sin \alpha$$

Can you do this calculation of the torque without referring to a specific coordinate system, just manipulating the original form for $I(\vec{\omega})$? Yes, but it's not much help:

$$\vec{\omega} \times \vec{L} = \vec{\omega} \times I(\vec{\omega}) = \vec{\omega} \times \int dm \, \vec{r} \times (\vec{\omega} \times \vec{r})$$

$$= \vec{\omega} \times \int dm \left[r^2 \vec{\omega} - \vec{r}(\vec{\omega} \cdot \vec{r})\right] = -\int dm \, (\vec{\omega} \times \vec{r})(\vec{\omega} \cdot \vec{r})$$

If you can find a use for this equation, you're welcome to it.

This calculation has a very practical application. When you have driven your car for a long time it may need adjustment. If you hit a curb with the wheel of your car you can knock the wheel out of alignment, and it will certainly need adjustment when you then feel strong vibrations in the car's steering wheel as you drive it. You then take the car in to the shop and ask a mechanic to adjust the wheels so that their angular velocity and angular momentum vectors are parallel, hoping thereby to eliminate the torque $\vec{\omega} \times \vec{L}$. Perhaps not. You may get better results if you ask for a wheel alignment.

A related problem will occur if the axis of rotation does not pass through the center of mass of the wheel. It is then out of balance and will also cause vibrations, calling for a wheel balancing.

Example

▶ What is the tensor of inertia of a ball about a point on its surface? In the problem 8.20 you will compute the moment of inertia of a uniform ball about an axis through its center. The result is $2MR^2/5$. Once you know this, you immediately know all the components of its inertia tensor because of the ball's symmetry: this same moment of inertia appears all along the matrix diagonal and all the off-diagonal elements are zero. It is a multiple of the unit matrix. What are the components when the origin is a point on the surface of the ball? Choose the coordinate system so that the z-axis passes through the center and use the parallel axis theorem.
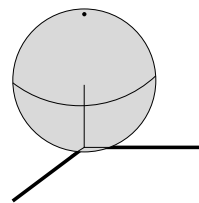
$$I(\vec{\omega}) = I_{\text{cm}}(\vec{\omega}) + M\vec{r}_{\text{cm}} \times (\vec{\omega} \times \vec{r}_{\text{cm}})$$

$$\longrightarrow \quad (I) = \frac{2MR^2}{5} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + M \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{pmatrix}_{\text{at center}}$$

$$= \frac{2MR^2}{5} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + M \begin{pmatrix} R^2 & 0 & 0 \\ 0 & R^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \frac{MR^2}{5} \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Kinetic Energy

When a rigid body is rotating about a fixed axis, what is its kinetic energy? It is a straight-forward calculation.

$$K = \frac{1}{2} \int v^2 \, dm = \frac{1}{2} \int dm \, (\vec{\omega} \times \vec{r})^2 = \frac{1}{2} \int dm \, (\vec{\omega} \times \vec{r}) \cdot (\vec{\omega} \times \vec{r})$$

$$= \frac{1}{2} \int dm \, \vec{\omega} \cdot [\vec{r} \times (\vec{\omega} \times \vec{r})] = \frac{1}{2} \vec{\omega} \cdot I(\vec{\omega}) \tag{8.33}$$

The sole vector identity that you need for this is that you can interchange the dot and the cross product in the triple scalar product: $\vec{A} \cdot \vec{B} \times \vec{C} = \vec{A} \times \vec{B} \cdot \vec{C}$, where $\vec{C} = \vec{\omega} \times \vec{r}$.

If $\vec{L}$ and $\vec{\omega}$ are aligned, so that $\vec{L} = I\vec{\omega}$, then this kinetic energy is $\frac{1}{2}I\omega^2$.

What if a body is both rotating and moving? For example rolling motion. The calculation is similar, again adding all the pieces of kinetic energy in the body. Let $\vec{\omega}$ be the rotation rate about the moving origin, and let $\vec{v}_0$ be the velocity of this origin with respect to you. The velocity of a mass dm is then $\vec{v}_0 + \omega \times \vec{r}$, where this $\vec{r}$ is from the moving origin to dm.
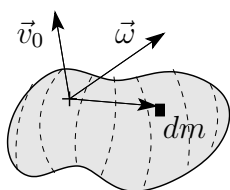
$$K = \frac{1}{2} \int dm \, v^2 = \frac{1}{2} \int dm \, [\vec{v}_0 + \omega \times \vec{r}]^2$$

$$= \frac{1}{2} \int dm \, v_0^2 + \frac{1}{2} \int dm \, (\omega \times \vec{r})^2 + \vec{v}_0 \cdot \vec{\omega} \times \int dm \, \vec{r} \tag{8.34}$$

$$= \frac{1}{2} m v_0^2 + \frac{1}{2} \vec{\omega} \cdot I(\vec{\omega}) + m\vec{v}_0 \cdot \vec{\omega} \times \vec{r}_{\text{cm}}$$

If the origin for computing the inertia tensor is the center of mass, then the final term is zero, and the total kinetic energy is in that case the sum of the rotational energy about the center of mass and the kinetic energy of a point mass *at* the center of mass.

In the same spirit, what is the angular momentum for an object that is both moving and rotating?

$$\vec{L} = \int dm\, \vec{r} \times \vec{v} = \int dm\, \vec{r} \times (\vec{v}_0 + \omega \times \vec{r}) = \left(\int dm\, \vec{r} \right) \times \vec{v}_0 + \int dm\, \vec{r} \times (\omega \times \vec{r})$$
$$= \vec{r}_{\mathsf{cm}} \times m\vec{v}_0 + I(\vec{\omega}) \tag{8.35}$$

This is the sum of the angular momentum of a point mass *at* the center of mass plus the rotational angular momentum *about* the center of mass.

Having looked at kinetic energy and angular momentum from different origins, what about torque? The torque on a piece of mass dm is $\vec{r} \times d\vec{F}$. If you choose a different origin, so that the position of the old origin is $\vec{R}$ with respect to the new origin, then the coordinate of dm is

$$\vec{r}_{\mathsf{new}} = \vec{r} + \vec{R} \qquad \mathsf{so} \qquad \vec{\tau}_{\mathsf{new}} = \int \vec{r}_{\mathsf{new}} \times d\vec{F} = \int (\vec{r} + \vec{R}) \times d\vec{F} = \vec{\tau}_{\mathsf{old}} + \vec{R} \times \int d\vec{F}$$

If the total force is zero, the torque is independent of the origin.

8.4 Principal Axes

Aside from spinning with zero torque, there are other reasons to be interested in lining up the angular velocity and the angular momentum; it makes calculations easier. When calculating the components of the inertia tensor for the example of two point masses, Eq. (8.28), the matrix would have been diagonal if you had chosen the coordinate system differently. If $\hat{y}$ is along the line connecting the masses then $\vec{\omega}$ would have had two non-zero components, but the matrix of inertia components would have been diagonal.

The integral for the inertia components is the same as before, being a sum of two terms, but this time the sum is easier.

$$I_{xx} = m_1 r_1^2 + m_2 r_2^2, \qquad I_{zz} = (m_1 r_1^2 + m_2 r_2^2)$$

All other components zero
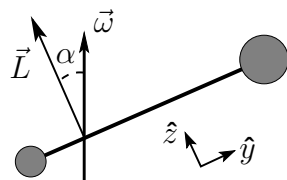
The equation $\vec{L} = I(\vec{\omega})$ translates into components as

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = (m_1 r_1^2 + m_2 r_2^2) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \omega \sin \alpha \\ \omega \cos \alpha \end{pmatrix}$$
$$\implies \vec{L} = (m_1 r_1^2 + m_2 r_2^2)\hat{z}\, \omega \cos \alpha$$

This agrees exactly with Eq. (8.29). Look and see.

When you compute the components of I, the definition is $I(\vec{e}_i) = I_{ji}\vec{e}_j$ (recall the implied sum). If there is a direction such that $\vec{\omega}$ and $\vec{L}$ are parallel, choose $\vec{e}_1$ along that direction. Then $\vec{L} = I(\vec{\omega})$ and $I(\vec{e}_1) = $ (a multiple of)$\vec{e}_1 = I_{11}\vec{e}_1$. The first column of the matrix has exactly one non-zero entry in the upper left corner of the matrix. The matrix is symmetric, so the upper row has only a single entry as well.

If you can find other directions along which $\vec{\omega}$ is parallel to $\vec{L}$, you can use those directions for basis vectors too. If there are three such directions, then you have a complete basis and the matrix is diagonal. Can you do this? Yes. In the case of the inertia tensor it's guaranteed.

Eigenvectors and Eigenvalues

The statement that $\vec{\omega}$ is parallel to $\vec{L}$ means that one vector is a multiple of the other: $\vec{L} = \lambda\vec{\omega}$. This is the equation

$$I(\vec{\omega}) = \lambda\vec{\omega} \tag{8.36}$$

If this is satisfied, then $\vec{\omega}$ is an "eigenvector" of I, and λ is an "eigenvalue". Several questions arise.

	inertia tensor	all tensors
1. Do these eigenvectors always exist?	yes	yes
2. Can you always make a basis out of them?	yes	no
3. If you can make a basis, is it orthogonal?	yes	no
4. How do you find them?	Therein lies a tale.	

Translate the problem into components. In the notation of Eq. (8.19) the last equation (8.36) becomes

$$\begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \lambda \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}$$

Assuming that you've already done the integrals to know the components of I, there are four unknowns in these three equations: λ and the three components of $\vec{\omega}$. It doesn't sound promising. There is always one solution, $\vec{\omega} = 0$, so all three components of this vector are zero and λ is arbitrary. That is not a very interesting solution though, and you can't start to construct a basis out of it. Are there any non-zero solutions? No, at least not for arbitrary λ. But, there are always certain values of λ for which there is a non-zero solution for $\vec{\omega}$.

> Now would be a good idea to review section 0.10, especially the material on simultaneous equations. Also reread the material leading to equation (3.59).

Rewrite this set of equations by moving everything to one side.

$$\left[\begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Further rewrite it as

$$\begin{pmatrix} I_{xx} - \lambda & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} - \lambda & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} - \lambda \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{8.37}$$

This matrix acting on the components of $\vec{\omega}$ gives the zero vector. Under almost all circumstances that requires $\vec{\omega}$ to be zero, but there can be an exception. That is that the matrix be singular— its determinant is zero. If you want a non-zero solution (and you do), then the determinant of the coefficients of the three linear equations must be zero*. That is an algebraic equation for λ, a cubic equation, and it will always have a solution. In the case at hand, the tensor of inertia, it will always have three real solutions. For each solution λ (an eigenvalue), there is a corresponding $\vec{\omega}$ (an eigenvector).

There are a few general results to prove about this process, but first I'll carry through the solution in a special case, one that doesn't entail as much algebra as the general case, and the cubic equation will be easy to solve. The mass is confined to a plane, and that will define the x-y coordinates.

---

* Go through this statement for a 1×1 matrix with a one element column matrix. What does it translate to?

Example

▶ Four masses, all the same, are at the four corners of a square: $(0, 0)$, $(0, a)$, (a, a), $(a, 0)$, all at $z = 0$. Now to compute the nine components of the inertia tensor, there are really just three integrals,

$$\int dm\, x^2 = 2ma^2, \quad \int dm\, y^2 = 2ma^2, \quad \text{and} \quad -\int dm\, xy = -ma^2$$
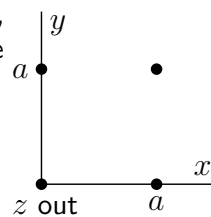
Any term involving z gives zero, so the defining equation (8.24) produces the components

$$(I) = \begin{pmatrix} 2ma^2 & -ma^2 & 0 \\ -ma^2 & 2ma^2 & 0 \\ 0 & 0 & 4ma^2 \end{pmatrix}$$

The eigenvector equation $I(\vec{\omega}) - \lambda\vec{\omega} = 0$, expressed in components as (8.37) is

$$\begin{pmatrix} 2ma^2 - \lambda & -ma^2 & 0 \\ -ma^2 & 2ma^2 - \lambda & 0 \\ 0 & 0 & 4ma^2 - \lambda \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Before setting the determinant of this matrix to zero, make a change of variables: $\lambda = ma^2\lambda'$. This saves a lot of writing.

$$\det ma^2 \begin{pmatrix} 2 - \lambda' & -1 & 0 \\ -1 & 2 - \lambda' & 0 \\ 0 & 0 & 4 - \lambda' \end{pmatrix} = 0 = (ma^2)^3(4 - \lambda')\left[(2 - \lambda')^2 - 1\right]$$

with roots $\lambda' = 4,\ 3,\ 1$.

For each of these eigenvalues, compute the eigenvector.

$$\lambda' = 4 \longrightarrow \begin{pmatrix} -2 & -1 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \longrightarrow \begin{matrix} -2\omega_x - \omega_y = 0 \\ -\omega_x - 2\omega_y = 0 \\ 0 = 0 \end{matrix}$$

The first pair have the unique solution $\omega_x = \omega_y = 0$, the last equation offers no constraint, so ω_z is arbitrary and the eigenvector is $\omega_z\hat{z}$. That is, any multiple of $\hat{z}$.

$$\lambda' = 3 \longrightarrow \begin{pmatrix} -1 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \longrightarrow \begin{matrix} -\omega_x - \omega_y = 0 \\ -\omega_x - \omega_y = 0 \\ \omega_z = 0 \end{matrix}$$

This has zero z-component for $\vec{\omega}$, and $\omega_x = -\omega_y$. The eigenvector is then $\vec{\omega} = \hat{x} - \hat{y}$ or any multiple of it.

$$\lambda' = 1 \longrightarrow \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \longrightarrow \begin{matrix} \omega_x - \omega_y = 0 \\ -\omega_x + \omega_y = 0 \\ 3\omega_z = 0 \end{matrix}$$

This time the eigenvector is any multiple of $\hat{x} + \hat{y}$.

Picture these motions. For the $\lambda' = 1$ case, the rotation is about the line from the origin to the far corner of the square. Only two masses are moving and they are symmetrically placed with respect to the rotation axis at a distance $r_\perp = a/\sqrt{2}$. The angular momentum is obviously in the same direction, and the moment of inertia about this axis is $2m(a/\sqrt{2})^2 = ma^2$. That's precisely the eigenvalue λ for this case.
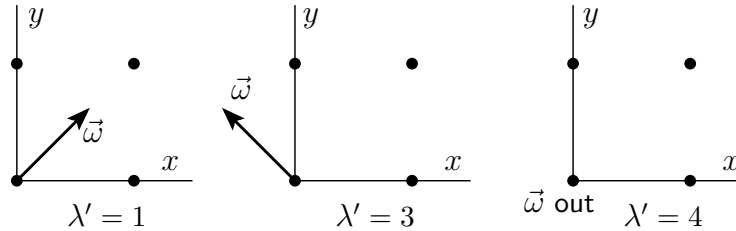
Fig. 8.12

For the $\lambda' = 3$ case, the rotation axis is still in the x-y plane, but along $\hat{y} - \hat{x}$, perpendicular to the preceding one. Now three masses are moving, and the rotational inertia about this axis is $2m(a/\sqrt{2})^2 + m(a\sqrt{2})^2 = 3ma^2$ as promised. This direction still has enough symmetry that it is easy to believe that $\vec{L}$ is in the same direction as $\vec{\omega}$.

For $\lambda' = 4$, the eigenvector is along $\hat{z}$, and again three masses are in motion, with moment $2ma^2 + m(a\sqrt{2})^2 = 4ma^2$. This is a sufficiently non-symmetric case that it is not immediately obvious that the angular momentum is in the same direction as $\vec{\omega}$, but do a couple of cross products and you can easily persuade yourself that it's true. Again, the eigenvalue $\lambda = 4ma^2$ is this moment of inertia about the z-axis.

The basis $\hat{x}$, $\hat{y}$, $\hat{z}$, in which I computed the components of I in this example were arbitrary. I picked them for their obvious convenience in computing the answer. What if I pick another basis, one consisting of the three orthogonal eigenvectors that I just found? I'll call them $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$ as a common notation used for basis vectors.

$$\vec{e}_1 = (\hat{x} + \hat{y})/\sqrt{2}, \qquad \vec{e}_2 = (\hat{y} - \hat{x})/\sqrt{2}, \qquad \vec{e}_3 = \hat{z} \tag{8.38}$$

Does it matter that these are unit vectors? No, as long as they are independent that is good enough. You could drop the $\sqrt{2}$ and compute the matrix for I, but for other manipulations you should be consistent. In keeping with the change I will use subscripts 1, 2, 3 instead of x, y, z on the components

$$I(\vec{e}_1) = \sum_j I_{j1}\vec{e}_j = I_{11}\vec{e}_1 + I_{21}\vec{e}_2 + I_{31}\vec{e}_3 = \lambda_1\vec{e}_1$$

This determines the first column of the matrix for I in this basis. The other columns are found the same way.

$$I(\vec{e}_2) = \sum_j I_{j2}\vec{e}_j = I_{12}\vec{e}_1 + I_{22}\vec{e}_2 + I_{32}\vec{e}_3 = \lambda_2\vec{e}_2$$

$$I(\vec{e}_3) = \sum_j I_{j3}\vec{e}_j = I_{13}\vec{e}_1 + I_{23}\vec{e}_2 + I_{33}\vec{e}_3 = \lambda_3\vec{e}_3$$

The matrix for I is now

$$(I) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} ma^2 & 0 & 0 \\ 0 & 3ma^2 & 0 \\ 0 & 0 & 4ma^2 \end{pmatrix}$$

and in this basis the matrix is diagonal. That can provide significant advantages in manipulation for further calculations. Notice that you would get the same matrix even if you do not put in the $\sqrt{2}$ factors in Eq. (8.38).

Is it always easy to find the eigenvectors? No, this example is very easy and you didn't really have to confront a cubic equation to solve for the eigenvalues. Alas, it is not typical.

8.5 Properties of Eigenvectors

There are some general results to be proved about these eigenvectors and eigenvalues, and the methods of proof used here will show up in several other important contexts, so they are worth learning. First an identity about the inertia tensor.

$$I(\vec{\omega}) = \int dm\, \vec{r} \times (\vec{\omega} \times \vec{r}) = \int dm\, \left[r^2 \vec{\omega} - \vec{r}(\vec{\omega} \cdot \vec{r}) \right]$$

$$\vec{\omega}_1^{\,*} \cdot I(\vec{\omega}_2) = \vec{\omega}_1^{\,*} \cdot \int dm\, \left[r^2 \vec{\omega}_2 - \vec{r}(\vec{\omega}_2 \cdot \vec{r}) \right]$$

$$= \int dm\, \left[r^2 \vec{\omega}_1^{\,*} \cdot \vec{\omega}_2 - \vec{\omega}_1^{\,*} \cdot \vec{r}(\vec{\omega}_2 \cdot \vec{r}) \right] = I(\vec{\omega}_1^{\,*}) \cdot \vec{\omega}_2 \tag{8.39}$$

What is a complex conjugation doing here when all these vectors are supposed to be real? And what does it mean anyway? Here I am allowing for the possibility that these may be complex in order to prove that they aren't. How can a vector be complex? In the usual basis, just allow the components to be complex numbers. You won't represent it by a single arrow any more (you could use two), but you won't need to do that anyway.

This symmetry property of the inertia tensor,

$$\vec{\omega}_1^{\,*} \cdot I(\vec{\omega}_2) = I(\vec{\omega}_1^{\,*}) \cdot \vec{\omega}_2 \tag{8.40}$$

plays a key role in much of this analysis. Very similar identities appear in very different-looking contexts such as differential equations, and there they will lead to results much like the ones here. A tensor satisfying this equation is called "Hermitian" or "symmetric" depending on whether a physicist or a mathematician respectively is talking. It parallels the results in the equations (7.53) to (7.55).

In the calculation of eigenvalues and eigenvectors in the example of the preceding section, the eigenvalues (roots of a cubic) were real and the eigenvectors are orthogonal. Is that a general property? For tensors that satisfy the identity derived in Eq. (8.40) it is. To prove this the manipulations are simple but not obvious. Assume that there are two eigenvectors:

$$I(\vec{\omega}_1) = \lambda_1 \vec{\omega}_1 \qquad \text{and} \qquad I(\vec{\omega}_2) = \lambda_2 \vec{\omega}_2$$

Take the scalar product of the first with $\vec{\omega}_2^{\,*}$ and the second with $\vec{\omega}_1^{\,*}$.

$$\vec{\omega}_2^{\,*} \cdot I(\vec{\omega}_1) = \lambda_1 \vec{\omega}_2^{\,*} \cdot \vec{\omega}_1 \qquad \text{and} \qquad \vec{\omega}_1^{\,*} \cdot I(\vec{\omega}_2) = \lambda_2 \vec{\omega}_1^{\,*} \cdot \vec{\omega}_2$$

Take the complex conjugate of the second equation:

$$\vec{\omega}_1 \cdot I(\vec{\omega}_2^{\,*}) = \lambda_2^{\,*} \vec{\omega}_1 \cdot \vec{\omega}_2^{\,*}$$

Subtract this from the first of the preceding equations.

$$\vec{\omega}_2^{\,*} \cdot I(\vec{\omega}_1) - \vec{\omega}_1 \cdot I(\vec{\omega}_2^{\,*}) = (\lambda_1 - \lambda_2^{\,*})\vec{\omega}_1 \cdot \vec{\omega}_2^{\,*}$$

The identity Eq. (8.40) says that the left side of this equation is zero, and the right side is

$$0 = \left(\lambda_1 - \lambda_2^* \right) \vec{\omega}_2^* \cdot \vec{\omega}_1 \tag{8.41}$$

Now take the special case that $\vec{\omega}_1 = \vec{\omega}_2$. These two vectors are the same vector, so necessarily the two eigenvalues are the same also

$$0 = \left(\lambda_1 - \lambda_1^* \right) \vec{\omega}_1^* \cdot \vec{\omega}_1$$

The second factor is not zero, and this implies that the first factor must be. That in turn says that λ is real. With real λ and real components of the inertia tensor, all the components of $\vec{\omega}$ are real too.

Now Eq. (8.41) is

$$0 = \left(\lambda_1 - \lambda_2 \right) \vec{\omega}_2 \cdot \vec{\omega}_1 \tag{8.42}$$

and this implies that for two different eigenvalues, *i.e.* $\lambda_1 \neq \lambda_2$, the corresponding eigenvectors are orthogonal, as exemplified in the example of Figure 8.12: $(\hat{z}) \perp (\hat{y} - \hat{x}) \perp (\hat{x} + \hat{y})$.

These eigenvalues are the moments of inertia about the axis defined by the eigenvector, and they are of course non-negative. Of course? Yes, for an eigenvector,

$$\lambda \omega^2 = \vec{\omega} \cdot I(\vec{\omega}) = \int dm \left[r^2 \omega^2 - (\vec{r} \cdot \vec{\omega})^2 \right] = \omega^2 \int dm \, r_\perp^2$$

where $r_\perp$ is the perpendicular distance from dm to the axis defined by $\vec{\omega}$. This integral will be zero only in the special case that all the mass is on the axis — when the body is a line. In all other cases it is positive. This final integral is the moment of inertia about the $\vec{\omega}$ direction.

8.6 Dynamics
The basic equation for rotating objects is $\vec{\tau} = d\vec{L}/dt$, Eq. (8.6). Now $\vec{L}$ is nothing more than a tensor expression involving $\vec{\omega}$.

$$\vec{\tau} = \frac{d}{dt} I(\vec{\omega}) = I \left(\frac{d\vec{\omega}}{dt} \right) + \frac{dI}{dt}(\vec{\omega})$$

Written this way all the components of the inertia tensor are time-dependent, and that's not a fruitful approach. The trick is to transform to the rotating coordinate system in which the body is at rest. Then the the tensor I is constant. From the equation (5.6),

$$\vec{\tau} = \frac{d\vec{L}}{dt} = \dot{\vec{L}} + \vec{\omega} \times \vec{L} = I(\dot{\vec{\omega}}) + \vec{\omega} \times I(\vec{\omega})$$

and this again uses the chapter five convention that the **dot now refers to the time derivative in the rotating system**. In this rotating system, you can pick the coordinates so that the components of I are a diagonal matrix — the basis using eigenvectors of I.

$$\vec{\tau} = I(\dot{\vec{\omega}}) + \left(\hat{x}\omega_x + \hat{y}\omega_y + \hat{z}\omega_z \right) \times \left(I_{xx}\hat{x}\omega_x + I_{yy}\hat{y}\omega_y + I_{zz}\hat{z}\omega_z \right)$$

$$\begin{aligned}
\tau_x &= I_{xx}\dot{\omega}_x + \omega_y\omega_z(I_{zz} - I_{yy}) \\
\tau_y &= I_{yy}\dot{\omega}_y + \omega_z\omega_x(I_{xx} - I_{zz}) \\
\tau_z &= I_{zz}\dot{\omega}_z + \omega_x\omega_y(I_{yy} - I_{xx})
\end{aligned} \tag{8.43}$$

These are the Euler equations. (As if he didn't have enough things named for him. He even used to be on the Swiss ten franc note.)

Example

▶ If your automobile tire is misaligned, its axis of rotation doesn't line up with its axis of symmetry. Call the z-axis the symmetry axis of the tire, then $I_{xx} = I_{yy}$. As you drive at constant velocity what torque does this cause on your tire? In the inertial system $\vec{\omega}$ is a constant because it is defined by the axle of the car. What happens to $\vec{\omega}$ in the rotating system? Again, Eqs. (5.6) and (5.13):

x-axis

z-axis

α

$\vec{\omega}$
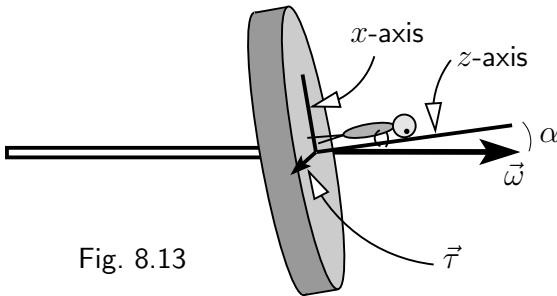
Fig. 8.13

$\vec{\tau}$

$$\frac{d\vec{\omega}}{dt} = \dot{\vec{\omega}} + \vec{\omega} \times \vec{\omega} = \dot{\vec{\omega}} = 0 \qquad (8.44)$$

and this means that it is constant in the rotating system too. The equations (8.43) are now

$$\tau_x = \omega_y \omega_z (I_{zz} - I_{yy}), \qquad \tau_y = \omega_z \omega_x (I_{xx} - I_{zz}), \qquad \tau_z = 0$$

If the misalignment between $\vec{\omega}$ and the axis of the tire is the angle α then pick the x-y coordinates (rotating) so that $\vec{\omega} = \omega \hat{z} \cos\alpha - \omega \hat{x} \sin\alpha$.

$$\tau_x = 0, \qquad \tau_y = \omega^2 \cos\alpha \sin\alpha (I_{zz} - I_{xx}), \qquad \tau_z = 0$$

This is a constant torque about the y-axis, so why does it cause the wheel to shake? This is the *rotating* system remember. In the inertial system (the driver) this torque vector, $\tau_y \hat{y}$, is spinning about the axle at a rate ω, and the size of this torque varies as the square of ω.

How can you picture this? Imagine yourself standing on this disk (tire?) and hanging on to the z-axis as hard as you can. Your feet are straddling the x-axis, which is stationary between your legs. If you look straight ahead you see the angular velocity vector standing still in front of you ($\dot{\vec{\omega}} = 0$). *But* if you look in the distance you may get dizzy from watching the world spinning around. If you look down and to your right you will see the torque vector at your feet, pointing along the y-axis. (Notice that $I_{zz} > I_{xx}$.) And where is $\vec{L}$?

$$\vec{L} = I(\vec{\omega}) = I(\omega \hat{z} \cos\alpha - \omega \hat{x} \sin\alpha) = \omega(I_{zz} \hat{z} \cos\alpha - I_{xx} \hat{x} \sin\alpha)$$

This points in a direction closer to the z-axis than the $\vec{\omega}$ vector does. As you look straight ahead you can see it fixed in front of you on the near side of $\vec{\omega}$. Remember, $\dot{\vec{L}} = 0$. I didn't draw this vector because the picture is cluttered enough already.

Free Rotation

What if there is no torque, and the object is free to rotate in space like a frisbee or a planet. The equations (8.43) (with $\vec{\tau} = 0$) are differential equations for $\vec{\omega}$. They can be solved, but except for the case of a symmetric rigid body the solution is non-linear and tough. For that reason pick $I_{xx} = I_{yy}$.

$$0 = I_{xx}\dot{\omega}_x + \omega_y\omega_z(I_{zz} - I_{xx})$$
$$0 = I_{xx}\dot{\omega}_y + \omega_z\omega_x(I_{xx} - I_{zz}) \qquad (8.45)$$
$$0 = I_{zz}\dot{\omega}_z + 0$$

The last line says that ω_z is a constant. That in turn makes the first two equations linear equations for ω_x and ω_y. You know how to solve those: $e^{\alpha t}$.

$$\omega_x(t) = Ae^{\alpha t}, \qquad \omega_y(t) = Be^{\alpha t}$$

$$0 = I_{xx}\alpha A + \omega_z(I_{zz} - I_{xx})B$$
$$0 = I_{xx}\alpha B - \omega_z(I_{zz} - I_{xx})A$$

A non-zero solution for A and B requires that the determinant vanish.

$$I_{xx}^2\alpha^2 + \omega_z^2(I_{zz} - I_{xx})^2 = 0 \;\longrightarrow\; \alpha = \pm i\,\omega_z\frac{I_{zz} - I_{xx}}{I_{xx}} = \pm i\,\omega' \qquad (8.46)$$

Put these back into one of the equations for A and B to get

$$0 = I_{xx}A\left(\pm i\,\omega_z\frac{I_{zz} - I_{xx}}{I_{xx}}\right) + \omega_z(I_{zz} - I_{xx})B \;\longrightarrow\; \pm iA + B = 0$$

A solution is now
$$\omega_x = Ae^{i\omega' t} + A'e^{-i\omega' t}, \qquad \omega_y = -iAe^{i\omega' t} + iA'e^{-i\omega' t}$$

Chose some initial conditions:

$$\omega_x(0) = \omega_0, \qquad \omega_y(0) = 0 \quad \text{then} \quad \omega_x(t) = \omega_0\cos\omega' t, \qquad \omega_y(t) = \omega_0\sin\omega' t \qquad (8.47)$$

The rotation axis precesses about the z-axis at an angular velocity ω'.

Example

▶ Earth. It is freely spinning in space and it is slightly ellipsoidal so that the moments of inertia aren't the same about all axes. The equatorial bulge makes $I_{zz} > I_{xx} = I_{yy}$. If the angular velocity of the planet exactly lined up with its axis of symmetry, the constant ω_0 in Eq. (8.47) would be zero. It would be too much of a coincidence for the alignment to be perfect and it isn't. It misses by an amount such that if you go to the North pole and look for the angular velocity vector you will find it several meters away. Then it wanders around the pole at a rate $\omega' \approx 2\pi/400$ days. Its motion is not as regular as this rigid body analysis would lead you to believe, but the Earth isn't perfectly rigid either.

That the Earth isn't rigid should lead to damping of this oscillation within years or centuries, but it's still here. It has been a puzzle what keeps this *Chandler wobble* going, but recent analysis points to fluctuating pressure on the bottom of the ocean as a likely source of the excitation.

This precession can give geologists information about the interior of the planet. Equation (8.46) tells you about $I_{zz} - I_{xx}$ and that gives some constraints on the distribution of mass within the Earth. And not just Earth; Mars too. One of the measuring devices sent to Mars looked at that planet's wobble, and that says something about the interior structure of Mars. Why hasn't the wobble been completely damped in the case of Mars? After all, it has no oceans to excite the oscillation. Unknown. Maybe the idea that the oceans cause it on Earth is wrong. You will have to do a search of the current literature on the subject to get some ideas of the complexity of the problem.

Stability of Rotations

What happens when you toss a hammer or a tennis racket up and set it spinning? The answer depends very much on how you do it, and the motion can be very smooth or very wild. If you don't have either a hammer or a racket handy, perhaps you have a heavy rubber band. Then you can wrap it around a book so that the pages don't open up when you toss *it* up and spinning. Depending on the axis about which it is spinning you will get very different results.

To analyze this, start again from Euler's equations (8.43). This time assume that the three moments of inertia are different from each other, *but* assume that the angular velocity vector is almost along one of the principle axes. This calculation will involve a series expansion and a linearization. Assume that the rotation is almost along the z-axis. The torque is zero because you have tossed the object up and it is then free to rotate as it will.

$$\vec{\omega}(t) = \omega_0 \hat{z} + \vec{\epsilon}(t), \qquad \text{with} \qquad \epsilon \ll \omega_0$$

Start from the equations (8.43) with $\vec{\tau} = 0$, and as usual with these expansions, keep only the first order terms in ϵ.

$$0 = I_{xx}\dot{\epsilon}_x + \epsilon_y \omega_0 (I_{zz} - I_{yy})$$
$$0 = I_{yy}\dot{\epsilon}_y + \epsilon_x \omega_0 (I_{xx} - I_{zz})$$
$$0 = I_{zz}\dot{\epsilon}_z$$

Every place that an ϵ^2 appeared I dropped it, and the resulting equations are linear. In the third equation, both ω_x and ω_y are first order in ϵ so that killed its final term.

The third equation says that ϵ_z is a constant, and I may as well take it to be zero because a non-zero value would just be a redefinition of the original rotation rate ω_0 about the z-axis. The other two equations are linear differential equations for ϵ_x and ϵ_y. You know how to handle those: an exponential.

$$\epsilon_x = A e^{\alpha t}, \qquad \epsilon_y = B e^{\alpha t}$$

$$\text{then} \qquad \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} I_{xx}\alpha & \omega_0(I_{zz} - I_{yy}) \\ \omega_0(I_{xx} - I_{zz}) & I_{yy}\alpha \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}$$

To get a non-zero solution the determinant must vanish.

$$I_{xx}I_{yy}\alpha^2 - \omega_0^2(I_{xx} - I_{zz})(I_{zz} - I_{yy}) = 0 \tag{8.48}$$

The nature of these solutions depends on the sign of α^2. If it is negative then α is imaginary and there are oscillations about the $\hat{z}$ direction; if α^2 is positive then α is real and there is exponential movement away from there the $\hat{z}$-axis: that rotation is unstable.

$$\alpha^2 \propto (I_{xx} - I_{zz})(I_{zz} - I_{yy})$$

If z has the largest moment, $I_{zz} > I_{xx}$ and $I_{zz} > I_{yy}$,   then $\alpha^2 < 0$

and this says that α is imaginary and that the motion is stable. If I_{zz} is the *smallest* of the three moments you have the same result: α is imaginary and the motion is stable. It is only in the third case, when I_{zz} is intermediate between the other two moments of inertia, that α^2 is positive and α comes out to be real. That motion is unstable. Take a book and wrap a heavy rubber band about it. Now toss it in the air, spinning about one of the three axes of symmetry. Do it for each axis, and the difference in the results will be very obvious. See for example the YouTube video Solid Body Rotation, done in orbit.

There is more to learn about the stability of these rotations. What is the kinetic energy of rotation about each axis? Use Eq. (8.33) about each axis: $K = \frac{1}{2}\vec{\omega} \cdot \vec{L}$.

$$\vec{\omega} = \omega_0 \hat{x} \longrightarrow K = \frac{1}{2}I_{xx}\omega_0^2, \qquad \vec{\omega} = \omega_0 \hat{y} \longrightarrow K = \frac{1}{2}I_{yy}\omega_0^2, \qquad \vec{\omega} = \omega_0 \hat{z} \longrightarrow K = \frac{1}{2}I_{zz}\omega_0^2$$

The angular momentum in each case is easy because these are eigenvectors of the inertia tensor and are respectively

$$L_x = I_{xx}\omega_0, \qquad L_y = I_{yy}\omega_0, \qquad L_z = I_{zz}\omega_0$$

Write the energy in terms of these angular momenta.

$$\vec{\omega} = \omega_0 \hat{x} \longrightarrow K = \frac{L^2}{2I_{xx}}, \quad \vec{\omega} = \omega_0 \hat{y} \longrightarrow K = \frac{L^2}{2I_{yy}}, \quad \vec{\omega} = \omega_0 \hat{z} \longrightarrow K = \frac{L^2}{2I_{zz}}$$

This assumes the same magnitude for the angular momentum in every case, and these equations say that for a given angular momentum the rotation about the axis with the largest moment of inertia has the smallest kinetic energy.

Whether this object is a rigid body or not, if there are no torques on it the angular momentum is conserved. For a truly rigid body, both angular momentum and mechanical energy are conserved, but completely rigid bodies are a mathematical fiction, and any real object will flex slightly. This has an important consequence: If a satellite is set to rotating about one of its principle axes and if that is one of the two stable axes, then you might expect the rotation to remain about that axis forever. But... A tiny flexibility will cause a tiny friction—a way to dissipate energy. and the only truly stable rotation in this case will be about the axis with the largest moment of inertia, the one that has the smallest kinetic energy for the given angular momentum. This effect destabilizes rotations about the axis having the smallest moment.
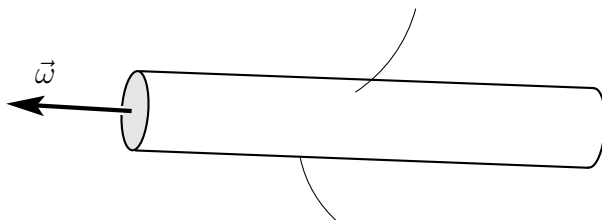
Fig. 8.14

Early in the space program the satellite Explorer-1 was placed in orbit around Earth and set spinning about its long axis, the one with the smallest moment, and one they thought would provide a stable rotation. The satellite however, had whip antennas with more than enough flexibility to destabilize this rotation and to set it tumbling about its really stable axis—end over end—much to the embarrassment of the engineers and physicists involved.

If you have seen the motion pictures 2001 and its sequel 2010, a space ship was left in orbit around Jupiter's moon Io at the end of the first picture, and when they returned in the sequel it was tumbling end-over-end. They made no comment in the movie, but they got the physics right. The same applies to the more recent movie, "Gravity". When Sandra Bullock was first cut loose from the spacecraft, she was tumbling wildly. After a brief time she was rotating smoothly head over feet. Was that really about the axis with the largest moment of inertia? Should it be sideways or as shown, front to back? That does depend on the mass of her backpack, and it appears that they did it correctly in the movie.

8.7 Perturbation Theory˙

If the equation for the eigenvalues of an inertia tensor is easy to solve then you're lucky. What if it isn't? At worst you may have to resort to solving it numerically, but there is an intermediate situation that comes up surprisingly often and that falls between these extremes. If the problem is *close* to one that is easily solved, then the method of perturbations is available. This involves a series expansion that can be very useful in both the quantitative and qualitative analysis of these problems.

For example, for a mass of uniform density and in the shape of a rectangular box, it is easy to find the eigenvectors and eigenvalues of the inertia tensor using the center of the box as the origin. The

principle axes are parallel to the edges of the box, and the eigenvalues are easily computed. What if you now add a small mass to the corner of the box? Suddenly the symmetry is lost and you don't have any intuitive feeling for the directions of the eigenvectors, *but* in some sense the eigenvectors ought to be close to the original ones that were parallel to the edges.

Example

▶ Compute the components of the inertia tensor in the example just mentioned. Then what are the eigenvectors and eigenvalues of the tensor? Call the sides of the box a, b, c. First compute $\int dm\, x^2$. After that it's all downhill. For dm use a slice of the whole mass:

$$dm = \frac{M}{abc} bc\, dx \longrightarrow \int dm\, x^2 = \int_{-a/2}^{a/2} \frac{M}{abc} bc\, dx\, x^2 = \frac{M}{a} \frac{x^3}{3} \bigg|_{-a/2}^{a/2} = \frac{Ma^2}{12}$$

The integrals of $y^2\, dm$ and $z^2\, dm$ just change a to b to c, and the off-diagonal elements are zero. This produces (I_0). For little m, it is just a point mass, so Eq. (8.24) produces the result without any integration. This is (I_1).

$$(I_0) = \frac{M}{12} \begin{pmatrix} b^2 + c^2 & 0 & 0 \\ 0 & a^2 + c^2 & 0 \\ 0 & 0 & a^2 + b^2 \end{pmatrix} \qquad (I_1) = \frac{m}{4} \begin{pmatrix} b^2 + c^2 & -ab & -ac \\ -ab & a^2 + c^2 & -bc \\ -ac & -bc & b^2 + c^2 \end{pmatrix}$$

The total tensor comes from both of the masses and the components are the sum of these two matrices. Suddenly, finding the eigenvectors and eigenvalues is hard. If you want a general answer for this case, look up the formula for the solution of a general cubic equation. You will then see why I'm not going to use it.

If $m \ll M$, the eigenvectors and eigenvalues will be close to those of I_0, and because that matrix is diagonal, you already know what these approximate values are. The next problem is to find the corrections to them.

Fig. 8.15

The idea is to assume that there is a power series expansion for these results and then to compute the terms of that expansion one at a time. This is

$$\begin{aligned} I &= I_0 + \epsilon I_1 \qquad \text{and} \qquad I(\vec{\omega}) = \lambda \vec{\omega}, \quad \text{with} \\ \vec{\omega} &= \vec{\omega}_0 + \epsilon \vec{\omega}_1 + \epsilon^2 \vec{\omega}_2 + \cdots \qquad \text{and} \qquad \lambda = \lambda_0 + \epsilon \lambda_1 + \epsilon^2 \lambda_2 + \cdots \end{aligned} \qquad (8.49)$$

This introduction of "ϵ" is a convenient way to keep track of the many terms. None of these $\vec{\omega}_{0,\,1,\,2,}$ ... or the similarly subscripted λ's are known. That doesn't matter because you will let the equations tell you what they are. Plug in:

$$(I_0 + \epsilon I_1)(\vec{\omega}_0 + \epsilon \vec{\omega}_1 + \cdots) = (\lambda_0 + \epsilon \lambda_1 + \cdots)(\vec{\omega}_0 + \epsilon \vec{\omega}_1 + \cdots)$$

Multiply everything and look for the various coefficients of powers of ϵ.

$$\begin{aligned} \epsilon^0 &: I_0(\vec{\omega}_0) = \lambda_0 \vec{\omega}_0 \\ \epsilon^1 &: I_0(\vec{\omega}_1) + I_1(\vec{\omega}_0) = \lambda_0 \vec{\omega}_1 + \lambda_1 \vec{\omega}_0 \\ \epsilon^2 &: I_0(\vec{\omega}_2) + I_1(\vec{\omega}_1) = \lambda_0 \vec{\omega}_2 + \lambda_1 \vec{\omega}_1 + \lambda_2 \vec{\omega}_0 \\ & \quad \cdots \end{aligned}$$

When you have two power series that equal each other, this requires that the individual terms must match too, and that is all that I did to extract the separate equations.
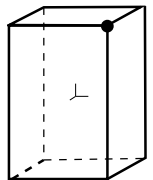
Now rearrange these equations to put them into a standard form.

$$I_0(\vec{\omega}_0) - \lambda_0\vec{\omega}_0 = 0 \tag{0}$$
$$I_0(\vec{\omega}_1) - \lambda_0\vec{\omega}_1 = \lambda_1\vec{\omega}_0 - I_1(\vec{\omega}_0) \tag{1}$$
$$I_0(\vec{\omega}_2) - \lambda_0\vec{\omega}_2 = \lambda_1\vec{\omega}_1 + \lambda_2\vec{\omega}_0 - I_1(\vec{\omega}_1) \tag{2}$$
$$\cdots \tag{8.50}$$

The 0th equation above is the one that you supposedly know how to solve. It's the one without the extra complicating term coming from the point mass on the corner. For the rest...

The clever method is to use Eq. (8.40), but now all the numbers are real: $\left[\vec{\omega}_1 \cdot I(\vec{\omega}_2) = I(\vec{\omega}_1) \cdot \vec{\omega}_2\right]$. Apply it to these perturbation equations. Take the scalar product of Eq. (8.50)(1) with $\vec{\omega}_0$.

$$\vec{\omega}_0 \cdot \left[I_0(\vec{\omega}_1) - \lambda_0\vec{\omega}_1\right] = \vec{\omega}_0 \cdot \left[\lambda_1\vec{\omega}_0 - I_1(\vec{\omega}_0)\right] \tag{8.51}$$

Apply the symmetry property of the inertia tensor to change the left side of this to

$$I_0(\vec{\omega}_0) \cdot \vec{\omega}_1 - \lambda_0\vec{\omega}_0 \cdot \vec{\omega}_1 = \left[I_0(\vec{\omega}_0) - \lambda_0\vec{\omega}_0\right] \cdot \vec{\omega}_1 = 0$$

This implies that the scalar product on the right side of Eq. (8.51) must be zero also. That is

$$0 = \vec{\omega}_0 \cdot \left[\lambda_1\vec{\omega}_0 - I_1(\vec{\omega}_0)\right] = \vec{\omega}_0 \cdot \lambda_1\vec{\omega}_0 - \vec{\omega}_0 \cdot I_1(\vec{\omega}_0) \longrightarrow \lambda_1 = \frac{\omega_0 \cdot I_1(\vec{\omega}_0)}{\vec{\omega}_0 \cdot \vec{\omega}_0}$$

and you have found the correction to the original eigenvalue without doing any additional matrix manipulation — only taking a scalar product. The magnitude of $\vec{\omega}_0$ cancels from this equation, so you can make its size anything you want. Translate this into the notation of matrices and you have

$$\vec{\omega}_0 = \begin{pmatrix} \omega_0 \\ 0 \\ 0 \end{pmatrix} \longrightarrow \lambda_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \frac{m}{4} \begin{pmatrix} b^2+c^2 & -ab & -ac \\ -ab & a^2+c^2 & -bc \\ -ac & -bc & b^2+c^2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{m}{4}(b^2+c^2) \tag{8.52}$$

as the first order correction to the first eigenvalue. You will easily see that the first order corrections to the other eigenvalues are the other diagonal elements of (I_1).

What is the first order correction to the eigenvector? $\vec{\omega}_1$? That requires solving Eq. (8.50)(1). Write it out in all its gory detail, and the the equations that you have to solve are surprisingly simple.

$$I_0(\vec{\omega}_1) - \lambda_0\vec{\omega}_1 = \lambda_1\vec{\omega}_0 - I_1(\vec{\omega}_0)$$

$$\frac{M}{12} \begin{pmatrix} b^2+c^2 & 0 & 0 \\ 0 & a^2+c^2 & 0 \\ 0 & 0 & a^2+b^2 \end{pmatrix} \begin{pmatrix} \omega_{1x} \\ \omega_{1y} \\ \omega_{1z} \end{pmatrix} - \frac{M}{12}(b^2+c^2)\begin{pmatrix} \omega_{1x} \\ \omega_{1y} \\ \omega_{1z} \end{pmatrix}$$

$$= \frac{m}{4}(b^2+c^2)\begin{pmatrix} \omega_0 \\ 0 \\ 0 \end{pmatrix} - \frac{m}{4}\begin{pmatrix} b^2+c^2 & -ab & -ac \\ -ab & a^2+c^2 & -bc \\ -ac & -bc & b^2+c^2 \end{pmatrix} \begin{pmatrix} \omega_0 \\ 0 \\ 0 \end{pmatrix} \tag{8.53}$$

$$\longrightarrow \frac{M}{12}\begin{pmatrix} 0 \\ (a^2-b^2)\omega_{1y} \\ (a^2-c^2)\omega_{1z} \end{pmatrix} = \frac{m}{4}\begin{pmatrix} 0 \\ -ab \\ -ac \end{pmatrix}\omega_0 \tag{8.54}$$

And here you have the equations for ω_{1y} and ω_{1z}. Their sizes are proportional to m/M, as befits their being small corrections.

Are you done here? I may be but you aren't. What does this result *look* like? If a, b, and c are something like 1, 2, 3 in some units, then how does the direction of the eigenvector change when this correction is added? Is it toward or away from m? And if the sides are instead 2, 3, 1, then does your answer to this change? Play around with special cases.

What about ω_{1x}, the other component of the correction. Take that to be zero, because including it would just redefine the original ω_0. (And there's a technical reason, because if you go into the higher order corrections it will really mess things up if you don't do this.)

Exercises

1 Three equal point masses m are at the corners of the right triangle $(x, y) = (0,0)$, $(a,0)$, $(0,b)$. Where is the center of mass?

2 Remove the masses from the preceding triangle and replace the sides with thin rods of constant linear mass density λ. Where is the center of mass?

3 Remove the rods from the preceding triangle and replace the area with a uniform sheet of constant area mass density σ. Where is the center of mass?

4 Write out the equations (8.4) explicitly for $2 + 2 = 4$ point masses and verify that it does what it says.

5 A tensor f is defined by the equation $f(\vec{v}) = \vec{v}$ for all $\vec{v}$. Compute the components of f.

6 A vector-valued function F acts on vectors in the plane, reversing the direction of each. Compute the components of this function, a 2×2 matrix.

7 With respect to the origin, what are the components of the tensor of inertia for a point mass m at the point $(x, y, z) = (a, 0, 0)$?

8 With respect to the origin, what are the components of the tensor of inertia for a point mass m at the point $(x, y, z) = (a, a, 0)$? and what if its coordinates are (a, a, a)?

9 Use the summation convention and notation of Eq. (0.56) to write the equations (8.17), (8.21), (8.22).

10 For the springs that appear in figure 8.7, call them k_1 left and right and k_2 top and bottom. What are the components of the f in $\vec{d} = f(\vec{F})$?

11 What are the dimensions of $I(\hat{x})$?

12 Starting at Eq. (8.38), drop the denominators $\sqrt{2}$ from the basis vectors, so that $\vec{e}_1 = \hat{x} + \hat{y}$ etc. Now compute the components of the same inertia tensor as done there.

13 Four point masses m are at the four corners of a square in the $z = 0$ plane and with $(x, y) = (a, a)$, $(a, -a)$, $(-a, -a)$, $(-a, a)$. Compute the components of the tensor of inertia with the basis $\hat{x}$, $\hat{y}$, $\hat{z}$.

14 For the same four masses as in the preceding exercise, compute all the components using the rotated basis $\hat{x}' = (\hat{x} + \hat{y})/\sqrt{2}$ and $\hat{y}' = (-\hat{x} + \hat{y})/\sqrt{2}$. $\hat{z}' = \hat{z}$.

15 In the preceding two exercises, you compared results in two bases, one of which is rotated $45°$ with respect to the other. What would the answer be if the basis is rotated by the angle $20.14°$ instead?

16 If a rigid body is both spinning and moving with no external forces, you expect its kinetic energy to stay constant. In Eq. (8.34) however the first *two* terms are constant, but the center of mass is moving, so the last term is always changing. What's wrong?

17 In the analysis of free rotation leading to Eq. (8.47), what happens if there is zero rotation about the z-axis?

18 A point mass m is at coordinates specified by the vector $\vec{r}$ and it is rotating about the axis $\vec{\omega}$. What is its velocity $\vec{v}$? What then is its angular momentum $\vec{r} \times m\vec{v}$? This is $\vec{L} = I(\vec{\omega})$, so what are the components of I? That is, compute $I(\hat{x})$, *etc.*

Problems

8.1 Explicitly write out the derivation leading to Eqs. (8.2) and (8.3) for the case of two masses.

8.2 Compute the volume of a solid hemisphere of radius R. (**a**) cylindrical coordinates. (**b**) spherical coordinates. (**c**) Find the center of mass of a solid hemisphere with uniform mass density.

8.3 Write out the proof in Eq. (8.4) for the special case of $4 = 2 + 2$ masses to verify the calculation.

8.4 Find the center of mass of the parabolic-shaped plane object in problem 2.54(b).

8.5 Suppose the rod in Figure 8.3 is not along the x-axis, but is at an angle α above it, still starting from the origin. Now calculate the angular momentum for the same $\vec{\omega}$.

8.6 Does $f(\vec{v}) = \vec{v} + \vec{c}$ define a tensor f, where $\vec{c}$ is a constant vector?
What about $f(\vec{v}) = \vec{c} \cdot \vec{v}$?
What about $f(\vec{v}) = \hat{z} \times \vec{v}$?
What about $f(\vec{v}) = c\vec{v}$?
What about $f(\vec{v}) = \hat{x}\,\hat{x} \cdot \vec{v}$?
What about $f(\vec{v}) = \hat{y}\,\hat{x} \cdot \vec{v}$?
What about $f(\vec{v}) = v^2\vec{v}$?
What about $f(\vec{v}) = \hat{v}$?
In each case, draw pictures that explain what the function does.

8.7 The force on a charge moving in a magnetic field depends on its velocity as $\vec{F} = q\vec{v} \times \vec{B}$. Show that this function of velocity, $f(\vec{v}) = \vec{v} \times \vec{B}$, defines a tensor and compute its components in terms of the components of $\vec{B}$. Write the result as a matrix. Pick a simple, special $\vec{v}$ and $\vec{B}$, to verify that your result works.

8.8 In two dimensions rotating vectors by a fixed angle α defines a tensor. Use the usual $\hat{x}$-$\hat{y}$ basis to compute the components of this tensor: $\vec{u} = f(\vec{v})$. Ans: in part, $f_{21} = +\sin\alpha$.

8.9 If an object is pivoted about some point and the only other forces on it are from a uniform gravitational field, show that the torque about the pivot is the same as if all the mass were concentrated at the center of mass. If the gravitational field is not uniform, show by a counterexample that this is false.

8.10 A meter stick is hanging, pivoted about its end, and you start to accelerate the support (the pivot) horizontally at a_0. Find the initial angular acceleration of the stick and find the initial linear acceleration of the other end of the stick. Ans: $\alpha = 3a_0/2\ell$

8.11 Show that the moment of inertia about *any* axis is $\vec{\omega} \cdot I(\vec{\omega})/\omega^2 = \int r_\perp^2 \, dm$, where the origin for the tensor of inertia is on the axis and $\vec{\omega}$ is in the direction along the axis. No components. Just use vector manipulation and the definition Eq. (8.7). And you will need a vector identity for the triple cross product.

8.12 Take the example in Figure 8.5 and repeat it in the language of tensor components in order to be certain that the results are the same.

8.13 Derive all the components in the equation (8.28).

8.14 For a uniform rigid rod rotating about an axis perpendicular to the rod, but passing through a point on the line with the rod, compute the rotational kinetic energy. The rotation axis can intersect the line of the rod anywhere. For fixed ω, find the position of the axis for which this kinetic energy is a minimum.

8.15 Derive the component $I_{xx} = \int dm\, y^2$ in Eq. (8.31) by direct integration.

8.16 For the functions in problem 8.6 that do define tensors, which ones are symmetric as defined in Eq. (8.39)? *I.e.* $\vec{v}_1^{\,*} \cdot f(\vec{v}_2) = \left(f(\vec{v}_1)\right)^* \cdot \vec{v}_2$.

8.17 For the functions in problem 8.6 that do define tensors, find their components with respect to some (clearly stated) basis of your choice. Ans: *e.g.* $\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

8.18 Start with the rectangular plate whose components were computed in Eq. (8.26). Move the plate up a distance c so that its corner is at the coordinates $(0, 0, c)$, though still parallel to the x-y plane. Now what are the components of its tensor of inertia?

8.19 Compute the components of the tensor of inertia of a thin, uniform spherical shell, mass m and radius R, about an axis through its center. *Note:* you must do this in your head! No complicated integrals allowed.

8.20 Compute the moment of inertia of a ball of uniform volume mass density about an axis through its center. Mass $= M$ and radius $= R$. Do this two ways.
(**a**) Straight-forward integration of $\int dm\, r_\perp^2 = \int dm\, (x^2 + y^2)$ in your choice of coordinates, probably cylindrical but maybe spherical or (least likely) rectangular.
(**b**) Compute $\int dm\, r^2$, where *this* r is the spherical coordinate. Then use this result to get the moment of inertia of part (a). Ans: $\frac{2}{5}MR^2$.

8.21 A uniform ball is rolling downhill without slipping and starting from rest. (**a**) Compute its kinetic energy as a function of the distance travelled by using the result of Eq. (8.34) with the origin at the center of the ball. (**b**) Compute its kinetic energy by noting that the point of contact is not moving, so that it is at any instant moving with pure rotation about that point. Find the moment of inertia about this new axis of rotation in order to get the kinetic energy with this origin. It will be nice if the answers agree.

8.22 In the example leading to Eq. (8.8), what torque is needed to maintain the angular velocity a constant? Ans: $\vec{\tau} = \frac{1}{24}ma^2\omega\hat{x} \sin 2\alpha$

8.23 At what point in the development from Eq. (8.14) to (8.19) did I use the fact that $\hat{x}$, $\hat{y}$, *etc.* are orthonormal?

8.24 There is nothing in this chapter that required you to multiply two matrices, though I hope you've seen the method before. *Where* did that rule for multiplication come from? It follows from the definition of components as summarized in Eq. (8.20). Let f and g be tensors so that their components are defined as $f(\vec{e}_i) = \sum_j f_{ji}\vec{e}_j$ and $g(\vec{e}_i) = \sum_j g_{ji}\vec{e}_j$. The composition of two functions is defined: $h = f \circ g$ means $h(\vec{v}) = f(g(\vec{v}))$. (This is just like Eq. (0.32), though here you're not doing calculus.) (**a**) Show that h is linear, satisfying Eqs. (8.9). (**b**) Compute $h(\vec{e}_i)$ using the equations (8.20) and (8.9) to express h_{mn} in terms of the components of f and of g. Ans: $h_{mn} = \sum_k f_{mk}g_{kn}$

8.25 Starting from the definition Eq. (8.20), find the components of the function in Figure Fig. 8.16. Assume that the z-axis is out of the page.

8.26 Two equal masses are placed at coordinates $(a, 0, a)$ and $(0, a, a)$. Find the components of the tensor of inertia about the origin in terms of the basis $\hat{x}$, $\hat{y}$, $\hat{z}$ in this coordinate system.

8.27 Derive the angle between $\vec{L}$ and $\vec{\omega}$ for the spinning disk, Eq. (8.32). The picture there implies that $\vec{L}$ is closer to the z-axis than $\vec{\omega}$. Show that this equation agrees with that statement.

8.28 (a) Find the components of the inertia tensor for a triangular plate as shown in the sketch. The plate is in the x-y plane. (b) Find the eigenvectors and eigenvalues of the tensor. Ans: (a) $I_{zz} = \frac{4}{12}Ma^2$ (b) eigenvalues: $\frac{1}{12}Ma^2 \cdot [1, 3, 4]$.
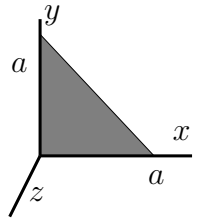
8.29 (a) Evaluate the components of the tensor of inertia of a uniform cube of mass m and side ℓ. The origin is at one corner. (b) What is $\vec{L}$ if $\vec{\omega}$ is through the origin and along the main diagonal of the cube? (c) What is the moment of inertia of the cube about the main diagonal? Ans: (c) $ma^2/6$

8.30 Find the components of the inertia tensor for a thin rod of mass m and length ℓ. One end is at the origin and the other is at the point (ℓ, θ, ϕ) of the spherical coordinate system. Ans: $I_{xx} = \frac{1}{3}m\ell^2(\sin^2\theta\sin^2\phi + \cos^2\theta)$

8.31 A solid, uniform, rectangular block of mass m has sides a, b, c. The moment of inertia about the long axis from one corner to its opposite is (perhaps) $(m/6)(a^2b^2 + b^2c^2 + c^2a^2)/(a^2 + b^2 + c^2)$. Analyze this proposed answer, trying to show that it is wrong. If you fail, then maybe it's right.
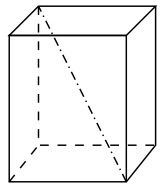
8.32 Derive the expression stated in the preceding problem.

8.33 Generalize the parallel axis theorem to find a relationship between two inertia tensors, neither of which is about the center of mass. This theorem is not so useful.

8.34 Apply the equation (8.40) to all the combinations of basis vectors $\hat{x}$, $\hat{y}$, and $\hat{z}$ to determine what it says about the components of the inertia tensor in this basis. Do this without using the computed form of the tensor in terms of m's and $\vec{r}$'s. Just Eq. (8.40) itself.

8.35 In previous chapters you have sometimes started with energy instead of force equations, knowing that you can differentiate the former to get the latter. There is an easy, elementary derivation of $\tau = I\alpha$ in the same spirit. You know that $K = \frac{1}{2}I\omega^2$ for rotation in the special case of a symmetric body spinning about a fixed axis. Attach it to an axle through its center and wrap a string around it. Now pull with constant force F for a distance x, and as a function of this distance pulled, get the work done. This work is the rotational kinetic energy: differentiate this equation with respect to time. Ans: $\tau = I\alpha$
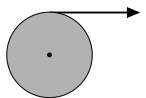
8.36 This figure is a plane sheet of metal of mass M and dimensions as shown. A coordinate system is set up in the lower left corner. (a) Compute the components of the tensor of inertia in this coordinate system. Find the eigenvectors of this tensor. (b) By how much would the coordinate system as given have to be rotated in order that the matrix of components be in diagonal form? You will of course check your result in all the special cases. Ans: $I_{zz} = \frac{2}{3}M[b^3 + ba^2 - a^3]/(2b - a)$
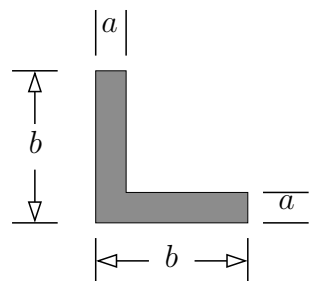
8.37 Parallel to the result in Eq. (8.40), take the tensor f, where $f(\vec{v}) = \vec{v} \times \vec{B}$ for fixed $\vec{B}$, only there are no integrals this time. What is the relation between $\vec{v}_1^* \cdot f(\vec{v}_2)$ and

$f(\vec{v}_1^*) \cdot \vec{v}_2$? Repeat problem 8.34 for this case to determine what it says about the components of this tensor.

8.38 ¨ Toss a disk into the air, so that it is horizontal and is spinning about its axis as it flies. Not exactly *on* axis though because unless you toss it perfectly you will see that it wobbles. Find the rate of wobbling and compare it to the rate of rotation of the disk itself. First, what is the relationship between ω_z and ω'? Which rotates faster, the plate or the wobble? That is, *interpret* the equations and describe exactly what you will see when you watch such a thin plate tossed into the air and wobbling about its axis. Try it; make sure that there is some sort of pattern imprinted on the disk so that you can see its rotation.

8.39 A thin, uniform rectangular plate has mass m, sides $2a$ and $2b$. Take the origin at the center of the plate.
(**a**) Find the components of the inertia tensor of the plate.
(**b**) If the rectangular plate is rotating about a diagonal with an angular speed ω find the angular momentum and the kinetic energy of the system.

8.40 ˙ Three equal masses m are placed at coordinates $(0, a, a)$, $(a, 0, a)$, and $(a, a, 0)$.
(**a**) Find the components of the tensor of inertia about the origin in terms of the basis $\hat{x}$, $\hat{y}$, $\hat{z}$ in this coordinate system.
(**b**) If the system is rotating as a rigid body about the z-axis through the origin, what is the angular momentum? Draw a picture of the vectors $\vec{\omega}$ and $\vec{L}$ in this system. If it's hard to do the drawing, then add some *words* to say what the directions are. In fact, put some words in anyway.
(**c**) Find any one eigenvector of I, and its corresponding eigenvalue. You may either derive it by computing it or you may guess an answer and demonstrate that it is correct.
(**c**) Having found one eigenvector and eigenvalue, find the other two. At worst, you have started with a cubic equation and know one root. Now factor it and then get the others. Or, you can try to find a more clever way.

8.41 ˙ A solid cylinder has mass m, radius R, and moment of inertia I. It is set spinning clockwise about its axis with an angular speed ω_0 and placed on the floor next to a wall as shown. The coefficient of kinetic friction between the floor and the cylinder is μ_1 and between the wall and the cylinder it is μ_2. The angular deceleration of the cylinder is claimed to be $\alpha = mgR\mu_1(1+\mu_2)/((1+\mu_1\mu_2)I) = 2\mu_1 g(1+\mu_2)/((1+\mu_1\mu_2)R)$. Is this plausible?
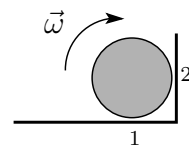
8.42 ˙ Derive the angular deceleration as stated in the preceding problem. Perhaps reread section 1.4 ?

8.43 ¨ A plane homogeneous plate of surface mass density σ is bounded by the logarithmic spiral $r = ke^{\alpha\phi}$ and the radii $\phi = 0$ and $\phi = \pi$. (**a**) Find the components of the inertia tensor about the origin. The plate is in the x-y plane. (**b**) Find the eigenvalues and eigenvectors of this tensor. Any special cases to check the validity of your solution?

8.44 Eight point masses m are placed at the eight corners of a cube of side a. The origin is at one corner of the cube and the x, y, and z axes run along edges of the cube. (**a**) Compute the components of the tensor of inertia in this coordinate system. (**b**) What is the moment of inertia about the long diagonal of the cube?

8.45 Find all eigenvectors and eigenvalues of a tensor with components $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Apply this result to the result of problem 8.7. Ans: $\hat{x} \pm i\hat{y}$, $\hat{z}$

8.46 Generalize problem 8.14 to an arbitrary body rotating about an arbitrary axis. Among the set of all axes parallel to some given one, derive the axis that presents the smallest kinetic energy for a fixed $\vec{\omega}$. Use the parallel axis theorem and Eq. (8.33). When you want to simplify this, *draw pictures.*

8.47 In Fig. 8.13 and assuming that the disk is thin, picture yourself as the person holding on to the z-axis. Where will you have to look to see the angular momentum vector? With respect to you, what is the motion of $\vec{\omega}$ and of $\vec{L}$? What are these motions with respect to a pedestrian?

8.48 The ice in Antarctica is up to a few kilometers thick in spots, averaging about 1.6 km. If it melts and spreads out over the surface of the Earth, estimate its effect on the number of seconds in a day and the cumulative effect for a year. Also, what would it do to sea level? Don't plug in any numbers until you've simplified all the algebra and so that you aren't subtracting two large numbers to get a small number.
Ans: $\frac{5}{3}T m/M \approx .5$ s or so per day. This answer assumes the Earth is a uniform ball. It isn't, so in which direction does that fact change the answer? Bigger or smaller?

8.49 The Sun rotates once in a little less that one Earth month. If the Sun collapses to a radius of 10 km while conserving angular momentum, what would its rotation period be? This is about the size of a neutron star, the remnant of an explosive stellar collapse. What would the speed of a point on its equator be? What would the collapsed star's density be? And of course not just a number. Find something to compare it to.

8.50 Compute the components of the inertia tensor for a solid cone of uniform mass density. The origin is the vertex. Mass: M, height: h, radius of base, R. What is its kinetic energy if it is rotating about its axis of symmetry with angular speed ω?
Ans: Diagonal: $I_{zz} = \frac{3}{10}MR^2$, $I_{xx} = I_{yy} = \frac{3}{5}Mh^2 + \frac{3}{20}MR^2$, $\quad K = \frac{3}{20}MR^2\omega^2$

8.51 The cone in the preceding problem is lying on a table and free to roll without slipping. Its vertex stays fixed, and the line of contact with the table is instantaneously at rest, so that defines the axis of rotation. What is its kinetic energy, expressed in terms of the given information and ω?
Ans: $\frac{1}{2}M\omega^2\left[\frac{3}{10}R^2 + \left(\frac{3}{5}h^2 + \frac{3}{20}R^2\right)R/h\right]/(1 + R^2/h^2)$

8.52 Use Eq. (8.34) to find the total kinetic energy of a solid cylinder rolling downhill without slipping. (**a**) Do this using an origin at the center of the cylinder. (**b**) Do this using an origin at the (moving) point of contact of the cylinder with the surface.

8.53 The equation (8.50)(1) usually has no solution. Pick one of the eigenvectors of I_0 and its corresponding eigenvalue, perhaps $\omega_0\left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$, or another. Plug this $\vec{\omega}_0$ into the right side of (8.50)(1) and see what happens when you try to solve for $\vec{\omega}_1$. Knowing that somehow there must be an answer, what does this tell you?

8.54 Derive Eq. (8.54) *without* using Eq. (8.52). Show that you wind up dividing by zero.

8.55 I recently installed a new bathroom sink, but it was installed incorrectly, so objects (including water) rolled quickly off it and onto the floor. To measure the angle, I used a cylindrical jar of strawberry preserves (height $h = 4.0''$, diameter $d = 2\frac{5}{8}''$). The distance from the back wall to the edge is $L = 21\frac{3}{8}''$. Placing the jar against the back wall and releasing it, the time it took to roll to the far edge was about 2.0 seconds. What was the angle at which the sink was tilted? Estimating the error range in the time to be $\pm 10\%$, what is the error range in the angle? Ans: $2.0 \pm 0.4°$ (It got fixed.)

8.56 (**a**) Repeat the calculations of Eqs. (8.52)-(8.54) for one of the other two eigenvectors, getting both the correction to the eigenvalue and for the eigenvector. (**b**) Take the scalar product of the $\vec{\omega}_0 + \vec{\omega}_1$

that you find with the $\vec{\omega}_0 + \vec{\omega}_1$ of the example, the one that started with $\omega_0 \hat{x}$. This dot product should be zero, at least to the order that you've calculated the vectors. That is, the product should be zero out to order m/M, but don't expect it to be any better than that unless you decide to compute the second order corrections and to include them.

8.57 The Earth has a bulge at the equator so that its diameter is about 41 km larger than it is from pole to pole. Model the Earth's tensor of inertia by estimating this bulge as a fat cylinder of thickness 20.5 km and height perhaps $1/3$ of Earth's diameter. Add this to the rest of the Earth, assumed to be a sphere, to get the total tensor of inertia. The Earth's rotational axis does not align perfectly with the polar axis, so this will cause a wobble. Compute it, and see how this simple model compares to the experimental value of about a 400 day period. The results will of course depend on your choice of density for this bulge.

8.58 Use the box example of Figure 8.15 (but without the extra mass) and force it to spin with angular speed ω about its long axis. What torque does it take to do this?

8.59 A right triangle has sides a, b, and c, and it has a uniform area mass density. What is the moment of inertia about an axis perpendicular to the plane and through its center of mass. I suggest that you start by computing its moment about an axis through a corner first. Ans: $\frac{1}{36}m(a^2 + b^2 + c^2)$

Special Relativity

.

The history of special relativity is fascinating: filled with partial answers, clever ideas, false starts, intellectual barriers, and some smart people who, given enough additional time, would have figured it all out. Einstein got there first. He cut through the complexities that had slowed or stalled other work and he found the underlying simplicity. That is where I will start.

Galileo was really the first to articulate one of the foundational concepts, and he did it three and a half centuries before Einstein was born. In his *Dialog on the Two Chief World Systems* he made the idea concrete by describing what you would see if you are in the main cabin below decks on some large ship, first when it is standing still, and then when you

> ...have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still.*

Put in modern terms,

1. Without looking outside the room, you can't tell how fast you're moving.
or 2. Only relative velocities are measurable, not absolute velocities.
or 3. The Results of Measurement depend only upon the relative velocity of observer and observed and not upon their absolute velocities.
   (And the third one is also Galileo's phrasing — surprisingly modern.)

Chapter five started with the Galilean transformation, which is the mathematical codification of the change to a moving coordinate system.

$$x' = x - v_0 t, \qquad y' = y, \qquad z' = z, \qquad t' = t \qquad\qquad \text{Eq. (5.1)}$$

In that description, as long as forces depend on at most the *relative* velocity of interacting objects, Newton's equations remain the same when switching to a moving observer. $\vec{F} = d\vec{p}/dt$ no matter how fast you move. The problem is that Newtonian Mechanics is not all. Maxwell found a set of equations that describe electromagnetism extraordinarily well, and they do *not* transform properly under the Galilean transformation. That means that something has to give: Newton, Galileo, or Maxwell.

When these difficulties arose in the late 1800's the natural response was to go with the familiar, keeping Newton and Galileo, but modifying Maxwell. The most famous approach was to say that Maxwell's electromagnetic equations are valid in only *one* coordinate system and no other: the "ether" theory.

That didn't work, and Einstein's approach kept Galileo and Maxwell, while forcing Newton to change. It all sounds so simple doesn't it? When one choice doesn't work, try another. This caricature of the history as I've sketched it in these few paragraphs has filled, and will further fill, whole books putting flesh on the details.

What is it about Maxwell's theory that caused a problem? The simple and most profound fact is that it produces a unique value for the speed of light in a vacuum. Not speed relative to something, just speed. The simplicity of the theory of special relativity is that you start from two assumptions:

---

* Translation: Stillman Drake

1. Only relative velocities are measurable
2. The speed of light in a vacuum is the same for everyone.

This assumes that Maxwell is right and that Galileo's verbal statement of the principle of relativity is right, but *not* the transformation equations written above. In particular the fourth one, $t' = t$, will be sacrificed, and that is the most startling and perplexing aspect of this theory. The concepts "simultaneity", or "at the same time", or "now" will become puzzles to be sorted out. In Galileo's day, or Maxwell's, and all the way up to Einstein's, this equation stating that time is the same for everyone was so obvious that it wasn't even written. It was just wrong.

What will replace the Galilean transformation of Eq. (5.1)? The answer is

$$x' = \frac{x - v_0 t}{\sqrt{1 - v_0^2/c^2}}, \qquad y' = y, \qquad z' = z, \qquad t' = \frac{t - v_0 x/c^2}{\sqrt{1 - v_0^2/c^2}}$$

where c is the speed of light in a vacuum. These are called the Lorentz transformation, but developing them will take some effort, concluding at Eq. (9.10). Notice that if $c \to \infty$, these reduce to the Galilean transformation.

A matter of typography: Most of the time I have tried to be careful about how I write the mathematics, sometimes probably to the point of being annoying. In this chapter there are so many velocities that I won't bother with the distinctions such as those between v (speed) and v_x (velocity component). I will leave it to you to figure out which is which from context, otherwise there would be too many sub-x's filling the pages, and by now you should be able to handle such an abuse of notation.

9.1 Time Dilation, Length Contraction

The first step is to derive an immediate implication of the assumptions of special relativity: that moving clocks run slow, which implies that measurements of time must be handled very carefully. There is a simple way to demonstrate this, one that is fairly standard in texts on relativity. It involves using a special kind of clock to analyze the effects of velocity. Because light in a vacuum is at the center of one of the assumptions of the theory, this clock is built precisely from light in a vacuum. Build an evacuated box with mirrors at the top and bottom and a source of light inside.

$L_0 = cT_0$

your view

cT

vT

your friend's view
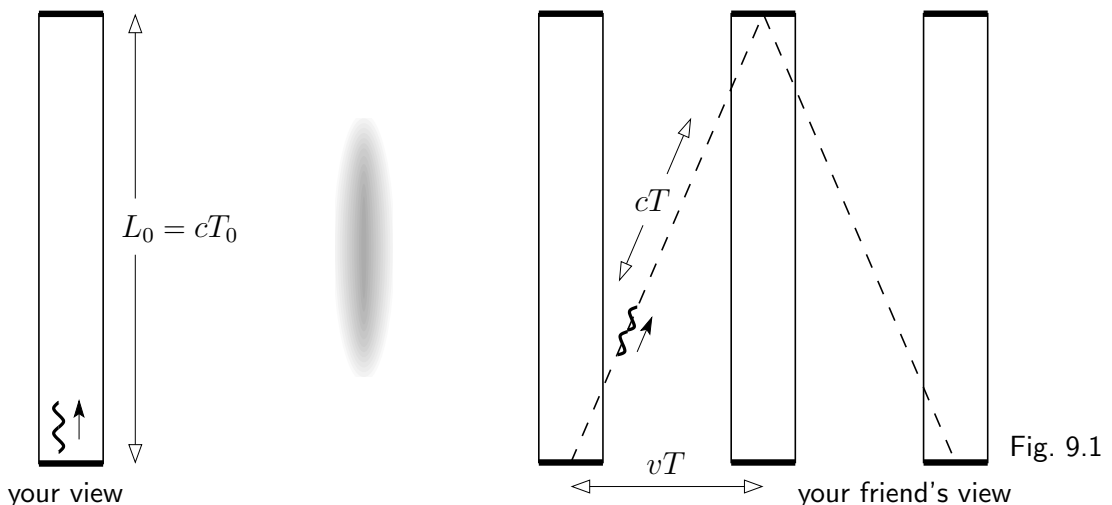
Fig. 9.1

In the sketch on the left, the light source starts a pulse of light moving straight up the tube until it hits the (perfect) mirror at the top. It bounces (straight) back down and hits the bottom (also perfect). The light keeps bouncing up and down, and that bouncing is the ticking of the clock. Any real clock will have imperfections of course, but the analysis of this idealized version is where to start.

What is the time between ticks? The length is L_0 and the speed of light is $c = 299792458\,\text{m/s}$, so $T_0 = L_0/c$ is the time for the light to go bottom to top or back.

As you observe the light pulse repeatedly hit the bottom and the top mirrors, a friend of yours is driving toward the left at (very) high speed v. He has no doubt that the light is hitting the mirrors, but says that the clock is moving right at high speed, so that during the time that the light pulse went from the bottom to the top, the clock itself has travelled some distance. If you say that he is moving left then he will say that the clock is moving right. That in turn implies that the light pulse will, according to your friend, be travelling along a diagonal path up and another diagonal path down. The second drawing shows three images of the clock as it moves right. When the light left the bottom mirror, when it hit the top one, and finally when it returns to the bottom.

If this were a discussion of Newtonian mechanics, and Maxwell hadn't been born yet, this would be easy to analyze. In the first picture the light pulse is moving at speed c. In the second picture, the clock *and the light* are moving right with a horizontal component of velocity v. The light has a vertical component of velocity c, so the total speed of the light according to your friend is $\sqrt{c^2 + v^2}$. All very simple, but that's not the way the universe is built. Instead, the speed of light is the same for everyone, and that is taken to be one of the fundamental assumptions of special relativity: c, not $\sqrt{c^2 + v^2}$.

What changes is the time measurements themselves. The previous paragraph tacitly assumed that both people measured velocities by the same methods and that the clocks each used to measure the time-of-flight of the light behaved the same way. That assumption has to go. In the original calculation of the ticking of the clock, $T_0 = L_0/c$. For the other person looking at the clock you cannot assume that the time to tick is the same; call it T instead (to be determined).

In the time of one tick, the clock moves a distance vT horizontally, *and the light moves a distance cT* along the diagonal. Remember the second assumption of the theory: light moves at the speed of light. There's a right triangle in this picture, and you can apply the Pythagorean Theorem to solve for T.
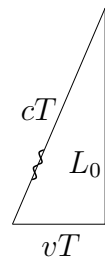
$$(vT)^2 + L_0^2 = (cT)^2 \quad \longrightarrow \quad (vT)^2 + (cT_0)^2 = (cT)^2$$

$$\longrightarrow \quad c^2 T_0^2 = (c^2 - v^2)T^2 \quad \longrightarrow \quad T^2 = \frac{c^2 T_0^2}{c^2 - v^2}$$

$$\longrightarrow \quad T = \frac{T_0}{\sqrt{1 - v^2/c^2}} = \gamma T_0 \tag{9.1}$$

and γ is a standard abbreviation for $1/\sqrt{1 - v^2/c^2}$.

This is time dilation. The moving clock takes longer to tick. The moving clock runs slow. Wait a minute. Isn't it your friend who's moving? Not the clock? Remember the *first* assumption of the theory: Galileo's idea that only relative velocities are measurable. That means that everyone is entitled to claim "I am standing still." "I am not moving; it is the rest of the universe that is moving." Everyone is right, including your friend, who says that the clock is moving right at speed v.

This is the first indication that these two innocuous-sounding assumptions can lead to surprising and puzzling consequences. When time has become slippery can space be far behind?

Point of Confusion

If your friend's clock is running slow by a factor 2, you see that it takes a longer-than-expected time between his ticks. If you see two events that occur $10\,\text{seconds}$ apart, does he say that they occurred $5\,\text{s}$ apart or $20\,\text{s}$? If these are the only two choices, then the answer is either or neither, depending on *where* the two events took place. See problem 9.9. After the whole Lorentz Transformation is available in a few more pages, then I hope this question and its mysterious sounding answer will be clearer.

Lengths

If clocks run slow, making time measurements complicated, then what about lengths? If you have

stopped at a train crossing and a long train passes in front of you, is the length of the train according to you the same as the length of the train according to a passenger on the train? No. To measure the length of the train, measure the time it takes to move past and multiply it by the speed: $L = vT_0$. (T_0 is the time that *you* read on *your* clock.) The question is now to find the length as measured by a passenger. Call the length and the time that the passenger measures L_0 and T instead of L and T. (L_0 is the length the passenger observes for the [to him, stationary] train.) The passenger doesn't need another clock to measure T, all that's needed is to be able to read *your* clock. If yours recorded T_0, and according to the passenger you are moving at speed v, then (again according to the passenger) your clock is running slow. The passenger then says that the time T for you to have moved from one end of the train to the other is greater than your measured T_0. The calibration factor between the clocks is $\sqrt{1 - v^2/c^2}$, so the passenger says that the time in which you go from one end of the train to the other is longer than the T_0 that you read on your slow clock. It is $T_0/\sqrt{1 - v^2/c^2}$. This is the passenger's T and the length as measured by the passenger is $L_0 = vT$. (Reread the above Point of Confusion.)

your view   passenger's view
Fig. 9.2
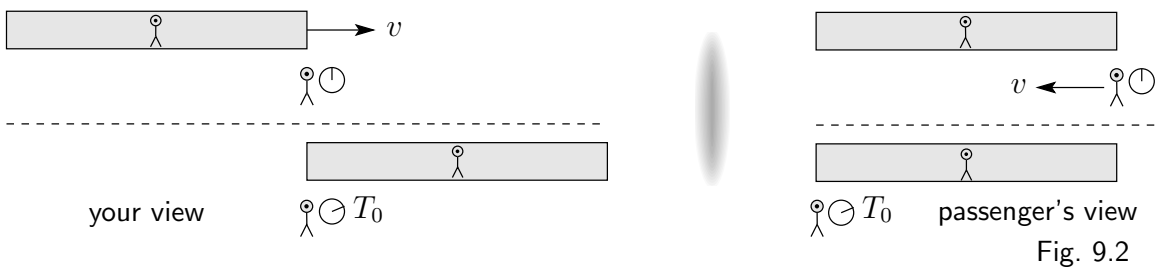
$$\text{you:} \quad L = vT_0$$

$$\text{passenger:} \quad L_0 = vT = vT_0/\sqrt{1 - v^2/c^2} = L/\sqrt{1 - v^2/c^2}$$

$$\text{length change:} \quad L = L_0\sqrt{1 - v^2/c^2} = L_0/\gamma \tag{9.2}$$

You see the train as shorter than the passenger does. This length contraction is a second peculiar feature of special relativity.

This derivation of length contraction (also called the Lorentz contraction) involves only the simplest algebra, but it requires careful attention to the details of who's looking at what and when. There is another way to get to this result, one that reverses these two difficulties. There the concepts are easy, but the algebra is more involved. The idea is simply to take the light clock of the previous two pages and turn it on its side, so that the light is going left-to-right and right-to-left as the clock moves right. It should read the same time as it did when it was upright, and that will let you figure out what its length is. See problem 9.6 for some suggestions about how to do this calculation (or better, you can try it on your own first and find the stumbling blocks yourself).

It is easy to mix up the applications of these dilation and contraction equations, so before applying either of Eqs. (9.1) or (9.2) you should think about the *qualitative* behavior of the time and distance measurements. If a clock takes two seconds to mark off a second, its owner will think that someone just did the 100 meter dash in five seconds. This means that you can't simply plug into formulas. The qualitative analysis should come first.

> A moving clock runs slow.
> The calibration factor between clocks is $\gamma = 1/\sqrt{1 - v^2/c^2}$.
> That factor is greater than one.
> So, does the factor go in the numerator or in the denominator?

Terminology: *Proper length* is the length of an object as measured by someone who says it is not moving. *Proper time* is the time interval between two events as measured by someone who says the occurred at the same place*—the clock is standing still. For the train example, L_0 is the proper length of the train, and T_0 is the proper time on my clock. There is no universal convention about how you place subscripts here, so don't assume that a sub-0 will always be proper. For *proper acceleration* see equation (9.23). [What would "proper velocity" be?]

If length contracts, shouldn't the L_0 in the original clock, Figure 9.1, contract? Should the L_0 in Eq. (9.1) be a variable too? The answer is No, but it is not only a legitimate question, it's an important one. In that derivation I assumed that both you and your friend used the same value for that number, the length perpendicular to the motion. Is that valid? The answer is Yes, but now you have to test it to be certain.

The proof is by contradiction. Assume the statement is false and show that it doesn't work. Take the light clock, supposedly length L_0. Your friend wants to measure it to see if it remains L_0 even when it moves past, so he uses a ruler of length L_0 and holds it up so the ruler and the clock just miss each other. To be certain of his measurement he attaches a pair of paint brushes to the two tips of the ruler, set so that they will mark the ends of the clock as it passes. If the clock shrinks because of its motion then the paint brush on one end of the ruler will mark the clock while the other end misses. The trouble is that paint leaves a permanent mark on the clock. You and your friend can get together over a beer after the experiment in order to compare data. If your clock has only one paint mark, then there is a lateral length contraction and *you were moving*. If it has two paint marks that are *not* at the ends, then he was moving. Either way you have violated the first axiom of the theory, that only relative velocities can be measured. What if lateral lengths expanded instead of contracting? You get into exactly the same contradiction, but interchanging who is moving.

Another way to picture this: If a railroad train is entering a tunnel through a mountain, and if the tunnel is just barely large enough to accommodate the width of the train when the train is moving slowly, what will happen when the train is moving rapidly? If there is *either* a lateral contraction or expansion because of the speed then someone, either the train engineer or the person who built the tunnel, will think that the train will still fit while the other will expect a crash. That is a qualitative difference, so they can't both be right.

Paradoxes

Under normal circumstances, *i.e.* before you started to study relativity, you may have an occasion to think that your watch keeps good time and that your friend's is running slow. Your friend, who thinks just as highly of his timepiece, would think that your watch is running fast. The calculation of time dilation leading to Eq. (9.1) says that your friend's clock runs slow because it's moving. If the first assumption of relativity is correct, that it's impossible to tell which of you is moving, then both of you will believe that the other's clock is running slow—and both be right. How can that happen?

The same paradox applies to lengths. If the moving train contracts, a passenger on that train can look out and say that another train you think to be standing still has shrunk. How can both be true? Resolving both of these problems will have to wait a few pages until the Lorentz transformation appears.

If you drive a long car and have a short garage, can you fit your car within the garage by driving extra fast? Someone standing outside and watching you try to do this would say yes; your car has shrunk, so it now fits inside. You, as the driver, see the garage approaching, only it is now even shorter than it was before. You're never going to fit inside now. Who's right?

If two people are approaching each other, one at $\frac{3}{4}c$ from the left and the other at $\frac{3}{4}c$ from the right, and you ask one of them how fast the other is moving, the answer will not be $\frac{3}{2}c$, but $\frac{24}{25}c$. A

---

* There is an obscure word for this: *collocal* is the spatial analog of simultaneous.

derivation of this will appear in Eq. (9.13).

Example

▶ Before going into some more analytical problems with relativity, a few elementary examples are useful. Cosmic rays are (mostly) very high energy protons that come from outer space and hit the Earth's atmosphere. When such an energetic proton collides with an oxygen or nitrogen nucleus it simply blasts the nucleus apart and in the process creates many thousands or millions of other new particles. A major component of the resulting debris is the muon. This particle is basically a massive, but unstable, electron; $m_\mu = 207 m_e$, and the mean lifetime of these muons is 2.2 microseconds. In this short time, even moving at nearly the speed of light, the mean distance such a particle goes before decaying is about

$$c\tau = (3 \times 10^8 \text{ m/s})(2.2\,\mu s) = (300 \text{ m}/\mu s)(2.2\,\mu s) = 660 \text{ m} = 0.66 \text{ km}$$

What does "mean life" mean? Radioactive decay follows statistical laws, and starting with a population of particles at time zero, the fraction of them left at time t is $e^{-t/\tau}$, or expressed in terms of the total population, $N(t) = N_0 e^{-t/\tau}$. To find the average value of the decay time, let dN be the change in this number in time dt, so that $-dN$ is the number that decayed between time t and $t + dt$. The mean time to decay is then a sum over all these times divided by the total number of particles:
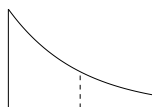
$$\frac{1}{N_0} \int t\,(-dN) = \frac{1}{N_0} \int_0^\infty -t\frac{dN}{dt}\,dt = \int_0^\infty \frac{1}{\tau} t e^{-t/\tau}\,dt = \tau$$

Starting from an altitude of 20 km and moving straight down, the fraction of these particles that would reach the Earth's surface is $e^{-20/0.66} \approx 10^{-13}$. Despite this, so many reach the surface that you will have a thousand or more of them passing through you each minute. This is not because there are so many cosmic rays that hit the Earth, but because of time dilation. The muons are moving so fast that the time dilation effect is large, and their lifetime is far longer than the two microseconds that they have when at rest. A factor of $1/\sqrt{1 - v^2/c^2} = 10$ is typical, and you might guess that such a factor would lead to a change in the number reaching the surface by about a factor of 10. Not so. This exponential changes the fraction hitting the surface from

$$e^{-ct/c\tau} = e^{-20\text{km}/0.66\text{km}} = 7 \times 10^{-14} \qquad \text{to} \qquad e^{-20/6.6} = 0.05 \tag{9.3}$$

That is a factor of about 10^{12}, and this is a major part of the background radiation that you live with all your life.

I used the "mean lifetime" to describe this and you may be more familiar with the "half-life". They are proportional. See problem 9.3 for an explanation of the relation between the two terms: $t_{1/2} = 0.693\,\tau$. Also, not all muons will have the same speed, so they won't all have the same time dilation factor. How do you handle that? See problem 9.8.

Example

▶ The Global Positioning System is so commonly used that it is standard equipment in many cars, and is available for personal use almost everywhere. It depends on precisely measuring the signals from several satellites in orbit at a radius of 26600 km (measured from the Earth's center). The orbital period is about half a day; its orbital speed is about 3.8 km/s. At this speed the time dilation effect is

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \approx 1 + \frac{v^2}{2c^2}, \qquad \text{with} \qquad \frac{v^2}{2c^2} = \frac{1}{2}\left(\frac{3.8}{300\,000}\right)^2 = 8 \times 10^{-11}$$

In one day $= 86\,400$ s, the clock error would be $86\,400$ s $\times\, 8 \times 10^{-11} = 7\mu$s. At the speed of light, the distance error from this is then 7μs $\cdot\, 300\,000$ km/s $= 2$ km. If no correction had been made for this effect the GPS would have been worthless. The effects from general relativity have a sign opposite from this but are even larger, giving a combined drift that, if uncorrected for these relativistic effects, would be about 10 km/day.

9.2 Space-Time Diagrams

As with almost everything else in life, and by that of course I mean physics, the ability to draw a sketch of your ideas is invaluable. Sometimes it will be a drawing and sometimes it is a graph. Here, the goal is to turn conceptually difficult questions about measurements with clocks and rulers into conceptually simpler problems in analytic geometry. When describing the motion of a particle, you can give its position as a function of time and graph that. For now, one dimension of space (x) will be enough; the y and z coordinates will come along later.

Dealing with objects that move near or at the speed of light makes it necessary to choose convenient units. The speed of light is 3×10^8 m/s $= 300$ m/μs $= 1$ ft/ns. The trick to make this easy is to measure time in microseconds and measure space in units of 300 meters. In other words, pick the units so that c has the numerical value one. If you prefer light-years and years or feet and nanoseconds, that's o.k. too. Any of these puts the the equation $x = ct$ at $45°$. Choose the time axis up* and the x-axis left and right, then the pictures will look like these:

Fig. 9.3

stationary   simultaneous $x = ct$ $x = vt$, $v < c$
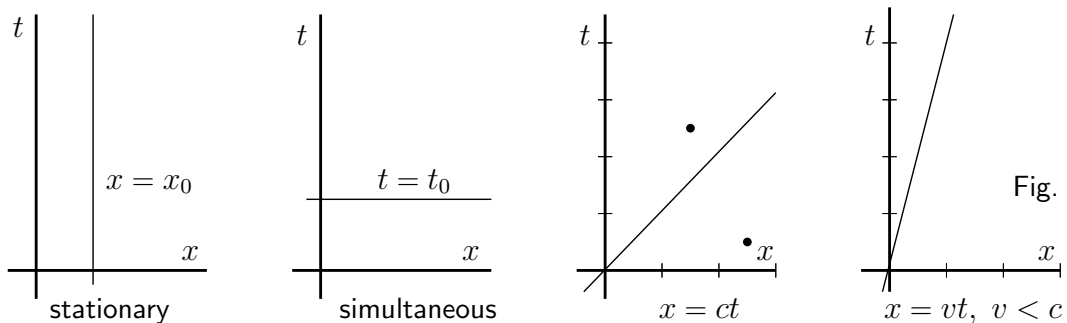
The second graph represents a set of events that occur at the same time, the equation is $t = t_0$.
The first and fourth graphs present constant velocity motion of a particle, one of them with $v = 0$ and the other with $v = 300$ m/4μs $= c/4$.
The third is a graph of the motion of a photon, $v = c$.
The t and x coordinate axes have equations that are respectively $x = 0$ and $t = 0$.
The two points in the third graph represent events at $(x, t) = (2.5, 0.5)$ and $(1.5, 2.5)$ in the units of 300 m, $1\,\mu$s.

But First, an Aside:

When dealing with ordinary two-dimensional rectangular coordinate x and y, sometimes you need to switch to a rotated coordinate system. If you're doing an elementary mechanics problem and $\vec{g}$ is down, you may choose x horizontal and y vertical, or you may not. Perhaps some other aspect of the problem, maybe a hill, suggests a different coordinate system with x' along the incline of the hill and y' perpendicular to it. What is the relationship between these coordinates? To find out, just express each pair of coordinates in terms of r and the angles. The angle the hill makes with the horizontal is α, and the single point P has coordinates (x, y) in one system and (x', y') in the other. The polar coordinate angle in the x-y system is ϕ, and that gives the first pair of equations. For the second pair

---

* Why is time up and space left and right? For the same reason that people in the U.S. drive on the right and those in the U.K. drive on the left. That's the custom. Besides, if the coordinates are x, y, and t, then laying out the x-y plane sort-of horizontally feels better; see page 310.

of equations in the x'-y' systen, use the polar coordinate angle $\phi - \alpha$.

$$x = r \cos \phi$$
$$y = r \sin \phi$$

$$x' = r \cos(\phi - \alpha)$$
$$y' = r \sin(\phi - \alpha)$$
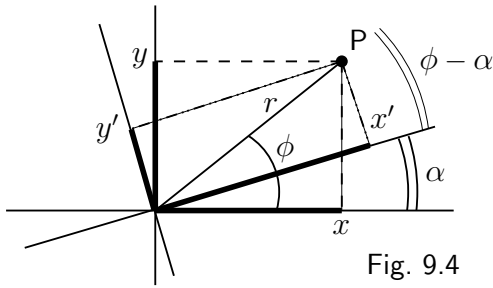
(9.4)

Fig. 9.4

$$\Longrightarrow \quad \begin{aligned} x' &= r \cos \phi \cos \alpha + r \sin \phi \sin \alpha = \quad x \cos \alpha + y \sin \alpha \\ y' &= r \sin \phi \cos \alpha - r \cos \phi \sin \alpha = -x \sin \alpha + y \cos \alpha \end{aligned}$$

(9.5)

These equations show how to relate the coordinates that two different people use to describe the same point, and all that was needed was a couple of standard trig identities. [Look at the exercises at the end of this chapter, page 329.]

A simple question: What are the equations for the various axes? A simple answer: the x-axis equation is $y = 0$ and the y-axis equation is $x = 0$. Similarly, the x'- and y'-axes have equations $y' = 0$ and $x' = 0$. If this seems trivial, just wait until Figure 9.7.

The reason for looking at these rotated coordinate systems is that something very much like this can be done with space-time coordinate systems. The transformations will not be rotations, but they will have many properties analogous to them. For space and time, the key step is to realize that you have to provide a precise statement of what these coordinates *mean.* You've already seen that times and lengths are no longer such simple concepts, so the coordinates that build on them have to be defined with care.

One method to define these coordinates uses tools that are conceptually very simple: a clock and a radar set. The radar set sends out electromagnetic pulses that can bounce from a distance object, and a clock times the reflected signal, telling how far away the object is.

Let t_1 be the time the radar pulse is sent and t_2 the time it returns. The total travel time of the pulse is $t_2 - t_1$, half in each direction, so the time at which the pulse hits the object is

$$t_1 + \frac{1}{2}(t_2 - t_1) = \tfrac{1}{2}(t_1 + t_2), \qquad \text{and the distance to it is} \qquad \frac{c}{2}(t_2 - t_1)$$

Translate this into the analytic geometry of space-time diagrams. The light travels at constant speed, so that going out its graph is a straight line in the picture. Returning, it's another straight line. The slopes are $dx/dt = \pm c$, and the lines pass through the points $(x = 0, t = t_1)$ and $(x = 0, t = t_2)$ respectively, so the equations are

$$x = -c(t - t_2)$$
$$x = \quad c(t - t_1)$$

$$\Longrightarrow \quad \begin{aligned} x_{\mathsf{E}} &= \tfrac{c}{2}(t_2 - t_1) \\ t_{\mathsf{E}} &= \tfrac{1}{2}(t_2 + t_1) \end{aligned}$$

(9.6)
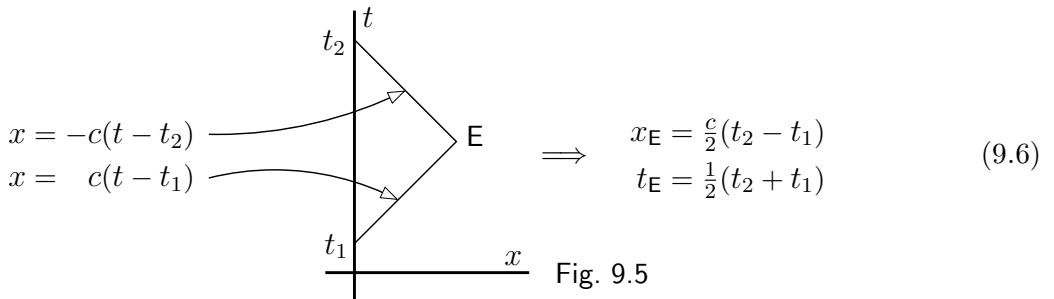
Fig. 9.5

The event at which the radar pulse hits the object is labeled E, and you solve those two simultaneous equations for x and t to get the solutions labeled x_{E} and t_{E}. Perhaps it is a police radar signal

bouncing from a speeding car. In any case, the space-time coordinates of the event are defined by the two measurements of time back at the radar set.

Now comes the interesting part: Bring your friend back into the picture, moving at velocity v. Using the *same definitions* to do the computation, what coordinates does your friend find for this same event? For now, make the velocity v a constant because it makes the mathematics far easier, and the motion is then a straight line in the picture, $x = vt$. Instead of having a separate radar and clock, which would have to be calibrated, simply use the same set for both. Then it is a question of computing what measurements the moving observer will get.

The outgoing and the returning radar pulses pass your friend at two points labeled a and b. The times at which they pass are t_a and t_b (you say). Computing each of those times is now finding the intersection of two lines — solving the two pairs of simultaneous equations on the left.

b: $x = -c(t - t_2)$
     $x = vt$

a: $x = c(t - t_1)$
     $x = vt$

$$t_b = ct_2/(c+v) = t_2/(1+v/c)$$

$$t_a = ct_1/(c-v) = t_1/(1-v/c)$$

(9.7)

Fig. 9.6

These times are when *you* think the radar pulse passed the moving observer as it goes out and bounces back. They are not the times the moving clock will read. That clock is running slow, remember, by a factor $\sqrt{1 - v^2/c^2}$, so the actual times on the moving clock as the pulse go by are the smaller times t_1' and t_2'. These numbers are what the moving observer will use to compute the coordinates of E.

$$t_1' = t_a\sqrt{1 - v^2/c^2} = \frac{t_1}{(1-v/c)}\sqrt{1 - v^2/c^2}$$

$$t_2' = t_b\sqrt{1 - v^2/c^2} = \frac{t_2}{(1+v/c)}\sqrt{1 - v^2/c^2}$$

(9.8)

Now all that's left is algebra, eliminating all the unwanted variables, $(t_1, t_2, t_a, t_b, t_1', t_2')$. The new coordinates of the event are

$$x_E' = \tfrac{c}{2}(t_2' - t_1')$$
$$t_E' = \tfrac{1}{2}(t_2' + t_1')$$

and solving Eqs. (9.6) for t_1 and t_2 $\longrightarrow$

$$t_1 = t_E - x_E/c$$
$$t_2 = t_E + x_E/c$$

(9.9)

$$x_E' = \frac{c}{2}\left[\frac{t_2}{(1+v/c)} - \frac{t_1}{(1-v/c)}\right]\sqrt{1 - v^2/c^2}$$

$$= \frac{c}{2}\left[\frac{t_E + x_E/c}{1+v/c} - \frac{t_E - x_E/c}{1-v/c}\right]\sqrt{} = \frac{1}{2}\left[\frac{(ct_E + x_E)(1-v/c)}{1-v^2/c^2} - \frac{(ct_E - x_E)(1+v/c)}{1-v^2/c^2}\right]\sqrt{}$$

$$= \frac{x_E - vt_E}{\sqrt{1 - v^2/c^2}}$$

$$t_E' = \frac{1}{2}\left[\frac{t_2}{(1+v/c)} + \frac{t_1}{(1-v/c)}\right]\sqrt{1 - v^2/c^2}$$

$$= \frac{1}{2}\left[\frac{t_E + x_E/c}{1+v/c} + \frac{t_E - x_E/c}{1-v/c}\right]\sqrt{} = \frac{1}{2}\left[\frac{(t_E + x_E/c)(1-v/c)}{1-v^2/c^2} - \frac{(t_E - x_E/c)(1+v/c)}{1-v^2/c^2}\right]\sqrt{}$$

$$= \frac{t_E - vx_E/c^2}{\sqrt{1 - v^2/c^2}}$$

Just as Eq. (9.5) referred to the point P without writing the letter, these look neater as

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}} = \gamma(x - vt) \qquad t' = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} = \gamma(t - vx/c^2)$$

$$y' = y \qquad\qquad\qquad\qquad\qquad z' = z \tag{9.10}$$

These Lorentz transformation equations are an algebraic codification of the two axioms written on the first page of this chapter. $\gamma = 1/\sqrt{1 - v^2/c^2}$ as before, and you often see $\beta = v/c$.

What about the other coordinates, y and z? They don't change, and the reason is nothing but the argument on page 302, showing that there is no contraction or expansion in the lateral direction.

This was a fair amount of algebra to arrive at these rather simple-looking equations, and if you look at other introductory texts on relativity you will find other, far less complicated derivations of the same two equations. Why then did I choose such an involved way to get to a result that another author may choose to do in a few lines? The first reason is that this method may be algebraically more complex, but it uses only the simple concept of determining how you measure x and t. A second reason is that this method is not restricted to motion at constant velocity; after all, if an observer can't make measurements while undergoing an acceleration even as small as g, then it's an awfully limited theory. There's no conceptual change in having an accelerated reference frame, just change the straight line, $x = vt$ in Figure 9.6 to a curved line. The real change is in the quantity of mathematics you have to carry out.* See problems 9.45 and 9.46.

As with the coordinate change Eq. (9.5) and the accompanying diagram, the picture of this transformation is important. As you can easily see, it is *not* a rotation, and the new axes don't look like those on page 305.

x-axis: $\quad t = 0$
t-axis: $\quad x = 0$
x'-axis: $t' = \gamma(t - vx/c^2) = 0$
t'-axis: $x' = \gamma(x - vt) = 0$
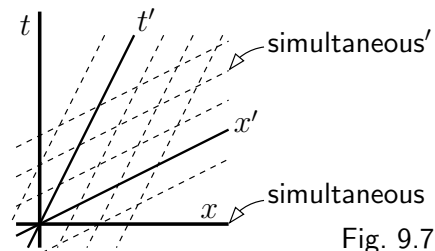
Fig. 9.7

The dashed lines fill in the coordinates as defined by the equations $x' = $ constant and $t' = $ constant. The picture has $v = c/2$.

Simultaneity

The phrase "at the same time" refers to events that occurred at the same value of t. The moving observer however uses different coordinates, so "at the same time" refers to the same values of t', the dashed lines parallel to the x'-axis in the above drawing. That simultaneity depends on the motion of the observer is the central radical departure in special relativity. It is this departure from common intuition that allows the paradoxes in the theory to be resolved.

How big is the effect? For an astronomical example, the Earth orbits the sun at a speed of about 30 km/s. This is $10^{-4}c$. At the distance of the star nearest the sun, about four light-years, how big a discrepancy does this make?

$$t' = 0 = t - vx/c^2, \quad\longrightarrow\quad t = vx/c^2 = 10^{-4}c \cdot 4\,\text{lt-yr}/c^2 = 4 \cdot 10^{-4}\,\text{yr} \cdot \frac{\pi \times 10^7\,\text{s}}{1\,\text{yr}} = 12000\,\text{s} \approx 3.5\,\text{hr}$$

---

* and despite what you may see in some books, this has nothing to do with general relativity.

This appears insignificant, but under some circumstances a much smaller speed can produce a very large effect. The current in an ordinary electric wire consists of electrons in motion. The average drift velocity of those electrons is surprisingly small, and in home wiring its magnitude is commonly less than 0.1 millimeters per second. How can such a small speed account for a large electric current? There are a *lot* of electrons. And what does this tiny speed have to do with relativity? That is what produces the magnetic field. How? Patience, only one more page.

Come back to the question of current in a moment. First, look further at simultaneity and use the Lorentz transformation to rederive length contraction. It's a consistency check, and it better give the same answer as before. Someone is moving past at a speed v and he claims to have an object of length L_0. Notice: He says that it isn't moving, so he gets to say that its proper length is L_0. What are the algebraic equations describing the front and back end of this object? Primed coordinates please.

$$x' = 0, \qquad x' = L_0 \qquad \text{for all } t' \text{—what could be simpler?} \tag{9.11}$$

I took the left end at zero to save algebra. The two lines of constant x' produce two parallel lines in the x-t coordinate system. Simply combine the Lorentz equations, Eq. (9.10) with the equation $x' = L_0$ and find the intersection of that line with the line $t = 0$. That is the dot in the figure.

$$\left.\begin{array}{l} x' = L_0 \\ t = 0 \\ x' = \gamma(x - vt) \end{array}\right\} \longrightarrow \quad \begin{array}{l} L_0 = \gamma x \\ x = L_0\sqrt{1 - v^2/c^2} \end{array}$$

The direct application of the transformation equations then reproduces the length contraction equation correctly. Can you do the same sort of analysis for time instead of length? Yes, *you* can; that's problem 9.5.

If you happen to have a ruler (stationary according to you) that is exactly this length $L = L_0\sqrt{1 - v^2/c^2}$ then what does your moving friend see? (You would call this L a proper length, but the symbol L_0 is already taken.) The equation for the right-hand end of your ruler is $x = L$, the dashed line in the next picture. What is the value of x' at the time $t' = 0$? *What do you expect it to be?* That is the dot in this picture. Again, the Lorentz transformation appears.

$$\begin{array}{cc} \begin{array}{l} t' = 0 \\ x = L \end{array} \quad \text{and} \quad \begin{array}{l} t' = \gamma(t - vx/c^2) \\ x' = \gamma(x - vt) \end{array}\end{array}$$

Eliminate x and t to get

$$t = vx/c^2 = vL/c^2 \longrightarrow x' = \gamma(L - v \cdot vL/c^2) = L\sqrt{1 - v^2/c^2}$$

and the Lorentz contraction appears again! Each person says that the other's rulers have shrunk.

Current

An ordinary copper wire carrying a current I has two parts: the stationary positive charges and the moving negative ones. There are cases such as fluorescent lights where both types of charge move, but that's an unneeded complication. In fact it doesn't matter whether the positive charges are moving right or the negative charges are moving left. Both represent a current to the right.* If it's convenient

---

* It wasn't until 1879, when Edwin Hall discovered the Hall effect, that it was possible to tell the difference, and the electron itself wasn't discovered until the late 1890s.

to simplify the explanations by assuming that the positive charges are moving right and the negative charges are stationary, that's o.k. It matters not which. The wire has no net charge, so the distance between the positive charges is on average the same as the distance between the negative ones. At $t = 0$, the x-axis, for every positive charge there is a negative one. In this picture the positive charges are moving right ($x = $ const. $+ vt$) and the negative ones are stationary ($x =$const.)
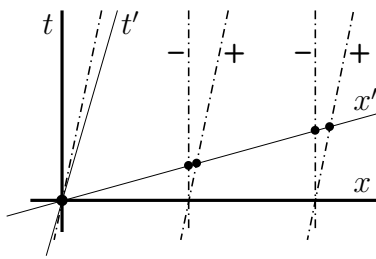
Positive and negative charges in a wire

Dashed lines are the charges.

Fig. 9.8

> Without solving a single equation, you can look at this picture and see that the moving observer, who says that the x'-axis is represents simultaneous points, concludes that the positive charges are farther apart than the negative ones. That implies that the moving observer sees a net negative charge density on the wire.

If the "moving observer" is a charge, it will feel a force because of the net charge on the wire. If that moving charge moves faster it will see more charge on the wire and will feel a larger force. You will notice that this force is perpendicular to the velocity vector. This velocity dependent force is there even though you, as a stationary observer, say that there is no electric field to push the charge; the wire is neutral. Instead you call the effect a "magnetic field", and conclude that the force it exerts is proportional to the velocity of the charge. This is how you derive magnetism from the combination of electric fields and relativity.

There are some details to fill in, allowing you to get the familiar "$q\vec{v} \times \vec{B}$" expression, but the essential ideas are already here. The consequence of the analysis is that just as space and time become intertwined in relativity, so do electric and magnetic fields. What is an $\vec{E}$ to me will be a combination of $\vec{E}$ and $\vec{B}$ to you. Similarly, my $\vec{B}$ will be your mixture of $\vec{B}$ and $\vec{E}$. The relevant equations are *not* the Lorentz transformation this time. They're quite different.

Future and Past

If simultaneity has become malleable, what happens to the meaning of the words future and past? Are they affected? Yes. According to a stationary observer, us, the future is defined by the inequality $t > 0$. The past is $t < 0$. Someone in motion uses the same concept, but it becomes $t' > 0$ and $t' < 0$ instead, and these are indicated by the various shaded regions in the drawings.
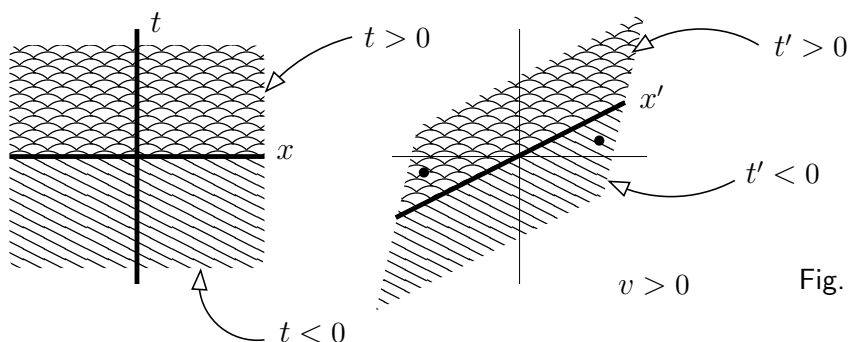
Fig. 9.9

The familiar part is defined by the horizontal line, the x-axis that separates our past from our future. The unfamiliar part is the past and future for the moving observer. $t' > 0$ is tilted because $t' = \gamma(t - xv/c^2)$, so the demarcation between past and future is the line $t = vx/c^2$. Remember the units I'm using to draw the pictures: $c = 300$ m/μs, the line $x = ct$ is at $45°$. $c = 1$ in units of 300 m and $1\,\mu$s, so the boundary between past and future will not get any steeper than $\pm 45°$. If the motion is in the opposite direction, $v < 0$, the tilt is reversed and goes from upper left to lower right.

This picture implies that what is in the future for me may be in the past for you and vice versa. The two dots in this Figure 9.9 represent events are exactly that way. Notice that the dot on the left is an event that occurred *before* the one on the right according to us, but it occurs *after* the one on the right for the moving observer. Their time-order is reversed. Does this mean that cause and effect can be reversed? No, because signals can't travel faster than light, and that implies that these two dots are far enough apart that light can't get from one to the other in the time between them.

There are two types of future and two types of past: relative and absolute. Absolute future is something that everybody agrees on. Relative future will mean different things to different people, depending on their relative velocities. The two dots above represent a case where future and past are relative to the observer, but remember that the x'-axis can tilt only up to $\pm 45°$ and no farther. There is a limit on just how far the relative future and past can be pushed.

Fig. 9.10
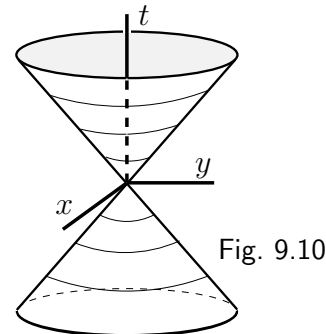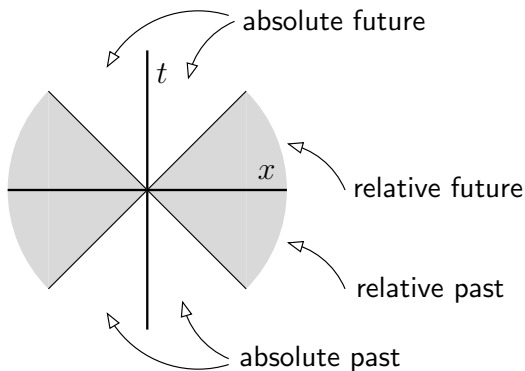
The events in the absolute future are in the future according to everyone, and they are also events that can be reached by a signal that starts at the origin. Such a signal has the equation $x = vt$ with $|v| \leq c$, and it can pass through any point within the absolute future. In the next picture, jumping up a dimension so that there are two spacial coordinates x and y accompanying t, the absolute future is within the upper part of the cone. This cone is the "light cone", $r = \sqrt{x^2 + y^2} = c|t|$.

The concepts of cause and effect involve the order of time. A cause precedes an effect. If the past and the future can become muddled because of relativity, then it puts some constraints on the way that things can interact. Something can cause an effect only if the effect is in the future of the cause *according to everyone*. This means that something that happens at the origin $(x = 0, t = 0)$ can only cause something on or within the future light cone, at a distance $r \leq ct$. Outside that cone, the past and future become relative concepts, and the relative past and future there are perhaps best called "elsewhere".

9.3 Relative Velocity

You are at rest (according to you). There are two objects, a moving observer with velocity v and some other object with velocity u. What is *their* relative velocity? This question will have different answers depending on just how you define the phrase "relative velocity". For one possible definition you can ask how you would measure their distance apart and how fast that distance is changing. In your time

interval Δt the moving observer goes a distance $v\Delta t$. The object goes a distance $u\Delta t$. The distance between those two has changed by the amount $u\Delta t - v\Delta t$. You then say that the relative velocity is

$$\frac{u\Delta t - v\Delta t}{\Delta t} = u - v$$

What could be simpler? This is exactly what you get if you know nothing about relativity, and it is not wrong. It is just not the definition of relative velocity needed here; there is a subtle difference in how it is stated.

The other way to define relative velocity: Ask the moving observer what velocity the other object has. *That will give a different answer,* not the one in the preceding equation. A slick and easy way to derive the equation for relative velocity starts from the basic definition of velocity as a derivative and then uses a common differentiation trick to manipulate it.

> x' and t' are functions of x and t (the Lorentz transformation).
> x is a function of t $(x = ut)$.
> So, x' and t' are functions of t.
> Compute dx'/dt' using t as a parameter, as in Eq. (0.37).

$$\frac{dx'}{dt'} = \frac{dx'/dt}{dt'/dt}, \quad \text{then} \quad \begin{array}{l} x' = \gamma(x - vt) \\ t' = \gamma(t - vx/c^2) \end{array} \implies \begin{array}{l} dx'/dt = \gamma(dx/dt - v) \\ dt'/dt = \gamma(1 - v(dx/dt)/c^2) \end{array} \tag{9.12}$$

$$\text{divide:} \quad u' = u'_x = \frac{dx'}{dt'} = \frac{\gamma(dx/dt - v)}{\gamma(1 - v(dx/dt)/c^2)} = \frac{u - v}{1 - \frac{vu}{c^2}} \tag{9.13}$$

Pay attention to this derivation. It carries through with little change in more complicated cases, such as in computing acceleration, Eq. (9.22) or superluminal speeds, Eq. (9.21). A probable point of confusion here: v is a number, not a function of time. What depends on t are x, x', t', and of course t itself.

If the object is a photon then $u = c$, and

$$\frac{c - v}{1 - vc/c^2} = \frac{c - v}{c - v}c = c \tag{9.14}$$

Everyone agrees about the vacuum speed of light. If instead, the observer and object are moving toward each other, say u positive and v negative, at $\frac{3}{4}c$, then the speed of one according to the other is

$$u' = \frac{\frac{3}{4}c + \frac{3}{4}c}{1 + (\frac{3}{4}c)^2/c^2} = \frac{\frac{3}{2}c}{1 + \frac{9}{16}} = \frac{24}{25}c$$

and this is not the $\frac{3}{2}c$ expected from a non-relativistic of velocity additon.

What about the other components of velocity? Recall that the Lorentz transformation for these other coordinates says that y and z don't change. Does that imply that u_y is the same for both? No, because t is different. Do the same calculation for y' as for x', using the same equations as in Eq. (9.12), but including $y' = y$. Notice that in the work of the last few lines, the symbol "u" is really "u_x". It's just that most of the development has been in one dimension of space, so it hasn't mattered. Sometimes however, you have to be a little more careful.

$$\frac{dy'}{dt'} = \frac{dy'/dt}{dt'/dt}, \quad \text{and} \quad \begin{array}{l} dy'/dt = dy/dt \\ dt'/dt = \gamma(1 - v(dx/dt)/c^2) \end{array}$$

divide: $\quad u'_y = \dfrac{dy'}{dt'} = \dfrac{dy/dt}{\gamma(1 - v(dx/dt)/c^2)} = \dfrac{u_y}{\gamma(1 - v(dx/dt)/c^2)} = \dfrac{u_y}{\gamma\left(1 - \frac{vu_x}{c^2}\right)}$ (9.15)

Even if $u_x = 0$, the y-component of velocity will change, decreasing by a factor of γ.

Example

▶ What if the original u_x and u_y represent the motion of a photon? What does the moving observer find for the speed of the photon? This is just a bunch of algebra, but it is worth checking. Maybe you can find a slicker way to avoid all these manipulations, maybe by doing problem 9.37.

$$u'_x = \frac{u_x - v}{1 - \frac{vu_x}{c^2}}, \qquad u'_y = \frac{u_y}{\gamma\left(1 - \frac{vu_x}{c^2}\right)}$$

$$
\begin{aligned}
u'^2_x + u'^2_y &= \frac{(u_x - v)^2 \gamma^2 + u_y^2}{\gamma^2\left(1 - \frac{vu_x}{c^2}\right)^2} = \frac{(u_x - v)^2 + u_y^2(1 - v^2/c^2)}{\left(1 - \frac{vu_x}{c^2}\right)^2} \\
&= \frac{(u_x - v)^2 + (c^2 - u_x^2)(1 - v^2/c^2)}{\left(1 - \frac{vu_x}{c^2}\right)^2} \qquad [\text{original speed} = c] \\
&= \frac{u_x^2 - u_x^2 + v^2 - v^2 - 2vu_x + c^2 + u_x^2 v^2/c^2}{\left(1 - \frac{vu_x}{c^2}\right)^2} = c^2
\end{aligned}
$$

(9.16)

9.4 Space-Time Intervals

In the Euclidean geometry expressed in the picture and equations (9.4), there is an invariant of the transformation: the distance to the origin, $\overline{OP}$. It is the square root of $x^2 + y^2$ or of $x'^2 + y'^2$, whichever coordinate system you choose. You can verify this algebraically by using the equations (9.5) directly.

$$x'^2 + y'^2 = (x \cos\alpha + y \sin\alpha)^2 + (-x \sin\alpha + y \cos\alpha)^2 = x^2 + y^2$$
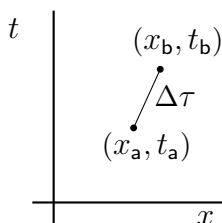
Multiply all the factors and collect the terms.

There is a similar relation for the geometry described by the Lorentz transformation, but it is *not* a geometry that Euclid would have recognized. Instead of the sum of squares, you use the difference of squares to compute the invariant. You also need factors of c in order to keep the dimensions straight, so the expression to examine is not $x^2 + y^2$, but $x^2 - c^2 t^2$. Instead of rotational transformations, use the Lorentz transformation Eq. (9.10).

$$x'^2 - c^2 t'^2 = \left(\gamma(x - vt)\right)^2 - c^2\left(\gamma(t - vx/c^2)\right)^2 = x^2 - c^2 t^2$$ (9.17)

Again, just multiply all the factors and collect terms.

In order to interpret what this means, it is better to interchange the terms and then to look at $t^2 - x^2/c^2$. Even better, use the interval between two events instead of the distance to the origin. Take the events to have coordinates (x_a, t_a) and (x_b, t_b), with the intervals

$$\Delta t = t_b - t_a \qquad \longrightarrow \qquad (\Delta t)^2 - (\Delta x)^2/c^2 = \Delta\tau^2$$

$$\Delta x = x_b - x_a$$

(9.18)

Let this line segment represent your friend's motion, so that he was at coordinate x_a at time t_a and in the time interval $\Delta t = t_b - t_a$ he moved to the point x_b. You can look at this and say that

the time interval is Δt, but how much time has elapsed on your friend's clock, remembering that it is running slow according to you. You know the calibration factor, and all that you need to know is his speed. That's easy; it is $v = \Delta x / \Delta t$. His own clock then reads an interval (the proper time interval) that is smaller than Δt by the standard calibration factor.

$$\Delta t \sqrt{1 - \frac{v^2}{c^2}} = \Delta t \sqrt{1 - \frac{(\Delta x / \Delta t)^2}{c^2}} = \sqrt{(\Delta t)^2 - (\Delta x)^2 / c^2} = \Delta \tau \qquad (9.19)$$

This invariant space-time interval that appeared in Eqs. (9.17) and (9.18) now has a simple interpretation: It is the proper time interval between two events, so *of course* it's the same for everyone — an invariant.

When you look at it from a reversed perspective, this proper time interval *must* be an invariant and *must* have the same value according to all observers. If you ask your friend what the time interval was between these two events, he gives you an answer. If someone else, moving three times as fast and in the other direction, asks him the same question he will give the same answer. This proper time interval has the same value for everyone, and everyone can compute its value by the same algorithm that you used in Eqs. (9.18) and (9.19).

Non-constant Velocity

The proper time interval between two events is $\Delta \tau$. The preceding calculation assumes however that your friend who went from (x_a, t_a) to (x_b, t_b) moved at constant velocity. What if that's not the case? You do the same thing that you do whenever you have a complicated problem: break it into little pieces and do successive approximations. In other words, use calculus.

Let the starting time for the curve be t_a and the ending time t_b as before. Pick $N - 1$ points along the curved path shown, dividing the curve into N parts: $t_a = t_0 < t_1 < \ldots < t_{N-1} < t_N = t_b$. The straight line approximation to the curve in the time interval $t_k < t < t_{k+1}$ has a proper time

$$\Delta \tau_k = \sqrt{(\Delta t_k)^2 - (\Delta x_k)^2 / c^2}$$

The total proper time from the start of the path to its end is the limit of the sum

Fig. 9.11

$$\lim_{\Delta t_k \to 0} \sum_0^{N-1} \Delta \tau_k = \int_a^b d\tau = \int_{t_a}^{t_b} \sqrt{(dt)^2 - (dx)^2 / c^2} = \int_{t_a}^{t_b} dt \sqrt{1 - v^2 / c^2}$$

with $v = dx/dt$.

The Twin Effect

The simplest example applying this integral uses paths that are straight line segments. You stay home at $x = 0$ and your friend leaves at high speed v. At some time he turns around and returns at the same speed. How do your ages compare? That is, compare your elapsed proper time to his elapsed proper time.

You: $\int d\tau = \int_0^T dt = T$

Him: $\int d\tau = \int_0^{T/2} dt \sqrt{1 - v^2/c^2} + \int_{T/2}^T dt \sqrt{1 - v^2/c^2} = T\sqrt{1 - v^2/c^2}$

(Hardly worthy of calculus.) If $v = 0.99c$ and $T = 50\,\text{yr}$ your friend returns having aged $50\sqrt{1-.99^2} = 7\,\text{years}$. This effect seems to upset many people— "That can't be." or "There must be some mistake." No, that's the way the universe is built. Twins will not be the same age, and one travelling identical twin can be 43 years younger than the other who stayed behind.

If the space-time path is not straight, but curved, the calculation is the same. Take a path with constant acceleration, $d^2x/dt^2 = -g$.* That path is a parabola, $x = A + Bt - gt^2/2$. Adjust the parameters so that this starts from the origin and returns at time T, then this is $x = gt(T-t)/2$. Now the proper time on the journey is

$$T_{\text{total}} = \int d\tau = \int_0^T dt\sqrt{1 - g^2(T-2t)^2/4c^2} = \int_{t=0}^{t=T} \frac{c}{g}du\sqrt{1-u^2}$$

where $u = (2t - T)g/2c$. Now let $u = \sin\theta$, then

$$T_{\text{total}} = \frac{c}{g}\int \cos^2\theta\, d\theta = \frac{c}{g}\int \frac{1}{2}[1 + \cos 2\theta]$$

$$= \frac{c}{2g}\left[\theta + \tfrac{1}{2}\sin 2\theta\right] = \frac{c}{2g}\left[\theta + \sin\theta\cos\theta\right]\Big|_{t=0}^{t=T}$$

$$= \frac{c}{g}\left[\sin^{-1}\frac{gT}{2c} + \frac{gT}{2c}\sqrt{1 - \left(\frac{gT}{2c}\right)^2}\right]$$

(9.20)

Naturally, the final step is to analyze this carefully. In particular, what can go wrong with this answer? This path is along the $\pm x$-direction only, so its motion is along a straight line, but it's commonly called a "*curved path*" because the picture is in space-time and not in space alone. The geometry has taken a step up. See problem 9.15.

In Euclidean geometry you are accustomed to the fact that the shortest distance between two points lies along a straight line, but in this (pseudo-Euclidean) space-time geometry the straight line produces the *longest* space-time interval between two events. That's easy to see because this integral, $\int d\tau$, is the same for all observers, and for the observer (your friend again) who says that the positions at t_a and at t_b are the same, the answer is $t_b - t_a$.

$$\int d\tau = \int_{t_a}^{t_b} dt = t_b - t_a \qquad \text{but any other path has} \qquad \int d\tau = \int_{t_a}^{t_b} dt\sqrt{1 - \frac{v^2}{c^2}}$$

and that final square root is always less than or equal to one. For the corresponding Euclidean calculation of the distance between two points you can choose your x-y coordinates so that both ends of the path have the same x-coordinate. Then the straight line path has length

$$\int d\ell = \int_{y_a}^{y_b} dy = y_b - y_a \qquad \text{but another path has}$$

$$\int d\ell = \int_{y_a}^{y_b} \sqrt{(dx)^2 + (dy)^2} = \int_{y_a}^{y_b} dy\sqrt{1 + \left(\frac{dx}{dy}\right)^2}$$

and this square root is always *greater than* or equal to one, implying that the straight line has the shortest length.

* Watch out for a distinction here: this is not proper acceleration. That will appear in section 9.6.

Your car's odometer will tell you how far your car has travelled over the course of all its journeys. If you drive your car from Toronto to Miami while a friend does the same but uses a different route, you will not be at all surprised to find that the odometers on the two cars will register different distances for the trips. Why then is it so baffling that a clock (which is an odometer for time) will register different *durations* for such a trip? And someone who stays home will register both different distances *and* different times. The bafflement lies not in the physics but in the psychology. You're not used to it.

9.5 Superluminal Speeds
Jets of high energy gas can be emitted from distant quasars or from violent objects in our own galaxy. Under the right circumstances the jet can *appear* to be moving at a speed greater than light. It is an illusion, but the analysis is interesting.
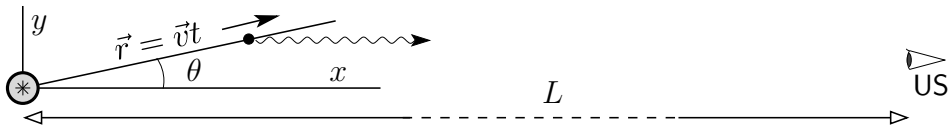
A ball of intensely hot gas is ejected at an angle θ to our line of sight, and it is moving at some speed v radially away from the source. Take $t = 0$ as the time it was ejected, then the x and y-coordinates of the ball as a function of time are

$$x = vt \cos\theta, \qquad y = vt \sin\theta$$

The center of the system is at a distance L away from us, and the light emitted at time t will have less distance to travel than did the light that was emitted at the start $(t = 0)$. The source is trying to catch up to the light, and the light emitted at time t will reach us at time

$$t_{\text{arrival}} = t + (L - vt \cos\theta)/c, \qquad \text{coming from the } y\text{-coordinate} \qquad y = vt \sin\theta$$

The derivative of the y-coordinate with respect to the arrival time is its apparent lateral velocity,

$$\frac{dy}{dt_{\text{arrival}}} = v_{\text{apparent}} = \frac{dy/dt}{dt_{\text{arrival}}/dt} = \frac{d(vt \sin\theta)/dt}{d(t + (L - vt\cos\theta)/c)/dt} = \frac{v \sin\theta}{1 - \frac{v}{c}\cos\theta} \tag{9.21}$$

If v is close to c and the angle is small, this can be larger than c. For example, if $v = 0.95c$ and $\theta = 0.1\,\text{radian} = 6°$ then this apparent speed is $1.73\,c$. You can easily verify that the maximum (over all theta) apparent speed is $v\gamma$ and that this is greater than c if $\beta = v/c > 1/\sqrt{2}$. Notice that the only aspect of relativity that shows up in this calculation is the statement that we all agree on the same speed for light and that its speed is independent of the motion of the source of the light.

Other examples of this sort of analysis appear in problems 9.18 and 9.19.

9.6 Acceleration
You *did* read section 0.2 didn't you? You also did problems $0.7 - 0.16$.
How does acceleration behave when you switch to a moving observer? In chapter five it was simple. In Eq. (5.2) everyone agreed on the value of the acceleration. There, $d^2x'/dt'^2 = d^2x/dt^2$, but that no longer holds because the equation relating times for the two observers is no longer $t' = t$.

When an object has an acceleration that is $d^2x/dt^2 = du/dt$ according to you, what is it according to someone moving at velocity v? Look back carefully at the derivation of relative velocity in Eq. (9.12). Understand it there and the following will be just some more algebra, but without any

new conceptual complications. As in that previous derivation, v is a constant. Here $\dot{u} = du/dt$:

$$\frac{du'}{dt'} = \frac{du'/dt}{dt'/dt}$$

$$\frac{du'}{dt} = \frac{d}{dt}\frac{u-v}{1-uv/c^2} = \frac{\dot{u}}{1-uv/c^2} - \frac{u-v}{(1-uv/c^2)^2}\cdot\left(\frac{-v\dot{u}}{c^2}\right) = \dot{u}\frac{1-v^2/c^2}{(1-uv/c^2)^2}$$

$$\frac{dt'}{dt} = \frac{d}{dt}\gamma(t-vx/c^2) = \gamma(1-vu/c^2)$$

$$\frac{du'}{dt'} = \frac{\dot{u}(1-v^2/c^2)/(1-uv/c^2)^2}{\gamma(1-uv/c^2)} = \frac{\dot{u}(1-v^2/c^2)^{3/2}}{(1-uv/c^2)^3} \tag{9.22}$$

If $v = 0$ then this reduces to $\dot{u}$, as it must. Also look at what happens as v becomes close to $+c$. In that case, unless u is also close to $+c$, the moving observer says that the object is already moving at very close to the speed of light, so it *can't* speed up much. Its acceleration must be small and that is exactly what this equation says. The numerator is $(1-v^2/c^2)^{3/2}$.

For an important special case of this equation, take $v = u$. That means that the new acceleration, du'/dt', is being computed by the person who says that the object is (momentarily) at rest. That is the definition of proper acceleration.

$$a_{\text{proper}} = \frac{\dot{u}(1-u^2/c^2)^{3/2}}{(1-uu/c^2)^3} = \frac{\dot{u}}{(1-u^2/c^2)^{3/2}} = \gamma^3\dot{u} \tag{9.23}$$

If you are taking an interstellar journey in a ship accelerating at what you perceive to be a constant a, then what does someone remaining behind on Earth say about your motion? In other words, what is the equation to describe constant proper acceleration; what replaces $at^2/2$? For a more realistic example, the acceleration of a charged particle in a uniform electric field will be described by constant proper acceleration. If you think that's not obvious, you're right.

This is a differential equation, and for a constant proper a, and recalling section 0.2 and problem 0.16

$$a = \frac{\dot{u}}{(1-u^2/c^2)^{3/2}} \longrightarrow a\,dt = \frac{du}{(1-u^2/c^2)^{3/2}} \longrightarrow at = \int_0^u \frac{du}{(1-u^2/c^2)^{3/2}}$$

Let $u = c\tanh\omega$ then (with $\omega = 0$ being $t = 0$)

$$at = \int \frac{c\,\text{sech}^2\,\omega\,d\omega}{(1-\tanh^2\omega)^{3/2}} = c\int\cosh\omega\,d\omega = c\sinh\omega \tag{9.24}$$

What is x now (starting at $x = x_0$)?

$$x - x_0 = \int_0^t u\,dt = \int c\tanh\omega\,dt = c\int\tanh\omega\,\frac{c}{a}\cosh\omega\,d\omega = \frac{c^2}{a}\int\sinh\omega\,d\omega$$

$$= \frac{c^2}{a}\cosh\omega\,\Big|_{t=0}^{t=t} = \frac{c^2}{a}(\cosh\omega - 1) \tag{9.25}$$

This is a particularly simple expression if you choose the initial value of x to be $x_0 = c^2/a$. Then
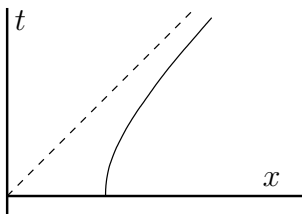
$$x = \frac{c^2}{a}\cosh\omega \qquad \text{and} \qquad t = \frac{c}{a}\sinh\omega$$

$$x^2 - c^2t^2 = \frac{c^4}{a^2}\left[\cosh^2\omega - \sinh^2\omega\right] = \frac{c^4}{a^2} \tag{9.26}$$

This is an equation for a hyperbola, replacing the non-relativistic parabola. Another way to say this is

$$x = x_0 + \frac{at^2}{2} \qquad \text{becomes} \qquad x = \sqrt{\frac{c^4}{a^2} + c^2 t^2} + x_0 - \frac{c^2}{a} \qquad (9.27)$$

If you're skeptical, do a series expansion about $t = 0$ to check it.
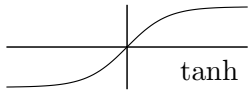
9.7 Rapidity
The parameter ω showed up as a change of variables in Eq. (9.24), appearing simply as a way to do an integral. It is really much more than that. Go back to the Lorentz transformation itself, Eq. (9.10), and make the same change of variable from v to ω.

$$v/c = \tanh \omega, \qquad \sqrt{1 - v^2/c^2} = \sqrt{1 - \tanh^2 \omega} = \text{sech } \omega, \qquad \gamma = 1/\text{sech } \omega = \cosh \omega \qquad (9.28)$$

$$\begin{aligned} x' &= \gamma(x - vt) = \cosh \omega \, (x - ct \tanh \omega) \\ t' &= \gamma(t - vx/c^2) = \cosh \omega \, (t - (x/c) \tanh \omega) \end{aligned} \qquad \Longrightarrow \qquad \begin{cases} x' = x \cosh \omega - ct \sinh \omega \\ t' = t \cosh \omega - (x/c) \sinh \omega \end{cases} \qquad (9.29)$$

This variable ω is called the rapidity.

What does velocity addition look like in terms of this variable?

$$\frac{v_1 + v_2}{1 + v_1 v_2/c^2} = c \frac{\tanh \omega_1 + \tanh \omega_2}{1 + \tanh \omega_1 \tanh \omega_2} = c \tanh(\omega_1 + \omega_2) \qquad (9.30)$$

$$\longrightarrow \qquad \omega_\text{total} = \omega_1 + \omega_2$$

If you're not sure about this identity for the hyperbolic tangent of a sum, multiply numerator and denominator by $\cosh \omega_1 \cosh \omega_2$ and refer to section 0.2 and problem 0.15. These equations (9.29) look much more like the equations for rotations, Eq. (9.5), than do the form of the Lorentz transformations that use v. The rapidity variable behaves the same way that the angle variable does in rotations, and it is a parameter that is far more convenient and natural than v in most applications. You expect the sum of two angles to be the total angle, but v doesn't do that. Rapidity however does. Cf. problem 9.10.

9.8 Energy and Momentum
It should be no surprise that energy and momentum do not behave the same way as in Newtonian mechanics. First the results, and leave the derivation until section 9.11, after you've used them for a while. For the moment, no potential energy; see Eq. (9.36) for that.

$$E^2 = p^2 c^2 + m^2 c^4, \qquad \vec{p} = \frac{E}{c^2} \vec{v} \qquad (9.31)$$

These are the exact, no approximation, cover-all-cases equations relating energy, momentum, velocity, and mass. Anything else is either a special case or an approximation. This energy E is not the kinetic energy, but it now includes mass as another form of energy. To see this, look at what happens for zero velocity.

$$\vec{v} = 0 \quad \longrightarrow \quad \vec{p} = 0 \quad \longrightarrow \quad E^2 = 0^2 c^2 + m^2 c^4 \quad \longrightarrow \quad E_0 = mc^2$$

Here E_0 is a common notation for the value of the energy when the velocity is zero — the rest energy.

Instead, what if the speed is small? In this case E is still close to the rest value, mc^2, so the momentum is

$$\vec{p} \approx \frac{mc^2}{c^2} \vec{v} = m\vec{v}$$

and this reduces to the ordinary classical result. Now what about the kinetic energy? That's the the energy due to motion, and it is the *difference* between the energy E and the rest energy $E_0 = mc^2$. For small speeds, the momentum is small, so this allows a series expansion. Rearrange the terms into the form $(1 + \text{small})^n$.

$$K = E - E_0 = \sqrt{p^2 c^2 + m^2 c^4} - mc^2 = mc^2 \left(1 + p^2/m^2 c^2\right)^{1/2} - mc^2$$

$$= mc^2 \left(1 + \frac{1}{2} \frac{p^2}{m^2 c^2} + \cdots\right) - mc^2 = \frac{p^2}{2m} + \cdots \tag{9.32}$$

In this small speed approximation, $\vec{p} = m\vec{v}$, then $p^2/2m = (mv)^2/2m = mv^2/2$ and it recovers the familiar, non-relativistic result. You also see what the word "small" means in order to allow this approximation: $p \ll mc$, and that in turn is equivalent to $v \ll c$.

These manipulations assumed that the mass is non-zero. What happens if $m = 0$? The first of the equations (9.31) is then

$$E^2 = p^2 c^2 + 0 \quad \longrightarrow \quad E = cp \tag{9.33}$$

That in turn makes the second of those equations

$$\vec{p} = \frac{E}{c^2} \vec{v} = \frac{cp}{c^2} \vec{v} = p \frac{\vec{v}}{c}$$

The magnitude of this equation is $|\vec{p}| = p = pv/c$, and that implies $v = c$. A mass zero particle must move at the speed of light. (For example, light.) That the mass of a particle is zero does not imply that its energy or momentum is zero. That's a misconception carried over from Newtonian mechanics, and it does not apply here. That light, a particle of zero mass, carries energy is clear enough — just step outside and into the sunshine. That it carries momentum is not so obvious because from Eq. (9.33) the momentum is $p = E/c$, and on a human scale that's not much. For an electron however it can be a lot, and the force light exerts on that scale can be significant. This pressure from sunlight can have a noticeable effect on small objects such as dust particles and asteroids. The force on an asteroid occurs because the sunlight hits one side of the asteroid while the other is in shade. It can even measurably affect the motion of asteroids though a surprising mechanism called the Yarkovsky effect, discussed in section 9.10. At high intensity the force exerted by light can be used to compress matter by huge factors, firing 192 high intensity lasers from different directions, aimed at a small sample. This has been used in the hope that it can lead to controlled nuclear fusion, with release of enough energy to create a commercially feasible source of power. (Well, maybe someday.)*

You've usually seen kinetic energy and momentum expressed in terms of velocity, and you can do that here simply by algebraic manipulation of the equations (9.31). Eliminate p between the two equations and for $m \neq 0$ the result is

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} = mc^2 \gamma, \qquad \text{then} \qquad \vec{p} = \frac{E}{c^2} \vec{v} = \frac{m\vec{v}}{\sqrt{1 - v^2/c^2}} = m\vec{v}\gamma \tag{9.34}$$

As in Eq. (9.32) you can use the binomial expansion to get

$$E = mc^2 \gamma = mc^2 + \frac{1}{2} mv^2 + \cdots$$

---

* See the Wikipedia entry for National Ignition Facility, and find its relationship with Star Trek.

What are energy and momentum expressed in terms of rapidity? As with the Lorentz transformation and the velocity addition formula, the energy and momentum become especially simple in terms of ω. From Eq. (9.28)

$$E = mc^2\gamma = mc^2\cosh\omega, \qquad p = mc\frac{v}{c}\gamma = mc\tanh\omega\cosh\omega = mc\sinh\omega \qquad (9.35)$$

The first energy-momentum equation of (9.31) is now a trigonometric identity.

$$E^2 - p^2c^2 = (mc^2)^2\cosh^2\omega - (mc^2)^2\sinh^2\omega = m^2c^4$$

What identity is the second energy-momentum equation?
And where does potential energy fit into this?

$$E_{\text{total}} = E_{\text{kinetic}} + E_{\text{mass}} + E_{\text{potential}} = E + E_{\text{potential}}$$

This means that when you add this complication, the mass-energy equation is

$$\left(E_{\text{total}} - E_{\text{potential}}\right)^2 - p^2c^2 = m^2c^4 \qquad (9.36)$$

We won't be using this, but if someday you need to understand the fine structure of the hydrogen atom, come back here. It is the total energy that you are looking for in that application.

9.9 Applications

At the atomic level particles are not conserved. Protons, electrons, and of course photons can be created where there were none before. Such processes still must obey other conservation laws: energy, momentum, angular momentum, and a few others, but mass is *not* one of the conserved quantities. The most obvious example of this is radiation. An atom in an excited state in a vacuum with no other atoms around can emit a photon. That's a particle that didn't exist before the radiative process occurred. The photon carries energy and momentum ($E = pc$), so the atom will have lower energy than it did before, and it will recoil by carrying momentum opposite in direction to the emitted photon. The atom's mass is less than before it emitted the photon.

An electron can collide with another electron, and the two electrons then go off in different directions. If the collision occurs at high enough energy the result can be the creation of other particles. Start with two electrons, and after the collision you can have three electrons and a positron (an anti-electron with the same mass as the electron but opposite charge). This pair creation process can however happen only if the incoming electron has high enough energy. Simply look at the mass-energy before and after the creation process: You have $2mc^2$ before the collision and $4mc^2$ after the collision. You might think that the incoming electron would then need $2mc^2$ worth of kinetic energy in order for there to be enough total energy to create two new particles having rest energy mc^2 each. Not that simple. When you add conservation of momentum to the mix, the equations become a little more involved and you will in the end, need $6mc^2$ of kinetic energy to create an electron-positron pair, problem 9.34.

When a positron comes near to an electron they attract each other and will often form a sort of atom, called positronium. It's like a hydrogen atom but replacing the proton by a positron. In a short time, about 10^{-10} s, the particles can annihilate each other and emit two* photons. What energy do these have? If the positronium formed at rest then conservation of momentum implies that the two photons must have opposite momenta, $\vec{p}_1 = -\vec{p}_2$. Their energies are then the same and must each be

---

* Other, modes are possible too, with three or more rarely, more photons.

one half the original energy: $2mc^2/2 = mc^2$. For the electron, $mc^2 = 511\,000$ eV, or about 0.5 MeV. For comparison the photons in a medical X-ray are about $50\,000$ eV and in visible light the photon energy is about 2 eV.

How much energy does it take to create a positron-electron pair when a high energy photon collides with an electron? Before the collision the photon has momentum of magnitude p and energy $E = pc$. The electron has zero momentum and energy mc^2. After the collision there are to be two electrons and one positron moving off in various directions and carrying various amounts of energy. To find the *minimum* energy possible, the three particles should be moving very slowly with respect to each other — in fact you don't want them moving away from each other at all. The minimum final state is three particles all having the same momentum and moving together. Call the momentum of *each* final particle $\vec{p}'$ with energy E'.

before   after

$$
\begin{array}{ll}
1. \quad E = pc \ \text{(Eq. (9.33))} & 3. \quad \vec{p} = 3\vec{p}' \ \text{(cons. momentum)} \\
2. \ E + mc^2 = 3E' \ \text{(cons. energy)} & 4. \ E'^2 = p'^2 c^2 + m^2 c^4 \ \text{(Eq. (9.31))}
\end{array}
\tag{9.37}
$$

These are four equations in four unknowns; solve for the initial energy E, though first find p'.

$$
E = 3p'c \ \longrightarrow \ 3p'c + mc^2 = 3E' \ \longrightarrow \ (3p'c + mc^2)^2 = (3E')^2 = 9E'^2 = 9p'^2 c^2 + 9m^2 c^4
$$

$$
\longrightarrow \ 6p'mc^3 + m^2 c^4 = 9m^2 c^4 \ \longrightarrow \ p' = \tfrac{4}{3}mc \ \longrightarrow \ E = 3 \cdot \tfrac{4}{3}mc^2 = 4mc^2 \approx 2 \, \text{MeV}
\tag{9.38}
$$

This is far more than the $2mc^2$ needed simply to account for the created mass-energy of the new positron and electron because much (half) the original kinetic energy of the photon goes into the kinetic energy of the final particles and not just into their masses.

The calculation doesn't change on switching from electrons and positrons to protons and anti-protons. The same concepts apply, where the anti-proton is a negatively charged version of a proton. A photon can hit a proton and create a p-p̄ pair. The minimum energy that the photon must have is again $4mc^2$, but this time it is the proton's mass, almost 2000 times that of the electron. The standard equations to represent these two reactions are

$$
\gamma + e^- \rightarrow e^- + e^- + e^+ \qquad \text{and} \qquad \gamma + p \rightarrow p + p + \bar{p}
\tag{9.39}
$$

p-p collisions

A proton moving at high speed can collide with another proton and create more particles. Cosmic rays are the high energy protons described briefly on page 303, and their collisions with atmospheric nuclei produce showers of new particles. Look at the simplest case, where two new particles are created: $p + p \rightarrow p + p + p + \bar{p}$. How much energy does it take to allow this? The same ideas of energy and momentum conservation apply here too.

The picture of this process, again at the minimum possible energy is like the one in the preceding two paragraphs, except for the addition of one more particle of momentum $\vec{p}'$ on the right and the fact that the photon on the left is now a proton of momentum $\vec{p}$. What changes in the equations (9.37)?
1. The first equation, $E = pc$, is replaced by Eq. (9.31).
2. In the second and third equations "3" becomes "4".
When you solve the equations, following the same procedure, the result is $E = 7mc^2$, so the required initial kinetic energy is $6mc^2$. I leave the algebra to you, problem 9.34.

photon emission

When an atom emits a photon, both energy and momentum are conserved. When a nucleus emits a photon, the same applies. The difference is that in the first case it's called light and in the second case a gamma-ray. The initial state is an object of mass m and it emits a photon resulting in an object of smaller mass m_1. Is the photon's energy given by this mass difference times c^2? No, except sometimes pretty nearly yes.

Write out the details assuming that the initial system is at rest. Use f for the final state of the atom and γ for the photon

$$\text{initial state: } E = mc^2, \quad p = 0 \qquad \text{final state: } E = E_\mathsf{f} + E_\gamma, \quad p = -p_\mathsf{f} + p_\gamma$$

Write the conservation equations together with the general energy-momentum relations in the first column. The manipulations of these four equations form the second column.

$$
\begin{aligned}
mc^2 &= E_\mathsf{f} + E_\gamma \\
0 &= -p_\mathsf{f} + p_\gamma \\
E_\mathsf{f}^2 - c^2 p_\mathsf{f}^2 &= m_1^2 c^4 \\
E_\gamma &= c p_\gamma
\end{aligned}
\quad \longrightarrow \quad
\begin{aligned}
E_\mathsf{f}^2 - \left(c p_\gamma\right)^2 &= m_1^2 c^4 \\
E_\mathsf{f}^2 - E_\gamma^2 &= m_1^2 c^4 \\
\left(mc^2 - E_\gamma\right)^2 - E_\gamma^2 &= m_1^2 c^4 \\
m^2 c^4 - 2mc^2 E_\gamma &= m_1^2 c^4
\end{aligned}
\quad \longrightarrow \quad
E_\gamma = \frac{(m^2 - m_1^2)c^2}{2m}
$$

If the mass change is very small, $m - m_1 \ll m$, then

$$E_\gamma = \frac{(m + m_1)(m - m_1)c^2}{2m} \approx \frac{(2m)(\Delta m)c^2}{2m} = \Delta m c^2$$

This case is typical of the emission of light from an atom. The emitted photon has the energy determined solely by the mass difference between the initial and the final state, and the recoil energy of the atom is usually negligible. In the case of photon emission from nuclei—gamma rays—the same approximation is usually still good, though in delicate measurements it can be quite significant. The energy of gamma rays can be a million times the typical light radiation from atoms, but it is still small compared to the mc^2 of the nucleus. At the opposite extreme, starting with positronium and end with two photons, $m_1 = 0$ and

$$E_\gamma = \frac{m^2 c^2}{2m} = \frac{1}{2} m c^2$$

In this case there *is* no final atom, only another photon, so the recoil carries half the energy.

Colliders

In the collision calculations presented here, there is a massive inefficiency at high energies. *E.g.* in p-p collisions you need $6mc^2$ of kinetic energy just to create two new particles of total mass-energy $2mc^2$. That happens because the process has to conserve momentum, and that takes up a lot of the available energy. At still higher energies with still more complicated processes it only gets worse. The largest particle accelerators in the world (LHC in Geneva, Tevatron near Chicago) circumvented this problem by turning the laboratory into a "zero-momentum frame". Fire two protons at each other with the same speed but opposite direction, and the reaction $p + p \to p + p + p + \bar{p}$ requires only $2mc^2$ of kinetic energy to create the two new particles, not 6. The total momentum here is zero, so it involves no kinetic energy.

Having particles moving at nearly the speed of light, being aimed at each other and expecting a collision, is a problem. Much easier would be to have two baseball pitchers throwing fast balls toward each other and asking that the balls collide in midair. Still, it is done (with protons anyway).

9.10 Yarkovsky Effect

The concept of this phenomenon is simple, but it took a clever civil engineer, working in his spare time, to think of it. Light can affect the motion of small asteroids. Sunlight carries momentum, so it will apply a force directly on any object in its path, pushing it radially outward. Once you know that mass zero particles (photons) carry momentum, you pretty much expect it, but the Yarkovsky effect is a different and more surprising one. As an asteroid orbits the sun and rotates on its axis, it will be heated during its day and cooled during its night.
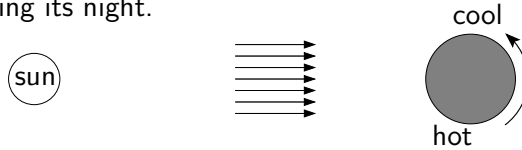
All objects radiate, and a hotter object radiates more. Radiation carries momentum ($p = E/c$), so the hot part of an asteroid will be emitting more momentum than the other parts of its surface. Conservation of momentum implies that this provides a net force on the asteroid in the direction away from the hot part of the surface. Surely that is negligibly small! No. It has been measured for some asteroids, and the important point is that even though it is small, it is unceasing. Over millenia, or even decades for smaller asteroids, it becomes a significant effect. For still another surprising effect in the same spirit as this, see section 9.13 for the Poynting-Robertson effect.

For a hot surface, how much momentum is emitted from it? The power per area of the radiation emitted from a hot surface is

$$\frac{dP}{dA} = \frac{dE}{dt\,dA} = \varepsilon\sigma T^4$$
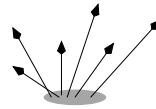
Here, σ is the Stefan-Boltzmann constant and ε is the emissivity of the surface. For an ideal black-body surface, $\varepsilon = 1$, but most surfaces fall well short of that. $\sigma = 5.67 \times 10^{-8}\,\text{Wm}^{-2}\text{K}^{-4}$.

This power is radiated in all directions above the surface, but in any given direction the amount received by a detector will vary with the angle at which it sees the emitting area. If you look at a piece of white paper straight on and then hold it at an angle so that it is almost edge on, the amount of light that enters you eye from the paper will be very different. It is proportional to the solid angle* that the paper subtends at your eye. That in turn is proportional to the cosine of the angle from the normal to the surface.

$$dP = \frac{1}{\pi}\varepsilon\sigma T^4 \cdot dA_{\text{source}} \cos\theta \cdot d\Omega$$

$$d\Omega = \frac{dA_{\text{detector}}}{r^2} = \sin\theta\,d\theta\,d\phi$$
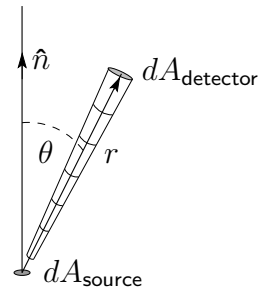
How much momentum does this element of area (dA_{source}) emit? For each chunk of energy dE that it emits, there is a momentum of magnitude $dp = dE/c$. The vector sum of all the $d\vec{p}$ is along the normal to the area, so

$$dF_n = \hat{n} \cdot \frac{d\vec{p}}{dt} = \int \frac{1}{c}dP\cos\theta = \int_{\theta=0}^{\theta=\pi/2} \frac{1}{\pi c}\varepsilon\sigma T^4 \cos^2\theta \sin\theta\,d\theta\,d\phi\,dA_{\text{source}} = \frac{2}{3c}\varepsilon\sigma T^4 dA_{\text{source}}$$

---

* See Mathematical Tools, section 8.13

The part of the asteroid facing the sun gets hot. That hot region then rotates so that the side at right angles to the sun is still hot, and it radiates and cools. It keeps cooling until it gets all the way around to face the sun again. That cooler side radiates much less ($\propto T^4$ recall), so the radiation on the hotter side wins the battle and pushes the asteroid sideways in its orbit. This is one of the many effects that makes the tracking of near-Earth asteroids far more complicated than a naïve analysis would imply.

Pioneer Anomaly
A variation on this effect is the most likely explanation for a long-standing puzzle in space flight. In 1972 and 1973 the Pioneer spacecraft were sent on long voyages to survey the outer planets, which they did extraordinarily well. They are now headed out of the Solar System, and are far beyond Pluto. The orbits have been carefully tracked, and they show the odd behavior that, even after taking account of all the known forces, the ships have an additional accelerations of almost 9×10^{-10} m/s^2. It doesn't sound like much, but it is well beyond the plausible errors from known causes. A recent (publ: PRL July 2012) explanation that fits this data quite well is that it is an effect like Yarkovsky's: heat radiation with resulting recoil from the emitted photons. The heat in this case comes from plutonium-238 that is present as a power source for the craft, and detailed calculations imply that the exterior heating is not uniform over the surface of the craft. There is enough asymmetry to account nicely for the observed anomalous acceleration, resolving this decades-old question.[*]

9.11 Conservation Laws[.]
How to derive these equations for energy and momentum? There is a famous paper by Eugene Wigner (more often cited than read) that provides the most complete, rigorous, and thorough analysis of this question and of many closely related questions. It is also *way* beyond anything I'm prepared to do here.

Instead of trying to demonstrate a general solution to the question, I will stay in one dimension and try to address the question of what sort of conservation laws are even *possible*. What constraints does relativity place on them. To do this, the whole transformation apparatus won't be needed, only the velocity addition law.

$$\frac{v - u}{(1 - vu/c^2)} = v' \quad \longleftrightarrow \quad \omega_v - \omega_u = \omega' \qquad \text{Eqs. (9.13), (9.30)}$$

Look back at the derivation leading to Eq. (1.16) for the idea behind this analysis. The basic question will be: What sort of conservation equation will work for both a stationary and a moving observer? Examine the collision of two masses m_1 and m_2 with velocities v_1 and v_2. After the collision they have masses m_3 and m_4 and velocities v_3 and v_4. Can masses change? Of course. In a hard impact a piece of m_1 can chip off and get stuck to the other mass, and in the collisions of elementary particles, new particles can be created.

$$m_1, \vec{v}_1 \qquad m_2, \vec{v}_2 \qquad \qquad m_3, \vec{v}_3 \qquad m_4, \vec{v}_4$$
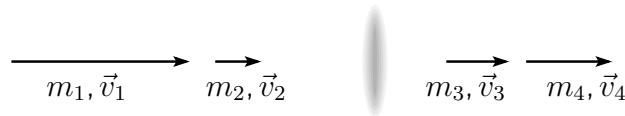
In the left picture two particles are about to collide, and the right picture is after the collision.

Non-relativistic Conservation
Before attacking the relativistic problem, go back to classical Newtonian mechanics and see what happens there. Instead of the Lorentz transformation it will be the Galilean transformation, the first

---

[*] A general interest book on this anomaly and why it was considered important is "The Pioneer Detectives" by Konstantin Kakaes. And there's also Wikipedia.

equation in this chapter and the first equation in chapter five. If someone is moving at velocity u with respect to you, those coordinates are

$$x' = x - ut \quad \longrightarrow \quad \frac{dx'}{dt} = \frac{dx}{dt} - u \quad \longrightarrow \quad v' = v - u$$

This is the ordinary equation for relative velocity that you are most familiar with. A conservation law for this collision will be of the form

$$m_1 f(v_1) + m_2 f(v_2) = m_3 f(v_3) + m_4 f(v_4) \tag{9.40}$$

where f is some function yet to be determined. You expect it to describe energy or momentum, but don't assume that.

When someone is trying to catch up with the motion, chasing with velocity u after the particles, this will be

$$m_1 f(v_1 - u) + m_2 f(v_2 - u) = m_3 f(v_3 - u) + m_4 f(v_4 - u) \tag{9.41}$$

Now differentiate with respect to u and then set $u = 0$.

$$-m_1 f'(v_1) - m_2 f'(v_2) = -m_3 f'(v_3) - m_4 f'(v_4)$$

This immediately says that if $mf(v)$ is conserved, then so is $mf'(v)$. That in turn implies that $mf''(v)$ is conserved. ... Does this mean that there are an infinite number of conservation laws? If this happens then you're in trouble because the sole solution to this system of many equations in four unknowns (v_1, v_2, v_3, v_4) is that no collision can occur, and the right side of Eq. (9.40) must identically match the left side. Nothing can happen.

If you have four unknowns (or fewer) on the right, you can have only four or fewer independent equations for them. That means that this method of generating new conservation laws by differentiation must come to a stop. There are various ways that can happen. Either
1. the sequence terminates in a constant so that the next step gives zero, or
2. the sequence repeats.
3. the masses are zero, making Eq. (9.41) useless. But, remove the m's, leaving the f's alone and you have a possibility. The preceding two options can then be reexamined.

Case 1: The last "f" in the sequence is a constant. Make it $f \equiv 1$. The f just before that is v and the one before that is v^2. (Look familiar?) Could there be a v^3 or v^4?

$$f = 1 \quad \longrightarrow \quad m_1 + m_2 = m_3 + m_4 \qquad\qquad f = v^2 \quad \longrightarrow \quad m_1 v_1^2 + m_2 v_2^2 = m_3 v_3^2 + m_4 v_4^2$$
$$f = v \quad \longrightarrow \quad m_1 v_1 + m_2 v_2 = m_3 v_3 + m_4 v_4 \qquad f = v^3 \quad \longrightarrow \quad m_1 v_1^3 + m_2 v_2^3 = m_3 v_3^3 + m_4 v_4^3$$

The first is conservation of mass though it does not require that $m_3 = m_1$ or $m_3 = m_2$. The second and third are the familiar conservation of momentum and kinetic energy. The fourth is an impossibility. If all four of these held, you would have four equations for four unknowns, and all of the factors in the final state would be determined *a priori.* The obvious solution is all there is — nothing happens and all the initial masses and velocities come out unchanged. Of course in a particular instance not all of these laws may hold. You *can* have an inelastic collision for which kinetic energy is not conserved, then only the two conservation laws, p and m, remain. The relativistic version of this is even more tightly constrained.

Case 2. You have an equation such as $f' = \alpha f$ or $f'' = \alpha f$ or $f''' = \alpha f$ or maybe even $f''' = \alpha_1 f'' + \alpha_2 f$. These have familiar exponential solutions such as e^v or the like. But wait! The expression e^v makes no sense dimensionally. What is the exponential of $5 \, \text{m/s}$? To make this work you

would need an additional constant to handle the units, something such as e^{v/v_0}. Classical Newtonian mechanics has no such built-in speed parameter, so you can't have such a conservation law, which brings us to

Relativistic Conservation,
which does. The procedure for the relativistic case is the same, but you may get involved with a lot more algebra if you don't plan ahead. If you use the variables v_1, v_2, *etc.*, you will fill the page(s) with algebra. Use rapidity instead, and remember the equation (9.30), $\omega_{\text{total}} = \omega_1 + \omega_2$, that describes velocity addition (or in this case subtraction).

The proposed conservation law is in the form

$$m_1 f(\omega_1) + m_2 f(\omega_2) = m_3 f(\omega_3) + m_4 f(\omega_4)$$

If someone chases after these particles with a rapidity ω_0 this becomes

$$m_1 f(\omega_1 - \omega_0) + m_2 f(\omega_2 - \omega_0) = m_3 f(\omega_3 - \omega_0) + m_4 f(\omega_4 - \omega_0) \tag{9.42}$$

This looks exactly the same as Eq. (9.41) except for changing the symbols. That implies that all the mathematics that follows is the same too. There are then three possible types of conservation laws with possible forms for f being either ω^2, ω, 1, or the sort of exponentials that result from solving equations such as $f'' = f$, or zero mass.

This time the first case is not possible because we know that mass is not conserved. There are high energy reactions such as

$$e^- + e^+ \longrightarrow p + \bar{p}$$
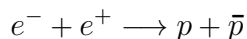
in which a positron hits an electron and converts to a proton and an anti-proton. If $f \equiv 1$ is disallowed, then its precursors ω and ω^2 can't be valid either. The second case is now possible however, because there *is* a built-in speed parameter — c.

In the same way that there can be no more than three conservation laws in the non-relativistic case, there can be no more than two here. From conservation laws alone you can't determine anything about the final masses, because experimentally you can have anything from a totally elastic collision to something such as the $e^- e^+ \to p\bar{p}$ reaction above to its reverse, $p\bar{p} \to e^- e^+$. This means that f'' is not independent of f and f'.

Could you have a differential equation such as $f'' = f + f'$? No. If conservation laws hold, then they should hold just as well if you reversed the coordinate system, changing every ω to $-\omega$. This differential equation doesn't satisfy that criterion. The f' term behaves opposite from the other two. The only linear, constant coefficient, second order differential equation that works is $f'' = \pm \alpha^2 f$ for some α. The solutions are then respectively even and odd. Why not first order? You expect two conservation laws not one, because in the non-relativistic limit that is what you have.

$$f'' = \pm \alpha^2 f \quad \longrightarrow \quad \begin{array}{ll} (+)\ f(\omega) = \cosh \alpha\omega, & f'(\omega) = \alpha \sinh \alpha\omega \\ (-)\ f(\omega) = \cos \alpha\omega, & f'(\omega) = \alpha \sin \alpha\omega \end{array} \tag{9.43}$$

The circular functions are unexpected; are they possible? They would have the feature that as one increases the other decreases, and in the non-relativistic limit nothing like that occurs. On that basis alone, look at the other equations.

Are there any restrictions on α? No, it's completely arbitrary and I can make it anything that I want. Back when the rapidity ω appeared in the equations (9.25) and (9.28) it could just as easily have been $v/c = \tanh 2\omega$ and nothing would have changed. In the former instance ω was a change of variable in an integral, and the result of an integration can't depend on your choice of how to do

it. In the second instance a change of ω to $\alpha\omega$ would just be carried through the next few equations intact, accomplishing nothing. So, just make it $\alpha = 1$ and be done with it. This is like choosing to work in radians when you use circular functions. You could work in degrees, but why carry along the extra factor of $\pi/180$?

One constraint is the non-relativistic limit. If $p = A\sinh\omega$ and $E = B\cosh\omega$, then for small $\omega \approx v/c$,

$$p \approx A\omega \approx Av/c \qquad \text{and} \qquad E \approx B[1 + \omega^2/2] \approx B[1 + v^2/2c^2]$$

This p should match mv and the non-constant part of E should match $mv^2/2$, so

$$A/c = m \qquad \text{and} \qquad m = B/c^2 \longrightarrow p = mc\sinh\omega, \quad E = mc^2\cosh\omega$$

These are precisely the equations (9.35).

The quotient of these equations relates E to p without m:

$$\frac{p}{E} = \frac{mc\sinh\omega}{mc^2\cosh\omega} = \frac{1}{c}\tanh\omega \tag{9.44}$$

For a photon, moving at the speed of light ($\omega \to \infty$), this says that $E = cp$. and that is just that third allowed way to have a consistent conservation law. There is no mass in the equation. Of course, it is an experimental question whether the photon has zero mass or not. The experimental upper limit on its mass is something like 10^{-20} electron masses, and some claim even tighter bounds.

Can you get this $E = cp$ relation directly, without taking a limit as $m \to 0$? Yes, but this time, remove the m's from the conservation equation and then chase after the particle as in Eq. (9.42). If you allow the equation $f'' = \alpha f$ as before, you get into a tangle. What then if $f' = \alpha f$? The solution to this is $f = e^{\alpha\omega}$. This says that the same Eq. (9.42) has $e^{\alpha\omega_0}$ as an overall factor. With $\alpha = -1$ this time, this is the Doppler factor! (See problem 9.25.) The two conservation laws are the same as you get by starting from Eq. (9.44). Notice also that this analysis implies that chasing after a photon makes the energy change by exactly the same factor ($e^{-\omega}$) as does the frequency. This has enormous implications in quantum mechanics.

Does this analysis show that *anything* is conserved? No. It does show that you don't have very much choice in the form of conservation laws if they do exist. There are other ways to obtain these results, but the real one, the derivation that nails it down so it is inescapable, is the 1939 paper by Wigner. (Good luck!)

In the non-relativistic case, you could have conservation of momentum in an inelastic collision, for which kinetic energy is not conserved. Here, because $f'' = f$, you can go in either direction, getting conservation of energy out of conservation of momentum. $\frac{d}{d\omega}\cosh\omega = \sinh\omega$, but $\frac{d}{d\omega}\sinh\omega = \cosh\omega$ too, so you can't have one without the other.

9.12 Energy-Momentum transformations

What happens to energy and momentum when switching to a moving observer? Answer: They transform using the Lorentz transformation, just as t and x do. You can show this using the same method as in Eq. (9.43), changing the rapidity and using the trigonometric identities in Eq. (0.11).

A mass m has rapidity ω_1, so its energy and momentum are

$$E = mc^2\cosh\omega_1, \qquad p = mc\sinh\omega_1$$

Someone is chasing it with rapidity ω, so that person will see that the mass has rapidity $\omega_1 - \omega$ and that the energy is (problem 0.15)

$$E' = mc^2\cosh(\omega_1 - \omega) = mc^2\big[\cosh\omega_1\cosh\omega - \sinh\omega_1\sinh\omega\big]$$
$$= E\cosh\omega - pc\sinh\omega$$

Similarly he sees that its momentum is

$$p' = mc \sinh(\omega_1 - \omega) = mc\big[\sinh\omega_1 \cosh\omega - \cosh\omega_1 \sinh\omega\big]$$
$$= p\cosh\omega - \tfrac{1}{c}E\sinh\omega$$

These are identical to the Lorentz transformation equations (9.29), and you should compare this deriva-
tion to the derivation of coordinate rotations in Eq. (9.4). To allow for more general expressions, put
in all the x, y, and z coordinates. Take the transformation velocity along x as usual.

$$
\begin{aligned}
ct' &= ct\cosh\omega - x\sinh\omega & & & E' &= E\cosh\omega - cp_x\sinh\omega \\
x' &= x\cosh\omega - ct\sinh\omega & x &\longleftrightarrow cp_x & cp'_x &= cp_x\cosh\omega - E\sinh\omega \\
y' &= y & ct &\longleftrightarrow E & cp'_y &= cp_y \\
z' &= z & & & cp'_z &= cp_z
\end{aligned}
\tag{9.45}
$$

  The invariant space-time interval $\Delta t^2 - \Delta x^2/c^2 = \Delta\tau^2$ from Eqs. (9.17) and (9.18) is precisely
reflected in the equation $E^2 - p^2c^2 = m^2c^4$ from Eq. (9.31). With more coordinates, y and z behave
just as p_y and p_z do, being unchanged by a Lorentz transformation along x, and $E^2 - c^2\vec{p}^2$ is still
invariant. With all these c's floating around, you can see why people who do this sort of thing for a
living commonly pick the units so that c is one.

Example
▶ Light is aimed in the $+y$-direction and you are moving in the $+x$-direction. According to you, what
direction is the light coming from? (If you are in a moving car and see someone throw a rock straight
across the road, what direction does it appear to move according to you?)
  A stationary observer (1) says that you have velocity v along $+x$ and the
photon has energy E and momentum $p_y = +E/c$, with $p_x = p_z = 0$. Now to
determine what *you* (2) should expect to see. First of course, you see that you are
at rest (with respect to yourself). For the photon, apply the Lorentz transformation
equations, Eq. (9.45). Here I have to be careful about the notation, because this
is not just one dimension. For the photon, the components of momentum in the
two systems are
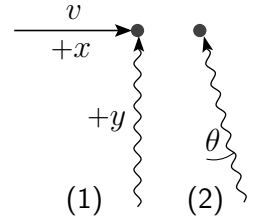
Fig. 9.12

$$
\begin{aligned}
\text{(1) stationary observer:} \quad & p_x = 0, \qquad cp_y = +E \\
\text{(2)} \quad\text{according to you:} \quad & E' = E\cosh\omega - cp_x\sinh\omega = E\cosh\omega \\
& cp'_x = cp_x\cosh\omega - E\sinh\omega = -E\sinh\omega \\
& cp'_y = cp_y = E
\end{aligned}
\tag{9.46}
$$

Check: is the energy momentum relation still correct?

$$E'^2 - c^2p'^2_x - c^2p'^2_y = (E\cosh\omega)^2 - (E\sinh\omega)^2 - E^2 = 0 \tag{9.47}$$

This is just what it is supposed to be for a massless particle. From what direction is the photon coming
according to you?

$$\frac{-p'_x}{p'_y} = \tan\theta = \frac{E\sinh\omega}{E} = \sinh\omega$$

It is coming from an angle such that it appears to be from a direction slightly in front of you. For
small speeds, $\sinh\omega \approx \tanh\omega = v/c$, and for the case of the Earth's orbital speed around the sun
this is $v/c = 30/300000 = 10^{-4}$. Convert this from radians to degrees to get $0.0057° = 21$ seconds of

angle, and stars will appear to shift their position back and forth by as much as this amount every six months, though not all at the same time because they aren't all in the same direction. What does this do for the star Polaris? For precise astronomical work, 21 seconds is huge. It is called stellar *aberration* of light, and it is large enough that it was observed as far back as the late 1600s. That was in the time that the parallax of stars had not yet been measured, and early attempts to make those measurements were confused by the existence of this effect. You can see why because the parallax of even the nearest star to our Sun is less than one second of angle. This stellar aberration however did provide a way to estimate the speed of light to respectable accuracy. That was two centuries before Einstein, but for this calculation a non-relativistic analysis was sufficient. At the other extreme, at very high speeds, $\omega \gg 1$ and $\theta \to \pi/2$; the light will then appear to you to be coming from almost straight ahead. See problem 9.43.

9.13 Poynting-Robertson Effect
When light is emitted from the Sun, it is moving radially outward. Well, almost radial, because the Sun is not a point, and when you're close to it there is a spread of angles proportional to the Sun's diameter. The effect I want to describe here shows that even if you neglect this small spread of angles and assume that the light is moving purely radially, then a dust particle orbiting the Sun will experience a tangential force from the light. This will slow the dust and cause it to fall inward. This effect sounds as if it is trivial and negligible, but it is important in understanding a major observation about our Solar System: Except for planets, comets, and asteroids, it is mostly empty. There's little dust. *Why?*

In Eq. (9.46) you have an equation stating the x-component of the momentum of a photon as seen by a moving observer: $p'_x = -(E/c)\sinh\omega$. Now the moving observer is a speck of dust orbiting the Sun, and in the frame of the dust—that's frame (2) in the sketch—the outward moving photon will appear to have this component of momentum against the direction of the orbital motion. That provides a force slowing the orbiting dust particle.

For each photon of energy E that hits the particle, Eq. (9.46) says that the x-component of its momentum will have magnitude $(E/c)\sinh\omega$, and this $\approx Ev/c^2$ for the orbital speeds that we are dealing with. Let $P_\odot$ be the total power radiated by the Sun. The distance to the sun is r and the radius of the particle is R. In time Δt an energy ΔE hits the particle, and that is

$$\Delta E = P_\odot \Delta t \cdot \frac{\pi R^2}{4\pi r^2} \quad \longrightarrow \quad \text{momentum transfered} = \Delta p_x = \Delta E \cdot \frac{v}{c^2}$$

The momentum per time that is transferred to slow the dust is then the retarding force:

$$\frac{\Delta p_x}{\Delta t} = F_{\text{PR}} = P_\odot \cdot \frac{\pi R^2}{4\pi r^2}\frac{v}{c^2} \tag{9.48}$$

How does this compare to the gravitational force from the Sun, GMm/r^2? How does it compare to the outward force by the light, $(P_\odot/4\pi r^2)\pi R^2/c$? Do either of these two questions matter? To answer the last question first: no. Both are radial forces and don't change the (mean) speed of the orbiting particle. This new force is opposite $\vec{v}$, and it has exactly the same form as a viscous drag, tending to cause the dust to lose energy and spiral inward.

Stay with circular orbits because it makes the computations easier. The basic equations are then

$$F_{\text{grav}} = \frac{GMm}{r^2} = m\frac{v^2}{r} \qquad \text{and} \qquad E = \frac{1}{2}mv^2 - \frac{GMm}{r} = -\frac{GMm}{2r}$$

A tangential force such as Eq. (9.48) gives a power $\vec{F} \cdot \vec{v} = -Fv$, so

$$\frac{dE}{dt} = -F_{\text{PR}}\,v \quad \longrightarrow \quad \frac{d}{dt}\left(-\frac{GMm}{2r}\right) = -\frac{P_\odot}{4\pi r^2}\pi R^2 \frac{v^2}{c^2}$$

$$\longrightarrow \ +\frac{GMm}{2r^2}\frac{dr}{dt} = -\frac{P_\odot}{4\pi r^2}\pi R^2\frac{v^2}{c^2} = -\frac{P_\odot R^2}{4r^2c^2}\cdot\frac{GM}{r}$$

$$\text{rearrange} \ \longrightarrow \ \frac{dr}{dt} = -\frac{P_\odot R^2}{2mc^2 r} = -\frac{\alpha}{r} \ \longrightarrow \ r(t) = \sqrt{r_0^2 - 2\alpha t} \tag{9.49}$$

The last step was just separation of variables to solve the differential equation for $r(t)$, with r_0 being the distance from the Sun at time zero. The units check. The mass of the particle varies as the cube of its radius, R, so the factor α is proportional to one over the radius of the particle, which implies that larger particles are affected less. Good. I believe that. Put in some numbers, using the Earth's orbital distance as a starting point, and choose a plausible density and size.

$$P_\odot = 3.8\times10^{26}\,\text{Watt}, \quad \rho = 2000\,\text{kg}/\,\text{m}^3, \quad R = 10^{-6}\,\text{m}, \quad r_0 \approx 1\,\text{AU} = 150\,\text{Gm},$$

$$m = \tfrac{4}{3}\pi R^3\rho \ \longrightarrow \quad \alpha = \frac{P_\odot}{\tfrac{8}{3}\pi R\rho c^2} = \frac{3.8\times10^{26}}{\tfrac{8}{3}\pi\cdot10^{-6}\cdot2000\cdot9\times10^{16}} = 2.5\times10^{11}\ \text{m}^2/\text{s}$$

The time to drop into the Sun is $r_0^2/2\alpha = 5\times10^{10}\,\text{s} \approx 1400\,\text{year}$. You see in this last equation that the time, $r_0^2/2\alpha$, is proportional to R, so if you change the particle radius from one micrometer to one millimeter the time to reach the center becomes more than 10^6 years, which is *still* small compared to the age of the Solar System.

The early solar system is believed to have condensed from a cloud of gas and dust, so the question was: why is there so little dust now? This calculation is the beginning of understanding the solution, but there's a lot more to it. In the early stages, the dust was probably thick enough to be mostly opaque. That means that this Poynting-Robertson effect would have been suppressed except very near to the sun. Today, we see comets that leave dust behind in their orbits, and asteroids collide with each other, replenishing the dust supply. Then too, the radiation pressure from the Sun will apply a force that depends on the area of the particle, so watch out for that. You see that it is now becoming a more involved problem to analyze how much dust there should be, but this seemingly tiny relativistic effect remains central to understanding this major feature of the solar system. Look at problem 9.48 for a start.

Exercises

1 Check special cases of Eqs. (9.5) to verify they they work: $0, \ \pm\pi/4, \ \pm\pi/2, \ \pi$. Various points.

2 In Eqs. (9.5), solve for x and y.

3 Using the picture above Eqs. (9.5) and taking $\alpha = 45°$, graph the curve $x'y' = 1$. Then express the equation in terms of x and y.

4 The Earth's gravitational field is g. Express it in the units light-years/year^2.

5 Two events occur at the space-time coordinates

$$x_1 = 0, \quad t_1 = 0 \qquad \text{and} \qquad x_2 = 2400\,\text{m}, \quad t_2 = 5\,\mu\text{s}$$

Use the Lorentz transformation to determine a speed v such that either (a) the two events are simultaneous or (b) the two events occurred at the same position. Solve for v in *each* case and show that one solution is possible and the other is not. Recall: the speed of light is $c = 300\,\text{m}/\mu\text{s}$.

6 You have two masses of one kg each, and one of them is here in your hand while the other is one light-year away. The center of mass is then one half light-year away. Now quickly move the one in your hand by one meter toward the other mass and the center of mass instantly moves by one half meter. Doesn't this contradict the idea that things can't move faster than light?

7 Can positronium decay by emitting a single photon? $e^- + e^+ \to \gamma$. Write the equations for conservation of energy and of momentum and verify that this is impossible, showing that they have no solution.

8 The π^+ meson, an unstable particle, lives on the average, about 2.6×10^{-8} s (measured in its own frame of reference) before decaying.
(**a**) If such a particle is moving with respect to the laboratory with a speed of 0.8c, what lifetime is measured in the laboratory?
(**b**) What average distance, measured in the laboratory, does the particle move before decaying?

9 A photon of energy E is emitted by an atom of mass m, which recoils in the opposite direction. Assuming the atom can be treated nonrelativistically, compute the recoil velocity of the atom. Hence, show that the recoil speed is much smaller than c whenever E is much smaller than the rest energy mc^2 of the atom.

10 A space ship if moving past you left to right. It has two clocks at the front and back, synchronized according to the ship's captain. As it moves past you, which clock has the earlier and which the later time according to you?

11 At what speed is the momentum of a particle twice as great as the result obtained from the nonrelativistic expression mv?

12 A particle is observed to go (on average) 6,000 meters in 100 μs before decaying. If its mean life at rest is 80 milliseconds, what was v/c? (assume no acceleration).

13 At what speed is the rest energy of a particle equal to its kinetic energy?

14 Two objects are moving toward you at speed v. One coming from your left and the other from your right. According to you what is their relative speed, and what is the maximum value that it can have?

15 In Eq. (9.22), look more closely at the case for which the moving observer "v" has velocity very close to $+c$ and then again when it is close to $-c$. For both of these, what does the equation say if u is close to $+c$ and to $-c$? (Four cases)

16 Expand Eqs. (9.24) and (9.25) to get t and x for small times. Eliminate ω between them to get $x(t)$ for small times.

17 Derive Eqs. (9.34) from Eqs. (9.31).

18 The non-relativistic analysis starting from Eq. (9.40) led to conservation of energy and momentum. You can have an inelastic collision with conservation of momentum but not kinetic energy, but can you have a collision with conservation of kinetic energy but not of momentum?

19 Equation (9.47) verifies one of the two energy-momentum relationships Eq. (9.31). What about the other one?

20 Draw the pictures in Figure 9.9, but now use the primed coordinate system for both.

21 Derive Eq. (9.17) using rapidity, Eq. (9.29).

Problems

9.1 In the picture of the Lorentz transformation, and in the units for which $c = 1$, show that the angle between the x and x' axes is $\tan^{-1} v$. The same for the angle between the t and t' axes.

9.2 Solve the Lorentz transformation equations as expressed in Eq. (9.29), for x and t.

9.3 On page 303 the mean lifetime of a muon was stated to be $\tau = 2.2\,\mu\text{s}$. That means that starting with N_0 muons, the number of muons as a function of time is $N(t) = N_0 e^{-t/\tau}$. Perhaps you are more familiar with the half-life, the time in which one-half of the muons decay, so that $N(t) = N_0 2^{-t/t_{1/2}}$. What is the relation between the mean life, τ, and the half-life, $t_{1/2}$, and what is the muon's (proper) half-life? Ans: $1.5\,\mu\text{s}$

9.4 A light source at $x = 0$ is emitting light to the right at frequency $f = 1/T$ where T is the period of the wave. Draw a space-time diagram displaying the x-t plots for several wave crests. For convenience, assume that one of the crests has coordinate $x = 0$ at time $t = 0$. You are moving to the right with velocity v, so draw the picture of your motion in the same diagram. Find the time interval *according to you* at which two successive wave crest move past you. You will have to find some intersections of lines and apply the Lorentz transformation. What frequency do you receive? If $v \ll c$, then approximately what is your result? Compare it to the Doppler shift equations you've seen before for sound. Power series expansions will be appropriate. Ans: $f' = f\sqrt{(1 - v/c)/(1 + v/c)}$
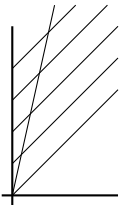
9.5 Starting at Eq. (9.11) you see how two people can each observe that the other's rulers have shrunk. Now follow the same technique and show how each can observe that the other's clock runs slow. The action this time will be along the t and t' axes instead of x and x'.

9.6 Another way to derive the length contraction factor of Eq. (9.2): just turn the light clock on its side. It will keep the same time, and that fact will determine its length. The picture shows the clock at times $t = 0$, $t = t_1$, and $t = t_2$. (Notice that it takes the light less time to return than to go; the ticks are asymmetric.) The dashed line is the coordinate $x = 0$, and the unknown length of the clock is L. You know that the time for the light to return, t_2, is $2T_0/\sqrt{1 - v^2/c^2} = 2(L_0/c)/\sqrt{1 - v^2/c^2}$. What is the x-coordinate of the light at t_1 (computed two ways)? The same question when it returns. Now eliminate t_1 and solve for L. And how does t_1 compare to t_2 and is that plausible?
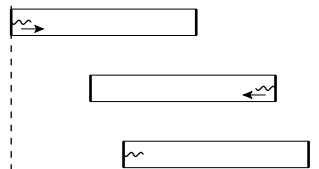
9.7 When an object of proper length L moves toward you, its length as measured by you will be only $L' = L\sqrt{1 - v^2/c^2}$. But, when you *look* at it your eye registers the light that hits your eye *at one instant*. Think of a rod aimed (almost) at your head. To reach your eye at a single specified time (maybe $t = 0$ or some other T_0) the light from the far end of the object must have left earlier than the light from the near end, and at that earlier time it was farther away. This introduces an apparent stretching of the length. How long does the object appear to be to your eye? Use a space-time diagram to assist in doing this, drawing all the paths involved, of both the front and back ends of the object and the light (think of the photons) from both. Remember that when you draw the front end of the object, that is a line in the space-time graph. The back end of the object is another line parallel to the first. Ans: $L\sqrt{(1 + v/c)/(1 - v/c)}$

9.8 The muons that come from high energy cosmic ray collisions with the upper atmosphere do not all have the same energy. Let $f(\gamma) = dN/d\gamma$ describe this spread as a function of γ, $(1 < \gamma < \infty$ and

$E = \gamma mc^2$). Make the same assumption as in the text, that they all hit the atmosphere at the same height ($h = 20\,\text{km}$), move straight down, and have the mean life at rest so that $c\tau = h_0 = 0.66\,\text{km}$. What fraction of the incident particles reach the ground? Now to evaluate the integral, assume that $f(\gamma) = A/\gamma^2$ where A is a constant. First, what is A for a total of N_0 muons? What is the result if f is proportional to $1/\gamma^3$ instead? Ans: $\gamma^{-3} \implies \frac{2}{a^2}\left[1 - e^{-a}(1 + a)\right] = 0.002$

9.9 You see two events that occur at times t_1 and t_2, time interval $\Delta t = t_2 - t_1$. They are at positions x_1 and x_2 and $\Delta x = x_2 - x_1$. (**a**) If your friend is moving by at velocity v, What time interval will he measure for the two events? (**b**) If $\Delta x = 0$ and you measure the time interval as $10\,\text{s}$ and his $\gamma = 2$, what time interval does your friend measure? (**c**) If $\Delta x/c\Delta t = w < 1$, how fast must your friend be moving in order that he gets the same value of Δt that you do? Ans: $v/c = 2w/(1 + w^2)$

9.10 The expressions Eq. (9.29) look much neater than the original Lorentz equations. For an analogy in the opposite direction, express the rotation equations (9.5) in terms of the variable s instead of α: $\tan \alpha = s$
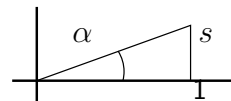
9.11 Draw the space-time figure on page 309 as seen by your friend, who of course says that you are the one moving to the left. What does your friend conclude about the total charge density that *you* see?

9.12 Analyze the problem leading to Eq. (9.3) from the viewpoint of someone riding along with the muon instead of someone standing the the Earth's surface.

9.13ˑ Your friend's automobile is too long to fit inside a garage, so you suggest that he drive it very fast so that it shrinks and then it will fit within the length of the garage. He does this, but later complains that you were wrong—the garage shrank and since he didn't fit in before, then he certainly didn't fit while the garage was shorter. Instead, he had major repair bills. (**a**) Who was right? Draw *careful* space-time diagrams of everything. To be specific when crafting your picture, say that the proper length of the car is 5 meters; the proper length of the garage is $4\frac{1}{2}$ meters; you recommend that he drive at $v = 0.6c$. (**b**) At the time (according to your friend) at which the front of the car is hitting the back of the garage, where is the door to the garage (according to him)? (**c**) Same questions, but now according to you.

9.14 An unstable star is surrounded by a spherical shell of dust at a large distance R from the center of the star. The star explodes. After a time R/c the light reaches the shell of dust and illuminates it, with the light scattering in all directions. We view this from a large distance, so what will we see? Which light do we see first? In terms of Δt, the time after the first light seen, what is the picture?
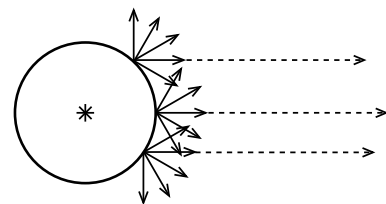
Ans: Ring or disk, radius $\left[2cR\Delta t - (c\Delta t)^2\right]^{1/2}$

9.15 Analyze the results in Eq. (9.20) thoroughly. There are many things to explore, and you *have* done the preliminary exercise 4 haven't you? Is τ_{total} really less than T? And when does the whole calculation go wrong? And if g is the gravitational field at the Earth's surface, what are the largest values that T and τ can have? What is the smallest possible ratio of τ/T over all possible accelerations and times? Is the expansion for small T useful, keeping terms out to the first correction?

9.16 (**a**) Verify that the maximum apparent speed (over all theta) in Eq. (9.21) is $v\gamma$. Also that this maximum is greater than c if $v > c/\sqrt{2}$. (**b**) For what angle does this maximum occur if β is close to one? Ans: $\theta \approx \sqrt{2(1 - \beta)}$

9.17 A star is a distance $L = 4$ lt-yrs away from Earth. A ship travels at an unknown speed v from here to there, and the pilot ages a time $T = 3$ yrs in the process. What is v? Does your algebraic result make sense if T is very large or if T is very small? Ans: $v = 0.8c$ for the numerical value.

9.18 If a star is orbiting a black hole so that the star's speed is a substantial fraction of the speed of light, what does the star's orbit *look* like to us? Assume that the orbit is a circle and that we view it edge on. (**a**) First, take the case for which the star's speed is $v \ll c$ and plot y versus t_{arrival} where the coordinates are as in section 9.5. (**b**) Now what does the plot look like if the star's speed is much larger, say $v = 0.5c$ or $0.9c$? Suggestion: As in section 9.5, get $y(t)$ and get $t_{\text{arrival}}(t)$. (Perhaps better: $y(\omega t)$, where $\theta = \omega t$ and $v/c = \omega R/c$.) Now plot both y and $\omega t_{\text{arrival}}$ versus ωt. Then you're ready to do the plot you want.

9.19 If the speed of light depended on the motion of its source the way the velocity of a ball thrown from a moving car depends on the velocity of the car, what would a binary star system look like? Look at the same case as in the preceding problem where a star is in a circular orbit around another star (taken to be fixed) and we see its orbit on edge. You don't have to assume that the star's speed is such a large fraction of c in order to discover something interesting, only that you are viewing it from astronomical distances (which of course, you are). Remember that when you want to solve an equation, a graph can help.

9.20 A generalization of problem 9.4 does not assume that the source is moving directly toward or away from you. Instead, the source of the light is moving by, and the angle in the drawing indicates the direction you have to look to see the light coming toward you. f is the source's frequency when at rest. $v\Delta t$ it the distance the source moves in (your) Δt. Check the results at $\theta = 0, \pm\pi/2$. Ans: $f' = f/(\gamma(1 - \frac{v}{c}\sin\theta))$
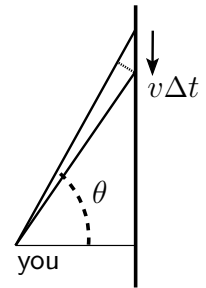
9.21 A very high energy photon can create many more than one positron-electron pair during a collision. Suppose that Eq. (9.39) is $\gamma + \text{e}^- \rightarrow \text{e}^- + N\text{e}^- + N\text{e}^+$, so that N pairs are created. What minimum kinetic energy must the incoming photon have? Show that for large enough energy the number of pairs that can be created varies approximately as the square root of the photon energy. And did you verify that your result is ok for $N = 1$? Ans: $E = 2N(N + 1)mc^2$

9.22 Derive Eqs. (9.34).

9.23 Use a series expansion to show the relation between the two expressions for $x(t)$ in Eq. (9.27).

9.24 In the graph at Eq. (9.26) what is the distance (a time really) between the hyperbola and the "$x = ct$" line, measured vertically for large time?

9.25 Take the result of problem 9.4 and express its answer in terms of the rapidity variable ω. *Don't* stop with the first expression that you get. You must simplify it, and if the result isn't *really* simple, work harder. Ans: $e^{-\omega}$

9.26 Let $d\tau$ be the proper time interval as in Eq. (9.19), then (**a**) what are $dt/d\tau$ and $dx/d\tau$? (**b**) Call them u_0 and u_1 respectively and compute the difference of their squares, putting in appropriate factors of c in order to make the units come out right. (**c**) Now what are $m\,dt/d\tau$ and $m\,dx/d\tau$?

9.27 What is the composition of two Lorentz transformations? That is, Eq. (9.10) gives (x', t') in terms of (x, t). The same equations will then provide (t'', x'') in terms of (x', t'), though with a possibly different velocity u. What then are (x'', t'') in terms of (x, t)? Ans: Eq. (9.13) (up to signs)

9.28 Do the preceding problem, but now using the form of the Lorentz transformation stated in Eq. (9.29). Ans: Eq. (9.30).

9.29 Use the results of the problem 9.26 to show directly why energy and momentum obey the same Lorentz transformation laws as time and position, Eqs. (9.10) and (9.45).

9.30 From Eq. (9.26) you can compute $dx/dt = (dx/d\omega)/(dt/d\omega)$. Show that this is $v = c \tanh \omega$.

9.31 Express the answer to part (c) of problem 9.9 in terms of rapidity and then explain why the result should be obvious. You may have to look up a hyperbolic identity.

9.32 ˙(**a**) What is the proper time as integrated along the path of constant proper acceleration in Eq. (9.26)? The ω variable makes the integration easier. (**b**) For an acceleration of $a = 1$ light-year/year$^2 \approx g$, what is the distance travelled in a proper time of 1 year? 5 years? 24 years? Ans: $\tau = c\omega/a$, $x = (c^2/a)(\cosh \omega - 1) \longrightarrow x = (c^2/a)(\cosh(a\tau/c) - 1)$. 5 yr $\rightarrow$ 73 lt-yr.

9.33 For constant proper acceleration starting from rest, what is the velocity as a function of time? Do it two ways: $dx/dt = (dx/d\omega)/(dt/d\omega)$ and directly from Eq. (9.27).

9.34 ˙Complete the setup and do the algebra to find the minimum kinetic energy required for the pair production reaction $p + p \rightarrow p + p + p + \bar{p}$ as in the discussion immediately after Eq. (9.39). What is the in terms of electron-Volts? Ans: $K_{\text{min}} = 6mc^2 \approx 6$ GeV

9.35 Two particles, each of mass m are headed straight toward each other at the same speed $v = c \tanh \omega$. In terms of ω, how fast is one moving with respect to the other? What is the energy of one as viewed from the other? What should the answer be in the non-relativistic case, and does your result agree for $v \ll c$ ($\omega \ll 1$)?

9.36 ˙(**a**) The equation (9.20) represents constant acceleration d^2x/dt^2, but what acceleration will the person taking this journey feel? That is, what is the proper acceleration for this motion? (**b**) For the case of constant proper acceleration as in Eq. (9.26), what is the acceleration d^2x/dt^2? In both cases, how does the answer behave when you push toward small or large speed?

9.37 ˙(**a**) Translate the velocity transformation equation Eq. (9.13) into the notation of rapidity. (**b**) Now try the calculation of Eq. (9.16) in terms of rapidity.

9.38 Plot the two graphs in Eq. (9.27) on the same graph and to the same scale. Take $x_0 = c^2/a$. At what point on your graph does the non-relativistic expression exceed the speed of light?

9.39 (**a**) What happens when you try to catch up to a photon? Part of the answer is Eq. (9.14), but what happens to the energy and to the momentum of a photon when you chase it? Or when it chases you? (See Eq. (9.45) for a start.) If you work this out in terms of ω, translate the result to v and vice versa. (**b**) In quantum mechanics you find that the energy of a photon is proportional to the frequency of the corresponding wave. What then is the relation between the two frequencies? This is another derivation of the Doppler effect. See problems 9.4 and 9.25.

9.40 Use the energy and momentum transformation equations to express the quantity $E'^2 - p'^2 c^2$ in terms of E and p. Since you know the value of this for $p = 0$, you now have it for all p and E.

9.41¨In chapter seven, problem 7.34, the question was to find the effect on the wave equation of changing to a moving coordinate system. The result is stated there, and it messes up the equation.

Now ask if there is another transformation, not the Galilean one, that will leave the wave equation alone. That is, try a change of variables

$$x' = \alpha x + \beta t, \qquad t' = \gamma x + \delta t$$

(**a**) And demand that the transformed equation look exactly the same as the original but in the variables x' and t' and with the same wave speed as before. (**b**) What if the transformed equation has the same form as the original, but the wave speed can be different?

9.42 An atomic nucleus with mass m emits a photon that you measure to have energy E_γ, and the nucleus recoils as a result. Start from the exact conservation of energy and momentum relations to find the mass of the recoiling nucleus. From the resulting equation, what is the largest value that E_γ can be? What does the result give when the photon energy is small (but not zero)?
Ans: $m'^2 = m^2 - 2mE_\gamma/c^2$

9.43 Someone moving to the right with velocity v emits a photon at an angle θ with respect to the forward direction. That is, θ according to him. (**a**) What is the angle that you will observe? One approach is to use the transformation equations (9.45), together with the relations for p_y and p_z stated there. (**b**) If photons are emitted with equal probability in all directions according to the moving emitter, then within what angle from the front will you detect half the photons? (**c**) If v is very close to c, approximately what is this angle? And what is it for $v/c = 0.99$? *I.e.* Hint: If $\cos^{-1}(1-x) = y$, solve this for x in the case of small y.
Ans: (a) $\cos\theta_{you} = \left(\cos\theta + \frac{v}{c}\right)/\left(1 + \frac{v}{c}\cos\theta\right)$, (b) $\theta_{1/2} = \cos^{-1}(v/c)$, (c) $\theta_{1/2} \approx \sqrt{2(1-v/c)}$, $8°$.

9.44 For a particle moving very close to the speed of light and having a total energy E, (**a**) show that the difference between its speed and that of a mass zero particle is very nearly $c - v = \frac{c}{2}\left(mc^2/E\right)^2$. (**b**) If a photon does not in fact have zero mass, but is about 10^{-20} electron masses, then in a race that started shortly after the big bang, (about 13.8 billion years ago), by how much distance will an X-ray photon beat a photon of visible light? Take the energies to be $50\,000\,\text{eV}$ and $2\,\text{eV}$. (**c**) In the same race, what wavelength photon would arrive one second behind either of these?
Ans: (b) About $\frac{1}{2}$ mm   (c) λ about 1.2 light-years

9.45 The Lorentz transformation was derived assuming that the moving observer has constant velocity. If instead, the moving observer is accelerating, you can still use a clock-radar definition of his coordinates in the same way as in Eq. (9.7). There is no change in the concepts, only a (big) change in the amount of algebra. An equation for constant proper acceleration is Eq. (9.26), and the corresponding equation showing proper time for the accelerated observer is problem 9.32. If you use the construction for a coordinate system following the same clock-radar procedure as for the standard Lorentz transformation, Eqs. (9.6)-(9.10), you get
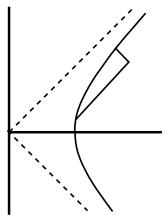
$$x = ke^{x'/k}\cosh(ct'/k) \qquad ct = ke^{-x'/k}\sinh(ct'/k), \qquad k = c^2/a$$

a is the constant proper acceleration. Here x' and t' are the coordinates assigned to an event by the accelerated observer, and (x, t) are the coordinates according to the one who stayed behind. *E.g.* $x' = 0$ is the hyperbola that represents the accelerated observer. Notice that when $t' = 0$ and $x' = 0$ you have $x = k$ and $t = 0$. That means that the two observers don't have the same origin.
(**a**) Show that the path of this accelerated person is the same as in Eq. (9.26).
(**b**) Near the apex of the hyperbola (x' near k and t' small) what are x and t?
(**c**) What do these transformation equations become when the acceleration $a \to 0$?

(**c'**) Now how do you modify the equations so that this has a more useful limit? (Change the origin.)

(**d**) Solve these equations for x' and t'. (multiply them and take the difference of squares).

(**e**) If the accelerated observer look back to see what's behind, what does he see? There are a lot of stars out there; draw the space-time paths of the stars and of the photons that leave them and reach him. In particular, if $a = g$ and if at time $t = 0$ a star at a distance 4 light-years back (at $x = -3$ light-years) blows up, what does this accelerated observer see?

9.46 ¨ Derive the equations in the preceding problem.

9.47 Use the equations of problem 9.45, and determine how far a traveller would go at constant acceleration g in (the traveller's) time 1 year, 10 years, 20 years, 30 years. Use the units of light-years for distance and years for time. The first question: What is x' for the traveller?

9.48 ¨¨¨ A small particle is orbiting the sun and its rotational angular velocity is parallel to its orbital angular velocity, show that the Yarkovsky effect will push the dust in a direction opposite to that of the Poynting-Robertson effect, and how do the sizes of these two effects compare? Can they balance?

Coupled Oscillators

.

Read section 0.10 again

In section 3.9 you saw what happened when two harmonic oscillators are coupled to each other. The result was that two different frequencies appeared in the same system along with two correspondingly different modes of oscillation. With three masses in a line there are three frequencies. If you have 10^{23} atoms bound together in a crystal and oscillating together in three dimensions, will there be correspondingly many modes of oscillation? (Yes, 3×10^{23}.) This last example is important in understanding the quantum theory of the specific heat of solids. If you skipped that part of chapter three, then it will be prudent to go back there first because it forms a gentler introduction to understanding modes of oscillation.

The first example above led to equations (3.57).
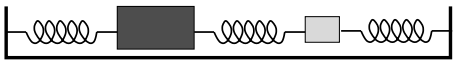
Fig. 10.1

$$m_1\ddot{x}_1 = -k_1 x_1 - k_2(x_1 - x_2)$$
$$m_2\ddot{x}_2 = -k_3 x_2 - k_2(x_2 - x_1)$$

$$(10.1)$$

Another example has three masses that are allowed to move in the plane. Take the coordinates to be measured from their points of equilibrium. That is, place these masses on a table at rest then you measure $\vec{r}_1$, $\vec{r}_2$, and $\vec{r}_3$ from these initial positions

$$m_1\ddot{\vec{r}}_1 = -k_{12}(\vec{r}_1 - \vec{r}_2) - k_{13}(\vec{r}_1 - \vec{r}_3)$$
$$m_2\ddot{\vec{r}}_2 = -k_{12}(\vec{r}_2 - \vec{r}_1) - k_{23}(\vec{r}_2 - \vec{r}_3)$$
$$m_3\ddot{\vec{r}}_3 = -k_{13}(\vec{r}_3 - \vec{r}_1) - k_{23}(\vec{r}_3 - \vec{r}_2)$$
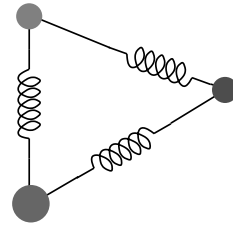
Fig. 10.2

When these are moved from their equilibrium positions and released, the motions will have various frequencies of vibration about their original position. There have six components altogether, so you could have up to six frequencies. There will actually be fewer in this case, as in this circumstance some of them are the same. Change this to three dimensions, letting the system move freely in space and you have the same equations with nine coordinates: a water molecule is an example of such a system. The forces between the atoms don't come from springs, but the basic picture and much of the mathematics is the same as in the real, quantum mechanical description.

Electric circuits provide even more complex possibilities. Self-inductance plays a role like mass. Self- and mutual-capacitances play the role of (reciprocal) spring constants. Mutual inductance has no analog in mass; it would be something like a term where the acceleration of one mass directly affects the acceleration of another *without* springs or capacitors in between. The currents in various parts of the circuits are the coordinates, the analogs of the x's.
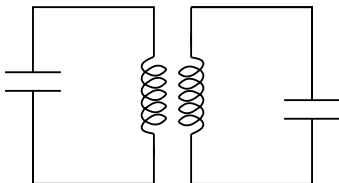
Fig. 10.3

$$L_{11}\frac{d^2 I_1}{dt^2} + \frac{I_1}{C_1} + L_{12}\frac{d^2 I_2}{dt^2} = 0$$
$$L_{22}\frac{d^2 I_2}{dt^2} + \frac{I_2}{C_2} + L_{21}\frac{d^2 I_1}{dt^2} = 0$$

$$(10.2)$$

L_{11} is the self inductance of coil one. $L_{12} = L_{21}$ is the mutual inductance of the two coils, and the capacitances are C_1 and C_2. Mutual capacitance will add an I_2 term to the first equation and an I_1 term to the second. This (L) matrix is probably not the way you've seen inductance stated before, but it is the more systematic way to write the equations. Compare problem 3.59 for the other notation.

What all of these have in common is a linear combination of the coordinates and of their derivatives adding to zero. The use of matrices provides a compact and powerful notation to do all of these problems at once. Use x_1, x_2, *etc.* for the coordinates and write these as a column matrix with N elements.

$$\mathbf{M\ddot{x}} = -\mathbf{Kx}, \qquad \text{for } N = 2, \qquad \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = - \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad (10.3)$$

For Eq. (10.1) these are

$$\mathbf{M} = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}, \qquad \mathbf{K} = \begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad (10.4)$$

and for Eq. (10.2) they are

$$\mathbf{M} = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}, \qquad \mathbf{K} = \begin{pmatrix} 1/C_1 & 0 \\ 0 & 1/C_2 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$$

The equation (10.3) compactly represents a set of linear constant coefficient differential equations, and the way to solve them is the same as if there were no matrices involved: use an exponential solution. This works because differentiating an exponential leaves it alone and it then cancels from the equations. Use $e^{\alpha t}$ or $e^{i\omega t}$ as you prefer because you expect this to oscillate. If α comes out to be real or ω turns out imaginary, either you've made a mistake or you have a surprise and the system is unstable.

The solutions of the general case or of such a specific case as Eq. (10.1) follow the same pattern.

$$\mathbf{x} = \mathbf{x}_0 e^{\alpha t}, \quad \text{then} \quad \mathbf{M\ddot{x}} = -\mathbf{Kx} \quad \text{is} \quad \mathbf{Mx}_0\alpha^2 = -\mathbf{Kx}_0 \implies (\alpha^2\mathbf{M} + \mathbf{K})\mathbf{x}_0 = 0$$

Here, $\mathbf{x}_0$ is a constant column matrix, and the final one of these equations says that the matrix $\alpha^2\mathbf{M}+\mathbf{K}$ on a non-zero vector gives zero. That implies that this matrix is singular; its determinant is zero, as in Eq. (3.59) or (0.55) or for that matter Eq. (3.1). *That* provides the algebraic equation with which to determine the α's. The examples here use 2×2 matrices, but for an example using 1×1 matrices look at Eq. (3.2), where in that case $\mathbf{x}_0$ is A.

The particular example (10.1) is, in this notation

$$\left[\alpha^2 \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} + \begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{pmatrix} \right] \begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad (10.5)$$

This is the same as the equations (3.62), and the statement that the determinant is zero is

$$\det \left(\alpha^2\mathbf{M} + \mathbf{K} \right) = \left(\alpha^2 m_1 + k_1 + k_2 \right) \left(\alpha^2 m_2 + k_2 + k_3 \right) - k_2^2 = 0 \qquad (10.6)$$

This is a quadratic equation in α^2, but it's messy. If there are three masses instead of two, even if they're oscillating along a straight line as these are, you will get a cubic equation instead of a quadratic. Unless there's some special symmetry that allows a simplification, these problems rapidly becomes intractable. Fortunately, there often are such symmetries.

For the circuit in Eq. (10.2) the corresponding equations are

$$\left[\alpha^2 \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} + \begin{pmatrix} 1/C_1 & 0 \\ 0 & 1/C_2 \end{pmatrix} \right] \begin{pmatrix} I_{10} \\ I_{20} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{10.7}$$

and

$$\left(\alpha^2 L_{11} + 1/C_1 \right) \left(\alpha^2 L_{22} + 1/C_2 \right) - \alpha^4 L_{12}^2 = 0$$

There are some general theorems about these equations that will apply to any such system, and they provide a guide to looking for a solution. First though I'll require that the matrices **M** and **K** satisfy certain properties:

M and **K**   are real and symmetric, $M_{ij} = M_{ji}$, $K_{ij} = K_{ji}$

M and **K**   are positive definite: $\tilde{\mathbf{x}}^* \mathbf{M} \mathbf{x} > 0$ and $\tilde{\mathbf{x}}^* \mathbf{K} \mathbf{x} > 0$ for all non-zero **x** $\tag{10.8}$

You can check, in the special case coming from Eq. (10.1), that the symmetry properties are satisfied. The *positive definite* requirement is needed to be sure that both the kinetic energy and the potential energy are positive. If this doesn't hold, then the oscillator is unstable. Of course if you're trying to stand a pencil on its tip, that instability is exactly what you expect and it isn't wrong. Just don't expect to use the results of this chapter without modification.

First a quick review of the notation and of the pertinent parts of the subject of matrices. Look back at section 0.10 for a running start.*

$\widetilde{\mathbf{M}}$ is the transpose of **M**. $(\widetilde{\mathbf{M}})_{ij} = M_{ji}$ $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies \widetilde{\mathbf{M}} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ $\tag{10.9}$

$\mathbf{M}^*$ complex conjugates each element, and $\widetilde{\mathbf{M}}^* = \mathbf{M}^\dagger$ is the adjoint

$$\widetilde{\mathbf{MN}} = \widetilde{\mathbf{N}}\widetilde{\mathbf{M}} \qquad \text{and} \qquad (\mathbf{MN})^\dagger = \mathbf{N}^\dagger \mathbf{M}^\dagger$$

To prove the first of these equations just write out the meaning of each side.

→ Remember the summation convention! (page 13) ←

$(\mathbf{MN})_{ij} = M_{ik} N_{kj}$ and the transpose of this interchanges i and j, so

$$(\widetilde{\mathbf{MN}})_{ij} = M_{jk} N_{ki} = N_{ki} M_{jk} = (\widetilde{\mathbf{N}})_{ik} (\widetilde{\mathbf{M}})_{kj} = (\widetilde{\mathbf{N}}\widetilde{\mathbf{M}})_{ij} \tag{10.10}$$

The corresponding equation for the adjoint adds some complex conjugations among the Eqs. (10.10).

If $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$,   then $\mathbf{x}^\dagger = (x_1^* \quad x_2^*)$ $(\mathbf{Mx})^\dagger = \mathbf{x}^\dagger \mathbf{M}^\dagger$ $\tag{10.11}$

and $\mathbf{x}^\dagger \mathbf{M} \mathbf{y}$ is a number whose complex conjugate is $\mathbf{y}^\dagger \mathbf{M}^\dagger \mathbf{x}$ $\tag{10.12}$

If $\mathbf{x} = \mathbf{y}$, this says $(\mathbf{x}^\dagger \mathbf{M} \mathbf{x})^* = \mathbf{x}^\dagger \mathbf{M}^\dagger \mathbf{x}$ $\tag{10.13}$

How do you derive these equations? The same way as in Eq. (10.10).

> Before you proceed, you should explicitly write out the details of the last half dozen lines for the 2×2 case in order to see exactly what is happening. That is, do problems 10.1 and 10.2 for $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\mathbf{N} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$.

---

   * Use **M** for the whole matrix; use M_{ij} or $(\mathbf{M})_{ij}$ for the single number sitting in the i-j position: $M_{\text{(which row), (which column)}}$.

10.1 Energy

What is the energy for these equations? If you are dealing with the example of Eqs. (10.1) or (10.2) you already know the kinetic and potential energies (at least for (10.1)), though your memory of the energies in the circuit may be hazy. What are the corresponding answers for Eq. (10.3)? That presents a different sort of problem because the **M**'s and **K**'s could be almost anything.

To answer this, go back to how energy appeared in the first place. It started with a differential equation ($\vec{F} = m\vec{a}$). Then came integration with either a dx or $\cdot d\vec{r}$ or whatever the coordinates were. In the present, more general case, the coordinates form the column matrix **x**. Look back to sections 1.3, 2.3, and 4.2. In the third one in particular, the power was $\vec{F} \cdot \vec{v}$ in Eq. (4.21).

$$\vec{F} = m\vec{a} \quad \longrightarrow \quad \vec{F} \cdot \vec{v} = m \frac{d\vec{v}}{dt} \cdot \vec{v} = \frac{d\left(mv^2/2\right)}{dt} \qquad \text{Eq. (4.21)}$$

OR, writing the one-dimensional case and working backwards,

$$E = \frac{1}{2}mv_x^2 + U(x) \quad \longrightarrow \quad \frac{dE}{dt} = \frac{m}{2}\frac{d}{dt}v_x^2 + \frac{dU(x)}{dt}$$
$$= \frac{m}{2}\frac{dv_x^2}{dv_x}\frac{dv_x}{dt} + \frac{dU}{dx}\frac{dx}{dt} = mv_x a_x + \frac{dU}{dx}v_x = 0$$

The last expression is just $(ma_x - F_x)v_x = 0$.

In the present context $\vec{v}$ is replaced by $\dot{\mathbf{x}}$ and the dot product with $\vec{v}$ is multiplication by the row matrix $\dot{\tilde{\mathbf{x}}}$

What can replace $mv^2/2$ for a system of masses obeying $\mathbf{M\ddot{x}} = -\mathbf{Kx}$? The mass is now a matrix and the velocity v_x or $\vec{v}$ becomes $\dot{\mathbf{x}}$, a column matrix. You still want kinetic energy to be a scalar. There aren't very many options to construct a scalar from a column matrix and a square matrix.

$$\dot{\tilde{\mathbf{x}}}\mathbf{M}\dot{\mathbf{x}}, \qquad \dot{\mathbf{x}}^\dagger \mathbf{M}\dot{\mathbf{x}} \qquad \text{are about it.}$$

You need a column matrix on the right and a row matrix on the left. That means that there is either a tilde or a dagger there. The only difference between these two forms is a complex conjugation in the second instance. Remember however, that the use of complex exponentials for simple harmonic motion is a trick that works *only* because the differential equations involved are linear ones. An x^2 or x^3 in the equation and would prevent using complex exponentials because of the cross terms. Energy is not a linear function of the coordinates. It involves squares and that means that you *must* use the real form of the solution in these calculations. That in turn means that the transpose and the dagger become identical. It is better in this case to stay with the real notation and avoid the temptation to use complex numbers. In the one dimensional oscillator, with a solution $x = Ae^{i\omega_0 t}$, then is $\frac{1}{2}m\dot{x}^*\dot{x} = \frac{1}{2}m\omega_0^2 A^2$ the kinetic energy? (No, in case there is any doubt.)

$$\text{Try} \qquad E = \frac{1}{2}\dot{\tilde{\mathbf{x}}}\mathbf{M}\dot{\mathbf{x}} + \frac{1}{2}\tilde{\mathbf{x}}\mathbf{Kx} \qquad \text{and compute} \qquad \frac{dE}{dt} \qquad (10.14)$$

to see what happens. Would it be more satisfying to start from the equations of motion and derive energy directly from them? Probably, but this is easier.

Do the derivatives.

$$\frac{d}{dt}\left[\frac{1}{2}\dot{\tilde{\mathbf{x}}}\mathbf{M}\dot{\mathbf{x}} + \frac{1}{2}\tilde{\mathbf{x}}\mathbf{Kx}\right] = \frac{1}{2}\ddot{\tilde{\mathbf{x}}}\mathbf{M}\dot{\mathbf{x}} + \frac{1}{2}\dot{\tilde{\mathbf{x}}}\mathbf{M}\ddot{\mathbf{x}} + \frac{1}{2}\dot{\tilde{\mathbf{x}}}\mathbf{Kx} + \frac{1}{2}\tilde{\mathbf{x}}\mathbf{K}\dot{\mathbf{x}}$$

Already the middle two terms cancel because they have the common factor $\dot{\tilde{\mathbf{x}}}$ on the left of $\mathbf{M}\ddot{\mathbf{x}}+\mathbf{K}\mathbf{x} = 0$. The first and last terms require only a little manipulation. The identity in Eq. (10.11) is $(\mathbf{M}\mathbf{x})\tilde{} = \tilde{\mathbf{x}}\tilde{\mathbf{M}}$, and by assumption, $\tilde{\mathbf{M}} = \mathbf{M}$ and $\tilde{\mathbf{K}} = \mathbf{K}$. The rest of this equation then becomes

$$\frac{d}{dt}\left[\tfrac{1}{2}\dot{\tilde{\mathbf{x}}}\mathbf{M}\dot{\mathbf{x}} + \tfrac{1}{2}\tilde{\mathbf{x}}\mathbf{K}\mathbf{x}\right] = \tfrac{1}{2}\dot{\tilde{\mathbf{x}}}\mathbf{M}\dot{\mathbf{x}} + \tfrac{1}{2}\tilde{\mathbf{x}}\mathbf{K}\dot{\mathbf{x}} = \tfrac{1}{2}(\mathbf{M}\ddot{\mathbf{x}})\tilde{}\,\dot{\mathbf{x}} + \tfrac{1}{2}(\mathbf{K}\mathbf{x})\tilde{}\,\dot{\mathbf{x}} = \tfrac{1}{2}\left[\mathbf{M}\ddot{\mathbf{x}} + \mathbf{K}\mathbf{x}\right]\tilde{}\,\dot{\mathbf{x}} = 0$$

Of course, when you reduce this to a single mass it is the usual $\tfrac{1}{2}mv^2 + \tfrac{1}{2}kx^2$.

What is this energy for the mass-spring system in Eq. (10.1)?

$$\mathbf{M} = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}, \qquad \mathbf{K} = \begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{pmatrix}$$

$$E = \tfrac{1}{2}\begin{pmatrix} \dot{x}_1 & \dot{x}_2 \end{pmatrix}\begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} + \tfrac{1}{2}\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \tfrac{1}{2}\left(m_1\dot{x}_1^2 + m_2\dot{x}_2^2\right) + \tfrac{1}{2}\left((k_1 + k_2)x_1^2 + (k_2 + k_3)x_2^2 - 2k_2x_1x_2\right)$$

$$= \tfrac{1}{2}\left(m_1\dot{x}_1^2 + m_2\dot{x}_2^2\right) + \tfrac{1}{2}\left(k_1x_1^2 + k_3x_2^2 + k_2(x_1 - x_2)^2\right) \tag{10.15}$$

and this is just what you would write down by looking at the physical system directly — three springs compressing, two masses moving, and an energy in a pear tree.

The kinetic energy and the potential energy for a spring are both supposed to be positive numbers. *That* is the reason for the positive definiteness assumptions in Eq. (10.8).

10.2 Normal Modes

The various frequencies of oscillation correspond to different patterns in the way that the masses move. These are called *normal modes*. What's normal about them and are other oscillations abnormal? Here, normal means perpendicular. The different oscillations are normal to each other, and normal means that the scalar product of one mode with another mode will be zero. Which scalar product? Probably not the one you are thinking of.

Notation

There will be a lot of subscripts here, and some of them will refer to different *vectors* while others will refer to different *components* of vectors. It is easy to get them confused. There's no accepted standard way to handle this, so I'll make up my own.

> Subscripts a, b, c, or the like will represent different vectors, and the
> subscripts i, j, k, *etc.* will represent components of a particular vector.

Basic Theorem

This important result comes without solving a single differential or algebraic equation. It involves only some matrix manipulation to derive the necessary result about normal modes. For the equation $(\alpha^2\mathbf{M} + \mathbf{K})\mathbf{x} = 0$, let $\mathbf{x}_a$ and $\mathbf{x}_b$ be two non-zero solutions for two values of α found by setting the determinant to zero. For now I don't need the values of the roots α, just that they exist. Because these are linear differential equations, complex notation is allowed.

$$(\alpha_a^2\mathbf{M} + \mathbf{K})\mathbf{x}_a = 0 \qquad\qquad (\alpha_b^2\mathbf{M} + \mathbf{K})\mathbf{x}_b = 0$$

Multiply the first of these by $\mathbf{x}_b^\dagger$ and the second by $\mathbf{x}_a^\dagger$.

$$\mathbf{x}_b^\dagger(\alpha_a^2\mathbf{M} + \mathbf{K})\mathbf{x}_a = 0 \qquad\qquad \mathbf{x}_a^\dagger(\alpha_b^2\mathbf{M} + \mathbf{K})\mathbf{x}_b = 0 \tag{10.16}$$

Both of these are numbers. Take the complex conjugate of the second one and use Eq. (10.12).

$$\left[\mathbf{x}_a^\dagger(\alpha_b^2\mathbf{M} + \mathbf{K})\mathbf{x}_b\right]^* = 0 \qquad \longrightarrow \qquad \mathbf{x}_b^\dagger(\alpha_b^{*2}\mathbf{M}^\dagger + \mathbf{K}^\dagger)\mathbf{x}_a = 0 \tag{10.17}$$

Recall the assumption that both $\mathbf{M}$ and $\mathbf{K}$ are real and symmetric, so $\mathbf{M}^\dagger = \mathbf{M}$ and $\mathbf{K}^\dagger = \mathbf{K}$. That means that I can combine the equations (10.16) and (10.17) as

$$\mathbf{x}_b^\dagger(\alpha_a^2\mathbf{M} + \mathbf{K})\mathbf{x}_a = 0 \qquad \text{and} \qquad \mathbf{x}_b^\dagger(\alpha_b^{*2}\mathbf{M} + \mathbf{K})\mathbf{x}_a = 0 \tag{10.18}$$

Now subtract these equations and the $\mathbf{K}$ terms cancel.

$$\boxed{(\alpha_a^2 - \alpha_b^{*2})\mathbf{x}_b^\dagger\mathbf{M}\mathbf{x}_a = 0} \tag{10.19}$$

This is the result I was aiming for. From this equation I can deduce two useful consequences, all without solving any equations.

For the first result take the special case in which the two solutions are the same solution. $\mathbf{x}_a = \mathbf{x}_b$ and $\alpha_a = \alpha_b$.

$$(\alpha_a^2 - \alpha_a^{*2})\mathbf{x}_a^\dagger\mathbf{M}\mathbf{x}_a = 0$$

One of my assumptions (and here's why) was that the product $\mathbf{x}_a^\dagger\mathbf{M}\mathbf{x}_a > 0$. $\mathbf{M}$ is positive definite. This number can't vanish unless the $\mathbf{x}$ itself is zero. The only way that the product of two numbers can be zero is if one of the two numbers is itself zero, so this implies that the other factor

$$(\alpha_a^2 - \alpha_a^{*2}) = 0$$

That is, α^2 is real. It doesn't say if it's positive or negative, but at least it's not complex. In fact it will be negative for these equations, and that is easy to show by one further simple manipulation. Start from the same equation and this time take the product with $\mathbf{x}_a$ itself:

$$(\alpha_a^2\mathbf{M} + \mathbf{K})\mathbf{x}_a = 0 \quad \longrightarrow \quad \mathbf{x}_a^\dagger(\alpha_a^2\mathbf{M} + \mathbf{K})\mathbf{x}_a = 0 \quad \longrightarrow \quad \alpha_a^2 = -\left(\mathbf{x}_a^\dagger\mathbf{K}\mathbf{x}_a / \mathbf{x}_a^\dagger\mathbf{M}\mathbf{x}_a\right) \tag{10.20}$$

Both the numerator and the denominator are positive, and that's all it takes to make this negative. You can now write it as $\alpha_a^2 = -\omega_a^2$.

Now go back to the general case

$$(\alpha_a^2 - \alpha_b^2)\mathbf{x}_b^\dagger\mathbf{M}\mathbf{x}_a = 0 \tag{10.21}$$

The α^2's are real now so I can drop a complex conjugation. If the two α^2's are different then this time the first factor cannot be zero and that means that the second factor must be zero.

$$\text{If } \alpha_a^2 \neq \alpha_b^2, \qquad \text{then} \qquad \mathbf{x}_b^\dagger\mathbf{M}\mathbf{x}_a = 0 \tag{10.22}$$

This is a scalar product. This is the sense in which the modes are "normal". Their scalar product vanishes. The mass matrix defines a way to multiply the two vectors, $\mathbf{x}_a$ and $\mathbf{x}_b$, and it is by using this particular definition of scalar product that you get the nice theorem. Go back to the material on waves and look closely at the equations (7.53)–(7.55). In a seemingly distant context you are doing essentially the same manipulations as here. In those equations you used partial integration. Here you used matrix transposition, but there are close parallels.* Also look back to section 8.5 where you saw

---

* This parallel is much closer than you may think. The book Linear Differential Operators by Lanczos, mentioned in the bibliography page iv, explains and exploits this link.

the orthogonality of the eigenvectors of the tensor of inertia. The proof there is essentially the same as for normal modes here, just with different notation.

What happens if $\alpha_a^2 = \alpha_b^2$? Can that ever happen? In one-dimensional problems of the sort you encounter here, it's not common, but in realistic three-dimensional problems it is unavoidable. Fortunately there's a cure. You can always choose linear combinations of these ("degenerate") vectors so that they *are* orthogonal, then the orthogonality property will hold for all your eigenvectors. See for example problem 10.41.

The particular values of the parameters α^2 that come from equations such as Eq. (10.6) are called "characteristic values" or "eigenvalues"[†] of the system. The corresponding solutions for **x** are then called characteristic vectors or or more commonly, eigenvectors. In this language, the equation (10.21) states that eigenvectors corresponding to distinct eigenvalues are orthogonal.

10.3 Scalar Products

What makes this a scalar product? Is it legitimate to call $\mathbf{x}_2^\dagger \mathbf{M} \mathbf{x}_1$ a scalar product? Yes. To see why, it is necessary to understand what a scalar product *is*.

The fundamental definition of a scalar product is that it is a scalar-valued function of two vector variables and that it must satisfy a certain set of requirements. Call the function f for now, so that $f(\vec{v}_1, \vec{v}_2)$ is a scalar and f has these four properties:

$$
\begin{aligned}
&1.\ f(\vec{v}_1, \vec{v}_2 + \vec{v}_3) = f(\vec{v}_1, \vec{v}_2) + f(\vec{v}_1, \vec{v}_3)\\
&2.\ f(\vec{v}_1, \alpha \vec{v}_2) = \alpha\, f(\vec{v}_1, \vec{v}_2)\\
&3.\ f(\vec{v}_1, \vec{v}_2) = f(\vec{v}_2, \vec{v}_1)^*\\
&4.\ f(\vec{v}, \vec{v}) \geq 0, \ \text{and} \ f(\vec{v}, \vec{v}) = 0 \ \text{if and only if the vector } \vec{v} \text{ is itself the zero vector.}
\end{aligned}
\tag{10.23}
$$

$$1, 2 : \leftrightarrow \text{"linear in 2}^{\text{nd}}\text{ arg"} \qquad 3 : \leftrightarrow \text{"conjugate symmetric"} \qquad 4 : \leftrightarrow \text{"positive definite"}$$

Any function that satisfies these requirements is called a scalar product. If the scalars are all real then the complex conjugation in the third requirement is unnecessary. Does the familiar dot product satisfy them?

$f(\vec{A}, \vec{B}) = \vec{A} \cdot \vec{B} = AB\cos\theta$ is certainly a scalar.

4. $\vec{A} \cdot \vec{A} \geq 0$ and it is never zero except when $\vec{A}$ is itself zero.

3. $\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A}$

2. $\vec{A} \cdot (\alpha \vec{B}) = \alpha(\vec{A} \cdot \vec{B})$ \qquad If $\alpha < 0$ remember that $\cos(\theta + \pi) = -\cos\theta$.

1. This is the only part that takes any effort. Draw the picture of $\vec{A} \cdot (\vec{B} + \vec{C})$ assuming that everything is in a single plane. That one is not so hard to prove — simply write out the meaning of each term. With a little more thought the same picture works for the three dimensional case too.
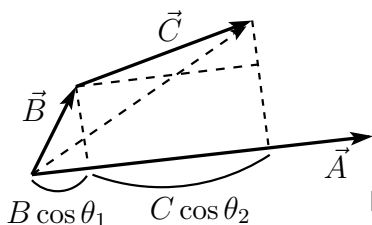
$$\vec{A} \cdot (\vec{B} + \vec{C}) \overset{?}{=} \vec{A} \cdot B + \vec{A} \cdot \vec{C}$$
$$|\vec{B} + \vec{C}|\cos\theta_3 = B\cos\theta_1 + C\cos\theta_2$$

Fig. 10.4

[†] This is an odd, sesquitranslated word. It started life in English as characteristic value, then it was translated into German as Eigenwerte, then half-translated back as eigenvalue. There it stuck.

The projections of $\vec{B}$ and of $\vec{C}$ on $\vec{A}$ sum to the projection of $\vec{B} + \vec{C}$ on $\vec{A}$. I didn't draw the angles explicitly because it would make the sketch too cluttered; you can easily figure out which angle is which.

Express displacement vectors in the plane in terms of rectangular coordinates, so that $\vec{r}_1 = x_1\hat{x} + y_1\hat{y}$ and similarly for $\vec{r}_2$, the standard scalar product is $\vec{r}_1 \cdot \vec{r}_2 = x_1 x_2 + y_1 y_2$. Suppose however that x is measured in meters and y in centimeters. This product would have to be modified to account for this, looking something like $\vec{r}_1 \cdot \vec{r}_2 = x_1 x_2 + k y_1 y_2$, where $k = 10^{-4}\,\text{m}^2/\text{cm}^2$. Go back to section 3.10 and now you can easily understand what it was about. Also, look at chapter six of Mathematical Tools for Physics for much more on the subject, with many more examples.

The example Eq. (10.22) is like this, although the mass matrix came naturally out of the computation instead of being an artifice of changing units. That (10.22) satisfies the requirements to be a scalar product (Eq. (10.23)) is easy to check, and is essentially a translation of the assumptions that I made about $\mathbf{M}$ in Eq. (10.8):

$$\text{"real, symmetric" (10.8)} \longrightarrow \text{property 3 (10.23)}$$
$$\text{"positive definite" (10.8)} \longrightarrow \text{property 4 (10.23)}$$

10.4 Initial Values

Just as in section 3.9, Eq. (3.68), you can apply initial conditions to these results, resulting in sometimes unexpected results. The matrix methods used here will organize the calculations so that they become simple to set up. (Though I make no promise that they will never be tedious.) The initial values here are conditions on $\mathbf{x}$ and $\dot{\mathbf{x}}$.

These differential equations are linear and homogeneous, meaning that the sum of solutions is a solution and a constant times a solution is one too. The key result about these special forms of matrices $\mathbf{M}$ and $\mathbf{K}$, the ones satisfying the requirements in Eq. (10.8), is that such eigenvectors "span the space". Every vector can be written as the sum of these eigenvectors. The *proof* that this is true given the stated properties of $\mathbf{M}$ and $\mathbf{K}$ is something that I will leave to appropriate math books.

Let $\mathbf{x}_a$ ($a = 1, 2, \ldots N$) be eigenvectors, so that $-\omega_a^2 \mathbf{M}\mathbf{x}_a + \mathbf{K}\mathbf{x}_a = 0$. This second order differential equation will require two initial conditions, position and velocity. Call them $\mathbf{x}_{\text{initial}}$ and $\dot{\mathbf{x}}_{\text{initial}}$. The general solution is a linear combination of all the $\mathbf{x}_a$, and this is

$$\mathbf{x}(t) = \sum_{a=1}^{N} \beta_a \mathbf{x}_a \cos \omega_a t + \sum_{a=1}^{N} \gamma_a \mathbf{x}_a \sin \omega_a t \tag{10.24}$$

where $\alpha_a^2 = -\omega_a^2$. At time $t = 0$ the initial position and initial velocities are

$$\sum_{a=1}^{N} \beta_a \mathbf{x}_a = \mathbf{x}_{\text{initial}} \qquad \text{and} \qquad \sum_{a=1}^{N} \gamma_a \omega_a \mathbf{x}_a = \dot{\mathbf{x}}_{\text{initial}}$$

These are $2N$ equations in $2N$ unknowns. Instead of brute force to solve them, simply use the orthogonality of the eigenvectors. Take the scalar product of both sides of the first of these equations with $\mathbf{x}_b$. Remember, the scalar product you need is the one that makes the vectors orthogonal, Eq. (10.22). Also, remember the notation: $\mathbf{x}_b$ is the bth eigenvector, but $(\mathbf{x})_i$ or x_i or $(\mathbf{x}_a)_i$ is the ith component of a single vector $\mathbf{x}$ or $\mathbf{x}_a$.

$$\mathbf{x}_b^\dagger \mathbf{M} \sum_{a=1}^{N} \beta_a \mathbf{x}_a = \sum_{a=1}^{N} \beta_a \mathbf{x}_b^\dagger \mathbf{M}\mathbf{x}_a = \beta_b \mathbf{x}_b^\dagger \mathbf{M}\mathbf{x}_b = \mathbf{x}_b^\dagger \mathbf{M}\mathbf{x}_{\text{initial}}$$

$$\beta_b = \frac{\mathbf{x}_b^\dagger \mathbf{M}\mathbf{x}_{\text{initial}}}{\mathbf{x}_b^\dagger \mathbf{M}\mathbf{x}_b} \tag{10.25}$$

All but one term in that sum on **a** from one to N is zero. That one term is a = b.
For the initial velocity, repeat this for the second of the initial value equations.

$$\sum_{a=1}^{N} \gamma_a \omega_a x_i = \dot{x}_{\text{initial}} \longrightarrow \gamma_b = \frac{x_b^\dagger M \dot{x}_{\text{initial}}}{\omega_b x_b^\dagger M x_b} \tag{10.26}$$

Now you have all the factors, β_a and γ_a in the equation (10.24).

Go back to the problem in section 3.9 and repeat that calculation using this formalism, going all the way through to Eq. (3.68), but in the process figuring out what **M**, **K**, and all the orthogonality properties are for this 2×2 case.

Other standard methods of solving differential equations apply here too, including the method of Green's functions as in section 3.11. This lets you add a forcing function and then easily solve for the results. Or, if not easily then at least it reduces the problem to evaluating an integral. See problems 10.23 − 10.25 and 10.35 − 10.37.

10.5 Tuned Mass Dampers

This is a very practical application of normal modes used for providing protection against earthquakes, for minimizing oscillations in the crankshaft of an automobile engine, for stabilizing tall smokestacks against buffeting caused by vortex shedding of the wind, ... more.

A very tall building will respond to an earthquake by oscillating, as every building is necessarily elastic to some extent. They are, after all, made of ordinary matter, and everything has some elasticity. A tall building can sway a disconcertingly large distance when an earthquake hits, and it becomes a very large oscillator. If any part of the earthquake's vibration is near to the natural frequency of such an oscillation there can be a resonant effect large enough to destroy the entire structure.

How do you protect against this? The simplest answer of course is not to build tall buildings. Another answer is to make them so rigid and so strong that no plausible earthquake can cause severe damage. This first is wasteful of land, and the second is wasteful of money. Then too there's always one more, still larger quake yet to come.

Is there a better way? Some clever engineers came up with the idea of placing a large mass at the top of the building and configuring it so that it can oscillate horizontally. If the building starts to sway, then this mass can be constructed so that it will move to minimize the building's motion. The details will involve some careful calculations to optimize the effect, but that's what mechanical engineers are paid to do.

To see the idea of what happens, take a simpler model, one that you now have some experience with: two masses coupled by springs. Apply an oscillating force, the earthquake, to m_1, which represents the building. The tuned mass damper is m_2, and it is attached only to m_1, so that it is not in contact with the applied force from the ground. Both masses have damping given by the terms in $\dot{x}$.
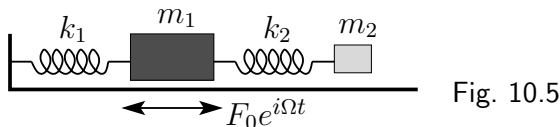
Fig. 10.5

$$m_1 \ddot{x}_1 = -k_1 x_1 - k_2(x_1 - x_2) - b_1 \dot{x}_1 + F_0 e^{i\Omega t}$$
$$m_2 \ddot{x}_2 = \qquad - k_2(x_2 - x_1) - b_2 \dot{x}_2$$

Translate this into matrix notation.

$$M\ddot{x} + B\dot{x} + Kx = f \tag{10.27}$$

$$M = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}, \qquad K = \begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix}$$

The inhomogeneous solution has $x_1 \propto e^{i\Omega t}$ and $x_2 \propto e^{i\Omega t}$. Then this equation, (10.27), becomes

$$\begin{pmatrix} -m_1\Omega^2 + k_1 + k_2 + b_1 i\Omega & -k_2 \\ -k_2 & -m_2\Omega^2 + k_2 + b_2 i\Omega \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} F_0 \\ 0 \end{pmatrix} e^{i\Omega t}$$

The solution involves nothing more than inverting a 2×2 matrix:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix} \longrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\det} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} e \\ f \end{pmatrix}, \qquad \det = ad - bc \qquad (10.28)$$

$$x_1(t) = \frac{F_0}{\det}\big(-m_2\Omega^2 + k_2 + b_2 i\Omega \big)e^{i\Omega t}, \qquad x_2(t) = \frac{F_0}{\det} k_2 e^{i\Omega t}$$

$$\det = \big(-m_1\Omega^2 + k_1 + k_2 + b_1 i\Omega \big)\big(-m_2\Omega^2 + k_2 + b_2 i\Omega \big) - k_2^2 \qquad (10.29)$$

The amplitudes of these oscillations are simply the magnitudes of the coefficients of $e^{i\Omega t}$ in each $x(t)$. Choose the parameters so that the motion of m_1 is greatly reduced, and graph the magnitude of its motion, $|x_1|$, versus Ω. In Figure 10.6 the $k_2 = 0$ graph has $x_2(t) \equiv 0$, and is equivalent to having no second mass present. The $k_2 = 0.1$ graph shows how effectively this much smaller tuned mass reduces the motion of m_1.

$$m_1 = 1 \qquad m_2 = 0.1$$
$$k_1 = 1 \qquad k_2 = 0.1$$
$$b_1 = 0.04 \qquad b_2 = 0.032$$

Fig. 10.6
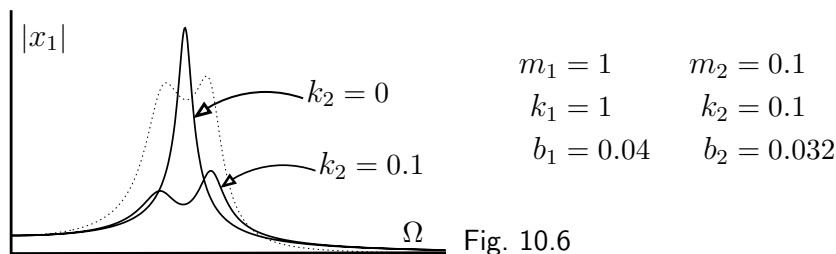
The parameters for the graph are picked simply to illustrate the effect, and in a real situation you will have to choose them much more carefully in order to achieve your goal. Then too, what is that dotted graph in the picture? It is the magnitude of the motion for the second mass, $|x_2|$, plotted to the same scale, and this graph implies that though I may have minimized the motion of the building itself, the cost is to have the mass damper swing back and forth by almost as much as the building would have swayed anyway. Maybe this is acceptable, but clearly this design requires some more analysis.

10.6 Chain of Masses·

A string is made of molecules. Many molecules. How can you model the motion of a string in terms of the motion of these constituents? Picture a line of masses m connected by very light cords. The cords are elastic and hold the masses in line, but the masses can move up and down by a small amount. Ignore gravity.

Fig. 10.7
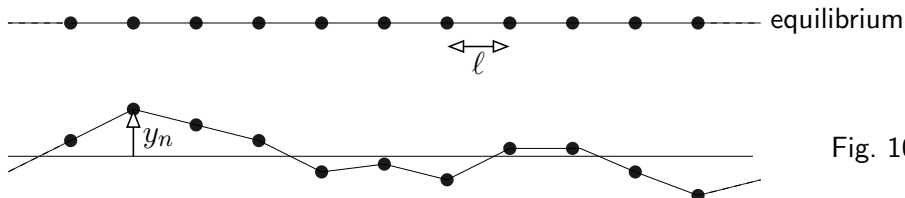
Let the equilibrium distance between the masses be ℓ, so that the x-coordinate of the nth mass is $x_n = n\ell$. The vertical coordinate of the nth mass is y_n, and as usual the departure from equilibrium is assumed to be small, with $|y_n - y_{n-1}| \ll \ell$. Let the tension in the connecting cords be T, then the y-component of force on the nth mass is

$$F_{y,n} = -T(y_n - y_{n-1})/\ell - T(y_n - y_{n+1})/\ell$$

This is just like the force calculation on a continuous string, as done in section 7.1. T is the magnitude of the force the cord exerts, and the factor multiplying it is the trigonometric $\sin\theta \approx \tan\theta$ factor giving the vertical component of the force.

This gives the differential equations

$$m\frac{d^2 y_n}{dt^2} + \frac{T}{\ell}\big(- y_{n-1} + 2y_n - y_{n+1}\big) = 0 \tag{10.30}$$

Translate this into the matrix notation, **M** and **K**.

$$\mathbf{M} = m\begin{pmatrix} \ddots & & & & \\ & 1 & 0 & 0 & 0 \\ & 0 & 1 & 0 & 0 \\ & 0 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 1 \\ & & & & \ddots \end{pmatrix} \qquad \mathbf{K} = \frac{T}{\ell}\begin{pmatrix} \ddots & & & & \\ & 2 & -1 & 0 & 0 \\ & -1 & 2 & -1 & 0 \\ & 0 & -1 & 2 & -1 \\ & 0 & 0 & -1 & 2 \\ & & & & \ddots \end{pmatrix} \tag{10.31}$$

How big are these matrices? As big as the chain. There's another, equally important question: What happens *at* the ends of the chain? How is the chain tied off? There are several possible ways to do it. The first and seemingly the simplest is to tie the ends down so that they don't move. A second is to tie the ends to massless rings, as in problem 7.28. There is still a third way, and that is to wrap the chain around in a circle and attach the leftmost mass to the rightmost mass as if it's just another mass. You have a choice among these boundary conditions.
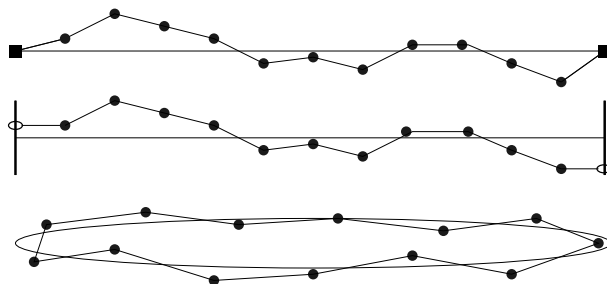
Fig. 10.8

Think of this third way as wrapping the cord around a smooth cylinder so that it is pulled taut and the masses are free to move up and down along the surface of the cylinder. If this third way looks peculiar and perhaps artificial, maybe it is, but it is also the simplest mathematically, and in real problems involving crystal structure, it's the method that is by far the most commonly used. In the study of solid state physics it is the only choice you will ever see. Of course in that case, the wrapping around is done in all three dimensions, so it's a little harder to picture, though the mathematics is no different.

There is another question that I skipped over in setting up the matrices (10.31). *What happens in the corners of the matrices where there are the diagonal dots?* How do the matrices end? This question and the choices of physical boundary conditions are really the same question in different guises.

Case 1: There are N masses with index $n = 1, \ldots, N$ and the y-coordinates y_0 and y_{N+1} are tied down to zero.

→ *Explicitly write out all three equations (10.30) for the case $N = 3$.* ←

You see, *because you just wrote out* the case $N = 3$ explicitly, that the **K** matrix is

$$\mathbf{K} = \frac{T}{\ell} \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & \text{zeros} & \\ & -1 & 2 & -1 & & \\ & & -1 & \ddots & & \\ & \text{zeros} & & & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \tag{10.32}$$

The second case is just as easy *if* you write out the $N = 3$ differential equations again. Remember now that the rings on the end, y_0 and y_{N+1} have zero mass, so they cannot have any vertical component of force. Think of these as extra coordinates 0 and 4 that obey different equations: $y_0 = y_1$ and $y_4 = y_3$

$$\mathbf{K} = \frac{T}{\ell} \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & \ddots & & \\ & & & & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix} \tag{10.33}$$

To figure out the third case, you should write out the differential equations for $N = 4$. Now, with the cord going around in a circle, the Nth (4th) mass is directly attached to the 1st mass. Again, that's like having extra coordinates 0 and 5 with $y_5 \equiv y_1$ and $y_0 \equiv y_4$.

$$\mathbf{K} = \frac{T}{\ell} \begin{pmatrix} 2 & -1 & & & & -1 \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & \ddots & & \\ & & & & 2 & -1 \\ -1 & & & & -1 & 2 \end{pmatrix} \tag{10.34}$$

All three of these **K** matrices satisfy the requirements of Eq. (10.8), though proving the positive definite requirement, $\mathbf{x}^\dagger \mathbf{K} \mathbf{x} > 0$, takes some slightly more serious knowledge of linear algebra.

Now to find the solutions to the equations of motion, and to get the normal modes: With **x** proportional to $e^{-i\omega t}$, the algebraic equations to solve are $(-\omega^2 \mathbf{M} + \mathbf{K})\mathbf{x} = 0$. (Why minus $i\omega t$? This sign is arbitrary, but its convenience will be clear within a few pages.) Trying to set the determinant of this $N \times N$ matrix to zero and solving an Nth-degree equation is more-or-less hopeless, but there's an easier way.

The easiest of the three possibilities that I've written is the third one. This uses "periodic boundary conditions", and it is simpler because in this case it is easy to use complex exponentials in the calculations, while in the other cases you will need to use trigonometric functions. This form of boundary condition may be less familiar now, but its use is so common that you should get used to it. The way to solve this system is to use a method very much like the one used to solve homogeneous, constant-coefficient differential equations—assume an exponential form. Here however the independent variable is not position or time (x or t), but the index n, and instead of using an e^x you use something more like e^n. Look at the nth line in the equation $(-\omega^2 \mathbf{M} + \mathbf{K})\mathbf{x} = 0$.

$$y_n(t) = q_n e^{-i\omega t} \quad \longrightarrow \quad -m\omega^2 q_n + (T/\ell)\big(-q_{n-1} + 2q_n - q_{n+1}\big) = 0 \tag{10.35}$$

This is of course nothing more that Eq. (10.30) for a solution proportional to $e^{-i\omega t}$. What makes the periodic boundary conditions easier to use than the others is that this equation applies to *all* the masses. There is no need for exceptions at the ends. What you do have is that the $(N+1)^{\text{th}}$ mass and the 1st mass are the same mass. (Or 0th and N^{th} — the same thing and usually easier to use.)

This is a linear, constant-coefficient *difference* equation, and you solve it using the methods so familiar with linear, constant-coefficient, *differential* equations, only here the independent variable is n instead of t. Assume an exponential solution. Is the exponential in the form p^n or $e^{i\beta n}$? Both work, and if $p = e^{i\beta}$ they are the same anyway. Anticipating that there may be waves here, I'll choose the second notation: $q_n = e^{i\beta n}$

$$-m\omega^2 e^{i\beta n} + (T/\ell)\big(-e^{i\beta(n-1)} + 2e^{i\beta n} - e^{i\beta(n+1)}\big) = 0$$
$$-m\omega^2 + (T/\ell)\big(-e^{-i\beta} + 2 - e^{i\beta}\big) = -m\omega^2 + (2T/\ell)\big(1 - \cos\beta\big) = 0 \qquad (10.36)$$

With a simple trigonometric identity, the last of these equations is

$$\omega^2 = \frac{4T}{m\ell}\sin^2(\beta/2) \ \longrightarrow\ \omega = \pm 2\sqrt{\frac{T}{m\ell}}\sin(\beta/2) \qquad (10.37)$$

Take ω to be positive always, allowing β to be positive or negative. This is a choice, but it is convenient.

How do you handle the condition that q_0 and q_N describe the same mass? It is very straight-forward, and it determines the allowed values for β.

$$q_N = q_0 \ \longrightarrow\ e^{i\beta N} = e^{i\beta 0} = 1 \ \longrightarrow\ \beta N = \text{ a multiple of } 2\pi$$
$$\longrightarrow\ \beta = 0,\ \pm\frac{2\pi}{N},\ \pm\frac{2\cdot 2\pi}{N},\ \pm\frac{3\cdot 2\pi}{N},\cdots \quad \text{(but see Eq. (10.39))} \qquad (10.38)$$

How far does this set of numbers go? It has to be finite because there are only a finite number of modes. Explore a special case first, and look at the shape that the function q_n takes. It is

$$q_n = e^{i\beta n} = \big\{1,\ e^{i\beta},\ e^{2i\beta},\ e^{3i\beta},\ \ldots\big\}$$

The $\beta = 0$ case is trivial — nothing moves, as Eq. (10.37) says that $\omega = 0$ in this case.* Now comes the first non-trivial case, where $\beta = 1\cdot 2\pi/N$, and here I take the real part to interpret it.

$$y_n(t) = q_n e^{-i\omega_1 t} = e^{i(2\pi n/N) - i\omega_1 t} \ \longrightarrow\ \cos\big((2\pi n/N) - \omega_1 t\big)$$

At $t = 0$, as n goes from zero to $N - 1$ and then to N (which is the same as zero), this cosine goes through one full cycle. At later times this cosine will move, and the peak will be defined by the point $\big((2\pi n/N) - \omega_1 t\big) = 0$. This is a wave, and it has the maximum wavelength possible along the loop. In this picture, $N = 10$ and $t = 0$.

$n = 0$ $n = 10 \longleftrightarrow n = 0$

Fig. 10.9

5    $\beta = 1\cdot 2\pi/10,\qquad t = 0$
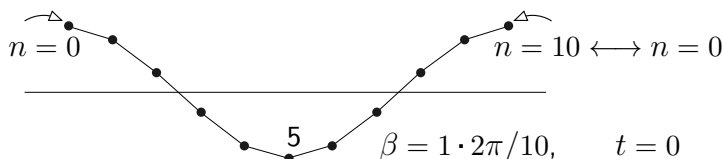
Draw a picture at a short time later, and the cosine will have moved to the right and around the loop. Take the time $\omega_1 t = 2\pi/5$ as an example

---

* Not really. Look back to Eq. (10.30) and see what happens if all the y_n have the same value.

$n = 0$ $n = 10 \longleftrightarrow n = 0$

Fig. 10.10

5   $\beta = 1 \cdot 2\pi/10, \qquad \omega_1 t = 2\pi/5$
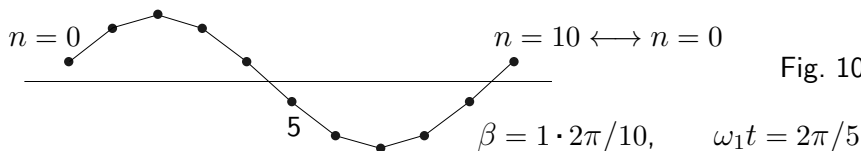
What if you choose the negative sign for β? Simple: The wave is moving in the opposite direction.

In Eq. (10.38) how large can β be? So far you've looked at only $\beta = 0$, $\pm 2\pi/N$. An easy case to test is $\beta = N \cdot 2\pi/N = 2\pi$. This is again trivial, just as $\beta = 0$ was, because $e^{i\beta n} = e^{i2\pi n} \equiv 1$. Eq. (10.37) tells you that $\omega = 0$ here too. Nothing moves.

What if β comes from an index one less than N? $\beta = (N-1)2\pi/N = 2\pi - 2\pi/N$. Then $e^{i\beta n} = e^{i(2\pi - 2\pi/N)n} = e^{-2\pi i n/N}$. This is identical to the case that $\beta = -2\pi/N$, and this wave is moving left, having the same wavelength as in these pictures, except for the direction of motion. It's not a new wave, only the same shape moving in the opposite direction. You can see from this reasoning that going down from zero or down from N gives the same set of solutions, and that means that the number of different solutions can be taken from the set $(0, 1, 2, \ldots, N-1)$ or from the set $(0, \pm 1, \pm 2, \ldots, \pm N/2)$ as you choose. *Not quite.* Does it matter if N is even or odd? You should try the cases $N = 3$ and $N = 4$, and see. You should in fact write out all the solutions for all the modes for $N = 2, 3$, and 4 in order to get a feeling for what they do.

Cutting through all the verbiage,

$$0 \le \beta < 2\pi \quad \text{in steps of} \quad 2\pi/N \qquad \text{(See the graph, Figure 10.12.)}$$
$$i.e. \quad \beta = 2\pi s/N, \quad s = 0, 1, 2 \ldots, (N-1) \tag{10.39}$$

How short can the wavelength be? That is when β is at the halfway point: $N/2$ for N even or $(N \pm 1)/2$ for N odd. For the pictures here, with $N = 10$, that will be $\beta = 5 \cdot 2\pi/10 = \pi$.

$$q_n = e^{i\pi n} \longrightarrow y_n(t) = e^{i\pi n - i\omega_5 t} \longrightarrow \cos(\pi n - \omega_5 t) \tag{10.40}$$

0 $10 \longleftrightarrow 0$

Fig. 10.11

5   $\beta = 5 \cdot 2\pi/10 = \pi, \quad t = 0$
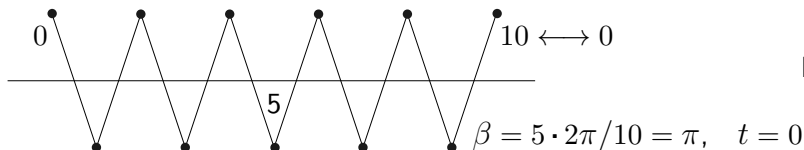
You can get an idea in this picture why this represents the shortest wavelength available in this system. I kept the vertical scale the same in all three drawings, even though it makes it appear very exaggerated in the third one. Which direction is it moving, left or right? Remember that small positive β gave motion to the right. If β is just less than N it moved left. This one is in the center, so what does *it* do? [What is the cosine of the difference of two angles in Eq. (10.40)?]

Look at the case in which there are *many* masses. This model can approximate a continuous string, but it will demonstrate some phenomena that don't appear in the simpler models of chapter seven. For example, in the equation (10.37), $\omega \propto \sin(\beta/2)$ implies that there is a maximum frequency that a wave can have. That doesn't happen with light or sound. Or does it? With sound, what if the frequency is so high that the corresponding wavelength is comparable to the distance between molecules? Sound doesn't make much sense any more, so there really *is* an upper limit for sound too.

• The letter "ℓ" is the distance between the masses, so $m/\ell = \mu$ is the average linear mass density.

• The wavelength of a wave is the distance between adjacent crests, so in the present context that is going from $n = 0$ to the value of n such that $\beta n = 2\pi$. The *distance* corresponding to this n is $n\ell$. Eliminate n and this says $\lambda = n\ell = 2\pi\ell/\beta$, which in turn means that the wave number is $k = 2\pi/\lambda = \beta/\ell$.

• In terms of the variables ω and k, the equation (10.37) is then

$$\omega^2 = \frac{4T}{m\ell}\sin^2(\beta/2) = \frac{4T}{\mu\ell^2}\sin^2\left(\tfrac{1}{2}k\ell\right) \tag{10.41}$$

The phase velocity of the wave satisfies

$$v^2 = \frac{\omega^2}{k^2} = \frac{4T}{\mu\ell^2}\frac{\sin^2\left(\tfrac{1}{2}k\ell\right)}{k^2}$$

For small values of $k\ell/2$, less than $1/2$ or so, the sine function is nearly linear, and this speed then agrees with $v^2 = T/\mu$ as in the case of a continuous string, Eq. (7.7). Put this in terms of the wavelength and it is $k\ell = 2\pi\ell/\lambda \lesssim 1$, or $\lambda \gtrsim 2\pi\ell$. One wavelength spans many masses so that the system starts to look like a continuous string.

Put the wave itself into the notation appropriate for waves.

$$y_n = e^{i(2\pi n/N)-i\omega_1 t} = e^{i(\beta n-\omega t)} = e^{i(k\ell n-\omega t)} = e^{i(kx-\omega t)}$$

This is exactly the standard form used in chapter seven. k is positive or negative depending on the direction of motion. This is also why I chose the oscillation to be proportional to $e^{-i\omega t}$ instead of $+i$.

Look back at the discussion about the several different definitions of wave velocity, section 7.10, Eq. (7.44). Recall the differences between two of the possible definitions of wave velocity: phase velocity and group velocity. Here they are

$$v_{\text{phase}} = \frac{\omega}{k} = \sqrt{\frac{T}{\mu}}\frac{\sin\tfrac{1}{2}k\ell}{\tfrac{1}{2}k\ell}, \qquad \text{and} \qquad v_{\text{group}} = \frac{d\omega}{dk} = \sqrt{\frac{T}{\mu}}\cos\tfrac{1}{2}k\ell \tag{10.42}$$

When the wavelength is long enough ($k\ell \ll 1$), both of these velocities become $\sqrt{T/\mu}$. At much higher frequencies, for which the wavelength gets down near the spacing ℓ, these become
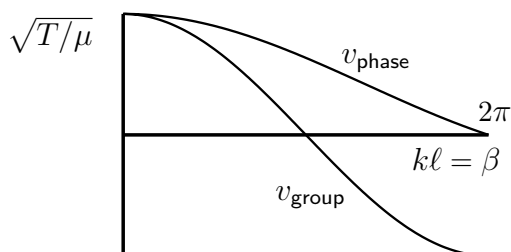
Fig. 10.12

What is β when $\lambda = \ell$? Notice that the group velocity can be in the opposite direction from the phase velocity. See www.falstad.com/dispersion/groupneg.html for animated images of this.

There are some curious puzzles about this system. Look again at the example used on the last few pages, $N = 10$, so that $\beta = 2\pi s/10$ for $0 \le s < N$ with s an integer. For $s = 5$ this is $\beta = \pi$, and $y_n = \cos(\pi n - \omega t) = \cos(\pi n)\cos(\omega t)$. The picture is Figure 10.11, and from Eq. (10.41) you can see that this is the highest frequency of all the modes of oscillation. But, this wave is not going anywhere; it is standing still. This is so despite the fact that the phase velocity is not zero. That's odd, but notice that the group velocity *is* zero at that point. This is another indication that the phase velocity is not as useful a concept as you may have been led to believe, even though it was probably the *only* wave velocity that in your first introduction to waves. The group velocity is negative for values of β above the middle, and this corresponds to the fact described a couple of paragraphs after Figure 10.10: When moving down from the maximum β, you get a left-moving wave.

10.7 Perturbation Theory˙

The special cases for which the matrices are simple and the problem is readily solvable are precisely that — special. What about the other cases for which a direct attack will lead to a morass of messy mathematics?

There is one set of circumstances for which there is hope. That occurs when the system is *almost* easy. That is, if the matrices are very close to ones for which there is an easy solution. (Or at least a known solution.)

The idea of perturbation theory is to write $\mathbf{M}$ and $\mathbf{K}$ each as the sum of two terms: one simple, one small. Then write $\mathbf{x}$ and ω^2 as a sum of terms, an infinite series typically, for which I can compute the first few terms. This can provide a good approximate solution to the full problem.

For an example of this, take the prototype problem, Eq. (10.1). If

$$m_1 = 1.01\,\text{kg}, \quad m_2 = 0.98\,\text{kg}, \quad k_1 = 10.0\,\text{N/m}, \quad k_3 = 10.2\,\text{N/m}, \quad \text{and} \quad k_2 = 5.0\,\text{N/m}$$

then the system is almost symmetric and its solution is close to that of the simple, symmetric system. But *how close?*

Use an expansion parameter ϵ as a convenient way to keep track of the powers in an expansion, and assume a series expansion for the unknowns $\mathbf{x}$ and ω^2.

$$\begin{aligned} \mathbf{M} &= \mathbf{M}_0 + \epsilon\mathbf{M}_1 & \mathbf{x} &= \mathbf{x}_0 + \epsilon\mathbf{x}_1 + \epsilon^2\mathbf{x}_2 + \cdots \\ \mathbf{K} &= \mathbf{K}_0 + \epsilon\mathbf{K}_1 & \omega^2 &= (\omega^2)_0 + \epsilon(\omega^2)_1 + \epsilon^2(\omega^2)_2 + \cdots \end{aligned} \tag{10.43}$$

Substitute these into the equations that I want to solve.

$$-\omega^2\mathbf{M}\mathbf{x} + \mathbf{K}\mathbf{x} = 0$$
$$-\big[(\omega^2)_0 + \epsilon(\omega^2)_1 + \epsilon^2(\omega^2)_2 + \cdots\big]\big[\mathbf{M}_0 + \epsilon\mathbf{M}_1\big]\big[\mathbf{x}_0 + \epsilon\mathbf{x}_1 + \epsilon^2\mathbf{x}_2 + \cdots\big]$$
$$+ \big[\mathbf{K}_0 + \epsilon\mathbf{K}_1\big]\big[\mathbf{x}_0 + \epsilon\mathbf{x}_1 + \epsilon^2\mathbf{x}_2 + \cdots\big] = 0$$

The only way that this can be valid for arbitrary ϵ is for the respective coefficients of the powers of ϵ to agree.

$$\begin{aligned} (\epsilon^0) \qquad & -(\omega^2)_0\mathbf{M}_0\mathbf{x}_0 + \mathbf{K}_0\mathbf{x}_0 = 0 \\ (\epsilon^1) \qquad & -(\omega^2)_0\mathbf{M}_0\mathbf{x}_1 - (\omega^2)_0\mathbf{M}_1\mathbf{x}_0 - (\omega^2)_1\mathbf{M}_0\mathbf{x}_0 + \mathbf{K}_0\mathbf{x}_1 + \mathbf{K}_1\mathbf{x}_0 = 0 \\ (\epsilon^2) \qquad & -(\omega^2)_0\mathbf{M}_0\mathbf{x}_2 - (\omega^2)_0\mathbf{M}_1\mathbf{x}_1 - (\omega^2)_1\mathbf{M}_0\mathbf{x}_1 - \\ & (\omega^2)_1\mathbf{M}_1\mathbf{x}_0 - (\omega^2)_2\mathbf{M}_0\mathbf{x}_0 + \mathbf{K}_0\mathbf{x}_2 + \mathbf{K}_1\mathbf{x}_1 = 0 \end{aligned} \tag{10.44}$$

It appears that I've made the problem worse, not better. I've replaced one difficult equation by an infinite number of difficult equations. But no. In what looks to be a hopeless tangle, there's some simplicity hidden. Besides, I'm never going to get around to using the third of these equations, only the first two. And why the frequency *squared* instead of the frequency? That's just the way it works, but see problem 7.16. Writing the frequency this way is unnecessarily cumbersome, so I will stop doing it. Instead of $(\omega^2)_1$ I will write ω_1^2.

Equation (ϵ^0) is the equation that I've already solved. It's the one that is supposedly simple. The first new equation that I have to examine is (ϵ^1).

$$-\omega_0^2\mathbf{M}_0\mathbf{x}_1 + \mathbf{K}_0\mathbf{x}_1 = \big(-\omega_0^2\mathbf{M}_0 + \mathbf{K}_0\big)\mathbf{x}_1 = +\omega_0^2\mathbf{M}_1\mathbf{x}_0 + \omega_1^2\mathbf{M}_0\mathbf{x}_0 - \mathbf{K}_1\mathbf{x}_0 \tag{10.45}$$

Look at the way that I rearranged this equation and notice that on the right-hand side the only thing that is unknown is ω_1^2, the term that gives the first order correction to the frequency. On the left side

there is something that looks like the simple, unperturbed equation. At this point I have to stop and present a theorem. It is such an easy theorem that you may wonder why I'd bother calling it such, but it is the key to solving all these equations. It breaks open the entire structure and makes it manageable.

Theorem: If (1) $\mathbf{H}$ is a matrix that satisfies $\mathbf{H}^\dagger = \tilde{\mathbf{H}}^* = \mathbf{H}$, (Hermitian) and (2) $\mathbf{H}$ has a null eigenvector — there is an $\mathbf{x}$ such that $\mathbf{Hx} = 0$ then for *any* column matrix $\mathbf{y}$,

$$\mathbf{x}^\dagger \mathbf{Hy} = 0 \tag{10.46}$$

That is, if $\mathbf{x}$ is an eigenvector with zero eigenvalue then $\mathbf{x}$ is orthogonal to the output of $\mathbf{H}$ for any input.

The proof involves writing it out.

$$\mathbf{x}^\dagger \mathbf{Hy} = \left(\mathbf{Hx}\right)^\dagger \mathbf{y} = 0 \tag{10.47}$$

It's as simple as that. To see it written in terms of components it is:

$$\mathbf{x}^\dagger \mathbf{Hy} = \sum_{i,j} x_i^* H_{ij} y_j = \sum_{i,j} \left(H_{ij}^* x_i\right)^* y_j = \sum_{i,j} \left(\tilde{H}_{ji}^* x_i\right)^* y_j = \sum_{i,j} \left(H_{ji} x_i\right)^* y_j = \left(\mathbf{Hx}\right)^\dagger \mathbf{y} \tag{10.48}$$

The matrix on the left side of Eq. (10.45) and that acts on $\mathbf{x}_1$ is precisely such a matrix. The (ϵ^0) equation says that $\mathbf{x}_0$ is a null eigenvector. That means that if I multiply the equation (ϵ^1) by $\mathbf{x}_0^\dagger$ then I must get zero.

$$\mathbf{x}_0^\dagger\left(-\omega_0^2 \mathbf{M}_0 \mathbf{x}_1 + \mathbf{K}_0 \mathbf{x}_1\right) = 0 = \mathbf{x}_0^\dagger\left[\omega_0^2 \mathbf{M}_1 \mathbf{x}_0 + \omega_1^2 \mathbf{M}_0 \mathbf{x}_0 - \mathbf{K}_1 \mathbf{x}_0\right]$$

There's no matrix inversion to do here, just a scalar product, and the result is that it determines the first order correction to the frequency, ω_1^2.

$$0 = \omega_1^2 \mathbf{x}_0^\dagger \mathbf{M}_0 \mathbf{x}_0 + \mathbf{x}_0^\dagger\left(\omega_0^2 \mathbf{M}_1 \mathbf{x}_0 - \mathbf{K}_1 \mathbf{x}_0\right) \qquad \text{so}$$
$$\omega_1^2 = \mathbf{x}_0^\dagger\left(-\omega_0^2 \mathbf{M}_1 \mathbf{x}_0 + \mathbf{K}_1 \mathbf{x}_0\right) / \mathbf{x}_0^\dagger \mathbf{M}_0 \mathbf{x}_0 \tag{10.49}$$

Computing the lowest order correction to the frequency (squared) is now just some matrix multiplication. Notice that what shows up in the denominator is the same scalar product as in Eq. (10.22).

Does all of this material seem familiar, something you've studied before? Return to section 7.13 and the development of perturbation theory for standing waves. There the mathematics were differential equations and here they are matrices, but the underlying structure of the calculations are the same. Here are some key equations from both calculations so that you can compare them.

Eq. (7.50)

$$\mu = \mu_0 + \epsilon\mu_1$$
$$F = F_0 + \epsilon F_1 + \epsilon^2 F_2 + \epsilon^3 F_3 + \cdots$$
$$\omega^2 = \omega_0^2 + \epsilon\omega_1^2 + \epsilon^2\omega_2^2 + \cdots$$

Eq. (10.43)

$$\mathbf{M} = \mathbf{M}_0 + \epsilon\mathbf{M}_1, \qquad \mathbf{K} = \mathbf{K}_0 + \epsilon\mathbf{K}_1$$
$$\mathbf{x} = \mathbf{x}_0 + \epsilon\mathbf{x}_1 + \epsilon^2\mathbf{x}_2 + \cdots$$
$$\omega^2 = \omega_0^2 + \epsilon\omega_1^2 + \epsilon^2\omega_2^2 + \cdots$$

If you've studied perturbation theory for the inertia tensor in section 8.7 then again, you are repeating the same basic ideas. The equations (8.49) look very much like (10.43). This material is sufficiently important that it *is* worth doing three times. Maybe more.

Eq. (7.52)(b)
$$TF_1'' + \mu_0\omega_0^2 F_1 = -\mu_1\omega_0^2 F_0 - \mu_0\omega_1^2 F_0$$

Eq. (10.45)
$$(-\omega_0^2\mathbf{M}_0 + \mathbf{K}_0)\mathbf{x}_1 = +\omega_0^2\mathbf{M}_1\mathbf{x}_0 + \omega_1^2\mathbf{M}_0\mathbf{x}_0 - \mathbf{K}_1\mathbf{x}_0$$

Eq. (7.54)
$$\int_0^L dx\, F_0(x)\big[TF_1'' + \mu_0\omega_0^2 F_1\big] = 0$$

Eq. (10.46)
$$\mathbf{x}^\dagger\mathbf{Hy} = 0$$

Eq. (7.57)
$$\omega_1^2 = -\omega_0^2 \int_0^L dx\, \mu_1(x) F_0(x)^2 \Big/ \mu_0 \int_0^L dx\, F_0(x)^2$$

Eq. (10.49)
$$\omega_1^2 = \mathbf{x}_0^\dagger\big(-\omega_0^2\mathbf{M}_1\mathbf{x}_0 + \mathbf{K}_1\mathbf{x}_0\big)\big/\mathbf{x}_0^\dagger\mathbf{M}_0\mathbf{x}_0$$

Example

▶ Take the case of two masses and three springs, Eq. (10.1). If this is the special symmetric case for which $k_1 = k_3$ and $m_1 = m_2$ then the modes and the frequencies are simple.

$$\omega_0^2 = \frac{k_1}{m},\ \mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \omega_0^2 = \frac{k_1 + 2k_2}{m},\ \mathbf{x}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad (10.50)$$

What happens if I stick a small wad of chewing gum on one of the masses? Now it's now not such a symmetric system and the result is more complicated. In this case you can actually solved for the frequencies, so that you can test this perturbation method on something that you can do explicitly. Let $m_2 = m_1 = m$ be changed to $m_2 = m + m_g$. What does the perturbation method say about the modified frequency?

$$\mathbf{M}_0 = \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix}, \qquad \mathbf{M}_1 = \begin{pmatrix} 0 & 0 \\ 0 & m_g \end{pmatrix}, \qquad \mathbf{K}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad (10.51)$$

The equation (10.49) now gives the correction to the frequency for each mode of oscillation.

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \longrightarrow \omega_1^2 =$$
$$(1\ \ 1)\left[-\frac{k_1}{m}\begin{pmatrix} 0 & 0 \\ 0 & m_g \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]\Big/ (1\ \ 1)\begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = -\frac{k_1}{m}\frac{m_g}{2m} \qquad (10.52)$$

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \longrightarrow \omega_1^2 =$$
$$(1\ \ -1)\left[-\frac{k_1 + 2k_2}{m}\begin{pmatrix} 0 & 0 \\ 0 & m_g \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right]\Big/ 2m = -\frac{k_1 + 2k_2}{m}\frac{m_g}{2m} \qquad (10.53)$$

Both frequencies are lowered, which is what you expect in adding some mass to the system.
      The exact solution is stated in problem 10.26, allowing a check on these results.

Exercises

1 In the example at Eq. (10.1), take the case in which $k_1 = k_3 = 0$ and add the two equations. What does this result mean?

2 What happens if you add the three equations at Figure 10.2?

3 Just before Eq. (10.5) there is a reference to Eq. (3.1). How does this new material apply to that equation? Where's the matrix?

4 Verify the result of Eq. (10.10) for the 2×2 case, with $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and N having components e, f, g, h.

5 Paralleling the calculation in Eq. (10.15), find the energy for the circuit of Eq. (10.2)

6 Answer the question in the paragraph immediately after Eq. (10.40).

7 What happens to the result in Eq. (10.24) if all the normal mode solutions (the $\mathbf{x_a}$) are multiplied by two?

8 Explicitly verify Eq. (10.47) for the 2×2 case. That is, take the matrix $\mathbf{H} = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix}$ with null eigenvector $\mathbf{x} = \begin{pmatrix} e \\ f \end{pmatrix}$ and show that $\mathbf{x}^\dagger \mathbf{H} \mathbf{y}$ is zero no matter what $\mathbf{y}$ is. Note: you do not have to *solve* the eigenvalue equation; just use it.

9 Translate the equations (10.24) through (10.26) into the 1×1 case. That is, what are these equations for the simple harmonic oscillator of chapter three?

10 In the equation (10.28), look at the whole expression in front of the second $\begin{pmatrix} e \\ f \end{pmatrix}$. What is its determinant?

11 There are two vectors with components $(1, \ 1, \ 1)$, $(0, \ 1, \ -1)$, and a third is orthogonal to the first two. Find it.

Problems

10.1 Explicitly write out the steps of the derivation presented in Eq. (10.10) for the 2×2 case, with **M** and **N** being arbitrary. *etc.*

10.2 Derive the equations (10.11) and (10.12). Also, explicitly write out each of them term-by-term for the 2×2 matrix case.

10.3 For the special case Eq. (10.4), explicitly write out the equations (10.16), (10.17), (10.18), and (10.19)

10.4 If the mass matrix is proportional to $\left(\begin{smallmatrix}1 & 2\\3 & 4\end{smallmatrix}\right)$, and the scalar product of two column matrices is defined as in Eq. (10.22), which of the four requirements for a scalar product are violated? Which of the properties of Eq. (10.8) are violated?

10.5 Do the manipulations starting at Eq. (10.14) explicitly for the 2×2 case, verifying the results by longhand.

10.6 Set up the equations for the frequencies of oscillation of the coupled circuits, Eq. (10.2), and
(**a**) find the frequencies of oscillation.
(**b**) In the special case that the mutual inductance is zero, do your results reduce to the correct values?
(**c**) If the two capacitances have the equal values and the self-inductances do too, find the frequencies and modes of oscillation.

10.7 Write out the manipulations of Eq. (10.48) explicitly for the two by two case, with $\mathbf{H} = \left(\begin{smallmatrix}a & b\\c & d\end{smallmatrix}\right)$.

10.8 If you didn't do problem 7.16, now is a good time.

10.9 Verify that Eq. (10.22) satisfies the requirements for a scalar product.

10.10 Two pendulums have the same length and the same mass at their ends. They are attached to each other by a very light spring as in the drawing, and the spring is relaxed when both masses are hanging vertically. Set up the equations of motion and solve for the normal modes of oscillation and their frequencies.
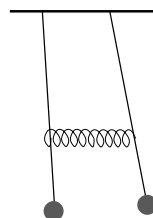
10.11 Use the general expression, Eq. (10.14), for energy in a coupled harmonic system, compute the energy for the coupled circuit in Eq. (10.2). This is *not* the energy that normally associated with capacitors and inductors as described in any book on electromagnetism. There you will see only q's and I's, with no dI/dt. Is something wrong? Go back and do problem 3.23 to answer this.

10.12 Solve for the vibrational modes of oscillation of a carbon dioxide molecule. CO_2 is linear, making the math a bit easier than for a water molecule. For this problem, assume that motion is confined to one dimension. Ans: One mode has frequency $\omega^2 = k(2m_o + m_c)/(m_o m_c)$

10.13 Repeat the preceding problem for the CO_2 molecule, but this time the molecule is rotating about an axis through the C-atom and perpendicular to the axis of the molecule. Assume that the angular velocity is fixed in space with magnitude Ω.

10.14 Repeat the preceding problem for the CO_2 molecule, with the molecule rotating about an axis through the C-atom and perpendicular to the axis of the molecule. Now add a touch of realism by noting that angular momentum is conserved: $L_z = 2mr^2\omega_z$, so that the angular velocity is not constant.

10.15 Use the methods of section 10.4 to solve the problem in section 3.9 leading to Eq. (3.68).

10.16 You can do the example in equations (10.50) and (10.51) another way. Rewrite the matrices in a more symmetric form before doing the perturbation calculation.

$$\mathbf{M}_0 = \begin{pmatrix} m + \frac{1}{2}m_\mathrm{g} & 0 \\ 0 & m + \frac{1}{2}m_\mathrm{g} \end{pmatrix}, \qquad \mathbf{M}_1 = \frac{1}{2}\begin{pmatrix} -m_\mathrm{g} & 0 \\ 0 & m_\mathrm{g} \end{pmatrix}, \qquad \mathbf{K}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The says that the perturbation involves subtracting some mass from one and adding it to the other. Now repeat the perturbation analysis. Show that to first order in m_g the two results agree. This way is easier.

10.17 In the example shown in Figure 10.2, assume that the masses are all the same, as are the springs. Without solving any equations, draw pictures that will describe the modes of oscillation. This is in the plane, two dimensions, so there are six coordinates and so six modes. In three of the cases you should be able to write the frequency corresponding to the mode exactly. In the other three cases you would have to do some algebra to find the result, but you need not. If you find seven possible solutions, you should be able to show that one is a linear combination of two others.

10.18 Two pendulums are coupled with a light spring. The two lengths of the pendulums are the same, but the masses are different. The spring is attached at a common distance ℓ from the support and the lengths of the pendulums are L. The masses are m_1 and m_2, and when the masses are hanging vertically the spring is unstretched. The solutions for the two oscillation frequencies are $\omega^2 = g/L$ and $\omega^2 = (g/L) + (k\ell^2/L^2\mu)$ where μ is the reduced mass. The respective amplitudes of motion for the two modes are $\begin{pmatrix} 1 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & -m_1/m_2 \end{pmatrix}$. Analyze these proposed solutions to see if they are plausible. Are the modes orthogonal? [It is easy to set up a demonstration of this apparatus and to see that the mass matrix is important in making these modes orthogonal.]
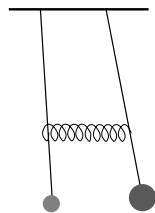
10.19 ˙Derive the solution for the preceding problem. Set up the equations of motion and find the frequencies and modes of oscillation. If you think that you're getting involved in solving a complicated quartic equation, look more closely.

10.20 Three masses, all equal to each other, are connected in a straight line between fixed walls as in the drawing Figure 10.1. All four springs are identical too. Without solving the equations, draw sketches showing what the normal modes will be. As a guide, remember the orthogonality relation, Eq. (10.22). It will help you in getting the shapes of the modes. One of the modes should be very easy to figure out exactly. Another one should be easy to do at least approximately. For the third, remember the orthogonality.
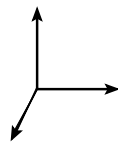
10.21 A string of N equal masses, m, are connected along a straight line by equal springs, k. The last springs on each end are attached to walls. (**a**) Write down the **M** and **K** matrices for this system. (**b**) If instead of being attached to walls, the system is pulled around into a large circle so that the masses are free to move only along the fixed circumference of the circle. The last spring is then attached to the first mass, making a system with circular symmetry. Write **M** and **K** now. (**c**) In both cases, do these matrices satisfy the real-symmetric requirements of Eq. (10.8)?

10.22 The expression for work in vector calculus notation is $W = \int d\vec{r}\cdot\vec{F}$, and the condition that this is independent of path and that $\oint d\vec{r}\cdot\vec{F} = 0$ is a condition on the partial derivatives of the components of $\vec{F}$ with respect to various coordinates. In the present context, $\vec{F} = -\mathbf{K}\mathbf{x}$ and

$$\oint d\vec{r}\cdot\vec{F} = -\oint d\tilde{\mathbf{x}}\mathbf{K}\mathbf{x}$$

If this is zero, what does this say about **K**?

10.23 In section 3.11 you find the Green's function for a simple harmonic oscillator. There is a similar result for coupled harmonic oscillators, only now the function will be a matrix of functions. Take the simplest special case of two equal masses and three springs (symmetric) and kick the first one of the masses, as in Eq. (3.72). Find the response of the two masses to this kick and then follow the pattern of the calculation leading to Eq. (3.73) to get the result for an arbitrary force applied to the first mass. You will get integrals giving you $x_1(t)$ and $x_2(t)$. Repeat this for a force on the second mass and finally write the whole result in matrices

$$\mathbf{x}(t) = \int_{-\infty}^{t} dt' \, \mathbf{G}(t - t')\mathbf{f}(t')$$

where $\mathbf{G}$ is a 2×2 matrix and $\mathbf{f}$ is a two element column matrix representing the forces. Does your answer make sense if the middle spring becomes either very weak or very strong?
Ans: $G_{11} = \frac{1}{2m}\left[\frac{1}{\omega}\sin\omega t + \frac{1}{\omega'}\sin\omega't\right]$

10.24 In the preceding problem you used the basis $\binom{1}{0}$ and $\binom{0}{1}$. What if you switch to the basis $\binom{1}{1}$ and $\binom{1}{-1}$? I.e., make the first kick equal on both masses $(\rightarrow, \rightarrow)$. Make the second kick opposite on the two masses $(\rightarrow, \leftarrow)$. Now what is the matrix $\mathbf{G}$? Again, does your answer make sense if the middle spring becomes either very weak or very strong?

10.25 (a) Use the result of problem 10.23 to to describe what happens if you apply the force $F_0 \sin\omega t$ to the first mass. (b) What is the result if the force is $F_0 \sin\omega't$? Here ω and ω' are the two natural frequencies for the system.

10.26 The exact solution for the frequencies found in the equations (10.52) and (10.53) are

$$\omega^2 = \left[(2m + m_g)(k_1 + k_2) \pm \sqrt{m_g^2(k_1 + k_2)^2 + 4m(m + m_g)k_2^2}\right]/2m(m + m_g)$$

Verify that a series expansion of the exact solution reproduces the two approximate solutions.

10.27 What will replace Eq. (10.24) if you use $e^{\pm i\omega t}$ instead of $\cos\omega t$ and $\sin\omega t$?

10.28 Like problem 10.10, but with three pendulums in a row. Set up the general equations and the algebraic equation for the frequencies. It is a cubic equation for ω^2, but you should be able to see one solution instantly. Factor that root out of the equation. Now you can see the shape of another mode easily and then compute its frequency. Factor that root out of the equation and you have a linear equation for the final ω^2. What is the shape of that mode, and are the three modes orthogonal?

10.29 In problem 10.19 you may have solved the problem exactly, but even if you didn't, and if the masses are not very different, use perturbation theory to find the oscillation frequencies approximately.

10.30 Use the methods of section 10.4 to solve problem 10.23.

10.31 Draw pictures like those in the figures 10.9–10.11 for the cases of two and of three masses: $N = 2, 3$.

10.32 What is the time dependence in Figure 10.11? I.e. draw the picture at later times.

10.33 Solve problem 3.59. Are the modes orthogonal? Is the scalar product positive definite? Don't just assume that it is; check it. If you don't know any fancy theorems to do this, then perhaps you can compute the eigenvectors and eigenvalues of $\mathbf{M}$, then deduce the result from there.

10.34 To differentiate a function numerically, using only the data at discrete, equally spaced points, the simplest method estimates the result by the "midpoint rule": $f'(x + \frac{1}{2}h) \approx [f(x+h) - f(x)]/h$. You can estimate the second derivative by first using this method to get $f'(x - h/2)$ and $f'(x+h/2)$. Then get the second derivative at x by applying the same midpoint rule and using these two results. Now go back to Eq. (10.30), divide it by ℓ, and see what it (approximately) is. You can even take the limit as $\ell \to 0$, and compare this result to the wave equation of chapter seven.

10.35 Generalize problem 10.23 to solve an arbitrary coupled oscillating system with an applied force: $\mathbf{M\ddot{x}} + \mathbf{Kx} = \mathbf{f}(t)$. Use the notation of section 10.4, and the answer is claimed to be

$$\mathbf{x}(t) = \int_{-\infty}^{t} dt'\, \mathbf{G}(t - t')\mathbf{f}(t'), \qquad \text{with} \qquad \mathbf{G}(t) = \sum_b \frac{\sin \omega_b t}{\omega_b \mathbf{x}_b^\dagger \mathbf{M x}_b}\mathbf{x}_b \mathbf{x}_b^\dagger$$

Check whether this is plausible. If you're not sure what $\mathbf{x}_b \mathbf{x}_b^\dagger$ means (and why should you be?), just put it in front of the $\mathbf{f}(t')$ and see what it looks like then.

10.36 Derive the result for the preceding problem. Sorting out the notation is probably the hardest part. Note: Do not assume that the matrix $\mathbf{M}$ is diagonal. This implies that if you start from rest and give an impulse to the first mass, then perhaps *all* the masses will immediately start moving. *E.g.* Eqs. (10.2). Play with the 2×2 case first to see what is going on, writing out all the M_{11}, M_{12}, ... explicitly. Go back to the derivation starting with Eq. (3.71), only now m is the matrix $\mathbf{M}$ and $1/m$ is the matrix $\mathbf{M}^{-1}$.

10.37 Use the result of problems 10.35 and 10.36 to derive the result for problem 10.23.

10.38 (**a**) If you haven't solved problem 3.80, do it now. Even if you have, look at it again and see how much easier it seems. I hope. (**b**) Compare your approximate answer to the exact answer stated in problem 10.26.

10.39 Examine the result stated in Eq. (10.29) and determine if it is plausible.

10.40 How do the phases of the motions of the two masses in Eq. (10.29) compare? The quotient $x_1(t)/x_2(t)$ cancels both the time dependence and the $\det$'s. Plot this relative phase versus Ω.

10.41 Three masses, all $= m$, are spaced equally around a wire and are able to slide along the wire. Springs connect the masses, all the same, and the masses can oscillate. Find the normal modes of this system. Did you do all of the preliminary exercises at the end of this chapter?
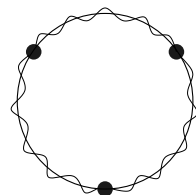
10.42 Return to problem 3.60 and analyze it using the tools of this chapter. What is the orthogonality relationship between the normal modes, and what do they look like?

10.43 In section 10.6 the boundary conditions on the masses required that there are a finite number of them, and the endpoints were connected in some way. Now assume that the chain is like the infinitely long string used at the beginning of chapter 7. It never stops. This means that you can more easily represent the solutions as running waves such as $e^{ikx - \omega t}$. (**a**) What is the relation between ω and k for this case? (**b**) Graph the phase and group velocities as in Figure 10.12. (**c**) What constraints are there now on the possible values of ω and k?

Index

.

2001, 287

•

aberration, 327
Abramowitz and Stegun, 146
absolute future, past, 310
abuse of notation, 299
accelerated coordinate system, 156, 172, 335
accelerated reference frame, 307
acceleration, **16**, 32, 34, 316
    polar coordinates 17
    proper 2, 316
    transformation 316
action, 34
Adelbereger, Eric, 37
adiabatic, 246
adiabatic invariant, 140
adjoint, 339
air resistance, 52, 64, 71, 74, 123, 142, 165, 181, 230
amplitude, 88, 93, 152, 241, 247
analyzing results, **59**, 66, 114, 149, 180, 314
angular momentum, 41, 149, 178, 193, 201, 222, 226, 267–287
angular velocity, 158, 159, 177, 182, 267, 275, 284
anharmonic oscillator, 108, 120
Antarctica, 296
anti-cyclonic, 166
anti-proton, 320, 325
aphelion, 191, 194, 213, 221, 222, 228
apparent speed, 332
apparent weight, 168
approximate orbits, 195
arbitrary function, 233, 259
Archimedes, 98
arithmetic-geometric mean, 220
artillery, 166
associative, 12, 19, 24
asteroid, 180, 224, 227, 318, 321, 322
astronomical unit, 75, 190, 222
atmosphere, 52, 138, 166, 226, 303, 331
Atwood machine, 45, 68, 121, 152
AU, 74, 190, 222

•

ballast, 89
baseball, 143, 321
basis, 12, 13, 16, 23, 43, 157, 180, 271, 278, 281
Bay of Fundy, 173
beats, 103
Beckham, David, 145
beer, 302
bend, 231
β, 307
binary star, 205, 333
binomial series, 2, 27, 29, 147, 235
birefringence, 245
black hole, 179, 332
blackboard bold, 26
boomerang, 162
boundary condition, 52, 64, 232, **239**–245, 253, 255, 260
    periodic 347, 348
brake, 58, 71, 175
breaking ball, 145
Brillouin, Leon, 246
bucket, 168, 183
Bullock, Sandra, 287
bullwhip, 230, 251, 258
bungee jump, 72

•

c is one, 327
Callaway, 52
capacitance, 88, 119, 337, 356
carbon monoxide, 120, 227
caress, 56
caricature, 298
center of mass, 203, 264, 270, 329
centrifugal:
    force, 159, 167–170, 181, 184
    potential 169, 193
chain, 346
chain rule, 14, 37, 46, 88, 233, 259
Chandler wobble, 285
characteristic value, 342
charge, 107, 132, 134, 138, 180, 292, 308–310
charge density, 108
circuit, 84, 88, 119, 337, 356