# Multivariable Calculus

The world is not one-dimensional, and calculus doesn't stop with a single independent variable. The ideas of partial derivatives and multiple integrals are not too different from their single-variable counterparts, but some of the details about manipulating them are not so obvious. Some are downright tricky.

## 8.1 Partial Derivatives
The basic idea of derivatives and of integrals in two, three, or more dimensions follows the same pattern as for one dimension. They're just more complicated.

The derivative of a function of one variable is defined as

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{8.1}$$

You would think that the definition of a derivative of a function of $x$ and $y$ would then be defined as

$$\frac{\partial f(x,y)}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x,y)}{\Delta x} \tag{8.2}$$

and more-or-less it is. The $\partial$ notation instead of $d$ is a reminder that there are other coordinates floating around that are temporarily being treated as constants.

In order to see why I used the phrase "more-or-less," take a very simple example: $f(x,y) = y$. Use the preceding definition, and because $y$ is being held constant, the derivative $\partial f/\partial x = 0$. What could be easier?

I don't like these variables so I'll switch to a different set of coordinates, $x'$ and $y'$:

$$y' = x + y \qquad \text{and} \qquad x' = x$$

What is $\partial f/\partial x'$ now?

$$f(x,y) = y = y' - x = y' - x'$$

Now the derivative of $f$ with respect to $x'$ is $-1$, because I'm keeping the other coordinate fixed. Or is the derivative still zero because $x' = x$ and I'm taking $\partial f/\partial x$ and why should that change just because I'm using a different coordinate system?

The problem is that the *notation* is ambiguous. When you see $\partial f/\partial x$ it doesn't tell you what to hold constant. Is it to be $y$ or $y'$ or yet something else? In some contexts the answer is clear and you won't have any difficulty deciding, but you've already encountered cases for which the distinction is crucial. In thermodynamics, when you add heat to a gas to raise its temperature does this happen at constant pressure or at constant volume or with some other constraint? The specific heat at constant pressure is not the same as the specific heat at constant volume; it is necessarily bigger because during an expansion some of the energy has to go into the work of changing the volume. This sort of derivative depends on type of process that you're using, and for a classical ideal gas the difference between the two molar specific heats obeys the equation

$$c_p - c_v = R$$

If the gas isn't ideal, this equation is replaced by a more complicated and general one, but the same observation applies, that the two derivatives $dQ/dT$ aren't the same.

In thermodynamics there are so many variables in use that there is a standard notation for a partial derivative, indicating exactly which other variables are to be held constant.

$$\left(\frac{\partial U}{\partial V}\right)_T \qquad \text{and} \qquad \left(\frac{\partial U}{\partial V}\right)_P$$

represent the change in the internal energy of an object per change in volume during processes in which respectively the temperature and the pressure are held constant. In the previous example with the function $f = y$, this says

$$\left(\frac{\partial f}{\partial x}\right)_y = 0 \qquad \text{and} \qquad \left(\frac{\partial f}{\partial x}\right)_{y'} = -1$$

This notation is a way to specify the *direction* in the $x$-$y$ plane along which you're taking the derivative.

## 8.2 Chain Rule

For functions of one variable, the chain rule allows you to differentiate with respect to still another variable: $y$ a function of $x$ and $x$ a function of $t$ allows

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} \qquad\qquad (8.3)$$

You can derive this simply from the definition of a derivative.

$$\frac{\Delta y}{\Delta t} = \frac{y\big(x(t+\Delta t)\big) - y\big(x(t)\big)}{\Delta t}$$
$$= \frac{y\big(x(t+\Delta t)\big) - y\big(x(t)\big)}{x(t+\Delta t) - x(t)} \cdot \frac{x(t+\Delta t) - x(t)}{\Delta t} = \frac{\Delta y}{\Delta x} \cdot \frac{\Delta x}{\Delta t}$$

Take the limit of this product as $\Delta t \to 0$. Necessarily then you have that $\Delta x \to 0$ too (unless the derivative doesn't exist anyway). The second factor is then the definition of the derivative $dx/dt$, and the first factor is the definition of $dy/dx$. The Leibnitz *notation* as written in Eq. (8.3) leads you to the required proof.

What happens with more variables? Roughly the same thing but with more manipulation, the same sort of manipulation that you use to derive the rule for differentiating more complicated functions of one variable (as in section 1.5).

$$\text{Compute} \quad \frac{d}{dt}f\big(x(t), y(t)\big)$$

Back to the $\Delta$'s. The manipulation is much like the preceding except that you have to add and subtract a term in the second line.

$$\frac{\Delta f}{\Delta t} = \frac{f\big(x(t+\Delta t), y(t+\Delta t)\big) - f\big(x(t), y(t)\big)}{\Delta t}$$
$$= \frac{f\big(x(t+\Delta t), y(t+\Delta t)\big) - f\big(x(t), y(t+\Delta t)\big) + f\big(x(t), y(t+\Delta t)\big) - f\big(x(t), y(t)\big)}{\Delta t}$$
$$= \frac{f\big(x(t+\Delta t), y(t+\Delta t)\big) - f\big(x(t), y(t+\Delta t)\big)}{x(t+\Delta t) - x(t)} \cdot \frac{x(t+\Delta t) - x(t)}{\Delta t}$$
$$+ \frac{f\big(x(t), y(t+\Delta t)\big) - f\big(x(t), y(t)\big)}{y(t+\Delta t) - y(t)} \cdot \frac{y(t+\Delta t) - y(t)}{\Delta t}$$
$$= \frac{\Delta f}{\Delta x} \cdot \frac{\Delta x}{\Delta t} + \frac{\Delta f}{\Delta y} \cdot \frac{\Delta y}{\Delta t}$$

In the first factor of the first term, $\Delta f / \Delta x$, the variable $x$ is changed but $y$ is not. In the first factor of the second term, the reverse holds true. The limit of this expression is then

$$\lim_{\Delta t \to 0} \frac{\Delta f}{\Delta t} = \frac{df}{dt} = \left(\frac{\partial f}{\partial x}\right)_y \frac{dx}{dt} + \left(\frac{\partial f}{\partial y}\right)_x \frac{dy}{dt} \tag{8.4}$$

If these manipulations look familiar, it's probably because they mimic the procedures of section 1.5. That case is like this one, with the special values $x \equiv y \equiv t$.

Example: (When you want to check out an equation, you should construct an example so that it reveals a lot of structure without requiring a lot of calculation.)

$$f(x, y) = Axy^2, \quad \text{and} \quad x(t) = Ct^3, \quad y(t) = Dt^2$$

First do it using the chain rule.

$$\begin{aligned}
\frac{df}{dt} &= \left(\frac{\partial f}{\partial x}\right)_y \frac{dx}{dt} + \left(\frac{\partial f}{\partial y}\right)_x \frac{dy}{dt} \\
&= (Ay^2)(3Ct^2) + (2Axy)(2Dt) \\
&= (A(Dt^2)^2)(3Ct^2) + (2A(Ct^3)(Dt^2))(2Dt) \\
&= 7ACD^2t^6
\end{aligned}$$

Now repeat the calculation by first substituting the values of $x$ and $y$ and then differentiating.

$$\begin{aligned}
\frac{df}{dt} &= \frac{d}{dt}\left[A(Ct^3)(Dt^2)^2\right] \\
&= \frac{d}{dt}\left[ACD^2t^7\right] \\
&= 7ACD^2t^6
\end{aligned}$$

What if $f$ also has an explicit $t$ in it: $f(t, x(t), y(t))$? That simply adds another term. Remember, $dt/dt = 1$.

$$\frac{df}{dt} = \left(\frac{\partial f}{\partial t}\right)_{x,y} + \left(\frac{\partial f}{\partial x}\right)_{y,t} \frac{dx}{dt} + \left(\frac{\partial f}{\partial y}\right)_{x,t} \frac{dy}{dt} \tag{8.5}$$

Sometimes you see the chain rule written in a slightly different form. You can change coordinates from $(x, y)$ to $(r, \phi)$, switching from rectangular to polar. You can switch from $(x, y)$ to a system such as $(x', y') = (x + y, x - y)$. The function can be expressed in the new coordinates explicitly. Solve for $x, y$ in terms of $r, \phi$ or $x', y'$ and then differentiate with respect to the new coordinate. OR you can use the chain rule to differentiate with respect to the new variable.

$$\boxed{\left(\frac{\partial f}{\partial x'}\right)_{y'} = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial x'}\right)_{y'} + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial x'}\right)_{y'}} \tag{8.6}$$

This is actually not a different equation from Eq. (8.4). It only looks different because in addition to $t$ there's another variable that you have to keep constant: $t \to x'$, and $y'$ is constant.
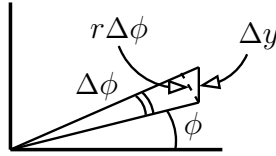
Example: When you switch from rectangular to plane polar coordinates what is $\partial f/\partial \phi$ in terms of the $x$ and $y$ derivatives?

$$x = r\cos\phi, \quad y = r\sin\phi, \quad \text{so}$$

$$\begin{aligned}
\left(\frac{\partial f}{\partial \phi}\right)_r &= \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial \phi}\right)_r + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial \phi}\right)_r \\
&= \left(\frac{\partial f}{\partial x}\right)_y (-r\sin\phi) + \left(\frac{\partial f}{\partial y}\right)_x (r\cos\phi)
\end{aligned}$$

If $f(x, y) = x^2 + y^2$ this better be zero, because I'm finding how $f$ changes when $r$ is held fixed. Check it out; it is. The equation (8.6) presents the form that is most important in many applications.

Example: What is the derivative of $y$ with respect to $\phi$ at constant $x$?

$$\left(\frac{\partial y}{\partial \phi}\right)_x = \left(\frac{\partial y}{\partial r}\right)_\phi \left(\frac{\partial r}{\partial \phi}\right)_x + \left(\frac{\partial y}{\partial \phi}\right)_r \left(\frac{\partial \phi}{\partial \phi}\right)_x$$

$$= [\sin \phi] \cdot \left[r \frac{\sin \phi}{\cos \phi}\right] + [r \cos \phi] \cdot 1 = r \frac{1}{\cos \phi} \tag{8.7}$$



You see a graphical interpretation of the calculation in this diagram: $\phi$ changes by $\Delta\phi$, so the coordinate moves up by $\Delta y$ ($x$ is constant). The angle between the lines $\Delta y$ and $r\Delta\phi$ is $\phi$ itself. This means that $\Delta y \div r\Delta\phi = 1/\cos \phi$, and that is precisely the preceding equation for $(\partial y/\partial \phi)_x$.

In doing the calculation leading to Eq. (8.7), do you see how to do the calculation for $(\partial r/\partial \phi)_x$? Differentiate the equation $x = r \cos \phi$ with respect to $\phi$.

$$x = r \cos \phi \;\rightarrow\; \left(\frac{\partial x}{\partial \phi}\right)_x = 0 = \left(\frac{\partial r}{\partial \phi}\right)_x \cos \phi + r \left(\frac{\partial \cos \phi}{\partial \phi}\right)_x = \left(\frac{\partial r}{\partial \phi}\right)_x \cos \phi - r \sin \phi$$

Solve for the unknown derivative and you have the result.

Another example: $f(x, y) = x^2 - 2xy$. The transformation between rectangular and polar coordinates is $x = r \cos \phi,\; y = r \sin \phi$. What is $(\partial f/\partial x)_r$?

$$\left(\frac{\partial f}{\partial x}\right)_r = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial x}\right)_r + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_r = (2x - 2y) + (-2x)\left(\frac{\partial y}{\partial x}\right)_r$$

$$\left(\frac{\partial y}{\partial x}\right)_r = \frac{(\partial y/\partial \phi)_r}{(\partial x/\partial \phi)_r} = \frac{r \cos \phi}{-r \sin \phi} = -\cot \phi \tag{8.8}$$

(Remember problem 1.49?) Put these together and

$$\left(\frac{\partial f}{\partial x}\right)_r = (2x - 2y) + (-2x)(-\cot \phi) = 2x - 2y + 2x \cot \phi \tag{8.9}$$

The brute-force way to do this is to express the function $f$ explicitly in terms of the variables $x$ and $r$, eliminating $y$ and $\phi$.

$$y = r \sin \phi = \sqrt{r^2 - x^2}, \qquad \text{then}$$
$$\left(\frac{\partial f}{\partial x}\right)_r = \frac{\partial}{\partial x}\left[x^2 - 2x\sqrt{r^2 - x^2}\right]_r$$

$$= 2x - 2\sqrt{r^2 - x^2} - 2x \frac{1}{\sqrt{r^2 - x^2}}(-x) = 2x + \frac{-2(r^2 - x^2) + 2x^2}{\sqrt{r^2 - x^2}} \tag{8.10}$$

You can see that this is the same as the equation (8.9) if you look at the next-to-last form of equation (8.10).

$$\frac{x}{\sqrt{r^2 - x^2}} = \frac{r \cos \phi}{\sqrt{r^2 - r^2 \cos^2 \phi}} = \cot \phi$$

Is this result reasonable? Look at what happens to $y$ when you change $x$ by a little bit. Constant $r$ is a circle, and if $\phi$ puts the position over near the right side (ten or twenty degrees), a little change in $x$ causes a big change in $y$ as shown by the rectangle. As drawn, $\Delta y / \Delta x$ is big and negative, sort of like the (negative) cotangent of $\phi$ as in Eq. (8.8).
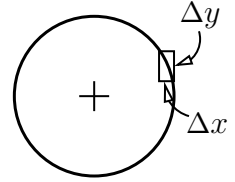
## 8.3 Differentials
For a function of a single variable you can write

$$df = \frac{df}{dx}dx \tag{8.11}$$

and read (sort of) that the infinitesimal change in the function $f$ is the slope times the infinitesimal change in $x$. Does this really make any sense? What is an infinitesimal change? Is it zero? Is $dx$ a number or isn't it? What's going on?

It *is* possible to translate this intuitive idea into something fairly simple and that makes perfectly good sense. Once you understand what it really means you'll be able to use the intuitive idea and its notation with more security.

Let $g$ be a function of *two* variables, $x$ and $h$.

$$g(x,h) = \frac{df(x)}{dx}h \quad \text{has the property that} \quad \frac{1}{h}\big|f(x+h) - f(x) - g(x,h)\big| \longrightarrow 0 \quad \text{as } h \to 0$$

That is, the function $g(x,h)$ approximates *very well* the change in $f$ as you go from $x$ to $x+h$. The difference between $g$ and $\Delta f = f(x+h) - f(x)$ goes to zero so fast that even after you've divided by $h$ the difference goes to zero.

The usual notation is to use the symbol $dx$ instead of $h$ and to call the function $df$ instead* of $g$.

$$df(x,dx) = f'(x)\,dx \quad \text{has the property that}$$
$$\frac{1}{dx}\big|f(x+dx) - f(x) - df(x,dx)\big| \longrightarrow 0 \quad \text{as } dx \to 0 \tag{8.12}$$

In this language $dx$ is just another variable that can go from $-\infty$ to $+\infty$ and $df$ is just a specified function of two variables. The point is that this function is useful because when the variable $dx$ is small $df$ provides a very good approximation to the increment $\Delta f$ in $f$.

What is the volume of the peel on an orange? The volume of a sphere is $V = 4\pi r^3/3$, so its differential is $dV = 4\pi r^2\,dr$. If the radius of the orange is $3\,\text{cm}$ and the thickness of the peel is $2\,\text{mm}$, the volume of the peel is

$$dV = 4\pi r^2\,dr = 4\pi(3\,\text{cm})^2(0.2\,\text{cm}) = 23\,\text{cm}^3$$

The whole volume of the orange is $\frac{4}{3}\pi(3\,\text{cm})^3 = 113\,\text{cm}^3$, so this peel is about 20% of the volume.

### Differentials in Several Variables
The analog of Eq. (8.11) for several variables is

$$df = df(x,y,dx,dy) = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy \tag{8.13}$$

---

* Who says that a variable in algebra must be a single letter? You would never write a computer program that way. $d\,\textit{Fred}^2/d\,\textit{Fred} = 2\,\textit{Fred}$ is perfectly sensible.

Roughly speaking, near a point in the $x$-$y$ plane, the value of the function $f$ changes as a linear function of the coordinates as you move a (little) distance away. This function $df$ describes this change to high accuracy. It bears the same relation to Eq. (8.4) that (8.11) bears to Eq. (8.3).

For example, take the function $f(x,y) = x^2 + y^2$. At the point $(x,y) = (1,2)$, the differential is

$$df(1,2,dx,dy) = (2x)\Big|_{(1,2)} dx + (2y)\Big|_{(1,2)} dy = 2dx + 4dy$$

so that

$$f(1.01, 1.99) \approx f(1,2) + df(1,2,.01,-.01) = 1^2 + 2^2 + 2(.01) + 4(-.01) = 4.98$$

compared to the exact answer, $4.9802$.

The equation analogous to (8.12) is

$$df(x,y,dx,dy) \qquad \text{has the property that}$$

$$\frac{1}{dr}\left|f(x+dx, y+dy) - f(x,y) - df(x,y,dx,dy)\right| \longrightarrow 0 \qquad \text{as } dr \to 0 \qquad (8.14)$$

where $dr = \sqrt{dx^2 + dy^2}$ is the distance to $(x,y)$. It's not that you will be able to do a lot more with this precise definition than you could with the intuitive idea. You will however be able to work with a better understanding of you're actions. When you say that "$dx$ is an infinitesimal" you can understand that this means simply that $dx$ is *any* number but that the equations using it are useful only for very small values of that number.

You can't use this notation for everything as the notation for the derivative demonstrates. The symbol "$df/dx$" does not mean to divide a function by a length; it refers to a well-defined limiting process. This notation is however constructed so that it provides an intuitive guide, and even if you *do* think of it as the function $df$ divided by the variable $dx$, you get the right answer.

Why should such a thing as a differential exist? It's essentially the first terms after the constant in the power series representation of the original function: section 2.5. But how to tell if such a series works anyway? I've been notably cavalier about proofs. The answer is that there is a proper theorem guaranteeing Eq. (8.14) works. It is that if both partial derivatives exist in the neighborhood of the expansion point and if these derivatives are continuous there, then the differential exists and has the value that I stated in Eq. (8.13). It has the properties stated in Eq. (8.14). For all this refer to one of many advanced calculus texts, such as Apostol's.*

### 8.4 Geometric Interpretation
For one variable, the picture of the differential is simple. Start with a graph of the function and at a point $(x,y) = (x, f(x))$, find the straight line that best approximates the function in the immediate neighborhood of that point. Now set up a new coordinate system with origin at this $(x,y)$ and call the new coordinates $dx$ and $dy$. In this coordinate system the straight line passes through the origin and the slope is the derivative $df(x)/dx$. The equation for the straight line is then Eq. (8.11), describing the differential.

$$dy = \frac{df(x)}{dx}dx$$

---

\* Mathematical Analysis, Addison-Wesley
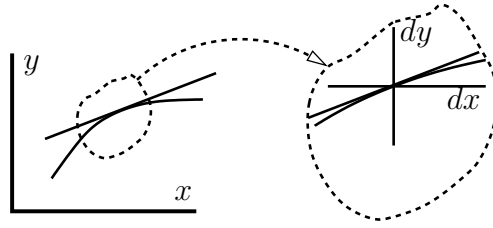
For two variables, the picture parallels this one. At a point $(x, y, z) = (x, y, f(x, y))$ find the *plane* that best approximates the function in the immediate neighborhood of that point. Set up a new coordinate system with origin at this $(x, y, z)$ and call the new coordinates $dx$, $dy$, and $dz$. The equation for a plane that passes through this origin is $\alpha\, dx + \beta\, dy + \gamma\, dz = 0$, and for this best approximating plane, the equation is nothing more than the equation for the differential, Eq. (8.13).

$$dz = \left(\frac{\partial f(x, y)}{\partial x}\right)_y dx + \left(\frac{\partial f(x, y)}{\partial y}\right)_x dy$$
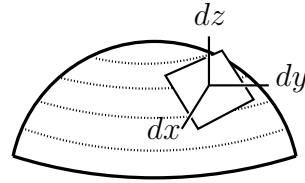


The picture is a bit harder to draw, but with a little practice you can do it.

For the case of three independent variables, I'll leave the sketch to you.

**Examples**

The temperature on the surface of a heated disk is given to be $T(r, \phi) = T_0 + T_1(1 - r^2/a^2)$, where $a$ is the radius of the disk and $T_0$ and $T_1$ are constants. If you start at position $x = c < a$, $y = 0$ and move parallel to the $y$-axis at speed $v_0$ what is the rate of change of temperature that you feel?

Use Eq. (8.4), and the relation $r = \sqrt{x^2 + y^2}$.

$$\frac{dT}{dt} = \left(\frac{\partial T}{\partial r}\right)_\phi \frac{dr}{dt} + \left(\frac{\partial T}{\partial \phi}\right)_r \frac{d\phi}{dt} = \left(\frac{\partial T}{\partial r}\right)_\phi \left[\left(\frac{\partial r}{\partial x}\right)_y \frac{dx}{dt} + \left(\frac{\partial r}{\partial y}\right)_x \frac{dy}{dt}\right]$$

$$= \left(-2T_1 \frac{r}{a^2}\right)\left[\frac{y}{\sqrt{x^2 + y^2}} v_0\right] = -2T_1 \frac{\sqrt{c^2 + v_0^2 t^2}}{a^2} \cdot \frac{v_0^2 t}{\sqrt{c^2 + v_0^2 t^2}}$$

$$= -2T_1 \frac{v_0^2 t}{a^2}$$
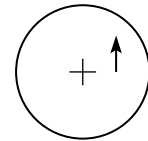


As a check, the dimensions are correct (are they?). At time zero, this vanishes, and that's what you should expect because at the beginning of the motion you're starting to move in the direction *perpendicular* to the direction in which the temperature is changing. The farther you go, the more nearly parallel to the direction of the radius you're moving. If you are moving exactly parallel to the radius, this time-derivative is easier to calculate; it's then almost a problem in a single variable.

$$\frac{dT}{dt} \approx \frac{dT}{dr}\frac{dr}{dt} \approx -2T_1 \frac{r}{a^2} v_0 \approx -2T_1 \frac{V_0 t}{a^2} v_0$$

So the approximate and the exact calculation agree. In fact they agree so well that you should try to find out if this is a lucky coincidence or if there some special aspect of the problem that you might have seen from the beginning and that would have made the whole thing much simpler.

**8.5 Gradient**

The equation (8.13) for the differential has another geometric interpretation. For a function such as $f(x,y) = x^2 + 4y^2$, the equations representing constant values of $f$ describe curves in the $x$-$y$ plane. In this example, they are ellipses. If you start from any fixed point in the plane and start to move away from it, the rate at which the value of $f$ changes will depend on the direction in which you move. If you move along the curve defined by $f = $ constant then $f$ won't change at all. If you move perpendicular to that direction then $f$ may change a lot.

> The gradient of $f$ at a point is the vector pointing in the direction in which $f$ is increasing most rapidly, and the component of the gradient along that direction is the derivative of $f$ with respect to the distance in that direction.

To relate this to the partial derivatives that we've been using, and to understand how to compute and to use the gradient, return to Eq. (8.13) and write it in vector form. Use the common notation for the basis: $\hat{x}$ and $\hat{y}$. Then let

$$d\vec{r} = dx\,\hat{x} + dy\,\hat{y} \qquad \text{and} \qquad \vec{G} = \left(\frac{\partial f}{\partial x}\right)_y \hat{x} + \left(\frac{\partial f}{\partial y}\right)_x \hat{y} \qquad (8.15)$$

The equation for the differential is now

$$df = df(x, y, dx, dy) = \vec{G} \cdot d\vec{r} \qquad (8.16)$$
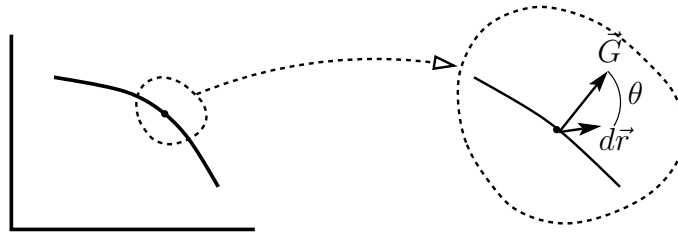


Because you know the properties of the dot product, you know that this is $G\,dr\cos\theta$ and it is largest when the directions of $d\vec{r}$ and of $\vec{G}$ are the same. It's zero when they are perpendicular. You also know that $df$ is zero when $d\vec{r}$ is in the direction along the curve where $f$ is constant. The vector $\vec{G}$ is therefore perpendicular to this curve. It is in the direction in which $f$ is changing most rapidly. Also because $df = G\,dr\cos 0$, you see that $G$ is the derivative of $f$ with respect to distance along that direction. $\vec{G}$ is the gradient.

For the example $f(x,y) = x^2 + 4y^2$, $\vec{G} = 2x\,\hat{x} + 8y\,\hat{y}$. At each point in the $x$-$y$ plane it provides a vector showing the steepness of $f$ at that point and the direction in which $f$ is changing most rapidly.
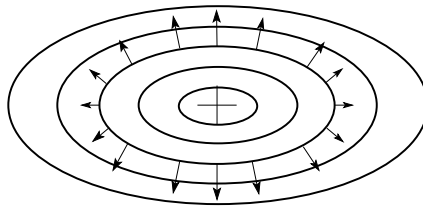


Notice that the gradient vectors are twice as long where the ellipses are closest together as they are at the ends where the ellipses are farthest apart. The function changes more rapidly in the $y$-direction.

The U.S.Coast and Geodetic Survey makes a large number of maps, and hikers are particularly interested in the contour maps. They show curves indicating the lines of constant altitude. When Apollo 16 went to the Moon in 1972, NASA prepared a similar map for the astronauts, and this is a small segment of that map. The contour lines represent 10 meter increments in altitude.*
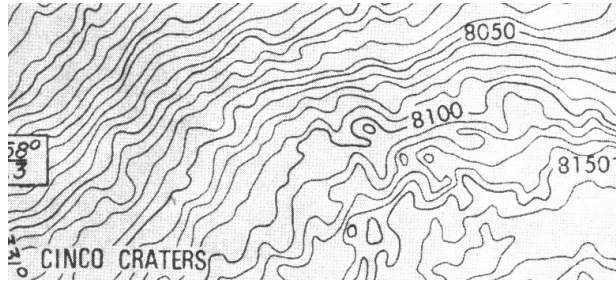


The gravitational potential energy of a mass $m$ near the Earth's (or Moon's) surface is $mgh$. This divided by $m$ is the gravitational potential, $gh$. These lines of constant altitude are then lines of constant potential, equipotentials of the gravitational field. Walk along an equipotential and you are doing no work against gravity, just walking on the level.

### 8.6 Electrostatics
The electric field can be described in terms of a gradient. For a single point charge at the origin the electric field is

$$\vec{E}(x, y, z) = \frac{kq}{r^2}\hat{r}$$

where $\hat{r}$ is the unit vector pointing away from the origin and $r$ is the distance to the origin. This vector can be written as a gradient. Because this $\vec{E}$ is everywhere pointing away from the origin, it's everywhere perpendicular to the sphere centered at the origin.

$$\vec{E} = -\text{grad}\,\frac{kq}{r}$$

You can verify this a several ways. The first is to go straight to the definition of a gradient. (There's a blizzard of minus signs in this approach, so have a little patience. It will get better.) This function is increasing most rapidly in the direction moving toward the origin. $(1/r)$ The derivative with respect to distance in this direction is $-d/dr$, so
$-d/dr(1/r) = +1/r^2$. The direction of greatest increase is along $-\hat{r}$, so $\text{grad}\,(1/r) = -\hat{r}(1/r^2)$. But the relation to the electric field has another $-1$ in it, so

$$-\text{grad}\frac{kq}{r} = +\hat{r}\frac{kq}{r^2}$$

There's got to be a better way.

Yes, instead of insisting that you move in the direction in which the function is *increasing* most rapidly, simply move in the direction in which it is changing most rapidly. The derivative with respect to distance in that direction is the component in that direction and the plus or minus signs take care of themselves. The derivative with respect to $r$ of $(1/r)$ is $-1/r^2$. That is the component in the direction $\hat{r}$, the direction in which you took the derivative. This says $\text{grad}\,(1/r) = -\hat{r}(1/r^2)$. You get the same result as before but without so much fussing. This also makes it look more like the familiar ordinary derivative in one dimension.

---

* history.nasa.gov/alsj/a16/ scan by Robin Wheeler

Still another way is from the Stallone-Schwarzenegger brute force school of computing. Put everything in rectangular coordinates and do the partial derivatives using Eqs. (8.15) and (8.6).

$$\left(\frac{\partial(1/r)}{\partial x}\right)_{y,z} = \left(\frac{\partial(1/r)}{\partial r}\right)_{\theta,\phi}\left(\frac{\partial r}{\partial x}\right)_{y,z} = -\frac{1}{r^2}\frac{\partial}{\partial x}\sqrt{x^2+y^2+z^2} = -\frac{1}{r^2}\frac{x}{\sqrt{x^2+y^2+z^2}}$$

Repeat this for $y$ and $z$ with similar results and assemble the output.

$$-\mathrm{grad}\frac{kq}{r} = \frac{kq}{r^2}\frac{x\,\hat{x}+y\,\hat{y}+z\,\hat{z}}{\sqrt{x^2+y^2+z^2}} = \frac{kq}{r^2}\frac{\vec{r}}{r} = \frac{kq}{r^2}\hat{r}$$

The symbol $\nabla$ is commonly used for the gradient operator. This vector operator will appear in several other places, the curl of a vector field will be the one you see most often.

$$\nabla = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z} \tag{8.17}$$

From Eq. (8.15) you have

$$\mathrm{grad}\,f = \nabla f \tag{8.18}$$

### 8.7 Plane Polar Coordinates

When doing integrals in the plane there are many coordinate systems to choose from, but rectangular and polar coordinates are the most common. You can find the element of area with a simple sketch: The lines (or curves) of constant coordinate enclose an area that is, for small enough increments in the coordinates, a rectangle. Then you just multiply the sides. In one case $\Delta x \cdot \Delta y$ and in the other case $\Delta r \cdot r\Delta\phi$.
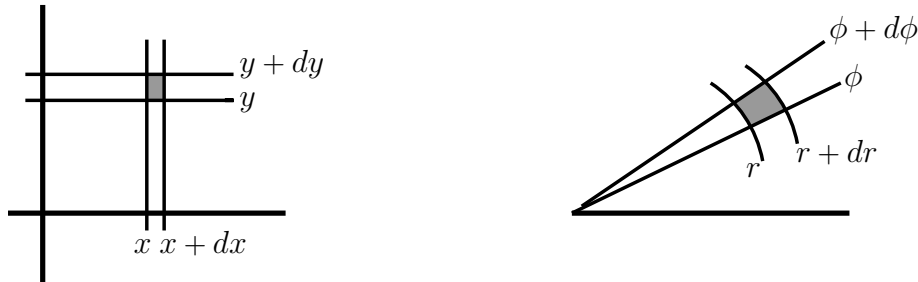


### Vibrating Drumhead

A circular drumhead can vibrate in many complicated ways. The simplest and lowest frequency mode is approximately
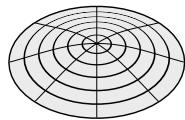


$$z(r,\phi,t) = z_0\left(1 - r^2/R^2\right)\cos\omega t \tag{8.19}$$

where $R$ is the radius of the drum and $\omega$ is the frequency of oscillation. (The shape is more accurately described by Eq. (4.22) but this approximation is pretty good for a start.) The kinetic energy density of the moving drumhead is $u = \frac{1}{2}\sigma(\partial z/\partial t)^2$. That is, in a small area $\Delta A$, the kinetic energy is $\Delta K = u\Delta A$ and the limit as $\Delta A \to 0$ of $\Delta K/\Delta A$ is the area-energy-density. In the same way, $\sigma$ is the area mass density, $dm/dA$.

What is the total kinetic energy because of this oscillation? It is $\int u\,dA = \int u\,d^2r$. To evaluate it, use polar coordinates and integrate over the area of the drumhead. The notation $d^2r$ is another notation for $dA$ just as $d^3r$ is used for a piece of volume.

$$\int u\, dA = \int_0^R r\, dr \int_0^{2\pi} d\phi\, \frac{\sigma}{2} z_0^2 \big((1 - r^2/R^2)\omega \sin \omega t\big)^2$$

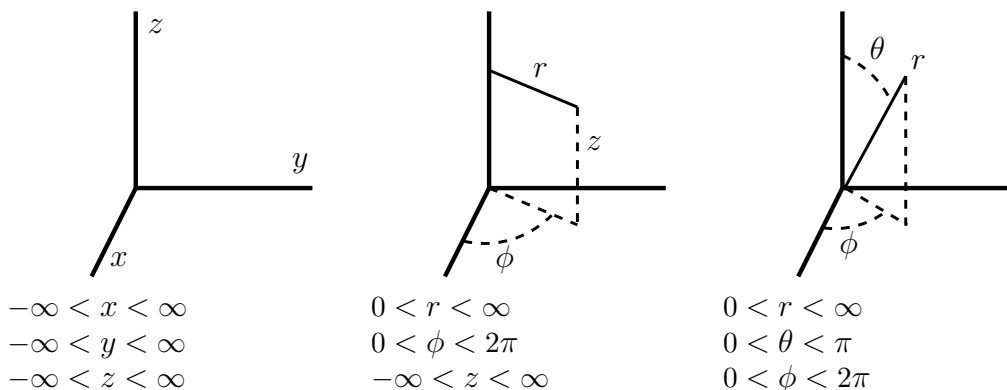$$= \frac{\sigma}{2} 2\pi z_0^2 \omega^2 \sin^2 \omega t \int_0^R dr\, r \big(1 - r^2/R^2\big)^2$$

$$= \sigma \pi z_0^2 \omega^2 \sin^2 \omega t \frac{1}{2} \int_{r=0}^{r=R} d(r^2) \big(1 - r^2/R^2\big)^2 \qquad (8.20)$$

$$= \sigma \pi z_0^2 \omega^2 \sin^2 \omega t \frac{1}{2} R^2 \frac{1}{3} \big(1 - r^2/R^2\big)^3 (-1)\Big|_0^{r=R}$$

$$= \frac{1}{6} \sigma R^2 \pi z_0^2 \omega^2 \sin^2 \omega t$$

See problem 8.10 and following for more on this.*

### 8.8 Cylindrical, Spherical Coordinates

The three common coordinate systems used in three dimensions are rectangular, cylindrical, and spherical coordinates, and these are the ones you have to master. When you need to use prolate spheroidal coordinates you can look them up.



$$-\infty < x < \infty \qquad\qquad 0 < r < \infty \qquad\qquad 0 < r < \infty$$
$$-\infty < y < \infty \qquad\qquad 0 < \phi < 2\pi \qquad\qquad 0 < \theta < \pi$$
$$-\infty < z < \infty \qquad\qquad -\infty < z < \infty \qquad\qquad 0 < \phi < 2\pi$$

The surfaces that have constant values of these coordinates are planes in rectangular coordinates; planes and cylinders in cylindrical; planes, spheres, and cones in spherical. In every one of these cases the constant-coordinate surfaces intersect each other at right angles, hence the name "orthogonal coordinate" systems. In spherical coordinates I used the coordinate $\theta$ as the angle from the $z$-axis and $\phi$ as the angle around the axis. In mathematics books these are typically reversed, so watch out for the notation. On the globe of the Earth, $\phi$ is like the longitude and $\theta$ like the latitude except that longitude goes 0 to $180°$ East and 0 to $180°$ West from the Greenwich meridian instead of zero to $2\pi$. Latitude is 0 to $90°$ North or South from the equator instead of zero to $\pi$ from the pole. Except for the North-South terminology, latitude is $90° - \theta$.

The volume elements for these systems come straight from the drawings, just as the area elements do in plane coordinates. In every case you can draw six surfaces, bounded by constant coordinates, and surrounding a small box. Because these are orthogonal coordinates you can compute the volume of the box easily as the product of its three edges.

In the spherical case, one side is $\Delta r$. Another side is $r\Delta\theta$. The third side is not $r\Delta\phi$; it is $r \sin\theta \Delta\phi$. The reason for the factor $\sin\theta$ is that the arc of the circle made at constant $r$ and constant

---

* For some animations showing these oscillations and others, check out
www.physics.miami.edu/nearing/mathmethods/drumhead-animations.html

$\theta$ is not in a plane passing through the origin. It is in a plane parallel to the $x$-$y$ plane, so it has a radius $r \sin \theta$.
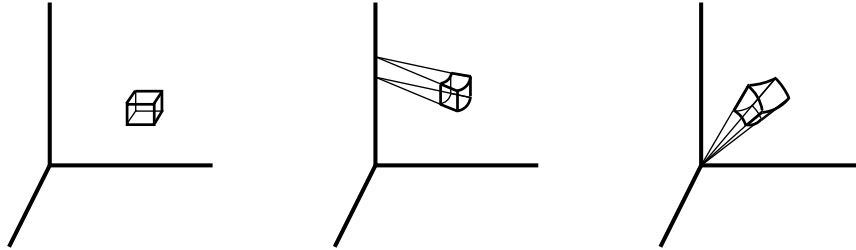
| | rectangular | cylindrical | spherical |
|---|---|---|---|
| volume | $d^3r = \quad dx\,dy\,dz$ | $r\,dr\,d\phi\,dz$ | $r^2 \sin\theta\,dr\,d\theta\,d\phi$ |
| area | $d^2r = \quad dx\,dy$ | $r\,d\phi\,dz \;$ or $\; r\,d\phi\,dr$ | $r^2 \sin\theta\,d\theta\,d\phi$ |

### Examples of Multiple Integrals

Even in rectangular coordinates integration can be tricky. That's because you have to pay attention to the limits of integration far more closely than you do for simple one dimensional integrals. I'll illustrate this with two dimensional rectangular coordinates first, and will choose a problem that is easy but still shows what you have to look for.

### An Area

Find the area in the $x$-$y$ plane between the curves $y = x^2/a$ and $y = x$.

$$\text{(A)} \int_0^a dx \int_{x^2/a}^x dy\, 1 \qquad \text{and} \qquad \text{(B)} \int_0^a dy \int_y^{\sqrt{ay}} dx\, 1$$
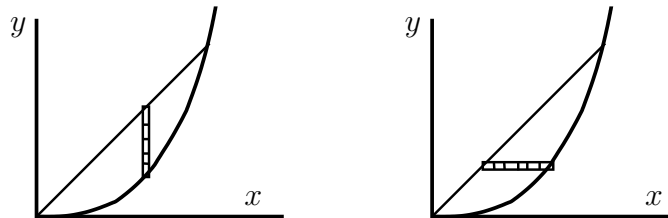
In the first instance I fix $x$ and add the pieces of $dy$ in the strip indicated. The lower limit of the $dy$ integral comes from the specified equation of the lower curve. The upper limit is the value of $y$ for the given $x$ at the upper curve. After that the limits on the sum over $dx$ comes from the intersection of the two curves: $y = x = x^2/a$ gives $x = a$ for that limit.

In the second instance I fix $y$ and sum over $dx$ first. The left limit is easy, $x = y$, and the upper limit comes from solving $y = x^2/a$ for $x$ in terms of $y$. When that integral is done, the remaining $dy$ integral starts at zero and goes up to the intersection at $y = x = a$.

Now do the integrals.

$$\text{(A)} \int_0^a dx \left[x - x^2/a\right] = \frac{a^2}{2} - \frac{a^3}{3a} = \frac{a^2}{6}$$

$$\text{(B)} \int_0^a dy \left[\sqrt{ay} - y\right] = a^{1/2}\frac{a^{3/2}}{3/2} - \frac{a^2}{2} = \frac{a^2}{6}$$

If you would care to try starting this calculation from the beginning, *without* drawing any pictures, be my guest.
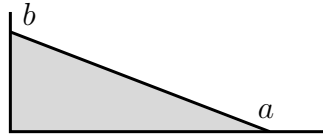
## A Moment of Inertia

The moment of inertia about an axis is $\int r_\perp^2 \, dm$. Here, $r_\perp$ is the perpendicular distance to the axis. What is the moment of inertia of a uniform sheet of mass $M$ in the shape of a right triangle of sides $a$ and $b$? Take the moment about the right angled vertex. The area mass density, $\sigma = dm/dA$ is $2M/ab$. The moment of inertia is then

$$
\int (x^2 + y^2)\sigma \, dA = \int_0^a dx \int_0^{b(a-x)/a} dy \, \sigma (x^2 + y^2) = \int_0^a dx \, \sigma \left[ x^2 y + y^3/3 \right]_0^{b(a-x)/a}
$$

$$
= \int_0^a dx \, \sigma \left[ x^2 \frac{b}{a}(a - x) + \frac{1}{3}\left(\frac{b}{a}\right)^3 (a - x)^3 \right]
$$

$$
= \sigma \left[ \frac{b}{a}\left(\frac{a^4}{3} - \frac{a^4}{4}\right) + \frac{1}{3}\left(\frac{b^3}{a^3}\frac{a^4}{4}\right) \right]
$$

$$
= \frac{1}{12}\sigma \left(ba^3 + ab^3\right) = \frac{M}{6}\left(a^2 + b^2\right)
$$

The dimensions are correct. For another check take the case where $a = 0$, reducing this to $Mb^2/6$. But wait, this now looks like a thin rod, and I remember that the moment of inertia of a thin rod about its end is $Mb^2/3$. What went wrong? Nothing. Look again more closely. Show why this limiting answer ought to be less than $Mb^2/3$.

## Volume of a Sphere

What is the volume of a sphere of radius $R$? The most obvious approach would be to use spherical coordinates. See problem 8.16 for that. I'll use cylindrical coordinates instead. The element of volume is $dV = r \, dr d\phi dz$, and the integrals can be done a couple of ways.

$$
\int d^3r = \int_0^R r \, dr \int_0^{2\pi} d\phi \int_{-\sqrt{R^2 - r^2}}^{+\sqrt{R^2 - r^2}} dz = \int_{-R}^{+R} dz \int_0^{2\pi} d\phi \int_0^{\sqrt{R^2 - z^2}} r \, dr \qquad (8.21)
$$

You can finish these now, see problem 8.17.

## A Surface Charge Density

An example that appears in electrostatics: The surface charge density, $dq/dA$, on a sphere of radius $R$ is $\sigma(\theta, \phi) = \sigma_0 \sin^2 \theta \cos^2 \phi$. What is the total charge on the sphere?

The element of area is $R^2 \sin \theta \, d\theta \, d\phi$, so the total charge is $\int \sigma \, dA$,

$$
Q = \int_0^\pi \sin \theta \, d\theta \, R^2 \int_0^{2\pi} d\phi \, \sigma_0 \sin^2 \theta \cos^2 \phi = R^2 \int_{-1}^{+1} d\cos\theta \, \sigma_0 \left(1 - \cos^2 \theta\right) \int_0^{2\pi} d\phi \, \cos^2 \phi
$$

The mean value of $\cos^2$ is $1/2$. so the $\phi$ integral gives $\pi$. For the rest, it is

$$
\sigma_0 \pi R^2 \left[ \cos \theta - \frac{1}{3}\cos^3 \theta \right]_{-1}^{+1} = \frac{4}{3}\sigma_0 \pi R^2
$$

## Limits of Integration

Sometimes the trickiest part of multiple integrals is determining the limits of integration. Especially

when you have to change the order of integration, the new limits may not be obvious. Are there any special techniques or tricks to doing this? Yes, there is one, perhaps obscure, method that you may not be accustomed to.

*Draw Pictures.*

If you have an integral such as the first one, you have to draw a picture of the integration domain to switch limits.

$$\int_0^1 dy \int_y^{\sqrt{2-y^2}} dx\, f(x,y) \qquad \left[\int_0^1 dx \int_0^x dy + \int_1^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} dy\right] f(x,y) \quad (8.22)$$

Of course, once you've drawn the picture you may realize that simply interchanging the order of integration won't help, but polar coordinates may.

$$\int_0^{\sqrt{2}} r\, dr \int_0^{\pi/4} d\phi$$

## 8.9 Vectors: Cylindrical, Spherical Bases
When you describe vectors in three dimensions are you restricted to the basis $\hat{x}$, $\hat{y}$, $\hat{z}$? In a different coordinate system you should use basis vectors that are adapted to that system. In rectangular coordinates these vectors have the convenient property that they point along the direction perpendicular to the plane where the corresponding coordinate is constant. They also point in the direction in which the other two coordinates are constant. *E.g.* the unit vector $\hat{x}$ points perpendicular to the plane of constant $x$ (the $y$-$z$ plane); it also point along the line where $y$ and $z$ are constant.

Do the same thing for cylindrical coordinates. The unit vector $\hat{z}$ points perpendicular to the $x$-$y$ plane. The unit vector $\hat{r}$ points perpendicular to the cylinder $r =$ constant. The unit vector $\hat{\phi}$ points perpendicular to the plane $\phi =$ constant and along the direction for which $r$ and $z$ are constant. The conventional right-hand rule specifies $\hat{z} = \hat{r} \times \hat{\phi}$.

For spherical coordinates $\hat{r}$ points perpendicular to the sphere $r =$ constant. The $\hat{\phi}$ vector is perpendicular to the plane $\phi =$ constant and points along the direction where $r =$ constant and $\theta =$ constant and toward increasing coordinate $\phi$. Finally $\hat{\theta}$ is perpendicular to the cone $\theta =$ constant and again, points toward increasing $\theta$. Then $\hat{\phi} = \hat{r} \times \hat{\theta}$, and on the Earth, these vectors $\hat{r}$, $\hat{\theta}$, and $\hat{\phi}$ are up, South, and East.

## Solenoid
A standard solenoid is cylindrical coil of wire, so that when the wire carries a current it produces a magnetic field. To describe this field, it seems that cylindrical coordinates are advised. Until you know something about the field the most general thing that you can write is

$$\vec{B}(r,\phi,z) = \hat{r}\, B_r(r,\phi,z) + \hat{\phi}\, B_\phi(r,\phi,z) + \hat{z}\, B_z(r,\phi,z)$$

In a real solenoid that's it; all three of these components are present. If you have an ideal, infinitely long solenoid, with the current going strictly around in the $\hat{\phi}$ direction, (found only in textbooks) the use of Maxwell's equations and appropriately applied symmetry arguments will simplify this to $\hat{z}\,B_z(r)$.

**Gravitational Field**
The gravitational field of the Earth is simple, $\vec{g} = -\hat{r}\,GM/r^2$, pointing straight toward the center of the Earth. Well no, not really. The Earth has a bulge at the equator; its equatorial diameter is about $43$ km larger than its polar diameter. This changes the $\vec{g}$-field so that it has a noticeable $\hat{\theta}$ component. At least it's noticeable if you're trying to place a satellite in orbit or to send a craft to another planet.

A better approximation to the gravitational field of the Earth is

$$\vec{g} = -\hat{r}\frac{GM}{r^2} - G\frac{3Q}{r^4}\left[\hat{r}(3\cos^2\theta - 1)/2 + \hat{\theta}\cos\theta\sin\theta\right] \tag{8.23}$$

The letter $Q$ stands for the quadrupole moment. $|Q| \ll MR^2$, and it's a measure of the bulge. By convention a football (American football) has a positive $Q$; the Earth's $Q$ is negative. (What about a European football?)

**Nuclear Magnetic Field**
The magnetic field from the nucleus of many atoms (even as simple an atom as hydrogen) is proportional to

$$\frac{1}{r^3}\left[2\hat{r}\cos\theta + \hat{\theta}\sin\theta\right] \tag{8.24}$$

As with the preceding example these are in spherical coordinates, and the component along the $\hat{\phi}$ direction is zero. This field's effect on the electrons in the atom is small but detectable. The magnetic properties of the nucleus are central to the subject of nuclear magnetic resonance (NMR), and that has its applications in magnetic resonance imaging* (MRI).

**8.10 Gradient in other Coordinates**
The equation for the gradient computed in rectangular coordinates is Eq. (8.15) or (8.18). How do you compute it in cylindrical or spherical coordinates? You do it the same way that you got Eq. (8.15) from Eq. (8.13). The coordinates $r$, $\phi$, and $z$ are just more variables, so Eq. (8.13) is simply

$$df = df(r, \phi, z, dr, d\phi, dz) = \left(\frac{\partial f}{\partial r}\right)_{\phi, z} dr + \left(\frac{\partial f}{\partial \phi}\right)_{r, z} d\phi + \left(\frac{\partial f}{\partial z}\right)_{r, \phi} dz \tag{8.25}$$

All that's left is to write $d\vec{r}$ in these coordinates, just as in Eq. (8.15).

$$d\vec{r} = \hat{r}\,dr + \hat{\phi}\,r\,d\phi + \hat{z}\,dz \tag{8.26}$$

The part in the $\hat{\phi}$ direction is the *displacement* of $d\vec{r}$ in that direction. As $\phi$ changes by a small amount the distance moved is not $d\phi$; it is $r\,d\phi$. The equation

$$df = df(r, \phi, z, dr, d\phi, dz) = \text{grad}\,f \cdot d\vec{r}$$

combined with the two equations (8.25) and (8.26) gives $\text{grad}\,f$ as

$$\text{grad}\,f = \hat{r}\frac{\partial f}{\partial r} + \hat{\phi}\frac{1}{r}\frac{\partial f}{\partial \phi} + \hat{z}\frac{\partial f}{\partial z} = \nabla f \tag{8.27}$$

---

\* In medicine MRI was originally called NMR, but someone decided that this would disconcert the patients.

Notice that the units work out right too.

In spherical coordinates the procedure is identical. All that you have to do is to identify what $d\vec{r}$ is.

$$d\vec{r} = \hat{r}\, dr + \hat{\theta}\, r\, d\theta + \hat{\phi}\, r \sin\theta\, d\phi$$

Again with this case you have to look at the distance moved when the coordinates changes by a small amount. Just as with cylindrical coordinates this determines the gradient in spherical coordinates.

$$\operatorname{grad} f = \hat{r}\frac{\partial f}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial f}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi} = \nabla f \tag{8.28}$$

The equations (8.15), (8.27), and (8.28) define the gradient (and correspondingly $\nabla$) in three coordinate systems.

## 8.11 Maxima, Minima, Saddles

With one variable you can look for a maximum or a minimum by taking a derivative and setting it to zero. For several variables you do it several times so that you will get as many equations as you have unknown coordinates.

Put this in the language of gradients: $\nabla f = 0$. The derivative of $f$ vanishes in every direction as you move from such a point. As examples,

$$f(x, y) = x^2 + y^2, \qquad \text{or} \qquad = -x^2 - y^2, \qquad \text{or} \qquad = x^2 - y^2$$

For all three of these the gradient is zero at $(x, y) = (0, 0)$; the first has a minimum there, the second a maximum, and the third neither — it is a "saddle point." Draw a picture to see the reason for the name. The generic term for all three of these is "critical point."

An important example of finding a minimum is "least square fitting" of functions. How close are two functions to each other? The most commonly used, and in every way the simplest, definition of the distance (squared) between $f$ and $g$ on the interval $a < x < b$ is

$$\int_a^b dx\, |f(x) - g(x)|^2 \tag{8.29}$$

This means that a large deviation of one function from the other in a small region counts more than smaller deviations spread over a larger domain. The square sees to that. As a specific example, take a function $f$ on the interval $0 < x < L$ and try to fit it to the sum of a couple of trigonometric functions. The best fit will be the one that minimizes the distance between $f$ and the sum. (Take $f$ to be a real-valued function for now.)

$$D^2(\alpha, \beta) = \int_0^L dx\, \left(f(x) - \alpha \sin\frac{\pi x}{L} - \beta \sin\frac{2\pi x}{L}\right)^2 \tag{8.30}$$

$D$ is the distance between the given function and the sines used to fit it. To minimize the distance, take derivatives with respect to the parameters $\alpha$ and $\beta$.

$$\frac{\partial D^2}{\partial \alpha} = 2\int_0^L dx\, \left(f(x) - \alpha \sin\frac{\pi x}{L} - \beta \sin\frac{2\pi x}{L}\right)\left(-\sin\frac{\pi x}{L}\right) = 0$$

$$\frac{\partial D^2}{\partial \beta} = 2\int_0^L dx\, \left(f(x) - \alpha \sin\frac{\pi x}{L} - \beta \sin\frac{2\pi x}{L}\right)\left(-\sin\frac{2\pi x}{L}\right) = 0$$

These two equations determine the parameters $\alpha$ and $\beta$.

$$\alpha \int_0^L dx \, \sin^2 \frac{\pi x}{L} = \int_0^L dx \, f(x) \sin \frac{\pi x}{L}$$

$$\beta \int_0^L dx \, \sin^2 \frac{2\pi x}{L} = \int_0^L dx \, f(x) \sin \frac{2\pi x}{L}$$

The other integrals vanish because of the orthogonality of $\sin \pi x/L$ and $\sin 2\pi x/L$ on this interval. What you get is exactly the coefficients of the Fourier series expansion of $f$. The Fourier series is the best fit (in the least square sense) of a sum of orthogonal functions to $f$. See section 11.6 for more on this

   Is it a minimum? Yes. Look at the coefficients of $\alpha^2$ and $\beta^2$ in Eq. (8.30). They are positive; $+\alpha^2 + \beta^2$ has a minimum, not a maximum or saddle point, and there is no cross term in $\alpha\beta$ to mess it up.

   The distance function Eq. (8.29) is simply (the square of) the norm in the vector space sense of the difference of the two vectors $f$ and $g$. Equations(6.12) and (6.7) here become

$$\|f - g\|^2 = \langle f - g, f - g \rangle = \int_a^b dx \, |f(x) - g(x)|^2$$



shortest distance to the plane
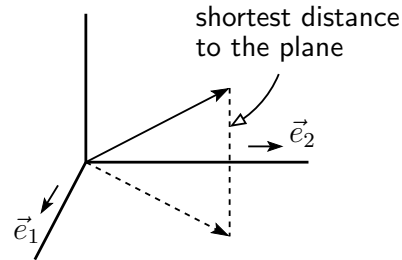
The geometric meaning of Eq. (8.30) is that $\vec{e}_1$ and $\vec{e}_2$ provide a basis for the two dimensional space

$$\alpha \vec{e}_1 + \beta \vec{e}_2 = \alpha \sin \frac{\pi x}{L} + \beta \sin \frac{2\pi x}{L}$$

The plane is the set of all linear combinations of the two vectors, and for a general vector not in this plane, the shortest distance to the plane defines the vector *in* the plane that is the best fit to the given vector. It's the one that's closest. Because the vectors $\vec{e}_1$ and $\vec{e}_2$ are orthogonal it makes it easy to find the closest vector. You require that the difference, $\vec{v} - \alpha\vec{e}_1 - \beta\vec{e}_2$ has only an $\vec{e}_3$ component. That is Fourier series.

**Hessian**
In this example leading to Fourier components, it's pretty easy to see that you are dealing with a minimum and not anything else. In other situations it may not be so easy. You may have a lot of variables. You may have complicated cross terms. Is $x^2 + xy + y^2$ a minimum at the origin? Is $x^2 + 3xy + y^2$? (Yes and No respectively.)

   When there's just one variable there is a simple rule that lets you decide. Check the second derivative. If it's positive you have a minimum; if it's negative you have a maximum. If it's zero you have more work to do. Is there a similar method for several variables? Yes, and I'll show it explicitly for two variables. Once you see how to do it in two dimensions, the generalization to $N$ is just a matter of how much work you're willing to do (or how much computer time you can use).

   The Taylor series in two variables, Eq. (2.16), is to second order

$$f(x + dx, y + dy) = f(x, y) + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dx \, dy + \frac{\partial^2 f}{\partial y^2}dy^2 + \cdots$$

Write this in a more compact notation in order to emphasize the important parts.

$$f(\vec{r} + d\vec{r}) - f(\vec{r}) = \nabla f \cdot d\vec{r} + \langle d\vec{r}, H\, d\vec{r}\rangle + \cdots$$

The part with the gradient is familiar, and to have either a minimum or a maximum, that will have to be zero. The next term introduces a new idea, the *Hessian*, constructed from all the second derivative terms. Write these second order terms as a matrix to see what they are, and in order to avoid a lot of clumsy notation use subscripts as an abbreviation for the partial derivatives.

$$\langle d\vec{r}, H\, d\vec{r}\rangle = (\,dx \quad dy\,) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} \qquad \text{where} \qquad d\vec{r} = \hat{x}\,dx + \hat{y}\,dy \qquad (8.31)$$

This matrix is symmetric because of the properties of mixed partials. How do I tell from this whether the function $f$ has a minimum or a maximum (or neither) at a point where the gradient of $f$ is zero? Eq. (8.31) describes a function of two variables even *after* I've fixed the values of $x$ and $y$ by saying that $\nabla f = 0$. It is a quadratic function of $dx$ and $dy$. Expressed in the language of vectors this says that $f$ has a minimum if (8.31) is positive no matter what the direction of $d\vec{r}$ is — $H$ is *positive definite*.

Pull back from the problem a step. This is a $2 \times 2$ symmetric matrix sandwiched inside a scalar product.

$$h(x, y) = (\,x \quad y\,) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad (8.32)$$

Is $h$ positive definite? That is, positive for all $x$, $y$? If this matrix is diagonal it's much easier to see what is happening, so diagonalize it. Find the eigenvectors and use those for a basis.

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{requires} \qquad \det \begin{pmatrix} a - \lambda & b \\ b & c - \lambda \end{pmatrix} = 0$$

$$\lambda^2 - \lambda(a + c) + ac - b^2 = 0 \implies \lambda = \Big[(a + c) \pm \sqrt{(a - c)^2 + b^2}\Big]/2 \qquad (8.33)$$

For the applications here all the $a$, $b$, $c$ are the real partial derivatives, so the eigenvalues are real and the only question is whether the $\lambda$s are positive or negative, because they will be the (diagonal) components of the Hessian matrix in the new basis. If this is a double root, the matrix was already diagonal. You can verify that the eigenvalues are positive if $a > 0$, $c > 0$, and $4ac > b^2$, and that will indicate a minimum point.

Geometrically the equation $z = h(x, y)$ from Eq. (8.32) defines a surface. If it is positive definite the surface is a paraboloid opening upward. If negative definite it is a paraboloid opening down. The mixed case is a hyperboloid — a saddle.

In this $2 \times 2$ case you have a quadratic formula to fall back on, and with more variables there are standard algorithms for determining eigenvalues of matrices, but I'll leave those to some other book.

## 8.12 Lagrange Multipliers
This is an incredibly clever method to handle problems of maxima and minima in several variables when there are constraints.

An example: "What is the largest rectangle?" obviously has no solution, but "What is the largest rectangle contained in an ellipse?" does.

Another: Particles are to be placed into states of specified energies. You know the total number of particles; you know the total energy. All else being equal, what is the most probable distribution of the number of particles in each state?

I'll describe this procedure for two variables; it's the same for more. The problem stated is that I want to find the maximum (or minimum) of a function $f(x,y)$ given the fact that the coordinates $x$ and $y$ must lie on the curve $\phi(x,y) = 0$. If you can solve the $\phi$ equation for $y$ in terms of $x$ explicitly, then you can substitute it into $f$ and turn it into a problem in ordinary one variable calculus. What if you can't?

Analyze this graphically. The equation $\phi(x,y) = 0$ represents one curve in the plane. The succession of equations $f(x,y) = $ constant represent many curves in the plane, one for each constant. Think of equipotentials.
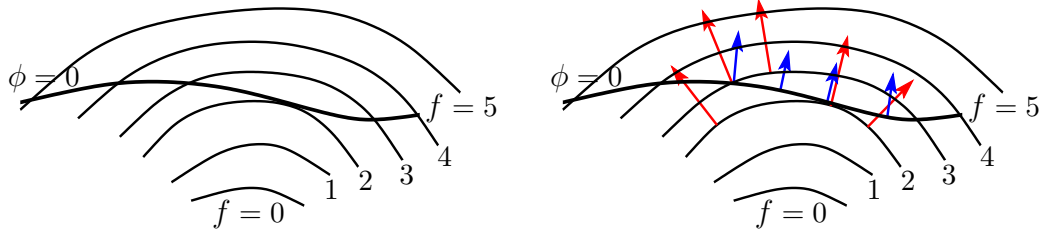


Look at the intersections of the $\phi$-curve and the $f$-curves. Where they intersect, they will usually cross each other. Ask if such a crossing could possibly be a point where $f$ is a maximum. Clearly the answer is no, because as you move along the $\phi$-curve you're then moving from a point where $f$ has one value to where it has another.

The one way to have $f$ be a maximum at a point on the $\phi$-curve is for the two curves to touch and not to cross. When that happens the values of $f$ will increase as you approach the point from one side and decrease on the other. That makes it a maximum. In this sketch, the values of $f$ decrease from 4 to 3 to 2 and then back to 3, 4, and 5. This point where the curve $f = 2$ touches the $\phi = 0$ curve is then a minimum of $f$ along $\phi = 0$.

To implement this picture so that you can compute with it, look at the gradient of $f$ and the gradient of $\phi$. The gradient vectors are perpendicular to the curves $f =$ constant and $\phi =$ constant respectively, and at the point where the curves are tangent to each other these gradients are in the same direction (or opposite, no matter). Either way one vector is a scalar times the other.

$$\nabla f = \lambda \nabla \phi \qquad\qquad (8.34)$$

In the second picture, the arrows are the gradient vectors for $f$ and for $\phi$. Break this into components and you have

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial \phi}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} - \lambda \frac{\partial \phi}{\partial y} = 0, \qquad \phi(x,y) = 0$$

There are three equations in three unknowns $(x, y, \lambda)$, and these are the equations to solve for the position of the maximum or minimum value of $f$. You are looking for $x$ and $y$, so you'll be tempted to ignore the third variable $\lambda$ and to eliminate it. Look again. This parameter, the Lagrange multiplier, has a habit of being significant.

**Examples of Lagrange Multipliers**
The first example that I mentioned: What is the largest rectangle that you can inscribe in an ellipse? Let the ellipse and the rectangle be centered at the origin. The upper right corner of the rectangle is at $(x,y)$, then the area of the rectangle is

$$\text{Area } = f(x,y) = 4xy,$$
$$\text{with constraint } \phi(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$
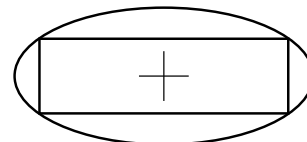
The equations to solve are now

$$\nabla(f - \lambda\phi) = 0, \qquad \text{and} \qquad \phi = 0, \qquad \text{which become}$$

$$4y - \lambda\frac{2x}{a^2} = 0, \qquad 4x - \lambda\frac{2y}{b^2} = 0, \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \tag{8.35}$$

The solutions to these three equations are straight-forward. They are $x = a/\sqrt{2}$, $y = b/\sqrt{2}$, $\lambda = 2ab$. The maximum area is then $4xy = 2ab$. The Lagrange multiplier turns out to be the required area. Does this reduce to the correct result for a circle?

    The second example said that you have several different allowed energies, typical of what happens in quantum mechanics. If the total number of particles and the total energy are given, how are the particles distributed among the different energies?

    If there are $N$ particles and exactly two energy levels, $E_1$ and $E_2$,

$$N = n_1 + n_2, \qquad \text{and} \qquad E = n_1E_1 + n_2E_2$$

you have two equations in two unknowns and all you have to do is solve them for the numbers $n_1$ and $n_2$, the number of particles in each state. If there are three or more possible energies the answer isn't uniquely determined by just two equations, and there can be many ways that you can put particles into different energy states and still have the same number of particles and the same total energy.

    If you're dealing with four particles and three energies, you can perhaps count the possibilities by hand. How many ways can you put four particles in three states? (400), (310), (301), (220), (211), *etc.* There's only one way to get the (400) configuration: All four particles go into state 1. For (310) there are four ways to do it; any one of the four particles can be in the second state and the rest in the first. Keep going. If you have $10^{20}$ particles you have to find a better way.

    If you have a total of $N$ particles and you place $n_1$ of them in the first state, the number of ways that you can do that is $N$ for the first particle, $(N - 1)$ for the second particle, *etc.* $= N(N-1)(N-2)\cdots(N - n_1 + 1) = N!/(N - n_1)!$. This is over-counting because you don't care which one went into the first state first, just that it's there. There are $n_1!$ rearrangements of these $n_1$ particles, so you have to divide by that to get the number of ways that you can get this number of particles into state 1: $N!/n_1!(N - n_1)!$ For example, $N = 4$, $n_1 = 4$ as in the (400) configuration in the preceding paragraph is $4!/0!4! = 1$, or $4!/3!1! = 4$ as in the (310) configuration.

    Once you've got $n_1$ particles into the first state you want to put $n_2$ into the second state (out of the remaining $N - n_1$). Then on to state 3.

    The total number of ways that you can do this is the product of all of these numbers. For three allowed energies it is

$$\frac{N!}{n_1!(N - n_1)!} \cdot \frac{(N - n_1)!}{n_2!(N - n_1 - n_2)!} \cdot \frac{(N - n_1 - n_2)!}{n_3!(N - n_1 - n_2 - n_3)!} = \frac{N!}{n_1!n_2!n_3!} \tag{8.36}$$

There's a lot of cancellation and the final factor in the denominator is one because of the constraint $n_1 + n_2 + n_3 = N$.

    Lacking any other information about the particles, the most probable configuration is the one for which Eq. (8.36) is a maximum. This calls for Lagrange multipliers because you want to maximize a complicated function of several variables subject to constraints on $N$ and on $E$. Now all you have to do is to figure out out to differentiate with respect to integers. Answer: If $N$ is large you will be able to treat these variables as continuous and to use standard calculus to manipulate them.

    For large $n$, recall Stirling's formula, Eq. (2.20),

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n} \qquad \text{or its log:} \qquad \ln(n!) \sim \ln\sqrt{2\pi n} + n\ln n - n \tag{8.37}$$

This, I can differentiate. Maximizing (8.36) is the same as maximizing its logarithm, and that's easier to work with.

$$\text{maximize } f = \ln(N!) - \ln(n_1!) - \ln(n_2!) - \ln(n_3!)$$
$$\text{subject to } n_1 + n_2 + n_3 = N \quad \text{and} \quad n_1 E_1 + n_2 E_2 + n_3 E_3 = E$$

There are two constraints here, so there are two Lagrange multipliers.

$$\nabla\big(f - \lambda_1(n_1 + n_2 + n_3 - N) - \lambda_2(n_1 E_1 + n_2 E_2 + n_3 E_3 - E)\big) = 0$$

For $f$, use Stirling's approximation, but not quite. The term $\ln\sqrt{2\pi n}$ is negligible. For $n$ as small as $10^6$, it is about $6 \times 10^{-7}$ of the whole. Logarithms are much smaller than powers. That means that I can use

$$\nabla\left(\sum_{\ell=1}^{3}\big(-n_\ell \ln(n_\ell) + n_\ell\big) - \lambda_1 n_\ell - \lambda_2 n_\ell E_\ell\right) = 0$$

This is easier than it looks because each derivative involves only one coordinate.

$$\frac{\partial}{\partial n_1} \to -\ln n_1 - 1 + 1 - \lambda_1 - \lambda_2 E_1 = 0, \ \textit{etc.}$$

This is

$$n_\ell = e^{-\lambda_1 - \lambda_2 E_\ell}, \quad \ell = 1,\, 2,\, 3$$

There are two unknowns here, $\lambda_1$ and $\lambda_2$. There are two equations, for $N$ and $E$, and the parameter $\lambda_1$ simply determines an overall constant, $e^{-\lambda_1} = C$.

$$C\sum_{\ell=1}^{3} e^{-\lambda_2 E_\ell} = N, \quad \text{and} \quad C\sum_{\ell=1}^{3} E_\ell e^{-\lambda_2 E_\ell} = E$$

The quantity $\lambda_2$ is usually denoted $\beta$ in this type of problem, and it is related to temperature by $\beta = 1/kT$ where as usual the Lagrange multiplier is important on its own. It is usual to manipulate these results by defining the "partition function"

$$Z(\beta) = \sum_{\ell=1}^{3} e^{-\beta E_\ell} \tag{8.38}$$

In terms of this function $Z$ you have

$$C = N/Z, \quad \text{and} \quad E = -\frac{N}{Z}\frac{dZ}{d\beta} \tag{8.39}$$

For a lot more on this subject, you can refer to any one of many books on thermodynamics or statistical physics. There for example you can find the reason that $\beta$ is related to the temperature and how the partition function can form the basis for computing everything there is to compute in thermodynamics. Especially there you will find that more powerful versions of the same ideas will arise when you allow the total energy and the total number of particles to be variables too.

**8.13 Solid Angle**
The extension of the concept of angle to three dimensions is called "solid angle." To explain what this is, I'll first show a definition of ordinary angle that's different from what you're accustomed to. When you see that, the extension to one more dimension is easy.

Place an object in the plane somewhere not at the origin. You are at the origin and look at it. I want a definition that describes what fraction of the region around you is spanned by this object. For this, draw a circle of radius $R$ centered at the origin and draw all the lines from everywhere on the object to the origin. These lines will intersect the circle on an arc (or even a set of arcs) of length $s$. *Define* the angle subtended by the object to be $\theta = s/R$.
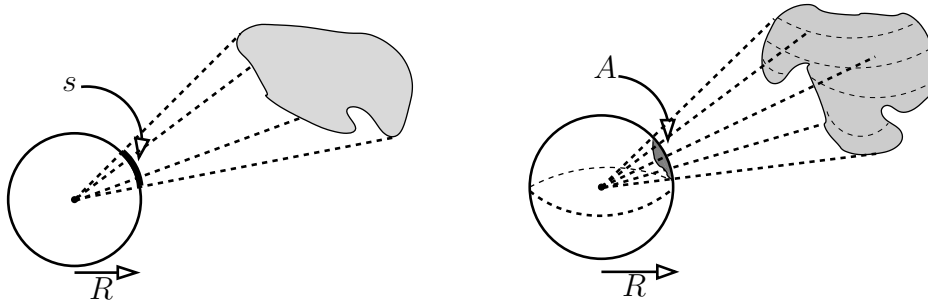


Now step up to three dimensions and again place yourself at the origin. This time place a sphere of radius $R$ around the origin and draw all the lines from the three dimensional object to the origin. This time the lines intersect the sphere on an area of size $A$. *Define* the solid angle subtended by the object to be $\Omega = A/R^2$. (If you want four or more dimensions, see problem 8.52.)

For the circle, the circumference is $2\pi R$, so if you're surrounded, the angle subtended is $2\pi R/R = 2\pi$ radians. For the sphere, the area is $4\pi R^2$, so this time if you're surrounded, the solid angle subtended is $4\pi R^2/R^2 = 4\pi$ sterradians. That is the name for this unit.

All very pretty. Is it useful? Only if you want to describe radiative transfer, nuclear scattering, illumination, the structure of the atom, or rainbows. Except for illumination, these subjects center around one idea, that of a "cross section."

**Cross Section, Absorption**
Before showing how to use solid angle to describe scattering, I'll take a simpler example: absorption. There is a hole in a wall and I propose to measure its area. Instead of taking a ruler to it I blindly fire bullets at the wall and see how many go in. The bigger the area, the larger the fraction that will go into the hole of course, but I have to make this quantitative to make it useful.

Define the flux of bullets: $f = dN/(dt\, dA)$. That is, suppose that I'm firing all the bullets in the same direction, but not starting from the same place. Pick an area $\Delta A$ perpendicular to the stream of bullets and pick a time interval $\Delta t$. How many bullets pass through this area in this time? $\Delta N$, and that's proportional to both $\Delta A$ and $\Delta t$. The limit of this quotient is the flux.
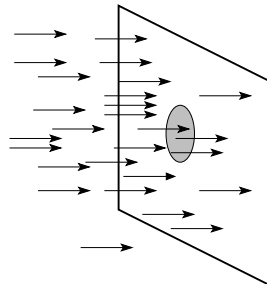


$$\lim_{\substack{\Delta t \to 0 \\ \Delta A \to 0}} \frac{\Delta N}{\Delta t \Delta A} = f \qquad (8.40)$$

Having defined the flux as a kind of density, call the (unknown) area of the hole $\sigma$. The rate at which these bullets enter the hole is proportional to the size of the hole and to the flux of bullets, $R = f\sigma$,

where $R$ is the rate of entry and $\sigma$ is the area of the hole. If I can measure the rate of absorption $R$ and the flux $f$, I have measured the area of the hole, $\sigma = R/f$. This letter is commonly used for cross sections.

Why go to this complicated trouble for a hole? I probably shouldn't, but to measure absorption of neutrons hitting nuclei this is precisely what you do. I can't use a ruler on a nucleus, but I can throw things at it. In this example, neutron absorption by nuclei, the value of the measured absorption cross section can vary from millibarns to kilobarns, where a barn is $10^{-24}\,\text{cm}^2$. The radii of nuclei vary by a factor of only about six from hydrogen through uranium ($\sqrt[3]{238} = 6.2$), so the cross section measured by bombarding the nucleus has little to do with the geometric area $\pi r^2$. It is instead a measure of interaction strength

## Cross Section, Scattering

There are many types of cross sections besides absorption, and the next simplest is the scattering cross section, especially the differential scattering cross section.



$$d\sigma = 2\pi b\, db$$
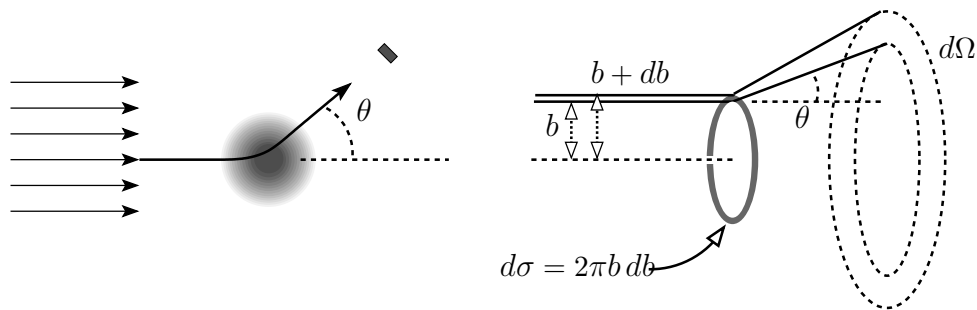
The same flux of particles that you throw at an object may not be absorbed, but may scatter instead. You detect the scattering by using a detector. (You were expecting a catcher's mitt?) The detector will have an area $\Delta A$ facing the particles and be at a distance $r$ from the center of scattering. The detection rate will be proportional to the area of the detector, but if I double $r$ for the same $\Delta A$, the detection rate will go down by a factor of four. The detection rate is proportional to $\Delta A/r^2$, but this is just the solid angle of the detector from the center:

$$\Delta \Omega = \Delta A/r^2 \tag{8.41}$$

The detection rate is proportional to the incoming flux and to the solid angle of the detector. The proportionality is an effective scattering area, $\Delta\sigma$.

$$\Delta R = f\Delta\sigma, \qquad \text{so} \qquad \frac{d\sigma}{d\Omega} = \frac{dR}{f\,d\Omega} \tag{8.42}$$

This is the differential scattering cross section.

You can compute this if you know something about the interactions involved. The one thing that you need is the relationship between where the particle comes in and the direction in which it leaves. That is, the incoming particle is aimed to hit at a distance $b$ (called the impact parameter) from the center and it scatters at an angle $\theta$, called of course the scattering angle, from its original direction. Particles that come in at distance between $b$ and $b + db$ from the axis through the center will scatter into directions between $\theta$ and $\theta + d\theta$.

The cross section for being sent in a direction between these two angles is the area of the ring: $d\sigma = 2\pi b\, db$. Anything that hits in there will scatter into the outgoing angles shown. How much solid angle is this? Put the $z$-axis of spherical coordinates to the right, so that $\theta$ is the usual spherical coordinate angle from $z$. The element of area on the surface of a sphere is $dA = r^2 \sin\theta d\theta d\phi$, so the

integral over all the azimuthal angles $\phi$ around the ring just gives a factor $2\pi$. The element of solid angle is then

$$d\Omega = \frac{dA}{r^2} = 2\pi \sin\theta \, d\theta$$

As a check on this, do the integral over all theta to get the total solid angle around a point, verifying that it is $4\pi$.

Divide the effective area for this scattering by the solid angle, and the result is the differential scattering cross section.

$$\frac{d\sigma}{d\Omega} = \frac{2\pi b \, db}{2\pi \sin\theta \, d\theta} = \frac{b}{\sin\theta}\frac{db}{d\theta}$$

If you have $\theta$ as a function of $b$, you can compute this. There are a couple of very minor modifications that you need in order to complete this development. The first is that the derivative $db/d\theta$ can easily be negative, but both the area and the solid angle are positive. That means that you need an absolute value here. One other complication is that one value of $\theta$ can come from several values of $b$. It may sound unlikely, but it happens routinely. It even happens in the example that comes up in the next section.

$$\frac{d\sigma}{d\Omega} = \sum_i \frac{b_i}{\sin\theta}\left|\frac{db_i}{d\theta}\right| \tag{8.43}$$

The differential cross section often becomes much more involved than this, especially the when it involves nuclei breaking up in a collision, resulting in a range of possible energies of each part of the debris. In such collisions particles can even be created, and the probabilities and energy ranges of the results are described by their own differential cross sections. You will wind up with differential cross sections that look like $d\sigma/d\Omega_1 \, d\Omega_2 \dots dE_1 \, dE_2 \dots$. These rapidly become so complex that it takes some elaborate computer programming to handle the information.

### 8.14 Rainbow

An interesting, if slightly complicated example is the rainbow. Sunlight scatters from small drops of water in the air and the detector is your eye. The water drops are small enough that I'll assume them to be spheres, where surface tension is enough to hold them in this shape for the ordinary small sizes of water droplets in the air. The first and simplest model uses geometric optics and Snell's law to figure out where the scattered light goes. This model ignores the wave nature of light and it does not take into account the fraction of the light that is transmitted and reflected at each surface.



$$\sin\beta = n\sin\alpha$$
$$\theta = (\beta - \alpha) + (\pi - 2\alpha) + (\beta - \alpha)$$
$$b = R\sin\beta$$
$$\tag{8.44}$$
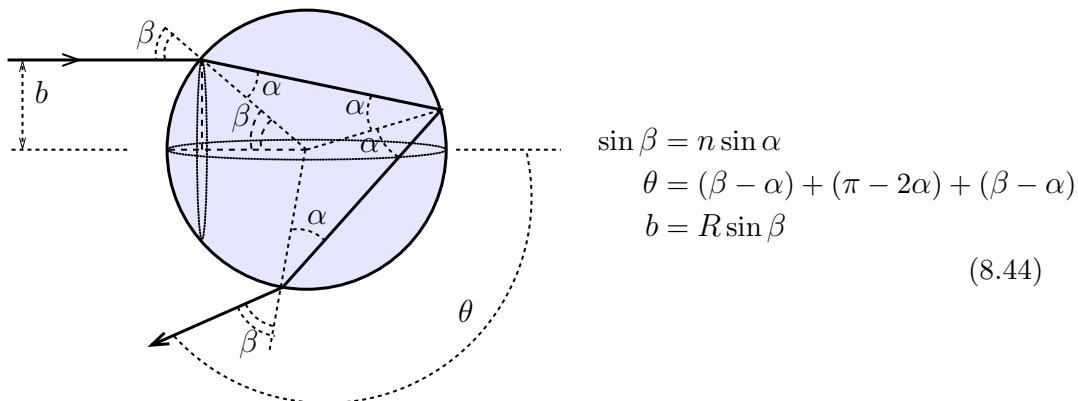
The light comes in at the indicated distance $b$ from the axis through the center of the sphere. It is then refracted, reflected, and refracted. Snell's law describes the first and third of these, and the

middle one has equal angles of incidence and reflection. The dashed lines are from the center of the sphere. The three terms in Eq. (8.44) for the evaluation of $\theta$ come from the three places at which the light changes direction, and they are the amount of deflection at each place. The third equation simply relates $b$ to the radius of the sphere.

From these three equations, eliminate the two variables $\alpha$ and $\beta$ to get the single relation between $b$ and $\theta$ that I'm looking for. When you do this, you find that the resulting equations are a bit awkward. It's sometimes easier to use one of the two intermediate angles as a parameter, and in this case you will want to use $\beta$. From the picture you know that it varies from zero to $\pi/2$. The third equation gives $b$ in terms of $\beta$. The first equation gives $\alpha$ in terms of $\beta$. The second equation determines $\theta$ in terms of $\beta$ and the $\alpha$ that you've just found.

The parametrized relation between $b$ and $\theta$ is then

$$b = R\sin\beta, \qquad \theta = \pi + 2\beta - 4\sin^{-1}\left(\frac{1}{n}\sin\beta\right), \qquad (0 < \beta < \pi/2) \qquad (8.45)$$

or you can carry it through and eliminate $\beta$.

$$\theta = \pi + 2\sin^{-1}\left(\frac{b}{R}\right) - 4\sin^{-1}\left(\frac{1}{n}\frac{b}{R}\right) \qquad (8.46)$$
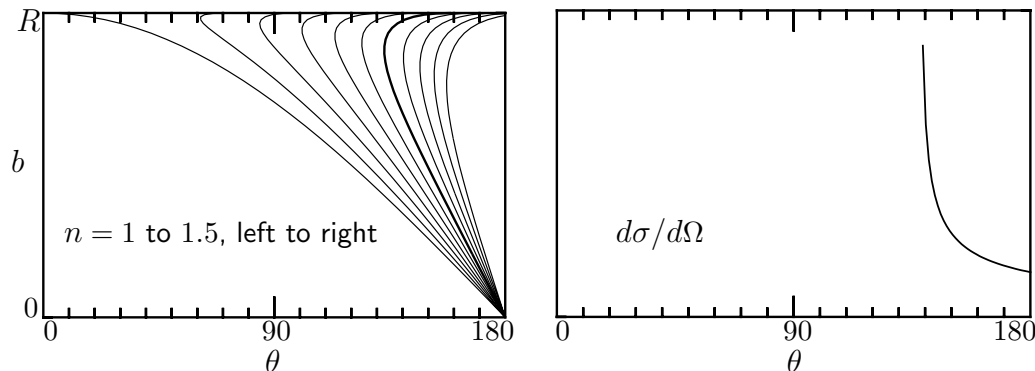
The derivative $db/d\theta = 1/[d\theta/db]$. Compute this.

$$\frac{d\theta}{db} = \frac{2}{\sqrt{R^2 - b^2}} - \frac{4}{\sqrt{n^2 R^2 - b^2}} \qquad (8.47)$$

In the parametrized form this is

$$\frac{db}{d\theta} = \frac{db/d\beta}{d\theta/d\beta} = \frac{R\cos\beta}{2 - 4\cos\beta/\sqrt{n^2 - \sin^2\beta}}$$

In analyzing this, it's convenient to have both forms, as you never know which one will be easier to interpret. (Have you checked to see if they agree with each other in any special cases?)



These graphs are generated from Eq. (8.45) for eleven values of the index of refraction equally spaced from 1 to 1.5, and the darker curve corresponds to $n = 1.3$. The key factor that enters the cross-section calculation, Eq. (8.43), is $db/d\theta$, because it goes to infinity when the curve has a vertical tangent. For water, with $n = 1.33$, the $b$-$\theta$ curve has a vertical slope that occurs for $\theta$ a little less than $140°$. *That* is the rainbow.

To complete this I should finish with $d\sigma/d\Omega$. The interesting part of the problem is near the vertical part of the curve. To see what happens near such a point use a power series expansion near there. Not $b(\theta)$ but $\theta(b)$. This has zero derivative here, so near the vertical point

$$\theta(b) = \theta_0 + \gamma(b - b_0)^2$$

At $(b_0, \theta_0)$, Eq. (8.47) gives zero and Eq. (8.46) tells you $\theta_0$. The coefficient $\gamma$ comes from the second derivative of Eq. (8.46) at $b_0$. What is the differential scattering cross section in this neighborhood?

$$b = b_0 \pm \sqrt{(\theta - \theta_0)/\gamma}, \qquad \text{so} \qquad db/d\theta = \pm \frac{1}{2\sqrt{\gamma(\theta - \theta_0)}}$$

$$
\begin{aligned}
\frac{d\sigma}{d\Omega} &= \sum_i \frac{b_i}{\sin\theta} \left| \frac{db_i}{d\theta} \right| \\
&= \frac{b_0 + \sqrt{(\theta - \theta_0)/\gamma}}{\sin\theta} \frac{1}{2\sqrt{\gamma(\theta - \theta_0)}} + \frac{b_0 - \sqrt{(\theta - \theta_0)/\gamma}}{\sin\theta} \frac{1}{2\sqrt{\gamma(\theta - \theta_0)}} \\
&= \frac{b_0}{\sin\theta\sqrt{\gamma(\theta - \theta_0)}} \approx \frac{b_0}{\sin\theta_0\sqrt{\gamma(\theta - \theta_0)}}
\end{aligned}
\qquad (8.48)
$$

In the final expression, because this is near $\theta - \theta_0$ and because I'm doing a power series expansion of the exact solution anyway, I dropped all the $\theta$-dependence except the dominant factors. This is the only consistent thing to do because I've previously dropped higher order terms in the expansion of $\theta(b)$.

Why is this a rainbow? (1) With the sun at your back you see a bright arc of a circle in the direction for which the scattering cross-section is very large. The angular radius of this circle is $\pi - \theta_0 \approx 42°$. (2) The value of $\theta_0$ depends on the index of refraction, $n$, and that varies slightly with wavelength. The variation of this angle of peak intensity is

$$\frac{d\theta_0}{d\lambda} = \frac{d\theta_0}{db_0} \frac{db_0}{dn} \frac{dn}{d\lambda} \qquad (8.49)$$

When you graph Eq. (8.48) note carefully that it is zero on the left of $\theta_0$ (smaller $\theta$) and large on the right. Large scattering angles correspond to the region of the sky underneath the rainbow, toward the center of the circular arc. This implies that there is much more light scattered toward your eye underneath the arc of the rainbow than there is above it. Look at your next rainbow and compare the area of sky below and above the rainbow.

There's a final point about this calculation. I didn't take into account the fact that when light hits a surface, some is transmitted and some is reflected. The largest effect is at the point of internal reflection, because typically only about two percent of the light is reflected and the rest goes through. The cross section should be multiplied by this factor to be complete. The detailed equations for this are called the Fresnel formulas and they tell you the fraction of the light transmitted and reflected at a surface as a function of angle and polarization.

This is far from the whole story about rainbows. Light is a wave, and the geometric optics approximation that I've used doesn't account for everything. In fact Eq. (8.43) doesn't apply to waves, so the whole development has to be redone. To get an idea of some of the other phenomena associated with the rainbow, see for example
www.usna.edu/Users/oceano/raylee/RainbowBridge/Chapter_8.html
www.philiplaven.com/links.html

**Exercises**

**1** For the functions $f(x, y) = Axy^2 \sin(xy)$, $x(t) = Ct^3$, $y(t) = Dt^2$, compute $df/dt$ two ways. First use the chain rule, then do explicit substitution and compute it directly.

**2** Compute $(\partial f/\partial x)_y$ and $(\partial f/\partial y)_x$ for

$$(a)\ f(x, y) = x^2 - 2xy + y^2, \qquad (b)\ f(x, y) = \ln(y/x), \qquad (c)\ f(x, y) = (y + x)/(y - x)$$

**3** Compute $df/dx$ using the chain rule for

$$(a)\ f(x, y) = \ln(y/x), \quad y = x^2, \qquad\qquad (b)\ f(x, y) = (y + x)/(y - x), \quad y = \alpha x,$$
$$(c)\ f(x, y) = \sin(xy), \quad y = 1/x$$

Also calculate the results by substituting $y$ explicitly and then differentiating, comparing the results.

**4** Let $f(x, y) = x^2 - 2xy$, and the polar coordinates are $x = r\cos\phi$, $y = r\sin\phi$. Compute

$$\left(\frac{\partial f}{\partial x}\right)_y, \quad \left(\frac{\partial f}{\partial y}\right)_x, \quad \left(\frac{\partial f}{\partial x}\right)_r, \quad \left(\frac{\partial f}{\partial y}\right)_r, \quad \left(\frac{\partial f}{\partial x}\right)_\phi, \quad \left(\frac{\partial f}{\partial y}\right)_\phi$$

**5** Let $f(x, y) = x^2 - 2xy$, and the polar coordinates are $x = r\cos\phi$, $y = r\sin\phi$. Compute

$$\left(\frac{\partial f}{\partial r}\right)_\phi, \quad \left(\frac{\partial f}{\partial \phi}\right)_r, \quad \left(\frac{\partial f}{\partial r}\right)_x, \quad \left(\frac{\partial f}{\partial \phi}\right)_x, \quad \left(\frac{\partial f}{\partial r}\right)_y, \quad \left(\frac{\partial f}{\partial \phi}\right)_y$$

**6** For the function $f(u, v) = u^3 - v^3$, what is the value at $(u, v) = (2, 1)$? Approximately what is its value at $(u, v) = (2.01, 1.01)$? Approximately what is its value at $(u, v) = (2.01, 0.99)$?

**7** Assume the Earth's atmosphere is uniform density and 10 km high, what is its volume? What is the ratio of this volume to the Earth's volume?

**8** For a cube 1 m on a side, what volume of paint will you need in order to paint it to a thickness of 0.2 mm? Don't forget to paint all the sides.

**9** What is $\operatorname{grad} r^2$? Do it in both rectangular and polar coordinates. Two dimensions will do. Are your results *really* the same?

**10** What is $\operatorname{grad}(\alpha x^2 + \beta y^2)$. Do this in both rectangular and polar coordinates. For the polar form, put $x$ and $y$ in terms of $r$ and $\phi$, then refer to Eq. (8.27) for the polar form of the gradient. Finally, compare the two results.

**11** The Moon has a radius about 1740 km and its distance from Earth averages about 384 000 km from Earth. What solid angle does the Moon subtend from Earth? What solid angle does Earth (radius 6400 km) subtend from the Moon?

**12** Express the cylindrical unit vectors $\hat{r}$, $\hat{\phi}$, $\hat{z}$ in terms of the rectangular ones. And vice versa.

**13** Evaluate the volume of a sphere by integration in spherical coordinates.

### Problems

**8.1** Let $r = \sqrt{x^2 + y^2}$, $x = A \sin \omega t$, $y = B \cos \omega t$. Use the chain rule to compute the derivative with respect to $t$ of $e^{kr}$. Notice the various checks you can do on the result, verifying (or disproving) your result.

**8.2** Sketch these functions* in plane polar coordinates:
(a) $r = a \cos \phi$   (b) $r = a \sec \phi$   (c) $r = a\phi$   (d) $r = a/\phi$   (e) $r^2 = a^2 \sin 2\phi$

**8.3** The two coordinates $x$ and $y$ are related by $f(x, y) = 0$. What is the derivative of $y$ with respect to $x$ under these conditions? [What is $df$ *along* this curve? And have you drawn a sketch?] Make up a test function (with enough structure to be a test but still simple enough to verify your answer independently) and see if your answer is correct. Ans: $-(\partial f/\partial x)/(\partial f/\partial y)$

**8.4** If $x = u + v$ and $y = u - v$, show that

$$\left( \frac{\partial y}{\partial x} \right)_u = -\left( \frac{\partial y}{\partial x} \right)_v$$

Do this by application of the chain rule, Eq. (8.6). Then as a check do the calculation by explicit elimination of the respective variables $v$ and $u$.

**8.5** If $x = r \cos \phi$ and $y = r \sin \phi$, compute

$$\left( \frac{\partial x}{\partial r} \right)_\phi \qquad \text{and} \qquad \left( \frac{\partial x}{\partial r} \right)_y$$

**8.6** What is the differential of $f(x, y, z) = \ln(xyz)$.

**8.7** If $f(x, y) = x^3 + y^3$ and you switch to plane polar coordinates, use the chain rule to evaluate

$$\left( \frac{\partial f}{\partial r} \right)_\phi, \quad \left( \frac{\partial f}{\partial \phi} \right)_r, \quad \left( \frac{\partial^2 f}{\partial r^2} \right)_\phi, \quad \left( \frac{\partial^2 f}{\partial \phi^2} \right)_r, \quad \left( \frac{\partial^2 f}{\partial r \partial \phi} \right)$$

Check one or more of these by substituting $r$ and $\phi$ explicitly and doing the derivatives.

**8.8** When current $I$ flows through a resistance $R$ the heat produced is $I^2 R$. Two terminals are connected in parallel by two resistors having resistance $R_1$ and $R_2$. Given that the total current is divided as $I = I_1 + I_2$, show that the condition that the total heat generated is a minimum leads to the relation $I_1 R_1 = I_2 R_2$. You don't need Lagrange multipliers to solve this problem, but try them anyway.

**8.9** Sketch the magnetic field represented by Eq. (8.24). I suggest that you start by fixing $r$ and drawing the $\vec{B}$-vectors at various values of $\theta$. It will probably help your sketch if you first compute the magnitude of $B$ to see how it varies around the circle. Recall, this field is expressed in spherical coordinates, though you can take advantage of its symmetry about the $z$-axis to make the drawing

---

   * See www-groups.dcs.st-and.ac.uk/~history/Curves/Curves.html for more.

simpler. Don't stop with just the field at fixed $r$ as I suggested you begin. The field fills space, so try to describe it.

**8.10** A drumhead can vibrate in more complex modes. One such mode that vibrates at a frequency higher than that of Eq. (8.19) looks approximately like

$$z(r, \phi, t) = A\, r\left(1 - r^2/R^2\right) \sin\phi \cos\omega_2 t$$

(**a**) Find the total kinetic energy of this oscillating drumhead.
(**b**) Sketch the shape of the drumhead at $t = 0$. Compare it to the shape of Eq. (8.19).
At the instant that the total kinetic energy is a maximum, what is the shape of the drumhead?
Ans: $\frac{\pi}{48}\sigma A^2 \omega_2^2 R^4 \sin^2 \omega_2 t$

**8.11** Just at there is kinetic energy in a vibrating drumhead, there is potential energy, and as the drumhead moves its total potential energy will change because of the slight stretching of the material. The potential energy density $(d\,\text{P.E.}/dA)$ in a drumhead is

$$u_\mathsf{p} = \frac{1}{2}T\left(\nabla z\right)^2$$

$T$ is the tension in the drumhead. It has units of Newtons/meter and it is the force per length you would need if you cut a small slit in the surface and had to hold the two sides of the slit together. This potential energy arises from the slight stretching of the drumhead as it moves away from the plane of equilibrium.
(**a**) For the motion described by Eq. (8.19) compute the total potential energy. (Naturally, you will have checked the dimensions first to see if the claimed expression for $u_\mathsf{p}$ is sensible.)
(**b**) Energy is conserved, so the sum of the total potential energy and the total kinetic energy from Eq. (8.20) must be a constant. What must the frequency $\omega$ be for this to hold? Is this a plausible result? A more accurate result, from solving a differential equation, is $2.405\sqrt{T/\sigma R^2}$.   Ans: $\sqrt{6T/\sigma R^2} = 2.45\sqrt{T/\sigma R^2}$

**8.12** Repeat the preceding problem for the drumhead mode of problem 8.10. The exact result, calculated in terms of roots of Bessel functions is $3.832\sqrt{T/\sigma R^2}$.   Ans: $4\sqrt{T/\sigma R^2}$

**8.13** Sketch the gravitational field of the Earth from Eq. (8.23). Is the direction of the field plausible? Draw lots of arrows.

**8.14** Prove that the unit vectors in polar coordinates are related to those in rectangular coordinates by

$$\hat{r} = \hat{x}\cos\phi + \hat{y}\sin\phi, \qquad \hat{\phi} = -\hat{x}\sin\phi + \hat{y}\cos\phi$$

What are $\hat{x}$ and $\hat{y}$ in terms of $\hat{r}$ and $\hat{\phi}$?

**8.15** Prove that the unit vectors in spherical coordinates are related to those in rectangular coordinates by

$$\hat{r} = \hat{x}\sin\theta\cos\phi + \hat{y}\sin\theta\sin\phi + \hat{z}\cos\theta$$
$$\hat{\theta} = \hat{x}\cos\theta\cos\phi + \hat{y}\cos\theta\sin\phi - \hat{z}\sin\theta$$
$$\hat{\phi} = -\hat{x}\sin\phi + \hat{y}\cos\phi$$

**8.16** Compute the volume of a sphere using spherical coordinates. Also do it using rectangular coordinates. Also do it in cylindrical coordinates.

**8.17** Finish both integrals Eq. (8.21). Draw sketches to demonstrate that the limits stated there are correct.

**8.18** Find the volume under the plane $2x + 2y + z = 8a$ and over the triangle bounded by the lines $x = 0$, $y = 2a$, and $x = y$ in the $x$-$y$ plane. Ans: $8a^3$

**8.19** Find the volume enclosed by the doughnut-shaped surface (spherical coordinates) $r = a \sin \theta$. Ans: $\pi^2 a^3 / 4$

**8.20** In plane polar coordinates, compute $\partial \hat{r} / \partial \phi$, also $\partial \hat{\phi} / \partial \phi$. This means that $r$ is fixed and you're finding the change in these vectors as you move around a circle. In both cases express the answer in terms of the $\hat{r}$-$\hat{\phi}$ vectors. Draw pictures that will demonstrate that your answers are at least in the right direction. Ans: $\partial \hat{\phi} / \partial \phi = -\hat{r}$

**8.21** Compute the gradient of the distance from the origin (in three dimensions) in three coordinate systems and verify that they agree.

**8.22** Taylor's power series expansion of a function of several variables was discussed in section 2.5. The Taylor series in one variable was expressed in terms of an exponential in problem 2.30. Show that the series in three variables can be written as

$$e^{\vec{h} \cdot \nabla} f(x, y, z)$$

**8.23** The wave equation is (a) below. Change variables to $z = x - vt$ and $w = x + vt$ and show that in these coordinates this equation is (b) (except for a constant factor). Did you *explicitly* note which variables are kept fixed at each stage of the calculation? See also problem 8.53.

$$(a) \quad \frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0 \qquad\qquad (b) \quad \frac{\partial^2 u}{\partial z \partial w} = 0$$

**8.24** The equation (8.23) comes from taking the gradient of the Earth's gravitational potential in an expansion to terms in $1/r^3$.

$$V = -\frac{GM}{r} - \frac{GQ}{r^3} P_2(\cos \theta)$$

where $P_2(\cos \theta) = \frac{3}{2} \cos^2 \theta - \frac{1}{2}$ is the second order Legendre polynomial. Compute $\vec{g} = -\nabla V$.

**8.25** In problem 2.25 you computed the electric potential at large distances from a pair of charges, $-q$ at the origin and $+q$ at $z = a$ ($r \gg a$). The result was

$$V = \frac{kqa}{r^2} P_1(\cos \theta)$$

where $P_1(\cos \theta) = \cos \theta$ is the first order Legendre polynomial. Compute the electric field from this potential, $\vec{E} = -\nabla V$. And sketch it of course.

**8.26** In problem 2.26 you computed the electric potential at large distances from a set of three charges, $-2q$ at the origin and $+q$ at $z = \pm a$ ($r \gg a$). The result was

$$V = \frac{kqa^2}{r^3} P_2(\cos\theta)$$

where $P_2(\cos\theta)$ is the second order Legendre polynomial. Compute the electric field from this potential, $\vec{E} = -\nabla V$. And sketch it of course.

**8.27** Compute the area of an ellipse having semi-major and semi-minor axes $a$ and $b$. Compare your result to that of Eq. (8.35). Ans: $\pi ab$

**8.28** Two equal point charges $q$ are placed at $z = \pm a$. The origin is a point of equilibrium; $\vec{E} = 0$ there. **(a)** Compute the potential near the origin, writing $V$ in terms of powers of $x$, $y$, and $z$ near there, carrying the powers high enough to describe the nature of the equilibrium point. Is $V$ maximum, minimum, or saddle point there? It will be easier if you carry the calculation as far as possible using vector notation, such as $|\vec{r} - a\hat{z}| = \sqrt{(\vec{r} - a\hat{z})^2}$, and $r \ll a$.
**(b)** Write your result for $V$ near the origin in spherical coordinates also.
Ans: $\frac{2q}{4\pi\epsilon_0 a}\left[1 + \frac{r^2}{a^2}\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)\right]$

**8.29** When current $I$ flows through a resistance $R$ the heat produced is $I^2 R$. Two terminals are connected in parallel by three resistors having resistance $R_1$, $R_2$, and $R_3$. Given that the total current is divided as $I = I_1 + I_2 + I_3$, show that the condition that the total heat generated is a minimum leads to the relation $I_1 R_1 = I_2 R_2 = I_3 R_3$. You can easily do problem 8.8 by eliminating a coordinate then doing a derivative. Here it's starting to get sufficiently complex that you should use Lagrange multipliers. Does $\lambda$ have any significance this time?

**8.30** Given a right circular cylinder of volume $V$, what radius and height will provide the minimum total area for the cylinder. Ans: $r = (V/2\pi)^{1/3}$, $h = 2r$

**8.31** Sometimes the derivative isn't zero at a maximum or a minimum. Also, there are two types of maxima and minima; local and global. The former is one that is max or min in the immediate neighborhood of a point and the latter is biggest or smallest over the entire domain of the function. Examine these functions for maxima and minima both inside the domains and on the boundary.

$$|x|, \ (-1 \leq x \leq +2)$$
$$T_0(x^2 - y^2)/a^2, \ (-a \leq x \leq a, \ -a \leq y \leq a)$$
$$V_0(r^2/R^2)P_2(\cos\theta), \ (r \leq R, \ 3 \text{ dimensions})$$

**8.32** In Eq. (8.39) it is more common to specify $N$ and $\beta = 1/kT$, the Lagrange multiplier, than it is to specify $N$ and $E$, the total energy. Pick three energies, $E_\ell$, to be 1, 2, and 3 electron volts.
**(a)** What is the average energy, $E/N$, as $\beta \to \infty$ ($T \to 0$)?
**(b)** What is the average energy as $\beta \to 0$?
**(c)** What are $n_1$, $n_2$, and $n_3$ in these two cases?

**8.33 (a)** Find the gradient of $V$, where $V = V_0(x^2 + y^2 + z^2)a^{-2}e^{-\sqrt{x^2+y^2+z^2}/a}$. **(b)** Find the gradient of $V$, where $V = V_0(x + y + z)a^{-1}e^{-(x+y+z)/a}$.

**8.34** A billiard ball of radius $R$ is suspended in space and is held rigidly in position. Very small pellets are thrown at it and the scattering from the surface is completely elastic, with no friction. Compute

the relation between the impact parameter $b$ and the scattering angle $\theta$. Then compute the differential scattering cross section $d\sigma/d\Omega$.
Finally compute the total scattering cross section, the integral of this over $d\Omega$.

**8.35** Modify the preceding problem so that the incoming object is a ball of radius $R_1$ and the fixed billiard ball has radius $R_2$.

**8.36** Find the differential scattering cross section from a spherical drop of water, but instead of Snell's law, use a pre-Snell law: $\beta = n\alpha$, without the sines. Is there a rainbow in this case? Sketch $d\sigma/d\Omega$ versus $\theta$.
Ans: $R^2 \sin 2\beta/\left[4\sin\theta|1 - 2/n|\right]$, where $\theta = \pi + 2(1 - 2/n)\beta$

**8.37** From the equation (8.43), assuming just a single $b$ for a given $\theta$, what is the integral over all $d\Omega$ of $d\sigma/d\Omega$?  Ans: $\pi b_{\text{max}}^2$

**8.38** Solve Eq. (8.47) for $b$ when $d\theta/db = 0$. For $n = 1.33$ what value of $\theta$ does this give?

**8.39** If the scattering angle $\theta = \frac{\pi}{2}\sin(\pi b/R)$ for $0 < b < R$, what is the resulting differential scattering cross section (with graph). What is the total scattering cross section? Start by sketching a graph of $\theta$ versus $b$.  Ans: $2R^2/\left[\pi^2 \sin\theta\sqrt{1 - (2\theta/\pi)^2}\right]$

**8.40** Find the signs of all the factors in Eq. (8.49), and determine from that whether red or blue is on the outside of the rainbow.  Ans: Look

**8.41** If it suddenly starts to rain small, spherical diamonds instead of water, what happens to the rainbow? $n = 2.4$ for diamond.

**8.42** What would the rainbow look like for $n = 2$? You'll have to look closely at the expansions in this case. For small $b$, where does the ray hit the inside surface of the drop?

**8.43** (**a**) The secondary rainbow occurs because there can be two internal reflections before the light leave the drop. What is the analog of Eqs. (8.44) for this case? (**b**) Repeat problems 8.38 and 8.40 for this case.

**8.44** What is the shortest distance from the origin to the plane defined by $\vec{A}\cdot(\vec{r} - \vec{r}_0) = 0$? Do this using Lagrange multipliers, and then explain why of course the answer is correct.

**8.45** The U.S. Post Office has decided to use a norm like Eq. (6.11)(2) to measure boxes. The size is defined to be the sum of the height and the circumference of the box, and the circumference is around the thickest part of the package: "length plus girth." What is the maximum volume you can ship if this size is constrained to be less than $130$ inches? For this purpose, assume the box is rectangular, not cylindrical, though you may expect the cylinder to improve the result. Assume that the box's dimensions are $a$, $a$, $b$, with volume $a^2 b$.
(**a**) Show that if you assume that the girth is $4a$, then you will conclude that $b > a$ and that you *didn't* measure the girth at the thickest part of the package.
(**b**) Do it again with the opposite assumption, that you assume $b$ is big so that the girth is $2b + 2a$. Again show that it is a contradiction.
(**c**) You have two inequalities that you must satisfy: girth plus length measured one way is less than $L = 130$ inches and girth plus length measured the other way is too. That is, $4a + b < L$ and $3a + 2b < L$. Plot these regions in the $a$-$b$ plane, showing the allowed region in $a$-$b$ space. Also plot some curves of constant volume, $V = a^2 b$. Show that the point of maximum volume subject to these

constraints is on the edge of this allowed region, and that it is at the corner of intersection of the two inequalities. This is the beginning of the subject called "linear programming."
Ans: a cube

**8.46** Plot $\theta$ versus $b$ in equation (8.45) or (8.46).

**8.47** A disk of radius $R$ is at a distance $c$ above the $x$-$y$ plane and parallel to that plane. What is the solid angle that this disk subtends from the origin? Ans: $2\pi\left[1 - c/\sqrt{c^2 + R^2}\right]$

**8.48** Within a sphere of radius $R$, what is the volume contained between the planes defined by $z = a$ and $z = b$? Ans: $\pi(b - a)\left(R^2 - \frac{1}{3}(b^2 + ab + a^2)\right)$

**8.49** Find the mean-square distance, $\frac{1}{V}\int r^2\, dV$, from a point on the surface of a sphere to points inside the sphere. Note: Plan ahead and try to make this problem as easy as possible. Ans: $8R^2/5$

**8.50** Find the mean distance, $\frac{1}{V}\int r\, dV$, from a point on the surface of a sphere to points inside the sphere. Unlike the preceding problem, this requires some brute force. Ans: $6R/5$

**8.51** A volume mass density is specified in spherical coordinates to be

$$\rho(r, \theta, \phi) = \rho_0\left(1 + r^2/R^2\right)\left[1 + \tfrac{1}{2}\cos\theta\sin^2\phi + \tfrac{1}{4}\cos^2\theta\sin^3\phi\right]$$

Compute the total mass in the volume $0 < r < R$. Ans: $32\pi\rho_0 R^3/15$

**8.52** The circumference of a circle is some constant times its radius $(C_1 r)$. For the two-dimensional surface that is a sphere in three dimensions the area is of the form $C_2 r^2$. Start from the fact that you know the integral $\int_{-\infty}^{\infty} dx\, e^{-x^2} = \pi^{1/2}$ and write out the following two dimensional integral twice. It is over the entire plane.

$$\int dA\, e^{-r^2} \qquad \text{using} \qquad dA = dx\, dy \qquad \text{and using} \qquad dA = C_1 r\, dr$$

From this, evaluate $C_1$. Repeat this for $dV$ and $C_2 r^2$ in three dimensions, evaluating $C_2$. Now repeat this in arbitrary dimensions to evaluate $C_n$. Do you need to reread chapter one? In particular, what is $C_3$? It tells you about the three dimensional hypersphere in four dimensions. From this, what is the total "hypersolid angle" in four dimensions (like $4\pi$ in three)? Ans: $2\pi^2$

**8.53** Do the reverse of problem 8.23. Start with the second equation there and change variables to see that it reverts to a constant times the first equation.

**8.54** Carry out the interchange of limits in Eq. (8.22). Does the drawing really represent the integral?

**8.55** Is $x^2 + xy + y^2$ a minimum or maximum or something else at $(0,0)$? Do the same question for $x^2 + 2xy + y^2$ and for $x^2 + 3xy + y^2$. Sketch the surface $z = f(x, y)$ in each case.

**8.56** Derive the conditions stated after Eq. (8.33), expressing the circumstances under which the Hessian matrix is positive definite.

**8.57** In the spirit of problems 8.10 *et seq.* what happens if you have a rectangular drumhead instead of a circular one? Let $0 < x < a$ and $0 < y < b$. The drumhead is tied down at its edges, so an appropriate function that satisfies these conditions is

$$z(x, y) = A\sin(n\pi x/a)\sin(m\pi y/b)\cos\omega t$$

Compute the total kinetic and the total potential energy for this oscillation, a function of time. For energy to be conserved the total energy must be a constant, so compute the frequency $\omega$ for which this is true. As compared to the previous problems about a circular drumhead, this turns out to give the exact results instead of only approximate ones. Ans: $\omega^2 = \pi^2 \frac{\mu}{T} \left[ \frac{n^2}{a^2} + \frac{m^2}{b^2} \right]$

**8.58** Repeat problem 8.45 by another method. Instead of assuming that the box has a square end, allow it to be any rectangular box, so that its volume is $V = abc$. Now you have three independent variables to use, maximizing the volume subject to the post office's constraint on length plus girth. This looks like it will have to be harder. Instead, it's much easier. Draw pictures! Ans: still a cube

**8.59** An asteroid is headed in the general direction of Earth, and its speed when far away is $v_0$ relative to the Earth. What is the total cross section for it's hitting Earth? It is not necessary to compute the complete orbit; all you have to do is use a couple of conservation laws. Express the result in terms of the escape speed from Earth.
Ans: $\sigma = \pi R^2 \left( 1 + (v_{\text{esc}}/v_0)^2 \right)$
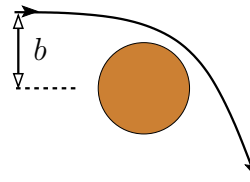
**8.60** In three dimensions the differential scattering cross section appeared in Eqs. (8.42) and (8.43). If the world were two dimensional this area would be a length instead. What are the two corresponding equations in that case, giving you an expression for $d\ell/d\theta$. Apply this to the light scattering from a (two dimensional) drop of water and describe the scattering results. For simplicity this time, assume the pre-Snell law as in problem 8.36.

**8.61** As in the preceding problem, but use the regular Snell law instead.

**8.62** This double integral is over the isosceles right triangle in the figure. The function to be integrated is $f(t') = \alpha t'^3$, BUT FIRST, set it up for an arbitrary $f(t')$ and then set it up again but with the order of integration reversed. In one of the two cases you should be able to do one integral without knowing $f$. Having done this, apply your two results to this particular $f$ as a test case that your work was correct. In the figure, $t'$ and $t''$ are the two coordinates and $t$ is the coordinate of the top of the triangle.