Chapter 7

Determinants and Integral Vector Calculus

7.1 Properties of Determinants

Recall that only square ($n \times n$) matrices have determinants. Here we shall extend the brief treatment in Sec. 2.5 to state the definition and principal algebraic properties of determinants of matrices of any size. Very few detailed proofs will be given. Most of the proofs would be rather tedious and, unlike those in the rest of linear algebra, not very instructive. A more detailed, but quite readable, treatment of determinants is Chapter 4 of B. Kolman, *Elementary Linear Algebra* (Macmillan, 1970).

Remark: The determinant of A^{t} (the transpose of A) is the same as the determinant of A. Therefore, any theorem about determinants remains true if "row" is replaced by "column" everywhere and vice versa.

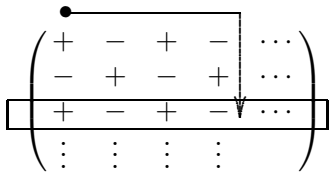
One way to define determinants goes as follows: The determinant of a 1×1 matrix (i.e., a number) is the number itself. The cofactor expansion (stated next) reduces the calculation of an $n \times n$ determinant to the calculation of n determinants of size $n-1$. Thus $\det A$ is defined *recursively* for square matrices of all sizes.

Algorithm (the Cofactor Expansion): If $A = \{a_{jk}\}$, then for any j,

$$\det A = \sum_{k=1}^{n} a_{jk}\,(-1)^{j+k} \det(\mathcal{A}_{jk}),$$

where $\mathcal{A}_{jk}$ is the matrix obtained by omitting the row and column of A containing a_{jk}. That is, you pick a row (the jth) and use the "checkerboard rule" to get the signs right:

$j = 3$:



$(-1)^{3+4} = -1$

(Always start with $+$ in the upper left corner. Then count modulo 2 (that is, "plus, minus, plus, minus, plus, ... ") along a right-angled path to the jk position.)

In accordance with the preceding remark, one can use a column as the basis for a cofactor expansion instead of a row.

Example: Calculate

$$D = \begin{vmatrix} 1 & 0 & -4 & 0 \\ 25 & 3 & 40 & -5 \\ -2 & 0 & 3 & 1 \\ 1 & 0 & -1 & 4 \end{vmatrix}.$$

SOLUTION: Because the second column is so simple, we expand around it. The checkerboard sign of the isolated "3" is $+$, so

$$D = 3 \begin{vmatrix} 1 & -4 & 0 \\ -2 & 3 & 1 \\ 1 & -1 & 4 \end{vmatrix}.$$

Now the top row is the simplest, and we get

$$D = 3\left[(1)\begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix} - (-4)\begin{vmatrix} -2 & 1 \\ 1 & 4 \end{vmatrix}\right]$$
$$= 3[12 - (-1) + 4(-8 - 1)] = 3(-23) = -69.$$

Each number $(-1)^{j+k} \det \mathcal{A}_{jk}$ is called a *cofactor*. Note that for given j and k, the symbol $\mathcal{A}_{jk}$ stands for a *whole matrix*, not just one element of a matrix. These smaller matrices (or their determinants, depending upon the author) are called *minors* of A; thus, the principal difference between a minor and a cofactor is the checkerboard sign. (Also, note that to minimize confusion we have used lower-case letters for the elements of A.)

An ugly feature of the cofactor definition of the determinant is that it is not obviously unique. We could expand the determinant in cofactors of any row, or of any column, and it needs to be proved that the result is the same in all these cases. One solution is to start from a more symmetrical, and theoretically preferable, definition, and show that the result is equal to any (hence all) of the cofactor expansions.

Definition: The *determinant* of the $n \times n$ matrix $A = \{a_{jk}\}$ is a sum of $n!$ terms, each of which contains as a factor exactly one element from each

row of A, and likewise contains exactly one element from each column of A. The sign of the term is $+$ or $-$ depending upon whether the number of interchanges needed to "unscramble" the column indices into the same order as the row indices is even or odd. That is, each term of the determinant has the form

$$\sigma(p) a_{1p(1)} a_{2p(2)} \cdots a_{np(n)},$$

where p is a *permutation* of the integers 1 through n (i.e., a one-to-one function from the set $\{1, \ldots, n\}$ onto itself); each permutation can be decomposed into elementary *transpositions* (interchanges of two elements), and $\sigma(p) = \pm 1$ consists of one factor of -1 for each such transposition.

To understand this definition, consider the low-dimensional cases that are already familiar and have been reviewed in Sec. 2.5.

1. When $n = 1$, the definition just reiterates that the determinant is the same as the matrix itself.

2. The rows and columns of a 2×2 matrix are labelled by the integers $\{1, 2\}$. There are only two ways of ordering these indices: the original order and the reverse. In the second case, one interchange is needed to unscramble the order back to the original. Therefore, the determinant is
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \, .$$

3. There are 6 ways to order the integers $\{1, 2, 3\}$:
$$1,2,3; \qquad 1,3,2; \qquad 2,1,3; \qquad 3,2,1; \qquad 2,3,1; \qquad 3,1,2.$$
The last two permutations are the *cyclic* ones that we have already encountered while studying how the three coordinates $\{x, y, z\}$ enter into cross-product and curl identities; they are *even* permutations, because 2 interchanges are needed to drag the displaced index back to its proper position at the opposite end. The identity permutation (the first in the list) is also even. The remaining three permutations are *odd*, because each of them involves a single interchange of a pair of indices. (In the case $(3, 2, 1)$, three steps are needed to unscramble if only interchanges of *adjacent* indices are considered, but only one step if jumps are allowed. In either case, the number is odd. More generally, it can be proved that the parity (evenness or oddness) is the

same for any successful unscrambling operation upon a permutation, although the actual number of interchanges may vary.) Therefore, the determinant is

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{cases} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ -a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \, , \end{cases}$$

which is the same as the result of the "criss-cross" or "basket-weaving" calculation of 3×3 determinants reviewed in Sec. 2.5.

Warning: The criss-cross method is *not applicable* to determinants of size 4×4 or larger. If you try to generalize the criss-cross diagram in Sec. 2.5 to a 4×4 matrix, you will get only 8 terms. But the determinant should contain $4! = 24$ terms. Closer inspection reveals that the criss-cross calculation produces only the terms that are either cyclic permutations of $(1, 2, 3, 4)$ or cyclic permutations of $(4, 3, 2, 1)$. Many other permutations, such as $(1, 3, 2, 4)$, are left out.

The permutation definition is useful in deducing general properties of determinants — for example, we will use it in Sec. 8.1 to establish the algebraic form of the equation for the eigenvalues of a matrix — but not very useful for calculations. The cofactor definition is easier to understand and to calculate, but it too becomes very tedious to implement in high dimensions. The next few properties on our list help to calculate determinants more efficiently.

Definition and Theorem: A *triangular* matrix is one with nothing but zeros below (or above) the main diagonal, such as $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. The determinant of a triangular matrix is the product of the elements on the diagonal. (This includes *diagonal* matrices as a special case.)

Theorem: Under the elementary row operations, the determinant of a square matrix changes in these ways:

(A) If a row is multiplied by a scalar, the determinant changes by that same scalar factor. [NOTE: If the entire $n \times n$ matrix is multiplied by a scalar r, the determinant is multiplied by r^n, not r.]

(B) If two rows are interchanged, the determinant changes sign. [**Corollary:** The determinant is 0 if two rows are equal, or even if two rows are proportional.]

(C) If a scalar multiple of one row is added to another row, the determinant does not change.

This theorem is very useful in evaluating large determinants. The strategy is to make the matrix triangular (or almost triangular) before expanding by cofactors.

Example: Calculate

$$D = \begin{vmatrix} 1 & 0 & 1 & 0 \\ 2 & 2 & 2 & 5 \\ 1 & 1 & 1 & 1 \\ 2 & 3 & 5 & 1 \end{vmatrix}.$$

SOLUTION: Since the first and third columns are almost alike, we subtract one from the other:

$$D = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 2 & 0 & 5 \\ 1 & 1 & 0 & 1 \\ 2 & 3 & 3 & 1 \end{vmatrix}.$$

Now we notice that the matrix is within inches of being triangular. Subtract 5 times the third row from the second row:

$$D = \begin{vmatrix} 1 & 0 & 0 & 0 \\ -3 & -3 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 2 & 3 & 3 & 1 \end{vmatrix} = (-3) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 2 & 3 & 3 & 1 \end{vmatrix}.$$

Now an interchange and another row subtraction make the matrix completely triangular:

$$D = (+3) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 3 & 3 & 1 \\ 1 & 1 & 0 & 1 \end{vmatrix} = (3) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 \\ 1 & 1 & 0 & 1 \end{vmatrix} = (3)(1)(1)(3)(1) = 9.$$

(Of course, when your goal is just to obtain the numerical answer as quickly as possible, you are free to stop the row reduction and switch to the cofactor method as soon as you have produced a row or column consisting mostly of zeros. Here we carried out the row reduction to the end, to demonstrate the method fully.)

The application of row operations to determinants differs from the solution of linear systems in two important respects. First, *column* operations of the same types are equally legal. Second, it is critically important to keep track of the numerical factors produced by operations of types (A) and (B).

Theorem: If $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ are n vectors in $\mathbf{R}^n$, then the determinant of the square matrix whose rows are the $\vec{x}$s is a linear function of each $\vec{x}_j$ when the other $n-1$ vectors are held fixed. (Such a function of several vector variables is called *multilinear*.)

REMARK: Parts (A) and (C) of the previous theorem are special cases of this one.

We long ago remarked the following:

Theorem: A^{-1} exists (i.e., A is nonsingular) if and only if $\det A \neq 0$.

The following simple fact is surprisingly difficult to prove directly. In the next section we shall give an indirect argument for it.

Theorem: $\det(AB) = (\det A)(\det B)$.

We close with an application of determinants that was once famous but is now in some disrepute:

Theorem (Cramer's Rule for solving linear equations): If A is nonsingular, the solution of the system $A\vec{x} = \vec{y}$ is

$$x_j = \frac{\det A_j}{\det A},$$

where the matrix A_j is obtained from A by replacing the jth column by $\vec{y}$.

Example: Solve $x + y = 3, \quad x - y = 5$. The rule gives

$$x = \frac{\begin{vmatrix} 3 & 1 \\ 5 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{-8}{-2} = 4, \qquad y = \frac{\begin{vmatrix} 1 & 3 \\ 1 & 5 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{2}{-2} = -1.$$

Theorem (Cramer's Rule for matrix inversion): If A is nonsingular, then

$$\left(A^{-1}\right)_{jk} = \frac{1}{\det A}\,(-1)^{j+k}(\det \mathcal{A}_{kj}).$$

That is, you divide the *transpose* of the matrix of cofactors by $\det A$. (In the 2×2 case, this is just the well-known rule, "Interchange the elements

on the main diagonal; change the signs of the off-diagonal elements without moving them; and divide by the determinant.") The transposed cofactor matrix is called the *adjugate* of A. (In some books it is called the *adjoint*, but nowadays that term is better reserved for the complex conjugate of the transpose of a matrix.)

Example: Find the inverse of the Jacobian matrix of the spherical coordinate system (see Exercise 4.2.3 and Sec. 5.5).

SOLUTION: The Jacobian is

$$J = \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}.$$

The determinant of J can be calculated (Exercise 7.3.6) to be $r^2\sin\theta$, which is familiar as the element of volume for integration in spherical coordinates (cf. Sec. 7.3 below). The top left element of J^{-1} is

$$\begin{aligned} \frac{\partial r}{\partial x} &= (r^2\sin\theta)^{-1}(-1)^0 \begin{vmatrix} r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ -r\sin\theta & 0 \end{vmatrix} \\ &= \frac{r^2\sin^2\theta\cos\phi}{r^2\sin\theta} = \sin\theta\cos\phi. \end{aligned}$$

(This equals x/r, which is what you would get by solving for r as $(x^2+y^2+z^2)^{\frac{1}{2}}$ and differentiating.) The middle element of the first column of J^{-1} is

$$\begin{aligned} \frac{\partial\theta}{\partial x} &= (r^2\sin\theta)^{-1}(-1)^1 \begin{vmatrix} \sin\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & 0 \end{vmatrix} \\ &= \frac{r\cos\theta\sin\theta\cos\phi}{r^2\sin\theta} = \frac{\cos\theta\cos\phi}{r}, \end{aligned}$$

and the middle element of the top row is

$$\begin{aligned} \frac{\partial r}{\partial y} &= (r^2\sin\theta)^{-1}(-1)^1 \begin{vmatrix} r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ -r\sin\theta & 0 \end{vmatrix} \\ &= \frac{r^2\sin^2\theta\sin\phi}{r^2\sin\theta} = \sin\theta\sin\phi. \end{aligned}$$

The tricky point is not to get these last two elements confused.* This should be enough of the calculation of J^{-1} to make the point; you can finish it as Exercise 7.3.6.

---

* Notice that physical dimensions dictate the location of the factors of r. This should serve as a check.

Remark: Cramer's rules are *less efficient* than row reduction (except for 2×2 matrices). The larger your matrix, the more advantageous row reduction is. Cramer's rules do have the advantage of providing *explicit formulas* for the answers, rather than just *procedures* which have to be applied to each particular case. This may be useful when you want to study the dependence of the answer on some parameter, for instance.

Additional examples

Example 1. Calculate the determinant of the matrix

$$A = \begin{pmatrix} 8 & 7 & 1 \\ 4 & 3 & 1 \\ -1 & -2 & 1 \end{pmatrix}.$$

Solution: We have already evaluated this determinant in two ways in Sec. 2.5. This time we'll apply the method of row operations to produce a triangular form. First transpose the first and the third rows, then add 4 times the first row to the second and 8 times the first row to the third to get

$$\det A = \begin{vmatrix} 8 & 7 & 1 \\ 4 & 3 & 1 \\ -1 & -2 & 1 \end{vmatrix} = - \begin{vmatrix} -1 & -2 & 1 \\ 4 & 3 & 1 \\ 8 & 7 & 1 \end{vmatrix} = - \begin{vmatrix} -1 & -2 & 1 \\ 0 & -5 & 5 \\ 0 & -9 & 9 \end{vmatrix}$$

$$= -5 \cdot 9 \begin{vmatrix} -1 & -2 & 1 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{vmatrix} = -45 \begin{vmatrix} -1 & -2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

Example 2. Find $\begin{vmatrix} 0 & a & a \\ a & 0 & a \\ a & a & 0 \end{vmatrix}$.

Solution: Every matrix element contains a factor a. Factor it out:

$$a^3 \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = a^3(0 + 1 + 1 - 0 - 0 - 0) = 2a^3.$$

Example 3. Verify that

$$\begin{vmatrix} \frac{x_1+x_2}{2} & \frac{y_1+y_2}{2} & 1 \\ \frac{x_1-x_2}{2} & \frac{y_1-y_2}{2} & 1 \\ x_1 & y_1 & 1 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.$$

SOLUTION: Add (-1) times the first row to the second and third rows to get a cofactor reduction to a 2×2 determinant:

$$\begin{vmatrix} \frac{x_1+x_2}{2} & \frac{y_1+y_2}{2} & 1 \\ \frac{x_1-x_2}{2} & \frac{y_1-y_2}{2} & 1 \\ x_1 & y_1 & 1 \end{vmatrix} = \begin{vmatrix} \frac{x_1+x_2}{2} & \frac{y_1+y_2}{2} & 1 \\ -x_2 & -y_2 & 0 \\ \frac{x_1-x_2}{2} & \frac{y_1-y_2}{2} & 0 \end{vmatrix} = 1 \begin{vmatrix} -x_2 & -y_2 \\ \frac{x_1-x_2}{2} & \frac{y_1-y_2}{2} \end{vmatrix}.$$

Now extract scalar factors, add the first row to the second, and interchange rows:

$$-\frac{1}{2} \begin{vmatrix} x_2 & y_2 \\ x_1 - x_2 & y_1 - y_2 \end{vmatrix} = -\frac{1}{2} \begin{vmatrix} x_2 & y_2 \\ x_1 & y_1 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.$$

Example 4. Verify that

$$\begin{vmatrix} 1+\lambda x+\mu x^2 & x & x^2 \\ 1+\lambda y+\mu y^2 & y & y^2 \\ 1+\lambda z+\mu z^2 & z & z^2 \end{vmatrix} = \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix},$$

where μ and λ are arbitrary real numbers.

SOLUTION: Calculate the determinant on the left side using elementary column operations. First, add $(-\lambda)$ times the second column to the first, then add $(-\mu)$ times the third column to the first, getting

$$\begin{vmatrix} 1+\lambda x+\mu x^2 & x & x^2 \\ 1+\lambda y+\mu y^2 & y & y^2 \\ 1+\lambda z+\mu z^2 & z & z^2 \end{vmatrix} = \begin{vmatrix} 1+\mu x^2 & x & x^2 \\ 1+\mu y^2 & y & y^2 \\ 1+\mu z^2 & z & z^2 \end{vmatrix} = \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}.$$

Example 5. Evaluate $\begin{vmatrix} \sin\alpha & \cos\alpha & 1 \\ \sin 2\alpha & \cos 2\alpha & 1 \\ \sin 3\alpha & \cos 3\alpha & 1 \end{vmatrix}$.

SOLUTION: Add (-1) times the first row to the second and third rows to get

$$\begin{vmatrix} \sin\alpha & \cos\alpha & 1 \\ \sin 2\alpha & \cos 2\alpha & 1 \\ \sin 3\alpha & \cos 3\alpha & 1 \end{vmatrix} = \begin{vmatrix} \sin\alpha & \cos\alpha & 1 \\ \sin 2\alpha - \sin\alpha & \cos 2\alpha - \cos\alpha & 0 \\ \sin 3\alpha - \sin\alpha & \cos 3\alpha - \cos\alpha & 0 \end{vmatrix}$$

$$= \begin{vmatrix} \sin 2\alpha - \sin\alpha & \cos 2\alpha - \cos\alpha \\ \sin 3\alpha - \sin\alpha & \cos 3\alpha - \cos\alpha \end{vmatrix}$$

$$= (\sin 2\alpha - \sin\alpha)(\cos 3\alpha - \cos\alpha) - (\sin 3\alpha - \sin\alpha)(\cos 2\alpha - \cos\alpha)$$

$$= \sin 2\alpha \cos 3\alpha - \sin\alpha \cos 3\alpha - \sin 2\alpha \cos\alpha + \sin\alpha \cos\alpha$$

$$- \sin 3\alpha \cos 2\alpha + \sin\alpha \cos 2\alpha + \sin 3\alpha \cos\alpha - \sin\alpha \cos\alpha$$

$$= (\sin 2\alpha \cos 3\alpha - \sin 3\alpha \cos 2\alpha) + (\sin 3\alpha \cos\alpha - \sin\alpha \cos 3\alpha)$$

$$+ (\sin\alpha \cos 2\alpha - \sin 2\alpha \cos\alpha)$$

$$= -\sin\alpha + \sin 2\alpha - \sin\alpha = 2\sin\alpha(\cos\alpha - 1).$$

Quadrature rules

Numerical analysis is a field that uses a tremendous amount of linear algebra. The most elementary and familiar problem of numerical analysis is *numerical integration*, also known as *quadrature*. Suppose that we need to know (approximately) a definite integral of some function f whose indefinite integral is too hard to calculate. (This may happen because f is given by a complicated formula; more often, it happens because we don't know any formula at all for f but can only compute or measure the values of f at a finite number of particular points.) Recall the standard approach from calculus courses: Divide up the interval of integration into small pieces on each of which the graph of f can be well approximated by a parabola; then integrate the quadratic function whose graph is the parabola. After a certain amount of algebra, this leads to *Simpson's rule*. The most obvious, but hardest, way to do this algebra is to find explicitly the formula for the parabola that passes through three points on the graph of f. Here we will show a much neater derivation, as a special case of a general method of deriving formulas for numerical integration.

Let us derive a formula for approximating $\int_{-1}^{1} f(x)\,dx$ in terms of the values of f at 3 points:

$$\int_{-1}^{1} f(x)\,dx \approx a_1 f(x_1) + a_2 f(x_2) + a_3 f(x_3).$$

We require that the formula give the exact answers for the three functions $f(x) = 1$, $f(x) = x$, and $f(x) = x^2$, and solve for the coefficients a_i. Since both the formula and the integral itself are linear functionals of f, the resulting formula will give the exact answer for *any linear combination* of the three basis functions; that is, we have a quadrature rule that is exact for all quadratic polynomials.

Following this strategy, we derive 3 equations in 3 unknowns:

$$\begin{aligned}
2 &= \int_{-1}^{1} 1\,dx = a_1 + a_2 + a_3\,, \\
0 &= \int_{-1}^{1} x\,dx = a_1 x_1 + a_2 x_2 + a_3 x_3\,, \\
\frac{2}{3} &= \int_{-1}^{1} x^2\,dx = a_1 {x_1}^2 + a_2 {x_2}^2 + a_3 {x_3}^2\,.
\end{aligned}$$

It is natural to solve these equations by Cramer's rule. The determinant that appears in the denominator is then

$$V \equiv \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{vmatrix}.$$

Determinants of this form (consisting of successive powers of a list of numbers) are called *Vandermonde determinants* and arise repeatedly in derivations in numerical analysis. They are the subject of Exercises 7.1.3–6. In particular, Exercise 7.1.3 shows that this particular determinant is

$$V = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

Now that the general point is clear (but the algebra threatens to become messy), let's retreat to the special case of Simpson's Rule:

$$x_1 = -1, \quad x_2 = 0, \quad x_3 = 1.$$

Then $V = 2$, and Cramer's rule gives

$$a_1 = \frac{1}{2}\begin{vmatrix} 2 & 1 & 1 \\ 0 & 0 & 1 \\ \frac{2}{3} & 0 & 1 \end{vmatrix} = \frac{1}{2}(-1)\left(-\frac{2}{3}\right) = \frac{1}{3},$$

$$a_2 = \frac{1}{2}\begin{vmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \\ 1 & \frac{2}{3} & 1 \end{vmatrix} = -\frac{1}{2}(2)\begin{vmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ \frac{1}{3} & 1 & 1 \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & \frac{4}{3} \end{vmatrix} \text{ (by row reduction)} \quad = \frac{4}{3},$$

$$a_3 = \frac{1}{2}\begin{vmatrix} 1 & 1 & 2 \\ -1 & 0 & 0 \\ 1 & 0 & \frac{2}{3} \end{vmatrix} = \frac{1}{3}.$$

So the proposed formula is

$$\int_{-1}^{1} f(x)\,dx \approx \frac{1}{3}\,f(x_1) + \frac{4}{3}\,f(x_2) + \frac{1}{3}\,f(x_3).$$

This is indeed Simpson's rule, or at least the basic special case of it. For the general case of a function known at an odd number of equally spaced

points $\{x_0 \equiv a, x_1, \ldots, x_n \equiv b\}$, one applies the basic rule to each interval $[x_{2j}, x_{2j+2}]$, mapping that interval to $[-1, 1]$ by an affine transformation of the independent variable. The scaling part of the affine transformation produces a chain-rule factor of $\Delta x \equiv (b-a)/n$. Adding all these little parabolic integrals gives the full Simpson rule,

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{3}\,[f(x_0) + 4f(x_1) + 2f(x_2) + \cdots + 4f(x_{n-1}) + f(x_n)].$$

Returning to the case of three general coordinates x_1 , x_2 , x_3 : Suppose that the numbers $f(x_j)$ must be obtained as empirical data, and that the right endpoint of the interval is located, say, inside a blast furnace, where making a measurement is very inconvenient. Then it would make sense to derive and use, instead of Simpson's rule, the formula with $x_3 = \frac{1}{2}$ (for instance) instead of 1. This just requires redoing our Cramer calculation with different numbers.

Another generalization is to integrals that are highly oscillatory, such as

$$\int_{-1}^{1} \sin(100\pi x)\, f(x)\,dx.$$

To approximate such an integral well by Simpson's rule would require computing a huge number of points, since each individual arc of the sine function would need to be approximated by a parabolic segment of length $2\Delta x$ (at most). It would make more sense to approximate just the function f, rather than the entire integrand, by quadratic polynomials (under the assumption that f itself is not as wiggly as the other factor). So we would look for a formula of the form

$$\int_{-1}^{1} \sin(100\pi x)\, f(x)\,dx \approx c_1 f(-1) + c_2 f(0) + c_3 f(1)$$

(called a *Filon rule*). Note that the Vandermonde determinant that arises here, and in fact the entire right-hand side of the linear system determining the c_j , is exactly the same as for Simpson's rule; only the nonhomogeneous terms will be different (involving integrals of the sine function). We could find the inverse of this *Vandermonde matrix* once and for all and use it to derive many different Simpson-like formulas with various integrand structures; for example, Filon rules with many different numbers playing the role of the 100π in the example. Exercise 7.1.13 is to derive a Filon rule for the simpler case of two points (generalizing the trapezoidal rule instead of Simpson's).

Exercises

7.1.1 If M is invertible, how is $\det(M^{-1})$ related to $\det M$?

7.1.2 Use Cramer's rule to solve

$$\begin{aligned} x + \ \ y &= a, \\ rx + sy &= b \end{aligned}$$

for x and y (with r, s, a, b as parameters).

7.1.3 Show that

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

7.1.4 Consider $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{vmatrix}$.

(a) Guess the answer by analogy with the previous exercise. (Set $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$. It may help to peek ahead to Exercise 7.1.6.)

(b) Calculate the determinant by the row-reduction method.

7.1.5 Find the quadrature rule of the form

$$\int_{-1}^{1} f(x)\,dx \approx a_1 f(x_1) + a_2 f(x_2) + a_3 f(x_3)$$

(exact for all quadratic f) assuming that the sample points are symmetrically spaced,

$$x_1 = -\xi, \quad x_2 = 0, \quad x_3 = +\xi,$$

but without setting $\xi = 1$. Verify that your formula reduces to Simpson's rule when $\xi = 1$.

7.1.6 Solve

$$\begin{vmatrix} 1 & x & x^2 & \dots & x^n \\ 1 & a_1 & a_1^2 & \dots & a_1^n \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 1 & a_n & a_n^2 & \dots & a_n^n \end{vmatrix} = 0,$$

where a_i are distinct real numbers for $i = 1, 2, \dots, n$. HINTS: (1) The determinant is equal to 0 if two rows are equal or proportional. (2) Use the fundamental theorem of algebra (see Secs. 4.1 and 8.1) to count how many roots the equation can have.

7.1.7 Find the determinants of these matrices:

(a) $\begin{pmatrix} z & z & z \\ -z & z & -z \\ z & -z & -z \end{pmatrix}$ (b) $\begin{pmatrix} b & -1 & b \\ 1 & b & -1 \\ b & 1 & b \end{pmatrix}$

7.1.8 Find the determinants of these matrices:

(a) $\begin{pmatrix} 0 & 0 & 0 & 4 \\ 0 & 0 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & -1 & -2 & -3 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 3 & 4 & 5 & 6 \end{pmatrix}$

7.1.9 *Cylindrical coordinates* are defined by

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = h. \tag{$*$}$$

We know of two ways to construct a normal vector to the cylinder $r = 2$:

(1) Find ∇r as a row of the inverse of the Jacobian matrix of the coordinate transformation ($*$). (See Sec. 5.5.)

(2) Set $r = 2$ in the formulas ($*$) and calculate $\frac{\partial \vec{r}}{\partial \theta} \times \frac{\partial \vec{r}}{\partial h}$. Since the two vectors are parallel to the surface, the cross product is perpendicular to the surface. (See Secs. 4.2 and 7.6.)

(a) Do both of those calculations and verify that they are consistent (up to a normalization factor).

(b) Explain why this will *always* work out right (for *any* 3-dimensional coordinate transformation and *any* curved coordinate surface in it). HINT: Relate Cramer's rule to the cross product.

7.1.10 Unfinished business from Sec. 5.4: Prove that the third condition of Theorem 4 of that section is indeed equivalent to the other two.

7.1.11 We know that the determinant of a diagonal or triangular matrix is the product of the elements on the main diagonal. Is the corresponding thing true of a "backwards diagonal" matrix, such as $\begin{pmatrix} 0 & 0 & 4 \\ 0 & 2 & 0 \\ -1 & 0 & 0 \end{pmatrix}$? (Be careful about *minus signs*. Start with 2×2 matrices, then consider 3×3, etc.)

7.1.12 Let x_1 and x_2 be two real numbers. Find a quadrature (numerical integration) formula of the form

$$\int_0^1 f(x)\,dx \approx c_1 f(x_1) + c_2 f(x_2)$$

by requiring that the formula give the right answer for $f(x) = 1$ and for $f(x) = x$. Then verify that your answer reduces to the usual trapezoidal rule when $x_1 = 0$ and $x_2 = 1$.

7.1.13 Find a quadrature formula (a version of *Filon's rule*) of the form

$$\int_0^1 \sin(100\pi x)\, f(x)\,dx \approx c_1 f(0) + c_2 f(1)$$

by requiring that the formula give the right answer for $f(x) = 1$ and for $f(x) = x$.

In the remaining exercises, calculate the given determinants.

7.1.14 $\begin{vmatrix} 0 & 1 & 0 & 0 \\ 2 & 9 & 1 & 2 \\ 3 & 8 & 1 & 5 \\ 4 & 7 & 1 & 4 \end{vmatrix}$

7.1.15 $\begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 5 & 1 & 1 \\ 2 & 7 & 2 & 3 \\ 4 & -2 & 5 & 5 \end{vmatrix}$

7.1.16 $\begin{vmatrix} 2 & 3 & 5 & -3 \\ 0 & 0 & 1 & 0 \\ 1 & 2 & 10 & -1 \\ 1 & 1 & -5 & 1 \end{vmatrix}$

7.1.17 $\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$

7.1.18 $\begin{vmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 5 & 7 \\ 0 & 0 & 6 & 8 \\ 2 & 4 & 0 & 0 \end{vmatrix}$

7.1.19 $\begin{vmatrix} a & b & c & d \\ 3 & -1 & 4 & 3 \\ 2 & -3 & 4 & 1 \\ 4 & -2 & 3 & 2 \end{vmatrix}$

7.1.20 $\begin{vmatrix} 2 & -1 & 5 & a \\ 4 & -3 & 4 & b \\ 3 & -2 & 2 & c \\ 5 & -4 & 4 & d \end{vmatrix}$

7.1.21 $\begin{vmatrix} 1 & 1 & 3 & a \\ 1 & 3 & 1 & b \\ 3 & 1 & 1 & c \\ 1 & 1 & 1 & d \end{vmatrix}$

7.2 Volume, Rotations, and All That: The Geometrical Significance of Determinants and Antisymmetry

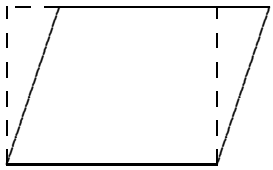
A full theory of n-dimensional volume goes far beyond the scope of this book. We shall build up an intuitive overview of the subject with a few remarks. In this section it is understood that we are working in $\mathbf{R}^n$ with the standard inner product, or in physical space equipped with an *orthonormal* basis, whenever we write formulas in terms of components of vectors.

The area (2-dimensional volume) of a rectangle is, by definition, the product of its length and width. Similarly, the volume of a 3-dimensional rectangular parallelepiped is the product of its three side lengths. Rectangular parallelepipeds in any dimension can be treated in the same way.

Roughly speaking, the volume of a more general region in $\mathbf{R}^n$ can be approximated by cutting up the region into nonoverlapping rectangular parallelepipeds ("bricks") and adding up their volumes. This decomposition

can be done only approximately, because in general the bricks will not fit snugly against the boundary of the region. By taking a limit of arbitrarily small bricks, however, we can bring their total volume arbitrarily close to what we mean intuitively by the volume of the region.

A particularly simple and pertinent type of region is the nonrectangular parallelepiped. In dimension 2, it is well known that the area of a generic parallelogram is the product of its base length and its height (measured perpendicularly to the base). This is easily seen by cutting a triangle off one side of the parallelogram and attaching it to the other, thereby producing a rectangle of the same area:



This principle extends to higher dimensions. The 3-dimensional case can be demonstrated by a deck of cards. The deck can be tilted into a nonrectangular parallelepipedal shape. The volume is still the same, since it is occupied by the same cardboard. The height of the deck is the same, and so is the area of the rectangular base. Therefore, the volume is still the product of the height by the area of the base.

What does this have to do with determinants? Everything. It can even be used as the definition of a determinant.*

Theorem: Any n vectors in $\mathbf{R}^n$, $\{\vec{r}_1, \ldots, \vec{r}_n\}$, determine a parallelepiped with one corner at the origin and the n vectors as adjacent edges. The volume of this parallelepiped is the absolute value of the determinant of the matrix whose columns are $\vec{r}_1, \ldots, \vec{r}_n$. (The volume is 0 precisely when the vectors are dependent — that is, they all lie in some lower-dimensional subspace.) The sign of this determinant depends only on the order of the vectors $\vec{r}_j$ and distinguishes "right-handed" from "left-handed" bases, the natural basis for $\mathbf{R}^n$ being taken to be right-handed by definition.

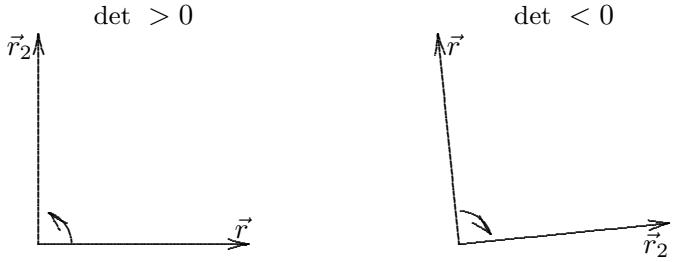
Remark: For brevity, we naturally refer to the determinant of the matrix whose columns are $\vec{r}_1, \ldots, \vec{r}_n$ as "the determinant of the vectors $\vec{r}_1, \ldots, \vec{r}_n$".

---

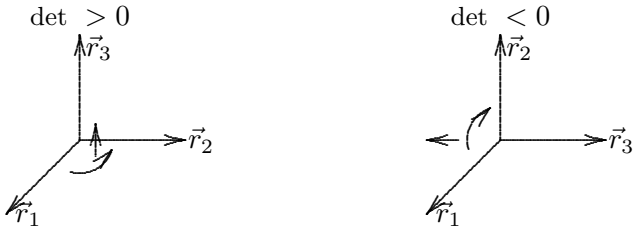
* J. Hannah, A geometric approach to determinants, *Amer. Math. Monthly* **103**, 401–409 (1996).

SKETCH OF PROOF: First imagine straightening up the parallelepiped (like "squaring" the deck of cards). This can be done by adding to each vector $\vec{r}_j$ a linear combination of the other $\vec{r}$s. By the theorem in Sec. 7.1 on the behavior of determinants under row operations, this does not change the determinant; by our previous argument, it does not change the volume, either. Second, the volume of the new rectangular parallelepiped is the product of the lengths of the new vectors that determine it. By another clause of the same theorem, the determinant of these vectors equals the product of their lengths times the determinant of the unit vectors that constitute the natural basis; but that determinant is 1.

We still need to explain the last part of the theorem. Given that determinants are fundamentally volumes, why are some determinants negative? To see why, notice first that all bases for $\mathbf{R}^2$ fall into two classes, depending on whether the second vector in the basis lies on the counterclockwise or the clockwise side of the first vector. (Here we are tacitly identifying the abstract space $\mathbf{R}^2$ with its conventional representation by a plane with a rightward-pointing horizontal $x \equiv x_1$ axis and an upward-pointing vertical $y \equiv x_2$ axis. Thus the natural basis, $\{\hat{\imath}, \hat{\jmath}\}$, is in the counterclockwise, or right-handed, class, by definition.)



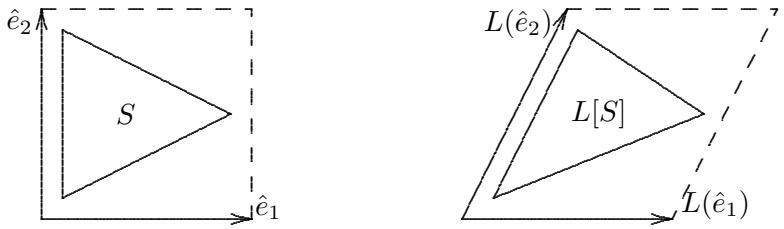
Similarly, in $\mathbf{R}^3$ the natural basis, $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$, is traditionally represented by a "right-handed" triple of vectors: If you sweep the fingers of your right hand from the x direction to the y direction, then your thumb will be pointing approximately in the z ($\equiv x_3$) direction, rather than the $-z$ direction.



Note that a left-handed basis can be converted to a right-handed one by interchanging two basis vectors. By the remaining clause of the theorem about row operations on determinants, this causes a change in the sign of the determinant of the basis vectors. This allows handedness to be defined in any dimension, and without reference to a graphical representation: A basis is right-handed if its determinant is positive, or, equivalently, if it is obtained by acting on the natural basis with a linear transformation $\mathbf{R}^n \to \mathbf{R}^n$ whose matrix has positive determinant; the others are left-handed.

The theorem can be rephrased this way: *The volume of an n-dimensional parallelepiped in $\mathbf{R}^n$ equals the absolute value of the determinant of the linear function that maps the elements of the natural basis for $\mathbf{R}^n$ onto the edge vectors of the parallelepiped.* Since an arbitrary region can be thought of as built out of infinitesimal parallelepipeds, the following generalized theorem is plausible (and true):

Theorem: If L is the linear function from $\mathbf{R}^n$ to $\mathbf{R}^n$ represented with respect to the natural basis by A, and S is a set in $\mathbf{R}^n$, then the volume of the image $L[S]$ equals $|\det A|$ times the volume of S.



If we accept this geometrical interpretation of determinants as the scaling factors of volumes under linear transformations, then it becomes obvious that the identity

$$\det(AB) = (\det A)(\det B)$$

must be true; whether we perform a linear mapping in one step or two, the volumes must dilate by the same factor.

Two vectors in $\mathbf{R}^3$ determine a parallelogram. The 3-dimensional volume of the parallelogram is, of course, zero, but (unless the two vectors are dependent) the parallelogram has a nonzero area as a region in the plane determined by the vectors. Such lower-dimensional volumes are also described by determinants:

Volume Theorem:

(1) In $\mathbf{R}^3$, the area of the parallelogram determined by two vectors $\vec{r}_1$ and $\vec{r}_2$ is the length (norm) of their cross product, $\|\vec{r}_1 \times \vec{r}_2\|$.

(2) For p vectors in $\mathbf{R}^n$, the p-dimensional volume of their parallelepiped is

$$|\det\{\vec{r}_j \cdot \vec{r}_k\}|^{\frac{1}{2}}.$$

(Here $\{\vec{r}_j \cdot \vec{r}_k\}$ is the matrix whose jk element is $\vec{r}_j \cdot \vec{r}_k$.)

REMARK: We omit the proof of (2). To see that (1) is a special case of (2), recall Exercise 2.5.10(a). To see that (1) is true, recall Exercise 2.5.10(b) and the elementary formula for the area of a parallelogram.

CAN THE CROSS PRODUCT BE GENERALIZED TO OTHER DIMENSIONS?

The beginning student of vectors must take on faith that the vector cross product is a specifically *three-dimensional* construction. In view of the great significance of the cross product in three-dimensional physics and geometry, one must wonder "what happens to it" in other dimensions. There are at least two different answers, which we'll examine in turn.

Theorem: In dimension 3, the determinant of three vectors $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ equals their *scalar triple product*, $\vec{x}_1 \cdot (\vec{x}_2 \times \vec{x}_3)$.

PROOF: Consider taking the dot product of $\vec{x}_1$ with the cross product as defined by the formula

$$\vec{x}_2 \times \vec{x}_3 = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{vmatrix}.$$

It is easy to see that the result is equivalent to substituting the coordinates of $\vec{x}_1$ in place of the unit vectors in the top row. This is the determinant of the three vectors (since it doesn't matter whether the vectors are written as rows or columns).

This principle extends immediately to n-dimensional space, once we realize that the appropriate number of vectors to deal with is $n-1$, not 2. Given $n-1$ vectors in $\mathbf{R}^n$, labeled $\{\vec{x}_2, \ldots, \vec{x}_n\}$, we can form a new vector by the determinantal construction

$$\perp(\vec{x}_2, \ldots, \vec{x}_n) \equiv \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \cdots & \hat{e}_n \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & & & \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix}.$$

This vector is orthogonal to each of the input vectors; it changes sign if two of those vectors are interchanged; it is linear as a function of each of them (multilinear as a function of all the vectors together). These properties follow from the corresponding properties of determinants, and they generalize the familiar properties of the cross product of 2 vectors in $\mathbf{R}^3$. Finally, taking the inner product with a final vector, $\vec{x}_1$, we get the determinant of the n vectors $\{\vec{x}_1, \ldots, \vec{x}_n\}$.

We can make use of this construction not only for $n \geq 4$ but also for $n = 2$. In that case we start with just one vector,

$$\vec{r}_2 \equiv (x_2, y_2) = x_2\hat{\imath} + y_2\hat{\jmath}.$$

The associated vector is

$$\perp(\vec{r}_2) \equiv \begin{vmatrix} \hat{\imath} & \hat{\jmath} \\ x_2 & y_2 \end{vmatrix} = y_2\hat{\imath} - x_2\hat{\jmath} = (y_2, -x_2),$$

and the associated inner product with $\vec{r}_1 = (x_1, y_1)$ is

$$\vec{r}_1 \cdot [\perp(\vec{r}_2)] = x_1y_2 - y_1x_2 = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix},$$

which is (up to sign) the area of the parallelogram formed by the two vectors. Notice that if we visualize our two 2-dimensional vectors as lying in 3-dimensional space, by adding on vanishing z components,

$$\vec{r}_1 = (x_1, y_1, 0), \qquad \vec{r}_2 = (x_2, y_2, 0).$$

then this quantity is the nonvanishing component of their cross product:

$$\vec{r}_1 \times \vec{r}_2 = \{\vec{r}_1 \cdot [\perp(\vec{r}_2)]\}\hat{k}.$$

Finally, recall that *complex numbers* can be thought of as pairs of real numbers:

$$z = x + iy = (x, y).$$

(Here i stands for $\sqrt{-1}$ and has nothing to do with the unit vector $\hat{\imath}$; it corresponds more to the vector $\hat{\jmath}$, instead.[†]) *Multiplication* of complex numbers

[†] Ironically, electrical engineers do write the square root of -1 as j, but not for this reason.

is defined by the usual distributive and associative laws together with the basic identity

$$i(x, y) = i(x + iy) = -y + ix = -\perp(x, y).$$

Thus the questions "Where does the cross product come from, and what happens to it in dimensions other than 3?" are intimately related to the questions "Where do complex numbers come from, and what happens to them in dimensions other than 2?", and to some extent these questions answer each other.

The cross product unmasked

The foregoing discussion can be summarized rather flippantly as saying that the cross product is defined only in dimension 3 because 3 is the only number n with the property that $n - 1 = 2$. A strong case can be made, however, that a more fundamental property of the integer 3 is

$$\frac{n(n-1)}{2} = n.$$

The point is that the number on the left of this equation is the dimension of the vector space of $n \times n$ antisymmetric matrices (Exercise 5.1.7).

Theorem: In dimension 3 (and only there!) there is a one-to-one correspondence (isomorphism) between *vectors* and *antisymmetric matrices*:

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \longleftrightarrow \begin{pmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{pmatrix}.$$

The linear function represented by the matrix is

$$L(\vec{r}) \equiv \vec{v} \times \vec{r}.$$

Furthermore, if we define an antisymmetric matrix from two vectors, $\vec{u}$ and $\vec{w}$, by

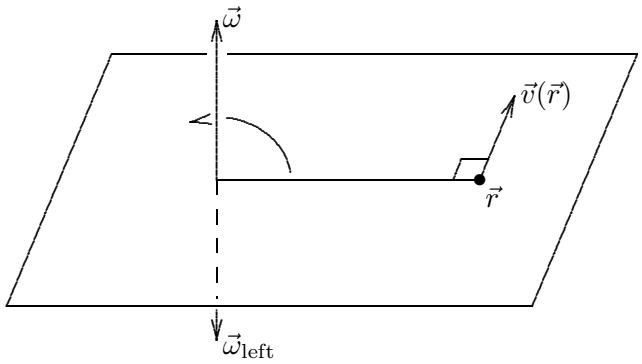
$$A_{jk} = u_k w_j - u_j w_k \,,$$

then the vector $\vec{v}$ corresponding to A is $\vec{u} \times \vec{w}$.

This is the true meaning of the vector cross product, and the best explanation of why the cross product exists only in three dimensions.

Whenever you see a cross product in a physical application, *there is an antisymmetric matrix hiding in the problem somewhere, disguised as a vector*. Such a "pseudovector" can be recognized by the fact that its *sign* is a matter of convention — it depends on the handedness of the coordinate system.

Examples of physical quantities represented by pseudovectors are *magnetic field*, *angular velocity*, *angular momentum*, and *torque*. All of these have to do with some kind of *turning* in the plane orthogonal to themselves. This is not surprising if you recall from Sec. 6.3 (Exercises 6.3.1–2) that an antisymmetric matrix is related to the derivative of a *rotation* matrix with respect to the angle of rotation. In the case of angular velocity $\vec{\omega}$ of a rotating rigid body, $\vec{v} = L(\vec{r}) = \vec{\omega} \times \vec{r}$ is the linear velocity of the point on the body located at $\vec{r}$.



Note that if $\vec{A}$ is a true vector, then $\nabla \times \vec{A}$ is a pseudovector, and *vice versa*. For example, the magnetic field $\vec{B}$ is a pseudovector, but according to one of Maxwell's equations,

$$\nabla \times \vec{B} = \frac{\partial \vec{E}}{\partial t} \, ,$$

which is a true vector. More generally, cross products (including curls) tend to occur in physical formulas in pairs, one $\times$ to create a pseudovector and a second one to get rid of it. A quantity of direct physical significance should not have an ambiguous sign! Pseudovectors are merely intermediate calculational conveniences.

For example, the *Lorentz force law* states that the force on a moving

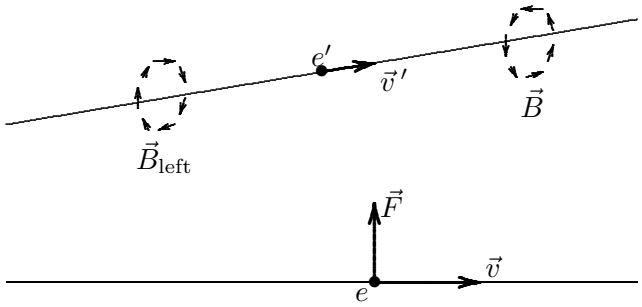
charge e due to a magnetic field[‡] $\vec{B}$ is

$$\vec{F} = e\vec{v} \times \vec{B},$$

where $\vec{v}$ is the velocity of the charge. But $\vec{B}$ itself may be calculated as a cross product. If the field is due to another moving charge, e', with velocity $\vec{v}\,'$, then

$$\vec{B} = e' \, \frac{\vec{v}\,' \times \vec{r}}{4\pi \|\vec{r}\|^3},$$

where $\vec{r}$ is the vector from e' to e. (This leads to the *Biot–Savart law* for the magnetic field surrounding an electric current.) Thus the formula for $\vec{F}$ in terms of the velocities and charges involves a double cross product. If we do the calculation in a left-handed coordinate system instead of a standard right-handed one, then $\vec{B}$ will change sign, but $\vec{F}$ (the quantity actually measured in the lab) will be unchanged.



Exercises

7.2.1 (a) Find the area of the parallelogram with edges $\vec{v}_1 = (2, 2)$ and $\vec{v}_2 = (1, -1)$.

(b) Do the vectors $\vec{v}_1$ and $\vec{v}_2$ (in that order) form a right-handed basis or a left-handed basis for $\mathbf{R}^2$, or not a basis at all?

7.2.2 (a) Calculate the determinant $\begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 1 \end{vmatrix}$.

(b) Does the set $\{ \vec{v}_1 = (2, 1, 1),\ \vec{v}_2 = (1, 2, -1),\ \vec{v}_3 = (1, -1, 1) \}$ form a left-handed or a right-handed basis for $\mathbf{R}^3$, or neither?

‡ Here we are talking about charges moving in a vacuum, and we use "natural" units (not SI), so there is no distinction between $\vec{B}$ and $\vec{H}$, and there are no factors of c or μ_0 in the equations.

7.2.3 Find the volume of the parallelepiped with the given vectors as edges.

(a) $\vec{v}_1 = (2,1,2),\ \vec{v}_2 = (1,1,1),\ \vec{v}_3 = (0,1,1)$.

(b) $\vec{v}_1 = (2,1,2),\ \vec{v}_2 = (3,1,5),\ \vec{v}_3 = (4,1,4)$.

7.2.4 Calculate the area of the parallelograms (in $\mathbf{R}^n$) with the given vectors as edges.

(a) $\vec{v}_1 = (1,2,1),\ \vec{v}_2 = (2,1,0)$.

(b) $\vec{v}_1 = (1,2,3,4),\ \vec{v}_2 = (1,1,1,1)$.

7.2.5

(a) Calculate the determinant $\begin{vmatrix} 0 & -2 & 0 \\ 4 & 1 & 4 \\ 2 & 0 & 1 \end{vmatrix}$.

(b) Let $L\colon \mathbf{R}^3 \to \mathbf{R}^3$ be represented by the matrix $\begin{pmatrix} 0 & -2 & 0 \\ 4 & 1 & 4 \\ 2 & 0 & 1 \end{pmatrix}$, and let S be the parallelepiped in Exercise 7.2.3(a). What is the volume of the image set $L[S]$?

7.2.6 Show that the determinant of a shear transformation (see Exercise 3.2.23) is always 1.

7.2.7 Show how the deformation of a deck of cards into a nonrectangular parallelepiped can be described by a 3-dimensional analogue of the shear transformation treated in Exercises 3.2.23 and 7.2.6.

7.2.8 Determinants are defined for *matrices*, but at some points in this section we have referred to the determinant of a *linear function*, thereby tacitly assuming that the determinant does not depend on which matrix representation of the function is used to calculate the determinant.

(a) Show that this assumption is justified for a linear function $L\colon \mathcal{V} \to \mathcal{V}$, with the understanding that the basis used for $\mathcal{V}$ in its role as domain is the same as the basis used for $\mathcal{V}$ in its role as codomain.

(b) Show why the assumption is *not* justified for a linear function $L\colon \mathcal{V} \to \mathcal{W}$, where $\mathcal{V}$ and $\mathcal{W}$ are two unrelated vector spaces that simply happen to have the same dimension (so that the determinants of the matrix representations of L are defined).

7.2.9 We know that when orthonormalizing a basis for $\mathbf{R}^3$ it is not necessary to find the last basis vector by the Gram–Schmidt algorithm: It is

more efficient just to take the cross product of the first two unit vectors. Describe precisely the corresponding procedure for $\mathbf{R}^4$.

7.2.10 Prove the two parts of the theorem unmasking the cross product.

7.3 Jacobi's Theorem: Changing Variables in Multiple Integrals

The determinant of the Jacobian matrix of a coordinate transformation is traditionally called just "the Jacobian" of the transformation. If $(x, y) = T(u, v)$, then an old notation for the Jacobian $\det T'$ is $\dfrac{\partial(x, y)}{\partial(u, v)}$. For example, in the most common curvilinear coordinate systems we have:

Polar coordinates: $$\frac{\partial(x, y)}{\partial(r, \theta)} = r.$$

Cylindrical coordinates: $$\frac{\partial(x, y, z)}{\partial(r, \theta, z)} = r.$$

Spherical coordinates: $$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = r^2 \sin\theta$$

(see Exercise 7.3.6). In third-semester calculus courses precisely these expressions are presented as the "weight factors" or "volume elements" that must be inserted into the integrand when a double or triple integral is evaluated in terms of the curved coordinates. In the calculus books these formulas are derived by looking at the geometry of the infinitesimal region generated by small changes Δr, ... in the curvilinear coordinates. Jacobi's theorem gives a more general and systematic approach to the change of variables in multiple integrals, less dependent on spatial visualization.

Jacobi's theorem states that the absolute value of the Jacobian determinant is the weight factor that comes in when a multiple integral is rewritten in terms of a new coordinate system. That is, just as one-dimensional integrals obey the rule

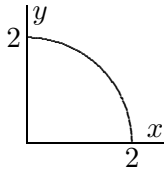
$$\int \dots \, dx = \int \dots \left|\frac{dx}{du}\right| du,$$

multidimensional integrals obey

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

(The one-dimensional formula is usually written without the absolute-value bars. In that case, a coordinate change with negative derivative produces an integral with limits in the abnormal order (the smaller number at the top); restoring the normal order causes a sign change equivalent to taking the absolute value of the derivative.)

An elementary example is the integration of a function over a quarter-circle:



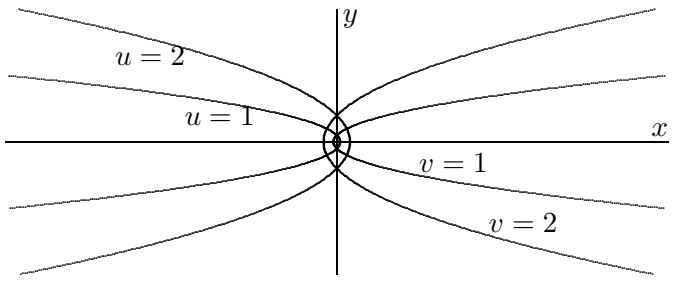
$$\int_{x=0}^{2} \int_{y=0}^{\sqrt{4-x^2}} (x^2 + y)\, dy\, dx$$
$$= \int_{\theta=0}^{\pi/2} \int_{r=0}^{2} (r^2 \cos^2\theta + r \sin\theta)\, r\, dr\, d\theta.$$

Here the $|\det T'|$ is the final factor r.

Here's a more unusual example: *Parabolic coordinates* are defined by

$$x = \frac{1}{2}(u^2 - v^2), \qquad y = uv.$$

The coordinate curves (or coordinate surfaces — they are the same thing in two dimensions) are parabolas with focus at the origin.



(a) Let's calculate $\dfrac{\partial(x,y)}{\partial(u,v)}$:

$$\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} u & -v \\ v & u \end{vmatrix} = u^2 + v^2.$$

(Recall that the vertical bars in this context merely indicate a determinant, not an absolute value. But our result is nonnegative, so it will not be necessary to take its absolute value when we use it in an integral.)

(b) Let's calculate $\dfrac{\partial(u,v)}{\partial(x,y)}$ as a function of u and v: The inverse-function theorem (Sec. 5.5) is designed for this. Since $\bigl(T^{-1}\bigr)' = (T')^{-1}$, we have

$$\det\Bigl[\bigl(T^{-1}\bigr)'\Bigr] = (\det T')^{-1} = \frac{1}{u^2+v^2}\,.$$

(c) Let's calculate the area of the region bounded by the four curves

$$u=1, \qquad u=2, \qquad v=1, \qquad v=2.$$

(Looking at the sketch, we notice that there are two such regions (or the region consists of two disconnected parts, depending on your terminology). We'll calculate the area of the upper region.) The area is $\iint 1\,dx\,dy$ where the upper and lower limits on the x integral are certain functions of y which we shall avoid calculating, and the limits on the y integral are certain numbers (coordinates of the intersections of the curves) which we shall also avoid calculating. Instead, we write

$$\begin{aligned} A &= \int_1^2 \int_1^2 \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv \\ &= \int_1^2 \int_1^2 (u^2+v^2)\,du\,dv = 2\int_1^2 \int_1^2 u^2\,du\,dv \\ &= \frac{14}{3} \quad \text{square units} \end{aligned}$$

(where the "units" are those in which x and y are measured). At the middle step we exploited the symmetry beween u and v to reduce two integrals to one.

General statement of Jacobi's theorem

1. Analytical point of view:

Let $T\colon \mathbf{R}^n \to \mathbf{R}^n$ be a coordinate transformation: $\vec{x} = T(\vec{u})$. Let $f\colon \mathbf{R}^n \to \mathbf{R}$ be a function to be integrated over a set X of values of the

Cartesian coordinates $\vec{x}$. Then $f \circ T$ is a function of $\vec{u}$, and the theorem states that

$$\int_X f(\vec{x})\, d^n x = \int_{T^{-1}(X)} \big(f \circ T\big)(\vec{u})\, |\det T'|\, d^n u,$$

where $T^{-1}(X)$ is the set of points in $\vec{u}$-space which is mapped into the set X in $\vec{x}$-space by the coordinate transformation.

The theorem carries two technical hypotheses:

(1) T is one-to-one from $T^{-1}(X)$ onto X.

(2) T' is defined and invertible (i.e., $\det T' \neq 0$) throughout $T^{-1}(X)$.

Although related, these are separate conditions. They are satisfied by any "good" coordinate transformation, as long as you don't try to use it in too large a region. For example, the polar coordinate system breaks down at the origin, where the Jacobian r equals 0; the Jacobian matrix of the inverse transformation is undefined there. If we cut out the origin and look at a ring-shaped region $R_1 < r < R_2$, the Jacobian matrix will be invertible throughout, but the coordinate mapping is not one-to-one: $\theta + 2\pi$ and θ label the same points. So, to get a good coordinate mapping we must cut out a path across the ring. (Nevertheless, Jacobi's theorem can still be used to evaluate integrals of smooth functions over a circle: Integrate over the keyhole-shaped region that is left after the surgery just described, then take a limit.)

2. Geometrical or physical point of view:

There is *only one function* in the problem; f and $f \circ T$ are *two representations* of it. We are really dealing with some abstract space M, which has two different coordinate systems defined on [a part of] it, mapping [some of] its points into elements of $\mathbf{R}^n$, or vice versa. (Example: When we calculate the mass of physical object by integrating its density over the region it occupies, the points in that region of physical space are not *really* triples of real numbers; and that is why we can change coordinates and represent them by different triples.) Let's use p as a variable standing for points in M. Then each coordinate system is a mapping of $\mathbf{R}^n$ into M,

$$p = \alpha(\vec{u}) = \beta(\vec{x}).$$

Therefore,

$$\vec{x} = \big(\beta^{-1} \circ \alpha\big)(\vec{u}); \qquad T = \beta^{-1} \circ \alpha.$$

Now if P is a set of points p in M, the integral $\int_P f(p)\,dp$ is defined in the abstract (by hypothesis) and can be calculated in either system.

$$\int_P f(p)\,dp = \int_{\alpha^{-1}(P)} \big(f\circ\alpha\big)(\vec{u})\,\sigma(\vec{u})\,d^n u = \int_{\beta^{-1}(P)} \big(f\circ\beta\big)(\vec{x})\,\rho(\vec{x})\,d^n x$$

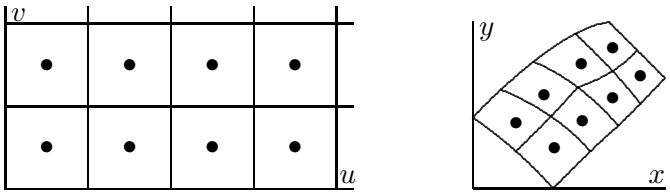
for some weight functions σ and ρ.

If M is flat physical space and $\vec{x}$ is the Cartesian coordinates, we usually identify p with $\vec{x}$, so that β is the identity mapping, $P = X$, $\rho \equiv 1$, $T = \alpha$, and $\sigma = |\det T'|$. Then the equation becomes Jacobi's theorem,

$$\int_P f(p)\,dp = \int_{\alpha^{-1}(P)} \big(f\circ\alpha\big)(\vec{u})\,|\det T'|\,d^n u = \int_X f(\vec{x})\,d^n x$$

More generally, Jacobi tells us that $d^n x = |\det T'|\,d^n u$, or $\sigma = |\det T'|\rho$.

Outline of the proof of Jacobi's theorem: Recall from the preceding section that a linear function with matrix A carries a parallelepiped with volume V into one with volume $|\det A|\,V$. Cut up $T^{-1}(X)$ into tiny rectangles with centers called $\vec{u}_j$. Take $A = T'(\vec{u}_j)$. Write both integrals as multidimensional Riemann sums and take limits. Remark: It would be circular to use Jacobi's theorem to *prove* the theorems about determinants and volumes in Sec. 7.2. As a practical matter, it is perfectly legitimate to use Jacobi's theorem to evaluate an integral over a parallelepiped; but as a proof, it is cheating.



Exercises

7.3.1 Find the area between the curves $u = 1$ and $u = 2$ in the elliptical coordinate system of Exercise 4.2.1. Sketch those curves and two curves of constant v.

7.3.2 Let $f(x, y) = y^2$ and let D be a region bounded by four curves $u = 1$, $u = 2$, $v = 2$, $v = 3$ in parabolic coordinates (defined in Exercise 4.2.2 or in the text of this section). Write $\int_D f(x, y)\,dx\,dy$ as a double integral with respect to u and v, and evaluate it.

7.3.3 Consider the coordinate system introduced in Exercise 5.5.6.

(a) Verify that the Jacobian is nonzero except at the origin.

(b) Find the area of the region bounded by the four curves $u = 1$, $v = 2$, $u = 3$, $v = 4$.

7.3.4 The equations

$$x = s\cosh\psi, \qquad y = s\sinh\psi$$

define *hyperbolic polar coordinates*, (s, ψ), in the wedge-shaped region of $\mathbf{R}^2$ defined by the inequality $x \geq |y|$. They are similar to ordinary polar coordinates except that lines of constant s are hyperbolas, not circles.

(a) Find the tangent vectors to the coordinate curves and the normal vectors to the coordinate hypersurfaces (which are curves in this two-dimensional situation). (Find all of these in terms of s and ψ. Recall that $\cosh^2\psi - \sinh^2\psi = 1$.) Sketch these vectors at the point $s = 1$, $\psi = 1$, along with some typical coordinate curves near that point.

(b) Calculate the area of the region bounded by the four curves

$$s = 1, \qquad s = 2, \qquad \psi = 0, \qquad \psi = 10.$$

Sketch the region.

7.3.5 In $\mathbf{R}^3$ introduce "slanted" cylindrical coordinates (r, θ, h) by

$$\begin{aligned} x &= r\cos\theta, \\ y &= r\sin\theta, \\ z &= h + r\cos\theta. \end{aligned}$$

(a) Calculate the tangent vectors to the coordinate curves and the normal vectors to the coordinate surfaces.

(b) Calculate the mass of a region bounded by the surfaces $r = 2$, $h = 0$, and $h = 1$, if the mass density is $\rho(r, \theta, h) = 10 + r^2$. (Sausage is sometimes sliced this way in delicatessens.)

7.3.6 Let J be the Jacobian matrix of the spherical coordinate transformation. Fill in the details of the calculation of $\det J$ and complete the calculation of J^{-1}, begun in Sec. 7.1.

7.3.7 A region is described in cylindrical coordinates by the inequalities

$$0 < r < 4, \quad 0 < \theta < 2\pi, \quad 2 < z < 5.$$

(a) Describe the region in words, and sketch it.

(b) The region is filled with electrical charge with density $\rho(r, \theta, z) = (1 + r^2)z$. Find the total charge.

7.3.8 A region is described in spherical coordinates by the inequalities

$$0 < r < 4, \quad 0 < \theta < \pi/2, \quad \pi/2 < \phi < \pi.$$

(a) Describe the region in words.

(b) Calculate the integral of $f(r, \theta, \phi) = z$ over the region (where z is the original Cartesian coordinate).

7.4 Surface Integrals — Definition

Like line integrals, surface integrals come in two varieties, scalar and vector, with the vector integrals being of greater interest. What this means is that the principal element in the *integrand* can be either a scalar function or a vector field; the *value* of the integral is a scalar in either case. Throughout this section we will be working in $\mathbf{R}^3$ (or the three-dimensional physical space that $\mathbf{R}^3$ represents).

Let us dispose of the scalar integrals quickly. In Sec. 6.3 we briefly looked at *scalar line integrals*, which answer such questions as *What is the total weight of a cable whose density is ρ pounds per foot?* Similarly, *scalar surface integrals* answer questions such as *What is the total weight of an automobile fender whose density is ρ pounds per square foot?* Here ρ may be a function of position on the fender. The key step in setting up such an integral is getting an analytic expression for the *surface area* of each infinitesimal piece of the surface (just as scalar line integrals hinge on finding the infinitesimal arc length). When that has been done, the correct integral expression for the weight (for instance) is obtained by inserting the density function as a factor in the integrand:

$$W = \iint_{\text{surface } S} \rho \, dS.$$

When we know how to express the *element of surface area*, dS, in terms of differentials of two coordinates on the surface, the integral is reduced to an ordinary double integral in the coordinate plane.

We postpone discussing how to get dS in general, but there are several geometrically simple surfaces for which the surface-area element is well known:

Sphere of radius R: $$dS = R^2 \sin\theta \, d\theta \, d\phi.$$

(Recall that in this book θ is the colatitude, measured down from the north pole.)

Cylinder of radius R: $$dS = R \, d\theta \, dz.$$

Plane normal to the y axis: $$dS = dx \, dz$$

(and similarly for planes normal to any other coordinate axis).

Plane normal to $(\cos\alpha, 0, \sin\alpha)$*:* $$dS = \sec\alpha \, dx \, dy$$

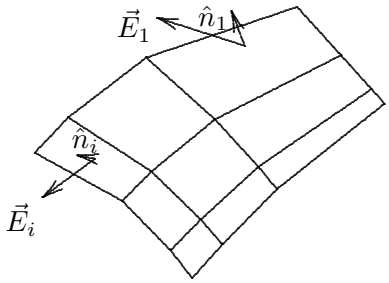
(see drawing near end of Sec. 7.6).

Surface integrals of vector fields,

$$I = \iint_S \vec{A}(\vec{r}) \cdot d\vec{S},$$

are much more interesting, since they have a very special relationship with the vector differential operations, expressed in the famous integral theorems of Green, Gauss, and Stokes. Our primary attention will be devoted to them. Normally, the terms "line integral" and "surface integral" refer to these vectorial integrals.

Such a surface integral answers a question such as *What is the flux of an electric field* $\vec{E}(\vec{r})$ *(or of the current of mass density in a fluid) through the surface S?* To answer that question, one would take the numerical value of the component of $\vec{E}(\vec{r})$ which is perpendicular to the surface, and "add it up" over the whole surface. More precisely: Cut the surface up into little pieces, each so small that it is nearly flat and $\vec{E}$ is nearly constant over it. With each piece (say the ith piece) associate a vector $\Delta\vec{S}_i$, equal to the surface area of the piece (a scalar) times a unit vector normal to that piece. Take



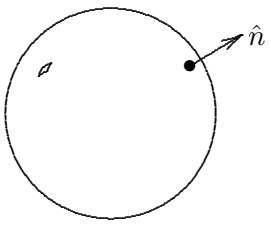
the dot product of $\Delta\vec{S}_i$ with the vector field on that piece, $\vec{E}(\vec{r}_i)$. Finally, take the limit of these *Riemann sums* as the size of the pieces goes to zero:

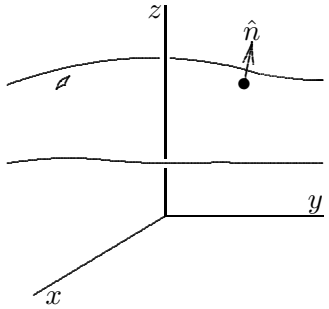
$$I = \iint_S \vec{E}(\vec{r}) \cdot d\vec{S} \equiv \lim_{\|\Delta\vec{S}_i\| \to 0} \sum_i \vec{E}(\vec{r}_i) \cdot \Delta\vec{S}_i \, .$$

This prescription contains an ambiguity: There are actually *two* unit normals at each point of the surface (pointing in opposite directions), and therefore part of the definition of the surface integral is a choice between them. In other words, just as the curve in a line integral must have a definite orientation (a distinction between the positive and the negative *direction* along it), so must a surface in $\mathbf{R}^3$ have a definite orientation (a distinction between its positive and its negative *side*). The following conventions are fairly standard: Let $\hat{n}(\vec{r})$ stand for the (still to be chosen) unit normal vector to S at the point $\vec{r}$. Then

(1) If S is *closed* (so that it is the boundary of a three-dimensional region), choose $\hat{n}$ to be the *outward* normal vector (so that $\int_S \vec{A} \cdot d\vec{S}$ is the flux of the vector field $\vec{A}$ *out* of the region).

(2) If S is *the graph of a function*, $z = f(x, y)$, choose $\hat{n}$ to be the *upward* normal. (See Exercise 7.4.5 for one prescription for calculating it.)





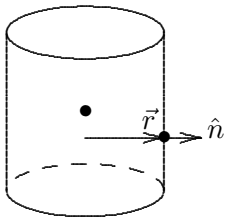
In more general situations, the normal vector must be explicitly specified. Naturally, one chooses it to be continuous over the surface, rather than jumping from one side to the other. (And one doesn't try to define the electrical flux through a Möbius strip, in case you're wondering.)

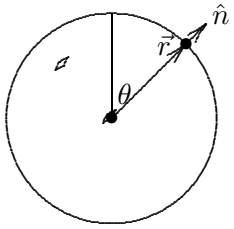
Again, for basic surfaces such as planes, cylinders, and spheres the normal vector at each point is taken to be well known. (In Sec. 7.6 we will give a prescription for calculating the normal to a more general parametrized surface.) Normals to planes were treated back in Sec. 1.2. On a cylinder around the z axis with radius R, the unit normal vector at the point $\vec{r}$ with cylindrical coordinates (R, θ, z) has Cartesian components $(\cos\theta, \sin\theta, 0)$, or

$$\hat{n} = \cos\theta\,\hat{\imath} + \sin\theta\,\hat{\imath}.$$

This is the same thing we called $\hat{r}$, the normal vector to the coordinate surface $r =$ constant, in Secs. 4.2 and 5.5. Geometrically, it is the unit vector in the direction of $\vec{r}_\perp$, the projection of $\vec{r}$ onto the x-y plane. Similarly, on a sphere centered at the origin with radius R, the unit normal vector at the point with spherical coordinates (R, θ, ϕ) is

$$\hat{n} = \frac{\vec{r}}{\|\vec{r}\|} = \sin\theta\cos\phi\,\hat{\imath} + \sin\theta\sin\phi\,\hat{\jmath} + \cos\theta\,\hat{k} = \hat{r}.$$





The upshot of this discussion is that the surface integral through a patch of surface, S, of a vector field, $\vec{A}(\vec{r})$, can be regarded as the scalar surface integral over S of the function $\vec{A}\cdot\hat{n}$:

$$\iint_S \vec{A}(\vec{r})\cdot d\vec{S} = \iint_S \vec{A}(\vec{r})\cdot\hat{n}(\vec{r})\,dS. \tag{G}$$

Another point of view, however, is to absorb $\hat{n}$ into the differential instead of the integrand. As we'll see in Sec. 7.6, when this is written out in Cartesian coordinates it amounts to interpreting $d\vec{S}$ formally as $\hat{\imath}\,dy\,dz + \hat{\jmath}\,dz\,dx + \hat{k}\,dx\,dy$:

$$\iint_S \vec{A}(\vec{r})\cdot d\vec{S} = \iint_S (A_x\,dy\,dz + A_y\,dz\,dx + A_z\,dx\,dy). \tag{A}$$

Each term then has meaning as an ordinary double integral, when the missing variable is expressed in terms of the other two via the equation of the surface (e.g., $A_x = A_x\big(x(y,z),y,z\big)$). Examples of setting up and evaluating such integrals in practice, either directly or after reinterpreting them as parameter integrals analogous to formula (P) for line integrals, will appear in Sec. 7.6.

However, the most salient fact about surface integrals of vector fields (in undergraduate physics and engineering, at least) is that one can usually *avoid* evaluating them directly, instead converting them to volume integrals (by Gauss's theorem) or line integrals (by Stokes's theorem), so we want to concentrate on those techniques first (see Sec. 7.5). Just as in Sec. 6.3 the examples and exercises concentrated on the arc-length formulation of line integrals, in this section we will consider only examples where the surface integral can be evaluated easily by the geometrical (surface-area integral) formula, (G). What was said there about line integrals is even more true here: Problems of this sort are either those with very simple and familiar surfaces, for which we know the surface area and normal vectors, or those with $\vec{A}\cdot\hat{n}$ constant, so that we can evaluate the integral just by multiplying by the total area of the surface.

Example 1: Find the flux of the vector field $\vec{A}(\vec{r}) = \vec{r} + \hat{k}$ through the sphere of radius 3 centered at the origin.

SOLUTION: It is most instructive to study the two terms in $\vec{A}$ separately. The element of surface area is $dS = 9\sin\theta\,d\theta\,d\phi$, and the normal vector is $\hat{n} = \frac{\vec{r}}{r} = \frac{1}{3}\vec{r}$. Thus for the first term the integrand is $\hat{n}\cdot\vec{r} = \frac{1}{3}r^2 = 3$. Integrating this over the sphere amounts to multiplying by the total surface area, $4\pi r^2 = 36\pi$. For the second term we have $\hat{n}\cdot\hat{k} = z = 3\cos\theta$, so the integral is

$$\int_0^{\pi} d\theta \int_0^{2\pi} d\phi\,(3\cos\theta)(9\sin\theta)\,d\theta\,d\phi.$$

This integral is actually zero, as can be seen either by performing the ϕ integration, or more physically by realizing that the flux through the bottom half of the sphere is exactly equal and opposite to the flux through the top half. Therefore, the total surface integral in this problem is equal to $3\times 36\pi = 108\pi$. (The units are the units of $\vec{A}$ itself, times the square of the unit of length.)

Example 2: Let S be the cylindrical surface

$$r = 1, \quad 0 \le \theta < 2\pi, \quad 0 < z < 3.$$

Calculate $\iint_S \vec{F} \cdot d\vec{S}$ when $\vec{F}(\vec{r}) = x\,\hat{\imath} + y\,\hat{\jmath} + (x^2+y^2)\,\hat{k}$.

SOLUTION: The element of surface area on the cylinder is $dS = r\,d\theta\,dz$, and the normal vector to the cylinder is $\hat{n} = \cos\theta\,\hat{\imath} + \sin\theta\,\hat{\jmath}$ (or $\frac{x}{r}\,\hat{\imath} + \frac{y}{r}\,\hat{\jmath}$). Therefore,

$$\vec{F} \cdot \hat{n} = x\cos\theta + y\sin\theta = r(\cos^2\theta + \sin^2\theta) = r,$$

so the integral is

$$\iint_S (\vec{F} \cdot \hat{n})\,dS = \int_0^{2\pi} d\theta \int_0^3 dz\,\{r^2\} = 2\pi \cdot 3 \cdot 1 = 6\pi.$$

Example 3: Evaluate $\iint_S \vec{B} \cdot d\vec{S}$ when $\vec{B}(\vec{r}) = y\,\hat{\imath}$ and S is the triangle with corners at $(0,0,1)$, $(1,0,0)$, and $(1,1,0)$.

SOLUTION: We can use x and y as coordinates on the surface. (One could equally well use x and z.) Since the plane (in which the triangle lies) sits at a 45° angle, the area element is $dS = \sec\frac{\pi}{4}\,dx\,dy = \sqrt{2}\,dx\,dy$. The normal vector at all points is $\hat{n} = \frac{1}{\sqrt{2}}(\hat{\imath} + \hat{k})$. (This is clear from geometry, but one could also find it by the method of Exercise 7.4.5: Regard the plane as the graph of $z = 1 - x$ and find the gradient $\nabla(x + z - 1) = \hat{\imath} + \hat{k}$; then normalize. Or, we could find two vectors parallel to the plane and take their cross product, as in Sec. 1.2.) Thus the integral is

$$\begin{aligned}\iint_S (\hat{n} \cdot \vec{B})\,dS &= \int_0^1 dx \int_0^x dy \frac{1}{\sqrt{2}}(\hat{\imath} + \hat{k}) \cdot (y\,\hat{\imath})\,\sqrt{2} \\ &= \int_0^1 \int_0^x y\,dy\,dx = \int_0^1 \frac{1}{2}\,x^2\,dx = \frac{1}{6}\,.\end{aligned}$$

The parametric approach to this integral, which is slightly simpler, will be presented in Sec. 7.6.

EUCLIDEAN GEOMETRY VERSUS DIFFERENTIAL FORMS

To summarize, both line integrals and surface integrals of vector fields can be defined in two ways. In each case there is a geometrical construction (labelled "(G)" in this book) that can be thought of as the most fundamental and intuitive from the physical point of view; but calculations based on this definition give the impression of being needlessly roundabout — and become difficult once one goes beyond the simplest geometrical situations.

In each case there is also an algebraic (coordinate-based) formulation (labelled "(A)") that is very convenient when the curve or surface of integration consists of pieces parallel to the coordinate axes or planes, but whose interpretation is somewhat subtle in more general situations. The basic formulas for line integrals are

$$\int_C \vec{A} \cdot d\vec{r} = \int_C (\vec{A} \cdot \hat{T})\, ds = \int_C (A_x\, dx + A_y\, dy + A_z\, dz)\,.$$

Those for surface integrals are

$$\iint_S \vec{A} \cdot d\vec{S} = \iint_S \vec{A} \cdot \hat{n}\, dS = \iint_S (A_x\, dy\, dz + A_y\, dz\, dx + A_z\, dx\, dy).$$

We saw that when C is presented as a parametrized curve, there is another formula for the line integral (labelled "(P)"). The corresponding construction for surface integrals has been postponed to Sec. 7.6, since it requires some machinery that we don't want to build up right now. (P) is the most general formulation, since both (G) and (A) are special cases of it. Furthermore, when one tries to apply prescription (G) or (A) to a particular problem outside the simplest geometrical situations, one is usually led inexorably to a parametric representation of the line or surface and hence to an instance of (P).

The (G) definitions make very essential use of the *inner product* that defines the *Euclidean geometry* of physical three-dimensional space. That is, the (G) formulas and the physical intuition behind them centrally involve the concepts of arc length, surface area, unit vectors, dot products, and perpendicularity. However, it is very striking that the (A) definitions make no overt reference to these Euclidean elements. How does this come about? The key point was observed (for line integrals) in Sec. 6.3: The calculation of the element of arc length involves a factor ($\|\vec{g}\,'\|$ if C is parametrized as $\vec{r} = \vec{g}(t)$) that also appears in the *denominator* in the calculation of the unit tangent vector and therefore cancels from the final formula. In surface integrals, similarly, there is a cancellation between the element of surface area and the denominator in the formula for the unit normal vector, as we will see in detail in Sec. 7.6. It is precisely this cancellation of irrelevant geometrical factors that makes parametric calculations simpler than explicit evaluations of the (G) formulas. For instance, in Example 2 of Sec. 6.3 the (G) calculation involved two useless factors of $\sqrt{5}$ that eventually disappeared, whereas

in the alternative (P) calculation the number $\sqrt{5}$ never appeared in the first place.

These observations suggest that line and surface integrals of vector fields are somehow more fundamental than our notions of length, area, and angle. Indeed, there is a modern mathematical theory of integration of appropriate quantities over p-dimensional hypersurfaces sitting in n-dimensional spaces, called the theory of *differential forms*, which is independent of the inner product, or "metric", of space. The familiar line, surface, and volume integrals in $\mathbf{R}^3$ are special cases of these generalized multidimensional integrals. This theory is necessary for integral calculus in dimensions greater than 3 or in nonflat spaces. For 3-dimensional Euclidean space, however, the classical Gibbs formalism of vector calculus is adequate (perhaps superior).

Despite its apparent independence from the concept of the inner product, a formula like $\int (A_x\,dx + A_y\,dy + A_z\,dz)$ should not be used to calculate a line integral in a nonorthogonal basis (or a non-Cartesian coordinate system) without careful attention to how the components of $\vec{A}$ are defined with respect to that new system of reference. If $\vec{A}$ is a gradient, or some other quantity that naturally transforms as a row vector (see the end of Sec. 4.5), then the naive (A) formula is indeed correct. (Furthermore, the apparent dot product $\vec{A} \cdot \vec{g}\,'$ in the (P) formula is just the matrix product of a row vector with a column vector, showing that (P) also is independent of the inner product.) But if $\vec{A}$ is a column vector, then since $d\vec{r}$ also behaves like a column vector, in a nonorthogonal coordinate system the inner product $\vec{A} \cdot d\vec{r}$ will acquire a nonstandard form as in the example near the end of Sec. 6.1.* Note that the force field $\vec{F}$ that appears in work integrals *is* naturally a row vector; if the force is conservative, then $\vec{F} = -\nabla V$, where V is the potential energy.

In the language of differential forms, a function whose value at each point is a row vector is called a "1-form". Just as the (A) formula for line integrals applies to vector fields whose values are row vectors, the (A) formula for surface integrals applies to "2-forms", which are functions whose values are antisymmetric matrices. In orthogonal coordinates these are related to column-vector fields through the theorem unmasking the vector cross

* If the only change in the basis vectors is an overall change in length scale, then the change in the matrix representing the inner product will be absorbed into a redefinition of the unit of length. After all, we have never specified whether the length of $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ is 1 centimeter or 1 meter.

product (Sec. 7.2). Magnetic fields are the foremost examples of fields that are "naturally" 2-forms without needing to go through such an identification.

Exercises

7.4.1 A tin can has height 15 units and radius 4 units. Write down an integral formula for the total mass of the material in the can, if the density (mass per unit area) is not uniform. HINTS: First choose a suitable coordinate system: A natural choice is to put the center of the bottom of the can at the origin, and the axis of the can along the positive z axis. Now write down an integral in cylindrical coordinates for the mass of each of the three pieces of metal making up the can, in terms of unspecified density functions.

7.4.2 Write down an integral formula for the flux of an electric field through the can in the previous exercise. HINT: Find formulas for the normal vector at every point on the surface, and take the needed dot products with the unspecified electric field, $\vec{E}(\vec{r})$. Cylindrical coordinates are advised. Specify clearly (at least in words) where $\vec{E}$ is to be evaluated in each term of your formula.

7.4.3 Let $\vec{H}(\vec{r})$ be a vector field pointing radially outward, with magnitude equal to r^N, where r is (as usual) the distance of $\vec{r}$ from the origin, and N is a constant. Thus, in spherical coordinates,

$$\vec{H} = r^N \hat{r} = r^{N-1}\vec{r}.$$

Find the flux of $\vec{H}$ through the sphere $r = R$, where R is a parameter. For what values of N, if any, is the integral independent of R?

7.4.4 Let $\vec{H}$ be as in the preceding problem. Find an expression for $\vec{H}$ in Cartesian coordinates, and use it to write down an integral for the flux of H through the square patch of plane surface defined by

$$x = 2, \quad -5 \le y \le 5, \quad -5 \le z \le 5.$$

7.4.5 Let the surface S be the graph of a function, $z = \psi(x, y)$. Construct a unit normal vector to S at each point by considering the (three-dimensional) gradient, ∇G, where $G(x, y, z) = z - \psi(x, y)$.

7.4.6 Find the flux of the constant vector field $\vec{A}(\vec{r}) = \hat{k} = (0, 0, 1)$ through the hemisphere

$$x^2 + y^2 + z^2 = 1, \quad z > 0.$$

7.5 The Integral Theorems of Vector Analysis

The first of this famous family is a fairly obvious generalization of the foremost theorem of elementary calculus:

Fundamental Theorem of Calculus for Line Integrals: Let $\vec{A} = \nabla f$ be a vector field that equals the gradient of a function throughout a region in $\mathbf{R}^n$. Let C be a curve that lies entirely in that region, with initial endpoint $\vec{r}_1$ and final endpoint $\vec{r}_2$. Then

$$\int_C \vec{A} \cdot d\vec{r} = \int_{\vec{r}_1}^{\vec{r}_2} \vec{A} \cdot d\vec{r} = f(\vec{r}_2) - f(\vec{r}_1).$$

REMARK: In particular, the theorem implies that the value of the integral is the same for all curves joining two given endpoints. This is not necessarily true for a vector field $\vec{A}$ that is not equal to the gradient of any function. (Unlike the fundamental theorem for scalar functions, this one is not applicable (even in principle) to every integration problem.)

The celebrated theorems of Green, Gauss, and Stokes* are analogues of the fundamental theorem for higher-dimensional integration regions. We need to start with some remarks about the variety of such regions that might be encountered. The discussion will be loose and intuitive; precise statements and proofs belong in a course on topology.

A curve of the sort described in the fundamental theorem has a *boundary* consisting of two endpoints, $\vec{r}_1$ and $\vec{r}_2$. There are other curves, such

* These three names are convenient monosyllabic labels for the three theorems but should not be taken too literally as accurate and complete history. The history has been reviewed by V. J. Katz, The history of Stokes' theorem, *Math. Mag.* **52**, 146–156 (1979). We need not get into the biographical and nationalistic details here; of greater scientific importance is the fact that the theorems were known and used in coordinate-and-component language for several generations before modern vector notation was developed.

as circles, that do not have boundaries in this sense. Such curves are called *closed*. It is fairly clear (at least if the curve does not intersect itself) that a closed curve in $\mathbf{R}^2$ *is itself the boundary* of some two-dimensional region, and that a closed curve in $\mathbf{R}^3$ can be the boundary of a patch of surface — in fact, that there are many different surfaces that "fill in" the curve in that way. (For example, the earth's equator is the boundary of the northern hemisphere of the earth's surface, the boundary of the southern hemisphere, and the boundary of a disk through the earth's center.)

Similarly, a patch of surface in $\mathbf{R}^3$ may have an edge, or boundary, or it may not. A sphere and a torus (doughnut surface) are both examples of surfaces without boundary, or *closed* surfaces. Again, a closed surface is expected to be the boundary of some three-dimensional region. The sketch used below to illustrate Gauss's theorem shows a three-dimensional region (volume) that has a hole through it (making it doughnut-like) and also a hole inside it (making it hollow). Because of the hollowness, the boundary of that solid region consists of *two* closed surfaces (or one closed surface consisting of two disconnected parts), an inner boundary with the structure (topology) of a sphere and an outer boundary that has the structure of a torus (because of the other hole). Similarly, the boundary of a surface may consist of more than one closed curve, as shown in the sketches used below to illustrate the Stokes and Green theorems.

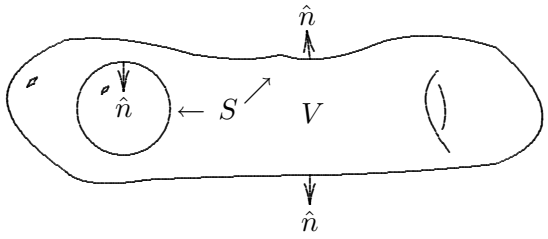
Now we are ready for the statements of the theorems. No proofs will be given here; since this is a book primarily on linear algebra, there is no space or time for them. The interested reader should consult the back third of a standard calculus textbook, or a book on rigorous advanced calculus.

Gauss's Theorem: In $\mathbf{R}^3$, let $\vec{A}(\vec{r})$ be a vector field, S be a *closed* surface, and V be the region interior to S. The normal vector on S points *outward* (away from V). Then[†]

$$\oint_S \vec{A} \cdot d\vec{S} = \int_V (\nabla \cdot \vec{A})\, d^3r.$$

[†] The differential in a three-dimensional volume integral has, according to taste and context, the various notations

$$d^3r = d^3x = dV = dx\, dy\, dz.$$



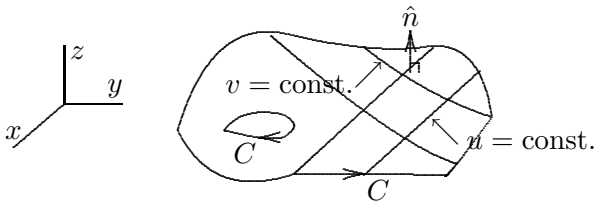
NOTATIONAL REMARK: The circle on the surface integral sign emphasizes the closedness of the surface. When one wants instead to emphasize the dimensions of the integration sets, one writes

$$\iint_S \vec{A} \cdot d\vec{S} = \iiint_V (\nabla \cdot \vec{A})\, d^3r.$$

There is actually a Gauss theorem (also called *divergence theorem*) in any dimension $\mathbf{R}^n$, relating the integral of the divergence (trace of the Jacobian matrix) of a vector field over an n-dimensional region to a flux-like integral of the field itself over the $(n-1)$-dimensional boundary of the region. Indeed, even for $p < n$ there is always such a theorem relating integrals over p-dimensional sets to integrals over their $(p-1)$-dimensional boundaries. When $p = 1$, this is just the vectorial fundamental theorem stated at the beginning of this section. The only other case we can state in the elementary Gibbsian language is $p = 2$, $n = 3$:

Stokes's Theorem: In $\mathbf{R}^3$, let $\vec{A}(\vec{r})$ be a vector field, S be a surface, C be the boundary of S. The directions of C and $\hat{n}$ are related this way: Call the side from which $\hat{n}$ emerges the "top" side of S; if one walks on the top of S along C, then S is on one's *left*. (In particular, the *outer* boundary of S goes around S in the *counterclockwise* direction as seen from the top side.) Then

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



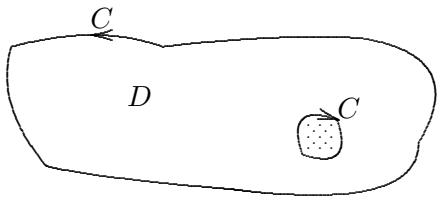
REMARK: In particular, the theorem implies that the value of the surface integral is the same for all surfaces with a given curve as boundary. This is not generally true for a surface integral of a vector field that is not the curl of another vector field.

LINGUISTIC REMARK: Note that the man's name was *Stokes*, not *Stoke*. Therefore there must be an *s before* the apostrophe. The *s* after the apostrophe is optional. (The same is true of *Gauss*, of course.)

An important special case of Stokes's theorem is where C and S lie entirely in a coordinate plane, say the x-y plane. Then $\hat{n} = \hat{k}$, the surface integral is an ordinary double integral, and the component of $\vec{A}$ perpendicular to the plane may be ignored. This is:

Green's Theorem: In $\mathbf{R}^2$, let $\vec{A}(\vec{r})$ be a vector field, D be a region, C be the boundary of D (oriented so that D is on the left as one traverses C). Then

$$\oint_C \vec{A} \cdot d\vec{r} = \iint_D \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) dx\, dy.$$



We shall not prove any of these theorems here, except to remark that they all boil down somehow to applications of the fundamental theorem of calculus to double or triple integrals. Our interest is rather in the geometrical and physical significance of the theorems, and in how they are used either to evaluate certain multidimensional integrals easily, or to elucidate the qualitative features of certain vector fields or certain regions in space.

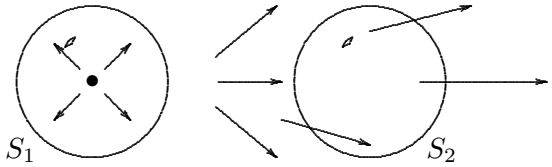
REMARK: In trying to remember and apply these theorems, it may help to keep in mind that the physical dimensions of the two sides of an equation must be the same, and therefore an extra integration on one side must be compensated by an extra differentiation. This means that you should instantly recognize such formulas as

$$\iint_S \vec{F} \cdot d\vec{S} = \oint_C \nabla \times \vec{F} \cdot d\vec{r} \quad \text{or} \quad \iint_S \vec{F} \cdot d\vec{S} = \iint_S \nabla \times \vec{F} \cdot d\vec{S}$$

as **wrong** (although students have been known to write them on tests).

Intuitive significance and integral definition of div and curl

$\nabla \cdot \vec{A}$ is a measure of the tendency of $\vec{A}$ to point outwards. Thus, this is an example of a vector field with positive divergence:

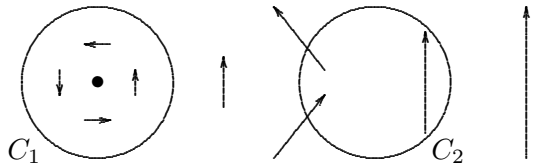


For instance, $\vec{A}(\vec{r}) = \vec{r}$ is a field that behaves in this way. By Gauss, the flux of $\vec{A}$ through any closed surface is positive in such a case. (In the drawing, that is obvious for the sphere on the left, where the flux at every point is positive, and plausible for the sphere on the right, where the flux out the right side is larger than the flux in through the left.) In fact, considering a very *small* surface (say a sphere) around a point $\vec{r}$, enclosing volume ΔV, leads to this formula for $\nabla \cdot \vec{A}$:

$$\nabla \cdot \vec{A} = \lim_{\Delta V \to 0} \frac{\oint_S \vec{A} \cdot d\vec{S}}{\Delta V} \, .$$

This can be used as a *definition* of $\nabla \cdot \vec{A}$; it has the advantage over the partial-derivative definition that it is manifestly independent of the Cartesian coordinate system used for the calculation of the partial derivatives.

Here is the corresponding statement for the curl: Let $\hat{n}$ be any unit vector. Then $\hat{n} \cdot (\nabla \times \vec{A})$ is a measure of the net tendency of $\vec{A}$ to curl (counterclockwise) around the axis $\hat{n}$. For example, this vector field (seen projected on the x-y plane) has a positive $\hat{k} \cdot (\nabla \times \hat{A})$:



For instance, $\vec{F}(\vec{r}) = -y\,\hat{\imath} + x\,\hat{\jmath}$ behaves this way. By Stokes, the line integral around any closed curve in a plane perpendicular to $\hat{n}$ is positive in such a case. (In the drawing, that is obvious for the circle on the left, and plausible for the one on the right.) Letting the curve be a small circle around $\vec{r}$, with area ΔS, we arrive at

$$\hat{n} \cdot (\nabla \times \vec{A}) = \lim_{\Delta S \to 0} \frac{\oint_C \vec{A} \cdot d\vec{r}}{\Delta S} \, .$$

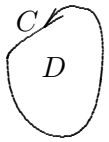
Indirect methods of evaluating line and surface integrals

The integral theorems can be used to convert one type of integral into another. Often the new type is simpler, even trivial, to evaluate. In other cases, the reformulation of the integral has an important physical interpretation; for example, "Gauss's Law" in physics relates the flux integral of the electric field through a closed surface to the volume integral of the electric charge over the region enclosed by the surface.

We shall give here a parallel discussion of line integrals and surface integrals. In each case, there are circumstances where the integral can be moved to a lower-dimensional set and other circumstances where it can be profitably moved to a higher-dimensional set.

A. *Line integrals* $I = \int_C \vec{A} \cdot d\vec{r}$

1. If C is *closed*, then the Green (or Stokes) theorem converts I to an area (or surface) integral over a region (or surface) bounded by C. (Of course, there is no guarantee that the surface integral will be easier to evaluate than the line integral; sometimes the contrary is true.) The integrand is $\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}$ in the 2-dimensional (Green) case, and $(\nabla \times \vec{A}) \cdot \hat{n}$ in the 3-dimensional (Stokes) case.



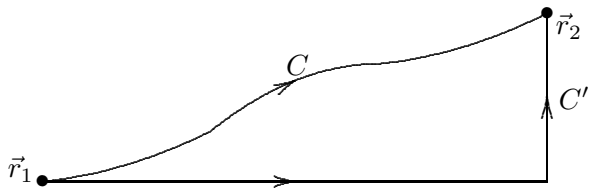
Example: Let $\vec{A} = y\hat{\imath} + x^2\hat{\jmath}$ and let C be the unit circle in the x-y plane. Then

$$\int_C \vec{A} \cdot d\vec{r} = \iint (2x - 1)\, dx\, dy$$
$$= \int_0^1 dr \int_0^{2\pi} d\theta\, r(2r\cos\theta - 1) = -2\pi \left[\frac{1}{2}\, r^2\right]_0^1 = -\pi.$$

2. If $\nabla \times \vec{A} = \vec{0}$ $\left(\text{in dimension 2, if } \frac{\partial A_y}{\partial x} = \frac{\partial A_x}{\partial y}\right)$, then:

   (a) The Stokes (or Green) theorem implies that the integration can be moved to a more convenient curve (e.g., a sequence of segments parallel to the axes) with the same endpoints. (It is essential that $\nabla \times \vec{A}$ be defined and zero *everywhere* on a surface

filling in between the two curves.)



EXAMPLE: Let $\vec{A} = y\,\hat{\imath} + x\hat{\jmath}$ and let C be the curve $y = x^8 + \frac{1}{100}\sin(\pi x)$ $(0 < x < 1)$. The endpoints of this curve, as you can easily check, are $(0,0)$ and $(1,1)$. Then

$$\int_C \vec{A}\cdot d\vec{r} = \int_{C_1} \vec{A}\cdot d\vec{r} + \int_{C_2} \vec{A}\cdot d\vec{r},$$

where C_1 is the straight path from $(0,0)$ to $(0,1)$ and C_2 is the straight path from $(0,1)$ to $(1,1)$. Write out the C_2 integral as

$$\int_{C_2} y\,dx + x\,dy.$$

On C_2 we have $x = 1$ and $dx = 0$ (cf. Example 1, Sec. 6.3), so the first term vanishes and the second is $\int_0^1 dy = 1$. On C_1 we have $y = 0$ and $dy = 0$, so *both* terms vanish. Thus the total line integral equals 1. (A three-dimensional example of this procedure will come up later in this section.)

(b) If $\vec{A} = \nabla f$ for some scalar function f, then

$$\int_{\vec{r}_1}^{\vec{r}_2} \vec{A}\cdot d\vec{r} = f(\vec{r}_2) - f(\vec{r}_1),$$

by the vectorial version of the fundamental theorem of calculus. In particular, the integral is *the same for all paths* C joining $\vec{r}_1$ and $\vec{r}_2$. (The relationship between the two conditions $\nabla\times\vec{A} = \vec{0}$ and $\vec{A} = \nabla f$ will be discussed in depth later; they are almost, but not quite, equivalent.)

EXAMPLE: If $f(x,y,z) = x^2 + 4y^2 - 2z^2$, then $\nabla f = 2x\,\hat{\imath} + 8y\,\hat{\jmath} - 4z\,\hat{k}$. Let C be *any* curve from $\vec{r}_{\rm i} = (1,1,1)$ to $\vec{r}_{\rm f} = (4,-1,1)$. Then

$$\begin{aligned}\int_C (2x\,\hat{\imath} + 8y\,\hat{\jmath} - 4z\,\hat{k})\cdot d\vec{r} &= f(4,-1,1) - f(1,1,1)\\ &= (16+4-2) - (1+4-2) = 15.\end{aligned}$$

3. In particular, C closed *and* $\nabla \times \vec{A} = \vec{0}$ everywhere inside it imply together that $\oint_C \vec{A} \cdot d\vec{r} = 0$.

EXAMPLE: The line integral of $\vec{F} = x^3\,\hat{\imath} + y^3\,\hat{\jmath} + z^3\,\hat{k}$ around any closed Rapidrudder pseudolemniscate vanishes. You do not know or care what a Rapidrudder pseudolemniscate looks like. (A more realistic application: The force $\vec{F}$ does no work on a particle that travels on *any* closed orbit. We need not solve the equations of motion for the orbit in order to know this.)

B. *Surface integrals* $I = \iint_S \vec{A} \cdot d\vec{S}$

1. If S is *closed*, then Gauss's theorem converts I to a volume integral over the region bounded by S:

$$I = \iiint_V \nabla \cdot \vec{A}\, d^3r.$$

This is an ordinary triple integral in $\mathbf{R}^3$ with a scalar integrand.

EXAMPLE: Let $\vec{F}(\vec{r}) = x\,\hat{\imath} + y\,\hat{\jmath} + (x^2 + y^2)\,\hat{k}$. Let's calculate $\iint_S \vec{F} \cdot d\vec{S}$, where the surface S is the "soup can" consisting of the cylinder

$$r = 1, \quad 0 \le \theta < 2\pi, \quad 0 < z < 3$$

and the top and bottom disks

$$0 \le r < 1, \quad 0 \le \theta < 2\pi, \quad z = 0 \text{ or } 3.$$

(The normal vector points outward, and (r, θ, z) are the standard cylindrical coordinates.) Well, in Sec. 6.4 we calculated the divergence of this vector field to be equal to 2 at all points. Therefore, by Gauss's theorem this integral is 2 times the volume of the can:

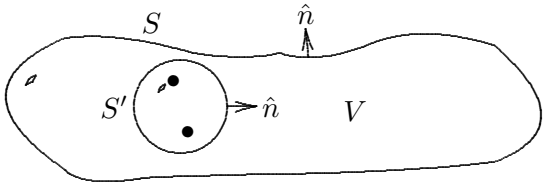
$$2 \cdot \pi r^2 \cdot h = 2 \cdot \pi \cdot 3 = 6\pi.$$

Incidentally, this provides us with an alternative way of solving Example 2 of Sec. 7.4, which dealt with the *open* surface consisting of the cylindrical part of the can alone: Since F_z doesn't depend on z, the fluxes through the top and the bottom of the can exactly cancel, so the flux through the open cylinder must be the same as the entire integral — 6π.

2. If $\nabla \cdot \vec{A} = 0$, then:

   (a) The Gauss theorem implies that the integration can be moved to a more convenient surface with the same boundary. (It is

essential that $\nabla \cdot \vec{A}$ be defined and zero *everywhere* in a region V filling in between the two surfaces.)



EXAMPLE: The integral over an arbitrary simple closed surface S (a "sack") equals the integral over a sphere S', if S can be continuously moved into S' without encountering any "holes" in space, where $\vec{A}$ is undefined or $\nabla \cdot \vec{A} \neq 0$. Any kind of dragons may live inside S'. FAMOUS EXAMPLE OF THE EXAMPLE: The electric field of a point charge q is

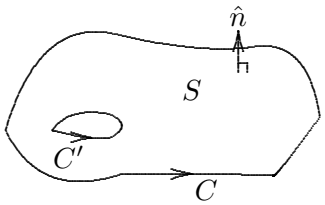
$$\vec{E} = \frac{q}{r^2}\,\hat{r}$$

(up to a constant factor depending on the system of units). Its flux through any sack surrounding the charge is $4\pi q$, as follows from choosing S' to be a small sphere centered on the charge. The charge itself occupies a "hole" in the foregoing sense; the sack may not be moved through it without changing the integral.

(b) If $\vec{A} = \nabla \times \vec{Z}$ for some vector field $\vec{Z}$, then the Stokes theorem says that

$$I = \oint_C \vec{Z} \cdot d\vec{r},$$

where C is the boundary of S with the counterclockwise orientation. (That is, as you stand on the positive side of S — i.e., with your head pointing in the direction of $\hat{n}$, rather than $-\hat{n}$ — and walk around C, you have S always on your left.) In particular, the integral is *the same for all surfaces* S having C as boundary. (The relationship between the two conditions $\nabla \cdot \vec{A} = 0$ and $\vec{A} = \nabla \times \vec{Z}$ will be discussed in depth later; they are almost, but not quite, equivalent.)



EXAMPLE: In $\mathbf{R}^3$, let

$$\vec{A} = \frac{-y}{x^2+y^2}\,\hat{\imath} + \frac{x}{x^2+y^2}\,\hat{\jmath}.$$

Then (cf. Exercise 7.5.8(a)) $\nabla \times \vec{A} = 0$ everywhere except on the z axis ($x = y = 0$), where both $\vec{A}$ and $\nabla \times \vec{A}$ are undefined. Let C be any closed curve that encircles the z axis exactly once, in a counterclockwise direction — for instance, the crimped ellipse

$$\vec{r}(t) = \begin{pmatrix} \cos t \\ 25 \sin t \\ \sin(20t) \end{pmatrix}.$$

Let C' be the unit circle in the x-y plane,

$$\vec{r}(t) = \begin{pmatrix} \cos t \\ \sin t \\ 0 \end{pmatrix}.$$

By Stokes's theorem, the integral around C is the same as the integral around the circle, which is easily evaluated to be 2π, by method (P) of Sec. 6.3 (cf. Exercise 7.5.8(b)).

3. In particular, S closed *and* $\nabla \cdot \vec{A} = 0$ everywhere inside it imply together that $\oint_S \vec{A} \cdot d\vec{S} = 0$.

FAMOUS EXAMPLE: The electric flux integral is 0 if there is no charge inside.

Now, please, go back and reread the statements — ignoring the examples — to appreciate the perfect parallelism between the line and surface theories.

PATH INDEPENDENCE AND POTENTIALS

We need to review some standard physics. Let's do so in the form of a problem.

Example: Let $\vec{F}(\vec{r}) = z\,\hat{\imath} + \hat{\jmath} + x\,\hat{k}$. (a) Show that a line integral $\int_{\vec{r}_\mathrm{i}}^{\vec{r}_\mathrm{f}} \vec{F}(\vec{r}) \cdot d\vec{r}$ is always independent of the path joining the initial and final points. (b) Find a potential energy function V such that $\vec{F} = -\nabla V$. (c) Explain why facts (a) and (b) are related.

SOLUTION: The curl of this vector field is everywhere zero:

$$\nabla \times \vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ z & 1 & x \end{vmatrix} = (1-1)\hat{\jmath} = \vec{0}.$$

This implies that the integrals are independent of path. By the fundamental theorem of calculus, the integral equals $-V(\vec{r}_{\rm f}) + V(\vec{r}_{\rm i})$. Clearly this is possible only if the integral is independent of path, so exhibiting the potential function is an alternative way of proving that. Conversely, if the integral is independent of path, we can integrate along any convenient path $C_{\vec{r}}$ from any fixed starting point (say $\vec{0}$) to $\vec{r}$ to define $-V(\vec{r})$. Alternatively, we can solve the three equations

$$\frac{\partial V}{\partial x} = -z, \quad \frac{\partial V}{\partial y} = -1, \quad \frac{\partial V}{\partial z} = -x$$

and fix the constants of integration by requiring that the three results for V be consistent. Either way, the answer is

$$V(x,y,z) = -xz - y$$

plus an arbitrary constant.

DETAILS OF METHOD 1: Let $C_{\vec{r}}$ consist of the three straight line segments from $(0,0,0)$ to $(0,y,0)$, to $(x,y,0)$, to (x,y,z). We need to decorate the integration variables with tildes to distinguish them from the (temporarily) fixed coordinates (x,y,z). We have

$$\begin{aligned} V(x,y,z) - V(0,0,0) &= -\int_{C_{\vec{r}}} (\tilde{z}\,\hat{\imath} + \hat{\jmath} + \tilde{x}\,\hat{k}) \cdot d\vec{\tilde{x}} \\ &= -\int (\tilde{z}\,d\tilde{x} + d\tilde{y} + \tilde{x}\,d\tilde{z}) \\ &= -(y + 0 + xz), \end{aligned}$$

so $V(x,y,z) = -xz - y + \text{constant}$. (In more detail: Along each segment, only the integration with respect to the variable that *changes* on that segment contributes, as we have seen in several similar examples previously. On the first segment we integrate 1 with respect to $\tilde{y}$ from 0 to y and get the answer y. On the second segment we integrate $\tilde{z} = 0$ with respect to $\tilde{x}$ and get 0. On the third segment we integrate $\tilde{x} = x$ with respect to $\tilde{z}$ from 0 to z and get xz.)

DETAILS OF METHOD 2: Since $\frac{\partial V}{\partial x} = -z$, we must have

$$V(x, y, z) = -zx + C_1(y, z),$$

where the "constant" of integration may depend on y and z. Similarly, we get

$$V(x, y, z) = -y + C_2(x, z), \qquad V(x, y, z) = -xz + C_3(x, y).$$

The only way the three answers can be consistent is if

$$C_1(y, z) = C_3(x, y) = -y + \text{constant}, \qquad C_2(x, z) = -xz + \text{constant}.$$

Thus V has the form claimed.

WARNING: A common error is to scramble Methods 1 and 2 together like this:

$$-\int (z\,dx + dy + x\,dz) = -(zx + y + xz) = -2xz - y.$$

This is **wrong**; the gradient of this function is not $-\vec{F}$! This is what Exercise 3.5.8 is about. When using Method 1, you must keep track of what $(\tilde{x}, \tilde{y}, \tilde{z})$ are on each segment of the curve; they are not always equal to (x, y, z). When using Method 2, you must set the three expressions for V equal to each other, *not add them.*

HOLES IN SPACE

Let's pursue the subject of the previous subsection further. In Sec. 6.4 we observed that the familiar calculus theorem on the equality of mixed second-order partials implies the following:

Theorem:

(1) The curl of a gradient is zero: If $\vec{A} = \nabla f$ for some f, then $\nabla \times \vec{A} = \vec{0}$. (More generally, in dimension n, the Jacobian matrix of $\vec{A}$ is symmetric.)

(2) The divergence of a curl is zero: If $\vec{B} = \nabla \times \vec{A}$ for some $\vec{A}$, then $\nabla \cdot \vec{B} = 0$.

Equally important is the question of whether the *converses* of those two identities are valid. That is,

- If $\nabla \times \vec{A} = 0$ (or the Jacobian matrix is symmetric), must there exist an f such that $\vec{A} = \nabla f$?
- If $\nabla \cdot \vec{B} = 0$, must there exist an $\vec{A}$ such that $\vec{B} = \nabla \times \vec{A}$?

The answers are *YES* if the given vector field in each case is defined and smooth everywhere in $\mathbf{R}^3$ and the curl or divergence condition is satisfied everywhere. In the first case, the "scalar potential" f is given by

$$f(\vec{r}) \equiv \int_{\vec{r}_0}^{\vec{r}} \vec{A}(\tilde{r}) \cdot d\tilde{r} + C \qquad (C = \text{arbitrary constant});$$

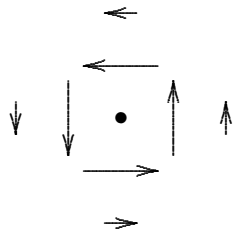
the line integral is independent of path by Stokes's theorem, and the fundamental theorem of calculus then shows that $\nabla f = \vec{A}$. In the second case, a suitable $\vec{A}$ is‡

$$\vec{A}(\vec{r}) \equiv -\vec{r} \times \int_0^1 \vec{B}(\lambda \vec{r})\, \lambda \, d\lambda.$$

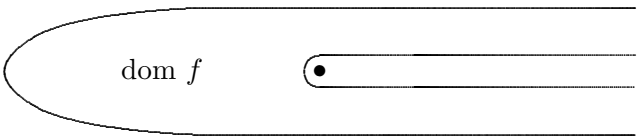
This answer is not unique: By virtue of the other identity, any gradient may be added to $\vec{A}$ without changing $\vec{B}$.

On the other hand, one answer or the other may be *NO* if the domain of the given vector field is a subset of $\mathbf{R}^n$ involving "holes". The most famous example is the two-dimensional vector field

$$A_x = \frac{-y}{x^2 + y^2}\,, \qquad A_y = \frac{x}{x^2 + y^2}\,.$$



Its two-dimensional curl, $\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}$, equals 0 everywhere except at the origin, where $\vec{A}$ and all its derivatives are undefined. (In contrast to the previous sketch of a field with positive curl, here the decrease in magnitude with distance from the origin precisely cancels the effect of the change of direction with angle.) In any region not encircling the origin,

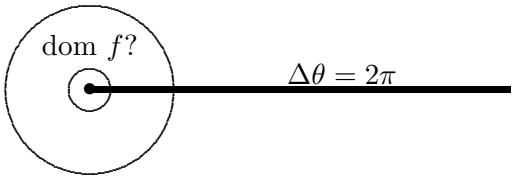


‡ L. Brand, The vector potential of a solenoidal vector, *Amer. Math. Monthly* **57**, 161–167 (1950).

$\vec{A}$ can be written as ∇f, where

$$f(\vec{r}) \equiv \tan^{-1} \frac{y}{x} + C' = \theta + C,$$

where θ is the angular variable in polar coordinates. As is well known, θ can't be smoothly defined everywhere in a region that does encircle the origin;



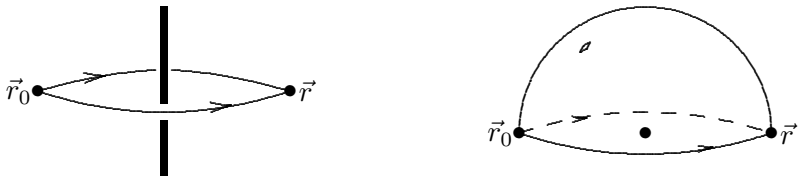
somewhere there must be a jump in θ of magnitude 2π. (See Exercise 7.5.8 for details.)

The impossibility of finding an f with the desired property (for this particular $\vec{A}$!) in such a ring-shaped region follows from the fact that $\int_C \vec{A} \cdot d\vec{r} = 2\pi \neq 0$ if the closed curve C encircles the origin once (in the counterclockwise direction). (If C is the unit circle, it is a routine exercise to evaluate this line integral by a technique discussed in the next section, using θ as the parameter.) If f existed, this integral would contradict the fundamental theorem of calculus, which tells us that

$$\int_{\vec{r}}^{\vec{r}} \vec{A} \cdot d\tilde{r} = f(\vec{r}) - f(\vec{r}) = 0.$$

Notice that if $\vec{A}$ were defined and $\nabla \times \vec{A} = 0$ everywhere inside C, then the integral would have to be 0 by Stokes's theorem. It is precisely the presence of the "hole" at the origin which creates the possibility that the line integral around a closed curve can be nonzero, or, equivalently, that the line integral of $\vec{A}$ between two points can depend on the path joining them (more specifically, on which side the path goes past the hole).

This example remains in force in dimension 3. (Define $A_z = 0$ with the same x and y components as before. Then $\nabla \times \vec{A} = 0$.) Then the "hole" is the entire z-axis. Since paths passing on opposite sides of the hole go together to form a closed curve which is *not* the boundary of a surface to which Stokes's theorem can be applied, it is possible for a line integral to depend upon path, so the line integral does not define a function f whose gradient is $\vec{A}$ everywhere.



On the other hand, a single point hole in $\mathbf{R}^3$ is no obstacle to defining f (for any $\vec{A}$ having zero curl everywhere except at that one point): Any closed curve can be filled in by a surface that avoids the hole, so Stokes says that the line integral around it must be 0.

However, point holes *can* interfere with the existence of a "vector potential" $\vec{A}$ such that $\nabla \times \vec{A} = \vec{B}$, given a "magnetic field" $\vec{B}$ such that $\nabla \cdot \vec{B} = 0$. The standard example is the famous *magnetic monopole*:

$$\vec{B}(\vec{r}) = \frac{1}{r^2}\,\hat{r}.$$

We have $\nabla \cdot \vec{B} = 0$ except at $\vec{r} = 0$. But $\iint_S \vec{B} \cdot d\vec{S} = 4\pi$ if S is a simple closed surface containing the origin; therefore, existence of an $\vec{A}$ would contradict Stokes's theorem. (The closed surface has no boundary, so the integral of $\vec{A}$ around the boundary is zero, not 4π. If you're uneasy about evaluating a line integral around a nonexistent boundary, think of S as the limiting case of a balloon, whose boundary is a very small curve at the neck of the balloon.)

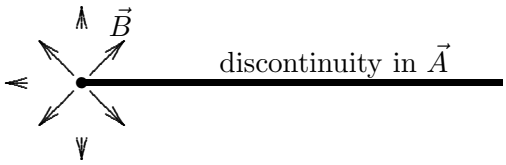
In analogy with the almost-everywhere-defined polar angle in the previous example, one can define an $\vec{A}$ which satisfies $\nabla \times \vec{A} = \vec{B}$ everywhere except on some curve connecting the origin to infinity. In the physical theory of magnetic monopoles, such a curve is called a *Dirac string*. An explicit example is

$$\vec{A}(\vec{r}) = \frac{\hat{k} \times \hat{r}}{r^2 \sin^2\theta}(1 - \cos\theta),$$

or equivalently

$$\vec{A} = \frac{1 - \cos\theta}{r \sin\theta}\,\hat{\phi}.$$

This is singular on the negative z axis, $\theta = \pi$.



This discussion has brought us to the brink of the subject called algebraic topology (specifically, homology and cohomology). Most pure mathematicians would probably say that this question of what types of holes prevent the construction of scalar and vector potentials is the only aspect of classical vector analysis that remains of research interest at the end of the 20th century.

Exercises

7.5.1 Write out Gauss's theorem in two dimensions and prove it equivalent to Green's theorem. HINT: Consider two vector fields, related so that the dot product of one with the tangent vector to the boundary is equal to the dot product of the other with the normal vector to the boundary. See the discussion in Sec. 6.2 of a two-dimensional analogue to the three-dimensional cross product.

7.5.2 The point of this problem is to compile a "cookbook" summary of the discussion of "holes in space". Suppose that $\vec{B}$ is a vector field defined in a region Ω in $\mathbf{R}^3$.

(a) State a list of conditions on $\vec{B}$ and Ω that together guarantee existence of a scalar function f such that $\vec{B} = \nabla f$ everywhere in Ω.

(b) State a list of conditions on $\vec{B}$ and Ω that together guarantee existence of a vector field $\vec{A}$ such that $\vec{B} = \nabla \times \vec{A}$ everywhere in Ω.

7.5.3 Explain the following statement:

> The identity $\nabla \times (\nabla f) = 0$ is necessary to prevent an inconsistency between Stokes's theorem and the fundamental theorem of calculus.

HINT: Consider two curves that have the same endpoints.

7.5.4 Explain the following statement:

> The identity $\nabla \cdot (\nabla \times \vec{B}) = 0$ is necessary to prevent an inconsistency between Stokes's theorem and Gauss's theorem.

HINT: Consider two surfaces that have the same curve as boundary.

7.5.5 Compute the (flux) integral of $\nabla \times \vec{F}$, where $\vec{F}(\vec{r}) = (x^2 z e^{2y}, yx^3, z)$, over the hemisphere $x^2 + y^2 + z^2 = 1$, $z \geq 0$ (with the upward normal vector field). Two methods:

(a) Use Stokes's theorem to reduce the problem to an example in Sec. 6.3.

(b) Use Gauss's theorem to replace the hemisphere by the disk at its base.

7.5.6 *(Continuation of Exercise 5.3.4)* Suppose that the Poisson equation $\nabla^2 \phi = \rho$ holds inside a sphere with center at the origin, and that the normal derivative $\frac{\partial \phi}{\partial r} = \hat{n} \cdot \nabla \phi$ equals 0 everywhere on the sphere. Show that the integral of ρ over the interior of the sphere must equal 0.

7.5.7 Let $\vec{B} = \nabla \times \vec{A}$ be a magnetic field, and S a surface. Consider the flux of $\vec{B}$ through the surface, $\iint_S (\nabla \times \vec{A}) \cdot d\vec{S}$.

(a) Describe a scenario in which Gauss's theorem would enable you to evaluate the flux integral.

(b) Describe a scenario in which Stokes's theorem would enable you to evaluate the flux integral.

7.5.8 Consider the vector field in $\mathbf{R}^2$,

$$\vec{A}(x, y) = \frac{-y}{x^2 + y^2}\,\hat{\imath} + \frac{x}{x^2 + y^2}\,\hat{\jmath}.$$

(a) Show that $\partial A_x / \partial y = \partial A_y / \partial x$ for all $(x, y) \neq (0, 0)$, and conclude that therefore $\mathbf{A}$ can be written as the gradient of a function throughout any region of the plane that does not encircle the origin.

(b) Evaluate $\int_C \vec{A} \cdot d\vec{r}$ around a circle C,

$$x = R\cos t, \quad y = R\sin t, \qquad 0 \leq t < 2\pi,$$

and conclude that $\mathbf{A}$ can't be written as the gradient of a function throughout a region of the plane that either contains or encircles the origin.

7.5.9 Find a potential function $V(\vec{r})$ so that $\vec{F}(\vec{r}) = -\nabla V(\vec{r})$, if

$$\vec{F}(x, y, z) = y^2 z\hat{\imath} + 2xyz\hat{\jmath} + xy^2\hat{k}.$$

In the remaining exercises, evaluate the work done in moving a particle along the given path through the given force field, and tell whether the work is actually independent of the path (with fixed endpoints). (Of course, it may be helpful to answer the second question first!)

7.5.10 $\vec{F}(\vec{r}) = y^2\hat{\imath} + z^2\hat{\jmath} + x^2\hat{k}$,

$$\vec{r}(t) = \big(\cos t, 3 + \sin t, (t-\pi)^2\big) \quad (0 < t < 2\pi).$$

7.5.11 $\vec{F}(\vec{r}) = \dfrac{\hat{r}}{r^2} = \dfrac{\vec{r}}{\|\vec{r}\|^3}$ (the Coulomb field of a point charge),

$$\vec{r}(t) = (1+t)\vec{r}_0 + (\sin t)\vec{r}_1 \quad (0 < t < \pi),$$

where $\vec{r}_0$ and $\vec{r}_1$ are nonzero and perpendicular to each other, but otherwise arbitrary. (Draw a sketch.)

7.5.12 $\vec{F}(\vec{r}) = y\hat{\imath} - x\hat{\jmath} + 0\hat{k}$, $\quad \vec{r}(t) = 2\cos t\,\hat{\imath} + 2\sin t\,\hat{\jmath} \quad (0 < t < 2\pi)$.

7.6 Direct Methods of Evaluating Line and Surface Integrals

As in the previous section, we shall treat line and surface integrals in exact parallel. (Unavoidably there will be some repetition of previous material, but here we shall finish and consolidate it.) Recall that we have two ways of writing each of these integrals, one that involves a dot product and a unit vector and one that does not. There are two classes of evaluation methods, corresponding to these two symbolic representations, (G) and (A). In general, both approaches lead to some kind of parametric (P) representation of the particular integral at hand, which can then be evaluated by standard calculus.

Direct methods I: The algebraic viewpoint

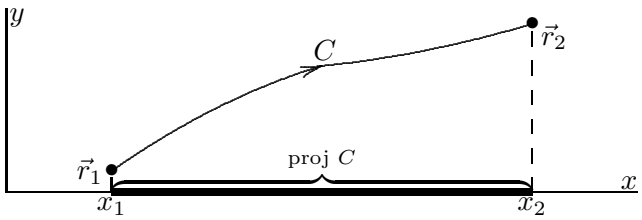
A. *Line integrals* $I = \int_C (A_x\,dx + A_y\,dy)$ (A)

For notational and pictorial simplicity we consider just the two-dimensional case.

1. If C is the graph of a function, $y = \psi(x)$, then the term $\int_C A_x\,dx$ is an ordinary integral over x, y being evaluated at the proper place:

$$\int_{\text{proj } C} A_x\big(x, \psi(x)\big)\,dx = \int_{x_1}^{x_2} A_x\big(x, \psi(x)\big)\,dx,$$

where proj C is the projection of C onto the x axis. A more complicated curve (not the graph of a function) can be patched together from pieces like this. (E.g., a circle is two semicircles.) The A_y term can be projected onto the y axis similarly; usually this requires a *different* cutting of the curve into pieces.



EXAMPLE: Let $\vec{A}(\vec{r}) = x^2\,\hat{\imath} + (x+1)y\,\hat{\jmath}$ and let C be the parabolic arc

$$y = 1 - x^2, \quad -1 < x < \tfrac{1}{2}$$

(traversed from left to right). The term $\int_C A_x\,dx$ is

$$\int_{-1}^{1/2} x^2\,dx = \left.\frac{x^3}{3}\right|_{-1}^{1/2} = \frac{1}{24} - \frac{(-1)}{3} = \frac{3}{8}\,.$$

The term $\int_C A_y\,dy$ requires a more delicate analysis. Note that on this path y varies from 0 to 1 and then back to $\frac{3}{4}$. So we should split up the dy integral as

$$\int_0^1 (x+1)y\,dy + \int_1^{3/4} (x+1)y\,dy.$$

But x must be evaluated *on the curve*, and this makes x a different function of y in each term:

$$x = \begin{cases} -\sqrt{1-y} & \text{in first term,} \\ +\sqrt{1-y} & \text{in second term.} \end{cases}$$

Thus

$$\begin{aligned}
\int_C A_y\,dy &= \int_0^1 (1-\sqrt{1-y})\,y\,dy + \int_1^{3/4} (1+\sqrt{1-y})\,y\,dy \\
&= \int_0^{3/4} y\,dy - \int_0^{3/4} \sqrt{1-y}\,y\,dy - 2\int_{3/4}^1 \sqrt{1-y}\,y\,dy \\
&= \frac{y^2}{2}\Big|_0^{3/4} + \left[\frac{2}{3}(1-y)^{3/2} - \frac{2}{5}(1-y)^{5/2}\right]_0^{3/4} \\
&\qquad + 2\left[\frac{2}{3}(1-y)^{3/2} - \frac{2}{5}(1-y)^{5/2}\right]_{3/4}^1 \\
&= \frac{9}{32} - \Big[\cdots\Big]_{3/4} - \Big[\cdots\Big]_0 + 2\Big[\cdots\Big]_1 \\
&= \frac{9}{32} - \left(\frac{1}{12} - \frac{1}{80}\right) - \left(\frac{2}{3} - \frac{2}{5}\right) + 2(0-0) \\
&= \frac{9}{32} + \left(\frac{33}{80} - \frac{3}{4}\right) = -\frac{9}{160}\,.
\end{aligned}$$

Thus $I = \frac{3}{8} - \frac{9}{160} = \frac{51}{160}$.

2. However, it isn't necessary to project each term onto its own axis. For example, sticking with the representation $y = \psi(x)$, one can evaluate the A_y term as

$$\int_C A_y\,dy = \int_{x_1}^{x_2} A_y\big(x, \psi(x)\big)\,\psi'(x)\,dx.$$

EXAMPLE: This provides us with a much easier way of doing the previous example.

$$\begin{aligned}
\int_C (A_x\,dx + A_y\,dy) &= \int_{-1}^{1/2} [x^2 + (x+1)(1-x^2)\,\tfrac{d}{dx}(1-x^2)]\,dx \\
&= \int_{-1}^{1/2} (2x^4 + 2x^3 - x^2 - 2x)\,dx \\
&= \frac{51}{160} \quad \text{(after some routine algebra).}
\end{aligned}$$

3. *Parametrization of the curve.* Regard C as the range of a function of a new variable, t: $x\,\hat{\imath} + y\hat{\jmath} = \vec{g}(t)$. Let t increase in the direction of the

arrow on C. Let t_1 and t_2 be the parameter values at the endpoints: $g(t_1) = \vec{r}_1$, $g(t_2) = \vec{r}_2$. Then

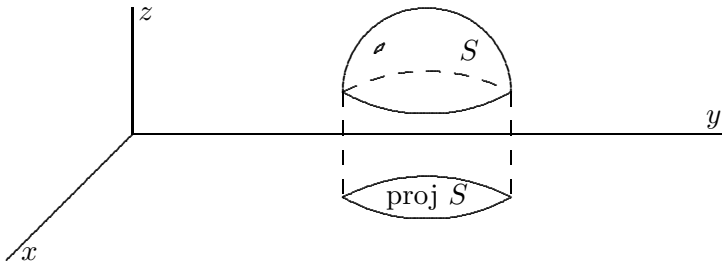
$$\int_C \vec{A} \cdot d\vec{r} = \int_{t_1}^{t_2} \left(A_x \frac{dx}{dt} + A_y \frac{dy}{dt} \right) dt$$
$$= \int_{t_1}^{t_2} \vec{A}\big(\vec{g}(t)\big) \cdot \vec{g}\,'(t)\, dt. \qquad \text{(P)}$$

Note that if you do this in the special case where t is one of the original Cartesian coordinates, you get integrals of the forms (A2) and (A1); the latter are special cases of the parametrization prescription.

EXAMPLES: We have already been doing this for some time. See Sec. 6.3 (end); Exercise 7.5.8(b); Exercise 7.5.10.

B. *Surface integrals* $I = \iint_S (A_x\, dy\, dz + A_y\, dz\, dx + A_z\, dx\, dy)$ (A)

We have not yet justified (or even defined) this expression. Here we shall present consistent procedures for giving it a definite numerical meaning; the next subsection will show why it is equivalent to our earlier definition of I in terms of the unit normal vector and the surface area.



1. If S is the graph of a function, $z = \psi(x, y)$, then the term $\iint_S A_z\, dx\, dy$ can be evaluated as an ordinary double integral over x and y, z being evaluated at the proper place:

$$\iint_{\text{proj } S} A_z\big(x, y, \psi(x, y)\big)\, dx\, dy.$$

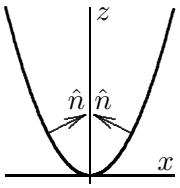
(If S is perpendicular to the x-y plane, the A_z term is zero.) A more complicated surface can be patched together from pieces like this. (For example, a sphere is two hemispheres.) The other two terms can be projected into the other two coordinate planes similarly; usually

each of the three terms requires a *different* way of cutting the surface into pieces.

EXAMPLE: Let S be the part of the graph of $z = x^2$ lying above the square $-1 < x < 1$, $-1 < y < 1$. Let us reduce the surface integral $\iint_S \vec{A}\cdot d\vec{S}$ to double integrals in the coordinate planes. The easy term is

$$\iint A_z\,dx\,dy = \int_{-1}^{1} dx \int_{-1}^{1} dy\, A_z(x, y, x^2).$$

The term $\iint A_x\,dy\,dz$ would better be done by the next method on our list, but to demonstrate the principles we shall treat it by regarding y and z as coordinates on S. To do that we need to solve for x: $x = \pm\sqrt{z}$. From the point of view of the y-z coodinate plane, there are really two surfaces, corresponding to the two signs of x. The normal vector (which points *into* the parabola) is in the positive x direction on the negative-x sheet, and vice versa.



Thus the positive sheet, as the graph of a function $x = \psi(y, z)$, must be regarded as upside down, so we have

$$\iint A_x\,dy\,dz = \int_0^1 dz \int_{-1}^{1} dy\, A_x(-\sqrt{z}, y, z) - \int_0^1 dz \int_{-1}^{1} dy\, A_x(+\sqrt{z}, y, z).$$

Finally, the surface projects onto a one-dimensional curve (of area 0) in the x-z plane, so

$$\iint A_y\,dz\,dx = 0.$$

You should check (Exercise 7.6.15) that this same result is obtained by the next method.

2. However, it isn't necessary to project each term onto its own plane. Sticking with the representation $z = \psi(x, y)$, one can do the $dy\,dz$ and $dz\,dx$ terms this way: We have

$$dz = \frac{\partial\psi}{\partial x}\,dx + \frac{\partial\psi}{\partial y}\,dy.$$

Substitute this for dz and use the mystical identities

$$(dx_j)^2 = 0 \qquad (j = 1,\, 2,\, 3),$$
$$dx_j dx_k = -\,dx_k\, dx_j \quad \text{if } j \neq k.$$

For example,

$$\begin{aligned}\iint A_x\, dy\, dz &= \iint A_x\, dy \left(\frac{\partial\psi}{\partial x}\, dx + \frac{\partial\psi}{\partial y}\, dy\right) \\ &= -\iint A_x\, \frac{\partial\psi}{\partial x}\, dx\, dy.\end{aligned}$$

Thus, in the case where S has a one-to-one projection onto the x-y plane,

$$I = \iint_{\text{proj } S} \left[-A_x\, \frac{\partial\psi}{\partial x} - A_y \frac{\partial\psi}{\partial y} + A_z\right] dx\, dy. \qquad (*)$$

(Note that the *cyclic* order ($x \to y \to z \to x \to \cdots$) of the differentials in each term of our original expression for I is critical to getting the signs right.) The projection onto any other coordinate plane can be worked out similarly. The mystical rules are part of the lore of differential forms; but from our point of view it is better to see this technique as a special case of the next one, parametrization.

EXAMPLE: If $\vec{B}(\vec{r}) = y\,\hat{\imath} + z\,\hat{k}$ and S is the part of the graph of $z = x^2 + y$ above the region $0 < x < 1$, $0 < y < 2$, then

$$\iint_S \vec{B}\cdot d\vec{S} = \iint_S [B_x\, dy\, dz + B_y\, dz\, dx + B_z\, dx\, dy] = \iint_S (y\, dy\, dz + z\, dx\, dy).$$

With $dz = 2x\, dx + dy$ and $dy\, dx = -dx\, dy$, $dy\, dy = 0$, we get

$$\begin{aligned}\iint_x (-2xy + x^2 + y)\, dx\, dy &= \int_0^2 dy \int_0^1 dx\, (-2xy + x^2 + y) \\ &= \int_0^2 [-x^2 y + \frac{1}{3}\, x^3 + xy]_0^1\, dy \\ &= \int_0^2 \left(\frac{1}{3}\right)\, dy = \frac{2}{3}\,.\end{aligned}$$

(Note that once you reduce the integral to an ordinary double integral with differentials $dx\, dy$, you can evaluate the iterated integral in either order without changing the sign again!)

3. *Parametrization of the surface.* Regard S as the range of a function of two new variables, u and v:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \vec{g}(u,v).$$

(The prototype is the parametrization

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} R\sin\theta\cos\phi \\ R\sin\theta\sin\phi \\ R\cos\theta \end{pmatrix}$$

of a sphere of radius R.) We insist that (u,v) form a right-handed coordinate system from the point of view of an observer standing "on top" of S; in other words,

$$\frac{\partial\vec{g}}{\partial u} \times \frac{\partial\vec{g}}{\partial v}$$

points in the direction of $\hat{n}$, not $-\hat{n}$. Let

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

with similar definitions of $\dfrac{\partial(z,x)}{\partial(u,v)}$ and $\dfrac{\partial(x,y)}{\partial(u,v)}$. (Recall that the vertical bars denote a determinant, which may be positive or negative.) Then

$$\begin{aligned} &\iint_S \vec{A}\cdot d\vec{S} \\ &\quad = \iint_{g^{-1}(S)} \left[A_x \frac{\partial(y,z)}{\partial(u,v)} + A_y \frac{\partial(z,x)}{\partial(u,v)} + A_z \frac{\partial(x,y)}{\partial(u,v)} \right] du\,dv. \qquad (P) \end{aligned}$$

(Note that this formula neatly fills in between the curve parametrization formula (essentially integration by substitution) for line integrals and Jacobi's theorem for volume integrals.)

Example 1: When prescription (B3) is carried out in the special case where u and v are two of the original Cartesian coordinates, it reduces to prescriptions (B1) and (B2). For instance,

$$\frac{\partial(y,z)}{\partial(x,y)} = \begin{vmatrix} 0 & 1 \\ \frac{\partial\psi}{\partial x} & \frac{\partial\psi}{\partial y} \end{vmatrix} = -\frac{\partial\psi}{\partial x}$$

is one of the determinants that arises when $(u, v) = (x, y)$. The "mystical identities" simply summarize the process of calculating the 2×2 Jacobian determinants. For another approach to the same integrals, see Exercise 7.6.13.

Example 2: For the sphere,

$$\begin{aligned} \frac{\partial x}{\partial \theta} &= R\cos\theta\cos\phi, & \frac{\partial x}{\partial \phi} &= -R\sin\theta\sin\phi, \\ \frac{\partial y}{\partial \theta} &= R\cos\theta\sin\phi, & \frac{\partial y}{\partial \phi} &= R\sin\theta\cos\phi, \\ \frac{\partial z}{\partial \theta} &= -R\sin\theta, & \frac{\partial z}{\partial \phi} &= 0. \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{\partial(y,z)}{\partial(\theta,\phi)} &= R^2\sin^2\theta\cos\phi, \\ \frac{\partial(z,x)}{\partial(\theta,\phi)} &= R^2\sin^2\theta\sin\phi, \\ \frac{\partial(x,y)}{\partial(\theta,\phi)} &= R^2\cos\theta\sin\theta. \end{aligned}$$

Therefore,

$$d\vec{S} = R^2\left[\sin^2\theta\cos\phi\,\hat{\imath} + \sin^2\theta\sin\phi\,\hat{\jmath} + \cos\theta\sin\theta\,\hat{k}\right]d\theta\,d\phi.$$

Substituting this expression into $\iint_S \vec{A}\cdot d\vec{S}$ tells one how to find the flux of $\vec{A}$ through a sphere (or a piece of a sphere) in terms of the Cartesian components of $\vec{A}$.

Incidentally, the mystical rules apply to curvilinear coordinates, too. Instead of writing down the Jacobian determinants, we could have done the calculation this way:

$$\begin{aligned} dx &= \frac{\partial x}{\partial \theta}\,d\theta + \frac{\partial x}{\partial \phi}\,d\phi = R\cos\theta\cos\phi\,d\theta - R\sin\theta\sin\phi\,d\phi, \\ dy &= R\cos\theta\sin\phi\,d\theta + R\sin\theta\cos\phi\,d\phi, \end{aligned}$$

etc. Therefore, if we interpret

$$(d\theta)^2 = 0 = (d\phi)^2, \qquad d\phi\,d\theta = -d\theta\,d\phi,$$

we get

$$\begin{aligned} dx\,dy &= (-R\sin\theta\sin\phi\,d\phi)(R\cos\theta\sin\phi\,d\theta) \\ &\qquad + (R\cos\theta\cos\phi\,d\theta)(R\sin\theta\cos\phi\,d\phi) \\ &= R^2\cos\theta\sin\theta\,d\theta\,d\phi, \end{aligned}$$

etc., in agreement with our previous result for the z component of $d\vec{S}$.

REMARK: The foregoing is not usually the most efficient way to do a flux integral over a part of a sphere; we did it just to demonstrate the general method in a familiar context. Since the unit normal vector, $\hat{n} = \hat{r}$, and the scalar element of surface area, $dS = R^2\sin\theta\,d\theta\,d\phi$, are well known for the sphere, it is usually easier to use the (G) formula as we did in Sec. 7.4.

DIRECT METHODS II: THE EUCLIDEAN GEOMETRIC VIEWPOINT

We have already discussed this for geometrically simple problems in Secs. 6.3 and 7.4. Now we'll see how, in more complicated problems, it leads once more to the parametric formulas (P).

A. *Line integrals* $I = \int_C \big(\vec{A}(\vec{r})\cdot\hat{T}(\vec{r})\big)\,ds$ (G)

Here the vector line integral is expressed as a *scalar* line integral with integrand $\vec{A}\cdot\hat{T}$ and integration variable s. The key step in setting up such an integral is getting an analytic expression for ds, the differential of the *arc length parameter* along the curve. Just as in the previous discussion, there are two ways to convert an arc-length integral to an ordinary integral: use one of the coordinates as a parameter, or introduce a new parameter.

1. If C is the graph of a function, $y = \psi(x)$, then

$$ds = \sqrt{1 + (\psi')^2}\,dx. \qquad (\dagger)$$

In dimension 3 this becomes

$$\sqrt{1 + \left(\frac{dy}{dx}\right)^2 + \left(\frac{dz}{dx}\right)^2}\,dx. \qquad (\ddagger)$$

2. If C is parametrized as $\vec{r} = \vec{g}(t)$, then

$$ds = \|\vec{g}\,'(t)\|\,dt = \sqrt{(g_x')^2 + (g_y')^2 + (g_z')^2}\,dt.$$

Note that (A1) is a special case of (A2). Note also that (A2) is quite understandable as a representation of the length of the "infinitesimal" piece of curve, $d\vec{r}$, corresponding to an interval of the variable t of length dt.

If you know the unit tangent vector $\hat{T}(\vec{r})$ by some independent method, these formulas for arc length reduce the line integral to an ordinary single-variable integral. If $\hat{T}$ is calculated from $\|\vec{g}\,'\|$ as in Sec. 6.3, then, as shown there, the square root factor will cancel and one recovers the parametric formula (P).

B. *Surface integrals* $I = \iint_S (\vec{A} \cdot \hat{n})\, dS$ (G)

Our task is to find expressions for the scalar surface-area element, dS, and the unit normal vector, $\hat{n}$.

Well, the previously mentioned vector

$$\frac{\partial \vec{g}}{\partial u} \times \frac{\partial \vec{g}}{\partial v}$$

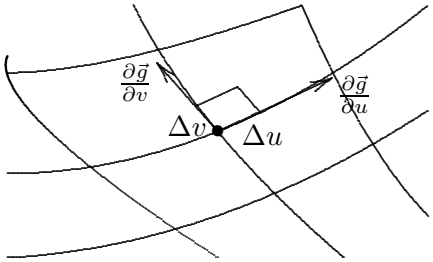
is normal to the surface, since the two factors in the cross product are vectors tangent to the surface. To get a unit normal (pointing in the agreed "top" direction) we need only divide by the length:

$$\hat{n}(\vec{r}) \equiv \frac{\frac{\partial \vec{g}}{\partial u} \times \frac{\partial \vec{g}}{\partial v}}{\left\| \frac{\partial \vec{g}}{\partial u} \times \frac{\partial \vec{g}}{\partial v} \right\|}\,. \tag{1}$$

In any particular case, the denominator will be the square root of some fairly complicated expression. Fortunately, we can usually avoid calculating it, as we shall now see.

To find dS, recall how we got the Jacobian determinant in the formula for ordinary multiple integrals in $\mathbf{R}^n$ in terms of curvilinear coordinates. Essentially it went back to the formula for the volume of a parallelepiped (or area of a parallelogram) as the determinant of the vectors that determine it (Sec. 7.2). The same principle applies to area in the surface. In the limit of small increments, the area of the patch of surface with coordinate values lying between (u, v) and $(u + \Delta u, v + \Delta v)$ — i.e., the "infinitesimal parallelogram" determined by the tangent vectors $\frac{\partial \vec{g}}{\partial u}\,\Delta u$ and $\frac{\partial \vec{g}}{\partial v}\,\Delta v$ — is the norm of the cross product of the tangent vectors. (See Sec. 7.2.) Thus the desired formula is

$$dS \equiv \left\| \frac{\partial \vec{g}}{\partial u} \times \frac{\partial \vec{g}}{\partial v} \right\| du\, dv. \tag{2}$$

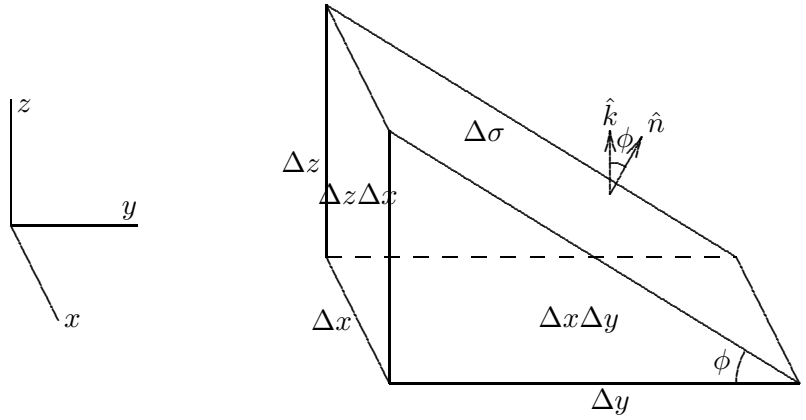


Putting these two formulas together, we see that in the vector surface element the normalization factor cancels out!

$$d\vec{S} \equiv \hat{n}\, dS = \left(\frac{\partial \vec{g}}{\partial u} \times \frac{\partial \vec{g}}{\partial v} \right) du\, dv. \tag{3}$$

When this is written out in Cartesian coordinates, one recovers the Jacobian (or mystical) rule of the preceding subsection. Rather than demonstrate this algebraically, we shall show geometrically that the Jacobian rule agrees with our definition of a vector surface integral as the scalar surface integral of the field's normal component.

Without any real loss of generality, consider the case that $n_x = 0$, and look at a small rectangular piece of surface, $\Delta\vec{S}$. Then the projection of that piece onto the x-y plane is a rectangle; call its sides Δx and Δy, so that its area is $\Delta x\, \Delta y$. Similarly, its x-z projection is a rectangle of area $\Delta z\, \Delta x$. Its projection onto the y-z plane, however, has area 0, *not* $\Delta y\, \Delta z$.



Let the area of $\Delta\vec{S}$ be $\Delta\sigma$, and the angle between $\hat{n}$ and the vertical be ϕ. Then the area $\Delta x\, \Delta y$ equals $\cos\phi\, \Delta\sigma$, and $\Delta z\, \Delta x = \sin\phi\, \Delta\sigma$. If $\vec{A}$ is perpendicular to the x-y plane ($\vec{A} = A_z\, \hat{k}$), then

$$\vec{A} \cdot \hat{n}\, \Delta\sigma = A_z \cos\phi\, \Delta\sigma = A_z\, \Delta x\, \Delta y.$$

(The two cosine factors have cancelled! Think of $\vec{A}$ as representing a flux of particles moving along parallel paths in the z direction. As the surface is tilted from the horizontal, the density of "hits" of the particles on the surface decreases. (This is why the polar regions are cooler than the tropics, when the particles are photons from the sun.) On the other hand, the area of the piece of surface corresponding to a given projected region in the x–y coordinate plane *increases* as the tilt increases. The two effects exactly cancel.)

Similarly, if $\vec{A} = A_y\,\hat{\jmath}$, then $\vec{A}\cdot\hat{n}\,\Delta\sigma = A_y\,\Delta z\,\Delta x$. Finally, if $\vec{A} = A_x\,\hat{\imath}$, then $\vec{A}\cdot\hat{n} = 0$. Since any vector field can be decomposed into three parts of the types we've considered, we can write in general (still assuming that $n_x = 0$)

$$\vec{A}\cdot\hat{n} = A_z\,\Delta x\,\Delta y + A_y\,\Delta z\,\Delta x.$$

If we did this for each term in a Riemann sum for a surface whose normal is always parallel to the y-z plane, we would establish that

$$\iint_S \vec{A}\cdot\hat{n}\,dS = \iint_S \big[A_x\,dy\,dz + A_y\,dz\,dx + A_z\,dx\,dy\big].$$

(When a parametrization is introduced and the three terms are evaluated by the 2×2 Jacobian prescription, the term $\iint A_x\,dy\,dz$ will be 0 in such a case. For example, if we take x and y as parameters,

$$\frac{\partial(y,z)}{\partial(x,y)} = \begin{vmatrix} 0 & 1 \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{vmatrix} = -\frac{\partial z}{\partial x} = 0,$$

because the surface is not sloping in the x direction. On the other hand, we have

$$\frac{\partial(x,y)}{\partial(x,y)} = 1, \qquad \frac{\partial(z,x)}{\partial(x,y)} = -\frac{\partial z}{\partial y} = \pm\tan\phi,$$

consistent with our observation that the areas of the projections on the other two planes are in the ratio $\sin\phi : \cos\phi$.)

Although this has been far from a complete proof — we said nothing about what happens if $\hat{n}$ wanders about instead of staying in one plane — it should suffice to convince you that the various prescriptions for defining and evaluating surface integrals are consistent.

Exercises

7.6.1 Use Gauss's theorem to evaluate $\iint_S \vec{F} \cdot d\vec{S}$ when $\vec{F}(\vec{r}) = x\hat{\imath} + y\hat{\jmath} = (x, y, 0)$ and S is the parabolic dome

$$z = 1 - x^2 - y^2, \qquad x^2 + y^2 < 1.$$

7.6.2 Evaluate the integral in the preceding exercise by a direct method.

7.6.3 Let $\vec{F}(x, y, z) = (y^2, 2x^2, z^2)$, and let C be the curve

$$(x, y, z) = \big(2\cos t, 2\sin t, 5 + \sin(10t)\big), \qquad 0 \le t \le 2\pi.$$

Calculate $\oint_C \vec{F} \cdot d\vec{r}$. HINT: Think of the curve as the jagged edge of an open tin can whose base lies in the plane $z = 0$. (Sketch it.)

7.6.4 By the method of your choice, calculate $\int_S \vec{F} \cdot d\vec{S}$ when S is the sphere of radius 2 centered at the origin in $\mathbf{R}^3$ (with outward-pointing normal), where (in two equivalent notations)

$$\vec{F}(\vec{r}) = y^2\,\hat{\jmath} + zy^2\,\hat{k} = (0, y^2, zy^2).$$

7.6.5 Repeat the preceding exercise with S equal to the boundary of the unit cube $0 < x < 1$, $0 < y < 1$, $0 < z < 1$.

7.6.6 Use the Stokes (or the Green) theorem to justify the following prescription for numerically calculating (approximately) the area of a region R in the plane:

Let $\{(x_k, y_k)\}$ be a sequence of n closely spaced points tracing out the boundary of R. The starting point, (x_0, y_0), is the same as the ending one, (x_n, y_n). Then

$$\text{Area of } R \approx \frac{1}{2}\left|\sum_{k=0}^{n-1}(x_k y_{k+1} - x_{k+1}y_k)\right|.$$

HINT: $x_k y_{k+1} - x_{k+1}y_k = -(x_{k+1} - x_k)y_k + x_k(y_{k+1} - y_k)$. Now think "Riemann sum!".

7.6.7 Let $\vec{A}(\vec{r}) = z\,\hat{k} = (0, 0, z)$. Use Gauss's theorem to compute the flux integral $\iint \vec{A} \cdot d\vec{S}$ over the sphere of radius 4 centered at the origin in $\mathbf{R}^3$ (with outward-pointing normal vector).

7.6.8 Evaluate the flux integral in Exercise 7.6.7 directly.

7.6.9 Still in the context of Exercise 7.6.7, calculate the flux $\iint_S \vec{A} \cdot d\vec{S}$ through the "polar cap" $\theta < \pi/3$. (Here θ is the polar angle (colatitude) defined by $z = R\cos\theta$, with $R = 4$ in this case.)

7.6.10 Evaluate $\iint_H \vec{F} \cdot d\vec{S}$ when H is the hemisphere $x^2 + y^2 + z^2 = 4$, $z > 0$ (with upward normal), and $\vec{F}(\vec{r}) = (0, 0, 1) = \hat{k}$ for all $\vec{r}$.

7.6.11 Evaluate $\int_C \vec{F} \cdot d\vec{r}$ when $\vec{F}(\vec{r}) = x^2\hat{\imath} + (y+1)\hat{\jmath} + z\hat{k}$ and C is:

(a) the ellipse $x = 5\cos t$, $y = 2\sin t$, $z = 4$ $(0 \le t < 2\pi)$.

(b) the line segment $x = 3t$, $y = 1 + 2t$, $z = t$ $(0 \le t < 2)$.

7.6.12 Use Stokes's theorem to calculate the flux of $\nabla \times \vec{A}$ through the triangle $0 < z < 1$, $y = z$, $z < x < 1$, if $\vec{A}(\vec{r}) = (x+y)\hat{\imath} + (y - x^2 + z)\hat{\jmath} + (z - y)\hat{k}$. (Orient the triangle so that the normal vector points upward.)

7.6.13 Let the surface S be the graph of a function, $z = \psi(x, y)$ (with some domain, proj S, in the x-y plane).

(a) Find the scalar surface area element, (2), and comment on the resemblance to the formulas for arc length, (†) and (‡).

(b) Find the unit normal vector, (1), and hence the vector surface area element, (3). (Compare with Exercise 7.4.5.)

(c) Use (a) and (b) to give an alternative derivation of formula $(*)$ for $\iint_S \vec{A} \cdot d\vec{S}$.

7.6.14 Do Example 3 of Sec. 7.4 by the parametric method.

7.6.15 Verify by the parametric method B2 that $\iint A_y \, dz\, dx = 0$ for the parabolic sheet used as the example for method B1.

7.6.16 Some books give this prescription for calculating surface area: Let the surface be given by the parametric equations

$$x = \varphi_1(u, v), \quad y = \varphi_2(u, v), \quad z = \varphi_3(u, v).$$

Then

$$dS = \sqrt{EG - F^2}\, du\, dv$$

where

$$E = \left(\frac{\partial \varphi_1}{\partial u}\right)^2 + \left(\frac{\partial \varphi_2}{\partial u}\right)^2 + \left(\frac{\partial \varphi_3}{\partial u}\right)^2,$$

$$G = \left(\frac{\partial \varphi_1}{\partial v}\right)^2 + \left(\frac{\partial \varphi_2}{\partial v}\right)^2 + \left(\frac{\partial \varphi_3}{\partial v}\right)^2,$$

$$F = \frac{\partial \varphi_1}{\partial u} \cdot \frac{\partial \varphi_1}{\partial v} + \frac{\partial \varphi_2}{\partial u} \cdot \frac{\partial \varphi_2}{\partial v} + \frac{\partial \varphi_3}{\partial u} \cdot \frac{\partial \varphi_3}{\partial v}.$$

Where did this formula come from? HINT: Use Exercise 2.5.10(a).

7.6.17 Derive the formula in the preceding exercise another way: Apply part (2) of the volume theorem in Sec. 7.2 (with $p = 2$) to the basic definition of surface area as the area of the infinitesimal parallelogram formed by two tangent vectors.

In the remaining exercises calculate $I \equiv \iint_S \vec{A} \cdot d\vec{S}$ with

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

7.6.18 S is the hemisphere $x^2 + y^2 + z^2 = 1, \quad z > 0$.

7.6.19 S is the cylinder $x^2 + y^2 = 9$, $-1 < z < 1$. (Note that this is just the curved part of the cylinder; the end faces are not included.)

7.6.20 S is the closed cylindrical surface consisting of the surface in the preceding problem together with the end faces at $z = \pm 1$ (in other words, the boundary of the 3-dimensional region $x^2 + y^2 < 9$, $-1 < z < 1$.)

7.6.21 S is the sphere $x^2 + y^2 + z^2 = 1$ (with the usual, outward unit normal vector).