DEPARTMENT OF MATHEMATICS
OREGON STATE UNIVERSITY

CALCULUS STUDY GUIDE
MTH 255

BY

TEVIAN DRAY

Revised edition
5 September 2007

LESSON SUMMARY

Study Guide	**MHG** (4th edition)	**Stewart** (5th edition)
Warmup	§13.1–§13.3	§12.2–§12.3
Lesson 1	(§14.3)	—
Lesson 2	—	—
Lesson 3	§14.4–§14.5	§14.6
Lesson 4	§17.3	§16.1
Lesson 5	§18.1–§18.2 (§17.1)	§16.2
Lesson 6	§18.3	§16.3
Lesson 7	—	—
Interlude	§13.4	§12.4
Lesson 8	(§17.5)	§16.6
Lesson 9	§19.1–§19.3	§16.7
Lesson 10	§20.1–§20.2	§16.9
Lesson 11	§18.4	§16.4
Lesson 12	§20.3–§20.4	§16.8
Lesson 13	(§20.1,§20.3)	§16.5
Lesson 14	§16.7	§15.9
REVIEW	§20.5	§16.10

Contents

INTRODUCTION

This is a Study Guide for MTH 255, Vector Calculus II. The Study Guide can in principle be used with any calculus text, although it was developed using both the fourth edition of **Calculus** by James Stewart (either the *Multivariable* or *Early Transcendentals* versions) and the third edition of **Multivariable Calculus** by William McCallum, Deborah Hughes-Hallett, Andrew Gleason, et al. A rough correspondence between sections in (more recent editions of) those texts and lessons in the Study Guide is provided on the inside front cover.

The Study Guide contains a summary of the key concepts of vector calculus, which are intended to *guide* you through the material in the book and lectures. You are strongly encouraged to at least skim the appropriate section of the Study Guide *before* the corresponding lecture, then (re)read it afterwards together with the text.

The Study Guide also contains material which is not in most texts, which may enhance your understanding of the underlying geometry of vector calculus. However, since every instructor is different, there is no guarantee that all the material in the Study Guide will be covered in class, nor that all the material covered in class will be in the Study Guide.

The Vector Calculus Bridge Project

The approach to vector calculus used in the Study Guide is an essential part of the Vector Calculus Bridge Project, whose homepage is:

$$\texttt{http://www.math.oregonstate.edu/bridge}$$

This approach emphasizes geometric visualization rather than (just) algebraic manipulation. Since most calculus textbooks emphasize the latter, you may occasionally feel that the Study Guide is speaking a different language than the textbook. To some extent, this is correct, and reflects cultural differences between different fields of science. Students in some disciplines, such as physics or electrical engineering, may find the approach used here especially familiar. But being comfortable with both algebraic and geometric representations is an important skill for everybody; it never hurts to be bilingual.

Homework

The only way to master the material in this course is to solve problems. Your instructor will most likely assign a few problems each week. ***Do not expect to succeed in this course if you only attempt the assigned problems.*** It is your responsibility to attempt as many problems as necessary on your own, both routine problems and more challenging problems, and to seek help as needed.

Review

Familiarity with the following topics will be assumed in this course: basic properties of vectors, the dot and cross products, partial differentiation, multivariable differentials and chain rule, and multiple integration (including polar coordinates). This course will provide you with an opportunity to consolidate your understanding of these topics; a brief review of some of this material appears in the first few lessons.

Other Resources

In addition to your instructor and TA, help is available in the Mathematics Learning Center (MLC) on the ground floor of Kidder Hall. OSU also has an (old) online vector calculus study guide whose current URL is: *

`http://www.math.oregonstate.edu/home/programs/undergrad/CalculusQuestStudyGuides/vcalc/255.html`

For those using other texts, a geometric approach to vector calculus can be found in

> William McCallum, Deborah Hughes-Hallett, Andrew Gleason, et al., **Multi-variable Calculus**, 4th edition Wiley, 2005.

which is also available combined with single variable material as

> Deborah Hughes-Hallett, William McCallum, Andrew Gleason, et al., **Calculus**, 4th edition Wiley, 2005.

A copy of (an earlier edition of) the latter version should be available in the library. Finally, there is an excellent informal book on vector calculus from the point of view of a scientist or engineer, entitled

> **Div, Grad, Curl, and All That; an informal text on vector calculus**, by H. M. Schey, 4th edition, Norton & Co., 2005.

Acknowledgments

DISCLAIMER: The online study guide has no official connection with this course.

NOTATION

WARNING: This Study Guide does NOT use the same conventions for spherical coordinates as those used in your text. Instead, the Study Guide uses the most commonly used choice in science and engineering applications, explicitly given by

$$x = r \sin\theta \cos\phi \qquad y = r \sin\theta \sin\phi \qquad z = r \cos\theta$$

You will often see ρ (or R) instead of r, and/or the roles of θ and ϕ interchanged. However, spherical coordinates should ALWAYS be given in the following order:

(radius, angle from z-axis, angle in xy-plane)

For us this is (r, θ, ϕ). BE CAREFUL to correctly translate the text's notation! You should feel free to use whichever convention you are most comfortable with, but you should be aware that other conventions exist, and that your text's choice is not widely used outside the mathematics community. *

We will also adopt the convention used in some science and engineering texts and use ϕ rather than θ for polar coordinates, so that the angle in the xy-plane is always ϕ. We therefore write

$$x = r \cos\phi \qquad y = r \sin\phi$$

Again, this is not the choice used by your textbook; be careful! *

---

*You can find out more about the reasons for these choices in our paper:

Tevian Dray and Corinne A. Manogue, *Spherical Coordinates*,
College Math. J. **34**, 168–169 (2003).

which you can also find online at:

http://www.math.oregonstate.edu/bridge/papers/spherical.pdf

The short answer is that most students will need to switch conventions at some point during their education, so this might as well be done sooner rather than later.

WARMUP

Vectors

A *vector* $\vec{w}$ is an arrow in space, having both a *magnitude* and a *direction*. Examples of vectors include the displacement from one point to another, and your velocity when moving along some path.

The *magnitude* of $\vec{w}$ is denoted by $|\vec{w}|$, also written $||\vec{w}||$. This is often casually called the "length" of $\vec{w}$, but that is only technically correct if $\vec{w}$ has units of length. The displacement between two points does indeed have a length, but to talk about your speed as being the "length" of your velocity vector is confusing at best.

A *unit vector* is a vector whose magnitude is 1. Given any vector $\vec{w}$, we get a unit vector by dividing $\vec{w}$ by its magnitude $|\vec{w}|$, that is, $\frac{\vec{w}}{|\vec{w}|}$ is a unit vector.

Note that unit vectors are dimensionless!

If you divide your velocity by your speed, both measured in miles per hour, the magnitude of the result is just plain 1, *not* "1 mph". What information is left? The direction you are going. We therefore define the *direction* of a vector $\vec{w}$ to be the unit vector in the same direction, which is just $\frac{\vec{w}}{|\vec{w}|}$.

As is customary in many branches of science, we will denote unit vectors with a hat, rather than an arrow. We therefore write the unit vectors in the x, y, and z directions respectively as $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$. There are other common notations for these vectors. For instance, you may see $\hat{\imath}$ written instead as any of $\vec{\imath}$, $\vec{a}_x$, $\vec{e}_x$, or $\hat{x}$. You may also see the notation w and $\hat{w}$ for the magnitude and direction, respectively, of a given vector $\vec{w}$, so that $w = |\vec{w}|$ and $\hat{w} = \frac{\vec{w}}{|\vec{w}|}$.

Any vector $\vec{a}$ in 3 dimensions can of course be expressed in terms of its components. For instance, we might have $\vec{a} = 2\,\hat{\imath} + 3\,\hat{\jmath} + 4\,\hat{k}$. But it is important to think of vectors as arrows in space, rather than as a list of 3 components. While these points of view are equivalent, the former is more geometric. One of the goals of this course is to develop this geometric intuition so that you can use it in applications. It is in part for this reason that, unlike many texts, we *always* include the basis vectors $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ when referring to the (Cartesian) components of a vector; we never write "$\vec{a} = \langle 2, 3, 4 \rangle$".

For an arbitrary vector $\vec{w}$ with unknown (Cartesian) components, we write

$$\vec{w} = w_x\,\hat{\imath} + w_y\,\hat{\jmath} + w_z\,\hat{k}$$

Note the use of subscripts to denote the components; these subscripts do *not* denote partial differentiation, and in fact for this reason we *never* use subscripts to denote derivatives.
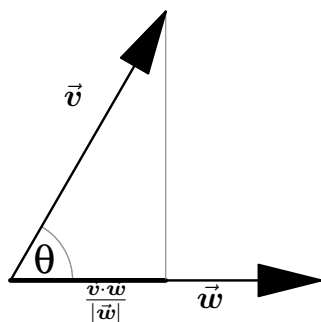
Dot Product

The dot product is fundamentally a projection, as shown in the figure above. In the figure, the length of the heavy line segment represents the component of $\vec{v}$ in the $\vec{w}$ direction. This defines the dot product of $\vec{v}$ with $\frac{\vec{w}}{|\vec{w}|}$; multiplying through by $|\vec{w}|$ and using some triangle trigonometry leads to the geometric formula

$$\vec{v} \cdot \vec{w} = |\vec{v}||\vec{w}| \cos \theta$$

for the dot product of $\vec{v}$ and $\vec{w}$. (Projecting $\vec{w}$ onto $\vec{v}$ instead leads to the same result.) An immediate consequence of this formula is that the dot product of a vector with itself gives the square of its magnitude, that is

$$\vec{v} \cdot \vec{v} = |\vec{v}|^2$$

In particular, taking the "square" of any unit vector yields 1, for example

$$\hat{\imath} \cdot \hat{\imath} = 1$$

Furthermore, since $\cos \frac{\pi}{2} = 0$, it follows immediately from the geometric definition that two vectors are orthogonal if and only if their dot product vanishes, that is

$$\vec{v} \perp \vec{w} \iff \vec{v} \cdot \vec{w} = 0$$

It is easy to see that the geometric formula reduces to the standard component form of the dot product. For example, if $\vec{v} = v_x \, \hat{\imath} + v_y \, \hat{\jmath}$ and $\vec{w} = w_x \, \hat{\imath} + w_y \, \hat{\jmath}$, then

$$
\begin{aligned}
\vec{v} \cdot \vec{w} &= (v_x \, \hat{\imath} + v_y \, \hat{\jmath}) \cdot (w_x \, \hat{\imath} + w_y \, \hat{\jmath}) \\
&= v_x w_x \, \hat{\imath} \cdot \hat{\imath} + v_y w_y \, \hat{\jmath} \cdot \hat{\jmath} + v_x w_y \, \hat{\imath} \cdot \hat{\jmath} + v_y w_x \, \hat{\jmath} \cdot \hat{\imath} \\
&= v_x w_x + v_y w_y
\end{aligned}
$$

The dot product of a vector with itself reduces to the Pythagorean theorem when written out in terms of components, for example

$$\vec{v} \cdot \vec{v} = |\vec{v}|^2 = v_x^2 + v_y^2$$

Note that in the above computation we simply "multiplied out" the original expression in terms of components, then used the multiplication table

$$\hat{\imath} \cdot \hat{\imath} = 1 = \hat{\jmath} \cdot \hat{\jmath}$$
$$\hat{\imath} \cdot \hat{\jmath} = 0 = \hat{\jmath} \cdot \hat{\imath}$$

This is a useful strategy which you are encouraged to practice.

A good test of your understanding of the dot product is try to use it to find the angle between the diagonal of a cube and one of its edges, or the angle between two diagonals.

Lesson 1:   Differentials

Given a function $f = f(x)$, the *differential* of f is defined to be

$$df = \frac{df}{dx}\,dx$$

This definition is not very satisfying. What, after all, is dx? There are several ways to make this precise. In the branch of mathematics known as differential geometry, dx and df are *differential forms*, but this is more sophistication than we need.

We can safely think of dx as an infinitesimal change in the x direction, and df as the corresponding infinitesimal change in f. Then the above equation shows how these infinitesimal changes are related. That's exactly what a derivative is, a ratio of infinitesimal changes!

One can also regard dx and df as "things which you integrate". You are already used to using differentials this way — just think about substituting $u = x^2$ into the following integral:

$$\int 2x \sin(x^2)\,dx = \int \sin(u)\,du$$

Didn't you use the relation $du = 2x\,dx$? Those are differentials!

Another way to think of differentials is as the numerators of derivatives, in the sense that "dividing" differentials leads to derivatives. For instance, if $u = x^2$ as above, and $x = \cos\theta$, then

$$du = 2x\,dx = 2x(-\sin\theta\,d\theta) = -2\cos\theta\sin\theta\,d\theta$$

and "dividing" through by $d\theta$ yields

$$\frac{du}{d\theta} = -2\cos\theta\sin\theta$$

which is just the chain rule, since $u = x^2 = \cos^2\theta$. In fact, using differentials makes the chain rule automatic!

Although it is hard to give a rigorous definition of differentials at this mathematical level, these two examples should convince you that they are a useful tool. Even if you view the use of differentials as nothing more than formal manipulation, you will always get correct answers — provided, of course, that you manipulate them correctly.

What about a small changes in a function of two variables, $f = f(x, y)$? We can approximate this change using the tangent plane to the graph of $z = f(x, y)$ which can be written in the form

$$\Delta z = \frac{\partial f}{\partial x}\,\Delta x + \frac{\partial f}{\partial y}\,\Delta y$$

where the partial derivatives are to be evaluated at the point where the tangent plane touches the graph. This equation is *linear*; planes are straight. And if you know how a plane is tilted in two directions, you know everything about it — it takes precisely two slopes, given by the partial derivatives, to determine the plane.

So we have

$$\Delta f \approx \Delta z$$

with Δz as above. As we zoom in, the approximation becomes better and better, so we write

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

A similar argument works for functions of 3 variables. Suppose that $f = f(x, y, z)$. Then small changes in, say, x will produce small changes in f, and similarly for small changes in y or z. But changing one variable while holding the others fixed is precisely what partial differentiation is all about, so it should come as no surprise that the differential of f is now

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz$$

Once again, the use of differentials makes the chain rule automatic. Suppose that x, y, z are functions of u. Then $dx = \frac{dx}{du}\,du$, etc.; substituting this into the above expression for df and "dividing" by du yields the standard chain rule formula for $\frac{df}{du}$, namely

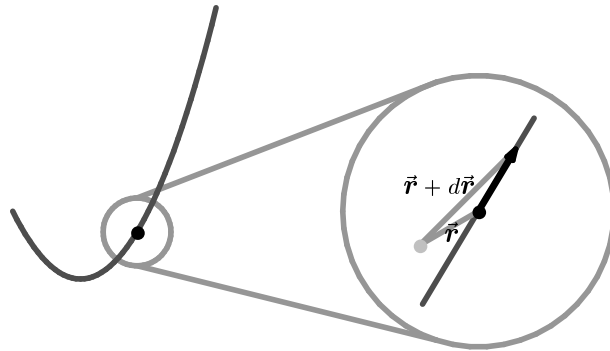$$\frac{df}{du} = \frac{\partial f}{\partial x}\frac{dx}{du} + \frac{\partial f}{\partial y}\frac{dy}{du} + \frac{\partial f}{\partial z}\frac{dz}{du}$$

More complicated cases, such as x, y, z being themselves functions of several variables, are only slightly more difficult once you understand the basic idea.

GOALS

- Know how to use differentials.

- Know the difference between differentials and derivatives.

Lesson 2:   Vector Differentials

The vector from the origin to the point (x, y, z) is

$$\vec{r} = x\,\hat{\boldsymbol{\imath}} + y\,\hat{\boldsymbol{\jmath}} + z\,\hat{\boldsymbol{k}}$$

and is called the *position vector*. A parametric curve can be described by viewing x, y, and z, and hence $\vec{r}$, as functions of some parameter u. It is instructive to draw a picture showing the small change $\Delta\vec{r} = \Delta x\,\hat{\boldsymbol{\imath}} + \Delta y\,\hat{\boldsymbol{\jmath}} + \Delta z\,\hat{\boldsymbol{k}}$ in the position vector between nearby points. Try it!

This picture is so useful that we will extend it one step further, and consider an infinitesimal change in position. What does this mean? Zoom in on the curve so much that it becomes straight, as illustrated in the figure above. Instead of $\Delta\vec{r}$, we will write $d\vec{r}$ for the vector between two points on the curve which are infinitesimally close together. For the moment, that's just notation.

Like any vector, $d\vec{r}$ can be expanded with respect to $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$, $\hat{\boldsymbol{k}}$; the components of $d\vec{r}$ are just the infinitesimal changes dx, dy, dz, in the x, y, and z directions, respectively, that is

$$d\vec{r} = dx\,\hat{\boldsymbol{\imath}} + dy\,\hat{\boldsymbol{\jmath}} + dz\,\hat{\boldsymbol{k}} \tag{2.1}$$

"Dividing" this by du leads to the derivative

$$\frac{d\vec{r}}{du} = \frac{dx}{du}\,\hat{\boldsymbol{\imath}} + \frac{dy}{du}\,\hat{\boldsymbol{\jmath}} + \frac{dz}{du}\,\hat{\boldsymbol{k}}$$

which could also have been constructed as a limit, namely

$$\frac{d\vec{r}}{du} = \lim_{\Delta u \to 0} \frac{\Delta\vec{r}}{\Delta u}$$

You may find the intuitive notion of $d\vec{r}$ as an infinitesimal vector displacement to be helpful in visualizing the geometry of vector calculus throughout this course. If you want to be more pedantic, you may view this as a purely formal manipulation, or as a definition of the vector differential $d\vec{r}$, that is

$$d\vec{r} = \frac{d\vec{r}}{du}\,du$$

By construction, $\frac{d\vec{r}}{du}$, and hence also $d\vec{r}$, is tangent to the curve — and if the parameter u happens to be time t, then $\frac{d\vec{r}}{dt}$ is of course just the velocity.

8
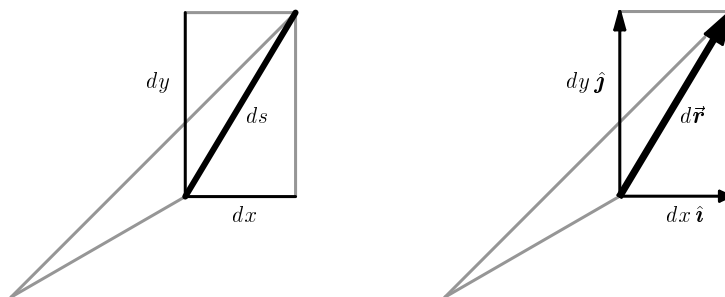
What is the infinitesimal distance ds between nearby points on the curve? Just the length of $d\vec{r}$! We have

$$ds = |d\vec{r}|$$

and "squaring" both sides leads to

$$ds^2 = d\vec{r} \cdot d\vec{r} = dx^2 + dy^2 + dz^2 \tag{2.2}$$

which is just the infinitesimal Pythagorean Theorem, and which is shown (in two dimensions) in the first drawing above. The vector version of the same result is shown in the second drawing; this is just (2.1) (restricted to two dimensions).

Other Coordinates

It is important to realize that $d\vec{r}$ and ds are defined geometrically, not by (2.1) and (2.2). To emphasize this coordinate-independent nature of $d\vec{r}$, it is useful to study $d\vec{r}$ in another coordinate system, such as polar coordinates (r,ϕ) in the plane. It is also natural to introduce basis vectors $\{\hat{r}, \hat{\phi}\}$ adapted to these coordinates, with $\hat{r}$ being the unit vector in the radial direction, and $\hat{\phi}$ being the unit vector in the direction of increasing ϕ. *

Determining the lengths of the sides of the resulting infinitesimal polar "rectangle" yields

$$d\vec{r} = dr\,\hat{r} + r\,d\phi\,\hat{\phi} \tag{2.3}$$

as shown in the second drawing below. Using (2.2) and again squaring both sides leads to

$$ds^2 = d\vec{r} \cdot d\vec{r} = dr^2 + r^2\,d\phi^2 \tag{2.4}$$

which is the infinitesimal Pythagorean Theorem in polar coordinates

A similar procedure can be used to determine $d\vec{r}$ in cylindrical and spherical coordinates.

---
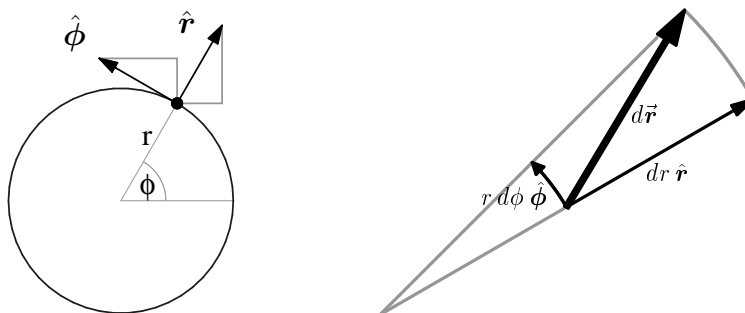
*One can of course relate $\hat{r}$ and $\hat{\phi}$ to $\hat{i}$ and $\hat{j}$, for instance using the first drawing in the figure at the bottom of the page. However, in most physical applications (as opposed to problems in calculus textbooks) this can be avoided by making an appropriate initial choice of coordinates.

9

GOALS

- Know how to derive the formula for the vector differential $d\vec{r}$ in standard coordinate systems (rectangular, polar, cylindrical, spherical).

- Know the resulting formulas for $d\vec{r}$; these are worth memorizing.

- Know the relationship between $d\vec{r}$ and infinitesimal arclength ds.

Lesson 3:   The Gradient

Take another look at the formula for the differential of a function of several variables:

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz$$

Each term is a product of two factors, labeled by x, y, and z. This looks like a dot product! Separating out the pieces, we have

$$df = \left(\frac{\partial f}{\partial x}\, \hat{\boldsymbol{\imath}} + \frac{\partial f}{\partial y}\, \hat{\boldsymbol{\jmath}} + \frac{\partial f}{\partial z}\, \hat{\boldsymbol{k}}\right) \cdot (dx\, \hat{\boldsymbol{\imath}} + dy\, \hat{\boldsymbol{\jmath}} + dz\, \hat{\boldsymbol{k}})$$

The last factor is just $d\vec{r}$, and you may recognize the first factor as the gradient of f, that is

$$\vec{\nabla} f = \frac{\partial f}{\partial x}\, \hat{\boldsymbol{\imath}} + \frac{\partial f}{\partial y}\, \hat{\boldsymbol{\jmath}} + \frac{\partial f}{\partial z}\, \hat{\boldsymbol{k}}$$

Putting this all together, we have the *master formula* *

$$df = \vec{\nabla} f \cdot d\vec{r} \tag{3.1}$$

Recall that df represents the infinitesimal change in f when moving to a "nearby" point. What information do you need in order to know how f changes? You must know where you started, which way you went, and something about how f behaves. The master formula organizes this information into two geometrically different pieces, namely the gradient, containing generic information about how f changes at the point in question, and the vector differential $d\vec{r}$, containing information about the particular change in position being made.

Like all derivative operators, the gradient is linear (the gradient of a sum is the sum of the gradients), and also satisfies a product rule

$$\vec{\nabla}(fg) = (\vec{\nabla} f)\, g + f\, (\vec{\nabla} g)$$

Differentials such as df are rarely themselves the answer to any physical question. So what good is the master formula? The short answer is that you can use it to answer *any* question about how f changes. Here are some examples.

1. *Suppose you are an ant walking in a puddle on a flat table. The depth of the puddle is given by $h(x,y)$. You are given x and y as functions of time t. How fast is the depth of water through which you are walking changing per unit time?*

This problem is asking for the derivative of h with respect to t. So divide the master formula by dt to get

$$\frac{dh}{dt} = \vec{\nabla} h \cdot \frac{d\vec{r}}{dt}$$

---

*This formula can in fact be taken as the *definition* of the gradient!

where $\vec{r}$ describes the particular path you are taking. The factor $\frac{d\vec{r}}{dt}$ is simply your velocity! This dot product is easy to evaluate, and yields the answer to the question.

(There are of course many ways to solve this problem; which method you choose may depend on how your path is described. It is often easiest to simply insert the given expressions for x and y in terms of t directly into h, then differentiate the resulting function of a single variable, thus calculating the left-hand-side directly.)

2. *You are another ant on the same surface, moving on a path with $y = 3x$. How fast is the depth changing compared with x?*

This problem is asking for the derivative of h with respect to x *as you move along the path*; note that this is the total derivative $\frac{dh}{dx}$, not the partial derivative $\frac{\partial h}{\partial x}$, which would only be appropriate if y were constant along the path. So divide the master formula by dx to get

$$\frac{dh}{dx} = \vec{\nabla}h \cdot \frac{d\vec{r}}{dx}$$

then use what you know $(y = 3x)$ to relate the changes in x and y $(dy = 3\,dx)$, so that

$$d\vec{r} = dx\,\hat{\imath} + dy\,\hat{\jmath} = dx\,\hat{\imath} + 3\,dx\,\hat{\jmath} = (\hat{\imath} + 3\,\hat{\jmath})\,dx$$

to obtain

$$\frac{d\vec{r}}{dx} = \hat{\imath} + 3\,\hat{\jmath}$$

Evaluating the dot product yields the answer to the question.

3. *You are still moving on the same surface, but now the question is how fast is the depth changing per unit distance along the path?*

This problem is asking for the derivative of h with respect to arclength ds. We can divide the master formula by ds, which leads to
$$\frac{dh}{ds} = \vec{\nabla}h \cdot \frac{d\vec{r}}{ds}$$
Unfortunately, it is often difficult to determine s; it is not always possible to express h as a function of s. On the other hand, all we need to know is that

$$ds = |d\vec{r}|$$

so that dividing $d\vec{r}$ by ds is just dividing by its length; the result must be a unit vector! Which unit vector? The one tangent to your path, namely the unit tangent vector $\hat{\boldsymbol{T}}$, so

$$\frac{dh}{ds} = \vec{\nabla}h \cdot \hat{\boldsymbol{T}} \tag{3.2}$$

Evaluating the dot product answers the question, without ever worrying about arclength.

12

Directional Derivatives

We have just seen that the derivative of f along a curve splits into two parts: a derivative of f (namely $\vec{\nabla} f$), and a derivative of the curve ($d\vec{r}/du$). But the latter depends only on the tangent direction of the curve at the given point, not on the detailed shape of the curve. This leads us to the concept of the *directional derivative* of f at a particular point $\vec{r} = \vec{r}_0 = \vec{r}(u_0)$ *along* the vector $\vec{v}$, which is traditionally defined as follows: *

$$D_{\vec{v}} f = \lim_{\epsilon \to 0} \frac{f(\vec{r}_0 + \epsilon \vec{v}) - f(\vec{r}_0)}{\epsilon}$$

According to the above discussion, this is just df/du along the tangent line, which is

$$D_{\vec{v}} f = \vec{\nabla} f \cdot \vec{v}$$

In Example 3 above, the left-hand-side of (3.2) is just the directional derivative of h in the direction $\hat{T}$, and could have been denoted by $D_{\hat{T}} h$.

What does the gradient mean geometrically? Let $\hat{w}$ be a unit vector, and consider

$$D_{\hat{w}} f = \vec{\nabla} f \cdot \hat{w} = |\vec{\nabla} f|\, |\hat{w}| \cos \theta = |\vec{\nabla} f| \cos \theta$$

which is clearly maximized by $\theta = 0$. Thus, the *direction* of $\vec{\nabla} f$ is just the direction in which f increases the fastest, and the *magnitude* of $\vec{\nabla} f$ is the rate of increase of f in that direction.

You can also visualize the gradient using the *level surfaces* on which $f(x, y, z) = $ const. (In 2-d there is the analogous concept of *level curves*, on which $f(x, y) = $ const.) Let $\vec{r} = \vec{r}(u)$ be any curve contained in a level surface of f, so that

$$f\left(\vec{r}(u)\right) = \text{const}$$

But then

$$\vec{\nabla} f \cdot \vec{v} = \frac{df}{du} = 0$$

Therefore, $\vec{\nabla} f$ is perpendicular to $\vec{v}$, and hence to every curve in the level surface. In other words, $\vec{\nabla} f$ is perpendicular to the level surfaces of f:

$$\vec{\nabla} f \perp \{f(x, y, z) = \text{const}\}$$

Other Coordinates

The master formula can be used to derive formulas for the gradient in other coordinate systems. We illustrate the method for polar coordinates.

In polar coordinates, we have

$$df = \frac{\partial f}{\partial r}\, dr + \frac{\partial f}{\partial \phi}\, d\phi$$

---

*It is often assumed that $\vec{v}$ is a unit vector, although this is not necessary.

and of course
$$d\vec{r} = dr\,\hat{\boldsymbol{r}} + r\,d\phi\,\hat{\boldsymbol{\phi}}$$
which is (2.3). Comparing these expressions with (3.1), we see immediately that we must have
$$\vec{\nabla}f = \frac{\partial f}{\partial r}\,\hat{\boldsymbol{r}} + \frac{1}{r}\frac{\partial f}{\partial \phi}\,\hat{\boldsymbol{\phi}} \tag{3.3}$$

Note the factor of $\frac{1}{r}$, which is needed to compensate for the factor of r in (2.3). Such factors are typical for the component expressions of vector derivatives in curvilinear coordinates.

Why would one want to compute the gradient in polar coordinates? Consider the computation of $\vec{\nabla}\left(\ln\sqrt{x^2+y^2}\right)$, which can done by brute force in rectangular coordinates; the calculation is straightforward but messy, even if you first use the properties of logarithms to remove the square root. Alternatively, using (3.3), it follows immediately that

$$\vec{\nabla}\left(\ln\sqrt{x^2+y^2}\right) = \vec{\nabla}\left(\ln r\right) = \frac{1}{r}\,\hat{\boldsymbol{r}}$$

GOALS

- Understand gradients and directional derivatives.

- Know the master formula.

Lesson 4:   Vector Fields

It is time to distinguish between several different vector-like objects. The arrow pointing from the origin to the point with (Cartesian) coordinates (a, b, c) is

$$\vec{w} = a\,\hat{\imath} + b\,\hat{\jmath} + c\,\hat{k}$$

This is a *vector*, which is said to have its *tail at the origin*, or to *live at the origin*. There is nothing special about the origin; vectors can live at any point.

A *vector function* (of one variable)

$$\vec{w}(u) = f(u)\,\hat{\imath} + g(u)\,\hat{\jmath} + h(u)\,\hat{k}$$

is a vector for each value $u = u_0$ of its argument. One example is a *parametric curve* $\vec{r}(u)$. We often identify the *position vector* $\vec{r} = x\,\hat{\imath} + y\,\hat{\jmath} + z\,\hat{k}$ with the *point* (x, y, z), and similarly for the corresponding vector function $\vec{r}(u)$. For example, we sometimes write $f(\vec{r})$ rather than $f(x, y, z)$. A parametric curve is a special vector function, in that it always lives at the origin. Many vector functions live instead on a parametric curve, such as the velocity $\vec{v}(u)$, which, for each value of u, lives at the point $\vec{r}(u)$ on the curve.
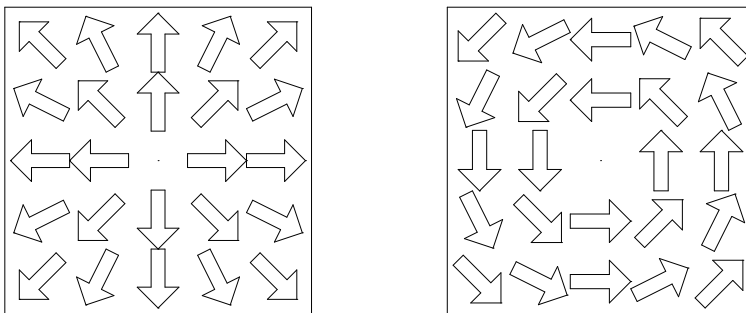
Finally, a *vector field*

$$\vec{F}(x, y, z) = F_x(x, y, z)\,\hat{\imath} + F_y(x, y, z)\,\hat{\jmath} + F_z(x, y, z)\,\hat{k}$$

assigns a vector $\vec{F}(x_0, y_0, z_0)$ to each point (x_0, y_0, z_0), which lives at (x_0, y_0, z_0). [*] An example of a vector field is the gradient of a function, $\vec{\nabla} f$. It is common practice to omit the explicit functional dependence, and write simply

$$\vec{F} = F_x\,\hat{\imath} + F_y\,\hat{\jmath} + F_z\,\hat{k}$$

as it is usually clear from the context what type of object $\vec{F}$ is. But be careful; the *vector* $\hat{\imath}$ (a single arrow at the origin) is different from the *constant vector field* $\hat{\imath}$ (an arrow at each point, all pointing in the x-direction).

In some settings, other bases are more appropriate. We have already introduced $\hat{r}$ and $\hat{\phi}$, the unit vectors pointing in the directions of increasing r and ϕ, respectively. *These are vector fields!* If desired, one can work out explicit formulas for these vector fields in terms of $\hat{\imath}$ and $\hat{\jmath}$, but these are not needed to draw pictures of these vector fields, which are shown below.

---

[*]In two dimensions, of course, $F_z = 0$, and $F_x = F_x(x, y)$, $F_y = F_y(x, y)$.

It is obvious geometrically that
$$\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{\phi}} = 0$$
at every point, so that $\{\hat{\boldsymbol{r}}, \hat{\boldsymbol{\phi}}\}$ is an orthonormal basis for vector fields in the plane. * We can thus expand a vector field $\vec{\boldsymbol{F}}$ in the plane with respect to either basis, yielding

$$\begin{aligned} \vec{\boldsymbol{F}} &= F_x\,\hat{\boldsymbol{\imath}} + F_y\,\hat{\boldsymbol{\jmath}} \\ &= F_r\,\hat{\boldsymbol{r}} + F_\theta\,\hat{\boldsymbol{\phi}} \end{aligned}$$

where F_r denotes the *r-component* of $\vec{\boldsymbol{F}}$, etc. Similar expressions hold for cylindrical and spherical coordinates.

Most mathematics textbooks always use the Cartesian (also called rectangular) basis $\{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$, which has the advantage of establishing a single formalism for all situations. But problems involving special symmetry, such as cylindrical or spherical symmetry, are usually much simpler if the computation is done in an adapted basis, such as those appropriate for cylindrical and spherical coordinates. Remarkably, essentially all of undergraduate engineering and physics can be handled by mastering these special cases! For this reason, engineering and physics texts typically emphasize the use of an adapted basis.

The bottom line is that all problems in this course *can* be solved using only rectangular basis vectors, but that many problems, both now and in your future coursework, are likely to be much easier if you learn how to work with other basis vectors. Ideally, you should learn both approaches, and know how they are related.

GOALS

- Know what a vector field is.

- Be able to sketch simple vector fields

- Be familiar with the basis vectors adapted to polar, cylindrical, and spherical symmetry.

---

*It is also *right-handed*, in the sense that $\{\hat{\boldsymbol{r}}, \hat{\boldsymbol{\phi}}, \hat{\boldsymbol{k}}\}$ is a right-handed basis in three dimensions.

Lesson 5:   Line Integrals

If you want to add up something along a curve, you need to compute a *line integral.* Common examples are the length of a curve, how much a wire weighs, and how much work is done when moving along a particular path.

We start with the last of these, namely finding the *work* W done by a force $\vec{F}$ in moving a particle along a curve C. We begin with the relationship

$$\text{work} = \text{force} \times \text{distance}$$

Suppose you take a small step $d\vec{r}$ along the curve. How much work was done? Since only the component along the curve matters, we need to take the dot product of $\vec{F}$ with $d\vec{r}$. Adding this up along the curve yields

$$W = \int_C \vec{F} \cdot d\vec{r}$$

So how do you evaluate such an integral? If you have an explicit parameterization $\vec{r} = \vec{r}(u)$, you can differentiate it in order to determine

$$d\vec{r} = \frac{d\vec{r}}{du}\, du \tag{5.1}$$

thus turning your line integral into an ordinary integral with respect to u.

This is essentially the strategy used in most texts: the distance you go is $ds = |d\vec{r}|$, the direction you are going is given by the unit tangent vector $\hat{T}$, the component of $\vec{F}$ in your direction is $\vec{F} \cdot \hat{T}$, and so the work done when moving a small distance is $\vec{F} \cdot \hat{T}\, ds$. But

$$\hat{T}\, ds = \frac{d\vec{r}}{ds}\, ds = d\vec{r}$$

so you wind up with the same integral as before. This is nice in theory, but in practice it is not always easy to write things explicitly in terms of the arclength s; it's usually easier to work with some other parameterization, which brings us back to (5.1).

But there is another strategy for doing line integrals, especially if you are not handed an explicit parameterization of the curve. Recall from (2.1) that the vector differential $d\vec{r}$ is given by

$$d\vec{r} = dx\,\hat{\imath} + dy\,\hat{\jmath} + dz\,\hat{k}$$

If you like, you can think of this as taking the differential of both sides of the expression for the position vector (in rectangular coordinates!), namely

$$\vec{r} = x\,\hat{\imath} + y\,\hat{\jmath} + z\,\hat{k}$$

But what if you don't have an explicit parameterization? *Use what you know!*

Suppose you want to find the work done by the force $\vec{F} = x\,\hat{\imath} + y\,\hat{\jmath}$ when moving along a given curve C. Curves can be specified in several different ways; let us consider some examples, all of which start at $(1,0)$ and end at $(0,1)$.

1. Consider first the parametric curve $\vec{r} = (1-u^2)\,\hat{\imath} + u\,\hat{\jmath}$. It is straightforward to compute

$$d\vec{r} = (-2u\,\hat{\imath} + \hat{\jmath})\,du$$

   Since a curve is described by a single parameter, in this case u, we write everything in terms of that parameter, obtaining

$$\int_C \vec{F} \cdot d\vec{r} = \int_0^1 \left((1-u^2)\,\hat{\imath} + u\,\hat{\jmath}\right) \cdot (-2u\,\hat{\imath} + \hat{\jmath})\,du$$
$$= \int_0^1 \left((1-u^2)(-2u) + u\right)du = \int_0^1 (2u^3 - u)\,du = 0$$

2. In physical applications, one is rarely given an explicit parameterization of the curve, but rather some other description. For instance, the curve just discussed might have been defined by the equation $x = 1 - y^2$. Finding the differential of both sides of this expression yields $dx = -2y\,dy$, and substituting into (2.1) leads to

$$d\vec{r} = (-2y\,\hat{\imath} + \hat{\jmath})\,dy$$

   The computation is exactly the same as before, using y instead of u.

3. Alternatively, one can solve for y, obtaining $y = \sqrt{1-x}$, then compute dy in terms of dx, then substitute into (2.1), obtaining

$$d\vec{r} = dx\,\hat{\imath} - \frac{dx}{2\sqrt{1-x}}\,\hat{\jmath}$$

   so that *

$$\int_C \vec{F} \cdot d\vec{r} = \int_1^0 \left(x - \frac{1}{2}\right)dx = 0$$

   It is important to realize that all of these methods work.

   Consider now the problem of trying to find the length of a quarter of a circle. What do you know? In polar coordinates, a circle is given by $r = $ constant, so that $dr = 0$. Inserting this fact into the expression (2.4) for arclength in polar coordinates, one immediately obtains

$$ds^2 = r^2\,d\phi^2$$

and finally

$$\int_C ds = \int_0^{\frac{\pi}{2}} r\,d\phi = \frac{\pi r}{2} \tag{5.2}$$

---

*Note that y *decreases* along the curve.

But what if you didn't remember (2.4)? The calculation is not much harder in rectangular coordinates: You know that $x = r \cos \phi$ and $y = r \sin \phi$, with $r = $ constant, so that $dx = -r \sin \phi \, d\phi$ and $dy = r \cos \phi \, d\phi$. Inserting this into (2.2) again leads to (5.2).

But what if you didn't even remember how to parameterize a circle, or, equivalently, how to use polar coordinates? Well, you still know that $x^2 + y^2 = r^2 = $ constant, so that $2x \, dx + 2y \, dy = 0$. Solving for dy and inserting this into (2.2) yields

$$ds^2 = \left(1 + \frac{x^2}{y^2}\right) dx^2 = \frac{r^2}{r^2 - x^2} dx^2$$

This leads to the (improper!) integral

$$\int_0^r \frac{dx}{\sqrt{1 - \frac{x^2}{r^2}}}$$

which is, of course, most easily computed via a trig substitution — or numerically — yielding the same answer.

The point is that using what you know will always yield correct answers — eventually.

A similar procedure can be used to determine the mass of a wire in the shape of the curve C if one knows the density ρ. In this case, the mass of a small piece of the wire is $\rho \, ds$, so the integral becomes

$$m = \int_C \rho \, ds$$

Another standard application of this type of line integral is to find the *center of mass* of a wire. This is done by averaging the values of the coordinates, weighted by the density ρ as follows:

$$\overline{x} = \frac{1}{m} \int_C x \rho \, ds$$

where m was defined earlier in the lesson. Similar formulas hold for $\overline{y}$ and $\overline{z}$; the center of mass is then the point $(\overline{x}, \overline{y}, \overline{z})$.

In a "use what you know" strategy, you may not be sure when to stop! A common error is to substitute for dy in terms of dx in $d\vec{r}$, but to forget to substitute for y in terms of x in $\vec{F}$. The rule of thumb is that you shouldn't start integrating until you have the integral in terms of a single parameter — including correctly determining the limits in terms of that parameter. Curves are one-dimensional!

Here is another example. Suppose you want to integrate $\int_C \vec{F} \cdot d\vec{r}$, where $\vec{F} = y \, \hat{\imath}$, and C is the line segment from $(1, 0)$ to $(0, -1)$. Start with expression (2.1) for $d\vec{r}$. What do you know? Well, the slope of the line segment is clearly $+1$, and its y-intercept is -1, so the equation of the line is $y = x - 1$. Taking the differential of both sides, $dy = dx$, so that $d\vec{r} = dx \, \hat{\imath} + dx \, \hat{\jmath}$. Finally, note that x runs from 1 to 0! Thus,

$$\int_C \vec{F} \cdot d\vec{r} = \int_1^0 (x - 1) \, \hat{\imath} \cdot (\hat{\imath} + \hat{\jmath}) \, dx = \int_1^0 (x - 1) \, dx = \left(\frac{x^2}{2} - x\right) \Big|_1^0 = \frac{1}{2}$$

19

Yes, of course, this particular curve is easy to parameterize, but this is not always the case. Note also how easy it was to get the limits right, and thus get the correct sign, simply by *always* starting with expression (2.1) for $d\vec{r}$, and integrating from your starting point to your final point.

This is important: Vector line integrals of the form $\displaystyle\int_C \vec{F} \cdot d\vec{r}$ are *directed* integrals; the sign of the answer depends on which way you traverse the curve. You will obtain the correct sign automatically if you integrate from the beginning point to the final point, without putting in any artificial signs. As in the last example, this may result in an integral which goes from a larger value of the integration variable to a smaller one.

This is *not* the case for line integrals with respect to arclength. Since $ds = |d\vec{r}|$, such integrals do not depend on which way the curve is traversed. Standard examples are arclength and mass, which must be positive! Other times you may have to think about it to get the sign right; the total charge on a wire could be positive or negative. One way to keep this straight is to remember that $ds = |d\vec{r}|$, which has an absolute value in it, which requires care with signs.

We summarize this discussion by writing

$$\int_{-C} f\, ds = + \int_{C} f\, ds \qquad \text{but} \qquad \int_{-C} \vec{F} \cdot d\vec{r} = - \int_{C} \vec{F} \cdot d\vec{r}$$

where $-C$ denotes the reversed curve. You will always get the sign right for vector line integrals if you are careful to put the limits in correctly (from starting point to ending point), whereas you may have to think about signs for the scalar line integrals.

GOALS

- Know the different kinds of line integrals and how to evaluate them.

Lesson 6:   Independence of Path

RECALL: $$\int_a^b \frac{df}{dx}\, dx = f\Big|_a^b$$

This is the *Fundamental Theorem of Calculus*, which just says that the integral of a derivative is the function you started with. We could also write this simply as

$$\int df = f$$

But recall the master formula (3.1), which says

$$df = \vec{\nabla} f \cdot d\vec{r}$$

Putting this all together, we get the *fundamental theorem for line integrals*, which says that

$$\int_C \vec{\nabla} f \cdot d\vec{r} = f\Big|_A^B$$

for any curve C starting at some point A and ending at some point B.

Notice that the right-hand-side does not depend on the curve C! * This behavior leads us directly to the notion of path independence. A line integral of the form $\int_C \vec{F} \cdot d\vec{r}$ is said to be *independent of path* if its value depends only on the endpoints A and B of the curve C, not on the particular curve connecting them. If a line integral is independent of path, we no longer need to specify the path C, so we instead write

$$\int_A^B \vec{F} \cdot d\vec{r}$$

If you know that a line integral is independent of path, you may *choose* a different path (with the same endpoints) which makes evaluating the integral as simple as possible!

The fundamental theorem implies that vector fields of the form $\vec{F} = \vec{\nabla} f$ are special; the corresponding line integrals are always independent of path. One way to think of this is to imagine the level curves of f; the change in f depends only on where you start and end, not on how you get there. These special vector fields have a name: A vector field $\vec{F}$ is said to be *conservative* if there exists a *potential function* f such that $\vec{F} = \vec{\nabla} f$.
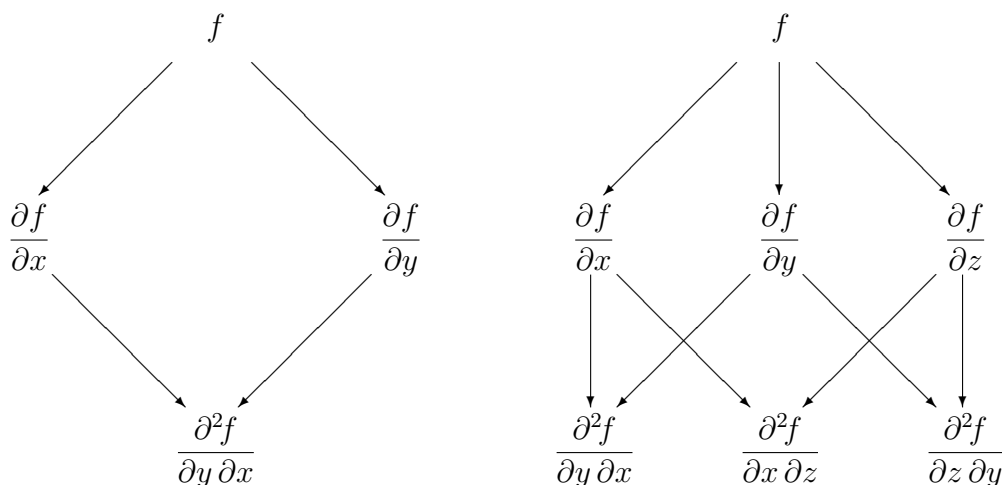
If $\vec{F}$ is conservative, then $\int_C \vec{F} \cdot d\vec{r}$ is independent of path; the converse is also true. But how do you know if a given vector field $\vec{F}$ is conservative? That's the next lesson.

GOALS

- Understand what it means for a line integral to be independent of path, and for a vector field to be conservative.

---

*There is some fine print here: the curve must lie in a connected region (one with no holes), on which $\vec{\nabla} f$ is defined everywhere. This is no problem if f is differentiable everywhere, which it often is, but there are important examples where this fails.

Lesson 7:   Finding Potential Functions

$$f$$

$$\frac{\partial f}{\partial x} \qquad\qquad \frac{\partial f}{\partial y}$$

$$\frac{\partial^2 f}{\partial y\,\partial x}$$

$$f$$

$$\frac{\partial f}{\partial x} \qquad \frac{\partial f}{\partial y} \qquad \frac{\partial f}{\partial z}$$

$$\frac{\partial^2 f}{\partial y\,\partial x} \qquad \frac{\partial^2 f}{\partial x\,\partial z} \qquad \frac{\partial^2 f}{\partial z\,\partial y}$$

We describe here a variation of the usual procedure for determining whether a vector field is conservative and, if it is, for finding a potential function.

It is helpful to make a diagram of the structure underlying potential functions and conservative vector fields. For functions of two variables, this is shown in the first diagram above. The potential function f is shown at the top. Slanted lines represent derivatives of f; derivatives with respect to x go to the left, while derivatives with respect to y go to the right. The second line thus gives the components of $\vec{\nabla} f$. The bottom line shows the mixed second derivatives, which can of course be computed in either order.

But what if we are not given f? Suppose we are given a vector field, such as
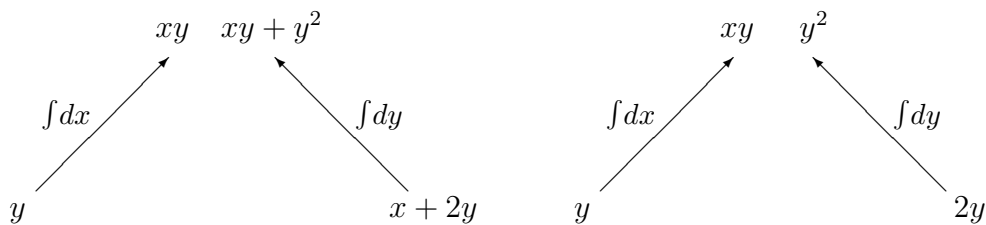
$$\vec{F} = y\hat{\imath} + (x + 2y)\,\hat{\jmath}$$

and need to determine whether it is conservative, that is, whether $\vec{F}$ is the gradient of some potential function f. This is the second line of the diagram! We could start by checking that the mixed derivatives agree. However, what we really want is the potential function; we should be moving up the diagram, not down. What happens if we simply integrate both components, as shown in the first diagram below? The potential function is clearly contained in the results of these two integrals; it is just a question of combining them correctly.

Furthermore, there is enough information here to determine whether $\vec{F}$ is conservative in the first place; there is no need to check the derivatives. For example, had we been given the vector field

$$\vec{H} = y\,\hat{\imath} + 2y\,\hat{\jmath}$$

and integrated its components, we would obtain the second diagram below. Simply by noticing that xy, a function of two variables, only occurs once, we see that $\vec{H}$ is not conservative.

$$xy \quad xy + y^2$$

$$\int dx \qquad\qquad \int dy$$

$$y \qquad\qquad\qquad x + 2y$$

$$xy \quad y^2$$

$$\int dx \qquad\qquad \int dy$$

$$y \qquad\qquad\qquad 2y$$

We describe this as a *murder mystery*. A crime has been committed by the unknown murderer f; your job is to find the identity of f by interviewing the witnesses. Who are the witnesses? The components of the vector field. What do they tell you? Well, you have to integrate ("interrogate") them! Now for the fun part.

If two witnesses say they saw someone with red hair, that doesn't mean the suspect has two red hairs! So if you get the same clue more than once, you only count it once.

On the other hand, some clues require corroboration. These witnesses were situated in such a way that each could only look in one direction. Thus, one witness, the x-component, only sees terms involving x, etc. If a clue contains more than one variable, it should have been seen by more than one witness! In fact, functions of n variables should occur precisely n times. In the case of the vector field $\vec{H}$, the clue xy was only seen by one witness, not both; somebody is lying! In short, clues must be consistent.

Here is the *Murder Mystery Method* in a nutshell:

- *Integrate:* Integrate the x-component with respect to x, etc.

- *Check consistency:* Functions of n variables must occur exactly n times. *
  (If the consistency check fails, the vector field is not conservative.)

- *Combine clues:* Use each clue once to determine the potential function.

The power of the murder mystery method is even more apparent in three dimensions. We encourage you to try to find a potential function for the vector field $\vec{G}$ defined by

$$\vec{G} = yz\,\hat{\boldsymbol{\imath}} + (xz + z)\,\hat{\boldsymbol{\jmath}} + (xy + y + 2z)\,\hat{\boldsymbol{k}}$$

using this method. The underlying structure is shown in the second diagram at the top of the previous page, where now y derivatives are shown going straight down, and z derivatives go to the right.

Consistency is traditionally checked by computing the last line of the appropriate diagram at the top of the previous page. We reiterate that this is *not* necessary with the Murder Mystery Method. We will however return to this diagram later, when discussing the curl of a vector field.

GOALS

- Be able to find potential functions.

- Be able to evaluate path-independent line integrals using potential functions.

---

*Checking consistency may not be straightforward. Are $-\cos 2x \cos^2 y$ and $\sin^2 x \cos 2y$ the same? We have $2\sin^2 x \cos^2 y = -\cos 2x \cos^2 y + \cos^2 y = \sin^2 x \cos 2y + \sin^2 x$, which shows that the "$xy$" parts of these functions agree. (If this cannot be done, consistency fails.) But how do we count the remaining functions of 1 variable? Recall that the x-witness can only reliably provide clues involving x! If rewriting one or more functions this way generates a new such term, count it as usual. If instead this generates a term involving the wrong variable(s), such as a clue provided by the x-witness which does not involve x, discard it as unsubstantiated; this amounts to allowing the constant of integration to depend on the other variables. This procedure does not depend on the manner in which particular functions are rewritten, and such manipulations are not usually needed for typical classroom examples.

INTERLUDE
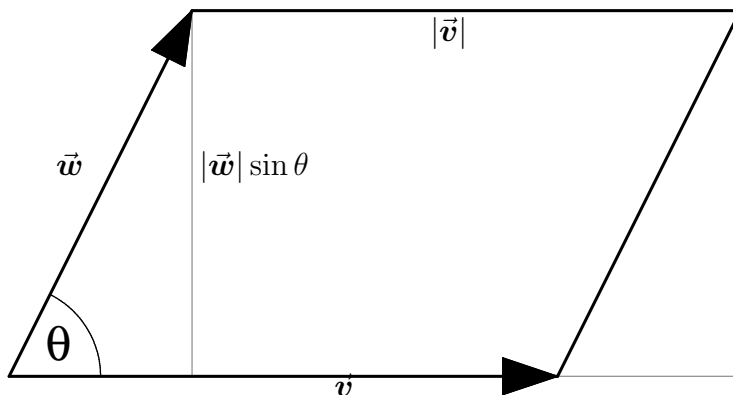
Cross Product

The cross product is fundamentally a directed area. The *magnitude* of the cross product is defined to be the area of the parallelogram, as shown in the figure above. This leads to the formula

$$|\vec{v} \times \vec{w}| = |\vec{v}||\vec{w}| \sin \theta$$

an immediate consequence of which is that

$$\vec{v} \parallel \vec{w} \iff \vec{v} \times \vec{w} = \vec{0}$$

The *direction* of the cross product is given by the right-hand rule, so that in the example shown $\vec{v} \times \vec{w}$ points out of the page. This implies that
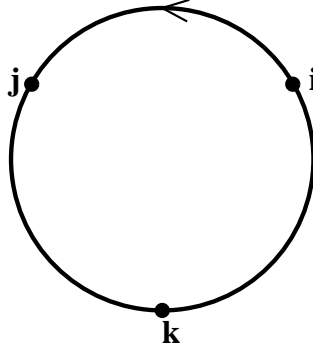
$$\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$$

so that the cross product is not commutative. * An important property of the cross product is that

$$\vec{v} \times \vec{v} = \vec{0}$$

which follows from any of the preceding three equations.

---

*The cross product also fails to be associative, since for example $\hat{\imath} \times (\hat{\imath} \times \hat{\jmath}) = -\hat{\jmath}$ but $(\hat{\imath} \times \hat{\imath}) \times \hat{\jmath} = \vec{0}$.

In terms of the standard orthonormal basis, the geometric formula quickly yields

$$\hat{\imath} \times \hat{\jmath} = \hat{k}$$
$$\hat{\jmath} \times \hat{k} = \hat{\imath}$$
$$\hat{k} \times \hat{\imath} = \hat{\jmath}$$

This cyclic nature of the cross product can be emphasized by abbreviating this multiplication table as shown in the figure at the top of the page. * Products in the direction of the arrow get a plus sign; products against the arrow get a minus sign.

Using an orthonormal basis such as $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$, the geometric formula reduces to the standard component form of the cross product. If $\vec{v} = v_x \hat{\imath} + v_y \hat{\jmath} + v_z \hat{k}$ and $\vec{w} = w_x \hat{\imath} + w_y \hat{\jmath} + w_z \hat{k}$, then

$$
\begin{aligned}
\vec{v} \times \vec{w} &= (v_x \hat{\imath} + v_y \hat{\jmath} + v_z \hat{k}) \times (w_x \hat{\imath} + w_y \hat{\jmath} + w_z \hat{k}) \\
&= (v_y w_z - v_z w_y) \hat{\imath} + (v_z w_x - v_x w_z) \hat{\jmath} + (v_x w_y - v_y w_x) \hat{k}
\end{aligned}
\tag{I.1}
$$

which is often written as the symbolic determinant

$$
\vec{v} \times \vec{w} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix}
\tag{I.2}
$$

We emphasize that this works in *any* (right-handed) orthonormal basis. In cylindrical coordinates, not only is

$$\hat{r} \times \hat{\phi} = \hat{z}$$

but cross products can be computed as

$$
\vec{v} \times \vec{w} = \begin{vmatrix} \hat{r} & \hat{\phi} & \hat{z} \\ v_r & v_\phi & v_z \\ w_r & w_\phi & w_z \end{vmatrix}
$$

where of course $\vec{v} = v_r \hat{r} + v_\phi \hat{\phi} + v_z \hat{z}$ and similarly for $\vec{w}$.

---

*This is really the multiplication table for the unit imaginary quaternions, a number system which generalizes the familiar complex numbers. Quaternions predate vector analysis, which borrowed the i, j, k notation for the rectangular basis vectors $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$. A more logical name would be $\hat{x}$, $\hat{y}$, $\hat{z}$, which is used by many physicists.

We encourage you to use (I.2), rather than simply memorizing (I.1). We also encourage you to compute the determinant as described below, rather than using minors; this tends to minimize sign errors. A 3×3 determinant can be computed in the form

$$\begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} \begin{matrix} \hat{\imath} & \hat{\jmath} \\ v_x & v_y \\ w_x & w_y \end{matrix}$$

where one multiplies the terms along each diagonal line, subtracting the products obtained along lines going down to the left from those along lines going down to the right. While this method works only for $(2 \times 2$ and$)$ 3×3 determinants, it emphasizes the cyclic nature of the cross product.

Another important skill is knowing when *not* to use a determinant at all. For simple cross products, such as $(\hat{\imath} + 3\,\hat{\jmath}) \times \hat{k}$, it is easier to use the multiplication table directly. It is also worth pointing out that the multiplication table and the determinant method generalize naturally to other orthonormal bases; all that is needed is to replace the rectangular basis $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ by the one being used (in the right order!).

A good problem emphasizing the geometry of the cross product is to find the area of the triangle formed by connecting the tips of the vectors $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$ (whose base is at the origin).

Lesson 8:   Surfaces

There are many ways to describe a surface. Consider the following following descriptions:

- the unit sphere;
- $x^2 + y^2 + z^2 = 1$;
- $r = 1$ (where r is the *spherical* radial coordinate);
- $x = \sin\theta\cos\phi$, $y = \sin\theta\sin\phi$, $z = \cos\theta$, with $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$;
- $\vec{r} = x\,\hat{\imath} + y\,\hat{\jmath} + z\,\hat{k}$, with x, y, z given as above;

all of which describe the same surface. Which representation is best for a given problem depends on the circumstances.

The simplest surfaces are those given by holding one of the coordinates constant. Thus, the xy-plane is given by $z = 0$. Its (surface) area element is $dA = (dx)(dy) = (dr)(r\,d\phi)$, as can easily be seen by drawing the appropriate small rectangle. The surface of a cylinder is nearly as easy, as it is given by $r = a$ in cylindrical coordinates, and drawing a small "rectangle" yields for the surface element

$$\text{cylinder}: \qquad dA = (a\,d\phi)(dz) = a\,d\phi\,dz$$

while a similar construction for the sphere given by $r = a$ in spherical coordinates yields

$$\text{sphere}: \qquad dA = (a\,d\theta)(a\sin\theta\,d\phi) = a^2 \sin\theta\,d\theta\,d\phi$$

The last expression can of course be used to compute the surface area of a sphere, which is *

$$\int\limits_{\text{sphere}} dA = \int_0^{2\pi} \int_0^\pi a^2 \sin\theta\,d\theta\,d\phi = 4\pi a^2$$

What about more complicated surfaces?

The basic building block for surface integrals is the infinitesimal area dA, obtained by chopping up the surface into small pieces. If the pieces are small parallelograms, then the area can be determined by taking the cross product of the sides!

Chopping up the surface is usually done by drawing two families of curves on the surface; surfaces are 2-dimensional. So compute $d\vec{r}$ on each family, take the cross product, and get the *vector surface element* in the form

$$d\vec{A} = d\vec{r}_1 \times d\vec{r}_2 \tag{8.1}$$

In order to determine the surface area, we need the magnitude of this expression, which is

$$dA = |d\vec{A}| = |d\vec{r}_1 \times d\vec{r}_2| \tag{8.2}$$

and which is called the *(scalar) surface element.* This should remind you of the corresponding expression for line integrals, namely $ds = |d\vec{r}|$.

---

*We write a single integral sign when talking about adding up "bits of area" (or "bits of volume"), reserving multiple integral signs for iterated single integrals. The notation $\iint dA$ is also common.

We illustrate this technique by computing the surface element for the paraboloid given by $z = x^2 + y^2$. We start with the basic formula for the vector differential $d\vec{r}$, namely

$$d\vec{r} = dx\,\hat{\imath} + dy\,\hat{\jmath} + dz\,\hat{k}$$

What do you know? The expression for z leads to

$$dz = 2x\,dx + 2y\,dy$$

In rectangular coordinates, it is natural to consider infinitesimal displacements in the x and y directions. In the x direction, y is constant, so $dy = 0$, and we obtain

$$d\vec{r}_1 = dx\,\hat{\imath} + 2x\,dx\,\hat{k}$$

Similarly, in the y direction, $dx = 0$, which leads to

$$d\vec{r}_2 = dy\,\hat{\jmath} + 2y\,dy\,\hat{k}$$

Putting this together, we obtain

$$d\vec{A} = d\vec{r}_1 \times d\vec{r}_2 = (\hat{\imath} + 2x\,\hat{k}) \times (\hat{\jmath} + 2y\,\hat{k})\,dx\,dy = (-2x\,\hat{\imath} - 2y\,\hat{\jmath} + \hat{k})\,dx\,dy$$

for the vector surface element, and

$$dA = \left| -2x\,\hat{\imath} - 2y\,\hat{\jmath} + \hat{k} \right|\,dx\,dy = \sqrt{1 + 4x^2 + 4y^2}\,dx\,dy$$

for the scalar surface element. Can you repeat this computation in cylindrical coordinates?

This construction emphasizes that "area" is really a *vector*, whose direction is perpendicular to the surface, and whose magnitude is the area. Note that there are always *two* choices for the direction; choosing one determines the *orientation* of the surface.

When using (8.1 and (8.2), it doesn't matter how you chop up the surface. It is of course possible to get the opposite orientation, for instance by interchanging the roles of $d\vec{r}_1$ and $d\vec{r}_2$. Rather than worrying too much about getting the "right" orientation from the beginning, it is usually simpler to check at the end whether the one you've got agrees with the requirements of the problem. If not, insert a minus sign.

Just as a curve is a 1-dimensional set of points, a surface is a 2-dimensional. When computing line integrals, it was necessary to write everything in terms of a single parameter before integrating. Similarly, for surface integrals one must write everything in terms of *two* parameters before starting to integrate.

A common technique, which we now discuss, is to introduce these two parameters at the beginning, thus representing the surface as a *parametric surface*. However, the expression (8.1) for $d\vec{A}$ makes no reference to the parameters! In fact, $d\vec{A}$ can be computed using *any* two (independent) infinitesimal displacements in the surface; no parameterization is needed! You will have to decide for yourself which approach you prefer. But real-world problems rarely come equipped with parameterizations, unlike those found in calculus texts.

Recall that a parametric curve can be written

$$\vec{r} = \vec{r}(u) = x(u)\,\hat{\imath} + y(u)\,\hat{\jmath} + z(u)\,\hat{k}$$

together with an appropriate domain for the parameter u. A *parametric surface* is similar, except there are now two parameters u,v (and an appropriate domain):

$$\vec{r} = \vec{r}(u, v) = x(u, v)\,\hat{\imath} + y(u, v)\,\hat{\jmath} + z(u, v)\,\hat{k}$$

Here are some examples of parametric surfaces (without domains):

$$
\begin{aligned}
\text{sphere}: \quad & \vec{r}(\theta, \phi) = a\sin\theta\cos\phi\,\hat{\imath} + a\sin\theta\sin\phi\,\hat{\jmath} + a\cos\theta\,\hat{k} \\
\text{cylinder}: \quad & \vec{r}(\phi, z) = a\cos\phi\,\hat{\imath} + a\sin\phi\,\hat{\jmath} + z\,\hat{k} \\
z = f(x, y): \quad & \vec{r}(x, y) = x\,\hat{\imath} + y\,\hat{\jmath} + f(x, y)\,\hat{k} \\
z = f(r, \phi): \quad & \vec{r}(r, \phi) = r\cos\phi\,\hat{\imath} + r\sin\phi\,\hat{\jmath} + f(r, \phi)\,\hat{k} \\
\text{surface of revolution}: \quad & \vec{r}(x, \phi) = x\,\hat{\imath} + f(x)\cos\phi\,\hat{\jmath} + f(x)\sin\phi\,\hat{k} \\
\text{(2-d) change of variables}: \quad & \vec{r}(u, v) = x(u, v)\,\hat{\imath} + y(u, v)\,\hat{\jmath}
\end{aligned}
$$

How do you find the surface element of a parametric surface? Use curves with u or v constant to chop up the surface. The sides of a small parallelogram are then

$$d\vec{r}_1 = \frac{\partial \vec{r}}{\partial u}\,du \qquad d\vec{r}_2 = \frac{\partial \vec{r}}{\partial v}\,dv$$

so that

$$d\vec{A} = d\vec{r}_1 \times d\vec{r}_2 = \left(\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right) du\,dv$$

and

$$dA = |d\vec{A}| = \left|\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}\right| du\,dv$$

We reiterate, however, that it is not always necessary to have an explicit parameterization in order to determine the surface element, as illustrated by the above computation for the paraboloid.

Here are some common surface elements:

$$
\begin{aligned}
\text{(plane)} \quad z = \text{const}: \quad & d\vec{A} = \pm\hat{k}\,dx\,dy \\
\text{(plane)} \quad z = \text{const}: \quad & d\vec{A} = \pm\hat{k}\,r\,dr\,d\phi \\
\text{(cylinder)} \quad r = \text{const}: \quad & d\vec{A} = \pm\hat{r}\,r\,d\phi\,dz \\
\text{(sphere)} \quad r = \text{const}: \quad & d\vec{A} = \pm\hat{r}\,r^2\sin\theta\,d\theta\,d\phi
\end{aligned}
$$

where we must distinguish between the the cylindrical unit vector $\hat{r} = (x\,\hat{\imath} + y\,\hat{\jmath})/r$ (with the *polar* coordinate r) and the spherical unit vector $\hat{r} = (x\,\hat{\imath} + y\,\hat{\jmath} + z\,\hat{k})/r$ (with the *spherical* coordinate r, sometimes written as ρ).

Finally, a word about notation. You will often see dS instead of dA, and $d\vec{S}$ instead of $d\vec{A}$; most authors use dA in the xy-plane.

GOALS

- Know how to compute surface elements.
- Be able to find the area of a surface.
- Be able to parameterize simple surfaces.

Lesson 9: Surface Integrals

The main results of the previous section were formulas for the surface elements $d\vec{A}$ and dA of a surface in terms of two infinitesimal displacements in the surface, that is

$$d\vec{A} = d\vec{r}_1 \times d\vec{r}_2$$

and

$$dA = |d\vec{A}| = |d\vec{r}_1 \times d\vec{r}_2|$$

which are (8.1 and 8.2), respectively. This allows us to integrate *any* function over a surface, by considering an integral of the form

$$\int_S f \, dA$$

The surface area of S is obtained as a special case by setting $f = 1$. Another example is the mass, which is obtained by setting $f = \rho$ where ρ denotes the density. The center of mass can be obtained from

$$\bar{x} = \frac{1}{m} \int_S x\rho \, dA$$

and similar formulas for $\bar{y}$ and $\bar{z}$; compare the analogous formulas for the mass and center of mass of a wire in terms of line integrals.

The *flux* of a vector field $\vec{F}$ across a surface S is the total amount of $\vec{F}$ pointing through S. An example which is easy to visualize is $\vec{F} = \rho \vec{v}$, where $\vec{v}$ is the velocity field of a fluid of density ρ. In this case, the flux measures how much of the fluid crosses the surface.

"Total amount" of course means to integrate, so we chop up our surface into small pieces. But how much of $\vec{F}$ "points through" a small piece of the surface? We only care about the component of $\vec{F}$ normal to the surface, so we need a dot product with a vector perpendicular to S. The flux through a small piece surely depends both on the size of $\vec{F}$ and the area of the piece. Putting this all together, the flux takes the form

$$\text{flux} = \int_S \vec{F} \cdot d\vec{A}$$

Again we see the fundamental role played by the *vector* surface element. Be careful to choose the correct orientation for S, depending on which flux is being computed: the flux up, the flux out, etc.

Another way to think about this is to first find the unit normal vector $\hat{n}$ to the surface, then use the dot product to find the normal component of $\vec{F}$, then integrate the resulting scalar field over the surface. These two viewpoints are related by

$$\hat{n} \, dA = d\vec{A}$$

If desired, you can compute $\hat{\boldsymbol{n}}$ itself from

$$\hat{\boldsymbol{n}} = \frac{d\vec{\boldsymbol{r}}_1 \times d\vec{\boldsymbol{r}}_2}{|d\vec{\boldsymbol{r}}_1 \times d\vec{\boldsymbol{r}}_2|} = \frac{\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}}{\left| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right|}$$

but this is not usually necessary. You should compare this formula with $\hat{\boldsymbol{T}} = \frac{d\vec{r}}{ds} = \frac{d\vec{r}}{du} \Big/ \frac{ds}{du}$.

GOALS

1. Know how to evaluate a surface integral of the form $\int\limits_{S} f \, dA$.

2. Know what the flux of a vector field through a surface is, and how to evaluate the corresponding surface integral.

Lesson 10: The Divergence Theorem

At any point P, we define the *divergence* of a vector field $\vec{F}$, which we write for now as $\text{div}\vec{F}$, to be the flux of $\vec{F}$ per unit volume "at" P, which is the (limiting value of the) flux per unit volume out of a small box around P.

Furthermore, we can chop up an arbitrary (closed) box into a suitable combination of (small) rectangular boxes. The outward fluxes through all interior sides will cancel, leaving just the flux out of the original box. Thus,

$$\int_{\text{box}} \vec{F} \cdot d\vec{A} = \int_{\text{inside}} \text{div}\vec{F} \ dV \tag{10.1}$$

This is the Divergence Theorem! But what is the divergence?

In order to find the divergence, we must compute the flux. Consider a (possibly large) rectangular box whose sides are parallel to the coordinate planes. What is the flux of $\vec{F}$ out of this box? Consider first the vertical contribution, namely the flux *up* through the top plus the flux *down* through the bottom. These two sides each have area $dA = dx\,dy$, so we get

$$\begin{aligned} \int_{\text{top+bottom}} \vec{F} \cdot d\vec{A} &= \iint_R \vec{F}\Big|_{\text{top}} \cdot \hat{k}\, dx\, dy - \iint_R \vec{F}\Big|_{\text{bottom}} \cdot \hat{k}\, dx\, dy \\ &= \iint_R \vec{F}\Big|_{\text{bottom}}^{\text{top}} \cdot \hat{k}\, dx\, dy \end{aligned}$$

where R denotes the common domain of the top and bottom faces in terms of x and y. Note the minus sign, which is really due to taking the dot product of $\vec{F}$ with $-\hat{k}$, which is the *outward* unit normal vector on the bottom face. But along any vertical line C from the bottom face to the top face, we have

$$\vec{F}\Big|_{\text{bottom}}^{\text{top}} \cdot \hat{k} = F_z\Big|_{\text{bottom}}^{\text{top}} = \int_C dF_z = \int_C \vec{\nabla}F_z \cdot d\vec{r} = \int_C \frac{\partial F_z}{\partial z}\, dz$$

where the second equality is the Fundamental Theorem of Calculus ("adding small changes yields the total change"), the next equality is the master formula (3.1), and the final equality follows since $d\vec{r} = dz\,\hat{k}$ for a vertical path.

Inserting this last expression into the one before results in

$$\int_{\text{top+bottom}} \vec{F} \cdot d\vec{A} = \iiint_V \frac{\partial F_z}{\partial z}\, dz\, dx\, dy$$

where V denotes the volume inside the box. Repeating this argument using the remaining pairs of faces, it follows that the total flux out of the box is

$$\text{total flux} = \int_{\text{box}} \vec{F} \cdot d\vec{A} = \int_V \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) dV$$

32

We can use this to compute the flux per unit volume through a small box around P. Suppose the box is small enough that the expression in parentheses is nearly constant. We then get

$$\text{total flux} \approx \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) \Big|_P (\text{volume of } V)$$

from which we see that the flux per unit volume, and hence the divergence, is given in rectangular coordinates by

$$\text{div}\vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \tag{10.2}$$

The formula (10.2) is often taken as the definition of the divergence, in which case the preceding calculation is used to prove the Divergence Theorem (10.1); we prefer to view (10.1) as the definition, from which (10.2) can be derived. This emphasizes the important fact that

$$\text{divergence} = \frac{\text{flux}}{\text{unit volume}}$$

For this reason, points where $\text{div}\vec{F} > 0$ are referred to as *sources*, and points where $\text{div}\vec{F} < 0$ are referred to as *sinks*.

It is often easier to work in coordinates adapted to the given problem. To use the Divergence Theorem in, say, spherical coordinates one must have an appropriate expression for the divergence. For example, a spherically symmetric, radial vector field of the form

$$\vec{H} = h(r)\,\hat{r} = h(r)\,\frac{\vec{r}}{r}$$

has divergence

$$\text{div}\vec{H} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 h(r) \right)$$

as can be checked by looking up the full formula for the divergence in spherical coordinates (e.g. in the book **Div, Grad, Curl, and All That**). The electric field of a point charge has this form. But where did this formula come from? You can derive it yourself by computing the flux of $\vec{H}$ out of a small box adapted to spherical coordinates, that is, whose sides are surfaces with one of the spherical coordinates held constant.

GOALS

- Be able to use the Divergence Theorem to evaluate integrals.

- Know the geometric interpretation of divergence.

Lesson 11:   Green's Theorem

The *circulation* of a vector field $\vec{F}$ around a closed loop C is simply the integral

$$\oint_C \vec{F} \cdot d\vec{r}$$

We will restrict ourselves in this lesson to the xy-plane; the circle on the integral sign then implies both that C is closed, and that the loop is oriented in a counterclockwise direction.

Consider a rectangular loop in the xy-plane, with sides parallel to the coordinate axes. What is the circulation of $\vec{F}$ around this loop? Consider first the two horizontal edges, on each of which $d\vec{r} = dx\,\hat{\imath}$, so that

$$
\begin{aligned}
\int_{\text{top+bottom}} \vec{F} \cdot d\vec{r} &= -\int_I \vec{F}\Big|_{\text{top}} \cdot \hat{\imath}\, dx + \int_I \vec{F}\Big|_{\text{bottom}} \cdot \hat{\imath}\, dx \\
&= -\int_I \vec{F}\Big|_{\text{bottom}}^{\text{top}} \cdot \hat{\imath}\, dx
\end{aligned}
$$

where I denotes the common domain of x along the top and bottom edges, and where the signs keep track of the counterclockwise orientation.

An argument similar to the one in the previous lesson shows that

$$\vec{F}\Big|_{\text{bottom}}^{\text{top}} \cdot \hat{\imath} = F_x\Big|_{\text{bottom}}^{\text{top}} = \int_{\text{bottom}}^{\text{top}} dF_x = \int_{\text{bottom}}^{\text{top}} \vec{\nabla} F_x \cdot d\vec{r} = \int_{\text{bottom}}^{\text{top}} \frac{\partial F_x}{\partial y}\, dy$$

so that

$$\int_{\text{top+bottom}} \vec{F} \cdot d\vec{r} = -\int\int_{\text{inside}} \frac{\partial F_x}{\partial y}\, dy\, dx$$

Repeating this argument for the remaining two sides, being careful with the orientation, leads to a similar term with x and y interchanged and with the opposite sign, so that the circulation around this loop is

$$\oint_{\text{loop}} \vec{F} \cdot d\vec{r} = \int_{\text{inside}} \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dA \tag{11.1}$$

This is Green's Theorem! We say more about the right-hand-side in the next lesson.

If $\vec{F}$ is conservative, then integrating $\vec{F}$ is independent of path, so that

$$\oint_C \vec{F} \cdot d\vec{r} = 0$$

over *any* simple closed curve C. In particular, in two dimensions we have

$$F_x\,\hat{\imath} + F_y\,\hat{\jmath} \text{ is conservative} \iff \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} = 0$$

which is often used as a test for conservative vector fields. Green's Theorem states that, even when $\vec{F}$ is *not* conservative, there is still a relationship between the circulation and the difference of the partial derivatives.

One of the most interesting consequences of Green's Theorem is that it can be used to compute the area of a region as a line integral along the boundary. If F_x and F_y are *any* functions such that

$$\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} = 1$$

then Green's Theorem says that

$$\int_C \vec{F} \cdot d\vec{r} = \int_{\text{inside}} dA = \text{area inside } C$$

Typical choices are

$$A = \oint_C x\,dy = -\oint_C y\,dx = \frac{1}{2}\oint_C x\,dy - y\,dx$$

It is therefore possible to determine the area of a region by walking around its boundary. Geographers use a gadget called a *planimeter* to do just that, determing the area of, say, Oregon by tracing its outline on a map. (Nowadays, of course, this is done by computer.)

GOALS

- Be able to use Green's Theorem to evaluate appropriate line integrals or double integrals.

- Be able to find area using Green's Theorem.

Lesson 12:   Stokes' Theorem

At any point P, we define the *curl* of a vector field $\vec{F}$, which we write for now as $\operatorname{curl}\vec{F}$, to be the vector which gives the circulation of $\vec{F}$ per unit area "at" P. That is, given any unit vector $\hat{n}$ at P, $(\operatorname{curl}\vec{F}) \cdot \hat{n}$ should be the (limiting value of the) circulation of $\vec{F}$ per unit area at P around a small loop with axis $\hat{n}$.

Suppose that the loop is in the xy-plane. But we know from (11.1) that

$$\text{circulation} = \oint_{\text{loop}} \vec{F} \cdot d\vec{r} = \int_{\text{inside}} \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dA$$

Assuming that the loop is small enough so that the expression in parentheses is nearly constant yields an expression for the circulation per unit area around the loop

$$\text{circulation} \approx \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \bigg|_{P} (\text{area inside loop})$$

from which we can compute the circulation per unit area. Since our loop has axis $\hat{k}$, this must be the $\hat{k}$-component of the curl, so we have

$$(\operatorname{curl}\vec{F}) \cdot \hat{k} = \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}$$

The rectangular expression for the full curl now follows by cyclic symmetry, yielding

$$\operatorname{curl}\vec{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \hat{\imath} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \hat{\jmath} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \hat{k}$$

You will learn an easier way to remember this formula in the next lesson. Remembering that in the xy-plane $d\vec{A} = \hat{k}\, dA$, it is not hard to see that the circulation around any rectangular loop takes the form

$$\oint_{\text{loop}} \vec{F} \cdot d\vec{r} = \int_{\text{inside}} \operatorname{curl}\vec{F} \cdot d\vec{A} \tag{12.1}$$

which of course reduces to Green's Theorem (11.1) in the xy-plane. Furthermore, we can chop up an arbitrary closed curve into (small) rectangular loops, and interior edges won't contribute to the total circulation, so that (12.1) holds for *any* loop; this is Stokes' Theorem.

There are a couple of subtleties here. First of all, the "inside" of a closed curve isn't well-defined in three dimensions. But Stokes' Theorem holds for *any* surface whose edge is the given loop. One way to visualize this is to think of the curve as the rim of a butterfly net; any butterfly net with the same rim is as good as any other.

It is also essential that the orientations of the curve and the surface appearing in Stokes' Theorem be compatible. When the normal vector $\hat{n}$ to the surface points *up*, we have already seen that the curve should be oriented counterclockwise *as seen from above*. More

generally, as you walk along the curve (facing forwards, that is, in the direction of $\hat{\boldsymbol{T}}$, with your head "up", that is, in the direction of $\hat{\boldsymbol{n}}$), the surface should always be on your *left*; this corresponds to the statement that $\hat{\boldsymbol{n}} \times \hat{\boldsymbol{T}}$ points *into*, rather than away from, the surface. Equivalently, a gear on the surface near the edge, rotating counterclockwise as seen from its axis $\hat{\boldsymbol{n}}$, should "push" the boundary curve in the direction $\hat{\boldsymbol{T}}$.

Notice that we have argued that Stokes' Theorem follows from Green's Theorem, while at the same time Green's Theorem is clearly a special case of Stokes' Theorem; these two theorems are in fact equivalent.

Finally, it is often easier to work in coordinates adapted to the given surface S, e.g. spherical coordinates when S is a sphere. To use Stokes' Theorem in such a situation, one must then have an expression for the curl in terms of the appropriate coordinates. Such formulas can be looked up in an appropriate reference, or derived directly by computing the circulation using "rectangular" loops adapted to the desired coordinates.

For example, finding $\operatorname{curl} \vec{\boldsymbol{G}}$ for

$$\vec{\boldsymbol{G}} = \frac{-y\,\hat{\boldsymbol{\imath}} + x\,\hat{\boldsymbol{\jmath}}}{x^2 + y^2 + z^2}$$

is quite messy in rectangular coordinates, whereas in spherical coordinates we have

$$\vec{\boldsymbol{G}} = \frac{\sin\theta}{r}\,\hat{\boldsymbol{\phi}}$$

and, after looking up the formula for the curl in spherical coordinates (e.g. in the book **Div, Grad, Curl, and All That**), we obtain

$$\operatorname{curl} \vec{\boldsymbol{G}} = \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\frac{\sin^2\theta}{r}\right)\hat{\boldsymbol{r}} - \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{r\sin\theta}{r}\right)\hat{\boldsymbol{\theta}} = \frac{2\cos\theta}{r^2}\,\hat{\boldsymbol{r}}$$

Noting that the (outward) unit normal vector to the sphere is just $\hat{\boldsymbol{r}}$, so it is now easy to compute the flux of the curl of $\vec{\boldsymbol{G}}$ over (part of) the sphere, that is, to calculate

$$\int_S \operatorname{curl} \vec{\boldsymbol{G}} \cdot d\vec{\boldsymbol{A}}$$

whereas to do this in rectangular coordinates (try it!) is a mess.

GOALS

- Be able to use Stokes' Theorem to evaluate integrals.

- Know the geometric interpretation of curl.

Lesson 13:   Curl and Divergence

RECALL: $\vec{\nabla} f = \frac{\partial f}{\partial x}\,\hat{\boldsymbol{\imath}} + \frac{\partial f}{\partial y}\,\hat{\boldsymbol{\jmath}} + \frac{\partial f}{\partial z}\,\hat{\boldsymbol{k}}$

It is convenient to think of $\vec{\nabla}$ as a differential operator (called "del"), written as

$$\vec{\nabla} = \frac{\partial}{\partial x}\,\hat{\boldsymbol{\imath}} + \frac{\partial}{\partial y}\,\hat{\boldsymbol{\jmath}} + \frac{\partial}{\partial z}\,\hat{\boldsymbol{k}}$$

Thus, $\vec{\nabla} f$ ("del f") is obtained by "multiplying" $\vec{\nabla}$ by f, being careful of course to interpret "multiplication" as partial differentiation.

The operator $\vec{\nabla}$ can also be used to express divergence and curl. We have seen that if

$$\vec{F} = F_x\,\hat{\boldsymbol{\imath}} + F_y\,\hat{\boldsymbol{\jmath}} + F_z\,\hat{\boldsymbol{k}}$$

then

$$\text{curl}\,\vec{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\hat{\boldsymbol{\imath}} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\hat{\boldsymbol{\jmath}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\hat{\boldsymbol{k}}$$

But this looks like a cross product! Rather than memorizing this formula, it is much easier to compute the curl directly using the determinant rule for cross product

$$\text{curl}\,\vec{F} = \vec{\nabla} \times \vec{F} = \begin{vmatrix} \hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}$$

Similarly, the divergence of $\vec{F}$ looks like a dot product, and can be written

$$\text{div}\,\vec{F} = \vec{\nabla} \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$

In both cases, care must be taken to correctly interpret "multiplication" as differentiation, as above.

This notation also helps you remember that the curl of a vector field is again a vector field, but the divergence of a vector field is a scalar field (i.e. a function), and of course the gradient of a function is a vector field.

There are several useful identities involving these operators. First of all, like all derivatives, they are linear. Second, there are the product rules

$$\begin{aligned} \vec{\nabla} \times (f\vec{G}) &= \vec{\nabla} f \times \vec{G} + f\,(\vec{\nabla} \times \vec{G}) \\ \vec{\nabla} \cdot (f\vec{G}) &= \vec{\nabla} f \cdot \vec{G} + f\,(\vec{\nabla} \cdot \vec{G}) \end{aligned}$$

which are both of the form:

> The derivative of a product is the derivative of the first factor times the second factor, plus the first factor times the derivative of the second factor.

where of course "derivative" and "times" must be interpreted appropriately. Be careful not to switch the order of the factors in the first identity!

Two further identities are

$$\vec{\nabla} \times \vec{\nabla} f = 0$$
$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{F}) = 0$$

The first identity follows from Stokes' Theorem, which we now write in the form

$$\oint_C \vec{F} \cdot d\vec{r} = \int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{A}$$

If $\vec{F} = \vec{\nabla} f$, it is conservative, so the left-hand-side must be zero by path independence for *any* closed curve C, which forces the *integrand* on the right-hand-side to vanish identically, which proves the identity.

To prove the second identity, suppose that *both* S_1 and S_2 meet the requirements of Stokes' Theorem for the same curve C. We can think of S_1 and S_2 as being two different butterfly nets (or soap bubbles) on the same wire frame C. Supposing that S_1 is above S_2, the *outward* unit normal to the closed surface $S = S_1 + S_2$ is the *upward* unit normal to S_1, but the *downward* unit normal to S_2. Thus,

$$\int_V \vec{\nabla} \cdot (\vec{\nabla} \times \vec{F}) \, dV = \int_S (\vec{\nabla} \times \vec{F}) \cdot \hat{\boldsymbol{n}}_{\text{out}} \, dA$$
$$= \int_{S_1} (\vec{\nabla} \times \vec{F}) \cdot \hat{\boldsymbol{n}}_{\text{up}} \, dA - \int_{S_2} (\vec{\nabla} \times \vec{F}) \cdot \hat{\boldsymbol{n}}_{\text{up}} \, dA$$
$$= \oint_C \vec{F} \cdot d\vec{r} - \oint_C \vec{F} \cdot d\vec{r} = 0$$

where the Divergence Theorem was used in the first step. This shows that the above integrals are zero over *any* volume V, forcing the integrand to vanish, which is the second identity.

The new notation for the divergence also allows us to write the Divergence Theorem as

$$\int_S \vec{F} \cdot d\vec{A} = \int_V \vec{\nabla} \cdot \vec{F} \, dV$$

There is also a 2-dimensional version of the Divergence Theorem, which is really a corollary of Green's Theorem, and which says

$$\oint_C \vec{F} \cdot \hat{\boldsymbol{n}} \, ds = \int_R \vec{\nabla} \cdot \vec{F} \, dA$$

where C is a closed curve in the plane, $\hat{\boldsymbol{n}}$ is its outward pointing unit normal vector, and R is the region inside the curve. The integral on the left defines flux in two dimensions.

Finally, a further useful combination is given by

$$\triangle f = \nabla^2 f = \vec{\nabla} \cdot (\vec{\nabla} f)$$

and is called the *Laplacian* of f.

GOALS

- Know how to find the divergence and curl of a vector field.

- Know the properties of divergence and curl.

Lesson 14: Change of Variables

      RECALL: $dA = dx\,dy = r\,dr\,d\phi$

Where did the factor of r come from in the above expression? Consider a "coordinate rectangle" bounded by curves of the form $r = $ constant, $\phi = $ constant. If this "rectangle" is small enough, its sides have length dr and $r\,d\phi$, so that its area is the product $(dr)(r\,d\phi)$. *

    For *orthogonal coordinates* in the plane, that is, one whose level curves (like $r = $ constant) are perpendicular, this method can be used to quickly determine the correct form of dA. This method also works fine for the volume element dV in 3 dimensions, yielding

$$dV = (dx)(dy)(dz) = (dr)(r\,d\phi)(dz) = (dr)(r\,d\theta)(r\sin\theta\,d\phi)$$

in rectangular, cylindrical, and spherical coordinates respectively.

    But what do we do if the coordinates are not orthogonal? We start by reviewing the 1-dimensional case, which is just the method of substitution.

      RECALL: $x = x(u)$ $\Longrightarrow$ $\displaystyle\int f\,dx = \int f\,\frac{dx}{du}\,du$

When making such a substitution, of course, one must replace x everywhere by u, so that both f and the integration limits must be rewritten in terms of u.

    Consider now a change of variables in 2 dimensions

$$x = x(u, v) \qquad y = y(u, v)$$

We can therefore write the position vector as

$$\vec{r} = \vec{r}(u, v) = x(u, v)\,\hat{\boldsymbol{\imath}} + y(u, v)\,\hat{\boldsymbol{\jmath}}$$

But this is just a special case of a parametric surface!

    We already know how to find the surface element of such surfaces. First compute

$$d\vec{A} = d\vec{r}_1 \times d\vec{r}_2 = \left(\frac{\partial\vec{r}}{\partial u} \times \frac{\partial\vec{r}}{\partial v}\right) du\,dv$$

and then take the magnitude of $d\vec{A}$ to obtain dA. We have

$$\frac{\partial\vec{r}}{\partial u} \times \frac{\partial\vec{r}}{\partial v} = \begin{vmatrix} \hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}} \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & 0 \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & 0 \end{vmatrix} = \hat{\boldsymbol{k}} \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix}$$

---

*Here and elsewhere in this lesson we are being somewhat informal. It is *not* true that the area of an infinitesimal rectangular rectangle is the same as the area of an infinitesimal polar rectangle. When we write, for instance, "$dx\,dy = r\,dr\,d\phi$", we really mean that the corresponding *integrals* are equal.

This determinant is important enough to have its own name; we define the *Jacobian* of the transformation from u, v to x, y to be *

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

Putting this all together, we obtain

$$dA = \left| \frac{\partial(x,y)}{\partial(u,v)} \right| du\, dv$$

Note that there are two sets of vertical bars here — this expression contains the *absolute value* of the Jacobian, since areas are positive.

Not surprisingly, when this procedure is applied to the polar coordinate transformation

$$x = r\cos\phi \qquad y = r\sin\phi$$

we get (check this yourself!)

$$\frac{\partial(x,y)}{\partial(r,\phi)} = r$$

so that, as we already knew, $dA = r\, dr\, d\phi$.

This not only tells us how to find area by doing an integral with respect to u and v (simply integrate the above expression), it tells us how to transform *any* integral in x and y to one involving u and v — hence the name "change of variables". We therefore have

$$\iint\limits_R f\, dx\, dy = \iint\limits_D f \left| \frac{\partial(x,y)}{\partial(u,v)} \right| du\, dv$$

where R is some region in the xy-plane, and D is that same region expressed in terms of u and v. We must also reinterpret f on the right as a function of u and v; this can be written formally as $f\big(x(u,v), y(u,v)\big)$, but this simply means to use what you know to rewrite all occurrences of x and y in terms of u and v.

To change variables in a double integral,

- Rewrite dA in terms of the new variables using the above formula.

- Rewrite the function being integrated in terms of the new variables.

- Change the limits from the old variables (R) to the new ones (D).

A similar formula holds for triple integrals. Given x, y, z as functions of u, v, w, the corresponding infinitesimal box is a parallelepiped with sides

$$d\vec{r}_1 = \frac{\partial\vec{r}}{\partial u}\, du \qquad d\vec{r}_2 = \frac{\partial\vec{r}}{\partial v}\, dv \qquad d\vec{r}_3 = \frac{\partial\vec{r}}{\partial w}\, dw$$

The volume of such a parallelepiped is given by the *triple product*

$$dV = \left| (d\vec{r}_1 \times d\vec{r}_2) \cdot d\vec{r}_3 \right| du\, dv\, dw$$

---

*Recall that the determinant of the transpose of a matrix is the same as the determinant of the matrix.

Defining the *Jacobian* for transformations of 3 variables to be

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$$

we obtain

$$dV = \left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| du\, dv\, dw$$

A useful example, found in most textbooks, is to use the Jacobian to derive the volume element dV in spherical coordinates, which you should already know. *(Many texts use a left-handed convention for spherical coordinates, which results in the determinant being negative. The absolute value signs occur in the change of variables formula precisely to prevent negative area and volume elements when transforming between right-handed and left-handed coordinate systems.)*

Knowing dV we can convert *any* integral using

$$\iiint\limits_{V} f\, dx\, dy\, dz = \iiint\limits_{W} f \left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| du\, dv\, dw$$

for appropriate regions V and W (and with f rewritten along the lines above). The above method for double integrals adapts easily to triple integrals.

In practice, one is often given u and v in terms of x and y, not the other way around. Rather than solving for x and y, which is often difficult, there is a simpler way. Inverting the roles of the two sets of variables in the derivation above, we must have not only

$$dx\, dy = \left| \frac{\partial(x,y)}{\partial(u,v)} \right| du\, dv$$

but also

$$du\, dv = \left| \frac{\partial(u,v)}{\partial(x,y)} \right| dx\, dy$$

from which we can conclude that

$$\frac{\partial(x,y)}{\partial(u,v)} = 1 \left/ \frac{\partial(u,v)}{\partial(x,y)} \right.$$

But the expression on the right can be computed without solving for x and y! This gives another way to compute the (left-hand) Jacobian in this case. A similar formula holds in 3 dimensions. If you use this method, don't forget to take the reciprocal!

GOALS
- Know the formulas for dA and dV in common coordinate systems (rectangular, polar, cylindrical, spherical).

- Know how to calculate dA and dV using Jacobians in 2 and 3 dimensions.

- Know how to use "inverse" Jacobians.

- Know how to use change of variables to evaluate double and triple integrals.

REVIEW

Most texts contain a very nice summary of the main theorems. Here is a partial list of topics you should review:

- **The Gradient:** You should understand what the gradient of a function means geometrically and physically. (Can you relate the gradient to the level sets of the function?)

- **Line Integrals:** You should be able to compute line integrals, determine whether vector fields are conservative, be able to find potential functions, and know the Fundamental Theorem for Line Integrals.

- **Surface Integrals:** You should know the surface elements for simple surfaces, and how to compute surface elements for other surfaces. You should be able to compute surface integrals, and understand flux.

- **The Divergence Theorem:** You should know how to apply the Divergence Theorem — and when to use it. (Is your surface closed?!) You should know what the divergence of a vector field means geometrically and physically. (Can you recognize in simple cases whether a vector field has nonzero divergence from a sketch?)

- **Stokes' Theorem:** You should know how to apply Stokes' Theorem — and when to use it. (Is your curve closed?!) You should know what the curl of a vector field means geometrically and physically. (Can you recognize in simple cases whether a vector field has nonzero curl from a sketch? Can you relate this to whether the vector field is conservative?)

- **Green's Theorem:** Green's Theorem is just a special case of Stokes' Theorem. You may wish to study the corollaries about area and the two vector versions of the theorem, one about (2-dimensional) flux and the other about work or circulation.

- **Change of Variables:** You should recognize simple changes of variables to use in particular problems, and should further know how to effect the change of variables once you have chosen (or are given) the variables. Such problems can be solved either using Jacobians or by viewing them as special cases of surface integrals.