The final exam will be cumulative: roughly 40% of the problems will be based on material covered prior to the midterm exam. For this material, see the separate midterm review document. The present document covers chapters 10 through 12 of the text.

CURVES IN SPACE: basic facts and definitions

- A vector-valued function \( r(t) = <x(t), y(t), z(t)> \) is called a space curve (or just a ‘curve’ if the context is clear). We assume that the derivatives of \( x(t), y(t), \) and \( z(t) \) do not all vanish simultaneously for any value of the parameter \( t \).

- The tangent vector to the curve is given by the first derivative: \( r'(t) = <x'(t), y'(t), z'(t)> \). The unit tangent vector is
  \[
  T(t) := \frac{r'(t)}{|r'(t)|}.
  \]

- The length of the parametric curve \( r(t) \) for \( a \leq t \leq b \) is
  \[
  L = \int_a^b |r'(t)| dt = \int_a^b \sqrt{x'^2(t) + y'^2(t) + z'^2(t)} \, dt.
  \]
  In principle, although this is difficult in practice, we can express the arc length \( s \) as a function of \( t \):
  \[
  s(t) = \int_a^t |r'(u)| \, du = \int_a^t \sqrt{x'^2(u) + y'^2(u) + z'^2(u)} \, du.
  \]
  Note that, by the fundamental theorem of calculus,
  \[
  \frac{ds}{dt} = \frac{d}{dt} \int_a^t |r'(u)| \, du = |r'(t)|.
  \]
  If \( t \) represents time, then \( \frac{ds}{dt} \) is the speed of the object, and is often denoted by \( v \) (lower case).

- Since the parametrization of a space curve is not unique, we standardize by defining as many quantities as possible in terms of the arc length \( s \). For instance, the unit tangent vector can be defined simply as
  \[
  T = \frac{dr}{ds} = \frac{dr}{dt} \frac{dt}{ds} = \frac{dr}{ds} = \frac{r'(t)}{|r'(t)|},
  \]
  which of course agrees with the definition given above.

- Since \( T \cdot T = 1 \), taking the derivative gives
  \[
  \frac{d}{ds}(T \cdot T) = 2 \frac{dT}{ds} \cdot T = 0,
  \]
  so \( \frac{dT}{ds} \) is orthogonal (or normal) to \( T \). The unit normal vector \( N \) is defined by
  \[
  \frac{dT}{ds} = \kappa N, \quad \text{where} \quad \kappa = \left| \frac{dT}{ds} \right| \quad \text{is the curvature.} \]
If the curvature vanishes, then $\frac{dT}{ds} = 0$, so the unit tangent vector is constant, and the curve doesn’t change direction – i.e., the curve is a straight line. Geometrically, $\frac{1}{\kappa}$ is the radius of the circle that best approximates the curve at the point in question.

- The curvature can be computed in terms of $t$ by
  \[
  \kappa(t) = \frac{|T'(t)|}{|r'(t)|} = \frac{|r'(t) \times r''(t)|}{|r'(t)|^3}.
  \]

- The *binormal* vector $B$ is defined by
  \[
  B = T \times N,
  \]
  so that $\{T, N, B\}$ form a positively oriented orthonormal frame at each point of the curve. One can show that the following equations, due to Frenet, hold for the derivatives:
  \[
  \frac{dT}{ds} = \kappa N, \\
  \frac{dN}{ds} = -\kappa T + \tau B, \\
  \frac{dB}{ds} = -\tau N
  \]
  \tau is called the *torsion* of the curve, and measures the rate at which the curve twists out of the plane determined by $T$ and $N$. In particular, if $\tau = 0$, then the binormal vector is constant, and the curve lies in a plane.

- The *acceleration* $r''(t)$ always lies in the plane spanned by $T$ and $N$:
  \[
  r''(t) = v'T + \kappa v^2 N.
  \]
  Note that if the speed is constant ($v' = 0$), then the acceleration is normal to the curve.

- We do not repeat the arguments here, but simply remind you that given Newton’s second law $F = ma$ and his law of gravitation $F = -\frac{GMm}{r^3} r$, one can deduce Kepler’s three laws of planetary motion:
  1. The planetary orbits are ellipses
  2. The radius vector joining a planet to the sun sweeps out equal areas in equal times (true for any central force).
  3. The period of the orbit is proportional to $a^{3/2}$, where $a$ is the length of the semi-major axis of the ellipse.

**PARAMETRIC SURFACES**

- In general, surfaces in $\mathbb{R}^3$ are given parametrically, by a vector-values function of two variables (the parameters):
  \[
  r(u, v) = (x(u, v), y(u, v), z(u, v)).
  \]
A standard example is the sphere of radius $a$, given parametrically by

\[
\begin{align*}
    x(\theta, \phi) &= a \cos(\phi) \sin(\theta) \\
    y(\theta, \phi) &= a \sin(\phi) \sin(\theta) \\
    z(\theta, \phi) &= a \cos(\theta)
\end{align*}
\]

Here $\theta$, the co-latitude, lies in $(0, \pi)$, and $\phi$, the longitude, in $(0, 2\pi)$. So $\mathbf{r}(\theta, \phi)$ maps an open rectangle in the parameter space onto the entire sphere except for the single meridian corresponding to the “international date line” (where $\phi$ would have to jump by $2\pi$). (Note that here, instead of following the text, $\theta$ and $\phi$ have been interchanged. This is actually the more common definition.)

- The graph of a function $f(x, y)$ of two variables is a parametric surface with parameters $x$ and $y$:

  \[
  \mathbf{r}(x, y) = < x, y, f(x, y) > .
  \]

A level set of a function of 3 variables can often be written as the graph of a function. For instance, the plane with equation $Ax + By + Cz = D$, with $C \neq 0$ can be written as the graph of $z = f(x, y) = (-1/C)(D - Ax - By)$.

- The graph of a function $y = f(x)$ revolved about the $x$ axis generates a surface of revolution which can be parametrized by $x$ and an angular variable $\theta$:

  \[
  \mathbf{r}(x, \theta) = < x, f(x) \cos(\theta), f(x) \sin(\theta) >
  \]

There are many examples of parametric surfaces scattered throughout your text.

FUNCTIONS OF SEVERAL VARIABLES

- A function $f(x, y)$ of two variables can be visualized in $\mathbb{R}^3$ by constructing its graph, which consists of all points of the form $< x, y, f(x, y) >$. The function can also be visualized by means of contour lines or level curves in the domain $D$ of the function. Such curves have the form

  \[
  \{ (x, y) : f(x, y) = c \}
  \]

for different values of the constant $c$. Topographical maps showing contours of equal height above sea level and weather maps showing isobars at the earth’s surface are two practical examples of this.

- A function $f(x, y, z)$ of three variables cannot be graphed, of course. But it can be partially visualized by means of level surfaces in $\mathbb{R}^3$; in analogy to the level curves above, the level surfaces are those surfaces in the domain of $f$ on which the function is constant.

- **Limits** of functions of several variables are a bit more complicated than those of a single variable, because, in order for the limit to exist, its value must be independent of the method of approach. A standard example is

  \[
  f(x, y) = \frac{x^2}{x^2 + y^2}.
  \]
a rational function of two variables which is continuous at all points in $\mathbb{R}^2$ except for $(0,0)$, where it is undefined. If the origin is approached along the $x$ axis ($y = 0$), then the limit appears to be 1. But if it’s approached along the $y$ axis, the limit appears to be 0. Since the result depends on the direction of approach, $\lim f(x,y)$ does not exist as $(x,y) \to (0,0)$.

- The definition of continuity is analogous to that for a function of a single variable: $f$ is continuous at $(a,b)$ if
  \[ \lim_{(x,y) \to (a,b)} f(x,y) = f(a,b). \]

- The partial derivative of $f(x,y)$ in the $x$ direction at the point $(a,b)$ is
  \[ \lim_{h \to 0} \frac{f(a + h, b) - f(a,b)}{h} \]
  provided that the limit exists. Notation: $\frac{\partial f}{\partial x}$, or $f_x$. Similarly, we can define $\frac{\partial f}{\partial y}$ by fixing $x = a$ and letting $y$ vary. Note that partial derivatives at a point can exist even though the function is not continuous there: if we define
  \[ f(x,y) = \begin{cases} \frac{x^2}{x^2 + y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases} \]
  then
  \[ \lim_{h \to 0} \frac{f(0, h) - f(0,0)}{h} = \lim_{h \to 0} \frac{0}{h} = 0, \]
  so $\frac{\partial f}{\partial y}(0,0) = 0$,
  even though, as we have seen, this function does not even have a limit at $(0,0)$.

- To compute $\frac{\partial f}{\partial x}$, we just regard $y$ as a constant, and differentiate $f$ as though it were a function of a single variable. The geometric interpretation of the partial derivative is similar as well: the plane $y = b$ is parallel to the $x$ axis and cuts the graph of $f$ in a curve which contains the point $< a, b, f(a, b)>$. The slope of the tangent line to this curve is $\frac{\partial f}{\partial x}(a,b)$.

- For a function of a single variable, the existence of the derivative means that the graph of the function has a tangent line at the point in question. The analog for a function of several variables is the existence of a tangent plane to the graph of $f$. Clearly the existence of both partial derivatives at the point in question is not sufficient to guarantee the existence of such a plane (why?). We need a stronger condition:

**Definition:** The function $f$ is differentiable at $(x_0, y_0)$ if we can write
\[
f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0) + f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \epsilon_1 \cdot \Delta x + \epsilon_2 \cdot \Delta y,
\]
where
\[
\lim_{(\Delta x, \Delta y) \to (0,0)} \epsilon_1 = \lim_{(\Delta x, \Delta y) \to (0,0)} \epsilon_2 = 0.
\]

If $f$ is differentiable at $(x_0, y_0)$, then, writing $x = x_0 + \Delta x$, $y = y_0 + \Delta y$, the function
\[
L(x, y) := f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)
\]
is called the linear approximation to $f$ at the point $(x_0, y_0)$. In words, the function $f$ is differentiable at $(x, y)$ if $f$ minus its linear approximation is quadratic or higher order in the quantities $\Delta x, \Delta y$. 

4
• **THEOREM:** If \( f_x \) and \( f_y \) are continuous at \((x_0, y_0)\), then \( f \) is differentiable at \((x_0, y_0)\).

• If \( f \) is differentiable at \((x_0, y_0)\), then

\[
z = L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)
\]

is the equation of the tangent plane to the graph of \( f \) at the point \((x_0, y_0, f(x_0, y_0))\).

Example: If \( f(x, y) = 3x^2y^3 - 2xy \), then at the point \((1, 2)\), we have \( f = 20 \), \( f_x = 44 \), and \( f_y = 34 \), so the linear approximation to \( f \) at \((1, 2)\) is \( L(x, y) = 20 + 44(x - 1) + 34(y - 2) \), and the equation of the tangent plane to the graph of \( f \) at \((1, 2, 20)\) is \( z = 20 + 44(x - 1) + 34(y - 2) \) or \( 44x + 34y - z = 92 \).

• Suppose \( \mathbf{r}(u, v) \) is the position vector of a point on a parametric surface, and suppose the derivatives \( \mathbf{r}_u = <x_u, y_u, z_u> \) and \( \mathbf{r}_v = <x_v, y_v, z_v> \) exist and are continuous at \((u_0, v_0)\).

These two vectors \( \mathbf{r}_u \) and \( \mathbf{r}_v \) are tangent to the surface at the point \( \mathbf{r}(u_0, v_0) \) and they are not parallel, so \( \mathbf{n} := \mathbf{r}_u \times \mathbf{r}_v \) is normal to the surface at that point. So the equation of the tangent plane to the surface is just

\[
(\mathbf{r} - \mathbf{r}(u_0, v_0)) \cdot \mathbf{n} = 0,
\]

where \( \mathbf{r} = <x, y, z> \) and \( \mathbf{r}(u_0, v_0) = <x_0, y_0, z_0> \).

In the simplest case, where the surface is the graph of some function \( f \), then \( \mathbf{r}(x, y) = <x, y, f(x, y)>, \mathbf{r}_x = <1, 0, f_x>, \mathbf{r}_y = <0, 1, f_y>, \) and \( \mathbf{n} = <-f_x, -f_y, 1> \), and the equation of the tangent plane to the surface is just the graph of the linear approximation (check this!).

**THE CHAIN RULE** The chain rule is the algorithm for computing derivatives of composite functions. There are many instances when such composite functions occur:

A curve on a parametric surface might be given by

\[
\mathbf{r}(t) = <x(u(t), v(t)), y(u(t), v(t)), z(u(t), v(t))>.
\]

Its tangent vector is given by

\[
\mathbf{r}'(t) = <\frac{dx}{dt} + x_u \frac{du}{dt}, \frac{dy}{dt} + y_u \frac{du}{dt}, \frac{dz}{dt} + z_u \frac{du}{dt} >.
\]

In a central force problem, the potential energy \( V \) depends only on the distance \( r \) from the center of force. Thus the potential energy at the point \((x, y, z)\) has the form \( V(r(x, y, z)) \). The chain rule gives

\[
\frac{\partial V}{\partial x} = V'(r) \frac{\partial r}{\partial x}.
\]

The temperature at a point \((x, y, z)\) is a scalar function which also generally depends on time \( t \). If a particle moves along a parametric curve in space, then the temperature \( T \) at its location is given by \( T(x(t), y(t), z(t), t) \). The derivative of \( T \) along the particle’s path is

\[
\frac{dT}{dt} = T_x \frac{dx}{dt} + T_y \frac{dy}{dt} + T_z \frac{dz}{dt} + T_t.
\]
Notice that when we write out the chain rule, we often omit the arguments of the functions in question. For instance, in the second example above, we should understand the expression given to be
\[ \frac{\partial V}{\partial x}(x, y, z) = V'(r(x, y, z)) \frac{\partial r}{\partial x}(x, y, z). \]

The “tree diagrams” discussed in your text are quite useful in this connection, and you are encouraged to use them.

**Evaluation of derivatives using the chain rule:** Suppose, in the third example above, you are explicitly given \( T \) as a function of \((x, y, z)\), and are also explicitly given \( x, y, \) and \( z \) as functions of \( t \). Then when you compute \( \frac{dT}{dt} \), your answer should be given in terms of the single variable \( t \). In particular, in \( \frac{\partial T}{\partial x} \), a function of \( x, y, \) and \( z \), the values of \( x, y, \) and \( z \) in terms of \( t \) should be computed.

**Implicit differentiation:** On the level surface \( F(x, y, z) = c \), we suppose that we can (in principle) solve for \( z \) in terms of \( x \) and \( y \). We then have \( F(x, y, z(x, y)) = c \). Taking the derivative of both sides with respect to \( x \), we find
\[ F_x + F_z z_x = 0, \text{ or } z_x = -\frac{F_z}{F_x}. \]

Clearly this only makes sense if \( F_z \neq 0 \). There is a similar expression for \( z_y \).

**Directional derivatives and gradients**

- In principle, we can take derivatives in any direction, not just in directions parallel to one of the coordinate axes: for any unit vector \( u = <\alpha, \beta> \) in the \( xy \) plane, and any point \((x_0, y_0)\) in the domain of \( f, f(x_0 + t\alpha, y_0 + t\beta) \) is a function of the single variable \( t \), and we define the directional derivative of \( f \) at \((x_0, y_0)\) in the direction \( u \) to be
  \[ D_u(f)(x_0, y_0) = \frac{d}{dt}f(x_0 + t\alpha, y_0 + t\beta)|_{t=0} \]

- Using the chain rule, we find
  \[ D_u(f)(x_0, y_0) = \alpha \frac{\partial f}{\partial x}(x_0, y_0) + \beta \frac{\partial f}{\partial y}(x_0, y_0), \]
  which we observe has the form of a dot product. We define the gradient of \( f \) at \((x_0, y_0)\) to be the vector
  \[ \nabla f(x_0, y_0) = <\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0)>, \]
  and write
  \[ D_u(f) = u \cdot \nabla f. \]

- Notice that if \( u = i \), then \( D_u(f) = \frac{\partial f}{\partial x} \). Similarly, \( D_j(f) = \frac{\partial f}{\partial y} = j \cdot \nabla f. \)
• Geometrically, $\mathbf{u} \cdot \nabla f$ is the slope of the tangent line to the curve determined by the intersection of the graph of $f$ and the plane parallel to the $z$ axis containing the line $<x_0, y_0> + t\mathbf{u}$.

• There are analogous corresponding definitions for functions of three or more variables.

• We have $\mathbf{D}_u f = |\nabla f| |\mathbf{u}| \cos(\theta) = |\nabla f| \cos(\theta)$, where $\theta$ is the angle between $\nabla f$ and $\mathbf{u}$. It’s clear that the magnitude of the directional derivative is greatest if $\cos(\theta) = 1$ (i.e. $\mathbf{u} \parallel \nabla f$ and pointing in the same direction). It achieves its minimal value if $\mathbf{u} \parallel \nabla f$ and points in the opposite direction. Thus $\nabla f(x_0, y_0)$ points in the direction of the greatest change of the function $f$ at the point $(x_0, y_0)$.

• To proceed in a direction in which $f$ does not change at all, one chooses a direction orthogonal to the gradient, since in this case, $\mathbf{D}_u f = \mathbf{u} \cdot \nabla f = 0$. In particular, it follows that the gradient vector is orthogonal to the level curves (or surfaces) of the function $f$. So if a surface $S$ in $\mathbb{R}^3$ is given by an equation of the form $f(x, y, z) = c$ (i.e., a level surface of $f$), then the normal to $S$ at some point $(x_0, y_0, z_0)$ on the surface is given by $\nabla f$ at this point.

For a surface in $\mathbb{R}^3$ given by the graph of a function $f(x, y)$, we can define $F(x, y, z) := z - f(x, y)$, so the surface is given by the equation $F = 0$, and the surface normal is $\nabla F = < -f_x, -f_y, 1 >$, in agreement with our previous result.

FINDING EXTREME VALUES OF FUNCTIONS OF SEVERAL VARIABLES

• Suppose $f$ is defined on some domain in $\mathbb{R}^3$ and $(x_0, y_0, z_0)$ is in this domain. If $f(x, y, z) \leq f(x_0, y_0, z_0)$ for all points $(x, y, z)$ in some ball around $(x_0, y_0, z_0)$, then $f$ is said to have a local maximum at $(x_0, y_0, z_0)$. Local minima are defined by reversing the inequality. If $f$ has a local extremum (either a local max or min) at $(x_0, y_0, z_0)$, and if $f$ is differentiable at $(x_0, y_0, z_0)$, then all directional derivatives of $f$ must vanish there. Equivalently, $\nabla f = \mathbf{0}$ at $(x_0, y_0, z_0)$, which is called a critical point of the function $f$. Note that there are three equations here: $f_x = f_y = f_z = 0$. So, to find candidates for local extrema, one sets $\nabla f = \mathbf{0}$ and solves the resulting three equations simultaneously. In general, this is rather difficult. Similar considerations apply to a function of two variables, except that here there are just two equations to be solved simultaneously.

• Just as in the case of a function of a single variable, the vanishing of the partial derivatives does not guarantee the existence of a local extremum. There is an analog to the second derivative test which we state for functions of two variables only:

SECOND DERIVATIVE TEST: Suppose $(x_0, y_0)$ is a critical point of $f$, and that all the second partial derivatives of $f$ exist and are continuous at this point. Define

$$D = f_{xx}f_{yy} - f_{xy}^2.$$ 

Then

- If $f_{xx} > 0$ at $(x_0, y_0)$, and $D > 0$ at $(x_0, y_0)$, $f$ has a local minimum at $(x_0, y_0)$;
- If $f_{xx} < 0$ and $D > 0$, then $f$ has a local maximum;
- If $D < 0$ at $(x_0, y_0)$, then $(x_0, y_0)$ is a saddle point.
- If $D = 0$ at the critical point, then the second derivative test fails.

- As in the case of a single variable, the existence of global or absolute extrema involves not only the value of $f$ at the critical points, but also the values of $f$ on the boundary of the domain.

- **Lagrange Multipliers:** The problem of finding the local extrema of $f$ when the function $f$ is subject to constraints is addressed through the use of Lagrange multipliers. Specifically, suppose it is desired to find the extrema of the function $f$ subject to the constraint that the point(s) must lie on the surface $g(x, y, z) = \text{constant}$. Examples: (a) Find the point on the surface $x^2 + 3y^2 + z^2 = 10$ (an ellipsoid) closest to the point (2, 3, 4). Here the function $f$ is the distance from an arbitrary point $(x, y, z)$ to the point (2, 3, 4), and the constraint is that $(x, y, z)$ must lie on the given ellipsoid. So $g = x^2 + 3y^2 + z^2$. (b) Minimize the potential energy function $V(x, y, z)$ for a particle constrained to lie on the surface $g(x, y, z) = \text{constant}$.

  In order for these problems to have a solution at some point $P$ on the surface, the vector $\nabla f$ must be orthogonal to the constraint surface $g = \text{constant}$ – if not, the vector projection of $\nabla f$ on to the surface at $P$ gives the directions of maximal increase and decrease of $f$ on the surface, and if these directions exist, then there can’t be a local extremum of $f$ at $P$. Since $\nabla g$ is certainly orthogonal to the surface at $P$, a necessary condition for an extremum is that the two gradient vectors be parallel. That is, we should have

  $\nabla f = \lambda \nabla g$

  for some non-zero scalar $\lambda$ (called a Lagrange multiplier). If the domains of $f$ and $g$ are in $\mathbb{R}^3$, then this vector equation consists of three scalar equations for the four unknowns $x, y, z$, and $\lambda$; the fourth equation is $g = \text{constant}$. So altogether, there are four equations in four unknowns. For a problem in $\mathbb{R}^2$, there would be three equations in the three unknowns $x, y$, and $\lambda$.

- In some cases, one can avoid the use of Lagrange multipliers by solving the equation $g = \text{constant}$ for (say) $z$ as a function of $(x, y)$, substituting this into $f$ to get a function of just $x$ and $y$, and trying to find the extrema of the resulting function of two variables.

**MULTIPLE INTEGRATION**

- A continuous function $f$ of several variables can be integrated over its domain in much the same manner as a function of a single variable: the domain (we suppose it’s a rectangle for convenience) is partitioned into smaller rectangles with areas $\Delta A_{ij} = \Delta x_i \Delta y_j$. An arbitrary point $(x_i^*, y_j^*)$ is picked from each rectangle and the Riemann sum

  $$\sum_{i=1}^{m} \sum_{j=1}^{n} f(x_i^*, y_j^*) \Delta A_{ij}$$
is formed. The limit of this sum, as \(|\Delta x_i|\) and \(|\Delta y_j|\) → 0, and \(m, n \to \infty\) is called the Riemann integral of \(f\) over the rectangle \(R\), and is denoted

\[
\int \int_R f \, dA \quad \text{or} \quad \int \int_R f(x, y) \, dx \, dy.
\]

This is often also called a *double* integral. Note that if \(f = 1\), then this integral computes the area of \(R\).

Analogously, triple and quadruple integrals can be defined over suitable regions of \(\mathbb{R}^3\) and \(\mathbb{R}^4\).

- Except in very special cases where the definition can be used, integrals are generally evaluated by reducing them to a sequence of one-dimensional (iterated) integrals. To motivate this, suppose \(f > 0\) in the rectangle \(a \leq x \leq b, \ c \leq y \leq d\), and we want to compute the volume of the region under the graph of \(f\).

For each \(x_0\) in \([a, b]\) we define the function \(A(x_0)\) to be the area of the cross-section of the region under the graph cut out by the plane \(x = x_0\). We then drop the subscript, let \(x\) vary, and find that

\[
V = \text{volume} = \int \int_R f \, dA = \int_{x=a}^{x=b} A(x) \, dx.
\]

But \(A(x)\) is just the area under the graph of \(f(x, y)\), with \(x\) held constant:

\[
A(x) = \int_{y=c}^{y=d} f(x, y) \, dy, \quad \text{or} \quad V = \int_{x=a}^{x=b} \int_{y=c}^{y=d} f(x, y) \, dy \, dx.
\]

This last integral is called an *iterated* integral. It is evaluated from inside to outside: that is, the integration is done first with respect to \(y\) and then with respect to \(x\). Iterated integrals can be used whether or not \(f\) is positive. Moreover, in principle, it makes no difference how we slice up the domain of \(f\): in the case above, we could have found the volume by adding up the volumes of thin slices perpendicular to the \(y\) axis:

\[
V = \int_{y=c}^{y=d} A(y) \, dy = \int_{y=c}^{y=d} \int_{x=a}^{x=b} f(x, y) \, dx \, dy.
\]

Clearly, the two different iterated integrals must give the same answer \((V)\) - this is the content of Fubini’s theorem.

The point of using iterated integrals is that we can evaluate the single integrals (sometimes) using the fundamental theorem of calculus.

- Of course, we can do iterated integrals over more complicated domains.

Example: \(R\) is the triangle formed by the lines \(y = 2, x = 1\) and \(y = -2x + 2\). Suppose we want to find the integral of \(f(x, y) = 2x^2y\) over \(R\). We have

\[
\int \int_R f \, dA = \int_{x=0}^{x=1} \int_{y=-2x+2}^{y=2} 2x^2y \, dy \, dx = \int_{y=0}^{y=2} \int_{x=1-(y/2)}^{x=1} 2x^2y \, dx \, dy = \frac{2}{15}.
\]
There are a number of variations on this theme, mostly involving different specifications of the region $R$. See your text.

- Sometimes, it is more natural to integrate a function in a different coordinate system. For example, integrating a function depending only on $r = \sqrt{x^2 + y^2}$ might be easier in polar coordinates. When using another coordinate system, it’s necessary to check the form of “$dA$” in the new system. For instance, in a polar coordinate grid, the region bounded by $\theta_0 \leq \theta \leq \theta_1$, $r_0 \leq r \leq r_1$ has the approximate area $\Delta A \approx r\Delta r\Delta \theta$, where, for instance, $\Delta r = r_1 - r_0$. So an integral in polar coordinates will have the form

$$\int \int_{R} f(r, \theta) \ r dr d\theta,$$

and this can be evaluated in the usual fashion as an iterated integral.

**Example:** Evaluate $\int \int_{R} f(x, y) \ dx dy$, where $R$ is the unit circle and $f(x, y) = e^{x^2 + y^2}$.

**Solution:** substituting $x = r \cos \theta$, $y = r \sin \theta$ into $f$, and changing to polar coordinates, we obtain

$$\int \int_{R} f(r \cos \theta, r \sin \theta) \ r dr d\theta = \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=1} re^{r^2} \ dr d\theta$$

$$= \int_{\theta=0}^{\theta=2\pi} \left[\left(e^{r^2}/2\right)\right]_0^1 \ d\theta$$

$$= \pi(e - 1).$$

- **Some applications**

  - For an inhomogeneous thin plate, with mass density $\rho$ given in (say) grams/cm$^2$, the mass of the plate is found by integrating the mass density over the region $R$ occupied by the plate:

  $$m = \int \int_{R} \rho \ dA.$$

  - The **moments** of this plate about the axes are

  $$M_x = \int \int_{R} \rho y \ dA; \quad M_y = \int \int_{R} \rho x \ dA.$$

  The two moments describe the tendency of the plate to rotate about the corresponding axes. Notice that the moments depend on the choice of coordinates; in fact, we can make both of these moments vanish by choosing the origin correctly. The **center of mass** of the plate has the coordinates

  $$(\bar{x}, \bar{y}) = \left(\frac{M_y}{m}, \frac{M_x}{m}\right).$$

  If new coordinates are chosen according to $x^* = x - \bar{x}, y^* = y - \bar{y}$, then in this system, we find $(\bar{x}^*, \bar{y}^*) = (0, 0)$. The physical meaning of this is that the plate can be balanced at the center of mass: there is no tendency to rotate about either axis if the axes pass through $(\bar{x}, \bar{y})$. 

10
– The second moments or the moments of inertia cannot be transformed away by using different coordinates: They are defined by

\[ I_x = \int \int_R \rho y^2 \, dA; \quad I_y = \int \int_R \rho x^2 \, dA. \]

The moment of intertia about the origin, \( I_0 \), is just \( I_x + I_y = \int \int_R (x^2 + y^2) \rho \, dA \).

The moment \( I_0 \) plays, in some sense, the same rôle for rotating bodies as mass does for linear motion.

– There are many, many other examples of multiple integrals in physics and engineering: The total electric charge in a region is obtained by integrating the charge density over the region; the electric field at some point \( P \) is obtained by adding up (i.e. integrating) the forces acting on a unit charge located at \( P \). The linear momentum of an extended body is obtained by dividing the body into small regions, computing the momenta of these regions in an elementary way, and adding things up (integrating).

- **Surface area** Suppose we’re given a parametric surface \( r(u,v) \) in \( \mathbb{R}^3 \). A small rectangle with sides \( \Delta u, \Delta v \) in the domain of \( r \) gets mapped to a curvilinear figure in \( \mathbb{R}^3 \) whose area is (approximately) given by

\[ \Delta A \approx |r_u \times r_v| \Delta u \Delta v. \]

Muttering the usual incantations and adding things up over the whole surface gives

\[ A = \int \int_R |r_u \times r_v| \, dudv, \]

where \( R \) is the domain of \( r \) in the \( uv \) plane.

Example: If the surface is given by the graph of a function of two variables, then

\[ r(x,y) = <x,y,f(x,y)>, \quad r_x = <1, 0, f_x>, \quad r_y = <0, 1, f_y>. \]

Then \( r_x \times r_y = <-f_x, -f_y, 1> \), so

\[ |r_x \times r_y| = \sqrt{1 + f_x^2 + f_y^2}, \]

and the surface area of the graph is given by

\[ A = \int \int_R \sqrt{1 + f_x^2 + f_y^2} \, dxdy. \]

Note the similarity to the formula for arc length of the graph of a function of a single variable.