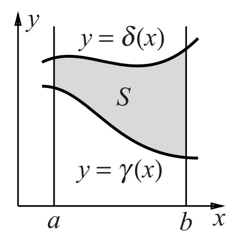
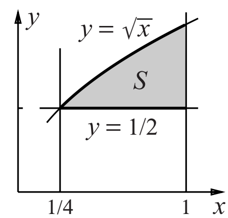
Chapter 9
Evaluating Double Integrals

Abstract Although the definition of the integral reflects its origins in scientific problems, its evaluation relies on a considerable range of mathematical concepts and tools. Most fundamental is the change of variables formula; the single-variable version (“u-substitution”) is perhaps the core technique of integration in the introductory calculus course. By contrast, the method of iterated integrals has no single-variable analogue; it evaluates a double integral by “partial integration” of one variable at a time. This chapter connects double and iterated integrals, establishes the change of variables formula, and discusses Green’s theorem as a tool for evaluating double integrals and as a reason for orienting them.

9.1 Iterated integrals

Partial integration





We define iterated integrals in their own terms, independently of double integrals. First, suppose that S is the region in the (x,y)-plane that lies between the graphs $y = \gamma(x)$ and $y = \delta(x)$ when $a \le x \le b$. We assume γ and δ are continuous and $\gamma(x) \le \delta(x)$ everywhere on this interval; we can write

$$S: \quad \begin{array}{c} a \le x \le b, \\ \gamma(x) \le y \le \delta(x). \end{array}$$

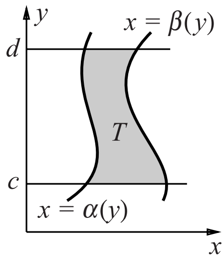
Now let $f(x,y)$ be a continuous function on S; for each x in $[a,b]$, compute the *“partial integral” of $f(x,y)$ with respect to y* from $\gamma(x)$ to $\delta(x)$:

$$F_2(x) = \int_{\gamma(x)}^{\delta(x)} f(x,y)\,dy, \quad a \le x \le b.$$

This is a continuous function of x. As an example, let $f(x,y) = x^2y^3$ and let S be the region between $y = \gamma(x) = 1/2$ and $y = \delta(x) = \sqrt{x}$ when $1/4 \le x \le 1$. Then the partial integral is

J.J. Callahan, *Advanced Calculus: A Geometric View*, Undergraduate Texts in Mathematics, DOI 10.1007/978-1-4419-7332-0_9,

$$F_2(x) = \int_{1/2}^{\sqrt{x}} x^2y^3\,dy = \frac{x^2y^4}{4}\bigg|_{1/2}^{\sqrt{x}} = \frac{x^2(\sqrt{x})^4}{4} - \frac{x^2/16}{4} = \frac{x^4}{4} - \frac{x^2}{64}.$$

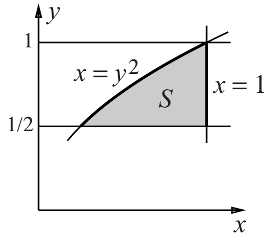


We can reverse the roles of the two variables if we start with a region T described in the following way,

$$T: \quad \begin{array}{c} c \le y \le d, \\ \alpha(y) \le x \le \beta(y). \end{array}$$

Then, for each y in the interval $[\alpha,\beta]$, the partial integral of $f(x,y)$ with respect to x from $\alpha(y)$ to $\beta(y)$ is

$$F_1(y) = \int_{\alpha(y)}^{\beta(y)} f(x,y)\,dx, \quad c \le y \le d.$$



This is continuous if α and β are. It can happen that a particular region can be described both ways. This is true for our example above:

$$S: \begin{array}{c} 1/4 \le x \le 1, \\ 1/2 \le y \le \sqrt{x}; \end{array} \quad \text{and also} \quad S: \begin{array}{c} 1/2 \le y \le 1, \\ y^2 \le x \le 1. \end{array}$$

Therefore

$$F_1(y) = \int_{y^2}^{1} x^2y^3\,dx = \frac{x^3y^3}{3}\bigg|_{y^2}^{1} = \frac{y^3 - y^9}{3},$$

so the two partial integrals of x^2y^3 are certainly different; they are even functions of different variables.

Iterated integrals

Now integrate each of the partial integrals $F_2(x)$ or $F_1(y)$ over its own domain:

$$\int_a^b F_2(x)\,dx = \int_a^b \left(\int_{\gamma(x)}^{\delta(x)} f(x,y)\,dy \right) dx,$$

$$\int_c^d F_1(y)\,dy = \int_c^d \left(\int_{\alpha(y)}^{\beta(y)} f(x,y)\,dx \right) dy.$$

Note that in each we have performed a repeated, or iterated, integration of the original function $f(x,y)$, first with respect to one variable and then the other. These are the **iterated integrals** of $f(x,y)$.

The iterated integrals have the same value

To illustrate, let us return to our example $f(x,y) = x^2y^3$ over the region S. We have

$$\int_{1/4}^{1} \left(\int_{1/2}^{\sqrt{x}} x^2y^3\,dy \right) dx = \int_{1/4}^{1} \left(\frac{x^4}{4} - \frac{x^2}{64} \right) dx = \frac{x^5}{20} - \frac{x^3}{192}\bigg|_{1/4}^{1} = \frac{459}{10240}$$

when the iteration is performed in one order, and

$$\int_{1/2}^{1} \left(\int_{y^2}^{1} x^2y^3\,dx \right) dy = \int_{1/2}^{1} \left(\frac{y^3 - y^9}{3} \right) dy = \frac{y^4}{12} - \frac{y^{10}}{30}\bigg|_{1/2}^{1} = \frac{459}{10240}$$

when it is performed in the other. These calculations suggest a general result: the two iterated integrals are always equal (Corollary 9.4, below). As we show, this happens because the iterated integrals, taken in either order, equal the double integral. Here is the statement and proof when the domain S is a rectangle.

Theorem 9.1. *Suppose $f(x,y)$ is continuous on the rectangle R defined by $a \le x \le b$, $c \le y \le d$; then*

$$\iint_R f(x,y)\,dA = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy.$$

Proof. We prove the double integral equals the first of the two iterated integrals; to show it also equals the second, interchange x and y. Let

$$F_2(x) = \int_c^d f(x,y)\,dy;$$

then we show

$$\iint_R f(x,y)\,dA = \int_a^b F_2(x)\,dx$$

by proving that, for any $\varepsilon > 0$,

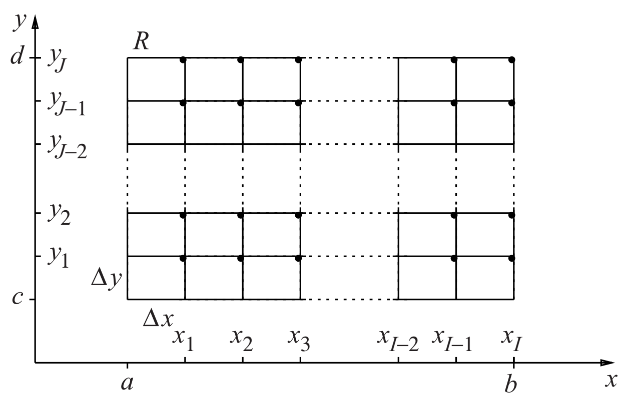
$$\left| \iint_R f(x,y)\,dA - \int_a^b F_2(x)\,dx \right| < \varepsilon.$$

We begin by subdividing R with a grid of congruent rectangles. For positive integers I and J, let

$$\Delta x = \frac{b-a}{I}, \quad \Delta y = \frac{d-c}{J},$$

and then define

$$\begin{aligned} x_1 &= a + \Delta x, & y_1 &= c + \Delta y, \\ x_i &= x_{i-1} + \Delta x, \quad i = 2, \ldots, I, & y_j &= y_{j-1} + \Delta y, \quad j = 2, \ldots, J. \end{aligned}$$



Each point (x_i, y_j) is the upper right corner of its cell. The mesh size of this grid, $\delta = \sqrt{(\Delta x)^2 + (\Delta y)^2}$, can be made as small as we wish by choosing both I and J sufficiently large.

Now suppose $\varepsilon > 0$ is given. Because f is continuous on R, it is integrable there (Theorem 8.35, p. 305). Consequently, all Riemann sums constructed using a grid with a sufficiently small mesh will be arbitrarily close to the value of the double integral of f over R. Therefore, we can choose I and J so large that

$$\left| \iint_R f(x,y)\,dA - \sum_{i=1}^{I} \sum_{j=1}^{J} f(x_i,y_j)\,\Delta x\,\Delta y \right| < \frac{\varepsilon}{3}.$$

The continuous function F_2 is likewise integrable over $[a,b]$, so Riemann sums for its integral are also arbitrarily close to the value of the integral if the partition of the x-interval $[a,b]$ is fine enough. Therefore, by increasing the size of I, if necessary, we can make the inequality

$$\left| \int_a^b F_2(x)\,dx - \sum_{i=1}^{I} F_2(x_i)\,\Delta x \right| < \frac{\varepsilon}{3}$$

hold as well. Finally, each $F_2(x_i)$ is itself an integral,

$$F_2(x_i) = \int_c^d f(x_i,y)\,dy, \quad i = 1,\dots,I,$$

and thus has Riemann sum approximations. Therefore, after a sufficiently large I has been fixed, we can then increase the size of J, if necesary, so that all Riemann sums for each of the integrals $F_2(x_1), \dots, F_2(x_I)$ will be arbitrarily close to the value of that integral:

$$\left| F_2(x_i) - \sum_{j=1}^{J} f(x_i,y_j)\,\Delta y \right| < \frac{\varepsilon}{3(b-a)}, \quad i = 1,\dots,I.$$

Now consider the inequality we seek to prove. As often happens in such a proof, we begin with a telescoping sum,

$$\begin{aligned}
\iint_R f(x,y)\,dA - \int_a^b F_2(x)\,dx = {} & \iint_R f(x,y)\,dA - \sum_{i=1}^{I} \sum_{j=1}^{J} f(x_i,y_j)\,\Delta x\,\Delta y \\
& + \sum_{i=1}^{I} \left(\sum_{j=1}^{J} f(x_i,y_j)\,\Delta y - F_2(x_i) \right) \Delta x \\
& + \sum_{i=1}^{I} F_2(x_i)\,\Delta x - \int_a^b F_2(x)\,dx,
\end{aligned}$$

and then apply the triangle inequality:

$$\left| \iint_R f(x,y)\,dA - \int_a^b F_2(x)\,dx \right| \le \left| \iint_R f(x,y)\,dA - \sum_{i=1}^{I}\sum_{j=1}^{J} f(x_i,y_j)\,\Delta x\,\Delta y \right|$$
$$+ \left| \sum_{i=1}^{I} \left(\sum_{j=1}^{J} f(x_i,y_j)\,\Delta y - F_2(x_i) \right) \Delta x \right|$$
$$+ \left| \sum_{i=1}^{I} F_2(x_i)\,\Delta x - \int_a^b F_2(x)\,dx \right|.$$

For I and J large enough, the first term on the right is bounded by $\varepsilon/3$, and so is the third term; we claim the same is true for the second. We have

$$\left| \sum_{i=1}^{I} \left(\sum_{j=1}^{J} f(x_i,y_j)\,\Delta y - F_2(x_i) \right) \Delta x \right| \le \sum_{i=1}^{I} \left| \sum_{j=1}^{J} f(x_i,y_j)\,\Delta y - F_2(x_i) \right| \Delta x$$
$$< \sum_{i=1}^{I} \frac{\varepsilon}{3(b-a)}\,\Delta x = \frac{\varepsilon}{3(b-a)} \cdot I\Delta x = \frac{\varepsilon}{3},$$

as claimed. By what has been said above, this proves the theorem. □

Corollary 9.2 *In the iterated integration of a continuous function with constant limits of integration, the order of integration can be reversed.* □

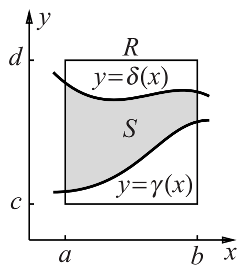
Order of integration

Theorem 9.3. *Let S be the region defined by $a \le x \le b$, $\gamma(x) \le y \le \delta(x)$, where $\gamma(x)$ and $\delta(x)$ are continuous functions of x on $[a,b]$. Let $f(x,y)$ be continuous on S; then*

$$\iint_S f(x,y)\,dA = \int_a^b \left(\int_{\gamma(x)}^{\delta(x)} f(x,y)\,dy \right) dx.$$

Proof. This theorem is similar to Theorem 9.1, and can be proven in essentially the same way. We begin by constructing a rectangle

$$R: \begin{array}{l} a \le x \le b, \\ c \le y \le d, \end{array}$$



where $c \le \gamma(x)$ and $\delta(x) \le d$, for all x in $[a,b]$. Now R contains S, and if we extend $f(x,y)$ in the usual way by having $f(x,y) = 0$ outside S, then f is continuous everywhere in R, except (in general) on the graphs $y = \gamma(x)$ and $y = \delta(x)$.

By Theorem 8.10 (p. 284), these graphs form a set Z of area zero. Because f is continuous on $R \setminus Z$, it is integrable on R by Theorem 8.35 (p. 305). Moreover, because S and $R \setminus S$ are nonoverlapping sets (on each of which f is integrable), we have

$$\iint\limits_R f(x,y)\,dA = \iint\limits_S f(x,y)\,dA + \iint\limits_{R\setminus S} f(x,y)\,dA = \iint\limits_S f(x,y)\,dA,$$

because the integral is additive (Theorem 8.27) and $f = 0$ on $R \setminus S$.

Now fix x; then $f(x,y)$ is a bounded function of y on the interval $[c,d]$, and is continuous except at the two points where $y=\gamma(x)$ and $y=\delta(x)$. Therefore, the partial integral of f with respect to y over $[c,d]$ exists and equals the integral over the smaller interval $[\gamma(x),\delta(x)]$ (because $f=0$ outside that smaller interval). Let

$$F_2(x)=\int_c^d f(x,y)\,dy=\int_{\gamma(x)}^{\delta(x)} f(x,y)\,dy$$

denote the common value; then $F_2(x)$ is a continuous function of x on $[a,b]$. To prove the theorem it is sufficient to show that

$$\iint_R f(x,y)\,dA=\int_a^b\left(\int_c^d f(x,y)\,dy\right)dx=\int_a^b F_2(x)\,dx.$$

Although this equation does not follow directly from the statement of Theorem 9.1 (because f is not continuous everywhere on R), we can show that it does follow from the proof.

Cover R with a grid of rectangles whose width is $\Delta x=(b-a)/I$ and whose height is $\Delta y=(d-c)/J$, and define (x_i,y_j) for $i=1,\ldots,I$, $j=1,\ldots,J$, as in that proof (cf. p. 319). With these choices we can now construct the various Riemann sums that appear in the following inequality, taken from the same proof:

$$\begin{aligned}\left|\iint_R f(x,y)\,dA-\int_a^b F_2(x)\,dx\right| \le{}& \left|\iint_R f(x,y)\,dA-\sum_{i=1}^{I}\sum_{j=1}^{J} f(x_i,y_j)\,\Delta x\,\Delta y\right|\\ &+\left|\sum_{i=1}^{I}\left(\sum_{j=1}^{J} f(x_i,y_j)\Delta y-F_2(x_i)\right)\Delta x\right|\\ &+\left|\sum_{i=1}^{I}F_2(x_i)\,\Delta x-\int_a^b F_2(x)\,dx\right|.\end{aligned}$$

Now choose $\varepsilon>0$. Then, because f is integrable on R, because $f(x_i,y)$ is integrable with respect to y for each $i=1,\ldots,I$, and because F_2 is integrable on $[a,b]$, we can choose I and J large enough that each of the terms on the right is less than $\varepsilon/3$. Because $\varepsilon>0$ is arbitrary, the left-hand side of the inequality must equal zero. By what has been said above, this completes the proof. □

Interchanging the variables

The theorem holds with the roles of x and y reversed. That is, if $f(x,y)$ is continuous over the region $T: c\le y\le d$, $\alpha(y)\le x\le\beta(y)$, then

$$\iint_T f(x,y)\,dA=\int_c^d\left(\int_{\alpha(y)}^{\beta(y)} f(x,y)\,dx\right)dy.$$

This implies the following corollary.

Corollary 9.4 *Suppose $f(x,y)$ is continuous over a region S that has the alternate descriptions*

$$S: \begin{array}{c} a \le x \le b, \\ \gamma(x) \le y \le \delta(x), \end{array} \qquad S: \begin{array}{c} c \le y \le d, \\ \alpha(y) \le x \le \beta(y); \end{array}$$

then

$$\int_a^b \left(\int_{\gamma(x)}^{\delta(x)} f(x,y)\,dy \right) dx = \int_c^d \left(\int_{\alpha(y)}^{\beta(y)} f(x,y)\,dx \right) dy.$$

Proof. Both iterated integrals equal the double integral $\iint_S f(x,y)\,dA$. □

Here are two common ways to write an iterated integral that dispense with the large parentheses: Notation

$$\int_a^b \left(\int_{\gamma(x)}^{\delta(x)} f(x,y)\,dy \right) dx = \int_a^b \int_{\gamma(x)}^{\delta(x)} f(x,y)\,dy\,dx = \int_a^b dx \int_{\gamma(x)}^{\delta(x)} f(x,y)\,dy;$$
$$\int_c^d \left(\int_{\alpha(y)}^{\beta(y)} f(x,y)\,dx \right) dy = \int_c^d \int_{\alpha(y)}^{\beta(y)} f(x,y)\,dx\,dy = \int_c^d dy \int_{\alpha(y)}^{\beta(y)} f(x,y)\,dx.$$

Most often, we use the first; the order of dx and dy indicates the order in which the partial integrations are to be carried out.

The gravitational field by iterated integrals

A good example of the way we can evaluate a double integral with iterated single integrals is provided by the gravitational field of a square plate (pp. 270–272) at a point above the center of the plate:

$$\text{field at } a = \iint_S \frac{-a\,dA}{(x^2+y^2+a^2)^{3/2}}, \qquad S: \begin{array}{c} 0 \le x \le R, \\ 0 \le y \le R. \end{array}$$

(In this expression we have used $4G\rho = 1$ and we have written the element of area as dA.) As an iterated integral, the field is

$$\text{field at } a = \int_0^R \int_0^R \frac{-a\,dy}{(x^2+y^2+a^2)^{3/2}}\,dx.$$

The first integration, with respect to y, can be done with the pullback substitution $y = \sqrt{x^2+a^2}\tan\theta$ (see Exercise 9.1); the result is

$$\int_0^R \frac{-a\,dy}{(x^2+y^2+a^2)^{3/2}} = \left. \frac{-ay}{(x^2+a^2)(x^2+y^2+a^2)^{1/2}} \right|_0^R$$
$$= \frac{-aR}{(x^2+a^2)(x^2+R^2+a^2)^{1/2}}.$$

A closed-form formula for field strength

The antidifferentiation needed for the second integration is more readily done with a table or a computer algebra system:

$$\begin{aligned}\text{field at } a &= \int_0^R \frac{-aR\,dx}{(x^2+a^2)(x^2+R^2+a^2)^{1/2}}\\ &= -\arctan\left(\frac{Rx}{a\sqrt{x^2+R^2+a^2}}\right)\Bigg|_0^R = -\arctan\left(\frac{R^2}{a\sqrt{2R^2+a^2}}\right).\end{aligned}$$

This gives us a closed-form expression for the field that can shed light on the numerical results we found in Chapter 8. First note (see Exercise 9.2) that

$$\frac{R^2}{a\sqrt{2R^2+a^2}} = \frac{R^2}{aR\sqrt{2}\sqrt{1+(a/2R)^2}} = \frac{R}{a\sqrt{2}} + O(a/R) \text{ as } a/R \to 0;$$

from this we obtain the approximation

$$\text{field at } a \approx -\arctan\left(\frac{R}{a\sqrt{2}}\right).$$

Comparing estimates of field strength

The following table gives the field strength (with $R = 32$) as determined by a Riemann sum (the numerical estimate from Chapter 8.1), by the approximation immediately above, and by the complete formula derived from the iterated integrals.

a	Numerical Estimate	$-\arctan\left(\frac{R}{a\sqrt{2}}\right)$	$-\arctan\left(\frac{R^2}{a\sqrt{2R^2+a^2}}\right)$
0.2	$-1.561\,957\ldots$	$-1.561\,957\,722$	$-1.561\,957\,636$
0.1	$-1.566\ldots$	$-1.566\,376\,938$	$-1.566\,376\,927$
0.05	$-1.568\ldots$	$-1.568\,586\,622$	$-1.568\,586\,620$

We also noted in Chapter 8 that the computations suggest the field becomes constant as the size of the plate increases, that is, as $R \to \infty$. The formula confirms this: because

$$\lim_{R\to\infty} \frac{R^2}{a\sqrt{2R^2+a^2}} = \lim_{R\to\infty} \frac{R}{a\sqrt{2}} = \infty,$$

we can say

$$\text{field of infinite plate at height } a = -\arctan(\infty) = -\frac{\pi}{2} = -1.570\,796\,327,$$

a value that is independent of the height a.

The circular plate in Cartesian coordinates

To continue our illustration, let us evaluate the field of a circular plate using Cartesian coordinates. Of course polar coordinates lead to a simpler evaluation; however, we have already done that (cf. pp. 273–275), and got the result

$$\text{field at } a = 2\pi G\rho\left(\frac{a}{\sqrt{R^2+a^2}} - 1\right).$$

Cartesian coordinates give us the chance to use iterated integrals to compare calculations.

We can define the region occupied by the circular plate as

$$S: \quad \begin{array}{c} -R \leq x \leq R, \\ -\sqrt{R^2 - x^2} \leq y \leq \sqrt{R^2 - x^2}; \end{array}$$

therefore,

$$\text{field at } a = \iint_S \frac{-G\rho a \, dA}{(x^2 + y^2 + a^2)^{3/2}} = -G\rho a \int_{-R}^{R} \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{dy}{(x^2+y^2+a^2)^{3/2}} \, dx.$$

We have already (p. 323) computed the inner antiderivative, and found

$$\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{dy}{(x^2+y^2+a^2)^{3/2}} = \frac{y}{(x^2+a^2)(x^2+y^2+a^2)^{1/2}} \Bigg|_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}}$$

$$= \frac{2\sqrt{R^2-x^2}}{(x^2+a^2)(R^2+a^2)^{1/2}}.$$

Consequently (again resorting to tables or a computer algebra system to find the antiderivative),

$$\text{field at } a = \frac{-2G\rho a}{\sqrt{R^2+a^2}} \int_{-R}^{R} \frac{\sqrt{R^2-x^2}}{x^2+a^2} \, dx$$

$$= \frac{-2G\rho a}{\sqrt{R^2+a^2}} \left(-\arctan\left(\frac{x}{\sqrt{R^2-x^2}}\right) + \frac{\sqrt{R^2+a^2}}{a} \arctan\left(\frac{x\sqrt{R^2+a^2}}{a\sqrt{R^2-x^2}}\right)\right)\Bigg|_{-R}^{R}$$

Because

$$\arctan\left(\frac{x}{\sqrt{R^2-x^2}}\right)\Bigg|_{x=\pm R} = \arctan(\pm\infty) = \pm\frac{\pi}{2},$$

and likewise

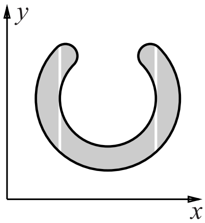
$$\arctan\left(\frac{x\sqrt{R^2+a^2}}{a\sqrt{R^2-x^2}}\right)\Bigg|_{x=\pm R} = \arctan(\pm\infty) = \pm\frac{\pi}{2},$$

we have

$$\text{field at } a = \frac{-2G\rho a}{\sqrt{R^2+a^2}} \left(-\frac{\pi}{2} + \frac{\sqrt{R^2+a^2}}{a} \cdot \frac{\pi}{2} - \frac{\pi}{2} + \frac{\sqrt{R^2+a^2}}{a} \cdot \frac{\pi}{2} \right)$$

$$= \frac{2\pi G\rho a}{\sqrt{R^2+a^2}} \left(1 - \frac{\sqrt{R^2+a^2}}{a} \right) = 2\pi G\rho \left(\frac{a}{\sqrt{R^2+a^2}} - 1 \right),$$

in agreement with the computation using polar coordinates.



Theorem 9.3 allows us to reduce a double integral to iterated single integrals; however, it holds only for a restricted class of regions. For example, the horseshoe-shaped region shown in the margin does not meet the restriction (at least for Cartesian coordinates). Nevertheless, it is the union of a finite number of nonoverlapping sets (e.g., the five whose boundaries are shown by the white lines) that, separately, do meet the restriction. The following theorem asserts that this is enough.

Theorem 9.5. *Suppose $f(x,y)$ is continuous on a bounded region R that is the union of nonoverlapping sets $S_1, \dots, S_p, T_1, \dots, T_q$ of the form*

$$S_i: \begin{array}{c} a_i \le x \le b_i, \\ \gamma_i(x) \le y \le \delta_i(x), \end{array} \qquad T_j: \begin{array}{c} c_j \le x \le d_j, \\ \alpha_j(y) \le x \le \beta_j(y); \end{array}$$

then

$$\iint_R f(x,y)\,dA = \int_{a_1}^{b_1}\int_{\gamma_1(x)}^{\delta_1(x)} f(x,y)\,dy\,dx + \cdots + \int_{a_p}^{b_p}\int_{\gamma_p(x)}^{\delta_p(x)} f(x,y)\,dy\,dx$$
$$+ \int_{c_1}^{d_1}\int_{\alpha_1(y)}^{\beta_1(y)} f(x,y)\,dx\,dy + \cdots + \int_{c_q}^{d_q}\int_{\alpha_q(y)}^{\beta_q(y)} f(x,y)\,dx\,dy.$$

Proof. Because $R = S_1 \cup \cdots \cup S_p \cup T_1 \cup \cdots \cup T_q$ is a decomposition into nonoverlapping sets, we have

$$\iint_R f(x,y)\,dA = \iint_{S_1} f(x,y)\,dA + \cdots + \iint_{S_p} f(x,y)\,dA$$
$$+ \iint_{T_1} f(x,y)\,dA + \cdots + \iint_{T_q} f(x,y)\,dA$$

by the additivity of the integral (Theorem 8.27, p. 298). The result now follows by reducing each double integral on the right to the appropriate iterated integral. □

9.2 Improper integrals

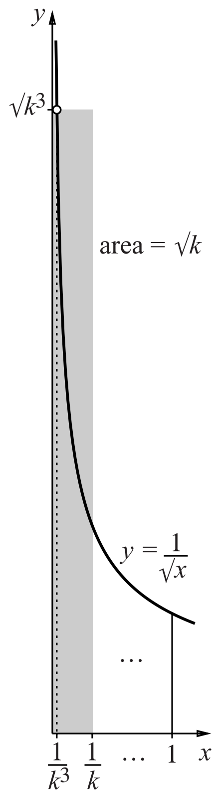
The integral of a function is defined using values of the function and the sizes of small regions, so it is natural to deal only with bounded functions over closed bounded regions. However, scientific and mathematical questions just as naturally involve unbounded functions and unbounded regions, so it is important to extend the process of integration to these more general settings. Such extensions are called *improper integrals*; they are evaluated as limits of "proper" integrals.

A function without a *proper* integral

By a **proper** integral we mean one whose value is determined, in principle, as a limit of Riemann sums. For example,

$$\int_0^1 \frac{dx}{\sqrt{x}}$$

is not a proper integral in this sense. Of course, in any Riemann sum for $f(x) = 1/\sqrt{x}$ we must avoid the point $x = 0$ because $f(0)$ is undefined. However, this is not the heart of the problem. To see what is, subdivide the interval $(0, 1]$ into equal subintervals of length $\Delta x = 1/k$, for any positive integer k. Then form a Riemann sum Σ whose first term is



$$f(1/k^3) \cdot \Delta x = \frac{1}{\sqrt{1/k^3}} \cdot \frac{1}{k} = \sqrt{k}.$$

For the remaining terms in Σ, make any valid choices. Because those terms are all positive, we have

$$\Sigma \geq \sqrt{k} \to \infty \text{ as } k \to \infty.$$

Because these Riemann sums do not converge, there is no integral. The problem is not that $f(0)$ is undefined but that $f(x)$ is unbounded on that first interval $(0, \Delta x]$. (We have already used this example in a slightly different form for a similar purpose on page 299.) Theorem 8.36 (p. 307) confirms this; it says that if $f(x)$ were bounded on $(0, 1]$, it would be integrable on $[0, 1]$.

On any smaller interval $[a, 1] \subset (0, 1]$, $f(x)$ is bounded and continuous and therefore integrable. Because this integral gives the area under the part of the graph that lies above the interval $[a, 1]$ on the x-axis, and because that area increases monotonically as $[a, 1] \to (0, 1]$, it seems reasonable to define

$$\int_0^1 \frac{dx}{\sqrt{x}} = \lim_{a \to 0} \int_a^1 \frac{dx}{\sqrt{x}}$$

if the values on the right converge to a finite limit. In fact,

$$\int_a^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\,\Big|_a^1 = 2\big(1 - \sqrt{a}\big) \to 2.$$

Thus we can say the improper integral "converges" and has the value 2:

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2.$$

The need for improper integrals

Is this argument unnecesarily elaborate and painstaking? It would appear that we could just write

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\,\Big|_0^1 = 2$$

and get the correct value. However, this computation uses the fundamental theorem of calculus, which says that

$$\int_a^b f(x)\,dx = F(x)\,\Big|_a^b$$

when $f(x)$ is continuous on $[a,b]$ and $F'(x) = f(x)$ there. But in our case the fundamental theorem fails to apply, because $f(x) = 1\sqrt{x}$ is not continuous on $[0,1]$. To integrate $1/\sqrt{x}$ over $(0,1]$, we must extend the definition of *integral* in some fashion; the ordinary, or "proper," integral does not exist.

Unbounded functions on bounded domains

More generally, to define the improper integral of a function $f(x)$ that is continuous but unbounded on the open interval $a < x < b$, first take $a < \alpha < \beta < b$ and compute the ordinary integral

$$I(\alpha,\beta) = \int_\alpha^\beta f(x)\,dx$$

as a function of its endpoints α and β. (The integral exists because f is bounded and continuous on $[\alpha,\beta]$.) If the values $I(\alpha,\beta)$ have a finite limit as $\alpha \to a$ and $\beta \to b$, then the **improper integral** converges and its value is that limit:

$$\int_a^b f(x)\,dx = \lim_{\substack{\alpha\to a\\ \beta\to b}} \int_\alpha^\beta f(x)\,dx.$$

More generally, if $f(x)$ is continuous on the closed interval $[a,b]$ except for the points $c_1 < c_2 < \cdots < c_k$ at which it becomes unbounded (and $a \le c_1$, $c_k \le b$), then we define the **improper integral**

$$\int_a^b f(x)\,dx = \int_a^{c_1} f(x)\,dx + \int_{c_1}^{c_2} f(x)\,dx + \cdots + \int_{c_k}^b f(x)\,dx$$

if all the intermediate improper integrals on the right converge.

Note that the intermediate improper integrals must converge separately and independently of each other. For example, the integral of $1/x$ over $[-1,1]$ is improper because $1/x$ is unbounded as $x \to 0$, so we must write

$$\int_{-1}^1 \frac{dx}{x} = \int_{-1}^0 \frac{dx}{x} + \int_0^1 \frac{dx}{x}.$$

But this fails to converge, because neither improper integral on the right converges:

$$\int_{-1}^0 \frac{dx}{x} = \lim_{\beta\to 0} \int_{-1}^\beta \frac{dx}{x} = \lim_{\beta\to 0} \ln|\beta| = -\infty,$$

$$\int_0^1 \frac{dx}{x} = \lim_{\alpha\to 0} \int_\alpha^1 \frac{dx}{x} = \lim_{\alpha\to 0} -\ln\alpha = +\infty.$$

If we were to link the two intermediate integrals in the following way,

$$\int_{-1}^1 \frac{dx}{x} = \lim_{\alpha\to 0}\left(\int_{-1}^{-\alpha} \frac{dx}{x} + \int_\alpha^1 \frac{dx}{x}\right),$$

we would reach the false conclusion

$$\int_{-1}^{1} \frac{dx}{x} = \lim_{\alpha \to 0} \left(\ln \alpha - \ln \alpha \right) = 0.$$

Unbounded domains

There are improper integrals for unbounded domains as well as for unbounded functions. If we try to calculate a Riemann sum for a function over an unbounded domain, one of the cells must have infinite size. However, there is a natural way to define an improper integral. Assuming that f is bounded and integrable on every finite subinterval of $a \leq x < \infty$ or $-\infty < x \leq b$, respectively, we set

$$\int_a^\infty f(x)\,dx = \lim_{B \to \infty} \int_a^B f(x)\,dx, \quad \int_{-\infty}^b f(x)\,dx = \lim_{A \to -\infty} \int_A^b f(x)\,dx.$$

Thus, for example,

$$\int_0^\infty \frac{dx}{1+x^2} = \lim_{B \to \infty} \int_0^B \frac{dx}{1+x^2} = \lim_{B \to \infty} \arctan x \Big|_0^B = \lim_{B \to \infty} \arctan B = \frac{\pi}{2}.$$

We sometimes find a sequence like this abbreviated to

$$\int_0^\infty \frac{dx}{1+x^2} = \arctan x \Big|_0^\infty = \arctan \infty = \frac{\pi}{2},$$

but we must always understand that the briefer calculation depends on the validity of the longer one.

Improper double integrals

We can now turn from single to double integrals. Suppose R is a closed bounded region with area in $\mathbb{R}^2$ and Z a set of area zero. If $f(x,y)$ is bounded and continuous on $R \setminus Z$, then we know f is "properly" integrable over R (Theorem 8.35, p. 305). But suppose we allow f to become unbounded on $R \setminus Z$ while remaining continuous there; can we define the *improper* integral of f over R?

Monotonic sequences of subregions

Single integrals suggest that we consider a monotonically increasing sequence $S_1 \subseteq S_2 \subseteq \cdots$ of closed subsets of $R \setminus Z$ for which $A(S_k) \to A(R)$ as $k \to \infty$. On each S_k, f is continuous and bounded, so it has a "proper" integral

$$I_k = \iint_{S_k} f(x,y)\,dA.$$

If the sequence $I_1, I_2, \ldots$ has a finite limit, I, we would like to say that the improper integral of f over R converges and has the value I.

However, in any definition that involves choices (as this does with the sequence $S_1, S_2, \ldots$), we must make certain that the outcome does not depend on the choices made. Thus, if $T_1 \subseteq T_2 \subseteq \cdots$ is another sequence of closed subsets of $R \setminus Z$ for which $A(T_m) \to A(R)$, and

$$J_m = \iint_{T_m} f(x,y)\,dA,$$

then we must verify that the sequence $J_1, J_2, \ldots$ also has a finite limit, J, and then that $J = I$.

Here is an example that illustrates how much variability there can be in the outcome. Consider $f(x,y) = 1/x$ on the unit square $R : -1 \le x,y \le 1$. Of course, f is undefined on the y-axis $Z : x = 0$, and is continuous but unbounded on $R \setminus Z$.

Let V_k be the infinite strip $-1/k < x < 1/k$ that is centered symmetrically on the y-axis, and let $S_k = R \setminus V_k$. The sets S_k are nested monotonically and $A(S_k) = 4 - 4/k \to 4 = A(R)$; furthermore,

$$I_k = \iint_{S_k} \frac{dA}{x} = \int_{-1}^{1}\int_{-1}^{-1/m} \frac{dx}{x}\,dy + \int_{-1}^{1}\int_{1/m}^{1} \frac{dx}{x}\,dy = 0$$

by symmetry. In fact,

$$\int_{-1}^{1}\int_{-1}^{-1/m} \frac{dx}{x}\,dy = -\int_{-1}^{1}\int_{1/m}^{1} \frac{dx}{x}\,dy;$$

that is, the contributions to I_k from the left-half plane and the right-half plane exactly cancel.

By contrast, let W_m be the *asymmetric* strip $-1/m < x < 1/m^2$, and let $T_m = R \setminus W_m$. Now $A(T_m) = 4 - 2/m - 2/m^2$, so we still have $A(T_m) \to A(R)$ as $m \to \infty$. But this time the cancellation is incomplete; the integral over T_m reduces to

$$J_m = \iint_{T_m} \frac{dA}{x} = \int_{-1}^{1}\int_{1/m^2}^{1/m} \frac{dx}{x}\,dy = 2\big(\ln 1/m - \ln 1/m^2\big) = 2\ln m.$$

(See the exercises.) Because $J_m \to \infty$, the two sets of integrals do not converge the same way, so the improper integral fails to exist.

We can attribute the variability of the outcomes to the way f changes sign on R. If we replace f by $|f|$ then that variability disappears. In fact, the integrals

$$\widehat{I}_k = \iint_{S_k} |f(x,y)|\,dA \quad\text{and}\quad \widehat{J}_m = \iint_{T_m} |f(x,y)|\,dA$$

are now both unbounded monotonic increasing sequences of numbers.

A comparison with infinite series

Compare what is happening with the integrals to what can happen with certain infinite series. For example,

$$1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \tfrac{1}{5} - \tfrac{1}{6} + \tfrac{1}{7} - \tfrac{1}{8} + \tfrac{1}{9} + \cdots = \ln 2;$$

that is, the sequence of partial sums $1,\ 1 - \frac{1}{2},\ 1 - \frac{1}{2} + \frac{1}{3},\ \ldots$ has the limiting value $\ln 2$. But a rearrangement of the terms can change the sum:

$$1 + \tfrac{1}{3} - \tfrac{1}{2} + \tfrac{1}{5} + \tfrac{1}{7} - \tfrac{1}{4} + \tfrac{1}{9} + \tfrac{1}{11} - \tfrac{1}{6} + \cdots = \tfrac{3}{2}\ln 2.$$

(Instead of strictly alternating positive and negative terms, the new series includes two positive terms for every negative one; see the exercises.) Choosing the order of terms here is analogous to choosing how the subsets S_k and T_m expand to fill out $R \setminus Z$. In both cases, different choices lead to different outcomes. Finally, replacing

f by $|f|$ is analogous to making all terms in the series be positive. In this case we get the harmonic series:

$$1+\tfrac{1}{2}+\tfrac{1}{3}+\tfrac{1}{4}+\tfrac{1}{5}+\tfrac{1}{6}+\tfrac{1}{7}+\tfrac{1}{8}+\tfrac{1}{9}+\cdots=\infty.$$

Conditional convergence

The harmonic series diverges; its sequence of partial sums is monotonically increasing and unbounded. Because the alternating series for $\ln 2$ converges but the related series of absolute values (the harmonic series) does not, we say the series for $\ln 2$ is **conditionally convergent**. By contrast, both the alternating (geometric) series

$$1-\tfrac{1}{2}+\tfrac{1}{4}-\tfrac{1}{8}+\tfrac{1}{16}-\tfrac{1}{32}+\tfrac{1}{64}-\tfrac{1}{128}+\cdots=\tfrac{2}{3}$$

and the corresponding series of absolute values

$$1+\tfrac{1}{2}+\tfrac{1}{4}+\tfrac{1}{8}+\tfrac{1}{16}+\tfrac{1}{32}+\tfrac{1}{64}+\tfrac{1}{128}+\cdots=2$$

Absolute convergence

converge, so we say the alternating series for $\frac{2}{3}$ is **absolutely convergent**. An absolutely convergent series is more robust: rearranging its terms does not change its value.

The improper integral we define is the analogue of an absolutely convergent series; its value will not change when we change the way the region $R\setminus Z$ is "filled up" by closed subsets S_k or T_m.

Theorem 9.6. *Let R be a closed bounded set with area, let Z be a set with area zero, and let $S_1,S_2,\ldots$ be a monotonic increasing sequence of closed subsets of $R\setminus Z$ for which $A(S_k)\to A(R)$ as $k\to\infty$. Suppose $f(x,y)$ is continuous but unbounded on $R\setminus Z$, but*

$$\widehat{I}_k=\iint_{S_k}|f(x,y)|\,dA\le B$$

for some bound B and for all k. Then the numbers

$$I_k=\iint_{S_k}f(x,y)\,dA$$

have a finite limit I as $k\to\infty$. Furthermore, the value of I is independent of the way the closed subsets S_k are chosen.

Proof. To show that various limits exist we use the *Cauchy convergence criterion*: the sequence $a_1,a_2,\ldots$ of real numbers has a finite limit if and only if, for every $\varepsilon>0$, there is an $N=N(\varepsilon)$ such that

$$i,j>N \implies |a_i-a_j|<\varepsilon.$$

The criterion says that the limit exists if all the a_i are arbitrarily close to one another when i is sufficiently large; for a proof see a text on analysis.

We first show that the integrals I_k converge for a particular collection of sets S_k. Suppose $i>j$; then $S_i\supseteq S_j$ so we have

$$\iint\limits_{S_i} f(x,y)\,dA = \iint\limits_{S_j} f(x,y)\,dA + \iint\limits_{S_i \setminus S_j} f(x,y)\,dA,$$

and similarly for $|f(x,y)|$. Because $|f(x,y)| \geq 0$, the sequence $\widehat{I}_1, \widehat{I}_2, \ldots$ is monotonic increasing; by hypothesis, it is bounded above so it has a finite limit. Therefore, by the Cauchy convergence criterion, we know that for any $\varepsilon > 0$, we can find an N for which

$$\widehat{I}_i - \widehat{I}_j = |\widehat{I}_i - \widehat{I}_j| < \varepsilon$$

whenever $i > j > N$. But

$$\widehat{I}_i - \widehat{I}_j = \iint\limits_{S_i} |f(x,y)|\,dA - \iint\limits_{S_j} |f(x,y)|\,dA = \iint\limits_{S_i \setminus S_j} |f(x,y)|\,dA,$$

so

$$\iint\limits_{S_i \setminus S_j} |f(x,y)|\,dA < \varepsilon.$$

For any closed set Q in $R \setminus Z$, we have

$$\left| \iint_Q f(x,y)\,dA \right| \leq \iint_Q |f(x,y)|\,dA.$$

Therefore, when $i > j > N$ we have

$$|I_i - I_j| = \left| \iint_{S_i} f(x,y)\,dA - \iint_{S_j} f(x,y)\,dA \right| = \left| \iint\limits_{S_i \setminus S_j} f(x,y)\,dA \right| \leq \iint\limits_{S_i \setminus S_j} |f(x,y)|\,dA < \varepsilon.$$

By the Cauchy convergence criterion, the sequence $I_1, I_2, \ldots$ converges to a finite limit.

Now let $T_1 \subseteq T_2 \subseteq \cdots$ be another sequence of closed sets with $A(T_m) \to A(R)$. We claim

$$\widehat{J}_m = \iint_{T_m} |f(x,y)|\,dA \leq B$$

for the same bound B. The foregoing proof would then imply that the sequence

$$J_m = \iint_{T_m} f(x,y)\,dA$$

also has a limit, J.

To prove the claim, let T be any one of the sets T_m. We know $f(x,y)$ is bounded on T: $|f(x,y)| \leq M$ for some M (that depends on T). Because $T \setminus (T \cap S_k) = T \setminus S_k$, we have

$$\iint\limits_{T} g(x,y)\,dA - \iint\limits_{T \cap S_k} g(x,y)\,dA = \iint\limits_{T \setminus S_k} g(x,y)\,dA,$$

where $g(x,y)$ stands for either $f(x,y)$ or $|f(x,y)|$. Therefore,

$$\left| \iint_T g(x,y)\,dA - \iint_{T\cap S_k} g(x,y)\,dA \right| \le \iint_{T\setminus S_k} |g(x,y)|\,dA$$

$$= \iint_{T\setminus S_k} |f(x,y)|\,dA \le M\cdot A(T\setminus S_k)$$

by Corollary 8.30. Now $A(T\setminus S_k)\to 0$ as $k\to\infty$, because $T\setminus S_k\subseteq R\setminus S_k$ and we have

$$A(T\setminus S_k)\le A(R\setminus S_k)=A(R)-A(S_k)$$

by Lemma 8.3, and $A(S_k)\to A(R)$ by hypothesis. It follows that

$$\iint_T f(x,y)\,dA = \lim_{k\to\infty}\iint_{T\cap S_k} f(x,y)\,dA \quad\text{and}\quad \iint_T |f(x,y)|\,dA = \lim_{k\to\infty}\iint_{T\cap S_k} |f(x,y)|\,dA.$$

Using the second equation and $T\cap S_k\subseteq S_k$, we find

$$\iint_T |f(x,y)|\,dA = \lim_{k\to\infty}\iint_{T\cap S_k} |f(x,y)|\,dA \le \lim_{k\to\infty}\iint_{S_k} |f(x,y)|\,dA \le B,$$

proving the claim and showing that the limit J exists.

To prove that $I=J$, we begin by noting that

$$\left|\iint_T f(x,y)\,dA - \iint_{T\cap S_j} f(x,y)\,dA\right| = \lim_{i\to\infty}\left|\iint_{T\cap S_i} f(x,y)\,dA - \iint_{T\cap S_j} f(x,y)\,dA\right|$$

$$= \lim_{i\to\infty}\left|\iint_{T\cap(S_i\setminus S_j)} f(x,y)\,dA\right|$$

$$\le \lim_{i\to\infty}\iint_{S_i\setminus S_j} |f(x,y)|\,dA \le \varepsilon,$$

a result that holds for all $j>N$, where N was the number provided by the Cauchy convergence criterion for the sequence $\widehat{I}_1,\widehat{I}_2,\dots$. (In particular, this N is independent of the choice of T.) The initial equality uses

$$\iint_T f(x,y)\,dA = \lim_{i\to\infty}\iint_{T\cap S_i} f(x,y)\,dA.$$

Furthermore, because the sequence $\widehat{J}_1, \widehat{J}_2, \ldots$ also converges, there is a number L for which

$$|\widehat{J}_l - \widehat{J}_m| < \varepsilon$$

for all $l > m > L$. Reversing the roles of S_k and T_m we can therefore conclude that

$$\left| \iint\limits_{S} f(x,y)\,dA - \iint\limits_{S \cap T_m} f(x,y)\,dA \right| \leq \varepsilon$$

for every $m > L$ and any closed subset S of $R \setminus Z$.

Finally, the telescoping sum

$$\begin{aligned} I - J = I - \iint\limits_{S_k} f(x,y)\,dA + \iint\limits_{S_k} f(x,y)\,dA - \iint\limits_{S_k \cap T_m} f(x,y)\,dA \\ + \iint\limits_{S_k \cap T_m} f(x,y)\,dA - \iint\limits_{T_m} f(x,y)\,dA + \iint\limits_{T_m} f(x,y)\,dA - J \end{aligned}$$

leads to the triangle inequality

$$\begin{aligned} |I - J| \leq \left| I - \iint\limits_{S_k} f(x,y)\,dA \right| + \left| \iint\limits_{S_k} f(x,y)\,dA - \iint\limits_{S_k \cap T_m} f(x,y)\,dA \right| \\ + \left| \iint\limits_{S_k \cap T_m} f(x,y)\,dA - \iint\limits_{T_m} f(x,y)\,dA \right| + \left| \iint\limits_{T_m} f(x,y)\,dA - J \right|. \end{aligned}$$

If we choose k and m sufficiently large, each of the four terms on the right will be bounded by ε, so $|I - J| \leq 4\varepsilon$. Because $\varepsilon > 0$ is arbitrary, $I = J$. □

Improper integral of an unbounded function

Definition 9.1 *Suppose R is a closed bounded set with area, Z a set with area zero, and S_k is a monotonic increasing collection of closed subsets of $R \setminus Z$ for which $A(S_k) \to A(R)$ as $k \to \infty$. Suppose $f(x,y)$ is continuous but unbounded on $R \setminus Z$, and the integrals*

$$\iint_{S_k} |f(x,y)|\,dA$$

are uniformly bounded in k. Then the ***improper integral of f over R*** *is*

$$\iint_R f(x,y)\,dA = \lim_{k \to \infty} \iint_{S_k} f(x,y)\,dA.$$

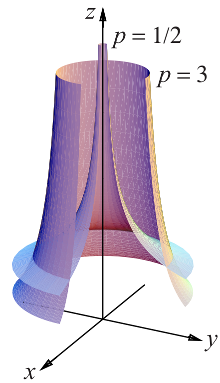
When the improper integral exists, we often say that it **converges**.

How "unbounded" can a function be and still have a convergent improper integral? For example,

$$f(x,y) = \frac{1}{r^p}, \quad r = \sqrt{x^2 + y^2}$$

is unbounded near the origin when $p > 0$; for which values of p does

$$\iint\limits_{x^2+y^2\leq 1} \frac{dA}{r^p}$$



converge? The figure suggests $1/r^p$ becomes unbounded more rapidly as p becomes larger; thus, the integral—thought of as the volume under the graph—is more likely to converge for smaller values of p. To answer the question, let S_k be the set of points (x,y) in the ring $1/k^2 \leq x^2+y^2 \leq 1$. Then, changing to polar coordinates, we find

$$I_k = \iint_{S_k} \frac{dA}{r^p} = \int_0^{2\pi}\int_{1/k}^{1} \frac{r\,dr}{r^p}\,d\theta = 2\pi \frac{r^{2-p}}{2-p}\bigg|_{1/k}^{1} = \frac{2\pi}{2-p}(1+k^{p-2}).$$

This formula does not allow $p = 2$, so we deal with that case separately.

First, the sequence $I_1, I_2, \ldots$ has a finite limit as $k \to \infty$ only if $k^{p-2} \to 0$, that is, only if $p-2 < 0$, or $p < 2$. If $p = 2$, then

$$I_k = \iint_{S_k} \frac{dA}{r^2} = \int_0^{2\pi}\int_{1/k}^{1} \frac{dr}{r}\,d\theta = 2\pi \ln r\bigg|_{1/k}^{1} = 2\pi \ln k \to \infty.$$

Thus the improper integral converges if and only if $0 < p < 2$, in which case it has the value

$$\iint\limits_{x^2+y^2\leq 1} \frac{dA}{r^p} = \frac{2\pi}{2-p}.$$

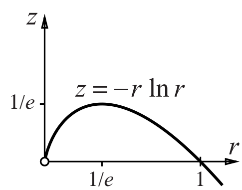
(Of course, this formula also gives the value of the "proper" integral that exists for all $p \leq 0$.) Our example has the following useful generalization.

Testing for absolute convergence

Because $|f(x,y)| = 1/r^p = f(x,y)$, the improper integral of f converges "absolutely" or not at all; we do not need to test separately whether the integrals of $|f(x,y)|$ are uniformly bounded, as stipulated in Theorem 9.6.

For another example, take $g(x,y) = \ln r$. It is also unbounded near the origin, but because $\ln r < 0$ when $0 < r < 1$, we should first consider $|g(x,y)| = -\ln r$. Thus, on the disks $S_k : 1/k^2 \leq x^2+y^2 \leq 1$ that we just used,

$$\widehat{I}_k = \iint\limits_{1/k^2\leq x^2+y^2\leq 1} |g(x,y)|\,dA = \int_0^{2\pi}\int_{1/k}^{1} -r\ln r\,dr\,d\theta.$$



The function $z = -r\ln r$ is continuous and bounded by $1/e$ on $0 < r \leq 1$, so we have

$$\int_{1/k}^{1} -r\ln r\,dr \leq \frac{1}{e}\left(1-\frac{1}{k}\right) \leq \frac{1}{e},$$

implying the uniform bound $\widehat{I}_k \leq 2\pi/e$ for all k. Therefore, by Theorem 9.6, the improper integral

$$\iint\limits_{x^2+y^2\leq 1} \ln\sqrt{x^2+y^2}\,dA$$

converges. (It converges to $-\pi/2$; see Exercise 9.22).

Unbounded regions

An integral will also be improper when its domain R is unbounded. As in the earlier case of an unbounded function, choose a monotonic increasing sequence of closed bounded subsets $S_1 \subseteq S_2 \subseteq \cdots$ of R. Because it no longer makes sense to require $A(S_k) \to A(R)$ (because the area of R may be infinite), we achieve what we really want—that the sets S_k eventually cover R—by stipulating instead that each closed bounded subset of R be contained in some S_k. As in the earlier case, we also require absolute convergence.

Theorem 9.7. *Let R be an unbounded set in $\mathbb{R}^2$, Z a set with area zero, and $S_1, S_2, \ldots$ a monotonic increasing sequence of closed bounded subsets of R such that, given any closed bounded subset W of R, $S_k \supseteq W$ for some k. Suppose $f(x,y)$ is bounded and continuous on $R \setminus Z$, and*

$$\widehat{I}_k = \iint_{S_k} |f(x,y)|\,dA \leq B$$

for some bound B and for all k. Then the numbers

$$I_k = \iint_{S_k} f(x,y)\,dA$$

have a finite limit I as $k \to \infty$. Furthermore, the value of I is independent of the way the closed subsets S_k are chosen.

Proof. This proof has many parallels with the previous one; we focus on the points where the two differ. To begin, because f and $|f|$ are bounded and continuous on $R \setminus Z$, the same is true on each $S_k \setminus Z$, so f and $|f|$ are integrable on each S_k. This is enough to establish, as in the earlier proof, that the sequence $I_1, I_2, \ldots$ has a finite limit.

The next step is to consider a second monotonic increasing sequence of closed bounded subsets $T_1, T_2, \ldots$ that exhaust R the same way the sequence $S_1, S_2, \ldots$ does. Each T_m is a closed bounded subset of R; thus it is entirely contained in some S_k, by hypothesis. Hence,

$$\widehat{J}_m = \iint_{T_m} |f(x,y)|\,dA \leq \iint_{S_k} |f(x,y)|\,dA \leq B.$$

As noted in the earlier proof, this implies that the integrals

$$J_m = \iint_{T_m} f(x,y)\,dA$$

converge to a finite limit, J.

To prove that $I = J$, the earlier proof first established that

$$\left| \iint\limits_{T_m} f(x,y)\,dA - \iint\limits_{T_m \cap S_k} f(x,y)\,dA \right| \leq \varepsilon$$

for all $k > N$, and uniformly for all T_m. There, the key step was that

$$\iint\limits_{T_m} f(x,y)\,dA = \lim_{i\to\infty} \iint\limits_{T_m \cap S_i} f(x,y)\,dA;$$

here, this holds for the simple reason that $T_m \cap S_i = T_m$ for all i sufficiently large. Reversing the roles of S_k and T_m we likewise conclude that

$$\left| \iint\limits_{S_k} f(x,y)\,dA - \iint\limits_{S_k \cap T_m} f(x,y)\,dA \right| \leq \varepsilon$$

for every $m > L$ and for all S_k. Then $|I - J| \leq 4\varepsilon$ as before. □

Improper integral over an unbounded domain

Definition 9.2 *When the conditions of the previous theorem are met, then the* ***improper integral of f over R converges*** *to*

$$\iint_R f(x,y)\,dA = \lim_{k\to\infty} \iint_{S_k} f(x,y)\,dA.$$

The normal density function of statistics

In Chapter 1, we met one of the standard examples of an improper integral over an unbounded interval:

$$\iint_{\mathbb{R}^2} e^{(-x^2-y^2)/2}\,dA = 2\pi.$$

We used this to evaluate another improper integral,

$$\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\,dx = \sigma\sqrt{2\pi},$$

that relates to the density function of the normal probability distribution of statistics.

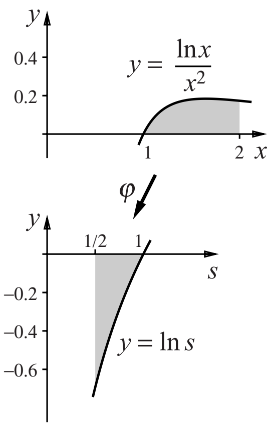
9.3 The change of variables formula

The formula for single integrals

This book begins with the change of variables formula for single integrals. It says that when there is an invertible pullback function $x = \varphi(s)$, then

$$\int_a^b f(x)\,dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(s))\,\varphi'(s)\,ds.$$

Our goal here is to state and prove the analogous formula for double integrals. As we show in a moment, the single integral takes orientation into account; however, we have not yet defined double integrals with orientation. At this stage, therefore, we must suppress the information about orientation in the single-integral formula to carry through an analogy. Here is an example that illustrates both the problem and the solution.



Let $f(x) = \ln x/x^2$; we can see by eye that the value of the integral

$$\int_1^2 \frac{\ln x}{x^2}\,dx$$

is positive but less than 0.2. (The vertical scales of the graphs shown in the margin have been doubled for clarity.) To find the value using the change of variables formula, consider the pullback $x = \varphi(s) = 1/s$. Then $\varphi'(s) = -1/s^2$, and

$$\int_1^2 \frac{\ln x}{x^2}\,dx = \int_1^{1/2} s^2 \ln\frac{1}{s}\left(\frac{-ds}{s^2}\right) = \int_1^{1/2} -\ln\frac{1}{s}\,ds = \int_1^{1/2} \ln s\,ds.$$

Note that the new integrand, $\ln s$, is *negative* on the new interval $[1/2, 1]$ More significantly, the new integration is carried out in the *negative* sense, from 1 backwards to 1/2. These two "negative" aspects of the new integral combine to produce a positive value,

$$\int_1^{1/2} \ln s\,ds = s\ln s - s\Big|_1^{1/2} = \tfrac{1}{2}\ln\tfrac{1}{2} - \tfrac{1}{2} - (\ln 1 - 1) = \tfrac{1}{2}(1 - \ln 2) \approx 0.15.$$

The effect of reversing orientation

What we see in this example always happens when φ is *orientation-reversing*: $\varphi'(s) < 0$ changes the sign of the integrand, and the oriented interval $a \to b$ is transformed into the oriented interval $\varphi^{-1}(a) \leftarrow \varphi^{-1}(b)$; that is, $a < b$ implies $\varphi^{-1}(a) > \varphi^{-1}(b)$.

An orientation-free change of variables formula

We therefore need to reformulate the way we write a single integral so as to suppress this information about orientation. The unoriented version of the change of variables formula has the following form,

$$\int_I f(x)\,dx = \int_{\varphi^{-1}(I)} f(\varphi(s))\,|\varphi'(s)|\,ds.$$

In this formula, I stands for the *unoriented* set of real numbers x that lie between a and b, inclusive, and $\varphi^{-1}(I)$ is the unoriented set of real numbers s for which $\varphi(s)$ lies in I. The integral over the *unoriented* domain I is defined in terms of the usual integral, as follows.

$$\int_I f(x)\,dx = \begin{cases} \displaystyle\int_a^b f(x)\,dx & \text{if } I = [a,b], \\[2ex] \displaystyle\int_b^a f(x)\,dx & \text{if } I = [b,a]. \end{cases}$$

Before we prove that the new, unoriented, change of variables formula is the correct modification of the original one, let us verify that it works on the example $f(x) = \ln x/x^2$ with $x = \varphi s = 1/s$. Because $\varphi([1,2]) = [1/2,1]$ and $|\varphi'(s)| = +1/s^2$,

$$\int\limits_{[1,2]} \frac{\ln x}{x^2}\,dx = \int\limits_{[1/2,1]} s^2 \ln\frac{1}{s}\left(\frac{ds}{s^2}\right) = \int_{1/2}^{1} -\ln s\,ds = s - s\ln s\Big|_{1/2}^{1} = \tfrac{1}{2}(1-\ln 2).$$

To prove that the oriented change of variables formula leads to the unoriented one, let us assume $a < b$; we can use a similar argument if $b < a$. If φ is orientation-preserving, then $|\varphi'(s)| = \varphi'(s)$ and $\varphi^{-1}(a) < \varphi^{-1}(b)$, so

$$\int\limits_{\varphi^{-1}(I)} f(\varphi(s))\,|\varphi'(s)|\,ds = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(s))\,\varphi'(s)\,ds = \int_a^b f(x)\,dx = \int\limits_I f(x)\,dx.$$

If φ is orientation-reversing, then $|\varphi'(s)| = -\varphi'(s)$ and $\varphi^{-1}(b) < \varphi^{-1}(a)$, so

$$\begin{aligned}\int\limits_{\varphi^{-1}(I)} f(\varphi(s))\,|\varphi'(s)|\,ds &= \int_{\varphi^{-1}(b)}^{\varphi^{-1}(a)} f(\varphi(s))\,\big(-\varphi'(s)\,ds\big)\\ &= \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(s))\,\varphi'(s)\,ds = \int_a^b f(x)\,dx = \int\limits_I f(x)\,dx.\end{aligned}$$

The new formula holds in both cases.

We can now formulate an analogous change of variables formula for double integrals. Let $f(x,y)$ be a continuous function on a domain D in $\mathbb{R}^2$, and assume that D is a closed bounded set that has area. Then the integral

Change of variables for double integrals

$$\iint_D f(x,y)\,dx\,dy$$

exists (Theorem 8.35, p. 305). We use "$dx\,dy$" here in place of "dA" as a way to keep track of the variables that appear in different integrals. If the change of variables is given by the pullback substitution

$$\boldsymbol{\varphi}: \begin{cases} x = x(s,t), \\ y = y(s,t), \end{cases}$$

then we show that the integral is transformed by

$$\iint\limits_D f(x,y)\,dx\,dy = \iint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(\boldsymbol{\varphi}(s,t))\,\big|\det \mathrm{d}\boldsymbol{\varphi}_{(s,t)}\big|\,ds\,dt.$$

In particular, $|\det \mathrm{d}\boldsymbol{\varphi}_{(s,t)}|$ corresponds to $|\varphi'(s)|$; this is the absolute value of the *Jacobian*,

$$J_{\boldsymbol{\varphi}}(s,t) = \det \mathrm{d}\boldsymbol{\varphi}_{(s,t)} = \begin{vmatrix} \dfrac{\partial x}{\partial s}(s,t) & \dfrac{\partial x}{\partial t}(s,t) \\ \dfrac{\partial y}{\partial s}(s,t) & \dfrac{\partial y}{\partial t}(s,t) \end{vmatrix} = \frac{\partial(x,y)}{\partial(s,t)}.$$

Consequently,

$$\iint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(\boldsymbol{\varphi}(s,t))\left|\det \mathrm{d}\boldsymbol{\varphi}_{(s,t)}\right| ds\,dt = \iint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(x(s,t),y(s,t))\left|\frac{\partial(x,y)}{\partial(s,t)}\right| ds\,dt$$

gives us an alternate expression for the transformed integral that is useful in the following work. Moreover, the notation suggests that, when the variables themselves change, then

$$dx\,dy \ \text{ changes to } \ \left|\frac{\partial(x,y)}{\partial(s,t)}\right| ds\,dt.$$

For the moment, this is just a mnemonic. Before proving the theorem that establishes the change of variables formula for double integrals, let us first explore some examples.

Most familiar is the change to polar coordinates that we have already used several times:

$$\boldsymbol{\varphi} : \begin{cases} x = r\cos\theta, \\ y = r\sin\theta; \end{cases} \qquad \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r \geq 0;$$

$$\iint\limits_{D} f(x,y)\,dx\,dy = \iint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta.$$

Changing variables to simplify domains

With single integrals, we change variables to simplify the integrand. With double integrals, there is a second reason: to simplify the domain. One example is an integral with circular symmetry; it is often recast into polar coordinates. A second example is the integral

$$\iint_D x^2 + y^2 \, dx\,dy,$$

where the domain D is the curvilinear quadrilateral in the first quadrant whose points (x,y) satisfy the inequalities

$$D: \begin{array}{c} 1 \leq x^2 - y^2 \leq 2, \\ 1 \leq 2xy \leq 3. \end{array}$$

The sides of D are hyperbolic arcs. The quadratic map

$$\mathbf{g} : \begin{cases} u = x^2 - y^2, \\ v = 2xy, \end{cases}$$

that we analyzed on pages 116–121 straightens these arcs. For example, the hyperbola $x^2 - y^2 = 1$ in the first quadrant of the (x,y)-plane becomes the line $u = 1$ in the first quadrant of the (u,v)-plane. The quadrilateral D becomes the rectangle

$$R = \mathbf{g}(D): \begin{array}{l} 1 \le u \le 2, \\ 1 \le v \le 3. \end{array}$$

Unfortunately, $\mathbf{g}$ is a *push-forward* map, not a *pullback*, so the change of variables formula does not apply directly. But $\mathbf{g}$ is invertible on the first quadrant, and $\mathbf{g}^{-1}$ does indeed pull back (x,y) to (u,v), so we let $\mathbf{g}^{-1}$ play the role of $\boldsymbol{\varphi}$ and write

$$\iint_D f(x,y)\,dx\,dy = \iint_{\mathbf{g}(D)} f(\mathbf{g}^{-1}(u,v)) \left|\det d\mathbf{g}^{-1}_{(u,v)}\right| du\,dv.$$

To evaluate the right-hand side we can use formulas for the components of $\mathbf{g}^{-1}(u,v)$ that appear in Exercise 4.13 (p. 144). However, we can actually determine everything we need without recourse to those formulas. To begin,

$$u^2 = x^4 - 2x^2y^2 + y^4 \text{ and } v^2 = 4x^2y^2,$$

so $u^2 + v^2 = x^4 + 2x^2y^2 + y^4 = (x^2 + y^2)^2$ and

$$f(x,y) = x^2 + y^2 = \sqrt{u^2 + v^2} = f(\mathbf{g}^{-1}(u,v)).$$

Next,

$$\det d\mathbf{g}^{-1} = \frac{1}{\det d\mathbf{g}} = \frac{1}{4(x^2 + y^2)} = \frac{1}{4\sqrt{u^2 + v^2}} > 0,$$

so the integrand of the transformed integral is just

$$f(\mathbf{g}^{-1}(u,v)) \left|\det d\mathbf{g}^{-1}_{(u,v)}\right| = \sqrt{u^2 + v^2}\,\frac{1}{4\sqrt{u^2 + v^2}} = \frac{1}{4}.$$

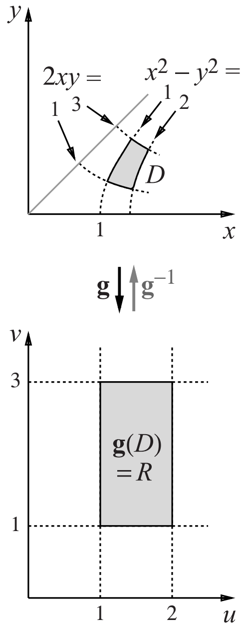
Therefore, the change of variables formula gives

$$\iint_D x^2 + y^2\,dx\,dy = \iint_R \frac{1}{4}\,du\,dv = \frac{1}{4}\,\text{area}\,R = \frac{1}{2}.$$

(Of course we could have tried to evaluate the original double integral by reducing it to iterated integrals, but they lack the simplicity and elegance of the transformed double integral.)

We can now formulate the change of variables formula for double integrals under a general push-forward substitution $\mathbf{g}$ *that has an inverse*:

$$\mathbf{g}: \begin{cases} u = u(x,y), \\ v = v(x,y); \end{cases} \qquad \mathbf{g}^{-1}: \begin{cases} x = x(u,v), \\ y = y(u,v). \end{cases}$$



Change of variables with a push-forward

If we write the Jacobian as

$$J_{\mathbf{g}^{-1}}(u,v) = \det \mathrm{d}\mathbf{g}^{-1} = \frac{\partial(x,y)}{\partial(u,v)},$$

then the change of variables formula takes the form

$$\iint\limits_{D} f(x,y)\,dx\,dy = \iint\limits_{\mathbf{g}(D)} f(x(u,y),y(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv.$$

Note that we have expressed the transformed integral in terms of the component functions $x(u,v)$ and $y(u,v)$ of $\mathbf{g}^{-1}$. But usually only the components of $\mathbf{g}$ itself are given; it may be difficult or impossible to find closed-form expressions for $x(u,v)$ and $y(u,v)$. This can make it impractical to transform an integral by a push-forward substitution.

Areas via the change of variables formula

The change of variables formula for double integrals also gives us a way to determine areas. To continue the last example, we have

$$\operatorname{area} D = \iint_D 1\;dx\,dy = \iint_R \frac{1}{4\sqrt{u^2+v^2}}\;du\,dv.$$

One way to continue is to convert the double integral into an iterated integral:

$$\iint_R \frac{1}{4\sqrt{u^2+v^2}}\;du\,dv = \frac{1}{4}\int_1^3 \int_1^2 \frac{du}{\sqrt{u^2+v^2}}\;dv.$$

This can be evaluated using a computer algebra system (or a table of integrals). We can also use the pullback substitution $u = v\cdot\sinh(s)$ (see Exercise 1.16, p. 23) to get

$$\int_1^2 \frac{du}{\sqrt{u^2+v^2}} = \operatorname{arcsinh}\frac{u}{v}\bigg|_1^2 = \operatorname{arcsinh}\frac{2}{v} - \operatorname{arcsinh}\frac{1}{v}.$$

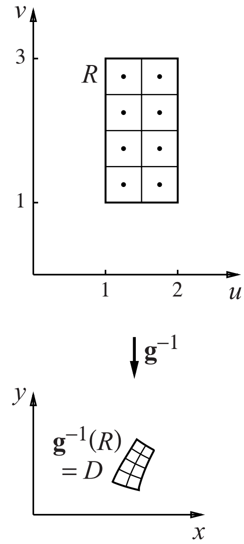
Integration by parts (when $v > 0$) gives

$$\begin{aligned}\int 1\cdot\operatorname{arcsinh}\frac{a}{v}\,dv &= v\cdot\operatorname{arcsinh}\frac{a}{v} - \int\frac{-a}{\sqrt{v^2+a^2}}\,dv\\ &= v\cdot\operatorname{arcsinh}\frac{a}{v} + a\cdot\operatorname{arcsinh}\frac{v}{a},\end{aligned}$$

so we get

$$\begin{aligned}&\int_1^3 \left(\operatorname{arcsinh}\frac{2}{v} - \operatorname{arcsinh}\frac{1}{v}\right) dv\\ &\quad = v\cdot\operatorname{arcsinh}\frac{2}{v} + 2\cdot\operatorname{arcsinh}\frac{v}{2} - v\cdot\operatorname{arcsinh}\frac{1}{v} - \operatorname{arcsinh} v\bigg|_1^3 \approx 0.820\,853.\end{aligned}$$

Finally, incorporating the factor $1/4$, we find $\operatorname{area} D \approx 0.205\,213$.



We can also get an approximation to area$D$ by approximating the value of the double integral by a Riemann sum. The integrand $J(u,v) = 1/4\sqrt{u^2+v^2}$ is, of course, the local area magnification factor for the map $\mathbf{g}^{-1} : R \to D$. Therefore, we can estimate the area of D as follows. Divide R into small rectangular cells Q of area $\Delta u\,\Delta v$; multiply that area by the value of J at the center of C to approximate the area of the image $\mathbf{g}^{-1}(Q)$ in D; add the results. The table below does this with R partitioned into eight squares of area 1/4. The accumulated sum of $J\,\Delta u\,\Delta v$ is tallied in the right column; it yields the estimate 0.204 806 for the area of D (cf. Exercise 8.2, p. 313).

u	v	$J(u,v)$	Sum
1.25	1.25	0.141 421	0.035 355
1.25	1.75	0.116 248	0.064 417
1.25	2.25	0.097 129	0.086 994
1.25	2.75	0.082 761	0.109 390
1.75	1.25	0.116 248	0.138 451
1.75	1.75	0.101 102	0.163 705
1.75	2.25	0.087 706	0.185 632
1.75	2.75	0.076 696	0.204 806

Area magnification for nonlinear maps

For a *linear* map L, $|\det L|$ is the magnification factor for areas (by Theorem 8.22, p. 292):

$$A(L(S)) = |\det L|\,A(S),$$

when S is any subset of the plane that has area. (Here *area* is nonnegative; in the following section we consider oriented regions that are assigned negative area.) For a *nonlinear* map $\boldsymbol{\varphi}(s,t)$, the connection between $A(\boldsymbol{\varphi}(S))$ and $A(S)$ is not so simple, but the change of variables formula still allows us to write

$$A(\boldsymbol{\varphi}(S)) = \iint_S |\det d\boldsymbol{\varphi}_{(s,t)}|\,ds\,dt = \iint_S |J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt.$$

Using the language of set functions (see below, p. 352), we show how this equation makes $|J_{\boldsymbol{\varphi}}(s,t)|$ the local area magnification factor for $\boldsymbol{\varphi}$. Here, it is the crucial "base case" of the change of variables formula for double integrals. We state it now as a theorem.

Jacobian as local area magnification factor

Theorem 9.8. *Let Ω be a bounded open set in $\mathbb{R}^2$, and let $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^2$ be a continuously differentiable map that has a continuously differentiable inverse $\boldsymbol{\varphi}^{-1} : \boldsymbol{\varphi}(\Omega) \to \Omega$. Suppose the set S has area and its closure $\overline{S} = S \cup \partial S$ lies within Ω; then $\boldsymbol{\varphi}(S)$ has area and*

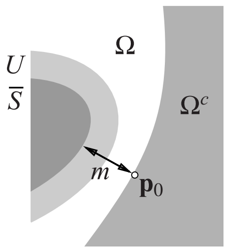
$$A(\boldsymbol{\varphi}(S)) = \iint_S |J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt,$$

where $J_{\boldsymbol{\varphi}}(s,t)$ is the Jacobian of $\boldsymbol{\varphi}$ at (s,t).

Proof. The proof is simple in principle; it follows an argument given by J. Schwartz [16]. First, partition S into small pieces Q_i. On Q_i, choose a representative value for the area magnification factor $|J_{\boldsymbol{\varphi}}| = |\det \mathrm{d}\boldsymbol{\varphi}|$. Then the area of the image $\boldsymbol{\varphi}(Q_i)$ is approximately $|J_{\boldsymbol{\varphi}}|A(Q_i)$, and the sum of such terms approximates the area of $\boldsymbol{\varphi}(S)$.

However, the details of the proof are numerous and lengthy; they involve several steps that we write as separate lemmas. We first need an open set U containing $\overline{S}$ on whose closure the functions $\boldsymbol{\varphi}$, $\mathrm{d}\boldsymbol{\varphi}$, and $J_{\boldsymbol{\varphi}}$ are uniformly continuous.

Lemma 9.1. *There is an open set U for which $\overline{S} \subset U \subset \overline{U} \subset \Omega$.*



Proof. Let $\mathbf{p}$ be a point in $\mathbb{R}^2$; define the function

$$d(\mathbf{p}) = \min_{\mathbf{s}\in\overline{S}} \|\mathbf{p}-\mathbf{s}\|$$

that gives the distance from $\mathbf{p}$ to the closed set $\overline{S}$. Then $d(\mathbf{p}) = 0$ if and only if $\mathbf{p}$ is in $\overline{S}$; in particular, $d(\Omega^c) > 0$ because $\overline{S} \subset \Omega$. But Ω^c is closed and d is continuous, so $(\mathbf{p})$ attains its minimum $m > 0$ at some point $\mathbf{p}_0$ in Ω^c: $d(\Omega^c) \geq d(\mathbf{p}_0) = m > 0$. To complete the proof of the lemma, we can take

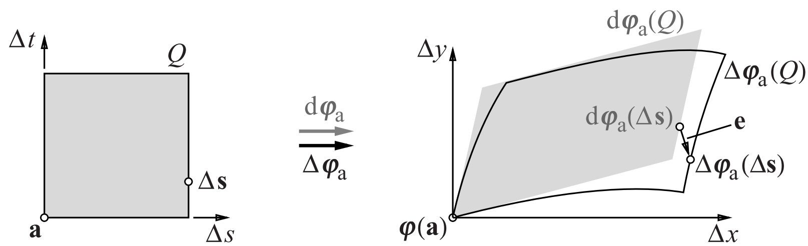
$$U = \{\mathbf{p} : d(\mathbf{p}) < m/2\}, \quad \overline{U} = \{\mathbf{p} : d(\mathbf{p}) \leq m/2\}. \qquad \square$$

The next lemma makes the first step in connecting $A(\boldsymbol{\varphi}(Q))$ to the integral of $|J_{\boldsymbol{\varphi}}|$ over Q, where Q is a square in one of the original grids $\mathcal{J}_k$ used to define Jordan content. It says that the outer area of the image $\boldsymbol{\varphi}(Q)$ is bounded by the maximum value of $|J_{\boldsymbol{\varphi}}|$ on Q.

Lemma 9.2. *For any given $\varepsilon > 0$, there is a positive integer K such that if $Q \subseteq \overline{U}$ is a square of $\mathcal{J}_k$ and $k \geq K$, then*

$$\overline{A}(\boldsymbol{\varphi}(Q)) \leq (M + O(\varepsilon))A(Q),$$

where M is the maximum value of $|J_{\boldsymbol{\varphi}}|$ on Q.



Proof. The idea of the proof is to compare the action of $\boldsymbol{\varphi}$ to the action of its linear approximation $\mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}$ taken at the lower-left hand corner $\mathbf{a}$ of Q. In terms of local (or "window") coordinates

$$\Delta\mathbf{s} = \mathbf{s} - \mathbf{a} \;\text{ and }\; \Delta\mathbf{x} = \mathbf{x} - \boldsymbol{\varphi}(\mathbf{a})$$

centered at this corner and its image, the two maps are

$$\Delta\mathbf{x} = \Delta\boldsymbol{\varphi}_{\mathbf{a}}(\Delta\mathbf{s}) = \boldsymbol{\varphi}(\mathbf{a}+\Delta\mathbf{s}) - \boldsymbol{\varphi}(\mathbf{a}) \ \text{ and } \ \Delta\mathbf{x} = \mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}(\Delta\mathbf{s}).$$

Because $\mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}$ is linear, the image $\mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}(Q)$ is a parallelogram, and

$$A(\mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}(Q)) = |\det \mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}|A(Q) = |J_{\boldsymbol{\varphi}}(\mathbf{a})|A(Q) \le M \times A(Q).$$

We now use the continuity of $\boldsymbol{\varphi}$ and $\mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}$ to show that, when Q is sufficiently small, the two images $\mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}(Q)$ and $\Delta\boldsymbol{\varphi}_{\mathbf{a}}(Q)$ are close enough so that our bound on the area of the first leads to a (slightly larger) bound on the outer area of the second.

For any point $\Delta\mathbf{s}$, we want to determine the difference

$$\mathbf{e} = \Delta\boldsymbol{\varphi}_{\mathbf{a}}(\Delta\mathbf{s}) - \mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}(\Delta\mathbf{s}).$$

It is convenient to work with the component functions of $\Delta\boldsymbol{\varphi}_{\mathbf{a}}$ and $\mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}$:

$$\Delta\boldsymbol{\varphi}_{\mathbf{a}} : \begin{cases} \Delta x = x(a+\Delta s, b+\Delta t) - x(a,b), \\ \Delta y = y(a+\Delta s, b+\Delta t) - y(a,b), \end{cases} \qquad \mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}} : \begin{cases} \Delta x = x_s(a,b)\,\Delta s + x_t(a,b)\,\Delta t, \\ \Delta y = y_s(a,b)\,\Delta s + y_t(a,b)\,\Delta t. \end{cases}$$

Applying the law of the mean (Theorem 3.5, p. 75) to each component of $\Delta\boldsymbol{\varphi}_{\mathbf{a}}$, we get

$$\Delta\boldsymbol{\varphi}_{\mathbf{a}} : \begin{cases} \Delta x = x_s(a_1,b_1)\,\Delta s + x_t(a_1,b_1)\,\Delta t, \\ \Delta y = y_s(a_2,b_2)\,\Delta s + y_t(a_2,b_2)\,\Delta t, \end{cases}$$

where (a_1,b_1) and (a_2,b_2) are two properly chosen points on the line connecting (a,b) and $(a+\Delta s, b+\Delta t)$. This allows us to write $\mathbf{e} = \Delta\boldsymbol{\varphi}_{\mathbf{a}}(\Delta\mathbf{s}) - \mathrm{d}\boldsymbol{\varphi}_{\mathbf{a}}(\Delta\mathbf{s})$ as

$$\mathbf{e} : \begin{cases} \Delta x = \big(x_s(a_1,b_1) - x_s(a,b)\big)\,\Delta s + \big(x_t(a_1,b_1) - x_t(a,b)\big)\,\Delta t, \\ \Delta y = \big(y_s(a_2,b_2) - y_s(a,b)\big)\,\Delta s + \big(y_t(a_2,b_2) - y_t(a,b)\big)\,\Delta t. \end{cases}$$

To get a bound on $\mathbf{e}$, we use the continuity of $\mathrm{d}\boldsymbol{\varphi}_{\mathbf{s}}$ as a function of the point $\mathbf{s}$. On the closed bounded set $\overline{U}$ (Lemma 9.1), the four components x_s, x_t, y_s, y_t of $\mathrm{d}\boldsymbol{\varphi}_{\mathbf{s}}$ are uniformly continuous. Thus, for $\varepsilon > 0$ as given in the statement of the lemma, there is a $\delta > 0$ such that

$$\|(s_1 - s_2, t_1 - t_2)\| < \delta \quad \Longrightarrow \quad \|x_s(s_1,t_1) - x_s(s_2,t_2)\| < \varepsilon,$$

and likewise for x_t, y_s, and y_t. Now choose K so that the mesh size $\|\mathcal{I}_K\| = \sqrt{2}/2^K$ is less than δ (Definition 8.14, p. 291). Then, for any $k \ge K$, we have

$$\|\mathcal{I}_k\| = \frac{\sqrt{2}}{2^k} \le \frac{\sqrt{2}}{2^K} < \delta.$$

Now let $Q \subseteq \overline{U}$ be a square of $\mathcal{I}_k$, $k \ge K$; then

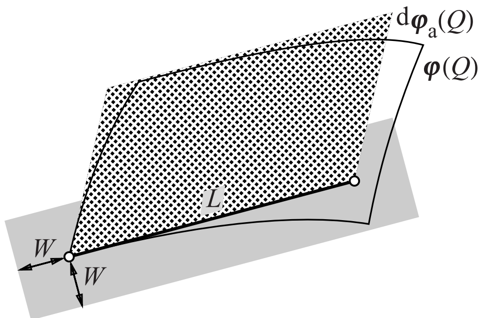
$$\|(a_1 - a, b_1 - b)\| < \delta \ \text{ and } \ \|(a_2 - a, b_2 - b)\| < \delta,$$

and $0 \le \Delta s, \Delta t \le 1/2^k$. Therefore,

$$\|\mathbf{e}\| : \begin{cases} |\Delta x| \le \left|x_s(a_1,b_1) - x_s(a,b)\right| \Delta s + \left|x_t(a_1,b_1) - x_t(a,b)\right| \Delta t < \dfrac{2\varepsilon}{2^k}, \\ |\Delta y| \le \left|y_s(a_2,b_2) - y_s(a,b)\right| \Delta s + \left|y_t(a_2,b_2) - y_t(a,b)\right| \Delta t < \dfrac{2\varepsilon}{2^k}, \end{cases}$$

so $\|\mathbf{e}\| < 2\varepsilon\sqrt{2}/2^k = W$.

Thus, every point of $\Delta\boldsymbol{\varphi}_{\mathbf{a}}(Q) = \boldsymbol{\varphi}(Q)$ lies either in the parallelogram $d\boldsymbol{\varphi}_{\mathbf{a}}(Q)$ or within the distance W from one of the four lines that make up its boundary. So $\boldsymbol{\varphi}(Q)$ is contained in the union of the parallelogram and four rectangles of length $L+2W$ and width $2W$, where L is the length of the longer side of the parallelogram. A bound on the length L will thus lead to a bound on the outer area of $\boldsymbol{\varphi}(Q)$.



To get a bound on L it is enough to consider the two sides of $d\boldsymbol{\varphi}_{\mathbf{a}}(Q)$ that meet at the corner $\boldsymbol{\varphi}(a,b)$. These are the vectors

$$d\boldsymbol{\varphi}_{\mathbf{a}}\begin{pmatrix} 1/2^k \\ 0 \end{pmatrix} = \frac{1}{2^k}\begin{pmatrix} x_s(a,b) \\ x_t(a,b) \end{pmatrix} \text{ and } d\boldsymbol{\varphi}_{\mathbf{a}}\begin{pmatrix} 0 \\ 1/2^k \end{pmatrix} = \frac{1}{2^k}\begin{pmatrix} y_s(a,b) \\ y_t(a,b) \end{pmatrix}.$$

Because $d\boldsymbol{\varphi}_{\mathbf{s}}$ is a continuous function of $\mathbf{s}$ and Q lies in the closed bounded set $\overline{U}$, the four component functions of $d\boldsymbol{\varphi}_{\mathbf{s}}$ are uniformly bounded on $\overline{U}$: for some $B > 0$,

$$|x_s(s,t)| \le B, \quad |x_t(s,t)| \le B, \quad |y_s(s,t)| \le B, \quad |y_t(s,t)| \le B,$$

for all (s,t) in $\overline{U}$. This implies $L \le B\sqrt{2}/2^k$; thus, for each of the four rectangles,

$$\text{area} = 2W\,(L+2W) \le \frac{4\varepsilon\sqrt{2}}{2^k}\left(\frac{B\sqrt{2}}{2^k} + \frac{4\varepsilon\sqrt{2}}{2^k}\right) = \frac{8\varepsilon(B+4\varepsilon)}{2^{2k}}.$$

Because $\varepsilon(B+4\varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$, we have

$$\overline{A}(\boldsymbol{\varphi}(Q)) \le A(d\boldsymbol{\varphi}_{\mathbf{a}}(Q)) + 4\frac{8\varepsilon(B+4\varepsilon)}{2^{2k}} \le M\,A(Q) + O(\varepsilon)\,A(Q). \qquad \square$$

Lemma 9.3. *If S has area and $\overline{S} \subset \Omega$, then*

$$\overline{A}(\boldsymbol{\varphi}(S)) \leq \iint_S |J_{\boldsymbol{\varphi}}(s,t)|\, ds\, dt.$$

Proof. To prove this lemma, we construct upper Darboux sums for $|J_{\boldsymbol{\varphi}}(s,t)|$ on S (cf. Definition 8.17, p. 300).

By the proof of Lemma 9.1, $\overline{S}$ is contained in the interior of a bounded closed set $\overline{U} \subset \Omega$ whose points lie within distance $m/2$ of $\overline{S}$. Fix $\varepsilon > 0$ and choose K to satisfy both the previous lemma and the condition that the mesh size $\|\mathcal{J}_K\|$ is less than $m/2$. Then, if $k \geq K$, any square Q of $\mathcal{J}_k$ that meets S will lie within $\overline{U}$. Let $Q_1, \ldots, Q_I$ be the squares of $\mathcal{J}_k$ that meet S. Set

$$M_i = \max_{(s,t) \in Q_i} |J_{\boldsymbol{\varphi}}(s,t)|, \quad i = 1, \ldots, I;$$

then, by the previous lemma,

$$\overline{A}(\boldsymbol{\varphi}(Q_i)) \leq (M_i + O(\varepsilon))A(Q_i), \quad i = 1, \ldots, I.$$

Because $S \subseteq Q_1 \cup \cdots \cup Q_I$ and thus $\boldsymbol{\varphi}(S) \subseteq \boldsymbol{\varphi}(Q_1) \cup \cdots \cup \boldsymbol{\varphi}(Q_I)$, we have

$$\begin{aligned} \overline{A}(\boldsymbol{\varphi}(S)) &\leq \sum_{i=1}^{I} \overline{A}(\boldsymbol{\varphi}(Q_i)) \leq \sum_{i=1}^{I} (M_i + O(\varepsilon))A(Q_i) \\ &= \sum_{i=1}^{I} M_i A(Q_i) + \sum_{i=1}^{I} O(\varepsilon) A(Q_i). \end{aligned}$$

The first term is the upper Darboux sum for $|J_{\boldsymbol{\varphi}}(s,t)|$ over S and the grid $\mathcal{J}_k$; the second is $O(\varepsilon)$ times the outer area of S over the same grid; that is,

$$\overline{A}(\boldsymbol{\varphi}(S)) \leq \overline{D}_{\mathcal{J}_k}(|J_{\boldsymbol{\varphi}}(s,t)|, S) + O(\varepsilon)\overline{A}_k(S).$$

Because $|J_{\boldsymbol{\varphi}}(s,t)|$ is integrable over S, the upper Darboux sums converge to the integral (and the outer areas to the area) as $k \to \infty$; thus

$$\overline{A}(\boldsymbol{\varphi}(S)) \leq \iint_S |J_{\boldsymbol{\varphi}}(s,t)|\, ds\, dt + O(\varepsilon)A(S).$$

This inequality holds for any $\varepsilon > 0$; therefore it continues to hold as $\varepsilon \to 0$ (and hence as $O(\varepsilon) \to 0$), so

$$\overline{A}(\boldsymbol{\varphi}(S)) \leq \iint_S |J_{\boldsymbol{\varphi}}(s,t)|\, ds\, dt. \qquad \square$$

Part of the assertion of the main theorem is that $\boldsymbol{\varphi}(S)$ has area (implying we are able to replace $\overline{A}(\boldsymbol{\varphi}(S))$ by $A(\boldsymbol{\varphi}(S))$ in the lemma just proven). The next two lemmas establish that $\boldsymbol{\varphi}(S)$ does indeed have area by showing that its boundary $\partial(\boldsymbol{\varphi}(S))$ has outer area equal to zero.

Lemma 9.4. $\partial(\boldsymbol{\varphi}(S)) \subseteq \boldsymbol{\varphi}(\partial S)$.

Proof. Let $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{s})$ be a boundary point of $\boldsymbol{\varphi}(S)$; we must show that $\mathbf{s}$ is a boundary point of S. We use the criterion established in Exercise 8.5 (p. 313): every open disk centered at a boundary point of a set T contains at least one point in T and one point not in T.

Let $D_1 \subset \Omega$ be an open disk centered at $\mathbf{s}$. By a corollary to the inverse function theorem (Corollary 5.3, p. 174), $\mathbf{x}$ is an interior point of $\boldsymbol{\varphi}(D_1)$, so there is an open disk D_2 centered at $\mathbf{x}$ for which $D_2 \subseteq \boldsymbol{\varphi}(D_1)$. But $\mathbf{x}$ is boundary point of $\boldsymbol{\varphi}(S)$, so D_2 contains a point $\mathbf{p}_2$ in $\boldsymbol{\varphi}(S)$ and another point $\mathbf{q}_2$ that is not in $\boldsymbol{\varphi}(S)$. But then the point $\mathbf{p}_1 = \boldsymbol{\varphi}^{-1}(\mathbf{p}_2)$ in D_1 is in S and the point $\mathbf{q}_1 = \boldsymbol{\varphi}^{-1}(\mathbf{q}_2)$ in D_1 is not in S. (Draw a picture.) □

Lemma 9.5. $\overline{A}(\partial(\boldsymbol{\varphi}(S))) = 0$.

Proof. Apply Lemma 9.3 to the zero-area set ∂S. If $|J_{\boldsymbol{\varphi}}(s,t)| \le B$ on ∂S; then

$$\overline{A}(\partial(\boldsymbol{\varphi}(S))) \le \overline{A}(\boldsymbol{\varphi}(\partial S)) \le \iint_{\partial S} |J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt \le B\,A(\partial S) = 0. \qquad \square$$

Corollary 9.9 $A(\boldsymbol{\varphi}(S)) \le \displaystyle\iint_S |J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt.$ □

Lemma 9.6. $A(\boldsymbol{\varphi}(S)) \ge \displaystyle\iint_S |J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt.$

Proof. The idea of the proof is to apply the previous arguments to the inverse map $\boldsymbol{\varphi}^{-1}$ and a set $T = \boldsymbol{\varphi}(R)$, where R has area and $R \cup \partial R \subset \Omega$. By Corollary 9.9, T has area. Furthermore,

$$\overline{T} = T \cup \partial T \subseteq \boldsymbol{\varphi}(R \cup \partial R) \subset \boldsymbol{\varphi}(\Omega),$$

so we can write

$$A(R) = A(\boldsymbol{\varphi}^{-1}(T)) \le \iint\limits_{\boldsymbol{\varphi}(R)} \left|J_{\boldsymbol{\varphi}^{-1}}(x,y)\right| dx\,dy.$$

Now let R be a square Q of the grid $\mathcal{J}_k$, and let μ be the minimum value of $|J_{\boldsymbol{\varphi}}(s,t)|$ on Q. Note that $\mu > 0$, because $d\boldsymbol{\varphi}_{\mathbf{s}}$ is invertible and a uniformly continuous function of $\mathbf{s}$ on Q. Using Corollary 4.13, page 138, we find

$$\left|J_{\boldsymbol{\varphi}^{-1}}(x,y)\right| = \frac{1}{|J_{\boldsymbol{\varphi}}(s(x,y),t(x,y))|} \le \frac{1}{\mu}$$

for all (x,y) in $\boldsymbol{\varphi}(Q)$. Therefore,

$$A(Q) \le \iint\limits_{\boldsymbol{\varphi}(Q)} \left|J_{\boldsymbol{\varphi}^{-1}}(x,y)\right| dx\,dy \le \frac{A(\boldsymbol{\varphi}(Q))}{\mu}.$$

or $A(\boldsymbol{\varphi}(Q)) \ge \mu\,A(Q)$.

Let $P_1,\ldots,P_J$ be the squares of the grid $\mathcal{J}_k$ that are entirely contained in S, and let

$$\mu_j = \min_{(s,t)\in P_j} |J_{\boldsymbol{\varphi}}(s,t)|, \quad j=1,\ldots,J.$$

Because $S \supseteq P_1 \cup\cdots\cup P_J$ and $\boldsymbol{\varphi}(S) \supseteq \boldsymbol{\varphi}(P_1)\cup\cdots\cup\boldsymbol{\varphi}(P_J)$, we have

$$A(\boldsymbol{\varphi}(S)) \geq \sum_{j=1}^{J} A(\boldsymbol{\varphi}(P_j)) \geq \sum_{j=1}^{J} \mu_j A(P_j) = \underline{D}_{\mathcal{J}_k}(|J_{\boldsymbol{\varphi}}|, S),$$

the lower Darboux sum for $|J_{\boldsymbol{\varphi}}|$ over S and the grid $\mathcal{J}_k$. In the limit as $k\to\infty$, the Darboux sum becomes the integral:

$$A(\boldsymbol{\varphi}(S)) \geq \iint_S |J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt. \qquad \square$$

This completes the proof of Theorem 9.8. $\square$

New grids from nonlinear maps

We have already constructed new integration grids from old ones using invertible linear maps (cf. pp. 287–293): if the grid $\mathcal{G}$ has cells Q_i and $L:\mathbb{R}^2\to\mathbb{R}^2$ is invertible, then the images $L(Q_i)$ form the cells of a new grid $\mathcal{H}=L(\mathcal{G})$. Furthermore, L determines a constant σ that relates the mesh sizes of $\mathcal{G}$ and $\mathcal{H}$: $\|\mathcal{H}\| = \sigma\,\|\mathcal{G}\|$. Theorem 9.8 creates the possibility of creating new grids using invertible nonlinear maps.

Let $\boldsymbol{\varphi}:\Omega\to\boldsymbol{\varphi}(\Omega)$ be continuously differentiable with a continuously differentiable inverse on $\boldsymbol{\varphi}(\Omega)$. Let $\mathcal{G}$ be a grid whose cells Q_i lie within Ω. By definition, the Q_i are closed nonoverlapping sets with area. By Theorem 9.8, the sets $P_i=\boldsymbol{\varphi}(Q_i)$ are likewise closed nonoverlapping sets with area. We define them to be the cells of the grid $\mathcal{H}=\boldsymbol{\varphi}(\mathcal{G})$.

Restricting a grid to a closed set

For a nonlinear map $\boldsymbol{\varphi}$, there is no general analogue to the scale factor σ. However, suppose S has area and is a closed subset of Ω. Then the sets

$$\widehat{Q}_i = Q_i\cap S \text{ and } \widehat{P}_i = P_i\cap\boldsymbol{\varphi}(S) = \boldsymbol{\varphi}(\widehat{Q}_i)$$

are closed nonoverlapping sets with area, so they constitute the cells of grids that we denote $\mathcal{G}_S$ and $\mathcal{H}_{\boldsymbol{\varphi}(S)}$, respectively. For these special grids, there is a natural link between their mesh sizes. To find it, note that the continuous map $\boldsymbol{\varphi}$ is uniformly continuous on S. Therefore, given any $\varepsilon>0$ there is a $\delta>0$ such that

$$\|\mathbf{s}_1-\mathbf{s}_2\|<\delta \quad\Longrightarrow\quad \|\boldsymbol{\varphi}(\mathbf{s}_1)-\boldsymbol{\varphi}(\mathbf{s}_2)\|<\varepsilon.$$

This implies

$$\|\mathcal{G}_S\|<\delta \quad\Longrightarrow\quad \|\mathcal{H}_{\boldsymbol{\varphi}(S)}\|<\varepsilon;$$

in other words, we can make $\|\mathcal{H}_{\boldsymbol{\varphi}(S)}\|$ as small as we wish by making $\|\mathcal{G}_S\|$ sufficiently small.

When calculating Riemann sums for a continuous function f defined only on a set S, we first define f to be zero outside S in order to allow for the possibility that f is evaluated on the part of a cell that lies outside S. Because f on this larger domain is usually not continuous, a delicate argument is needed to show that the Riemann sums converge. The grids $\mathcal{G}_S$ eliminate this problem, because their cells lie entirely in S. The following theorem shows that Riemann sums constructed with these restricted grids still converge to the integral.

Theorem 9.10. *Suppose f is continuous on a closed bounded set S that has area. Then Riemann sums constructed with grids $\mathcal{G}_S$ converge to the integral of f as $\|\mathcal{G}_S\| \to 0$.*

Proof. Let $\varepsilon > 0$ be given; we must find a $\delta > 0$ so that

$$\left| \iint_S f(x,y)\,dA - \sum_{i=1}^{I} f(x_i,y_i)A(\widehat{Q}_i) \right| < \varepsilon$$

for any grid $\mathcal{G}_S$ with $\|\mathcal{G}_S\| < \delta$, and for any point (x_i,y_i) in the cell $\widehat{Q}_i$ of $\mathcal{G}_S$, for each $i = 1,\ldots,I$.

Because f is uniformly continuous on S, we can choose $\delta > 0$ so that

$$\|(x,y) - (x',y')\| < \delta \quad \Longrightarrow \quad |f(x,y) - f(x',y')| < \frac{\varepsilon}{A(S)}.$$

Now let $\mathcal{G}_S$ be any grid for which $\|\mathcal{G}_S\| < \delta$. Then, because

$$\iint_S f(x,y)\,dA = \sum_{i=1}^{I} \iint_{\widehat{Q}_i} f(x,y)\,dA \ \text{ and } \ f(x_i,y_i)A(\widehat{Q}_i) = \iint_{\widehat{Q}_i} f(x_i,y_i)\,dA,$$

we have

$$\begin{aligned}\left| \iint_S f(x,y)\,dA - \sum_{i=1}^{I} f(x_i,y_i)A(\widehat{Q}_i) \right| &\le \left| \sum_{i=1}^{I} \iint_{\widehat{Q}_i} f(x,y)\,dA - \sum_{i=1}^{I} \iint_{\widehat{Q}_i} f(x_i,y_i)\,dA \right| \\ &\le \sum_{i=1}^{I} \iint_{\widehat{Q}_i} \left| f(x,y) - f(x_i,y_i) \right| dA < \frac{\varepsilon}{A(S)} \sum_{i=1}^{I} \iint_{\widehat{Q}_i} dA = \varepsilon. \qquad \square\end{aligned}$$

Change of variables in double integrals

We can use this result immediately, to prove the main formula on the *change of variables* in double integrals (by continuing to follow the argument of J. Schwartz in [16]).

Theorem 9.11 (Change of variables). *Let Ω be a bounded open set in $\mathbb{R}^2$, and let $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^2 : (s,t) \to (x,y)$ be a continuously differentiable map that has a continuously differentiable inverse $\boldsymbol{\varphi}^{-1} : \boldsymbol{\varphi}(\Omega) \to \Omega$. Suppose the function $f(x,y)$ is continuous on a closed set $D \subset \boldsymbol{\varphi}(\Omega)$ that has area; then*

$$\iint\limits_{D} f(x,y)\,dx\,dy = \iint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(x(s,t),y(s,t)) \left| \frac{\partial(x,y)}{\partial(s,t)} \right| ds\,dt.$$

Proof. By Theorem 8.35, page 305, $f(x,y)$ is integrable on D. By Theorem 9.8, $S = \boldsymbol{\varphi}^{-1}(D)$ has area and the function

$$f(x(s,t),y(s,t))\left|\frac{\partial(x,y)}{\partial(s,t)}\right|$$

is bounded and continuous on $S = \boldsymbol{\varphi}^{-1}(D)$, so it is integrable there. To prove that the two integrals in the statement of the theorem are equal, we show they differ by less than any preassigned $\varepsilon > 0$.

Let $\mathcal{G}_S$ be an arbitrary integration grid whose cells Q_i, $i = 1,\ldots,I$, partition S, and let $\mathcal{H}_D = \boldsymbol{\varphi}(\mathcal{G}_S)$ be the image grid; its cells $P_i = \boldsymbol{\varphi}(Q_i)$ partition $D = \boldsymbol{\varphi}(S)$. Let (s_i,t_i) be a point in Q_i, and let $(x_i,y_i) = \boldsymbol{\varphi}(s_i,t_i)$ be the corresponding point in P_i. Consider the following, obtained by applying the triangle inequality to a rather lengthy telescoping sum:

$$\begin{aligned}
&\left|\iint_D f(x,y)\,dx\,dy - \iint_S f(\boldsymbol{\varphi}(s,t))\,|J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt\right| \\
&\quad\le \left|\iint_D f(x,y)\,dx\,dy - \sum_{i=1}^{I} f(x_i,y_i)A(P_i)\right| \\
&\qquad + \left|\sum_{i=1}^{I} f(x_i,y_i)A(P_i) - \sum_{i=1}^{I}\iint_{Q_i} f(\boldsymbol{\varphi}(s_i,t_i))\,|J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt\right| \\
&\qquad + \left|\sum_{i=1}^{I}\iint_{Q_i} \big(f(\boldsymbol{\varphi}(s_i,t_i)) - f(\boldsymbol{\varphi}(s,t))\big)\,|J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt\right| \\
&\qquad + \left|\sum_{i=1}^{I}\iint_{Q_i} f(\boldsymbol{\varphi}(s,t))\,|J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt - \iint_S f(\boldsymbol{\varphi}(s,t))\,|J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt\right|.
\end{aligned}$$

Now consider each of the four terms on the right-hand side of the inequality.

The first term contains a Riemann sum for f on D and the grid $\mathcal{H}_D$. Because f is integrable, there is a $\delta_1 > 0$ that makes that term less than $\varepsilon/2$ whenever $\|\mathcal{H}_D\| < \delta_1$ and (x_i,y_i) is an arbitrary point in P_i. As we saw above (p. 349), we have $\|\mathcal{H}_D\| < \delta_1$ when $\|\mathcal{G}_S\| < \delta_2$ for some properly chosen $\delta_2 > 0$.

The second term is zero by Theorem 9.8:

$$A(P_i) = A(\boldsymbol{\varphi}(Q_i)) = \iint_{Q_i} |J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt, \quad i = 1,\ldots,I.$$

For the third term, first note that

$$\begin{aligned}
&\left|\sum_{i=1}^{I}\iint_{Q_i} \big(f(\boldsymbol{\varphi}(s_i,t_i)) - f(\boldsymbol{\varphi}(s,t))\big)\,|J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt\right| \\
&\qquad\qquad \le \sum_{i=1}^{I}\iint_{Q_i} \big|f(\boldsymbol{\varphi}(s_i,t_i)) - f(\boldsymbol{\varphi}(s,t))\big|\,|J_{\boldsymbol{\varphi}}(s,t)|\,ds\,dt.
\end{aligned}$$

Because $f(\boldsymbol{\varphi}(s,t))$ is uniformly continuous on S, there is a $\delta_3 > 0$ for which

$$\|(s,t)-(s',t')\| < \delta_3 \quad \Longrightarrow \quad \big|f(\boldsymbol{\varphi}(s,t)) - f(\boldsymbol{\varphi}(s',t'))\big| < \frac{\varepsilon}{2A(D)}.$$

Therefore, if $\mathcal{G}_S$ is any partition of S for which $\|\mathcal{G}_S\| < \delta_3$, then

$$\sum_{i=1}^{I} \iint_{Q_i} \big|f(\boldsymbol{\varphi}(s_i,t_i)) - f(\boldsymbol{\varphi}(s,t))\big|\, |J_{\boldsymbol{\varphi}}(s,t)|\, ds\, dt$$
$$< \frac{\varepsilon}{2A(D)} \sum_{i=1}^{I} \iint_{Q_i} |J_{\boldsymbol{\varphi}}(s,t)|\, ds\, dt = \frac{\varepsilon}{2A(D)} \iint_{S} |J_{\boldsymbol{\varphi}}(s,t)|\, ds\, dt = \frac{\varepsilon}{2}.$$

The last equality in this chain is provided by Theorem 9.8:

$$\iint_S |J_{\boldsymbol{\varphi}}(s,t)|\, ds\, dt = A(\boldsymbol{\varphi}(S)) = A(D).$$

The fourth term, like the second, is zero. Therefore, the two integrals differ by less than ε whenever the partition $\mathcal{G}_S$ satisfies $\|\mathcal{G}_S\| < \min \delta_2, \delta_3$. Because $\varepsilon > 0$ is arbitrary, the two integrals must be equal. □

Local area magnification

Because $|\det L|$ is the area magnification factor for a linear map L of the plane, we have

$$\frac{A(L(S))}{A(S)} = |\det L|$$

for any subset of the plane that has area. For the nonlinear map $\boldsymbol{\varphi}$, we introduce the set function (cf. pp. 310–312)

$$M(S) = A(\boldsymbol{\varphi}(S)) = \iint_S |J_{\boldsymbol{\varphi}}(s,t)|\, ds\, dt$$

By Theorem 8.39, page 312, the derivative of M is

$$M'(s,t) = |J_{\boldsymbol{\varphi}}(s,t)|$$

In other words, if S contains the point (s,t), then

$$\frac{A(\boldsymbol{\varphi}(S))}{A(S)} \approx |J_{\boldsymbol{\varphi}}(s,t)|$$

as closely as we wish by making the diameter of S (p. 291) sufficiently small. It is in this sense that we consider

$$|J_{\boldsymbol{\varphi}}(s,t)| = |\det \mathrm{d}\boldsymbol{\varphi}_{(s,t)}|$$

to be the **local area magnification factor** for $\boldsymbol{\varphi}$ when *area* is understood to be nonnegative Jordan content.

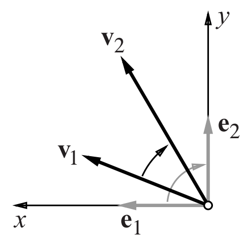
9.4 Orientation

In the next section, we introduce Green's theorem as an additional tool to evaluate double integrals. However, the integrals in Green's theorem are oriented. In this section, therefore, we say what it means for a 2-dimensional region to be oriented, and then define *oriented* double integrals. Finally, we extend the change of variables formula to oriented integrals.

Orientation of an ordered pair of vectors

Orientation in the plane involves, either explicitly or implicitly, comparison with the coordinate axes (or with the standard basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$ that determine them). Consider first an ordered pair of linearly independent vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ in $\mathbb{R}^2$. To compare $\{\mathbf{v}_1, \mathbf{v}_2\}$ with $\{\mathbf{e}_1, \mathbf{e}_2\}$, let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be the unique linear map for which $L(\mathbf{e}_1) = \mathbf{v}_1$, $L(\mathbf{e}_2) = \mathbf{v}_2$. Then we say the pair $\{\mathbf{v}_1, \mathbf{v}_2\}$ has the **same orientation** as $\{\mathbf{e}_1, \mathbf{e}_2\}$ if $\det L > 0$; otherwise, we say it has the **opposite orientation**. In particular, if we reverse the order of the vectors, orientation is reversed, as well: $\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\{\mathbf{v}_2, \mathbf{v}_1\}$ always have opposite orientations. We write $\{\mathbf{v}_2, \mathbf{v}_1\} = -\{\mathbf{v}_1, \mathbf{v}_2\}$. The components of $\mathbf{v}_1$ and $\mathbf{v}_2$ with respect to the standard basis determine the orientation of $\{\mathbf{v}_1, \mathbf{v}_2\}$. That is, if

$$\mathbf{v}_1 = \begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} v_{12} \\ v_{22} \end{pmatrix},$$



then the matrix of L in terms of the standard basis is

$$L = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}, \quad \text{and } \det L = v_{11}v_{22} - v_{12}v_{21}.$$

Positive and negative orientations

We also say that an ordered pair that has the same orientation as the standard basis is **positively oriented**; otherwise, we say it is **negatively oriented**. The figure in the margin helps make the point that orientation is always determined by reference to the coordinate axes: it is relative, not absolute. The pair $\{\mathbf{v}_1, \mathbf{v}_2\}$ illustrated is positively oriented.

Orientation and sense of rotation

The figure also shows that the ordering of a pair of linearly independent vectors implicitly defines a sense of rotation, namely, rotation from the first to the second through the smaller of the two angles determined by the vectors. "Sense of rotation" therefore gives us a second way to represent orientation. For example, we can confirm that the pair $\{\mathbf{v}_1, \mathbf{v}_2\}$ in the margin figure is positively oriented because it defines the same clockwise sense of rotation as the basis vectors.

Orienting a region in the plane

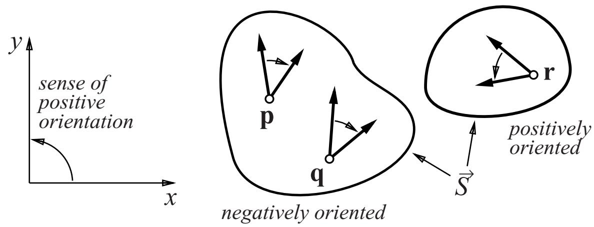
To orient a region S with area in $\mathbb{R}^2$, orient each point $\mathbf{p}$ of S by assigning to $\mathbf{p}$ an ordered pair of linearly independent vectors $\{\mathbf{v}_1(\mathbf{p}), \mathbf{v}_2(\mathbf{p})\}$ in such a way that both $\mathbf{v}_1(\mathbf{p})$ and $\mathbf{v}_2(\mathbf{p})$ vary continuously with $\mathbf{p}$ over S. To indicate that S has acquired an orientation, we write it as $\vec{S}$. (On page 7, we introduced a similar notation for curves: $\vec{C}$ denotes a curve C together with an orientation.) There are now two different definitions of orientation when $\vec{S}$ is a parallelogram, but they agree if we assign to each point of $\mathbf{v} \wedge \mathbf{w}$ the ordered pair $\{\mathbf{v}, \mathbf{w}\}$. For each point $\mathbf{p}$ in $\vec{S}$, let $L_{\mathbf{p}} : \mathbb{R}^2 \to \mathbb{R}^2$ be the unique linear map for which $L_{\mathbf{p}}(\mathbf{e}_i) = \mathbf{v}_i(\mathbf{p})$, $i = 1, 2$. Then, by what we said above, the function

$$\det L_{\mathbf{p}} = v_{11}(\mathbf{p})v_{22}(\mathbf{p}) - v_{12}(\mathbf{p})v_{21}(\mathbf{p})$$

varies continuously with $\mathbf{p}$ and is never zero. We write $-\vec{S}$ to denote $\vec{S}$ with its orientation reversed at every point. That is, if $\{\mathbf{v}_1(\mathbf{p}), \mathbf{v}_2(\mathbf{p})\}$ defines $\vec{S}$, then $-\{\mathbf{v}_1(\mathbf{p}), \mathbf{v}_2(\mathbf{p})\} = \{\mathbf{v}_2(\mathbf{p}), \mathbf{v}_1(\mathbf{p})\}$ defines $-\vec{S}$.

Theorem 9.12. *On any pathwise-connected component of $\vec{S}$ (i.e., a largest subset in which any two points can be joined by a continuous path in $\vec{S}$), all points have the same orientation.*

Proof. Let $\mathbf{p}$ and $\mathbf{q}$ be joined by the continuous path $\mathbf{s}(t)$, $a \le t \le b$, with $\mathbf{s}(a) = \mathbf{p}$ and $\mathbf{s}(b) = \mathbf{q}$. Then $\det L_{\mathbf{s}(t)}$ is a continuous and nonzero function of t on the interval $[a,b]$, so it cannot change sign. Therefore $\{\mathbf{v}_1(\mathbf{p}), \mathbf{v}_2(\mathbf{p})\}$ and $\{\mathbf{v}_1(\mathbf{q}), \mathbf{v}_2(\mathbf{q})\}$ have the same orientation. □



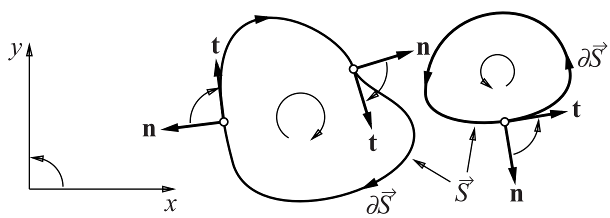
Sense of rotation at every point of $\vec{S}$

In the figure, $\vec{S}$ has two components, one with positive orientation and the other with negative. As the figure suggests, the orientation of $\vec{S}$ always defines a "sense of rotation" at each point of $\vec{S}$, and all points in any connected component of $\vec{S}$ have the same sense of rotation.

Definition 9.3 *Two assignments* $\mathbf{p} \mapsto \{\mathbf{w}_1(\mathbf{p}), \mathbf{w}_2(\mathbf{p})\}$, $\mathbf{p} \mapsto \{\mathbf{v}_1(\mathbf{p}), \mathbf{v}_2(\mathbf{p})\}$ *define the* ***same orientation of S*** *if the unique linear map* $M_{\mathbf{p}}(\mathbf{v}_i(\mathbf{p})) = \mathbf{w}_i(\mathbf{p})$, $i = 1,2$, *has* $\det M_{\mathbf{p}} > 0$ *for all* $\mathbf{p}$ *in S.*

Positive and negative orientations

Two different assignments of ordered pairs of vectors define the same orientation precisely when they give the same sense of rotation. Thus, each component of a region S can be given exactly two orientations, either agreeing or disagreeing with the sense of rotation of the coordinate axes. If S has k components, it has 2^k possible orientations. We say $\vec{S}$ is **positively oriented** if all its components are positively oriented, and is **negatively oriented** if all its components are negatively oriented. Usually, S has just one component.



If ∂S is a piecewise-smooth curve, then an orientation of S *induces* an orientation of ∂S. To see how, let $\vec{S}$ be oriented. As the previous definition indicates, there is considerable freedom in choosing the orientation vectors at a point. Thus, orient each point of $\partial\vec{S}$ where it is smooth by the pair of vectors $\{\mathbf{n},\mathbf{t}\}$, where $\mathbf{n}$ is the outward-pointing unit normal and $\mathbf{t}$ is one of the two unit tangents to $\partial\vec{S}$, choosing $\mathbf{t}$ so $\{\mathbf{n},\mathbf{t}\}$ gives the local sense of rotation on that component of $\vec{S}$. In the figure above, the pair $\{\mathbf{n},\mathbf{t}\}$ is oriented in the clockwise sense on one component and in the counterclockwise sense on the other. The choice of a tangent vector orients a piecewise-smooth curve. We use the tangent vector $\mathbf{t}$ to define the **orientation of ∂S induced by $\vec{S}$**. Orientation induced on a boundary

A map can also induce an orientation of its image. Let Ω be a bounded open set in $\mathbb{R}^2$, and let $\boldsymbol{\varphi}:\Omega\to\mathbb{R}^2$ be a continuously differentiable map that has a continuously differentiable inverse $\boldsymbol{\varphi}^{-1}:\boldsymbol{\varphi}(\Omega)\to\Omega$. Suppose $\vec{S}$ is an oriented set that has area and its closure $\overline{S}=S\cup\partial S$ lies within Ω. Suppose the ordered pair $\{\mathbf{v}_1(\mathbf{p}),\mathbf{v}_2(\mathbf{p})\}$ defines the orientation of $\vec{S}$ at the point $\mathbf{p}$. Then we orient the point $\boldsymbol{\varphi}(\mathbf{p})$ in $\boldsymbol{\varphi}(\vec{S})$ with the ordered pair of vectors Orientation induced on an image

$$\{d\boldsymbol{\varphi}_{\mathbf{p}}(\mathbf{v}_1(\mathbf{p})),d\boldsymbol{\varphi}_{\mathbf{p}}(\mathbf{v}_2(\mathbf{p}))\}.$$

Because $\boldsymbol{\varphi}$ is invertible, each image point $\boldsymbol{\varphi}(\mathbf{p})$ is assigned only one such pair. Because $d\boldsymbol{\varphi}_{\mathbf{p}}$ is invertible, the image vectors are linearly independent. Finally, because $\boldsymbol{\varphi}$ is continuously differentiable, the assignment varies continuously with $\boldsymbol{\varphi}(\mathbf{p})$.

In Chapter 4 we first observed informally that the sign of the Jacobian of a map determines whether it preserves or reverses orientation. Now that we have defined the orientation of a region, we can state this observation as a theorem and prove it. Orientation and the Jacobian

Theorem 9.13. *The regions $\vec{S}$ and $\boldsymbol{\varphi}(\vec{S})$ have the same orientation if and only if the Jacobian* $\det d\boldsymbol{\varphi}_{\mathbf{p}}$ *is everywhere positive.*

Proof. Suppose the ordered pair $\{\mathbf{v}_1(\mathbf{p}),\mathbf{v}_2(\mathbf{p})\}$ defines the orientation of $\vec{S}$ at $\mathbf{p}$; then $\{d\boldsymbol{\varphi}_{\mathbf{p}}(\mathbf{v}_1(\mathbf{p})),d\boldsymbol{\varphi}_{\mathbf{p}}(\mathbf{v}_2(\mathbf{p}))\}$ defines the induced orientation of $\boldsymbol{\varphi}(\vec{S})$ at $\boldsymbol{\varphi}(\mathbf{p})$. Let

$$L_{\mathbf{p}}(\mathbf{e}_i)=\mathbf{v}_i(\mathbf{p}) \text{ and } M_{\mathbf{p}}(\mathbf{e}_i)=d\boldsymbol{\varphi}_{\mathbf{p}}(\mathbf{v}_i(\mathbf{p})),\quad i=1,2,$$

define the linear maps that determine the orientations. Then $\boldsymbol{\varphi}(\vec{S})$ has the same orientation at $\boldsymbol{\varphi}(\mathbf{p})$ that $\vec{S}$ does at $\mathbf{p}$ if and only if the determinants $\det M_{\mathbf{p}}$ and $\det L_{\mathbf{p}}$ have the same sign. But

$$M_{\mathbf{p}}=d\boldsymbol{\varphi}_{\mathbf{p}}\circ L_{\mathbf{p}},\quad \det M_{\mathbf{p}}=\det d\boldsymbol{\varphi}_{\mathbf{p}}\det L_{\mathbf{p}},$$

so $\det M_{\mathbf{p}}$ and $\det L_{\mathbf{p}}$ have the same sign if and only if $\det d\boldsymbol{\varphi}_{\mathbf{p}}>0$. $\square$

We can now introduce *oriented* integrals, that is, double integrals defined over oriented regions. We begin with a closed, bounded, and *unoriented* subset S of the (x,y)-plane. Assume S has area and $f(x,y)$ is a function that is integrable over S; then we have the ordinary Riemann integral Oriented integrals

$$\iint_S f(x,y)\,dA$$

as defined in Chapter 8.3. This integral is monotonic in the sense that $f(x,y) \geq 0$ on S implies $I \geq 0$ (Theorem 8.28, p. 298).

Definition 9.4 *If $\vec{S}$ has either positive or negative orientation, then the **oriented integral of f over $\vec{S}$** is*

$$\iint_{\vec{S}} f(x,y)\,dx\,dy = \operatorname{sgn}\vec{S} \iint_S f(x,y)\,dA,$$

where $\operatorname{sgn}\vec{S} = +1$ when the orientation of $\vec{S}$ is positive and $\operatorname{sgn}\vec{S} = -1$ when it is negative.

The oriented integral uses $dx\,dy$ rather than as dA as the "element of area" in order to help convey orientation, in a way we explain below.

Properties of oriented integrals

The definition has several immediate consequences. First, because $-\vec{S}$ is negatively oriented when $\vec{S}$ is positively oriented, and vice versa, their oriented integrals have opposite signs:

$$\iint_{-\vec{S}} f(x,y)\,dx\,dy = -\iint_{\vec{S}} f(x,y)\,dx\,dy.$$

Second, an oriented integral over a positively oriented region $\vec{S}$ is monotonic:

$$\iint_{\vec{S}} f(x,y)\,dx\,dy \geq 0 \text{ if } f \geq 0 \text{ on } \vec{S}.$$

Third, we can define the *signed* area of a positively or negatively oriented region $\vec{S}$ as

$$\operatorname{area}\vec{S} = \iint_{\vec{S}} dx\,dy = \operatorname{sgn}\vec{S} \times A(S),$$

where

$$A(S) = \iint_S dA$$

is the ordinary area (i.e., the Jordan measure) of the unoriented region S. Oriented area reverses sign with the orientation of the region:

$$\operatorname{area}(-\vec{S}) = -\operatorname{area}(\vec{S}).$$

$dy\,dx = -dx\,dy$

Writing the element of area in an oriented integral as $dx\,dy$ rather than dA gives us the opportunity to convey differences in orientation. If we take $dx\,dy$ as representing the coordinate axes in their usual order (i.e., positive orientation), then $-dx\,dy$ and $dy\,dx$ should both represent the opposite order (negative orientation). Let us, therefore, adopt the symbolic convention

$$dy\,dx = -dx\,dy,$$

so that

$$\iint_{\vec{S}} f(x,y)\,dy\,dx = -\iint_{\vec{S}} f(x,y)\,dx\,dy = \iint_{-\vec{S}} f(x,y)\,dx\,dy$$

for any positively or negatively oriented region $\vec{S}$. In particular,

$$\iint_{\vec{S}} dy\,dx = -\iint_{\vec{S}} dx\,dy = -\operatorname{area}\vec{S}.$$

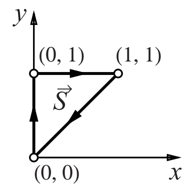
Order in iterated integrals

It is important to note that the sign change when switching from $dy\,dx$ to $dx\,dy$ in an oriented integral does not happen when we reverse the order of integration in iterated integrals (Corollary 9.2, p. 321). For example, if we integrate $f(x,y)$ over the rectangle $R : a \le x \le b, c \le y \le d$ and assume $\vec{R}$ is negatively oriented, then

$$\begin{aligned}\iint_{\vec{R}} f(x,y)\,dx\,dy &= -\iint_R f(x,y)\,dA \\ &= -\int_c^d \int_a^b f(x,y)\,dx\;dy = -\int_a^b \int_c^d f(x,y)\,dy\;dx.\end{aligned}$$

Example

As an example, let us compute the oriented integral of $f(x,y) = x^2$ over the triangle $\vec{S}$ with vertices $(0,0)$, $(0,1)$, and $(1,1)$, taken in that order. The boundary path indicates that $\vec{S}$ is negatively oriented. Because $f(x,y) \ge 0$ on $\vec{S}$, we therefore expect the value of the integral to be negative. We first express the oriented integral as an ordinary (unoriented) double integral, and then convert that to iterated integrals. For the last step, we can describe the unoriented set S by either of the following sets of inequalities:



$$S: \begin{array}{l} 0 \le x \le 1, \\ x \le y \le 1; \end{array} \qquad S: \begin{array}{l} 0 \le y \le 1, \\ 0 \le x \le y. \end{array}$$

Using the first set, we have

$$\begin{aligned}\iint_{\vec{S}} x^2\,dx\,dy &= -\iint_S x^2\,dA = -\int_0^1 \left(\int_x^1 x^2 dy\right) dx = -\int_0^1 x^2 y\Big|_x^1\, dx \\ &= -\int_0^1 (x^2 - x^3)\,dx = -\left(\frac{1}{3} - \frac{1}{4}\right) = -\frac{1}{12}.\end{aligned}$$

The second set leads to the same result.

Change of variables with orientation

In the *oriented* form of the change of variables formula for single integrals,

$$\int_a^b f(x)\,dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(s))\,\varphi'(s)\,ds,$$

the sign of the Jacobian $\varphi'(s)$ indicates whether the interval from a to b and the interval from $\varphi^{-1}(a)$ to $\varphi^{-1}(b)$ have the same or opposite orientations. Here is the analogous formula for oriented double integrals.

Theorem 9.14 (Oriented change of variables). *Let Ω be a bounded open set in $\mathbb{R}^2$, and let $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^2 : (s,t) \to (x,y)$ be a continuously differentiable map that has* *a continuously differentiable inverse* $\boldsymbol{\varphi}^{-1} : \boldsymbol{\varphi}(\Omega) \to \Omega$. *Suppose the function* $f(x,y)$ *is continuous on* $\vec{D} \subset \boldsymbol{\varphi}(\Omega)$, *a closed, oriented, and pathwise-connected region that has area; then*

$$\iint\limits_{\vec{D}} f(x,y)\,dx\,dy = \iint\limits_{\boldsymbol{\varphi}^{-1}(\vec{D})} f(x(s,t),y(s,t))\,\frac{\partial(x,y)}{\partial(s,t)}\,ds\,dt.$$

Proof. Because $\vec{D}$ has only one pathwise-connected component, all of its points have the same orientation (Theorem 9.12), so $\operatorname{sgn}(\vec{D})$ is defined. Furthermore, the Jacobian $J_{\boldsymbol{\varphi}}$ cannot change sign on $\vec{D}$, so all points of $\boldsymbol{\varphi}^{-1}(\vec{D})$ have the same orientation, and $\operatorname{sgn}\boldsymbol{\varphi}^{-1}(\vec{D}) = \operatorname{sgn}J_{\boldsymbol{\varphi}^{-1}}\,\operatorname{sgn}\vec{D} = \operatorname{sgn}J_{\boldsymbol{\varphi}}\,\operatorname{sgn}\vec{D}$. Thus (using $dA_{x,y}$ and $dA_{s,t}$ to denote the unoriented elements of area),

$$\begin{aligned}
\iint\limits_{\vec{D}} f(x,y)\,dx\,dy &= \operatorname{sgn}\vec{D}\iint\limits_{D} f(x,y)\,dA_{x,y}\\
&= \operatorname{sgn}\vec{D}\iint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(x(s,t),y(s,t))\left|\frac{\partial(x,y)}{\partial(s,t)}\right| dA_{s,t} \qquad \text{(Theorem 9.11)}\\
&= \operatorname{sgn}\vec{D}\ \operatorname{sgn}J_{\boldsymbol{\varphi}}\iint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(x(s,t),y(s,t))\,\frac{\partial(x,y)}{\partial(s,t)}\,dA_{s,t}\\
&= \operatorname{sgn}\boldsymbol{\varphi}^{-1}(D)\iint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(x(s,t),y(s,t))\,\frac{\partial(x,y)}{\partial(s,t)}\,dA_{s,t}\\
&= \iint\limits_{\boldsymbol{\varphi}^{-1}(\vec{D})} f(x(s,t),y(s,t))\,\frac{\partial(x,y)}{\partial(s,t)}\,ds\,dt. \qquad \square
\end{aligned}$$

Summary

The following summary points out parallels between the ways that elements of oriented single and double integrals transform under a change of variables ($\vec{I}$ is an oriented interval on the x-axis):

$$x \to x(s) \qquad dx \to \frac{dx}{ds}\,ds, \qquad \vec{I} \to \varphi^{-1}(\vec{I})$$

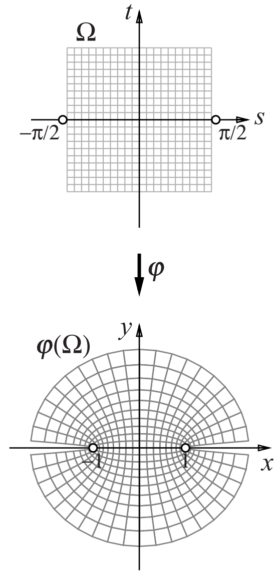
$$\begin{cases} x \to x(s,t), \\ y \to y(s,t), \end{cases} \qquad dx\,dy \to \frac{\partial(x,y)}{\partial(s,t)}\,ds\,dt, \qquad \vec{D} \to \boldsymbol{\varphi}^{-1}(\vec{D}).$$

Example 1: areas in curvilinear coordinates

To illustrate the use of the oriented change of variables formula, we first compute the signed area of a curvilinear quadrilateral specified by curvilinear coordinates $(s,t) \mapsto (x(s,t),y(s,t))$ in the (x,y)-plane. Let Ω be the infinite strip in the (s,t)-plane given by $-\pi/2 < s < \pi/2$, and let $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^2$ be

$$\boldsymbol{\varphi} : \begin{cases} x = \sin s \cosh t, \\ y = \cos s \sinh t; \end{cases} \qquad \mathrm{d}\boldsymbol{\varphi}_{(s,t)} = \begin{pmatrix} \cos s \cosh t & \sin s \sinh t \\ -\sin s \sinh t & \cos s \cosh t \end{pmatrix}.$$

Everywhere on Ω, $\boldsymbol{\varphi}$ is orientation-preserving:



$$\frac{\partial(x,y)}{\partial(s,t)} = \cos^2 s \cosh^2 t + \sin^2 s \sinh^2 t = \cos^2 s + \sinh^2 t > 0.$$

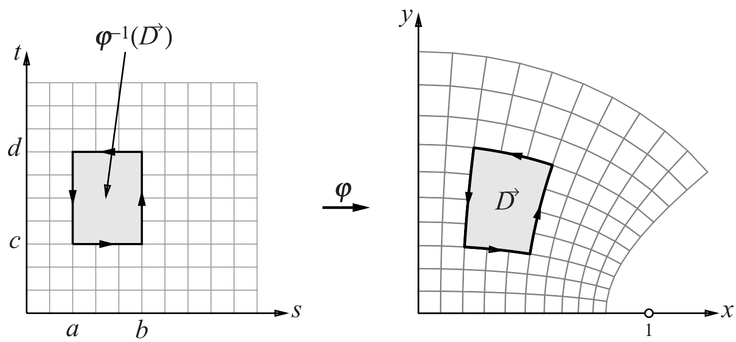
Note: $\cosh^2 t = 1 + \sinh^2 t$). In fact, $\boldsymbol{\varphi}$ is a *conformal* map (Definition 4.2, p. 118) that "flares out" Ω in such a way that the image of the vertical line $s =$ constant lies on the hyperbola

$$\frac{x^2}{\sin^2 s} - \frac{y^2}{\cos^2 s} = 1.$$

The focal points of this hyperbola are $(\pm f, 0)$, where $f^2 = \sin^2 s + \cos^2 s = 1$ (see Exercise 9.24). The hyperbolas for various s are thus *confocal*. The image of the horizontal line $t =$ constant lies on the ellipse

$$\frac{x^2}{\cosh^2 s} + \frac{y^2}{\sinh^2 s} = 1.$$

Its focal points are $(\pm f, 0)$, where $f^2 = \cosh^2 t - \sinh^2 t = 1$. Thus the ellipses and hyperbolas are all simultaneously confocal; the image of Ω is the entire plane minus the two rays $|x| \geq 1$ on the x-axis.



Let $\vec{D}$ be the positively oriented curvilinear quadrilateral in the (x,y)-plane bounded by

$$\frac{x^2}{\sin^2 a} - \frac{y^2}{\cos^2 a} = 1, \qquad \frac{x^2}{\cosh^2 c} + \frac{y^2}{\sinh^2 c} = 1,$$
$$\frac{x^2}{\sin^2 b} - \frac{y^2}{\cos^2 b} = 1, \qquad \frac{x^2}{\cosh^2 d} + \frac{y^2}{\sinh^2 d} = 1,$$

where $0 < a < b < \pi/2$ and $0 < c < d$. The rectangle $\boldsymbol{\varphi}^{-1}(\vec{D})$ is also positively oriented, and

$$\operatorname{area}\vec{D} = \iint_{\vec{D}} dx\,dy = \iint_{\boldsymbol{\varphi}^{-1}(\vec{D})} (\cos^2 s + \sinh^2 t)\,ds\,dt = \int_a^b \int_c^d (\cos^2 s + \sinh^2 t)\,ds\,dt.$$

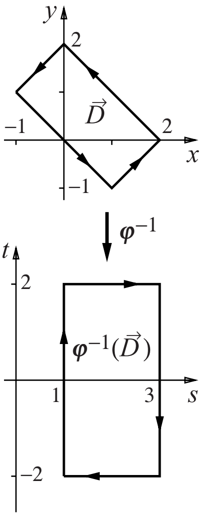
After some standard calculations (see Exercise 9.25), we find

$$\operatorname{area}\vec{D} = (d-c)\frac{\sin 2b - \sin 2a}{4} + (b-a)\frac{\sinh 2d - \sinh 2c}{4}.$$

Example 2

For a second example, let us find

$$\iint_{\vec{D}} \frac{(x-y)^2}{1+x+y}\,dx\,dy,$$



where $\vec{D}$ is the rectangle with vertices $(1,-1)$, $(2,0)$, $(0,2)$, and $(-1,1)$, taken in that order. Thus $\vec{D}$ is positively oriented; as an unoriented set, it is given by the inequalities

$$D: \quad \begin{array}{r} 0 \le x+y \le 2, \\ -2 \le x-y \le 2. \end{array}$$

The form of the integrand and the expressions in these inequalities suggest that we set

$$\boldsymbol{\varphi}^{-1}: \begin{cases} s = 1+x+y, \\ t = x-y; \end{cases} \qquad \frac{\partial(s,t)}{\partial(x,y)} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2.$$

Then $\boldsymbol{\varphi}^{-1}(\vec{D})$ is a simple rectangle with sides parallel to the coordinate axes:

$$\boldsymbol{\varphi}^{-1}(D): \quad \begin{array}{r} 1 \le s \le 3, \\ -2 \le t \le 2. \end{array}$$

Because the Jacobian is negative, $\boldsymbol{\varphi}^{-1}$ and $\boldsymbol{\varphi}$ both reverse orientation, so $\boldsymbol{\varphi}^{-1}(\vec{D})$ has negative orientation. To apply the change of variables formula, we need only

$$\frac{\partial(x,y)}{\partial(s,t)} = \frac{1}{\dfrac{\partial(s,t)}{\partial(x,y)}} = -\frac{1}{2}$$

(Corollary 4.13, p. 138). Therefore,

$$\iint_{\vec{D}} \frac{(x-y)^2}{1+x+y}\,dx\,dy = \iint_{\boldsymbol{\varphi}^{-1}(\vec{D})} \frac{t^2}{s}\left(-\frac{1}{2}\right)ds\,dt = +\frac{1}{2}\iint_{\boldsymbol{\varphi}^{-1}(D)} \frac{t^2}{s}\,dA.$$

In the last equality, we convert the oriented integral over $\boldsymbol{\varphi}^{-1}(\vec{D})$ into an ordinary double integral over the unoriented set $\boldsymbol{\varphi}^{-1}(D)$; the sign change occurs because $\boldsymbol{\varphi}^{-1}(\vec{D})$ is negatively oriented (Definition 9.4). To complete the computation, we write

$$\frac{1}{2}\iint\limits_{\boldsymbol{\varphi}^{-1}(D)} \frac{t^2}{s}\,dA = \frac{1}{2}\int_{-2}^{2} t^2\,dt \int_{1}^{3}\frac{ds}{s} = \frac{1}{2}\,\frac{t^3}{3}\bigg|_{-2}^{2}\,\ln s\bigg|_{1}^{3} = \frac{8\ln 3}{3}.$$

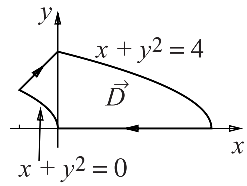
Our third example involves a similar integral, Example 3

$$\iint_{\vec{D}} \frac{(x-y)^2(1+2y)}{1+x+y^2}\,dx\,dy;$$

$\vec{D}$ is the negatively oriented region that satisfies the equalities

$$0 \le x+y^2 \le 4, \quad 0 \le y \le x+2.$$

The integrand is positive on $\vec{D}$, but $\vec{D}$ is negatively oriented; therefore we expect the value of the integral to be negative. Guided once again by the form of the integrand and the shape of $\vec{D}$, we change coordinates with



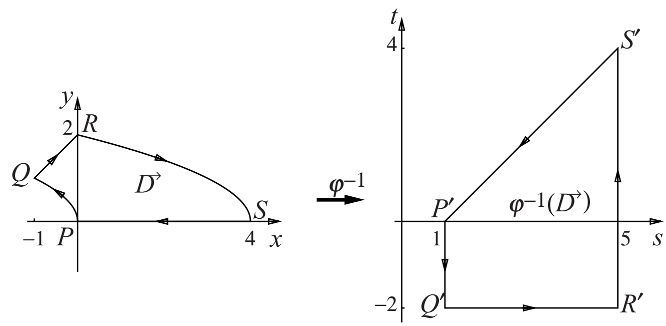
$$\boldsymbol{\varphi}^{-1}: \begin{cases} s = 1+x+y^2, \\ t = x-y; \end{cases} \qquad \frac{\partial(s,t)}{\partial(x,y)} = \begin{vmatrix} 1 & 2y \\ 1 & -1 \end{vmatrix} = -(1+2y).$$

Two of the factors in the integrand transform readily, but $1+2y$ has no simple expression in terms of s and t (but see Exercise 9.27). However, because $\partial(x,y)/\partial(s,t) = -1/(1+2y)$, we find

$$(1+2y)\,dx\,dy \to (1+2y)\,\frac{\partial(x,y)}{\partial(s,t)}\,ds\,dt = -ds\,dt$$

Therefore, because $\boldsymbol{\varphi}^{-1}$ is invertible on $1+2y>0$ (see Exercise 9.27), the change of variables formula is valid and we have

$$\iint\limits_{\vec{D}} \frac{(x-y)^2(1+2y)}{1+x+y^2}\,dx\,dy = -\iint\limits_{\boldsymbol{\varphi}^{-1}(\vec{D})} \frac{t^2}{s}\,ds\,dt,$$



where the image $\boldsymbol{\varphi}^{-1}(\vec{D})$ is the positively oriented trapezoid defined by $1 \le s \le 5$, $-2 \le t \le s-1$. Using iterated integrals and simple antiderivatives, we find

$$-\iint\limits_{\boldsymbol{\varphi}^{-1}(\vec{D})} \frac{t^2}{s}\,ds\,dt = -\int_1^5 \int_{-2}^{s-1} \frac{t^2}{s}\,dt\;ds = -\int_1^5 \frac{t^3}{3s}\bigg|_{-2}^{s-1} ds = -\frac{52}{9} - \frac{7}{3}\ln 5.$$

Essential factors in examples

Notice that the factor $1+2y$ was crucial; without it, the integral would not have been so easy to transform. In the integrals

$$\int (9+y+y^2)^{47}\,dy \quad \text{and} \quad \int (9+y+y^2)^{47}(1+2y)\,dy,$$

the same factor $1+2y$ plays the same role; the typographically simpler integral on the left is mathematically more ponderous. When examples are contrived for instructional purposes, they include such essential factors.

Local area multiplier with orientation

In Chapter 4, we defined the *local area multiplier* (or *magnification factor*) of a nonlinear map $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^2$ at a point (a,b) to be the area multiplier of the linear approximation $d\boldsymbol{\varphi}_{(a,b)}$ at that point (Definition 4.4, p. 115). In the previous section of this chapter, we justified that definition, at least up to sign. However, using the notions of orientation and signed area, we can now remove the sign restriction.

Theorem 9.15. *Suppose $\boldsymbol{\varphi}(s,t)$ is continuously differentiable on an open set U, and the Jacobian $J_{\boldsymbol{\varphi}}(a,b) \neq 0$ at some point (a,b) in U. Then*

$$\frac{\operatorname{area}\boldsymbol{\varphi}(\vec{S})}{\operatorname{area}\vec{S}} \to J_{\boldsymbol{\varphi}}(a,b)$$

as the diameter $\delta(\vec{S}) \to 0$, where the limit is taken over closed oriented sets $\vec{S}$ that have signed area and contain the point (a,b).

Proof. We show that the limit of the quotient of signed areas is the derivative of the ordinary set function (cf. pp. 310–312)

$$M(S) = \iint_S J_{\boldsymbol{\varphi}}(s,t)\,dA,$$

where S is $\vec{S}$ without its orientation.

The inverse function theorem (Theorem 5.2, p. 169) implies there is a smaller open set $\Omega \subseteq U$ containing (a,b) on which $\boldsymbol{\varphi}$ has a continuously differentiable inverse. Because $\delta(\vec{S}) \to 0$ in computing the limit of the quotient of signed areas, it is sufficient to restrict $\vec{S}$ to closed oriented subsets of Ω.

By taking Ω to be pathwise-connected, we can guarantee that the Jacobian $J_{\boldsymbol{\varphi}}(s,t)$ has constant sign on Ω. Hence, because $\vec{S}$ has signed area, the same is true of the image $\boldsymbol{\varphi}(\vec{S})$, and we can write (cf. p. 356)

$$\operatorname{area}\boldsymbol{\varphi}(\vec{S}) = \operatorname{sgn}\boldsymbol{\varphi}(\vec{S})\;A(\boldsymbol{\varphi}(S)),$$

where $\operatorname{area}\vec{D}$ denotes the signed area of the positively or negatively oriented region $\vec{D}$ (p. 356). From Theorem 9.8 and the fact that $\operatorname{sgn} J_{\boldsymbol{\varphi}}$ is well defined on S, we get

$$A(\boldsymbol{\varphi}(S)) = \iint_S |J_{\boldsymbol{\varphi}}(s,t)|\,dA = \operatorname{sgn} J_{\boldsymbol{\varphi}} \iint_S J_{\boldsymbol{\varphi}}(s,t)\,dA = \operatorname{sgn} J_{\boldsymbol{\varphi}}\, M(S).$$

Thus,

$$\operatorname{area} \boldsymbol{\varphi}(\vec{S}) = \operatorname{sgn} \boldsymbol{\varphi}(\vec{S}) \times A(\boldsymbol{\varphi}(S)) = \operatorname{sgn} \boldsymbol{\varphi}(\vec{S}) \times \operatorname{sgn} J_{\boldsymbol{\varphi}} \times M(S) = \operatorname{sgn} \vec{S} \times M(S);$$

$\operatorname{sgn} \boldsymbol{\varphi}(\vec{S})\ \operatorname{sgn} J_{\boldsymbol{\varphi}}\ M(S) = \operatorname{sgn} \vec{S}$ follows from the proof of Theorem 9.14. We also have $\operatorname{area} \vec{S} = \operatorname{sgn} \vec{S} \times A(S)$; thus it follows that

$$\frac{\operatorname{area} \boldsymbol{\varphi}(\vec{S})}{\operatorname{area} \vec{S}} = \frac{M(S)}{A(S)} \to M'(a,b) = J_{\boldsymbol{\varphi}}(a,b)$$

for positively or negatively oriented sets $\vec{S}$ that contain (a,b) and for which $\delta(S) \to 0$ (Theorem 8.39, p. 312). □

The theorem implies that $\operatorname{area} \boldsymbol{\varphi}(\vec{S}) \approx J_{\boldsymbol{\varphi}}(a,b)\ \operatorname{area} \vec{S}$ for any sufficiently small positively or negatively oriented region containing the point (a,b). For this reason we say that the Jacobian $J_{\boldsymbol{\varphi}}(a,b)$ is the *local signed area magnification factor* for the map $\boldsymbol{\varphi}$ at (a,b). $\operatorname{area} \boldsymbol{\varphi}(\vec{S}) \approx J_{\boldsymbol{\varphi}}\ \operatorname{area} \vec{S}$

The change of variables formulas we have established for double integrals and domains in $\mathbb{R}^2$ (Theorem 9.11 and Theorem 9.14) extend naturally to triple integrals and domains in $\mathbb{R}^3$. We state the extensions here with the understanding that they can be proved by adapting the proofs of the 2-dimensional versions. To help underscore the distinction between the oriented and unoriented cases, we use dV as the unoriented element of volume. Changing variables in triple integrals

Theorem 9.16 (Change of variables in $\mathbb{R}^3$). *Let Ω be a bounded open set in $\mathbb{R}^3$, and let $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^3 : (r,s,t) \to (x,y,z)$ be a continuously differentiable map that has a continuously differentiable inverse $\boldsymbol{\varphi}^{-1} : \boldsymbol{\varphi}(\Omega) \to \Omega$. Suppose the function $f(x,y,z)$ is continuous on $D \subset \boldsymbol{\varphi}(\Omega)$, a closed region that has volume; then*

$$\iiint\limits_{D} f(x,y,z)\,dV_{x,y,z} = \iiint\limits_{\boldsymbol{\varphi}^{-1}(D)} f(\boldsymbol{\varphi}(r,s,t)) \left|\frac{\partial(x,y,z)}{\partial(r,s,t)}\right| dV_{r,s,t}. \qquad \square$$

Theorem 9.17 (Oriented change of variables in $\mathbb{R}^3$).
Let Ω be a bounded open set in $\mathbb{R}^3$, and let $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^3 : (r,s,t) \to (x,y,z)$ be a continuously differentiable map that has a continuously differentiable inverse $\boldsymbol{\varphi}^{-1} : \boldsymbol{\varphi}(\Omega) \to \Omega$. Suppose the function $f(x,y,z)$ is continuous on $\vec{D} \subset \boldsymbol{\varphi}(\Omega)$, a closed, oriented, and pathwise-connected region that has volume; then

$$\iiint\limits_{\vec{D}} f(x,y,z)\,dx\,dy\,dz = \iiint\limits_{\boldsymbol{\varphi}^{-1}(\vec{D})} f(\boldsymbol{\varphi}(r,s,t))\,\frac{\partial(x,y,z)}{\partial(r,s,t)}\,dr\,ds\,dt. \qquad \square$$

9.5 Green's theorem

Green's theorem equates the double integral of a certain function over an oriented region in the plane to the path integral of a related expression over that region's boundary (with its induced orientation). Each integral can be used to evaluate the other. We state and prove several increasingly more general versions of Green's theorem, and then use the final version to extend the change of variables formula for oriented double integrals to settings in which the change of variables may not be invertible.

A special case of Green's theorem

The first version of Green's theorem is a special case involving a positively oriented region $\vec{S}$ that can be described in both of the following ways:

$$\vec{S}: \begin{array}{c} a \leq x \leq b, \\ \gamma(x) \leq y \leq \delta(x); \end{array} \qquad \vec{S}: \begin{array}{c} c \leq y \leq d, \\ \alpha(y) \leq x \leq \beta(y); \end{array}$$

we assume $\gamma(x)$ and $\delta(x)$ are continuous functions of x on $[a,b]$, and $\alpha(y)$ and $\beta(y)$ are continuous functions of y on $[c,d]$. The orientation on $\vec{S}$ induces (p. 355) an orientation on $\partial\vec{S}$.

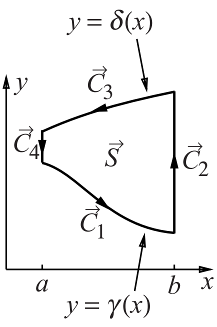
Theorem 9.18 (Green's theorem). *Suppose $P(x,y)$ and $Q(x,y)$ are continuously differentiable functions defined on the closure of the region $\vec{S}$, and $\partial\vec{S}$ has the orientation induced by $\vec{S}$; then*

$$\oint_{\partial\vec{S}} P\,dx + Q\,dy = \iint_{\vec{S}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy.$$

Proof. We assume $\vec{S}$ has positive orientation, and prove half of the equality,

$$\oint_{\partial\vec{S}} P\,dx = \iint_{\vec{S}} -\frac{\partial P}{\partial y}\,dx\,dy,$$

using the first description of $\vec{S}$. The other half, involving Q and $\partial Q/\partial x$, is done in a similar way using the second description of $\vec{S}$.



First consider the path integral. As the figure indicates, we can partition the oriented path $\partial\vec{S}$ into four segments (or fewer: either vertical segment $\vec{C}_2$ or $\vec{C}_4$ may reduce to a point). Neither vertical segment contributes to the path integral, because x is constant and $dx = 0$ there. Consequently,

$$\oint_{\partial\vec{S}} P\,dx = \int_{\vec{C}_1} P\,dx + \int_{\vec{C}_3} P\,dx.$$

On $\vec{C}_1$ and $\vec{C}_3$ we can use x itself as the parameter (but make x "run backwards" from b to a for $\vec{C}_3$); then

$$\int_{\vec{C}_1} P\,dx = \int_a^b P(x,\gamma(x))\;dx,$$

$$\int_{\vec{C}_3} P\,dx = \int_b^a P(x,\delta(x))\;dx = -\int_a^b P(x,\delta(x))\;dx,$$

and hence

$$\oint_{\partial\vec{S}} P\,dx = \int_a^b \big(P(x,\gamma(x)) - P(x,\delta(x))\big)\,dx.$$

That is, the path integral reduces to an ordinary single integral. We now show that the double integral reduces to the same ordinary single integral:

$$\begin{aligned}\iint_{\vec{S}} -\frac{\partial P}{\partial y}\,dx\,dy &= \int_a^b \int_{\gamma(x)}^{\delta(x)} -\frac{\partial P}{\partial y}(x,y)\,dy\;dx \\ &= \int_a^b -P(x,y)\Big|_{\gamma(x)}^{\delta(x)}\;dx = \int_a^b \big(P(x,\gamma(x)) - P(x,\delta(x))\big)\,dx.\end{aligned}$$

This completes half the proof; use a similar argument with Q and with the second description of $\vec{S}$ to prove the other half. □

Below we consider how Green's theorem can be used as a tool for evaluating double integrals. More commonly, though, it is a tool for evaluating path integrals, and we consider this use first.

Evaluating the path integral

Corollary 9.19 *Suppose $P = p(x)$ is a function of x alone, and $Q = q(y)$ a function of y alone; then*

$$\oint_{\partial\vec{S}} p(x)\,dx + q(y)\,dy = 0.$$

Proof. $Q_x - P_y = 0$, so $\iint_{\vec{S}} (Q_x - P_y)\,dx\,dy = 0$. □

Recall that $\Phi(x,y)$ is called a *potential* (cf. p. 25) for the vector field $\mathbf{F}(x,y) = (P(x,y),Q(x,y))$ if $\mathbf{F} = \operatorname{grad}\Phi$; that is, if $P = \partial\Phi/\partial x$, $Q = \partial\Phi/\partial y$.

Potential functions

Corollary 9.20 *Suppose the vector field $(P(x,y),Q(x,y))$ has a potential $\Phi(x,y)$ that has continuous second derivatives on $\vec{S}$; then*

$$\oint_{\partial\vec{S}} P\,dx + Q\,dy = 0.$$

Proof. $Q_x - P_y = \Phi_{yx} - \Phi_{xy} = 0$ on $\vec{S}$ when Φ has continuous second derivatives on $\vec{S}$. □

The second corollary is a generalization of the first when $p(x)$ and $q(y)$ are continuously differentiable, because then we can take

$$\Phi(x,y) = \int p(x)\,dx + \int q(y)\,dy.$$

Evaluating double integrals

Green's theorem can be used to evaluate a double integral by reducing it to a path integral. Specifically, given

$$\iint_{\vec{S}} f(x,y)\,dx\,dy, \quad \text{set } F(x,y) = \int f(x,y)\,dx.$$

That is, F is a "partial integral" of $f(x,y)$ with respect to x, which means only that $\partial F/\partial x = f$. If we now take $P(x,y)=0$ and $Q(x,y)=F(x,y)$, then Green's theorem gives

$$\iint_{\vec{S}} f(x,y)\,dx\,dy = \oint_{\partial\vec{S}} F(x,y)\,dy.$$

We can even write this as

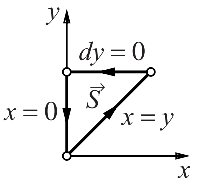
$$\iint_{\vec{S}} f(x,y)\,dx\,dy = \oint_{\partial\vec{S}} \left(\int f(x,y)\,dx\right) dy,$$

a kind of iterated integral in which one of the iterates is a path integral.

Example

To illustrate, let us compute (cf. the example on p. 357)

$$\iint_{\vec{S}} x^2\,dx\,dy$$



where $\vec{S}$ is the positively oriented triangle with vertices $(0,0)$, $(1,1)$, and $(0,1)$. We have

$$\iint_{\vec{S}} x^2\,dx\,dy = \oint_{\partial\vec{S}} \left(\int x^2\,dx\right) dy = \oint_{\partial\vec{S}} \frac{x^3}{3}\,dy.$$

The path $\partial\vec{S}$ has three segments, but the path integral vanishes along two of them: on the top, $dy=0$; on the vertical side, $x=0$. On the diagonal side, we can use $x=y$ as the parameter, so

$$\oint_{\partial\vec{S}} \frac{x^3}{3}\,dy = \int_0^1 \frac{y^3}{3}\,dy = \frac{1}{12}.$$

The indeterminacy of partial integration

Incidentally, when we convert the double integral of $f(x,y)$ over $\vec{S}$ into the path integral of

$$F(x,y) = \int f(x,y)\,dx$$

over $\partial\vec{S}$, the partial integral $F(x,y)$ is determined only up to an additive "constant" with respect to the integration variable x. Such a "constant" is, in fact, an arbitrary function of y. However, this indeterminacy has no effect on the outcome. Adding an arbitrary function of y to the partial integral $x^3/3$ in the last example, we find

$$\iint_{\vec{S}} x^2\,dx\,dy = \oint_{\partial\vec{S}} \left(\int x^2\,dx\right) dy = \oint_{\partial\vec{S}} \left(\frac{x^3}{3} + q(y)\right) dy = \oint_{\partial\vec{S}} \frac{x^3}{3}\,dy,$$

because Corollary 9.19 implies

$$\oint_{\partial\vec{S}} q(y)\,dy = 0.$$

We turn now to the task of extending Green's theorem to more general oriented domains $\vec{S}$. Our first step in this direction is to assume that $\vec{S}$ no longer admits both descriptions that we use in computing iterated integrals, but only one of them. For example, suppose we know only that

Green's theorem for more general regions

$$\vec{S}: \quad \begin{matrix} a \le x \le b, \\ \gamma(x) \le y \le \delta(x). \end{matrix}$$

Then our earlier proof of the equality

$$\oint_{\partial\vec{S}} P\,dx = \iint_{\vec{S}} -\frac{\partial P}{\partial y}\,dx\,dy$$

still holds, but we do need a new proof of the second half of the theorem,

$$\oint_{\partial\vec{S}} Q\,dy = \iint_{\vec{S}} \frac{\partial Q}{\partial x}\,dx\,dy,$$

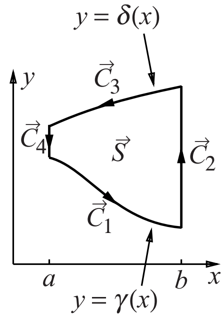
because that depended on the now-absent second description of $\vec{S}$.

To construct a new proof, let F be a "partial integral" of Q with respect to y:

$$F(x,y) = \int Q(x,y)\,dy \quad \text{or} \quad F_y(x,y) = Q(x,y).$$

Because we assume that Q has continuous first derivatives, F has continuous second derivatives, and $Q_x = F_{yx} = F_{xy}$. We can express the double integral of Q_x in terms of F:

$$\begin{aligned}\iint_{\vec{S}} Q_x(x,y)\,dx\,dy &= \int_a^b \int_{\gamma(x)}^{\delta(x)} F_{xy}(x,y)\,dy\,dx = \int_a^b F_x(x,y)\Big|_{\gamma(x)}^{\delta(x)}\,dx \\ &= \int_a^b \big(F_x(x,\delta(x)) - F_x(x,\gamma(x))\big)\,dx.\end{aligned}$$



We now show that the integral of Q over the path $\partial\vec{S}$ reduces to the same expression, making separate calculations on each of the four segments $\vec{C}_1, \vec{C}_2, \vec{C}_3, \vec{C}_4$. On $\vec{C}_1$ we can use x as the parameter with $y = \gamma(x)$ and $a \le x \le b$:

$$\int_{\vec{C}_1} Q\,dy = \int_a^b Q(x,\gamma(x))\,\gamma'(x)\,dx$$

Now consider $F(x,\gamma(x))$; because

$$\frac{d}{dx}F(x,\gamma(x)) = F_x(x,\gamma(x)) + F_y(x,\gamma(x))\,\gamma'(x) = F_x(x,\gamma(x)) + Q(x,\gamma(x))\,\gamma'(x),$$

the chain rule and the defining condition $F_y = Q$ give us

$$\int_a^b Q(x,\gamma(x))\,\gamma'(x)\,dx = \int_a^b \frac{d}{dx}F(x,\gamma(x))\,dx - \int_a^b F_x(x,\gamma(x))\,dx.$$

The first integral on the right equals $F(b,\gamma(b)) - F(a,\gamma(a))$, so

$$\int_{\vec{C}_1} Q\,dy = F(b,\gamma(b)) - F(a,\gamma(a)) - \int_a^b F_x(x,\gamma(x))\,dx.$$

On $\vec{C}_2$, $x = b$ and we can use y as the parameter with $\gamma(b) \le y \le \delta(b)$:

$$\int_{\vec{C}_2} Q\,dy = \int_{\gamma(b)}^{\delta(b)} Q(b,y)\,dy = \int_{\gamma(b)}^{\delta(b)} F_y(b,y)\,dy = F(b,\delta(b)) - F(b,\gamma(b)).$$

On $\vec{C}_3$, we can again take x as the parameter, but now $y = \delta(x)$ and we must integrate with respect to x from b to a:

$$\int_{\vec{C}_3} Q\,dy = \int_b^a Q(x,\delta(x))\,\delta'(x)\,dx = -\int_a^b Q(x,\delta(x))\,\delta'(x)\,dx.$$

Using $F(x,\delta(x))$ and an argument similar to the one for $\vec{C}_1$, we find

$$\int_{\vec{C}_3} Q\,dy = -F(b,\delta(b)) + F(a,\delta(a)) + \int_a^b F_x(x,\delta(x))\,dx.$$

On $\vec{C}_4$, $x = a$ and we can again use y as the parameter, but must now integrate from $\delta(a)$ to $\gamma(a)$:

$$\int_{\vec{C}_4} Q\,dy = \int_{\delta(a)}^{\gamma(a)} Q(a,y)\,dy = \int_{\delta(a)}^{\gamma(a)} F_y(a,y)\,dy = F(a,\gamma(a)) - F(a,\delta(a)).$$

Therefore,

$$\begin{aligned}\oint_{\partial\vec{S}} Q\,dy &= \int_{\vec{C}_1} Q\,dy + \int_{\vec{C}_2} Q\,dy + \int_{\vec{C}_3} Q\,dy + \int_{\vec{C}_4} Q\,dy \\ &= \int_a^b F_x(x,\delta(x))\,dx - \int_a^b F_x(x,\gamma(x))\,dx = \iint_{\vec{S}} Q_x\,dx\,dy\end{aligned}$$

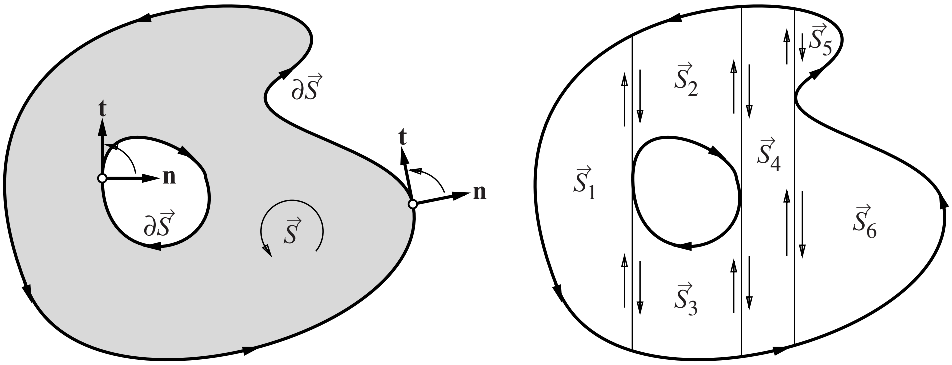
when all the cancellations are taken into account. □

The same arguments, *mutatis mutandis*, allow us to prove Green's theorem for a region $\vec{S}$ when we have only the single description

$$\vec{S}: \quad \begin{array}{c} c \le y \le d, \\ \alpha(y) \le x \le \beta(y). \end{array}$$

Green's theorem on more general domains

Our final version of Green's theorem uses oriented regions $\vec{S}$ that are finite unions of the two kinds we have already considered. In particular, we allow the boundary $\partial\vec{S}$ to have more than one component, although each component will still have the orientation induced by $\vec{S}$. Thus (cf. p. 355), if $\mathbf{n}$ is the outward-pointing unit normal at any smooth point of $\partial\vec{S}$, then the orienting unit tangent $\mathbf{t}$ for $\partial\vec{S}$ at that point is chosen so that the pair $\{\mathbf{n},\mathbf{t}\}$ agrees with the orientation of $\vec{S}$ itself. We assume, as in the figure below, that $\vec{S}$ is closed, bounded, and oriented, and that it can be subdivided into a finite number of nonoverlapping closed cells $\vec{S}_1,\ldots,\vec{S}_N$ with the same orientation as $\vec{S}$. As the figure suggests, this can often be accomplished with properly placed vertical or horizontal lines.



Combining the double integrals

To prove that Green's theorem holds on $\vec{S}$, consider separately the double integral and the path integral. For the double integral we have

$$\iint_{\vec{S}} (Q_x - P_y)\,dx\,dy = \iint_{\vec{S}_1} (Q_x - P_y)\,dx\,dy \;+\cdots+ \iint_{\vec{S}_N} (Q_x - P_y)\,dx\,dy$$

immediately, by Theorem 8.27, page 298.

Combining the path integrals

The path integrals combine in a more interesting way. If two cells $\vec{S}_i$ and $\vec{S}_j$ have a boundary segment $\vec{C}$ in common, then their outward normals point in opposite directions on $\vec{C}$, because $\vec{S}_i$ and $\vec{S}_j$ are on opposite sides of $\vec{C}$. Therefore, the orientation of $\vec{C}$ as part of $\partial\vec{S}_i$ is opposite its orientation as part of $\partial\vec{S}_j$, so the contributions that $\vec{C}$ makes to

$$\oint_{\partial\vec{S}_i} P\,dx + Q\,dy \quad\text{and}\quad \oint_{\partial\vec{S}_j} P\,dx + Q\,dy$$

exactly cancel. The only contributions that do not cancel are from those boundary segments $\vec{C}$ that $\vec{S}_i$ shares with $\vec{S}$ itself. By construction, the orientation of $\vec{C}$ as part of $\partial\vec{S}_i$ is the same as its orientation as part of $\partial\vec{S}$. Therefore, after all the cancellations are taken into account,

$$\oint_{\partial\vec{S}} P\,dx + Q\,dy = \oint_{\partial\vec{S}_1} P\,dx + Q\,dy \;+\cdots+ \oint_{\partial\vec{S}_N} P\,dx + Q\,dy.$$

Thus, because Green's theorem holds for each $\vec{S}_i$, it holds for $\vec{S}$:

$$\iint_{\vec{S}} (Q_x - P_y)\,dx\,dy = \oint_{\partial\vec{S}} P\,dx + Q\,dy. \qquad \square$$

Of course, if the orientation of $\partial\vec{S}$ is *opposite* the orientation induced by the orientation of $\vec{S}$, then

$$\oint_{\partial\vec{S}} P\,dx + Q\,dy = -\iint_{\vec{S}} (Q_x - P_y)\,dx\,dy.$$

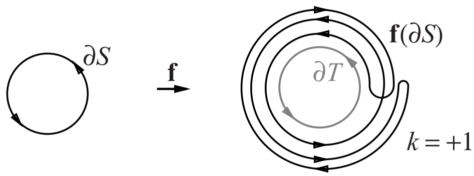
Change of variables via Green's Theorem

Using Green's theorem, we now establish yet another version of the change of variables formula. As with the previous version (Theorem 9.14), this one applies to oriented integrals. For that reason, it uses the Jacobian itself, rather than its absolute value. Unlike the previous version, it does not require the map $\mathbf{f}$ that changes variables to be 1–1. Therefore, because we no longer assume there is a 1–1 correspondence between points of S and points of $T = \mathbf{f}(S)$, we cannot assume that an orientation of S induces (cf. p. 355) an orientation of T.

Thus let S and $T = \mathbf{f}(S)$, and suppose $\vec{S}$ and $\vec{T}$ are independently oriented regions, with Green's Theorem holding on each. Assume $\partial\vec{S}$ and $\partial\vec{T}$ are simple, piecewise-smooth closed curves. The image $\mathbf{f}(\partial\vec{S})$ need not be simple; instead, we assume only that $\mathbf{f}(\partial\vec{S}) \subseteq \partial\vec{T}$ as sets, that $\partial\vec{S}$ and $\mathbf{f}(\partial\vec{S})$ have a common decomposition into smooth oriented curves, and

$$\mathbf{f}(\partial\vec{S}) = k\,\partial\vec{T}$$

as oriented paths. The integer k counts the number of times $\mathbf{f}$ wraps $\partial\vec{S}$ around $\partial\vec{T}$ in the positive direction, minus the number of times it wraps in the negative direction. In the figure below (where the images have been separated for clarity), $k = +1$.



To compensate for the possibility that $\mathbf{f}$ is not 1–1, we require that it now have continuous second derivatives.

Theorem 9.21 (Change of variables with Green's theorem).
Suppose $\mathbf{f}(s,t) = (x(s,t), y(s,t))$ has continuous second derivatives on a bounded open set Ω in $\mathbb{R}^2$. Let $\vec{S} \subset \Omega$ and $\vec{T} = \mathbf{f}(\vec{S})$ be closed, independently oriented sets whose boundaries $\partial\vec{S}$ and $\partial\vec{T}$ are simple closed curves. Assume that Green's theorem holds for both $\vec{S}$ and $\vec{T}$, that $\partial\vec{S}$ and $\mathbf{f}(\partial\vec{S})$ have common decompositions into smooth oriented curves, and that $\mathbf{f}(\partial\vec{S}) = k\,\partial\vec{T}$ as oriented paths. Then, for any continuous function $g(x,y)$ on $\vec{T}$,

$$k\iint_{\vec{T}} g(x,y)\,dx\,dy = \iint_{\vec{S}} g(x(s,t),y(s,t))\,\frac{\partial(x,y)}{\partial(s,t)}\,ds\,dt.$$

Proof. Because Green's theorem holds for the region $\vec{T}$, we can write

$$\iint_{\vec{T}} g(x,y)\,dx\,dy = \oint_{\partial\vec{T}} G(x,y)\,dy,$$

where $G(x,y)$ is some function for which $G_x(x,y) = g(x,y)$ (i.e., a "partial integral"). Because $k\,\partial\vec{T} = \mathbf{f}(\partial\vec{S})$, we have

$$k \iint\limits_{\vec{T}} g(x,y)\,dx\,dy = k \oint\limits_{\partial(\vec{T})} G(x,y)\,dy = \oint\limits_{k\,\partial\vec{T}} G(x,y)\,dy = \oint\limits_{\mathbf{f}(\partial\vec{S})} G(x,y)\,dy.$$

Now apply Exercise 4.37 (p. 149) to $\mathbf{f}$ to transform the last path integral:

$$\oint\limits_{\mathbf{f}(\partial\vec{S})} G(x,y)\,dy = \oint\limits_{\partial\vec{S}} G(x(s,t),y(s,t))(y_s\,ds + y_t\,dt) = \oint\limits_{\partial\vec{S}} G^* y_s\,ds + G^* y_t\,dt.$$

(Here $G^*(s,t) = G(x(s,t),y(s,t))$.) Applying Green's Theorem a second time, we transform this new path integral over $\partial\vec{S}$ back into a double integral, but now one over $\vec{S}$:

$$\oint_{\partial\vec{S}} G^* y_s\,ds + G^* y_t\,dt = \iint_{\vec{S}} \big((G^* y_t)_s - (G^* y_s)_t\big)\,ds\,dt.$$

The terms in the new double integral are

$$(G^* y_t)_s = \frac{\partial}{\partial s}\Big(G(x(s,t),y(s,t)) \cdot y_t(s,t)\Big) = (G_x^* x_s + G_y^* y_s)\,y_t + G^* y_{ts}$$

and

$$(G^* y_s)_t = \frac{\partial}{\partial t}\Big(G(x(s,t),y(s,t)) \cdot y_s(s,t)\Big) = (G_x^* x_t + G_y^* y_t)\,y_s + G^* y_{st}.$$

Therefore,

$$(G^* y_t)_s - (G^* y_s)_t = G_x^*\,(x_s y_t - x_t y_s) + G^*\,(y_{ts} - y_{st}).$$

The second term vanishes because $y_{ts} - y_{st} = 0$; this is where we need the hypothesis that the map $\mathbf{f}$ has continuous second derivatives. Finally, because $G_x^* = g(x(s,t),y(s,t))$ and $x_s y_t - x_t y_s$ is the Jacobian of $\mathbf{f}$, we have

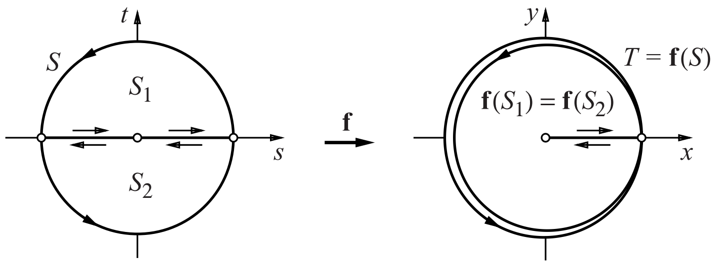
$$\iint_{\vec{S}} \big((G^* y_t)_s - (G^* y_s)_t\big)\,ds\,dt = \iint_{\vec{S}} g(x(s,t),y(s,t))\,\frac{\partial(x,y)}{\partial(s,t)}\,ds\,dt. \qquad \square$$

For our first example of a noninvertible change of variables, we use the quadratic map $\mathbf{f}: \mathbb{R}^2 \to \mathbb{R}^2$, Example 1: the quadratic map

$$\mathbf{f}: \begin{cases} x = s^2 - t^2, \\ y = 2st, \end{cases} \qquad J(s,t) = \frac{\partial(x,y)}{\partial(s,t)} = \begin{vmatrix} 2s & -2t \\ 2t & 2s \end{vmatrix} = 4(s^2 + t^2),$$

that we analyzed in Chapter 4 (cf. pp. 116–121). We saw there that $\mathbf{f}$ squares distances from the origin and doubles polar angles. Away from the origin, $J > 0$ and $\mathbf{f}$ is locally 1–1, by the inverse function theorem.

Let S be the unit disk in the (s,t)-plane; its image $T = \mathbf{f}(S)$ is the unit disk in the (x,y)-plane. Note that S_1, the upper half of S, already covers all of T, and so does the lower half, S_2. The images of the boundaries of S_1 and S_2 are the same; however, in the figure below they have been separated slightly, for clarity.



For the sake of illustration, let $\vec{S}$ have positive orientation and $\vec{T}$ negative. Then the image of $\partial\vec{S}$ wraps twice around $\partial\vec{T}$, but with the opposite orientation:

$$\mathbf{f}(\partial\vec{S}) = -2\,\partial\vec{T}.$$

For any continuous function $g(x,y)$ on $\vec{T}$, Theorem 9.21 asserts that

$$-2\iint_{\vec{T}} g(x,y)\,dx\,dy = \iint_{\vec{S}} g(s^2-t^2, 2st)\;4(s^2+t^2)\,ds\,dt.$$

For instance, if $g(x,y) \equiv 1$, then the assertion reduces to

$$-2\text{ area}\,\vec{T} = -2\iint_{\vec{T}} dx\,dy = \iint_{\vec{S}} 4(s^2+t^2)\,ds\,dt.$$

To verify this, note that $\vec{T}$ is a negatively oriented unit disk, so area $\vec{T} = -\pi$ and the left-hand side is $+2\pi$. The right-hand side has the same value, as we can see by making a change to polar coordinates:

$$\iint_{\vec{S}} 4(s^2+t^2)\,ds\,dt = \int_0^{2\pi} d\theta \int_0^1 4r^2\,r\,dr = 2\pi r^4\Big|_0^1 = 2\pi.$$

For a second instance, take $g(x,y) = x^2$; then we must verify that

$$-2\iint_{\vec{T}} x^2\,dx\,dy = \iint_{\vec{S}} (s^2-t^2)^2\,4(s^2+t^2)\,ds\,dt.$$

Because $\vec{T}$ is negatively oriented, we have

$$-2\iint_{\vec{T}} x^2\,dx\,dy = +2\iint_T x^2\,dA = 2\int_0^{2\pi}\int_0^1 r^2\cos^2\theta\; r\,dr\,d\theta$$
$$= 2\int_0^{2\pi}\cos^2\theta\,d\theta\int_0^1 r^3\,dr = 2\cdot\pi\cdot\frac{1}{4} = \frac{\pi}{2}.$$

The integral over $\vec{S}$ can also be evaluated by a change to polar coordinates in which $s = r\cos\theta$, $t = r\sin\theta$. Because

$$(s^2-t^2)^2 = (r^2\cos^2\theta - r^2\sin^2\theta)^2 = r^4\cos^2 2\theta,$$

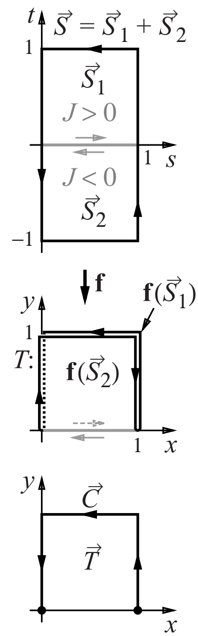
we find

$$\iint_{\vec{S}}(s^2-t^2)^2\,4(s^2+t^2)\,ds\,dt = \int_0^{2\pi}\cos^2 2\theta\int_0^1 4r^7\,dr = \pi\times\frac{1}{2} = \frac{\pi}{2}.$$

Example 2: a fold

For our second example of a change of variables that transforms integrals using Green's theorem, we take

$$\mathbf{f}:\begin{cases} x = s, \\ y = t^2; \end{cases} \qquad J(s,t) = \frac{\partial(x,y)}{\partial(s,t)} = \begin{vmatrix} 1 & 0 \\ 0 & 2t \end{vmatrix} = 2t.$$



The Jacobian $J(s,t)$ is positive in the upper half-plane and negative in the lower; it changes sign on the s-axis. The map $\mathbf{f}$ is a *fold* (cf. Exercise 4.21, p. 146). It folds the (s,t)-plane along the s-axis; points that are symmetrically placed across the s-axis have the same image.

Let $\vec{S}$ be the rectangle $0 \le s \le 1$, $-1 \le t \le 1$ with positive orientation. Where J changes sign, split $\vec{S}$ into two positively oriented nonoverlapping cells $\vec{S}_1$ ($t \ge 0$) and $\vec{S}_2$ ($t \le 0$), so that we can write $\vec{S} = \vec{S}_1 + \vec{S}_2$. The image $T = \mathbf{f}(S)$ is the unit square $0 \le x \le 1$, $0 \le y \le 1$; If we make $\vec{T}$ positively oriented, then

$$\vec{T} = \mathbf{f}(\vec{S}_1) = -\mathbf{f}(\vec{S}_2).$$

Notice that the image of the boundary, $\mathbf{f}(\partial\vec{S})$, is a proper subset of the boundary of $\vec{T}$; it does not "wrap around" $\partial\vec{T}$. In fact, we can write $\mathbf{f}(\partial\vec{S})$ as $\vec{C} - \vec{C}$, where $\vec{C}$ goes around three sides of $\vec{T}$. This implies

$$\mathbf{f}(\partial\vec{S}) = 0\times\partial\vec{T}$$

as an oriented path. Therefore, by Theorem 9.21,

$$\iint_{\vec{S}} g(s,t^2)\,2t\,ds\,dt = 0\times\iint_{\vec{T}} g(x,y)\,dx\,dy = 0$$

for any continuous function $g(x,y)$ defined on the unit square $\vec{T}$.

To see, from another perspective, why the integral of any function of the form $G(s,t) = 2t\,g(s,t^2)$ over $\vec{S}$ must automatically equal zero, we note two things. First, $G(s,t)$ is an odd function of t (i.e., $G(s,-t) = -G(s,t)$). Second, the region $\vec{S}$ is symmetric across the s-axis: the point (s,t) is $\vec{S}$ if and only if $(s,-t)$ is. Therefore, if we write

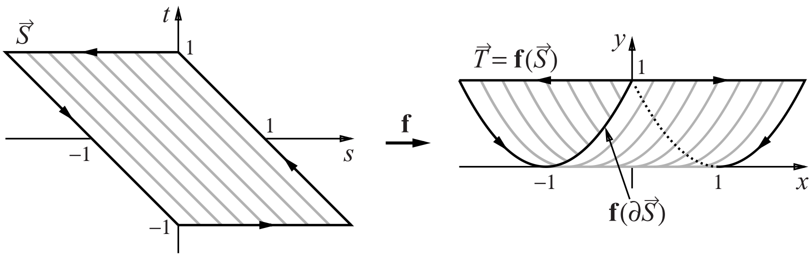
$$\iint_{\vec{S}} g(s,t^2)\,2t\,ds\,dt = \int_0^1 ds \int_{-1}^1 G(s,t)\,dt,$$

then we see that we must integrate an odd function of t over a t-interval that is symmetric about the origin. The value of such an integral is always zero (Exercise 9.26).

Example 3: folding a different region

Our third example again uses the fold map, but applies it to a region on which the boundary condition of Theorem 9.21 fails. The region is the parallelogram S shown below. Because parts of $\mathbf{f}(\partial S)$ lie in the interior of T, $\mathbf{f}(\partial S) \not\subseteq \partial T$. Therefore, no matter how $\vec{S}$ and $\vec{T}$ are oriented, no integer k can be found for which

$$\mathbf{f}(\partial\vec{S}) = k\,\partial\vec{T}.$$



Example 4: a pleat

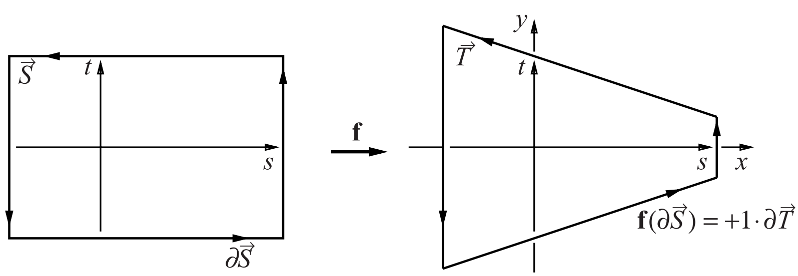
Our fourth example is the map

$$\mathbf{f}: \begin{cases} x = s, \\ y = t^3 - \frac{1}{3}st; \end{cases} \qquad J(s,t) = \frac{\partial(x,y)}{\partial(s,t)} = \begin{vmatrix} 1 & 0 \\ -\frac{1}{3}t & 3t^2 - \frac{1}{3}s \end{vmatrix} = 3t^2 - \tfrac{1}{3}s.$$

For reasons that soon become clear, this is called a **pleat**. Let $\vec{S}$ be the positively oriented rectangle $-1 \le s \le 2$, $-1 \le t \le 1$. Note that $\mathbf{f}$ preserves vertical lines, because $x = c$ when $s = c$. Horizontal lines are not preserved, but each is mapped to some straight line. Specifically, the horizontal line $t = k$ is mapped to the line $y = k^3 - kx/3$ (using $s = x$) with slope $-k/3$ and y-intercept k^3. Therefore, the image of $\vec{S}$ is the trapezoid

$$\vec{T}: \quad \begin{array}{c} -1 \le x \le 2, \\ \frac{1}{3}x - 1 \le y \le -\frac{1}{3}x + 1 \end{array}$$

that we define to be positively oriented. In that case, $\mathbf{f}(\partial\vec{S}) = +1 \times \partial\vec{T}$.



Let us compute the area of the trapezoid $\vec{T}$ first using a double integral provided by the change of variables formula, and then by elementary geometry. The integral gives

Area given by a double integral

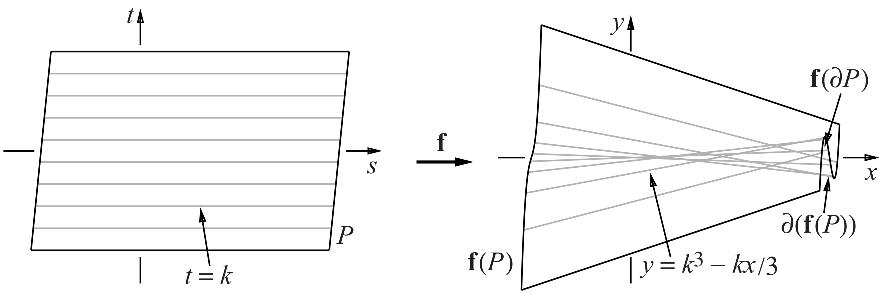
$$\begin{aligned}
\operatorname{area}\vec{T} &= +1\times\iint_{\vec{T}} dx\,dy = \iint_{\vec{S}}\left(3t^2-\tfrac{1}{3}s\right)ds\,dt\\
&= \int_{-1}^{2}\int_{-1}^{1}\left(3t^2-\tfrac{1}{3}s\right)dt\;ds = \int_{-1}^{2} t^3-\tfrac{1}{3}st\Big|_{-1}^{1}\;ds\\
&= \int_{-1}^{2}\left(2-\tfrac{2}{3}s\right)ds = 2s-\tfrac{1}{3}s^2\Big|_{-1}^{2} = 4-\tfrac{4}{3}-\left(-2-\tfrac{1}{3}\right)=5.
\end{aligned}$$

As a figure in elementary geometry, $\vec{T}$ has "height" $H=3$ and "bases" $B_1 = 2\frac{2}{3}$, $B_2=\frac{2}{3}$; therefore, we find once again that

$$\operatorname{area}\vec{T} = H\left(\frac{B_1+B_2}{2}\right)=5.$$

The map **f** has subtleties that, among other things, make the validity of the transformation of integrals more surprising than we might at first imagine. The family of horizontal lines in the figure below shows clearly that **f** makes a "pleat" in the target. The pleat is made up of two folds that come together at the origin of the target. Inside the pleat, each target point is the image of three points in the source; outside, only one.

f makes a pleat



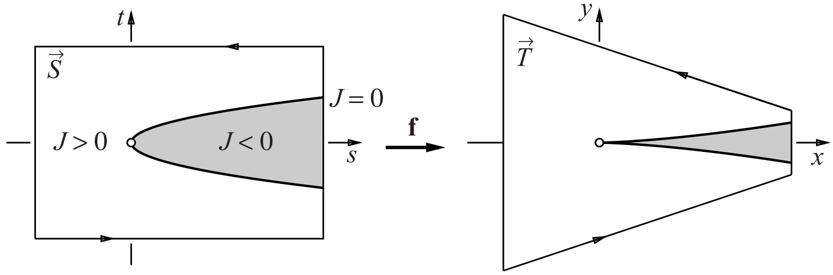
To make the pleating more apparent, we have sheared the rectangle into a parallelogram P. This makes the image of the right edge of P follow a cubic curve that shows how the pleat is folded. As a result, $\mathbf{f}(\partial P)\not\subseteq\partial(\mathbf{f}(P))$: part of that cubic curve lies in the interior of $\mathbf{f}(P)$, and the image of part of the interior of P lies on $\partial(\mathbf{f}(P))$. Although the boundary condition of Theorem 9.21 holds on S, it does not hold on P, even though the shear that changes S into P can be as slight as we wish.

The source is folded to make the pleat. Given any point in the source near that fold, there is another point (on the other side of the fold) that has the same image. In other words, **f** is never locally 1–1 at a fold point. But, by the inverse function theorem, **f** is locally 1–1 near any point (s,t) where $J(s,t)\neq 0$. Therefore, the fold points of **f** must occur where $J=0$, on the parabola $s=9t^2$. The pleat is the image

The fold locus and its image

of this parabola; that is, the pleat is the set of points $\mathbf{f}(9t^2,t)$. It therefore has the parametric equations

$$x = 9t^2,$$
$$y = t^3 - \tfrac{1}{3}9t^3 = -2t^3.$$

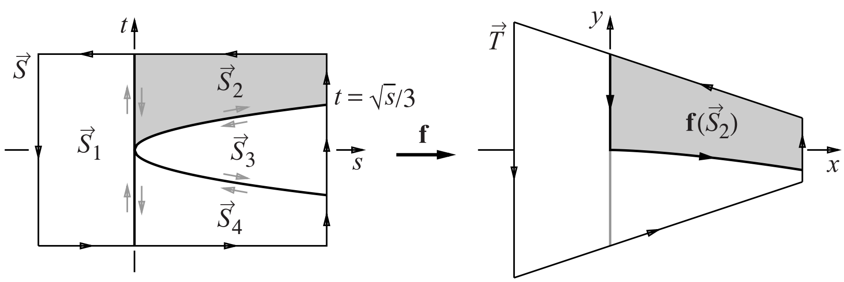


The pleat makes a cusp at the origin. We can see this afresh by setting $t = \sqrt[3]{y/-2}$, so that

$$x = 9\left(\sqrt[3]{y/-2}\right)^2 = \frac{9}{\sqrt[3]{4}}y^{2/3}.$$

This is a function whose graph is a cusp.

As we have seen, $\mathbf{f}$ is a threefold cover of the inside of the cusp. It would appear that calculating the area of the trapezoid $\mathbf{f}(S)$ by integrating over S would count the area inside the cusp three times, instead of just once. But note that $\mathbf{f}$ reverses the orientation of one of the three “sheets” that cover the inside of the cusp (namely, the inside of the parabola). Let us now see how this leads to the correct outcome.



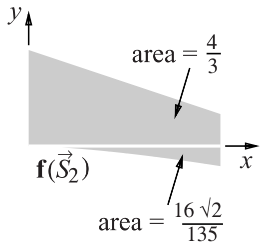
We partition S into the four nonoverlapping, positively oriented regions shown above, so that $\vec{S} = \vec{S}_1 + \vec{S}_2 + \vec{S}_3 + \vec{S}_4$. On each region, $\mathbf{f}$ is 1–1. The change of variables formula implies that the area of $\vec{T}$ is

$$\iint_{\vec{S}} J\,ds\,dt = \iint_{\vec{S}_1} J\,ds\,dt + \iint_{\vec{S}_2} J\,ds\,dt + \iint_{\vec{S}_3} J\,ds\,dt + \iint_{\vec{S}_4} J\,ds\,dt,$$

where $J = 3t^2 - \frac{1}{3}s$ in every case. The value of the first integral on the right is $2\frac{1}{3}$, the area of the trapezoid $\mathbf{f}(\vec{S}_1)$. For the second integral, we have

$$\iint_{\vec{S}_2} \left(3t^2 - \tfrac{1}{3}s\right) ds\,dt = \int_0^2 \int_{\sqrt{s}/3}^1 \left(3t^2 - \tfrac{1}{3}s\right) dt\; ds$$
$$= \int_0^2 t^3 - \tfrac{1}{3}st \Big|_{\sqrt{s}/3}^{1} \; ds = \int_0^2 \left(1 - \tfrac{1}{3}s + \tfrac{2}{27}s^{3/2}\right) ds$$
$$= s - \tfrac{1}{6}s^2 + \tfrac{2\cdot 2}{27\cdot 5}s^{5/2} \Big|_0^2 = 2 - \tfrac{4}{6} + \tfrac{16\sqrt{2}}{135} = \tfrac{4}{3} + \tfrac{16\sqrt{2}}{135}.$$

The part of $\mathbf{f}(\vec{S}_2)$ that lies in the first quadrant is a trapezoid whose area is $4/3$; the remaining part, the curvilinear triangle in the fourth quadrant, must therefore account for the remaining area, namely $16\sqrt{2}/135 \approx 0.17$. By symmetry, the fourth integral has the same value, and breaks down in the same way:



$$\iint_{\vec{S}_4} \left(3t^2 - \tfrac{1}{3}s\right) ds\,dt = \tfrac{4}{3} + \tfrac{16\sqrt{2}}{135}.$$

We see that $\mathbf{f}(\vec{S}_1)$ and the trapezoidal parts of $\mathbf{f}(\vec{S}_2)$ and $\mathbf{f}(\vec{S}_4)$ already completely cover $\vec{T}$, and their total area is $2\frac{1}{3} + 2\cdot\frac{4}{3} = 5$, the area of $\vec{T}$. The integrals over $\vec{S}_2$ and $\vec{S}_4$ therefore produce an excess of $+32\sqrt{2}/135$. With this in mind, consider the third integral:

$$\iint_{\vec{S}_3} \left(3t^2 - \tfrac{1}{3}s\right) ds\,dt = \int_0^2 \int_{-\sqrt{s}/3}^{\sqrt{s}/3} \left(3t^2 - \tfrac{1}{3}s\right) dt\; ds$$
$$= \int_0^2 t^3 - \tfrac{1}{3}st \Big|_{-\sqrt{s}/3}^{\sqrt{s}/3} \; ds = \int_0^2 -\tfrac{4}{27}s^{3/2}\, ds = -\tfrac{8}{135}2^{5/2}$$
$$= -\tfrac{32\sqrt{2}}{135}.$$

This is negative and exactly cancels the excess contributions made by $\vec{S}_2$ and $\vec{S}_4$. The sum of the four oriented integrals is 5.

Notice that although $\mathbf{f}(\partial\vec{S}) = \partial\vec{T}$ as sets and as oriented paths in Example 4, $\mathbf{f}$ is not a 1–1 map of $\partial\vec{S}$ to $\partial\vec{T}$: the image of ∂S "doubles back" on itself briefly inside the cusp. However, all we need (for the proof of the change of variables theorem to hold) is that $\mathbf{f}(\partial\vec{S})$ make a single traversal of $\vec{T}$ in the sense induced by $\vec{T}$, because then **f need not be 1–1**

$$\oint_{\mathbf{f}(\partial\vec{S})} P(x,y)\,dx + Q(x,y)\,dy = \oint_{\partial\vec{T}} P(x,y)\,dx + Q(x,y)\,dy$$

for any continuous integrands $P(x,y)$ and $Q(x,y)$.

Exercises

9.1. Use the pullback substitution $y = \sqrt{x^2+a^2}\tan\theta$ (here y is a function of θ; x is fixed) to show that

$$\int_0^R \frac{dy}{(x^2+y^2+a^2)^{3/2}} = \frac{R}{(x^2+a^2)(x^2+R^2+a^2)^{1/2}}.$$

9.2. Show that

$$\frac{1}{\sqrt{1+(a/2R)^2}} = 1 + O\big((a/R)^2\big) \quad \text{as } a/R \to 0$$

and then provide the details to show that

$$\frac{R^2}{a\sqrt{2R^2+a^2}} = \frac{R}{a\sqrt{2}} + O(a/R) \quad \text{as } a/R \to 0.$$

9.3. Determine the value of each of the following iterated integrals. (Note the different orders "$dy\,dx$" and "$dx\,dy$". Which integrals are equal and which are not? Is that what you expect?

a. $\displaystyle\int_0^3 \int_1^5 (x^2+xy^3)\,dy\,dx = \int_0^3 \left(\int_1^5 (x^2+xy^3)\,dy\right)dx$

b. $\displaystyle\int_1^5 \int_0^3 (x^2+xy^3)\,dx\,dy$

c. $\displaystyle\int_0^3 \int_1^5 (x^2+xy^3)\,dx\,dy$

d. $\displaystyle\int_1^5 \int_0^3 (x^2+xy^3)\,dy\,dx$

9.4. Evaluate:

a. $\displaystyle\int_{-1}^1 \int_{y-5}^{2y+1} xy\,dx\,dy$

b. $\displaystyle\int_0^4 \int_{x/2}^{\sqrt{x}} (y^3+x^2y)\,dy\,dx$

c. $\displaystyle\int_0^1 \int_{x^2}^{\sqrt{x}} 1\,dy\,dx$

d. $\displaystyle\int_0^1 \int_{y^2}^{\sqrt{y}} dx\,dy$

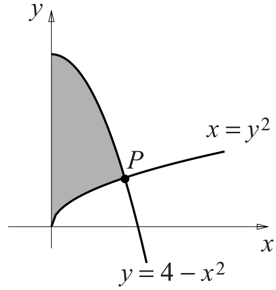
9.5. Sketch, in the (x,y)-plane, the domain of integration of each of integrals in Exercises 9.3 and 9.4.

9.6. Sketch each of the following regions in the (x,y)-plane and then describe it in the form

$$a \le x \le b, \quad \gamma(x) \le y \le \delta(x).$$

a. The unit circle (the circle of radius 1 centered at the origin).
b. The circle of radius 3 centered at $(5,-1)$.
c. The circle of radius R centered at (p,q).
d. The bounded region that lies between the graphs of $y = x^2$ and $y = 4$.

e. The bounded region in the first quadrant that lies between the graphs $y = x^2$ and $x = y^2$.
f. The triangle with vertices at $(0,0)$, $(5,5)$, and $(0,5)$.
g. The diamond-shaped (or lozenge-shaped) region whose vertices are at the points $(1,0)$, $(0,1)$, $(-1,0)$, and $(0,-1)$.
h. The region shown shaded in the margin. Write P as (p,q); you need not determine the values of p or q.



9.7. Describe each of the following regions in the form

$$c \le y \le d, \quad \alpha(y) \le x \le \beta(y).$$

a. The circle of radius R centered at $(x,y) = (p,q)$.
b. The bounded region that lies between the graphs of $y = x^2$ and $x = y^2$.
c. The triangle with vertices at $(0,0)$, $(5,5)$, and $(0,5)$.
d. The triangle with vertices at $(0,0)$, $(5,5)$, and $(5,0)$.
e. The region where $0 \le x \le 2$, $0 \le y \le x$. Sketch this!
f. The region where $0 \le x \le 2$, $x^3 \le y \le 10 - x$.

9.8. Reverse the order of integration in the following integral:

$$\int_0^2 \int_{x^3}^{10-x} f(x,y)\,dy\,dx.$$

That is, rewrite it an iterated integral with the integration done first with respect to x, and then with respect to y (i.e., "$dx\,dy$" instead of "$dy\,dx$").

9.9. By reversing the order of integration, show that

$$\int_0^1 \int_y^1 e^{x^2}\,dx\,dy = \frac{e-1}{2}.$$

What happens when you try to determine the integral directly, without reversing the order of integration?

9.10. In each of the following, reverse the order of integration and then calculate the new integral.

a. $\displaystyle\int_0^2 \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} dy\,dx$

b. $\displaystyle\int_0^1 \int_{x^2}^{\sqrt{x}} dy\,dx$

c. $\displaystyle\int_0^1 \int_x^{\sqrt[3]{x}} y^2\,dy\,dx$

d. $\displaystyle\int_0^4 \int_{x/2}^{\sqrt{x}} (y^3 + x^2 y)\,dy\,dx$

9.11. Evaluate the given double integral using iterated integrals.

a. $\displaystyle\iint_R (x^2 + xy^3)\,dA$, $\quad R$: rectangle with vertices $(0,1)$, $(3,1)$, $(3,5)$, $(0,5)$

b. $\iint_D xy\,dA$, D: bounded region between the graphs $y = x^2$ and $x = y^2$

c. $\displaystyle\iint\limits_{\substack{-1\le x\le 1\\ 0\le y\le 3}} y^2\,dA$

d. $\displaystyle\iint\limits_{\substack{0\le y\le 2\\ y\le x\le 4-y}} x\,dA$

e. $\iint_K x\,dA$, K : disk of radius 2 centered at $(0,0)$

f. $\iint_K x\,dA$, K : disk of radius r centered at (p,q)

g. $\iint_T xy\,dA$, T : triangle with vertices $(0,0)$, $(4,4)$, $(0,4)$

9.12. Express, as a double integral, the volume of the solid bounded by the planes $x = 0, y = 0, z = 0$, and $x+y+z = 4$. Then determine the volume by evaluating the integral.

9.13. The double integral

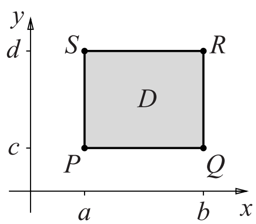
$$\iint\limits_{(x-p)^2+(y-q)^2\le R^2} H\,dA, \qquad H \text{ constant},$$

is the volume of a familiar solid shape. Describe the shape quite precisely and use that knowledge (rather than an iterated integral) to determine the value of the integral.

9.14. The double integral

$$\iint\limits_{x^2+y^2\le R^2} \sqrt{R^2-x^2-y^2}\,dA$$

is the volume of a familiar solid shape. What is the shape? Use that knowledge (rather that an iterated integral) to determine the value of the integral.



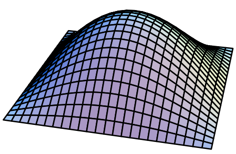
9.15. Show that

$$\iint_D \frac{\partial^2 f}{\partial x\,\partial y}\,dA = f(P) - f(Q) + f(R) - f(S),$$

where $P = (a,c)$, $Q = (b,c)$, $R = (b,d)$, and $S = (a,d)$ are the vertices of the rectangle D.

9.16. In Exercise 9.15, take $f(x,y) = xy$ and $a = c = 0$. Then show, using separate observations or arguments, that the expression $f(P) - f(Q) + f(R) - f(S)$ and the double integral both equal the area of D.

$z = \sin(x)\,\sin(y)$



9.17. The volume of the “ravioli” $z = \sin x \sin y$ shown in the margin is

$$\int_0^\pi \int_0^\pi \sin x \sin y\,dy\,dx.$$

Evaluate this integral using the result in Exercise 9.15. What is $f(x,y)$ in this case?

9.18. a. What is the mean, or average, value $\overline{z}$ (cf. Definition 3.1, p. 75) of the function $ax+by+c$ on the circle $x^2+y^2 \leq R^2$. Does the value depend on the coefficients a or b?

b. The double integral

$$\iint\limits_{x^2+y^2\leq R^2} (ax+by+c)\,dA$$

is the volume of a certain solid shape. Describe the shape and make a general sketch (take $c > R > 0$). Determine the volume directly from your knowledge of $\overline{z}$, without calculating the integral again.

c. Without calculating the following integral, explain why

$$\iint\limits_{x^2+y^2\leq R^2} (ax+by)\,dA = 0,$$

for $R > 0$ and for any a and b.

9.19. a. (This concerns a function of a single variable, and it appears here to provide a comparison with the result of the previous exercise.) What is the average value $\overline{y}$ of $y = mx+b$ on the interval $-R \leq x \leq R$?

b. Does $\overline{y}$ depend on m? Draw a graph of $y = mx+b$ that explains your answer. (To make a concrete graph, try $m = 1/2$ and $b = 3$.)

c. Your work on this and the previous exercise should now allow you to make a general statement about the average value of a linear function (of either one or two variables) on a domain that is symmetric with respect to the origin. What is the statement?

9.20. a. Show that the average value $\overline{z}$ of $z = \sqrt{R^2-x^2-y^2}$ on the disk $x^2+y^2 \leq R^2$ is $2R/3$.

b. Sketch the graph of $z = \sqrt{R^2-x^2-y^2}$ and describe it in words.

c. On the same axes, sketch the horizontal plane $z = \overline{z}$. This defines a cylinder over the disk $x^2+y^2 \leq R^2$ and the volume of this cylinder should equal the volume of the solid $z \leq \sqrt{R^2-x^2-y^2}$ over the same disk. Why? Do the volumes appear equal, or nearly so, in your sketch? (This fact was discovered by the Greek mathematician Archimedes (*c*. 287–212 B.C.E.); it implies that the volume of a ball of radius R is $4\pi R^3/3$.)

9.21. (Here is another single-variable problem that is provides a comparison with an analogous result for a function of two variables.)

a. Show that the average value $\overline{y}$ of $y = \sqrt{R^2-x^2}$ on the line $-R \leq x \leq R$ is $\pi R/4$.

b. Sketch the graph of $y = \sqrt{R^2-x^2}$ and describe it in words. On the same plane sketch the graph $y = \overline{y}$.

c. According to the definition of $\overline{y}$, the area of the rectangle under the horizontal line $y = \overline{y}$ should equal the area under the graph of $y = \sqrt{R^2 - x^2}$; why? Does your graph show this?

9.22. a. Show that $\displaystyle\iint_{\varepsilon^2 \le x^2+y^2 \le 1} \ln\sqrt{x^2+y^2}\, dA = -\pi(1-\varepsilon^2+2\varepsilon^2\ln\varepsilon)/2.$

b. Show the improper integral $\displaystyle\iint_{x^2+y^2\le 1} \ln\sqrt{x^2+y^2}\, dA$ converges to $-\pi/2$.

9.23. Use the pullback $x = \cos s$ to show $\displaystyle\int_{1/2}^{1} \frac{dx}{\sqrt{x^2-x^4}} = \ln(2+\sqrt{3})$.

9.24. The aim here is analyze the focal points of the hyperbola $x^2/a^2 - y^2/b^2 = 1$ and the ellipse $x^2/a^2 + y^2/b^2 = 1$.

a. By definition, a hyperbola is the locus of points for which the difference of the distances to two fixed points (its *focal points*) is a constant. Show that the focal points of the hyperbola are $\mathbf{p}_\pm = (\pm\sqrt{a^2+b^2}, 0)$ in the following way: Parametrize the part of the hyperbola for which $x > 0$ as $\mathbf{x} = (x,y) = (a\cosh t, b\sinh t)$, and set $D_\pm = \|\mathbf{x} - \mathbf{p}_\pm\|$. Show by direct computation that $D_\pm = \sqrt{a^2+b^2}\cosh t \mp a$ and thus that $D_- - D_+ = 2a$.
b. Conclude that the hyperbolas $x^2/\sin^2 s - y^2/\cos^2 s = 1$ (with s arbitrary) are confocal with focal points $(\pm 1, 0)$.
c. By definition, an ellipse is the locus of points for which the sum of the distances to the two focal points is a constant. Show that, when $a > b > 0$, the focal points of the ellipse are $\mathbf{p}_\pm = (\pm\sqrt{a^2-b^2}, 0)$. Adapt the aproach you took for the hyperbola.
d. Conclude that the ellipses $x^2/\cosh^2 s + y^2/\sinh^2 s = 1$ (with s arbitrary) are confocal with focal points $(\pm 1, 0)$.

9.25. Use $\cos^2 s + \sinh^2 t = \frac{1}{2}(\cos 2s + \cosh 2t)$ (for example) to compute

$$\int_a^b \int_c^d (\cos^2 s + \sinh^2 t)\, ds\, dt.$$

9.26. Suppose $f(t)$ is an odd function ($f(-t) = -f(t)$) that is integrable on the interval $-a \le t \le a$. Use the change of variable $t = -s$ to show

$$\int_{-a}^{0} f(t)\, dt = \int_{a}^{0} f(s)\, ds \quad \text{and hence} \quad \int_{-a}^{a} f(t)\, dt = 0.$$

9.27. The aim is to show that the map $\boldsymbol{\varphi}^{-1}$ is invertible on the half plane $y > -1/2$, where

$$\boldsymbol{\varphi}^{-1} : \begin{cases} s = 1 + x + y^2, \\ t = x - y. \end{cases}$$

a. Show that the image of the line $y=a$ under $\boldsymbol{\varphi}^{-1}$ is the line $t=s+b$, where $b=-1-a-a^2$. Show that, for any two values of $a<-1/2$, the image lines are different.

b. Show that $\boldsymbol{\varphi}^{-1}$ is 1–1 on each line $y=a$. Conclude $\boldsymbol{\varphi}^{-1}$ is 1–1 on the entire half plane $y>-1/2$.

9.28. Show that $\iiint_{\mathbf{s}(D)} f(x,y,z)\,dx\,dy\,dz =$

$$\iiint_D f(\rho\cos\theta\cos\varphi, \rho\sin\theta\cos\varphi, \rho\sin\varphi)\,\rho^2\cos\varphi\;d\rho\,d\theta\,d\varphi$$

is the change of variables formula for triple integrals under the spherical coordinate map $\mathbf{s}$ (Exercise 5.10, p. 178).

9.29. Determine the change of variables formula for fourfold integrals under the map σ (Exercise 5.25, p. 183) that is the analogue in $\mathbb{R}^4$ of the spherical coordinate change in $\mathbb{R}^3$.

9.30. Compute both the path integral and the double integral of Green's theorem for $P=xy$, $Q=y$, and R the unit square in the (x,y)-plane.

9.31. a. Use the result of Exercise 9.18.c to show that

$$\oint\limits_{x^2+y^2=R^2} f(x)\,dx + (ax^2+bxy+cy^2)\,dy = 0$$

for any function $f(x)$ and any values of a, b, and c.

b. Use the same idea to explain why

$$\oint\limits_{x^2+y^2=R^2} f(x)\,dx + (ax^2+bxy+cy^2+\alpha x+\beta y+\gamma)\,dy = \alpha\pi R^2,$$

when $x^2+y^2=R^2$ has positive orientation and $f(x)$, a, b, c, α, β, and γ are arbitrary.

9.32. Use Green's theorem to evaluate each of the following path integrals.

a. $\oint_C 5y\,dx + 2x\,dy$, $\quad C$: triangle with vertices $(1,5)$, $(9,2)$, $(8,8)$.

b. $\oint_{\vec{C}} (x^2-x^3)\,dx + (x^3+y^2)\,dy$, $\quad \vec{C}$: counterclockwise unit circle.

c. $\oint_{\vec{C}} ye^x\,dx + xe^y\,dy$, $\quad \vec{C}$: rectangle with vertices $(-1,1)$, $(7,1)$, $(7,5)$, $(-1,5)$.

9.33. Show that path integral $\oint_{\partial\vec{R}} x\,dy$ equals the area of the oriented region $\vec{R}$.

9.34. Show that the path integrals $\oint_{\partial \vec{R}} -y\,dx$ and $\frac{1}{2}\oint_{\partial \vec{R}} x\,dy - y\,dx$ both *also* equal the area of $\vec{R}$.

9.35. Let $\vec{D}$ be the elliptical "disk" $x^2/a^2 + y^2/b^2 \le 1$ with positive orientation.

a. Sketch $\vec{D}$ when $a = 5$, $b = 3$.
b. Find parametric equations $x = x(t)$, $y = y(t)$ for the boundary ellipse $\partial\vec{D}$.
c. Use $\oint_{\partial \vec{D}} x\,dy$ and the parametrizations of $\partial\vec{D}$ to show that $\text{area}\,D = \pi ab$. Note that if $b = a$ then $\vec{D}$ is a *circular* disk with radius $a = b$ and area πa^2.

9.36. Suppose that $H(x,y)$ is a **harmonic function**; that is, H satisfies the Laplace equation:

$$\frac{\partial^2 H}{\partial x^2} + \frac{\partial^2 H}{\partial y^2} = 0.$$

Show that $\oint_{\vec{C}} \frac{\partial H}{\partial y}\,dx - \frac{\partial H}{\partial x}\,dy = 0$ for any closed curve $\vec{C}$.

9.37. Show that, under the maps

$$\mathbf{q} : \begin{cases} x = s^3 - 3st^2, \\ y = 3s^2t - t^3, \end{cases} \qquad \mathbf{s} : \begin{cases} x = s^4 - 6s^2t^2 + t^4, \\ y = 4s^3t - 4st^3, \end{cases}$$

the positively oriented unit disk $\vec{S}$ covers the positively oriented unit disk $\vec{D}$ three and four times, respectively, and the analogous integrals

$$\iint_{\vec{S}} J_{\mathbf{q}}(s,t)\,ds\,dt \quad\text{and}\quad \iint_{\vec{S}} J_{\mathbf{s}}(s,t)\,ds\,dt$$

(where $J_{\mathbf{q}}$ and $J_{\mathbf{s}}$ are the Jacobians of $\mathbf{q}$ and $\mathbf{s}$) have the values 3π and 4π.

9.38. Let $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2 : (x,y) \to (u,v)$ be the quadratic map (p. 340), and let the **arrowhead** $A = \mathbf{g}(S)$ be the image of the square S: $0.2 \le x \le 1$, $0.2 \le y \le 1$.

a. Sketch A in the (u,v)-plane.
b. Show that $\partial(x,y)/\partial(u,v) = 1/4\sqrt{u^2+v^2}$.
c. Determine $\text{area}\,A = \iint_A du\,dv = \iint_S 4(x^2+y^2)\,dx\,dy = 3968/1875$; cf. Exercise 8.3, p. 313.
d. Show that $\iint_A \frac{du\,dv}{\sqrt{u^+v^2}} = \iint_S 4\;dx\,dy = 4\,\text{area}\,S = 2.56$.
e. Determine the moments (cf. Exercise 8.22, p. 316) of the *arrowhead* around the v- and u-axes:

$$\iint_A u\;du\,dv, \qquad \iint_A v\;du\,dv.$$

f. Determine $\displaystyle\iint_A \frac{du\,dv}{v}$ and $\displaystyle\iint_A \frac{du\,dv}{u^2+v^2}$.

9.39. Let D be the quarter-disk $0 \le x^2 + y^2 \le 1$, $0 \le x$, $0 \le y$ in the first quadrant, and let $\mathbf{g}$ be the quadratic map from the previous exercise. Determine

$$\operatorname{area}\mathbf{g}(D) = \iint_{\mathbf{g}(D)} du\,dv \text{ and also } \iint_{\mathbf{g}(D)} \sqrt{u^2+v^2}\,du\,dv.$$

The next four exercises are intended to explore the question: How 'infinite' is $1/r$ at the origin in various dimensions? That is, although $1/r$ is infinite, its integral may or may not be, depending on the dimension of the space in which the calculation is done. The following exercise asks you to explore the same question for $1/r^2$ and then to compare your two sets of results.

9.40. Let B^1 be the interval $[-1,1]$ on the x-axis. Let $r = \sqrt{x^2} = |x|$. Show that

$$\int_{B^1} \frac{1}{r}\,dx = \int_{-1}^{1} \frac{1}{|x|}\,dx = 2\int_0^1 \frac{1}{x}\,dx = \infty.$$

(Think of B^1 as the "unit ball" in one dimension.)

9.41. Let B^2 be the unit disk in the (x,y)-plane: $r^2 = x^2 + y^2 \le 1$. (Think of B^2 as the "unit ball" in two dimensions.) Is

$$\iint_{B^2} \frac{1}{r}\,dx\,dy$$

finite or infinite? Did you make a coordinate change to calculate the integral?

9.42. Let B^3 be the unit ball in (x,y,z)space: $r^2 = x^2 + y^2 + z^2 \le 1$. Use an appropriate change of variables to determine whether

$$\iiint_{B^3} \frac{1}{r}\,dx\,dy\,dz$$

is finite or infinite. Make a conjecture about the integral of $1/r$ over the unit ball in $\mathbb{R}^n$.

9.43. Integrate $1/r^2$ over the unit ball in $\mathbb{R}^n$, $n = 1,2,3$.

a. How does the finiteness of the integral of $1/r^2$ depend on n?
b. In each dimension, how does the integral of $1/r$ compare to the integral of $1/r^2$? Is there some sense in which $1/r$ is either "more infinite" or "less infinite" than $1/r^2$ at the origin?

9.44. Determine

$$\iiint_{x^2+y^2+z^2 \le 1} \frac{1}{1+x^2+y^2+z^2}\,dx\,dy\,dz.$$

Suggestion: Show that $\dfrac{A}{1+A} = 1 - \dfrac{1}{1+A}$, and then use this fact.

9.45. a. Determine

$$\iiint\limits_{R^2 \le x^2+y^2+z^2 \le (R+\Delta R)^2} dx\,dy\,dz,$$

where $R > 0$ is fixed and $\Delta R \ll R$ is small. (The domain here is called a "thin shell".)

b. Show that the integral, which is the volume of the thin shell, equals $4\pi R^2 \Delta R + O(\Delta R^2)$.

9.46. Determine

$$\iiint\limits_{R^2 \le x^2+y^2+z^2 \le (R+\Delta R)^2} \frac{1}{x^2+y^2+z^2}\, dx\,dy\,dz, \quad \Delta R \ll R.$$