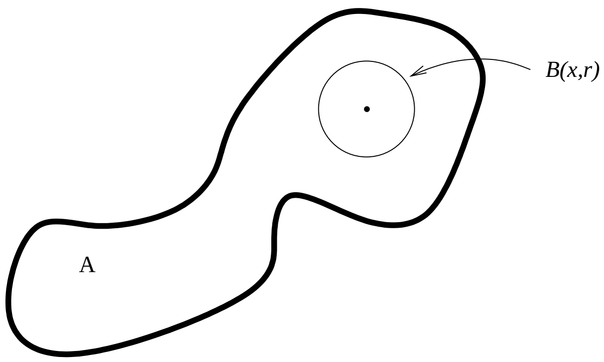
Chapter 10 Further investigation of integration

We have two exceedingly important things to discuss in this chapter: integration over general subsets of $\mathbb{R}^n$, and change of variables. We first discuss a characterization of contented sets.

A. **Topological background**

We have often hinted at many of the following concepts. It is now time to make certain we understand them completely, and have useful names and notations for them.

Recall the definition from Section 3A: if A is a subset of $\mathbb{R}^n$ and $x_0 \in A$, we say that x_0 is an *interior point* of A if there exists $r > 0$ such that the ball $B(x_0, r) \subset A$.



DEFINITION. The set of all interior points of A is called the *interior* of A, and is denoted int$A$ or int(A).

As we have defined a set to be *open* if and only if all its points are interior points, we see that A is open $\iff A = \text{int}A$.

Recall that we proved in Section 3A that the open ball $B(x, r)$ is itself an open set.

PROBLEM 10–1. Prove that every point of int$A$ is in the interior of int$A$. That is, $\text{int}(\text{int}A) = \text{int}A$.

PROBLEM 10–2. The preceding problem shows that int$A$ is an open set. Prove that it is the largest open subset of A, in the sense that if B is an open set and $B \subset A$, then $B \subset \text{int}A$.

PROBLEM 10–3. Prove that

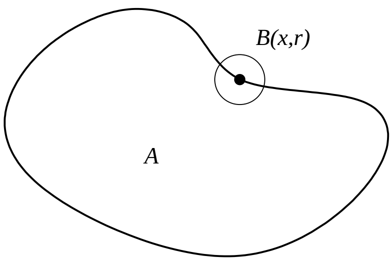
$$\begin{aligned} \operatorname{int}(A \cup B) &\supset \operatorname{int}A \cup \operatorname{int}B; \\ \operatorname{int}(A \cap B) &= \operatorname{int}A \cap \operatorname{int}B. \end{aligned}$$

Prove that the inclusion expressed above may be a strict one.

PROBLEM 10–4. Prove that the union and the intersection of two open sets are also open sets.

The notion which is dual to interior will now be discussed.

DEFINITION. Let $A \subset \mathbb{R}^n$ and $x_0 \in \mathbb{R}^n$. Then x_0 is a *closure point* of A if for every $0 < r < \infty$, $B(x_0, r) \cap A$ is not empty.



DEFINITION. The set of closure points of A is called the *closure* of A, and is denoted cl$A$ or cl(A).

DEFINITION. A set is said to be *closed* if it contains all its closure points.

PROBLEM 10–5. Prove that $\operatorname{cl}(\operatorname{cl}A) = \operatorname{cl}A$.

PROBLEM 10–6. Prove that the closed ball $\overline{B}(x, r)$ is a closed set.

PROBLEM 10–7. Devise the analog of Problem 10–2 and prove the result.

PROBLEM 10–8. Devise the analog of Problem 10–3 and prove the results.

PROBLEM 10–9. Prove that the union and the intersection of two closed sets are also closed sets.

PROBLEM 10–10. Interior and closure are truly dual notions, for

$$\mathrm{int}(\mathbb{R}^n - A) = \mathbb{R}^n - \mathrm{cl}A.$$

PROBLEM 10–11. Prove that A is open $\Longleftrightarrow$ $\mathbb{R}^n - A$ is closed.

WARNING. Don't misread this problem. A frequent error is to think that a set is open $\Longleftrightarrow$ it is not closed.

We have discussed int$A$ and cl$A$. A third set will be of great interest to us:

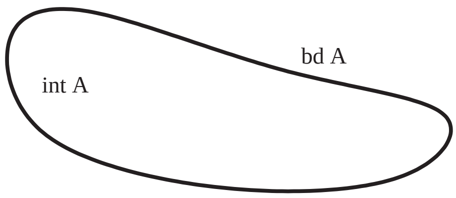
DEFINITION. For a given set $A \subset \mathbb{R}^n$, the *boundary* of A is the set

$$\begin{aligned} \mathrm{bd}A &= \mathrm{bd}(A) \\ &= \mathrm{cl}A - \mathrm{int}A. \end{aligned}$$

PROBLEM 10–12. Let $x_0 \in \mathbb{R}^n$ and $A \subset \mathbb{R}^n$. Prove that $x_0 \in \mathrm{bd}A \Longleftrightarrow$ for all $0 < r < \infty$ the ball $B(x_0, r)$ contains a point belonging to A and also a point not belonging to A.

PROBLEM 10–13. Prove that $\mathrm{bd}(A) = \mathrm{bd}(\mathbb{R}^n - A)$.

PROBLEM 10–14. Prove that $\mathbb{R}^n$ is the disjoint union of the three sets



$$\text{int}A, \quad \text{int}(\mathbb{R}^n - A), \quad \text{bd}A.$$

PROBLEM 10–15. Prove that

$$\text{bd}A = \text{cl}A \cap \text{cl}(\mathbb{R}^n - A).$$

PROBLEM 10–16. Prove that

$$\begin{aligned} A - \text{bd}A &= \text{int}A, \\ A \cup \text{bd}A &= \text{cl}A. \end{aligned}$$

PROBLEM 10–17. Prove that

$$\begin{aligned} \text{bd}(A \cup B) &\subset \text{bd}A \cup \text{bd}B, \\ \text{bd}(A \cap B) &\subset \text{bd}A \cup \text{bd}B, \\ \text{bd}(A - B) &\subset \text{bd}A \cup \text{bd}B. \end{aligned}$$

We conclude this section with an important

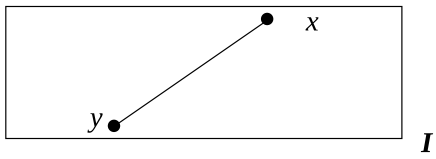
LEMMA. *Let $A \subset \mathbb{R}^n$ and let $I \subset \mathbb{R}^n$ be a special rectangle, and suppose that*

$$I \cap A \neq \emptyset \quad \text{and} \quad I - A \neq \emptyset.$$

Then

$$I \cap \text{bd}A \neq \emptyset.$$

PROOF. The hypothesis implies there exist points x, $y \in I$ such that $x \in A$ and $y \notin A$. Now the idea is to start at x and follow the line segment from x to y until we exit A permanently: so we define the set



$$S = \{t \mid 0 \leq t \leq 1, \quad (1-t)x + ty \in A\}.$$

Thus, S is a subset of the unit interval $[0,1]$. Of course $0 \in S$ and $1 \notin S$. We define $t_0 = \sup S$. Then $0 \leq t_0 \leq 1$, and we set

$$z = (1-t_0)x + t_0 y.$$

- If $t_0 = 0$, then $z \in A$ and the nearby points of $[x,y]$ are not in A, proving $z \in A - \text{int}A \subset \text{bd}A$.
- If $t_0 = 1$, then $z \notin A$ and there are nearby points of $[x,y]$ which are in A, proving $z \in \text{cl}A - A \subset \text{bd}A$.
- If $0 < t_0 < 1$, then any neighborhood of z contains points of $[x,y]$ which are in A $(t < t_0)$ and which are not in A $(t > t_0)$, proving $z \in \text{cl}A - \text{int}A = \text{bd}A$.

In all cases, $z \in \text{bd}A$.

QED

COROLLARY. *If A is a proper subset of $\mathbb{R}^n$* (i.e., *A is not empty and A is not all of $\mathbb{R}^n$*), *then* bd$A$ *is not empty.*

PROBLEM 10–18. Prove that in general a closed set is not equal to the closure of its interior. However, if A is the closure of an open set, then $A = \text{cl}(\text{int}A)$.

PROBLEM 10–19. Likewise, prove that if A is the interior of a closed set, then $A = \text{int}(\text{cl}A)$.

B. **Topological characterization of contentedness**

Now we arrive at a tremendous result, showing that the property of a set's being content can be expressed in terms of just the boundary of the set. The entire situation can be understood from the following equation.

THEOREM. *Let $A \subset \mathbb{R}^n$ be a bounded set. Then*

$$\overline{\mathrm{vol}}(A) \;=\; \underline{\mathrm{vol}}(A) + \overline{\mathrm{vol}}(\mathrm{bd}A).$$

PROOF. This equation will be achieved by proving two inequalities. First, suppose P and Q are elementary polygons such that

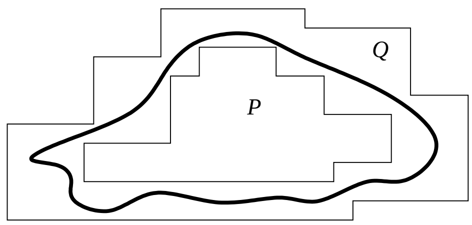
$$P \subset A \subset Q.$$

Then

$$\mathrm{int}P \subset \mathrm{int}A \subset \mathrm{cl}A \subset \mathrm{cl}Q,$$

and thus

$$\mathrm{bd}A \subset \mathrm{cl}Q - \mathrm{int}P.$$



Therefore

$$\begin{aligned}\overline{\mathrm{vol}}(\mathrm{bd}A) &\le \mathrm{vol}(\mathrm{cl}Q - \mathrm{int}P)\\ &= \mathrm{vol}(\mathrm{cl}Q) - \mathrm{vol}(\mathrm{int}P)\\ &= \mathrm{vol}(Q) - \mathrm{vol}(P),\end{aligned}$$

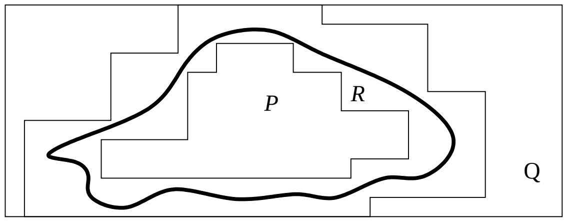
the last equality being due to the fact that the volume of a special rectangle is the same as the volume of its interior. It simply does not matter whether we use closed or open polygons. Now rearrange the inequality:

$$\mathrm{vol}(Q) \;\ge\; \mathrm{vol}(P) + \overline{\mathrm{vol}}(\mathrm{bd}A).$$

As P and Q are arbitrary,

$$\overline{\text{vol}}(A) \;\geq\; \underline{\text{vol}}(A) + \overline{\text{vol}}(\text{bd}A).$$

The reverse inequality is apparently somewhat more difficult to establish. Suppose that R is an arbitrary special polygon which contains bd$A$. Choose a special rectangle such that $R \subset I$. Using the edges of R we construct a partition of I in such a way that R itself consists



of the union of certain special rectangles belonging to the partition. Thus we may express I as a nonoverlapping union of the form

$$I \;=\; R \cup \bigcup_{j=1}^{N} I_j,$$

where each I_j is a special rectangle. The crucial observation to make is that each open rectangle intI_j is either entirely contained in A or entirely contained in $I - A$. For if intI_j contained a point of A and also a point of $I - A$, the lemma of the preceding section would imply that intI_j would contain a point of bd$A$, and thus of R.

We conclude that the sets in the partition of I can be grouped into a nonoverlapping union as follows:

$$I \;=\; R \cup P \cup Q,$$

where P and Q are special polygons such that

$$\text{int}P \subset A, \quad \text{int}Q \subset I - A.$$

Thus

$$\begin{aligned} \text{vol}(I) &= \text{vol}(R) + \text{vol}(P) + \text{vol}(Q) \\ &= \text{vol}(R) + \text{vol}(\text{int}P) + \text{vol}(\text{int}Q) \\ &\leq \text{vol}(R) + \underline{\text{vol}}(A) + \underline{\text{vol}}(I - A). \end{aligned}$$

Since the theorem on p. 9–34 implies that

$$\operatorname{vol}(I) = \overline{\operatorname{vol}}(A) + \underline{\operatorname{vol}}(I - A),$$

we conclude that

$$\overline{\operatorname{vol}}(A) \le \operatorname{vol}(R) + \underline{\operatorname{vol}}(A).$$

Finally, since R is an arbitrary special polygon containing bd$A$,

$$\overline{\operatorname{vol}}(A) \le \overline{\operatorname{vol}}(\mathrm{bd}A) + \underline{\operatorname{vol}}(A).$$

QED

We have as an immediate corollary the definitive result about contentedness,

THEOREM. *A bounded subset of* $\mathbb{R}^n$ *is contented* $\iff$ *its boundary has zero volume.*

PROBLEM 10–20. The theorem on p. 9–30 gives a very elementary proof that if A and B are contented, then so are $A \cap B$, $A \cup B$, and $A - B$. Give a different proof based on the theorem we have just proved.

REMARK. We have now seen two proofs of the fact that if we deal with the collection of contented subsets of $\mathbb{R}^n$, then the operations of taking finite intersections, finite unions, and differences produce contented sets. This is described in set theory by saying that the contented sets form an *algebra* of subsets of $\mathbb{R}^n$. The one finite set operation we can't actually include is that of complementation, but this is only because we require contented sets to be bounded.

C. **Integration over more general sets**

Up to this point we have been integrating only over special rectangles. It is extremely easy to generalize to integration over arbitrary sets.

DEFINITION. Suppose A is a bounded subset of $\mathbb{R}^n$, and $A \xrightarrow{f} \mathbb{R}$ is a bounded function defined on A. Let I be any special rectangle containing A, and define the new function $I \xrightarrow{f_e} \mathbb{R}$ by

$$f_e(x) = \begin{cases} f(x) & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

(The notation is intended to suggest that f has been *e*xtended to be zero outside of A.) Then f_e is a bounded function on I, and we define

$$\underline{\int}_A f = \underline{\int}_I f_e,$$
$$\overline{\int}_A f = \overline{\int}_I f_e.$$

(It is an easy matter to check that the definition of these new lower and upper integrals is independent of the choice of I; this same situation appeared also in Section 9H.)

In particular, we can now simply write

$$\underline{\int}_A 1 = \underline{\mathrm{vol}}(A),$$
$$\overline{\int}_A 1 = \overline{\mathrm{vol}}(A).$$

DEFINITION. We say f is *integrable* over A if these lower and upper integrals are the same, and we denote their common value as

$$\int_A f = \underline{\int}_A f = \overline{\int}_A f.$$

It is a straightforward task to use this definition effectively, so we will not have to say much about it. Here is an example, a generalization of the theorem on p. 9–33:

THEOREM. *Let A and B be bounded subsets of $\mathbb{R}^n$, and let f be integrable over B. Then*

$$\int_B f = \underline{\int}_{B\cap A} f + \overline{\int}_{B-A} f.$$

PROOF. Let I be any special rectangle containing B and define f_e to be the extension of f obtained by giving it the value zero outside B. Then f_e is an integrable function on I. Then Problem 9–12 implies immediately that

$$\int_I f_e = \underline{\int}_I f_e 1_{B\cap A} + \overline{\int}_I (f_e - f_e 1_{B\cap A}).$$

But we have the easy observations

$$f_e 1_{B\cap A} = \begin{cases} f & \text{on } B\cap A \\ 0 & \text{outside } B\cap A, \end{cases}$$

$$f_e - f_e 1_{B\cap A} = \begin{cases} f & \text{on } B - A \\ 0 & \text{outside } B - A. \end{cases}$$

Thus we conclude respectively

$$\underline{\int}_I f_e 1_{B\cap A} = \underline{\int}_{B\cap A} f,$$

$$\overline{\int}_I (f_e - f_e 1_{B\cap A}) = \overline{\int}_{B-A} f.$$

QED

PROBLEM 10–21. Suppose f is integrable over A and B is a contented subset of $\mathbb{R}^n$. Prove that f is integrable over $A\cap B$.

D. **Cavalieri's principle**

This old and famous principle enables us "to find simply and rapidly the volumes of various geometric figures." It's an immediate consequence of Fubini's theorem:

PROBLEM 10–22 (Cavalieri). Let $A \subset \mathbb{R}^n$ be contented, and for each $t \in \mathbb{R}$ let $A(t)$ be the "slice"

$$A(t) = \{y \in \mathbb{R}^{n-1} \mid (y,t) \in A\}$$

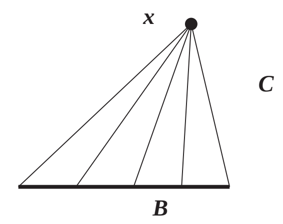
(thus $A(t)$ is a bounded subset of $\mathbb{R}^{n-1}$). Prove that

$$\mathrm{vol}_n(A) = \int_{-\infty}^{\infty} \underline{\mathrm{vol}}_{n-1}(A(t))dt = \int_{-\infty}^{\infty} \overline{\mathrm{vol}}_{n-1}(A(t))dt$$

(each of these integrals is of course extended only over a finite interval of $\mathbb{R}$, since $A(t)$ is empty for all sufficiently large $|t|$).

PROBLEM 10–23. This is a favorite problem of basic calculus texts. Find the volume of the intersection of two right circular cylinders in $\mathbb{R}^3$ which have the same radius r and whose axes are orthogonal. [You may assume the resulting set has the description $x_1^2 + x_3^2 \le r^2$, $x_2^2 + x_3^2 \le r^2$.]

PROBLEM 10–24. Let B be a contented subset of $\mathbb{R}^{n-1}$, regarded as the subset $\mathbb{R}^{n-1} \times \{0\}$ of $\mathbb{R}^n$. Let $x \in \mathbb{R}^n$, with n^{th} coordinate $x_n = h > 0$. Let C be the *cone* with base B and vertex x:



$$C = \{(1-t)y + tx \mid 0 \le t \le 1, \quad y \in B\}.$$

It is a fact that C is a contented subset of $\mathbb{R}^n$, but you need not prove that. Prove that

$$\text{vol}_n(C) = \frac{1}{n}\text{vol}_{n-1}(B)h.$$

(The volume of a cone in $\mathbb{R}^n$ is equal to $1/n$ times the altitude times the volume of the base.)

Later in the present chapter we shall compute the volumes of balls $B(x,r) \subset \mathbb{R}^n$ for all n. We can actually already compute these, at least in principle. We introduce the following

NOTATION. $\boxed{\alpha_n = \text{vol}_n(B(0,1))}$

It is an easy task to conclude that

$$\text{vol}_n(B(x,r)) = \alpha_n r^n.$$

PROBLEM 10–25. Use Cavalieri's principle to complete the following table:

$$\begin{aligned} \alpha_1 &= 2 \\ \alpha_2 &= \pi \\ \alpha_3 &= ? \\ \alpha_4 &= ? \\ \alpha_5 &= ? \\ \alpha_6 &= ? \end{aligned}$$

PROBLEM 10–26. Calculate the volume in $\mathbb{R}^3$ of the intersection of three cylinders: specifically,

$$\operatorname{vol}_3(\{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 \le 1, \quad x_1^2 + x_3^2 \le 1, \quad x_2^2 + x_3^2 \le 1\}).$$

PROBLEM 10–27. Calculate the 4-dimensional volume of the set

$$\{x \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + x_4^2 \le 1, \quad x_1^2 + x_3^2 + x_4^2 \le 1, \quad x_2^2 + x_3^2 + x_4^2 \le 1\}.$$

PROBLEM 10–28. Calculate the 4-dimensional volume of the set

$$\{x \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + x_4^2 \le 1, \quad x_1^2 + x_3^2 + x_4^2 \le 1\}.$$

PROBLEM 10–29. Calculate the 4-dimensional volume of the set

$$\{x \in \mathbb{R}^4 \mid x_1^2 + x_2^2 \le 1, \quad x_3^2 + x_4^2 \le 1\}.$$

PROBLEM 10–30. Calculate the 4-dimensional volume of the set

$$\{x \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + x_4^2 \le 1, \quad x_3^2 + x_4^2 \le 1\}.$$

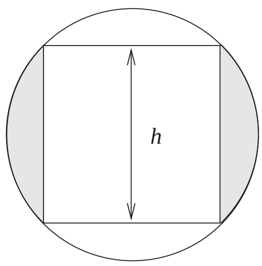
PROBLEM 10–31. For any dimension $n \ge 1$ and any real number $a > 0$, let

$$S_n(a) \;=\; \{x \in \mathbb{R}^n \mid 0 \le x_i \text{ for all } i \text{ and } x_1 + \cdots + x_n \le a\}.$$

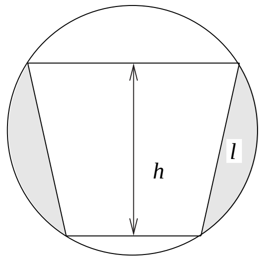
Calculate $\mathrm{vol}_n(S_n(a))$.

PROBLEM 10–32. A favorite calculus "stunt" is this problem: "A ball in $\mathbb{R}^3$ has a cylindrical hole bored from it, the axis of the cylinder coinciding with a diameter of the ball. The set that remains has height h. What is its volume."

a. Use Cavalieri to calculate the volume.
b. Why is this called a stunt?

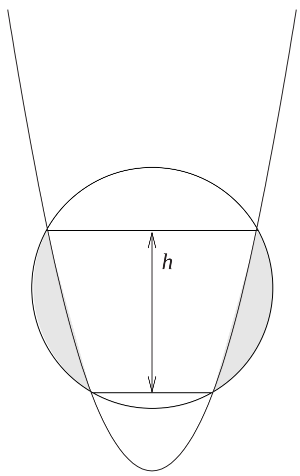


PROBLEM 10–33*. Generalize Problem 10–32 by assuming that the hole bored out comes from a right circular cone with axis coinciding with a diameter of the ball:



Calculate the remaining volume as a function of the altitude h and the slant height ℓ.

PROBLEM 10–34*. Generalize Problem 10–32 by assuming that the hole bored out is a paraboloid of revolution:



Calculate the remaining volume as a function of h.

E. **Elementary matrices**

We are now going to begin the final item of this chapter, the important change of variable formula for integrals on $\mathbb{R}^n$. We shall first treat the case of linear variable transformations. This will be done in the following section, and in preparation for that we need to present an interesting feature of $n \times n$ matrices which we have not mentioned heretofore.

As is our custom, when we think of a linear relation connecting x and $y \in \mathbb{R}^n$, we represent our points as column vectors and we write $y = Tx$, where T is of course an $n \times n$ real matrix. We certainly assume that T is invertible ($\det T \neq 0$), as we are going to be interested in a one-to-one correspondence between x and y.

We are particularly interested in two special kinds of what we shall call *elementary* matrices. In all cases below k and ℓ stand for fixed integers between 1 and n.

Multiplying. This is any matrix of the following form. Let $c \neq 0$. Let M be defined by

$$\begin{aligned} m_{ii} &= 1 \qquad \text{if } \ i \neq k, \\ m_{kk} &= c, \\ m_{ij} &= 0 \qquad \text{if } \ i \neq j. \end{aligned}$$

Note that $\underline{\det M = c}$ and that M^{-1} is also a multiplying matrix with c replaced by c^{-1}. Illustration ($k = 1$):

$$\begin{pmatrix} c & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Adding. In this case $k \neq \ell$ and let $c \in \mathbb{R}$. Let the matrix A be defined by

$$\begin{aligned} a_{ii} &= 1, \\ a_{kl} &= c, \\ a_{ij} &= 0 \qquad \text{for all other } \ (i,j). \end{aligned}$$

Note that $\underline{\det A = 1}$ and that A^{-1} is another adding matrix, with c replaced by $-c$. Illustration ($k = 1$, $l = 2$):

$$\begin{pmatrix} 1 & c & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Now notice the effect of multiplying a given matrix on the *left* by these two kinds of elementary matrices:

Multiplying. MT is the matrix obtained from T by multiplying row k by the number c.

Adding. AT is the matrix obtained from T by adding c times row l to row k.

PROBLEM 10–35. Describe what happens to T when T is multiplied on the *right* by an elementary matrix.

It is quite amazing at first glance that just these two types of elementary matrices suffice to generate all invertible matrices, but the proof is easy:

THEOREM. *Every invertible matrix can be expressed as a product of elementary matrices.*

PROOF. Let T be an arbitrary $n \times n$ invertible matrix. We first want to maneuver to obtain the entry 1 in the $(1,1)$ position. There are two possibilities. If $t_{11} \neq 0$, then there exists a unique multiplying matrix M such that MT has the desired property. If $t_{11} = 0$, then the fact that T is invertible implies that some $t_{i1} \neq 0$, and thus there exists an adding matrix A such that AT has the desired property. We describe both possibilities by saying that there exists an elementary matrix E such that ET has the form

$$ET = \begin{pmatrix} 1\underline{\hspace{3em}} \\ \underline{\hspace{3.5em}} \\ \vdots \\ \underline{\hspace{3.5em}} \end{pmatrix}.$$

Then there exists a sequence of adding matrices A_k such that

$$A_{n-1}A_{n-2}\dots A_1ET = \begin{pmatrix} 1 & \underline{\hspace{3em}} \\ 0 & \underline{\hspace{3em}} \\ \vdots & \vdots \\ 0 & \underline{\hspace{3em}} \end{pmatrix}.$$

By Problem 10–35, another sequence of adding matrices $\widetilde{A}_k$ can be found such that

$$A_{n-1}\dots A_1ET\widetilde{A}_1\dots\widetilde{A}_{n-1} = \begin{pmatrix} 1 & 0\dots0 \\ 0 & \underline{\hspace{3em}} \\ \vdots & \vdots \\ 0 & \underline{\hspace{3em}} \end{pmatrix}.$$

This new matrix we have manufactured contains an invertible $(n-1) \times (n-1)$ submatrix, to which we can apply the same procedure. This will not disturb the 1 or the 0's which are displayed above. Therefore, after n stages we arrive at a formula which can be expressed symbolically as

$$E_r E_{r-1} \dots E_2 E_1 T \; F_1 F_2 \dots F_{s-1} F_s \;=\; I,$$

where all the E's and F's are elementary matrices of the two kinds. Multiplying by their inverses in the correct way produces the formula

$$T \;=\; E_1^{-1} E_2^{-1} \dots E_r^{-1} F_s^{-1} \dots F_2^{-1} F_1^{-1}.$$

Since the inverses are also elementary matrices, the theorem is proved.

QED

Before presenting the integration application of this theorem about matrices, we pause to present an interesting proof of the profound multiplicative property of determinants. This fascinating proof is quite different from the proof we gave back in Section 3G.

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

PROOF. We are of course assuming A and B are $n \times n$ matrices. First we dispense with the easy singular case in which $\det A = 0$. In this case the columns of A are linearly dependent (see p. 3–36). As the columns of AB are themselves linear combinations of the columns of A, the columns of AB are also linearly dependent. Therefore $\det(AB) = 0$ and the desired equation holds.

Now assume $\det A \neq 0$. Then we know that A can be expressed as a product of elementary matrices,

$$A \;=\; E_1 \dots E_k.$$

But each of the two kinds of elementary matrices enjoys the elementary property

$$\det(EC) \;=\; \det E \det C,$$

as EC results from C by a row operation as described above. Thus we conclude that

$$\begin{aligned} \det(AB) &= \det(E_1 E_2 \dots E_k B) \\ &= \det E_1 \det(E_2 \dots E_k B) \\ &\;\vdots \\ &= \det E_1 \dots \det E_k \det B. \end{aligned}$$

Now use $B = I$ to conclude also that

$$\det A \;=\; \det E_1 \dots E_k.$$

Therefore,

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

QED

PROBLEM 10–36. Prove that the representation produced in the theorem can be achieved with no more than n^2 elementary matrices. Why is n^2 "obviously" the least number of elementary matrices that can be employed in general?

PROBLEM 10–37. Write out explicitly the representation of 2×2 matrices as products of elementary matrices.

PROBLEM 10–38. Let T represent a $90°$ rotation:

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Write T as a product of three adding matrices $T = E_1E_2E_3$, and draw a sketch of the result of operating step by step on the square $J = [0,1] \times [0,1]$. That is, sketch the parallelograms J, E_3J, E_2E_3J, and TJ.

F. **Linear changes of variables**

In this section we are going to derive and prove a tremendous formula for changing dummy variables *linearly* in a Riemann integral over $\mathbb{R}^n$. The key ingredient in this formula will be the determinant of the matrix which encodes the linear change of variables.

Before we launch into the actual statement and proof, let's examine a nice special case. Let T be an invertible $n \times n$ real matrix, and express T in the usual manner of displaying its columns:

$$T = (t_1 \; t_2 \; \dots \; t_n),$$

where t_j is of course the column vector

$$t_j = \begin{pmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{nj} \end{pmatrix}.$$

Denote by J the special rectangle in $\mathbb{R}^n$,

$$J = [0,1] \times \cdots \times [0,1].$$

Then we denote by TJ the set resulting from J by multiplying each member of J by T:

$$\begin{aligned} TJ &= \{Tx \mid x \in J\} \\ &= \left\{ T\sum_{j=1}^{n} x_j\hat{e}_j \mid \text{ each } \ 0 \le x_j \le 1 \right\} \\ &= \left\{ \sum_{j=1}^{n} x_j t_j \mid \text{ each } \ 0 \le x_j \le 1 \right\}. \end{aligned}$$

This set is therefore exactly the n-dimensional parallelogram with vertex 0 and edges $t_1, \dots, t_n$; see p. 8–1. We know from our *intuitive* treatment of volumes in Chapter 8 that

$$\text{vol}(TJ) = |\det T|.$$

That is,

$$\text{vol}(TJ) = |\det T| \ \text{vol}(J).$$

We are now going to *prove* from integration theory that this formula holds in utter generality, in that J can be replaced by any contented set in $\mathbb{R}^n$. In fact, we shall prove a great deal more even than this.

Before doing this, we clarify one small point. When one integrates on $\mathbb{R}^1$ and makes a change of variable $x = -y$, one may write in the time-honored way

$$\int_a^b f(x)dx = -\int_{-a}^{-b} f(-y)dy.$$

This is not very convenient for us, as our usual condition $a < b$ produces $-a > -b$ in the "integral" on the right side. Thus we much prefer to write

$$\int_a^b f(x)dx = \int_{-b}^{-a} f(-y)dy.$$

Notice the two things that occur with this viewpoint: (1) the change of variable $y = -x$ transforms the interval $[a, b]$ to the interval $[-b, -a]$, and (2) the *absolute value* of dy/dx is the correct change of "differential":

$$dx = dy.$$

The above point of view is the convenient one for n-dimensional integration. In fact, it is convenient to regard the functions we are integrating as defined on all of $\mathbb{R}^n$, but equal to zero outside some bounded set. Then we simply need not worry about the domain of integration, and we simply write the integral as

$$\int_{\mathbb{R}^n} f(x)dx.$$

Before starting the theorem, a lemma of a rather technical nature needs to be handled. It will help us in stating it if we employ a slight abuse of notation. If T is an $n \times n$ matrix, let us also denote the associated function from $\mathbb{R}^n$ to $\mathbb{R}^n$ by the same letter T; that is, T is the function whose value at $x \in \mathbb{R}^n$ is Tx. (See Problem 2–77 for a discussion of the distinction that usually should be made.) As a result, we can write $f \circ T$ for the composite funciton, the function whose value at x is $f(Tx)$.

You will see that virtually all the "hard" work involved in proving the desired result is contained in the following

LEMMA. *Let E be an elementary matrix, and let $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^n$ be a step function. Then $f \circ E$ is integrable, and*

$$\int_{\mathbb{R}^n} f(y)dy = |\det E| \int_{\mathbb{R}^n} f(Ex)dx.$$

PROOF. Since step functions are linear combinations of indicator functions of special rectangles, we may simply assume that f itself is such an indicator function:

$$f = 1_J, \qquad \text{where} \quad J = [a_1, b_1] \times \cdots \times [a_n, b_n].$$

Note that $(1_J \circ E)(x) = 1 \iff Ex \in J \iff x \in E^{-1}J$. Thus we have the simple formula

$$1_J \circ E = 1_{E^{-1}J}.$$

We see then that the formula to be proved is

$$\int_{\mathbb{R}^n} 1_J = |\det E| \int_{\mathbb{R}^n} 1_{E^{-1}J};$$

in other words,

$$\text{vol}(J) = |\det E| \text{vol}(E^{-1}J).$$

Of course we must also prove that $1_{E^{-1}J}$ is integrable, i.e., that $E^{-1}J$ is contented.

In case E is a *multiplying* matrix, this is quite simple. As a typical case suppose

$$E = \begin{pmatrix} a & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}, \qquad \text{where} \quad a < 0.$$

Then

$$E^{-1}J \;=\; \left[\frac{b_1}{a}, \frac{a_1}{a}\right] \times [a_2, b_2] \times \cdots \times [a_n, b_n],$$

so $E^{-1}J$ is itself also a special rectangle, thus contented, and

$$\frac{\mathrm{vol}(J)}{\mathrm{vol}(E^{-1}J)} \;=\; \frac{b_1 - a_1}{\frac{a_1}{a} - \frac{b_1}{a}} \;=\; -a \;=\; |\det E|.$$

The case in which E is an *adding* matrix has a bit of a twist (really, a *shear*). As a typical case suppose

$$E \;=\; \begin{pmatrix} 1 & & \\ & \ddots & \\ c & & 1 \end{pmatrix},$$

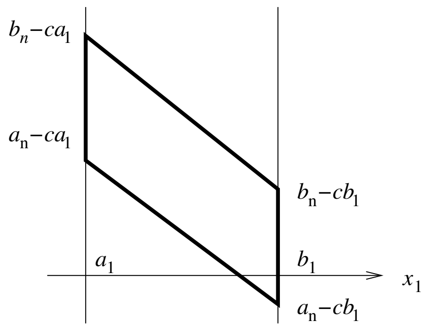
where the nonzero entry c is in the $(n, 1)$ position. Then $E^{-1}J$ is not a rectangle, but is instead described by noting that $x \in E^{-1}J \iff Ex \in J \iff$

$$\begin{aligned} a_1 \;\le\; & x_1 && \le\; b_1, \\ & \vdots && \\ a_{n-1} \le\; & x_{n-1} && \le\; b_{n-1}, \\ a_n \;\le\; & cx_1 + x_n && \le\; b_n. \end{aligned}$$

The last condition can be rewritten in the form

$$a_n - cx_1 \;\le\; x_n \;\le\; b_n - cx_1,$$

and $E^{-1}J$ can be represented by a sketch in the $x_1 - x_n$ plane:



First, this set is seen as the set contained between the graphs of two functions of $x_1, \dots, x_{n-1}$ (namely, the affine functions $a_n - cx_1$ and $b_n - cx_1$). As these functions are continuous, the set $E^{-1}J$ is contented. This is an application of a slight modification of the theorem of Section 9I. Finally, Cavalieri's principle gives the required formula for volume. The slice of $E^{-1}J$ for a given $x_1 = t$, $a_1 \le t \le b_1$, is the special rectangle

$$[a_2, b_2] \times \cdots \times [a_{n-1}, b_{n-1}] \times [a_n - ct, b_n - ct],$$

and the $(n-1)$-dimensional volume of this slice is

$$(b_2 - a_2) \cdots (b_n - a_n).$$

Thus Cavalieri gives us

$$\begin{aligned} \operatorname{vol}(E^{-1}J) &= \int_{a_1}^{b_1} (b_2 - a_2) \cdots (b_n - a_n) dt \\ &= (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n) \\ &= \operatorname{vol}(J). \end{aligned}$$

Since $\det E = 1$, the desired formula is proved.

QED

The hard work is over, and it is now a rather soft task to finish what we need to say about the linear case.

THEOREM. *Let T be a real $n \times n$ matrix with $\det T \neq 0$. Let $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ be a bounded function which is zero outside some bounded set. Then*

$$\overline{\int}_{\mathbb{R}^n} f(y)dy = |\det T| \overline{\int}_{\mathbb{R}^n} f(Tx)dx,$$

and likewise for lower integrals. Moreover, f is integrable $\iff$ $f \circ T$ is integrable, and in this case

$$\int_{\mathbb{R}^n} f(y)dy = |\det T| \int_{\mathbb{R}^n} f(Tx)dx.$$

PROOF. The proof divides naturally into two stages. First we assume $T = E$ is an elementary matrix. Then we consider any step function $\tau \ge f$. Clearly, $\tau \circ E \ge f \circ E$.

As we know from the lemma, $\tau \circ E$ is integrable. Thus

$$\begin{aligned} |\det E| \overline{\int} f \circ E &\leq |\det E| \overline{\int} \tau \circ E \\ &= |\det E| \int \tau \circ E \\ &= \int \tau \qquad \text{(also from the lemma).} \end{aligned}$$

Since $\tau \geq f$ is arbitrary, we obtain

$$|\det E| \overline{\int} f \circ E \leq \overline{\int} f. \tag{(*)}$$

That's only an inequality. But we can employ the trick of replacing f by $f \circ E^{-1}$ in $(*)$:

$$|\det E| \overline{\int} f \leq \overline{\int} f \circ E^{-1}.$$

Another ruse: replace E by E^{-1} in this inequality, and multiply both sides by $|\det E|$, to obtain

$$\overline{\int} f \leq |\det E| \overline{\int} f \circ E.$$

This is the reverse of $(*)$, so the theorem holds for the case of elementary matrices.

The second stage of the proof is just the observation that if the theorem holds for two matrices T_1 and T_2, then it also holds for their product. For

$$\overline{\int} f = |\det T_1| \overline{\int} f \circ T_1.$$

Then use the theorem for the matrix T_2 and the function $f \circ T_1$:

$$\begin{aligned} \overline{\int} f \circ T_1 &= |\det T_2| \overline{\int} f \circ T_1 \circ T_2 \\ &= |\det T_2| \overline{\int} f \circ (T_1 T_2). \end{aligned}$$

Since the determinant is multiplicative, these two equations combine to produce

$$\overline{\int} f = |\det T_1 T_2| \overline{\int} f \circ (T_1 T_2).$$

Finally, since any T is a product of elementary matrices, and the theorem is valid for elementary matrices, the theorem is valid in general.

QED

PROBLEM 10–39. Prove that if A is any bounded set in $\mathbb{R}^n$,

$$\overline{\text{vol}}(TA) \;=\; |\det T|\,\overline{\text{vol}}(A),$$

and likewise for lower volumes. Prove that A is contented $\iff TA$ is contented and in this case

$$\text{vol}(TA) \;=\; |\det T|\,\text{vol}(A).$$

PROBLEM 10–40. Prove that if $\Phi \in O(n)$, then

$$\overline{\text{vol}}(\Phi A) \;=\; \overline{\text{vol}}(A).$$

PROBLEM 10–41. Let A be the ellipsoid

$$\frac{x_1^2}{a_1^2} + \cdots + \frac{x_n^2}{a_n^2} \;\le\; 1,$$

where the semiaxes $a_1, \dots, a_n$ are any positive numbers. Prove that

$$\text{vol}(A) \;=\; \alpha_n\, a_1\, \dots\, a_n.$$

PROBLEM 10–42. Let M be a symmetric positive definite $n \times n$ matrix, and let A be the ellipsoid defined by

$$Mx \bullet x \;\le\; 1.$$

Prove that

$$\text{vol}(A) \;=\; \frac{\alpha_n}{\sqrt{\det M}}.$$

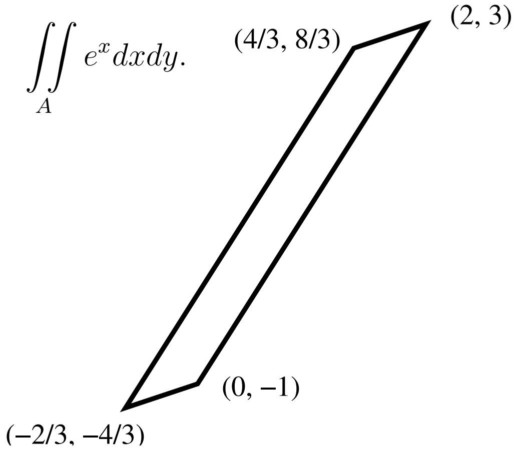
EXAMPLE. Here's a numerical example which nicely illustrates the power of these linear changes of variables. Let A be the parallelogram in $\mathbb{R}^2$ bounded by the four lines

$$\begin{aligned} y &= \frac{x}{2} - 1 \quad \text{and} \quad y = \frac{x}{2} + 2, \\ y &= 2x \quad \text{and} \quad y = 2x - 1. \end{aligned}$$

Compute

$$\iint_A e^x dxdy.$$

We do not even need to draw a sketch of A, but here is one nevertheless:



Now define new dummy variables in a rather obvious way:

$$\begin{cases} u = y - 2x, \\ v = y - \frac{x}{2}. \end{cases}$$

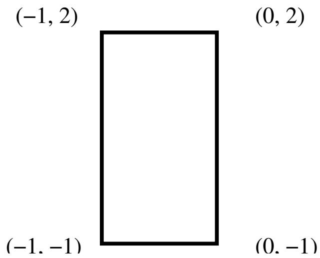
This is a linear change of variables, represented as

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Notice that the determinant of the displayed matrix is $-3/2$. The four lines that form A are expressed in the $u - v$ coordinates as

$$\begin{aligned} v &= -1 \quad \text{and} \quad v = 2, \\ u &= 0 \quad \text{and} \quad u = -1, \quad \text{respectively.} \end{aligned}$$

Here's the sketch in the $u - v$ plane:



We know that $dxdy$ should become $|\det T|dudv$, but T is really the inverse of the given matrix. Thus, $dxdy$ becomes $\frac{2}{3}dudv$. Finally, solving for x gives $x = \frac{2}{3}(v-u)$. Thus

$$\begin{aligned}\iint\limits_A e^x dxdy &= \int_{-1}^{2}\int_{-1}^{0} e^{2(v-u)/3}\,\frac{2}{3}dudv \\ &= \frac{2}{3}\int_{-1}^{2} e^{2v/3}dv \cdot \int_{-1}^{0} e^{-2u/3}du \\ &= \frac{2}{3}\cdot\frac{3}{2}\left(e^{4/3} - e^{-2/3}\right)\cdot\frac{3}{2}\left(e^{2/3}-1\right) \\ &= \frac{3}{2}\left(e^{4/3} - e^{-2/3}\right)\left(e^{2/3}-1\right).\end{aligned}$$

PROBLEM 10–43. Let A be the parallelogram in the $x-y$ plane with vertices $(1,-1)$, $(4,0)$, $(3,2)$, and $(0,1)$. Compute

$$\iint\limits_A y\ dxdy.$$

(Answer: 7/2) Can you now see how to find the area of A by inspection?

PROBLEM 10–44. Let A be the triangular region in the $x-y$ plane determined by the three lines $x-2y=1$, $x+y=4$, and $y=0$. Evaluate the integral

$$\iint\limits_A \sqrt{\frac{x+y}{x-2y}}dxdy$$

by using the coordinates $u = x+y$, $v = x-2y$.

(Answer: 17/9)

PROBLEM 10–45. Given $n+1$ points $p_0, p_1, \ldots, p_n$ in $\mathbb{R}^n$, the *simplex* determined by them is the set

$$S = \left\{ \sum_{j=0}^{n} t_j p_j \mid \text{all } t_j \geq 0 \text{ and } \sum_{j=0}^{n} t_j = 1 \right\}.$$

Thus a simplex in $\mathbb{R}^1$ is an interval, a simplex in $\mathbb{R}^2$ is a triangle, and a simplex in $\mathbb{R}^3$ is a tetrahedron. Show that S can also be expressed in the form

$$S = \left\{ p_0 + \sum_{j=1}^{n} t_j (p_j - p_0) \mid \text{all } t_j \geq 0 \text{ and } \sum_{j=1}^{n} t_j \leq 1 \right\}.$$

Using Problem 10–24, show that

$$\operatorname{vol}(S) = \frac{1}{n!} \left| \det(p_1 - p_0 \quad p_2 - p_0 \ \cdots \ p_n - p_0) \right|.$$

(Each p_j is written as a column vector.)

PROBLEM 10–46. Show that the result of the preceding problem can also be expressed as an $(n+1) \times (n+1)$ determinant:

$$\operatorname{vol}(S) = \frac{1}{n!} \left| \det \begin{pmatrix} p_0 & p_1 & \cdots & p_n \\ 1 & 1 & & 1 \end{pmatrix} \right|.$$

G. **The general formula for change of variables**

The purpose of the present section is the understanding of the nonlinear generalization of the change of variables formula of the preceding section. This is one of the truly great results of calculus. It is the final tool which we need in order to make us adept at handling integrals in several variables.

Important as it is, we choose to omit the proof. There are several reasons for this choice. First, knowing the proof will not help us at all in applying the result. Second, the context for providing an efficient proof is Lebesgue, not Riemann, integration. To be sure, we could give a proof using Riemann integration, but the technical details which would be required are actually far more involved than would be the case if we possessed the tool of Lebesgue integration; the proof in the Lebesgue context is actually quite elegant. Third, the formula we are going to present is a very natural guess based on our understanding of the linear case, so much so that the result is quite believable.

Thus our goal is to give a precise statement of the result, to provide several useful examples, and to provide exercises which illustrate the power of the result. We now give the exact hypothesis.

Let A_1 and A_2 be bounded open contented subsets of $\mathbb{R}^n$, and let

$$A_1 \xrightarrow{\Phi} A_2$$

be a C^1 bijection whose inverse is also of class C^1.

We shall clearly see in some of the applications how these assumptions might be relaxed in various ways.

THEOREM. *Given the above situation, suppose*

$$A_2 \xrightarrow{f} \mathbb{R}$$

is integrable. Then $f \circ \Phi$ is also integrable, and

$$\int_{A_2} f(y)dy \;=\; \int_{A_1} f(\Phi(x)) \mid \det \Phi'(x) \mid dx.$$

In this formula $\Phi'(x)$ is the Jacobian matrix defined in Section 2H.

REMARKS. The determinant of the Jacobian matrix of Φ is often called the *Jacobian determinant* of Φ. Notice that in the case of a linear function $y = Tx$, the Jacobian determinant is $\det T$. If we employ a sort of loose notation, regarding Φ as representing $y_1, \ldots, y_n$ as functions of $x_1, \ldots, x_n$, then the Jacobian determinant is precisely

$$\det \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{pmatrix}.$$

Notice also that in the linear case our result is precisely that of the preceding section. In that section the factor $|\det T|$ served as a type of magnification factor for volumes, as in the typical formula

$$\mathrm{vol}(TA) \;=\; |\det T| \;\mathrm{vol}(A)$$

of Problem 10–39.

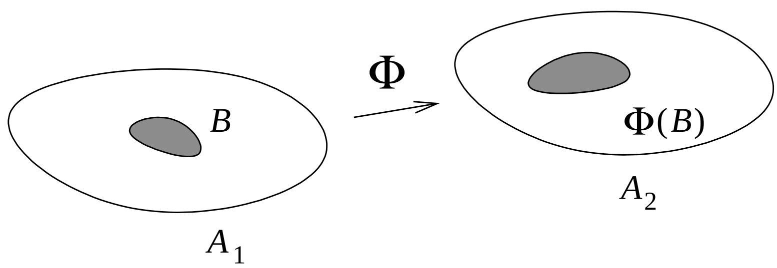
PROBLEM 10–47. Apply the theorem to the following situation: $B \subset A_1$ is a contented set and f is the indicator function on A_2 defined by

$$f(y) = 1 \Longleftrightarrow \Phi^{-1}(y) \in B.$$

Show that the resulting formula is

$$\mathrm{vol}(\Phi(B)) = \int_B |\det \Phi'(x)| dx.$$

As a result of this problem we see that Φ transforms subsets of A_1 to subsets of A_2 in such a manner that volumes of the transformed sets are calculated using integration over the initial set with a sort of *local* magnification factor $|\det \Phi'(x)|$. You should be comfortable with this generalization of the linear case and should feel that approximating B with special polygons and using the continuity of $|\det \Phi'(x)|$ could reasonably be expected to provide an actual proof.



POLAR COORDINATES. This is probably the most important of all examples. We use the usual formulas for polar coordinates in the $x - y$ plane,

$$\begin{cases} x = r\cos\theta, \\ y = r\sin\theta. \end{cases}$$

Here $0 < r < \infty$ and θ is restricted to some interval of length 2π. (Since $r > 0$ we are missing the origin, and if we think of $0 < \theta < 2\pi$ we are missing the positive x-axis. These are inconsequential details, as the missing sets have zero two-dimensional volume.) The Jacobian determinant is

$$\det \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r,$$

so that when we change variables we replace $dxdy$ by $rdrd\theta$. (Notice the dimensional correctness.) Thus

$$\iint f(x,y)dxdy = \int_0^{2\pi}\int_0^{\infty} f(r\cos\theta, r\sin\theta) r dr d\theta.$$

EXAMPLE. We can now give a quick derivation of a recursion formula for α_n, the volume of the unit ball in $\mathbb{R}^n$. We do this by modifying Cavalieri's principle to step down *two* dimensions at once. Thus

$$\begin{aligned}
\mathrm{vol}_n(B(0,R)) &= \int_{x_1^2+\cdots+x_n^2<R^2} dx_1\cdots dx_n \\
&= \iint\limits_{x_1^2+x_2^2<R^2} \left(\int_{x_3^2+\cdots+x_n^2<R^2-x_1^2-x_2^2} dx_3\cdots dx_n \right) dx_1dx_2 \\
&= \iint\limits_{x_1^2+x_2^2<R^2} \mathrm{vol}_{n-2}\left(B(0, \sqrt{R^2-x_1^2-x_2^2} \right) dx_1dx_2 \\
&= \iint\limits_{x_1^2+x_2^2<R^2} \alpha_{n-2}(R^2-x_1^2-x_2^2)^{\frac{n-2}{2}} dx_1dx_2 \\
&\overset{\text{polar coordinates}}{=} \int_0^{2\pi}\int_0^R \alpha_{n-2}(R^2-r^2)^{\frac{n}{2}-1} r dr d\theta \\
&= 2\pi\alpha_{n-2}\cdot\frac{1}{-2}\cdot\frac{(R^2-r^2)^{n/2}}{n/2}\Big|_{r=0}^{r=R} \\
&= \frac{2\pi}{n}\alpha_{n-2}R^n.
\end{aligned}$$

Thus

$$\boxed{\alpha_n = \frac{2\pi}{n}\alpha_{n-2}.}$$

We conclude therefore that

$$\alpha_n = \begin{cases} \frac{2\pi}{n}\cdot\frac{2\pi}{n-2}\cdots\frac{2\pi}{4}\pi & \text{if } n \text{ is even,} \\ \frac{2\pi}{n}\cdot\frac{2\pi}{n-2}\cdots\frac{2\pi}{3}2 & \text{if } n \text{ is odd.} \end{cases}$$

In other words,

$$\alpha_n = \begin{cases} \dfrac{\pi^{n/2}}{\frac{n}{2}\cdot\frac{n-2}{2}\cdots 1} & \text{if } n \text{ is even,} \\ \dfrac{\pi^{(n-1)/2}}{\frac{n}{2}\cdot\frac{n-2}{2}\cdots\frac{1}{2}} & \text{if } n \text{ is odd.} \end{cases}$$

THE GAUSSIAN INTEGRAL. We calculated this integral once before, in Section 9G, as an illustration of Fubini's theorem. We do it again here to illustrate polar coordinates. Again, as in Section 9G, define the number

$$A = \int_0^\infty e^{-x^2} dx.$$

Then

$$\begin{aligned} A^2 &= A\int_0^\infty e^{-x^2} dx \\ &= \int_0^\infty Ae^{-x^2} dx \\ &= \int_0^\infty \left(\int_0^\infty e^{-y^2} dy\right) e^{-x^2} dx \\ &= \int_0^\infty \left(\int_0^\infty e^{-x^2-y^2} dy\right) dx \\ &\overset{\text{Fubini}}{=} \int_0^\infty \int_0^\infty e^{-x^2-y^2} d\mathrm{vol}_2 \\ &\overset{\text{polar coordinates}}{=} \int_0^{\pi/2} \int_0^\infty e^{-r^2} r\,dr\,d\theta \\ &= \frac{\pi}{2}\int_0^\infty e^{-r^2} r\,dr \\ &= -\frac{\pi}{4}e^{-r^2}\Big|_0^\infty \\ &= \frac{\pi}{4}. \end{aligned}$$

Thus $A = \frac{1}{2}\sqrt{\pi}$. This was our second computation of the Gaussian integral. We'll see a third in Section H.

SPHERICAL COORDINATES FOR $\mathbb{R}^3$. As we discussed in Section 6D, we use

$$\begin{cases} x &= r\sin\varphi\cos\theta, \\ y &= r\sin\varphi\sin\theta, \\ z &= r\cos\varphi. \end{cases}$$

Here $0 < r < \infty$, $0 < \varphi < \pi$, and $0 < \theta < 2\pi$. The Jacobian determinant is (if we use the ordering r, φ, θ)

$$\det\begin{pmatrix} \sin\varphi\cos\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta \\ \sin\varphi\sin\theta & r\cos\varphi\sin\theta & r\sin\varphi\cos\theta \\ \cos\varphi & -r\sin\varphi & 0 \end{pmatrix}.$$

This can be calculated directly to be $r^2\sin\varphi$. An alternate point of view is that the columns of the matrix are orthogonal vectors whose lengths are, respectively, 1, r, and $r\sin\varphi$. The determinant is $\pm$ the volume of the rectangle with these three sides, and is thus $\pm r^2\sin\varphi$. We don't care about the sign in the context of integration, as we use the absolute value. Notice the dimensional correctness of the formula

$$dxdydz = r^2\sin\varphi dr d\varphi d\theta.$$

As an example we compute again the volume of the unit ball,

$$\begin{aligned} \alpha_3 &= \int_{x^2+y^2+z^2<1} dxdydz \\ &= \int_0^{2\pi}\int_0^{\pi}\int_0^1 r^2\sin\varphi dr d\varphi d\theta \\ &= \int_0^{2\pi} d\theta \cdot \int_0^{\pi}\sin\varphi d\varphi \cdot \int_0^1 r^2 dr \\ &= 2\pi\cdot 2\cdot\frac{1}{3} \\ &= 4\pi/3. \end{aligned}$$

PROBLEM 10–48. The formula for spherical coordinates in $\mathbb{R}^4$ is given in Problem 6–16. Show that

$$d\text{vol}_4 = r^3\sin^2\varphi_2\sin\varphi_1 dr d\varphi_1 d\varphi_2 d\theta.$$

Use this formula to recalculate α_4.

INTEGRATION OF SPHERICALLY SYMMETRIC FUNCTIONS. Our experience with $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{R}^4$ reveals a definite pattern. We could write down spherical coordinates for $\mathbb{R}^n$ in terms of r and $n-1$ angles; these angles are $\varphi_1, \dots, \varphi_{n-2}$, and θ. Here each $0 < \varphi_i < \pi$, while $0 < \theta < 2\pi$. These must result in a formula of the form

$$d\mathrm{vol}_n = r^{n-1} g(\varphi_1, \dots, \varphi_{n-2}) dr d\varphi_1 \dots d\varphi_{n-2} d\theta,$$

where we *could* calculate g if we wished.

Now suppose $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ is *spherically symmetric* with respect to the origin. This means that $f(x)$ depends only on $\|x\|$. By abuse of notation we choose to write $f(\|x\|)$. Then the integration formula becomes

$$\int_{\mathbb{R}^n} f(\|x\|) dx_1 \cdots dx_n = \int_0^\infty \int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} f(r) r^{n-1} g dr d\varphi_1 \dots d\varphi_{n-2} d\theta.$$

In this integration we can use Fubini's theorem to split the r integration and the angle integrations. The result is

$$\int_{\mathbb{R}^n} f(\|x\|) dx_1 \cdots dx_n = C \int_0^\infty f(r) r^{n-1} dr,$$

where C is a certain constant.

PROBLEM 10–49. Choose a certain example for f in the above formula to conclude that $C = n\alpha_n$. The result is therefore

$$\boxed{\int_{\mathbb{R}^n} f(\|x\|) dx_1 \cdots dx_n = n\alpha_n \int_0^\infty f(r) r^{n-1} dr.}$$

PROBLEM 10–50. The unit ball $B(0,1)$ can be inscribed in the cube $[-1,1] \times \cdots \times [-1,1]$. The ratio of the volume of the ball to that of the cube is $2^{-n}\alpha_n$. Prove that this ratio is a strictly decreasing function of n. Also prove that the ratio tends to zero as $n \to \infty$. In fact, prove that $\lim_{n\to\infty} \alpha_n = 0$.

H. **The gamma function**

This section is not actually about changing variables in integrals, but instead is a natural conclusion to what we have learned about the number α_n, the volume of the unit ball in $\mathbb{R}^n$. It is somewhat clumsy to have to use two different sorts of expressions for α_n depending on the parity of n, as we have done in the preceding section. We shall soon derive a single formula that works for all n.

The key lies in the introduction of a function which elegantly interpolates the factorial function, which is itself defined only for nonnegative integers. This function is rather artificially called the *gamma* function. The Greek letter gamma has no real significance in this context, just as the Greek letter pi in itself has no particular relation to circles.

DEFINITION. The *gamma function* is the function $(0,\infty) \xrightarrow{\Gamma} \mathbb{R}$ given by the expression

$$\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt, \quad 0 < a < \infty.$$

We remark that $0 < \Gamma(a) < \infty$. The defining integral is "improper" at infinity, but the exponential decay gives convergence; also the integral is improper at zero if $0 < a < 1$, but comparison with t^{a-1} gives convergence.

PROBLEM 10–51. Use an integration by parts to prove that

$$\boxed{\Gamma(a+1) = a\Gamma(a).}$$

It is clear that $\Gamma(1) = 1$, so that Problem 10–51 implies that

$$\boxed{\Gamma(a+1) = a! \qquad \text{if } a = 0, 1, 2, \dots .}$$

Thus we can assert that $\Gamma(a)$, defined for $0 < a < \infty$, *interpolates* the factorial function $(a-1)!$, defined only for $a = 1, 2, 3, \dots$.

PROBLEM 10–52. Use a change of variable to show that

$$\Gamma(a) = 2\int_0^\infty x^{2a-1}e^{-x^2}dx.$$

Conclude that

$$\boxed{\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.}$$

The equation for $\Gamma\left(\frac{1}{2}\right)$ is truly fascinating! Just think: a rather natural interpolation of the factorial function produces such an interesting outcome for the "factorial" of $-1/2$.

Now look at the formula for α_n, as given on p. 10–31. If n is even, the denominator is

$$\frac{n}{2} \cdot \frac{n-2}{2} \cdots 1 \;=\; \left(\frac{n}{2}\right)! \;=\; \Gamma\left(\frac{n}{2}+1\right).$$

If n is odd, it is instead

$$\begin{aligned}\frac{n}{2} \cdot \frac{n-2}{2} \cdots \frac{1}{2} \;&=\; \Gamma\left(\frac{n}{2}+1\right) \Big/ \Gamma\left(\frac{1}{2}\right) \\ &=\; \Gamma\left(\frac{n}{2}+1\right) \Big/ \sqrt{\pi}.\end{aligned}$$

Thus in all cases we have the single formula

$$\boxed{\alpha_n \;=\; \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.}$$

Incidentally, some books write $a! = \Gamma(a+1)$ for *all* $-1 < a < \infty$. With such a definition, then we have the elegant expression

$$\alpha_n \;=\; \frac{\pi^{n/2}}{(n/2)!}.$$

Our derivation of this formula for α_n was a bit clumsy in one sense: it used induction on n and produced two formulas depending on the parity of n. The following exercise produces α_n directly, not relying on induction. It even produces the value for $\Gamma(1/2)$ as a by-product.

PROBLEM 10–53. Consider the n-dimensional Gaussian function

$$e^{-\|x\|^2} = e^{-x_1^2-\cdots-x_n^2}.$$

In the following calculations do not worry about the improperness of any of the integrals. (This is in accord with the "conservation of mass" principle given on p. 9–25.)

a. Use Problems 10–49 and 52 to show that

$$\int_{\mathbb{R}^n} e^{-\|x\|^2} dx = n\alpha_n \Gamma(n/2)/2.$$

b. Use Fubini's theorem to show that

$$\int_{\mathbb{R}^n} e^{-\|x\|^2} dx = \prod_{k=1}^{n} \int_{-\infty}^{\infty} e^{-x_k^2} dx_k.$$

c. Combine the preceding formulas to achieve

$$\alpha_n = (\Gamma(1/2))^n \Big/ \Gamma\left(\frac{n}{2}+1\right).$$

d. Use the case $n = 2$ to show that $\Gamma(1/2) = \sqrt{\pi}$, and thus to obtain the desired formula for α_n.

THE BETA FUNCTION. There is another important special function, which is closely related to the gamma function. These functions deserve attention because they appear so often in applications. Again, the word "beta" has no special significance in this context. Incidentally, the letter B in the definition is to be regarded as an upper case beta.

DEFINITION. The *beta function* is the function $(0,\infty) \times (0,\infty) \xrightarrow{B} \mathbb{R}$ given by the expression

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt.$$

PROBLEM 10–54. Prove that the defining integral for $B(a,b)$ is finite if $a > 0$ and $b > 0$. Prove that $B(b,a) = B(a,b)$. Also prove that

$$B(a,b) \;=\; 2\int_0^{\pi/2} \sin^{2a-1}\theta \cos^{2b-1}\theta d\theta.$$

The next exercise gives a truly wonderful calculation of the beta function in terms of the gamma function.

PROBLEM 10–55. Explain the validity of each step in the following calculation (you may ignore the problems connected with the convergence of the integrals). For any $a > 0$, $b > 0$,

$$\begin{aligned}
\Gamma(a)\Gamma(b) &= \int_0^\infty s^{a-1}e^{-s}ds \cdot \int_0^\infty t^{b-1}e^{-t}dt \\
&= \int_0^\infty \int_0^\infty s^{a-1}t^{b-1}e^{-(s+t)}dsdt \\
&\overset{\text{why?}}{=} \int_0^\infty \int_t^\infty (x-t)^{a-1}t^{b-1}e^{-x}dxdt \\
&\overset{\text{why?}}{=} \int_0^\infty \int_0^x (x-t)^{a-1}t^{b-1}e^{-x}dxdt \\
&\overset{\text{why?}}{=} \int_0^\infty \int_0^1 (x-xy)^{a-1}(xy)^{b-1}e^{-x}xdydx \\
&\overset{\text{why?}}{=} \Gamma(a+b)B(a,b).
\end{aligned}$$

Thus we have found the formula

$$\boxed{B(a,b) \;=\; \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}}.$$

Notice from Problem 10–54 the special value $B(1/2,1/2) = \pi$. Setting $a = b = 1/2$ in the formula we have just derived gives therefore $\Gamma(1/2) = \sqrt{\pi}$. This marks our *third* derivation of the Gaussian integral.

The formula for B enables us to compute many difficult integrals very quickly. It doesn't help in computing *indefinite* integrals such as

$$\int \sin^{12}\theta d\theta;$$

these are sometimes called *incomplete* beta functions. But the *complete* beta function is merely

$$\begin{aligned}\int_0^{\pi/2} \sin^{12}\theta d\theta &= \frac{1}{2}B(13/2, 1/2)\\ &= \frac{\Gamma(13/2)\Gamma(1/2)}{2\Gamma(7)}\\ &= \frac{\frac{11}{2}\cdot\frac{9}{2}\cdot\frac{7}{2}\cdot\frac{5}{2}\cdot\frac{3}{2}\cdot\frac{1}{2}\sqrt{\pi}\cdot\sqrt{\pi}}{2\cdot 6!}\\ &= \frac{11\cdot 7\cdot 3\pi}{2^{11}}.\end{aligned}$$

PROBLEM 10–56. Use the trigonometric version of $B(a,a)$ (Problem 10–54) and the formula $\sin 2\theta = 2\sin\theta\cos\theta$ to derive the *duplication formula* of Legendre:

$$\Gamma(2a) = \frac{2^{2a-1}}{\sqrt{\pi}}\Gamma(a)\Gamma\left(a+\frac{1}{2}\right).$$

PROBLEM 10–57. Use the preceding problem to show that

$$\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right) = \pi\sqrt{2}.$$

There is another interesting formula for the gamma function, which we can now derive rather easily. It is based on the familiar representation of the exponential function,

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n.$$

Now consider the integral representation of $\Gamma(a)$, but instead of e^{-t} in the integrand insert the polynomial $(1-t/n)^n$. The resulting integral that we consider is

$$\Gamma_n = \int_0^n t^{a-1}\left(1-\frac{t}{n}\right)^n dt.$$

Replace t by ns to get

$$\Gamma_n = n^a \int_0^1 s^{a-1} (1-s)^n \, ds.$$

This integral can be evaluated explicitly simply by integration by parts n times. We can actually circumvent that by noting

$$\begin{aligned}
\Gamma_n &= n^a B(a, n+1) \\
&= n^a \frac{\Gamma(a)\Gamma(n+1)}{\Gamma(a+n+1)} \\
&= n^a \frac{\Gamma(a) n!}{(a+n)\cdots(a+1)a\Gamma(a)} \\
&= \frac{n^a n!}{(a+n)\cdots(a+1)a}.
\end{aligned}$$

It is not difficult to show that Γ_n has the limit $\Gamma(a)$ as $n \to \infty$. Thus we obtain

THEOREM. *For* $0 < a < \infty$,

$$\boxed{\Gamma(a) = \lim_{n\to\infty} \frac{n^a n!}{(a+n)\cdots(a+1)a}.}$$

To be completely honest, we need to provide a justification of the equation

$$\lim_{n\to\infty} \Gamma_n = \Gamma(a).$$

There are several ways to prove this, and here is an *ad hoc* one that has the advantage of being elementary. First note that

$$e^{-x} \geq 1 - x \qquad \text{for all } \ 0 \leq x \leq 1.$$

Thus

$$e^{-t} = \left(e^{-t/n}\right)^n \geq \left(1 - \frac{t}{n}\right)^n.$$

Therefore we immediately see that

$$\begin{aligned}
\Gamma(a) &\geq \int_0^n t^{a-1} e^{-t} dt \\
&\geq \int_0^n t^{a-1} \left(1 - \frac{t}{n}\right)^n dt \\
&= \Gamma_n.
\end{aligned}$$

Thus we need to establish some sort of reverse inequality.

One such inequality is this:

$$e^{-t} \leq \left(1 - \frac{t}{n}\right)^n + \frac{t}{n} \qquad \text{for } 0 \leq t \leq n.$$

To prove this define

$$g(t) = e^t \left[\left(1 - \frac{t}{n}\right)^n + \frac{t}{n}\right].$$

We need to show that $g(t) \geq 1$. Since $g(0) = 1$, it will suffice to prove that $g'(t) \geq 0$. But

$$\begin{aligned} g'(t) &= e^t \left[\left(1 - \frac{t}{n}\right)^n + \frac{t}{n} - \left(1 - \frac{t}{n}\right)^{n-1} + \frac{1}{n}\right] \\ &= e^t \left[\frac{t}{n}\left\{1 - \left(1 - \frac{t}{n}\right)^{n-1}\right\} + \frac{1}{n}\right] \\ &> 0. \end{aligned}$$

Now let $\epsilon > 0$ be arbitrary, and choose t_0 sufficiently large that

$$\int_{t_0}^{\infty} t^{a-1} e^{-t} dt \leq \epsilon.$$

We conclude that for all $n > t_0$

$$\begin{aligned} \Gamma(a) &\leq \epsilon + \int_0^{t_0} t^{a-1} e^{-t} dt \\ &\leq \epsilon + \int_0^{t_0} t^{a-1} \left(1 - \frac{t}{n}\right)^n dt + \int_0^{t_0} t^{a-1} \frac{t}{n} dt \\ &\leq \epsilon + \int_0^{n} t^{a-1} \left(1 - \frac{t}{n}\right)^n dt + \frac{t_0^{a+1}}{(a+1)n} \\ &= \epsilon + \frac{t_0^{a+1}}{(a+1)n} + \Gamma_n. \end{aligned}$$

Therefore, for all sufficiently large n,

$$\Gamma(a) < 2\epsilon + \Gamma_n.$$

We conclude that Γ_n has the limiting value $\Gamma(a)$.

PROBLEM 10–58. Use the above theorem to prove

$$\text{Wallis' formula : } \quad \frac{\pi}{2} = \lim_{n\to\infty} \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdots\frac{2n}{2n-1}\cdot\frac{2n}{2n+1}.$$

(This is usually written

$$\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots.)$$

I. **Notation for the Jacobian determinant**

There is a very fine notational device which aids the memory in handling changes of variables. Returning to the theorem of Section G, where $y = \Phi(x)$ denoted the variable change, we have the scale factor

$$\det \Phi'(x).$$

The clever notation for this quantity is

$$\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)}.$$

That is,

$$\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)} = \det(\partial y_i/\partial x_j).$$

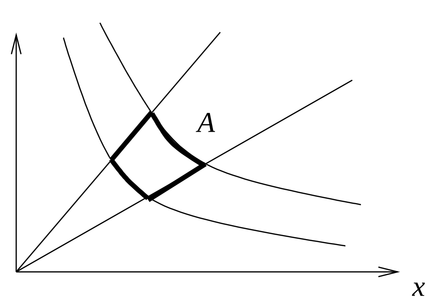
With this notation the formula for changing variables looks almost like that of the one-dimensional case (except for the absolute value sign):

$$\int_{y\in A_2} f(y)dy = \int_{x\in A_1} f(\Phi(x)) \left| \frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)} \right| dx.$$

PROBLEM 10–59. Show that

$$\frac{\partial(x_1,\ldots,x_n)}{\partial(y_1,\ldots,y_n)} = 1 \Big/ \frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)}.$$

EXAMPLE. Let A be the region in the first quadrant of the $x-y$ plane determined by the inequalities $x < y < 3x$ and $2 < xy < 4$. For this region we would likely be interested in a variable transformation of the sort



$$\begin{cases} u &= y/x, \\ v &= xy. \end{cases}$$

Then

$$\begin{aligned} \frac{\partial(u,v)}{\partial(x,y)} &= \det \begin{pmatrix} -yx^{-2} & x^{-1} \\ y & x \end{pmatrix} \\ &= -2y/x. \end{aligned}$$

Notice that we can express x, y explicitly in terms of the new variables as

$$\begin{cases} x &= \sqrt{\frac{v}{u}}, \\ y &= \sqrt{uv}. \end{cases}$$

Therefore the change of variables we obtain is

$$\begin{aligned} dxdy &= \frac{x}{2y} dudv \\ &= \frac{1}{2u} dudv, \end{aligned}$$

and we obtain for example

$$\begin{aligned} \int_A xdxdy &= \int_{\substack{1<u<3 \\ 2<v<4}} \sqrt{\frac{v}{u}} \frac{1}{2u} dudv \\ &= \frac{1}{2} \int_1^3 \int_2^4 \sqrt{v} u^{-3/2} dvdu \\ &= \frac{1}{2} \int_1^3 u^{-3/2} du \cdot \int_2^4 \sqrt{v} dv \\ &= \frac{1}{2} \cdot 2 \left(1 - \frac{1}{\sqrt{3}}\right) \cdot \frac{2}{3}(8 - 2\sqrt{2}) \\ &= \frac{4}{3} \left(1 - \frac{1}{\sqrt{3}}\right)(4 - \sqrt{2}). \end{aligned}$$

PROBLEM 10–60. Calculate the area of the set A in the preceding problem.

(Answer: log 3)

PROBLEM 10–61. For 2×2 matrices A regard $\det A$ as a function from $\mathbb{R}^4$ to $\mathbb{R}$, and then calculate

a. the average of $|\det A|$ over the cube $[0,1]^4$;

b. the average of $|\det A|$ over the cube $[-1,1]^4$.

(Answers: 13/54 and 10/27)

PROBLEM 10–62. Repeat the preceding problem, only with $|\det A|$ replaced with $(\det A)^2$.

(Answers: 7/72 and 2/9)

PROBLEM 10–63. For 3×3 matrices A regard $\det A$ as a function from $\mathbb{R}^9$ to $\mathbb{R}$, and then calculate

a. the average of $(\det A)^2$ over the cube $[0,1]^9$;

b. the average of $(\det A)^2$ over the cube $[-1,1]^9$.

(Answers: 5/144 and 1/9)

My student Stefan Allan recently showed me a problem from page 187 of Niven, Zuckerman, and Montgomery, *An Introduction to the Theory of Numbers, 5th edition, 1991*, and the following problem is based on that one.

PROBLEM 10–64. The *floor function* defined on $\mathbb{R}$ is defined by the expression

$$\lfloor t \rfloor = \text{ largest integer less than or equal to } t.$$

(Sometimes this is designated $[t]$.) Let $a \in \mathbb{R}^n$ have integer coordinates, and evaluate the integral

$$\int_{[0,1]^n} \lfloor a \bullet x \rfloor dx.$$

(HINT: change variables with $x_i = 1 - y_i$.)

PROBLEM 10–65. Under the same conditions as in the preceding problem, show that

$$\int_{[-1,1]^n} \lfloor a \bullet x \rfloor dx = -1.$$