Chapter 2 NOTIONS OF CALCULUS

One of the main methods of the modern approach to mathematics is the recognition of familiar concepts at work in unfamiliar settings. Thus, the ideas of linear algebra, originally introduced for the purpose of solving systems of equations, will be seen also to have relevance in the study of functions. In time it will be seen that many simple concepts of geometry permeate a lot of mathematics. Thus it is important to us, where possible, to try to isolate our concepts and set them in an initially very abstract situation in order to maximize their applicability. Of course, we can't just do that; we must have had some familiarity with the behavior of those concepts. For that we need examples. As we study these examples we can begin to recognize more and more clearly the essence of our concept. This gives rise to a (perhaps) tentative abstract proposal which requires further study of new examples, born out of our generalizations. This procedure, iterated over and over again, may take many generations and the best work of many mathematicians before a clear, precise and satisfactory definition is molded. So it has been with the limit notion, which was implicit in the early 17th century, which was in some sense formulated by Newton and Leibniz in the 18th century, but which did not take a final and comprehensible form until the late 19th century. We shall not try to encompass over two centuries of struggle in a few pages; we shall have to take some short cuts and we shall try (for obvious pedagogical reasons) to avoid the great confusion that is suffered during such development.

The basic technique of calculus is **approximation**. Let us give an illustration of how it goes. The problems of calculus are such that we are required to produce a function that has given properties. There are two aspects to this problem. There is the theoretical aspect: to be assured that there exists a solution to our problem, and the practical one: to describe a procedure for effectively computing that solution. These two aspects are inseparable. In fact, we make a sequence of attempts to solve the problem. If these attempts are good it will provide us with a sequence of functions successively providing better solutions to the problem. Then further study of the general form of these tentative solutions may provide a clue to the accurate solution.

Supposing we have a square of side length one unit in the plane (see Figure 2.1) consider this problem. Find a function f defined on the box which has prescribed values at the vertices and which satisfies this condition: For every point in the box and any rectangle with center at that point, the value of f at that point is the average of the values of f at the vertices of the rectangle. Now we can write these conditions more precisely:

$$f(0,0) = a \qquad f(0,1) = b \qquad f(1,0) = c \qquad f(1,1) = d$$

where a, b, c, d are given numbers. Further, for any (x, y) and (s, t) we must have

$$\begin{aligned} f(x, y) = \tfrac{1}{4}[&f(x-s, y-t) + f(x+s, y-t) + f(x+s, y+t) \\ &+ f(x-s, y+t)] \end{aligned} \tag{2.1}$$

Now, how do we find such a function? We compute, based on the given information, its value at certain points and try to see if we obtain a pattern.

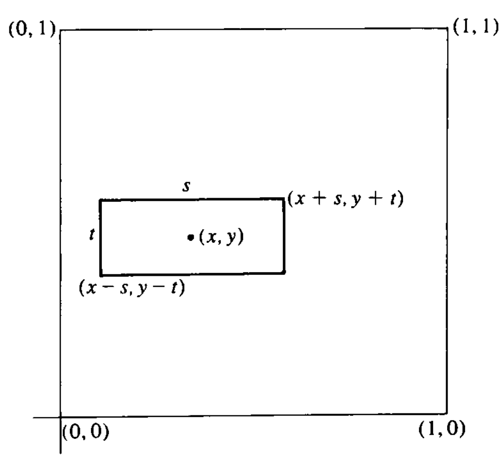

Figure 2.1

First of all, the value at the center of the box is easy to find

$$f(\tfrac{1}{2}, \tfrac{1}{2}) = \tfrac{1}{4}(a + b + c + d)$$

By (2.1) we can compute the value at the center points of the sides:

$(s = \frac{1}{2}, t = 0)$,

$$f(\tfrac{1}{2}, 0) = \tfrac{1}{4}[f(0,0) + f(1,0) + f(1,0) + f(0,0)] = \frac{a+b}{2}$$

Similarly, we obtain the values at all the other center points shown in Figure 2.2. Let us move to more complicated points, for example, the centers of the four squares in Figure 2.2. Since we know the values at all the relevant vertices, we may compute, by (2.1)

$$\begin{aligned} f(\tfrac{1}{4}, \tfrac{1}{4}) &= \tfrac{9}{16}a + \tfrac{3}{16}b + \tfrac{1}{16}c + \tfrac{3}{16}d \\ f(\tfrac{3}{4}, \tfrac{1}{4}) &= \tfrac{3}{16}a + \tfrac{9}{16}b + \tfrac{3}{16}c + \tfrac{1}{16}d \\ f(\tfrac{3}{4}, \tfrac{3}{4}) &= \tfrac{1}{16}a + \tfrac{3}{16}b + \tfrac{9}{16}c + \tfrac{3}{16}d \\ f(\tfrac{1}{4}, \tfrac{3}{4}) &= \tfrac{3}{16}a + \tfrac{1}{16}b + \tfrac{3}{16}c + \tfrac{9}{16}d \end{aligned} \tag{2.2}$$

Now we can see that we can break the given square into 16 squares and compute the values at the centers (points of the form $p/2^3$, $q/2^3$, and so

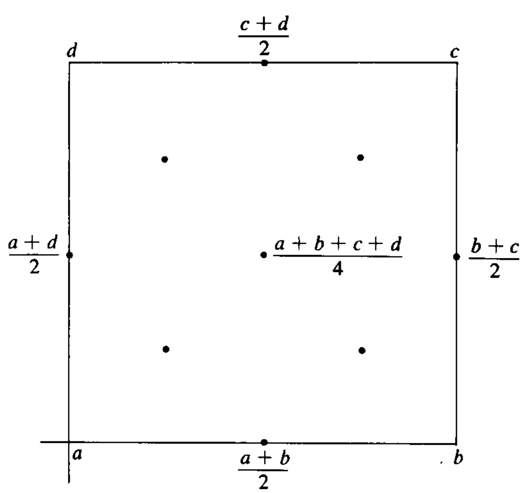

Figure 2.2

forth). We can successively compute the necessary values of f at all points of the form $p/2^n$, $q/2^n$. Since any point in the rectangle has points of this form arbitrarily near it, we surmise that by this tedious procedure, we will be able to approximate the value of f at any point. It is fair to guess then that a solution to our problem *exists* and that we have described a technique for computing its values. If we return to Equations (2.2) (or their successors at the next stage) we may be able to really find a formula for the solution. Now it turns out that Equations (2.2) may be rewritten as

$$f\left(\frac{p}{2^n}, \frac{q}{2^n}\right) = \frac{(2^n - p)(2^n - q)}{2^{2n}} a + \frac{p(2^n - q)}{2^{2n}} b + \frac{pq}{2^{2n}} c + \frac{q(2^n - p)}{2^{2n}} d \tag{2.3}$$

(the case $n = 2$; p $= 0, 1, 2, 3$; $q = 0, 1, 2, 3$). We can show by successively computing the values at centers of squares that (2.3) is valid for all n. Thus rewriting (2.3), we can assert that if (x, y) is of the form $(p/2^n, q/2^n)$ with p and q as integers, then

$$f(x, y) = (1 - x)(1 - y)a + x(1 - y)b + xyc + (1 - x)yd \tag{2.4}$$

Assuming that f is a well-behaved function this must then hold for all points (x, y). Finally, we can show by substituting into the required conditions that (2.4) gives the solution.

Our purpose in the present chapter is to discuss the theoretical concepts which remove the fuzziness in the above discussion. We shall expose the ideas *limit* and *continuity* in the setting of functions of many variables. We shall also present a review of the information from calculus which is necessary to the study of this text.

2.1 Convergence of Sequences

Before proceeding directly to the limit notion, a few words on the notion of a sequence are in order. Let X be any set. A sequence of points in X is an ordered collection $\{x_1, x_2, \ldots, x_n, \ldots\}$ of points in X, one for each positive integer. Another way of saying that is this: a sequence of points in X is given by a function $f: P \to X$, where we denote $f(n)$ by x_n. As a shorthand device we will often denote the sequence $\{x_1, x_2, \ldots, x_n, \ldots\}$ merely by its general term $\{x_n\}$.

Examples

1. $\{1, 2, 3, \ldots, n, \ldots\}$ $\qquad f(n) = n \qquad \{n\}.$

2. $\left\{-1, \frac{1}{2}, -\frac{1}{3}, \ldots, \frac{(-1)^n}{n}, \ldots\right\}$ $\qquad f(n) = \frac{(-1)^n}{n} \qquad \left\{\frac{(-1)^n}{n}\right\}.$

3. $\{10, 10^{1/2}, \ldots, 10^{1/n}, \ldots\}$ $\qquad f(n) = 10^{1/n} \qquad \{10^{1/n}\}.$

A subsequence of a given sequence $\{x_n\}$ is a sequence $\{y_n\}$ extracted from the ordered collection $\{x_1, \ldots, x_n, \ldots\}$. Thus, the collections

4. {odd-numbered x_n's} $= \{x_{2n-1}\}$,

5. {every fifth term in $\{x_n\}$} $= \{x_{5n}\}$,

6. $\{x_{p_n}\}$, where p_n is the nth prime,

7. $\{x_{g(n)}\}$, where g is a strictly increasing function on the positive integers,

are all subsequences of $\{x_n\}$, whereas

8. $\{x_1, x_1, \ldots, x_1, \ldots\}$ is *not* a subsequence.

The above description of subsequence is a bit vague. The phrase "extracted from" is picturesque but not too meaningful. Is the sequence

$$\{x_5, x_4, x_3, x_2, x_1, x_{10}, x_9, \ldots, x_6, \ldots, x_{5n}, x_{5n-1}, x_{5n-2}, x_{5n-3}, x_{5n-4}, \ldots\}$$

a subsequence of $\{x_n\}$? It isn't clear from the preceding paragraph. However, we should draw the line and exclude such new sequences. The essence of a subsequence will be that it consists of some of the x_n's, infinitely many of them, and collected *in the same order.* Now, to be really exacting, our notion of sequence itself is imprecise; we seem to have failed to say what it is. "An ordered collection" is not very satisfactory. We have already elaborated on that: "a sequence ... is given by a function $f: P \to X$ ... ." Yet, it is given by "...," but what is it? It turns out that this line of metaphysical questioning bogs down, and is in fact irrelevant. We have already found something which completely describes the sequence (the function $f: P \to X$), so why not *define* a sequence just as such a function? Indeed, when we do so, it becomes very easy to also define a subsequence.

Definition 1. Let X be a set. A **sequence** in X is a function $f: P \to X$. A **subsequence** of this sequence is another sequence $h: P \to X$, where $h = f \circ g$ and g is a strictly increasing function from P to P.

Thus, if $\{x_1, \ldots, x_n, \ldots\}$ is a sequence, this is in fact just another way of writing the function f, $f(n) = x_n$. If $f \circ g$ is a subsequence, we can enumerate it as $\{x_{g(1)}, x_{g(2)}, \ldots, x_{g(n)}, \ldots\}$.

The above definition is an illustration of a standard mathematical procedure of defining things. A concept is, mathematically, an object with such and such properties. Once we have stated the properties which we feel describe the concept, there is no need to further inquire what the object is; we simply *define it by those properties.* We now introduce the notion of convergence of a sequence of numbers (which we may take as complex numbers).

Definition 2. Let $\{z_n\}$ be a sequence of complex numbers. We say that the sequence **converges** if there is a $z \in C$ such that to every positive number $\varepsilon > 0$, there corresponds an integer N such that $|z_n - z| < \varepsilon$ for $n \geq N$. In this case we say $\{z_n\}$ **converges to** z, written $\lim_{n \to \infty} z_n = z$ or $\lim z_n = z$ or $z_n \to z$.

Said another way, the sequence $f: P \to C$ converges to $z \in C$ if, given any disk centered at z, the range of f on all but finitely many integers lies in that disk (see Figure 2.3).

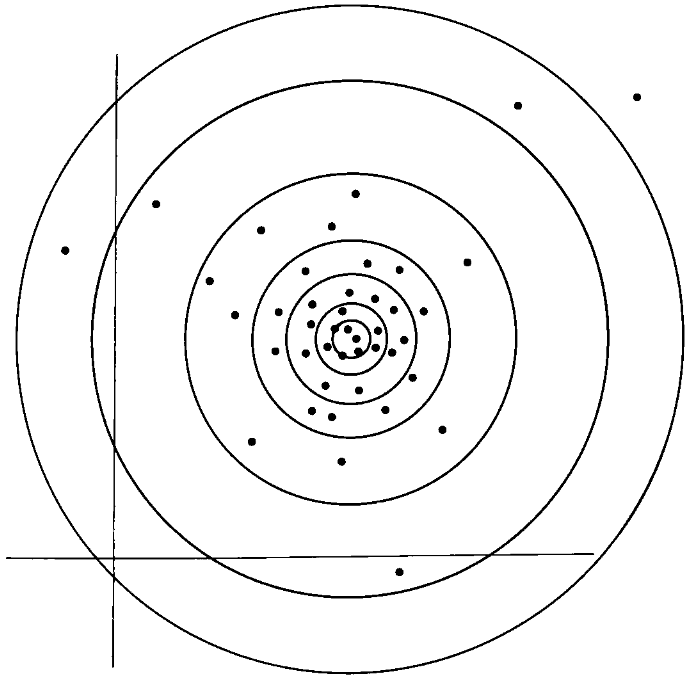

Figure 2.3

The following proposition asserts that a sequence cannot converge to more than one point, and gives necessary conditions for convergence, without reference to the limit point.

Proposition 1. *Suppose* $\lim z_n = z$.

(i) *if also* $\lim z_n = w$, *then* $w = z$,

(ii) *the sequence is bounded, that is, there is an* $M \geq 0$ *such that* $|z_n| \leq M$ *for all* n,

(iii) (*Cauchy criterion*) *for every* $\varepsilon > 0$, *there is an* $N \geq 0$ *such that* $|z_n - z_m| < \varepsilon$ *for all* $n, m \geq N$.

Proof.

(i) By the hypotheses, given $\varepsilon > 0$, there are N_1, N_2 such that $|z_n - z| < \varepsilon$ for $n \geq N_1$, $|z_n - w| < \varepsilon$ for $n \geq N_2$. Thus,

$$|z - w| \leq |z - z_{N_1+N_2}| + |z_{N_1+N_2} - w| < 2\varepsilon$$

Since the inequality $|z - w| \leq 2\varepsilon$ holds for all $\varepsilon > 0$, we must have $|z - w| \leq 0$, or $z = w$.

(ii) Taking $\varepsilon = 1$, there is an integer N such that $|z_n - z| < 1$ for $n \geq N$. Let $M = \max\{|z|, |z_1|, |z_2|, \ldots, |z_N|\} + 1$. Then if $n \leq N$, certainly $|z_n| \leq M$. If $n > N$, $|z_n| \leq |z_n - z| + |z| \leq 1 + |z| \leq M$.

(iii) Let $\varepsilon > 0$ be given. There is an integer N such that $|z_n - z| < \varepsilon/2$ for $n \leq N$. Thus, if $n, m \geq N$, we have

$$|z_n - z_m| < |z_n - z| + |z_m - z| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Condition (iii) is called a **criterion** for it implies convergence, as we shall see below. Notice that (ii) does not imply convergence. The sequence $\{(-1)^n\}$ is clearly bounded, but does not converge (it doesn't even satisfy the Cauchy criterion: $|(-1)^n - (-1)^{n+1}| = 2$ for *all* n).

Examples

9. $\lim (1/n) = 0$. Let $\varepsilon > 0$ be given, and choose the integer N so that $N > \varepsilon^{-1}$. Then, for $n \geq N$,

$$\left|\frac{1}{n} - 0\right| = n^{-1} \leq N^{-1} < \varepsilon$$

10. $\lim (i^n/n) = 0$. The proof is the same (see also Problem 2).

11.

$$\lim \frac{2}{1 + (1/n)} = 2$$

Let $\varepsilon > 0$ be given. Now,

$$\left|\frac{2}{1 + (1/n)} - 2\right| = 2\left|\frac{1}{1 + (1/n)} - 1\right| = 2\left|\frac{n - (n + 1)}{n + 1}\right| = \frac{2}{n + 1}$$

Thus, we need only take $N > 2\varepsilon^{-1}$ to verify the condition for convergence.

12.

$$\lim \frac{n}{n + 1} = 1 \quad \left|\frac{n}{n + 1} - 1\right| = \frac{1}{n + 1} < \varepsilon \qquad \text{if } n > \varepsilon^{-1}$$

13. $\lim h^n = 0$ for $0 \leq h < 1$. Since $h < 1$, there is an integer $K \geq 1$ such that $h \leq K/(K + 1)$. Since the sequence $n/(n + 1)$ is increasing, we have $h \leq n/(n + 1)$, all $n \geq K$. Now, we shall show by mathematical induction that $h^n \leq K/n$ for all n. The case $n = 1$ is clear since $K \geq 1$. Now, using the nth inequality we obtain the $(n + 1)$th:

$$h^{n+1} = h \cdot h^n \leq \frac{n}{n + 1} \cdot \frac{K}{n} \leq \frac{K}{n + 1}$$

Thus, if $\varepsilon > 0$ is given, let $N \geq K\varepsilon^{-1}$. For $n \geq N$, $|h^n - 0| = h^n \leq (K/n) < \varepsilon$.

The study of the convergence of complex sequences is easily reduced to that of real sequences by the following fact. A complex sequence converges if and only if the real and imaginary parts both converge.

Proposition 2. *Let $\{z_n\}$ be a sequence of complex numbers, $z_n = x_n + iy_n$. $\lim z_n = z = x + iy$ if and only if $\lim x_n = x$ and $\lim y_n = y$.*

Proof. Suppose $\lim z_n = z$. Then, given $\varepsilon > 0$ there is an N such that for $n \geq N$, $|z_n - z| < \varepsilon$. Since

$$|x_n - x| \leq |z_n - z| \quad \text{and} \quad |y_n - y| \leq |z_n - z|$$

we also have

$$|x_n - x| < \varepsilon \quad \text{and} \quad |y_n - y| < \varepsilon \qquad \text{for } n \geq N$$

so $\lim x_n = x$ and $\lim y_n = y$.

Conversely, given $\varepsilon > 0$, there are N_1, N_2 such that $n \geq N_1$ implies $|x_n - x| < \varepsilon/\sqrt{2}$, implies $|y_n - y| < \varepsilon/\sqrt{2}$. Then

$$|z_n - z| = [|x_n - x|^2 + |y_n - y|^2]^{1/2} < \left(\frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2}\right)^{1/2} = \varepsilon$$

We have so far been considering questions of this form: Given the sequence $\{z_n\}$ and the number z, is $\lim\limits_{n\to\infty} z_n = z$? A deeper problem is this: Given the sequence $\{z_n\}$, find if possible, a number z such that $\lim z_n = z$. A solution of such problems requires a more profound understanding of the real number system than we have so far needed. A question of *existence* is now involved. To resolve such questions we have to have explicit knowledge that there are many real numbers, whereas until now we have made use only of the existence of the numbers 0 and 1. The explicit knowledge desired here is that provided by the **axiom of the least upper bound** which roughly states that there are no gaps in the line or real numbers. A set S of real numbers is said to be **bounded from above** if there is a number M such that $x \leq M$ for every $x \in S$. If S is a set which is bounded above, it is conceivably useful to know the smallest number M which will serve as an upper bound. We shall refer to such a least upper bound of the set S by $\sup S$ (and $\inf S$ will denote the greatest lower bound if it exists). The axiom of the least upper bound asserts that for a set S which is bounded above, $\sup S$ *exists.* We shall state this same axiom in terms of sequences because in that form it is more appropriate to our present context.

Theorem 2.1. *Let $\{x_n\}$ be a decreasing sequence of real numbers; that is, $x_n \geq x_{n+1}$ for all n. If the set $\{x_n\}$ is bounded, the sequence converges.*

We have called this a theorem since it can be deduced from the axiom of the least upper bound (see Problem 1), which we can take as a *defining property of the real number system.* A consequence of this fact of existence for the real numbers is the fact that the Cauchy criterion (see Proposition

1(iii)) *is* a criterion for convergence. The proof goes like this: first we find a subsequence of the given sequence which is decreasing. An easy consequence of the Cauchy criterion is that the sequence is bounded. Thus, by Theorem 2.1 this subsequence has a limit x. It now follows from the Cauchy criterion that the full given sequence also has the limit x.

Theorem 2.2. *Let $\{x_n\}$ be a Cauchy sequence of real numbers. That is, for every $\varepsilon > 0$, there is an integer N such that $|x_n - x_m| < \varepsilon$ whenever $n, m \geq N$. Then there is an x such that* $\lim x_n = x$.

Proof. First, a Cauchy sequence is bounded. Let $\varepsilon = 1$; there is an N such that $|x_n - x_m| < 1$ for $n, m \geq N$. Then $M = \max\{|x_1|, \ldots, |x_N|\} + 1$ is a bound for $\{x_n\}$. Let $u_k = \sup\{x_n : n \geq k\}$. Clearly, the u_k are decreasing, and $u_k \geq -M$ for all k, so by Theorem 2.1 the sequence u_k converges, say $\lim u_k = u$. We shall show that also $\lim x_n = u$.

Let $\varepsilon > 0$. There are N_1, N_2 such that

$$|x_n - x_m| < \frac{\varepsilon}{3} \qquad \text{for } n, m \geq N_1$$

$$|u_k - u| < \frac{\varepsilon}{3} \qquad \text{for } k \geq N_2$$

Let $N = N_1 + N_2$. Since $u_N = \sup\{x_n : n \geq N\}$, there is an $n_0 \geq N$ such that $x_{n_0} \leq u_N + (\varepsilon/3)$. Then, combining all these inequalities, we have for $n \geq N$,

$$|x_n - u| \leq |x_n - x_{n_0}| + |x_{n_0} - u_N| + |u_N - u| < \varepsilon$$

Because of Proposition 2 that a complex sequence converges if its real and imaginary parts do, we can deduce the same theorem for complex sequences.

Corollary. *A Cauchy sequence of complex numbers converges.*

Proof. Problem 3.

● EXERCISES

1. What are the limits (when they exist) of these sequences:

(a) $\{n^2 - 4\}$ (c) $\left\{\dfrac{(-1)^n}{n^2}\right\}$

(b) $\{(n^2 - 4)^{-1}\}$ (d) $\{(-1)^n - (-1)^{n+1}\}$

(e) $\left\{\dfrac{n^2-2}{n^2+2}\right\}$ (g) $\left\{\dfrac{1}{n}\sin(n)\right\}$

(f) $\left\{n\sin\left(\dfrac{1}{n}\right)\right\}$ (h) $\left\{\dfrac{3n^3+2n^2+n+1}{n^3+1}\right\}$

2. If $\{x_n\}$ is a convergent sequence, then $\lim\limits_{n\to\infty}(x_{n+1}-x_n)=0$. Is this a criterion for convergence?
3. Suppose $\lim x_n = z$. Let $\{y_n\}$ be the sequence $\{x_{k+1}, x_{k+2}, \ldots\}$. Show that $\lim y_n = z$ also.
4. Let $\{s_n\}$, $\{t_n\}$ be two convergent sequences. Show that if they are convergent sequences with the same limit, then $\lim(s_n - t_n) = 0$. Is the converse true?
5. What is $\lim ((n + 1)^{1/2} - \sqrt{n})$?
6. Show that $\lim z_n = 0$ if and only if $\lim|z_n| = 0$.

• PROBLEMS

1. Let $\{x_n\}$ be a decreasing sequence of real numbers. Prove, using the least upper bound axiom, that if $\{x_n\}$ is bounded, it is convergent. Deduce also that an increasing bounded sequence is convergent.
2. Suppose that $\lim z_n = 0$ and $\{c_n\}$ is a bounded sequence. Prove that $\lim\limits_{n\to\infty} c_n z_n = 0$. If $\{z_n\}$ is a convergent sequence and $\{c_n\}$ is a bounded sequence, is $\{c_n z_n\}$ convergent? or bounded?
3. Deduce from the fact that Cauchy sequences of real numbers converge, that a Cauchy sequence of complex numbers is convergent.
4. Let $\{z_n\}$ be a sequence of complex numbers and $\{c_n\}$ a sequence of positive real numbers such that $|z_n| \le c_n$ for all $n \ge N_0$. Prove that if $\lim c_n = 0$, also $\lim z_n = 0$.
5. Suppose $\lim z_n = z$. Let $\{y_n\}$ be a subsequence of $\{z_n\}$. Then $\lim y_n = z$ also.
6. Let $\{s_n\}$, $\{x_n\}$, $\{t_n\}$ be three sequences of real numbers. Suppose that $s_n \le x_n \le t_n$ for all n.
   (a) Show that if $\lim s_n = \lim t_n = c$, then also $\lim x_n = c$.
   (b) Show that if $\lim s_n = c$ and $\lim(t_n - s_n) = 0$, then also $\lim x_n = c$.
7. (a) Let $1 > \delta > h > 0$. Show that there is an integer K such that for $n \ge K$, $(n/(n+1))\delta > h$.
   (b) Let $1 > h > 0$. Show that $\lim nh^n = 0$.
8. Suppose $\lim z_n = z$, $\lim w_n = w$. Show that
   (a) $\lim|z_n| = |z|$.
   (b) $\lim(z_n + w_n) = \lim z_n + \lim w_n$.
   (c) $\lim z_n w_n = \lim z_n \cdot \lim w_n$.

2.2 Series

A sequence may be formed term-by-term by adding a little bit to each term. In this case the limit, if it exists, will be an infinite sum. Such sequences are probably the most important kind, for in practice what we usually know about a sequence is the difference between two successive terms. This sequence is given to us as the sequence of sums of these differences.

Let $\{z_n\}$ be a given sequence. The **series** formed of this sequence is the sequence of sums

$$z_1, z_1 + z_2, \ldots, z_1 + \cdots + z_n = \sum_{i=1}^{n} z_i, \ldots$$

The **series converges** if the sequence of sums $\sum_{i=1}^{n} z_i$ converges; in this case the limit is denoted by $\sum_{i=1}^{\infty} z_i$.

Example

14. The **geometric series** $\sum_{n=0}^{\infty} z^n$. Let $S_N = \sum_{n=0}^{N} z^n$. Then

$$S_{N+1} = 1 + z + \cdots + z^n + z^{n+1} = S_N + z^{N+1}$$

Notice also that

$$S_{N+1} = 1 + z(1 + z + \cdots + z^n) = 1 + zS_N$$

These two equations give us the general term of the sequence explicitly:

$$1 + zS_N = S_N + z^{N+1}$$

or

$$S_N = \frac{1 - z^{N+1}}{1 - z} \qquad (z \neq 1) \tag{2.5}$$

Now, if $|z| < 1$, then $\lim S_N = (1 - z)^{-1}$, for

$$\left| S_N - \frac{1}{1 - z} \right| = \left| \frac{z^{N+1}}{1 - z} \right| = \frac{1}{|1 - z|} |z|^{N+1}$$

and $\lim |z|^N = 0$ (Example 13). So given $\varepsilon > 0$, we find N_0 so that $|z|^{n+1} < (|1 - z|)\varepsilon$ for $n \geq N_0$, and thus

$$\left| S_N - \frac{1}{1-z} \right| < \varepsilon \qquad \text{for } N \geq N_0$$

Notice that we cannot immediately determine whether or not the geometric series converges for $|z| \geq 1$ (of course, for $z = 1$, $S_N = N$, so the series diverges). In fact, the geometric series does *not* converge for $|z| \geq 1$, by application of the following proposition.

Proposition 3. *If $\sum_{n=0}^{\infty} z_n$ exists, then the general term must converge to zero; that is,* $\lim z_n = 0$.

Proof. Let $\{s_n\}$ be the sequence of partial sums. By hypothesis $\lim s_n$ exists. Let $\varepsilon > 0$ be given. There is an N such that for $n, m \geq N$, $|s_n - s_m| < \varepsilon$. Thus, for $n \geq N$, $|z_n| = |s_n - s_{n+1}| < \varepsilon$.

Thus,

$$\sum_{n=0}^{\infty} z^n = \begin{cases} \dfrac{1}{1-z} & \text{for } |z| < 1 \\ \text{diverges} & \text{for } |z| \geq 1 \end{cases} \tag{2.6}$$

For if $|z| > 1$, the general term is $|z^n|$. By Example 13, $\lim(1/|z|)^n = 0$, so $\{|z|^n\}$ gets arbitrarily large and does not converge. If $|z| = 1$, $|z|^n = 1$, and the sequence $\{1\}$ does not converge to zero.

By the way, the condition $\lim z_n = 0$ is *not sufficient* for the convergence of the series $\sum z_n$ as the following examples show.

Examples

15. Certainly the series $1 + 1 + \cdots + 1 + \cdots$ diverges. But we can rewrite this as

$$1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \cdots + \underbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}_{n} + \frac{1}{n+1} + \cdots$$

Here the general term tends to zero.

16. $\sum_{n=1}^{\infty} 1/n$ diverges. Let $s_N = \sum_{n=1}^{N} 1/n$. Then

$$S_{2N} - S_N = \frac{1}{N+1} + \cdots + \frac{1}{2N} \geq N\frac{1}{2N} = \frac{1}{2}$$

Thus, $\{S_N\}$ is not a Cauchy sequence.

The sum of a series of positive numbers is particularly easy to work with. For if $\{c_n\}$ is a sequence of positive numbers, then the series $\{\sum_{k=1}^{n} c_k\}$ is an increasing sequence, so by Theorem 2.1 (as rewritten in Problem 1) this sequence converges if and only if it is bounded.

Proposition 4. *Let $\{c_n\}$ be a sequence of nonnegative numbers. The following assertions are equivalent.*

(i) $\sum c_k$ *converges.*
(ii) $\{\sum_{k=1}^{n} c_k\}$ *is bounded.*
(iii) *For each $\varepsilon > 0$, there is an N such that for all $m \geq N$,*

$$\sum_{k=N}^{m} c_k < \varepsilon$$

The proof of the equivalence of (i) and (ii) is essentially given in the preceding paragraph. Part (iii) is just the Cauchy criterion restated for positive series (see Problem 11).

Examples

17. $\sum_{k=1}^{\infty} 1/n!$ converges. For $n! \geq 2^{n-1}$ for all n, so

$$\frac{1}{n!} < \frac{1}{2^{n-1}}$$

and thus for all N,

$$\sum_{n=1}^{N} \frac{1}{n!} \leq \sum_{n=0}^{N-1} \frac{1}{2^n} < 2$$

by (2.5).

18.

$$\sum_{n=1}^{\infty}\left(\frac{1}{n}-\frac{1}{n+\cos(1/n)}\right)$$

converges. For

$$\frac{1}{n}-\frac{1}{n+\cos(1/n)}\leq\frac{1}{n}-\frac{1}{n+1}$$

Thus, for all N

$$\sum_{n=1}^{N}\left(\frac{1}{n}-\frac{1}{n+\cos(1/n)}\right)\leq\sum_{n=1}^{N}\left(\frac{1}{n}-\frac{1}{n+1}\right)=1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\cdots+\frac{1}{N}-\frac{1}{N+1}<1$$

There is no such simple criterion as Proposition 4 for arbitrary series of complex (or real) numbers, and the question of convergence as well as computation of a limit can become extremely subtle. However, if for a given series the series formed of the absolute values converges, the situation is considerably clarified. Ordinarily we shall discuss the convergence of a series only in the happy circumstance that the corresponding series of absolute values converges.

Proposition 5. *Let $\{c_k\}$ be a sequence of complex numbers. If $\sum |c_n|$ converges, $\sum c_k$ also converges.*

Proof. Let t_n be the sequence of partial sums of $\sum |c_k|$ and s_n the partial sums of $\sum c_n$. Notice, for $m>n$

$$|s_m-s_n|=\left|\sum_{k=n+1}^{m}c_k\right|\leq\sum_{k=n+1}^{m}|c_k|=t_m-t_n$$

Thus, if $\{t_n\}$ is a Cauchy sequence, so also is $\{s_n\}$.

Definition 3. Let $\{c_k\}$ be a sequence of complex numbers. $\sum c_n$ is **absolutely convergent,** if $\sum |c_n|$ converges. If $\sum |c_n|$ diverges, but $\sum c_n$ converges, we say $\sum c_n$ is **conditionally convergent**.

There are such things as conditionally convergent sequences. In fact, $\sum_{n=1}^{\infty}(-1)^n/n$ converges. But as we have seen in Example 16 the series $\sum_{n=1}^{\infty}1/n$ of absolute values is divergent. It is easy to see that $\sum_{n=1}^{\infty}(-1)^n/n$

converges. Let $\{s_n\}$ be the sequence of partial sums. Then the subsequence $\{s_{2n}\}$ is decreasing, and bounded below by s_1, and the subsequence $\{s_{2n+1}\}$ is increasing, and bounded above by s_2. Thus, both these subsequences converge. Since

$$|s_{2n+1} - s_{2n}| < \frac{1}{n+1}$$

they have the same limit. It is easy to deduce that the full sequence also converges to that common limit. Here is the proof in a more general case (known as **Leibniz's theorem**).

Proposition 6. *Let $\{c_n\}$ be a decreasing sequence of positive numbers such that* $\lim c_n = 0$. *Then $\sum (-1)^n c_n$ converges.*

Proof. Let $s_n = \sum_{k=1}^{n} (-1)^k c_k$. We consider the sequences of even and odd partial sums separately. The sequence $\{s_{2n}\}$ is decreasing, since

$$s_{2(n+1)} - s_{2n} = c_{2n+2} - c_{2n+1} \leq 0$$

Similarly, the sequence of odd partial sums $\{s_{2n+1}\}$ is increasing. Furthermore these sequences are bounded, for, given any n,

$$s_1 \leq s_{2n+1} = s_{2n} - c_{2n+1} \leq s_{2n} \leq s_2$$

so $\{s_{2n}\}$ is bounded below by s_1 and above by s_2. The same is true for the sequence $\{s_{2n+1}\}$. Thus, by Theorem 2.1 $\lim_{n\to\infty} s_{2n} = s$, $\lim_{n\to\infty} s_{2n+1} = s'$ both exist. Furthermore, $s' - s = \lim s_{2n+1} - \lim s_{2n} = \lim(s_{2n+1} - s_{2n}) = \lim(c_{2n+1}) = 0$, so $s' = s$. Since both sequences, of odd partial sums and even partial sums converge and have the same limit, the whole sequence also converges to that limit.

Notice that this argument does not give any hint as to the value of $\sum (-1)^n/n$. Outside of the case of Proposition 4, there is no positive assertion that can be made about conditionally convergent series. In fact, they tend to behave very badly, as the following illustrative example shows.

Example

19. The sequence

$$\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \cdots + \underbrace{\frac{1}{2n} + \cdots + \frac{1}{2n}}_{2n} + \cdots$$

is the same as $1 + 1 + \cdots + 1 + \cdots$ and thus diverges. Since the general term is decreasing to zero, by Leibniz's theorem this series is conditionally convergent:

$$S = \frac{1}{2} - \frac{1}{2} + \frac{1}{4} - \frac{1}{4} + \frac{1}{4} - \frac{1}{4} + \cdots + \underbrace{\frac{1}{2n} - \frac{1}{2n} + \cdots - \frac{1}{2n}}_{n \text{ times}} + \cdots$$

The sequence of partial sums is

$$\frac{1}{2}, 0, \frac{1}{4}, 0, \frac{1}{4}, 0, \ldots, \frac{1}{2n}, 0, \ldots$$

and thus obviously converges to zero. However, we may now rearrange terms of the series so that it no longer converges! Consider the same series where in each group we first add the positive terms and then the negative terms:

$$\frac{1}{2} - \frac{1}{2} + \frac{1}{4} + \frac{1}{4} - \frac{1}{4} - \frac{1}{4} + \cdots + \underbrace{\frac{1}{2n} + \cdots + \frac{1}{2n}}_{n} \underbrace{- \frac{1}{2n} - \cdots - \frac{1}{2n}}_{n} + \cdots \tag{2.7}$$

The corresponding sequence of partial sums is

$$\frac{1}{2}, 0, \frac{1}{4}, \frac{1}{2}, \frac{1}{4}, 0, \ldots, \frac{n-1}{2n}, \ldots, \frac{1}{2n}, 0, \ldots$$

Thus, there is a subsequence: $\{\frac{1}{2}, \frac{1}{2}, \ldots\}$ and another: $\{0, 0, \ldots\}$ so we cannot have convergence of (2.7). We leave to the student (Exercise 9) to show that it can be further rearranged so that it once again converges, but this time to one!

No such foolishness holds for absolutely convergent series. We may attempt to sum the series in any way we please. If we arrive at a limit, it is the sum. In fact, if $\sum c_n$ is an absolutely convergent series we may sum first the positive terms, and then the negative terms; and $\sum c_n$ is the sum of these two sums. We conclude this section with the proof of these facts.

Proposition 7. *Let $\sum c_n$ be an absolutely convergent series of real numbers.*

(i) *Let*

$$c_k{}^+ = \begin{cases} c_k & \text{if } c_k \geq 0 \\ 0 & \text{if } c_k \leq 0 \end{cases} \qquad c_k{}^- = \begin{cases} -c_k & \text{if } c_k \leq 0 \\ 0 & \text{if } c_k \geq 0 \end{cases}$$

Then the sums $\sum c_k^+$, $\sum c_k^-$ converge and $\sum c_k = \sum c_k^+ - \sum c_k^-$.

(ii) (*Rearrangement.*) *Let g be a one-to-one mapping of the positive integers onto the positive integers. Then $\sum c_{g(n)} = \sum c_n$.*

(iii) (*Regrouping.*) *Let h be any strictly increasing function from P into P. Let*

$$d_n = \sum_{k=h(n-1)}^{h(n)} c_k$$

Then $\sum d_n = \sum c_n$.

Proof.

(i) Since the sequence $\{\sum_{k=1}^{n} |c_k|\}$ is bounded by absolute convergence, and

$$\sum_{k=1}^{n} |c_k| \geq \sum_{k=1}^{n} c_k{}^+, \sum_{k=1}^{n} c_k{}^-$$

the sequences $\sum_{k=1}^{n} c_k{}^+$, $\sum_{k=1}^{n} c_k^-$ are also increasing and bounded. Thus they converge to, say s, t respectively, by Theorem 2.1. We have to show that $\sum c_n = s - t$. Let $\varepsilon > 0$. Then there are N_1, N_2 such that for $n \geq N_1$,

$$\left| \sum_{k=1}^{n} c_k^+ - s \right| < \frac{\varepsilon}{2},$$

and for $n \geq N_2$,

$$\left| \sum_{k=1}^{n} c_k^- - t \right| < \frac{\varepsilon}{2}.$$

Then for $n \geq \max(N_1, N_2)$,

$$\left| \sum_{k=1}^{n} c_k - (s-t) \right| = \left| \sum_{k=1}^{n} c_k^+ - \sum_{k=1}^{n} c_k^- - (s-t) \right| < \varepsilon$$

(ii) Let g be a one-to-one map of P onto P. Then g^{-1} is defined and also maps P onto P into a one-to-one fashion. For each n, let $N_n = \max(g(1), \ldots, g(n))$.

Then

$$\sum_{k=1}^{n} |c_{g(k)}| \le \sum_{k=1}^{N_n} |c_k| \le \sum |c_k|$$

for all n, so the series $\sum c_{g(k)}$ is absolutely convergent. Similarly,

$$\sum_{k=1}^{n} c_{g(k)}^{+} \le \sum_{k=1}^{N_n} c_k^{+} \le \sum c_k^{+} \text{ and } \sum_{k=1}^{n} c_{g(k)}^{-} \le \sum c_k^{-}$$

for all n, so we have $\sum c_{g(k)}^{+} \le \sum c_k^{+}$, $\sum c_{g(k)}^{-} \le \sum c_k^{-}$. Applying the same reasoning but reversing the roles of the two series, we obtain the reverse inequalities so that in fact, $\sum c_{g(k)}^{+} = \sum c_k^{+}$ and $\sum c_{g(k)}^{-} = \sum c_k^{-}$. Thus, by part (i) we obtained the desired equality; that is, $\sum c_{g(k)} = \sum c_k$.

Part (iii) is actually true for any convergent series. Let $\sum c_n = c$, and the strictly increasing function h be given. Notice that $h(n) \ge n$ for all n. For $\varepsilon > 0$, there is an N such that

$$\left| \sum_{k=1}^{n} c_k - c \right| < \varepsilon$$

for all $n \ge N$. Thus, for

$$n \ge N, \sum_{k=1}^{n} d_n = \sum_{k=1}^{n} \sum_{j=h(k-1)}^{h(k)} c_j = \sum_{j=1}^{h(n)} c_j$$

and $h(n) \ge N$, so that

$$\left| \sum_{k=1}^{n} d_n - c \right| \le \left| \sum_{j=1}^{h(n)} c_j - c \right| < \varepsilon$$

• EXERCISES

7. Show that

$$\sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right)$$

converges.

8. What is

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{a_n}$$

where $a_{2n} = 2^n$, $a_{2n+1} = 3^n$?

9. Rearrange the series

$$\frac{1}{2}-\frac{1}{2}+\frac{1}{4}-\frac{1}{4}+\frac{1}{4}-\frac{1}{4}+\cdots+\underbrace{\frac{1}{2n}-\frac{1}{2n}+\cdots-\frac{1}{2n}}+\cdots$$

so it has the sum one.

10. Can the series $\sum (-1)^n/n$ be rearranged so as to have sum 10,000?

• PROBLEMS

9. Suppose $\sum z_n = z$ and $\sum w_n = w$. Show that $\sum (z_n + w_n) = z + w$.
10. Suppose
    (a) $\sum z_n$ and $\lim w_n$ exist. Does $\sum z_n w_n$ exist?
    (b) $\sum z_n$ and $\sum w_n$ exist. Does $\sum z_n w_n$ exist?
11. Prove that $\sum z_n$ converges if and only if for all $\varepsilon > 0$, there exists an $N > 0$ such that

$$\left|\sum_{k=N+1}^{n} z_k\right| < \varepsilon \qquad \text{for all } n \geq N$$

Deduce that Proposition 4(iii) is true.

2.3 Tests for Convergence

Since the theory of series is so important and the definition of convergence unwieldy, there has developed a large collection of tests (or criteria) for convergence which are more or less easy to apply in the relevant cases. We have already given some criteria for convergence.

(1) Cauchy criterion: $\sum c_n$ converges if and only if for every $\varepsilon > 0$, there is an integer N such that $|c_{n+1} + \cdots + c_m| < \varepsilon$ for all $m \geq n \geq N$.
(2) If the sequence $\{c_n\}$ decreases to zero, then $\sum (-1)^n c_n$ converges.
(3) If the sequence $\{c_n\}$ is nonnegative, $\sum c_n$ converges if and only if the sequence $\{\sum_{k=1}^{n} c_k\}$ of partial sums is bounded.

The last condition, which can be considered as a condition for absolute convergence, gives rise to the following criterion which is the basic one. The idea is to compare a given series with a known convergent one (if we suspect that it converges) or to a known divergent one (if we suspect that it diverges).

Examples

20. $\sum 1/n!$ converges, as we have seen in Example 17. There, we noticed that $1/n! \le 2^{-n+1}$, and since $\sum 2^{-n+1}$ is convergent, so is $\sum 1/n!$. For

$$\sum_{n=1}^{N}\left(\frac{1}{n!}\right) \le \sum_{n=1}^{N}\frac{1}{2^{n-1}} < \sum_{n=1}^{\infty}\frac{1}{2^{n}} \qquad \text{for all } N$$

21.

$$\sum_{n=1}^{\infty}\sin\left(\frac{5}{n}\right)$$

diverges. For if x is small enough, $\sin x \ge x/2$. Thus, there is an N such that if $n \ge N$,

$$\sin\left(\frac{5}{n}\right) \ge \frac{5}{2n}$$

and thus for $m \ge N$,

$$\sum_{n=1}^{m}\sin\left(\frac{5}{n}\right) = \sum_{n=1}^{N}\sin\left(\frac{5}{n}\right) + \frac{5}{2}\sum_{N+1}^{m}\frac{1}{n}$$

But we can make the last sum as large as we please by taking m large enough. Thus, $\sum_{n=1}^{m}\sin(5/n)$ is not bounded, and so it is not convergent.

22.

$$\sum_{n=0}^{\infty}\frac{1}{(1+i)^n}$$

is absolutely convergent. For $|1+i| = \sqrt{2}$, so for any m,

$$\sum_{n=0}^{m}\frac{1}{|1+i|^n} = \sum_{n=0}^{\infty}\frac{1}{(\sqrt{2})^n} < \infty \quad \text{since } \sqrt{2} > 1$$

The idea behind these examples is contained in the following theorem.

Theorem 2.3. **(Comparison Test)** *Let $\{c_n\}$ be a sequence of complex numbers. If there is a positive number K and an N, and a sequence $\{p_n\}$ of positive numbers such that*

(i) $|c_n| \le K p_n$, *for* $n \ge N$,

(ii) $\sum_{n=1}^{\infty} p_n < \infty$,

then $\sum c_n$ converges absolutely.

If instead, we have

(i)′ $|c_n| \ge K p_n$, *for* $n \ge N$,

(ii)′ $\sum_{n=1}^{\infty} p_n = \infty$,

then $\sum |c_n|$ diverges.

Proof. In the first case the sequence of partial sums is bounded.

$$\sum_{k=1}^{n} |c_k| = \sum_{k=1}^{N} |c_k| + \sum_{k=N+1}^{n} |c_k| \le \sum_{k=1}^{N} |c_k| + K \sum p_n < \infty$$

In the second case, the sequence of partial sums is unbounded.

$$\sum_{k=1}^{n} |c_k| = \sum_{k=1}^{N} |c_k| + \sum_{k=N+1}^{n} |c_k| \ge \sum_{k=1}^{N} |c_k| + \sum_{k=N+1}^{n} p_n$$

which is unbounded as $n \to \infty$.

Examples

23. $\sum_{n=0}^{\infty} z^n/n!$ converges absolutely for any complex z. Choose an integer N so that $N \ge 2|z|$. Then, for all n, $(N+n)! \ge (2|z|)^n$, so that

$$\frac{|z|^{N+n}}{(N+n)!} \le \frac{|z|^N}{2^n}$$

Since $\sum 1/2^n$ converges, so does $\sum |z|^n/n!$ by the comparison test. As a corollary result we obtain $\lim z^n/n! = 0$ for all z (this however could have been derived directly).

24. $\sum n^k z^n$ converges absolutely for all z, $|z| < 1$ and all integers k; and otherwise diverges. If $|z| \geq 1$, then $\lim n^k z \neq 0$, so the series can hardly converge. Now suppose $|z| < 1$. We want to prove the convergence by comparison with the geometric series, so we must account for the effect of the coefficients n^k. Note that $(n+1)/n \to 1$ as $n \to \infty$, thus also $(n+1)/n^k \to 1$ (Exercise 13). Let s be any number greater than 1. Then there is an N such that for all $n \geq N$,

$$\frac{(n+1)^k}{n^k} \leq s \quad \text{or} \quad (n+1)^k \leq sn^k$$

Thus, by induction we can conclude that, for all $n \geq 0$, $(N+n)^k \leq s^n N^k$. Thus, $(N+n)^k |z|^{N+n} \leq (s|z|)^n N^k |z|^N$. We should choose $s < 1/|z|$, so that $\sum_n (s|z|)^n < \infty$. With the choice then of s: $1 < s < 1/|z|$, we can apply the comparison test to obtain the convergence of our series $\sum n^k z^n$.

25. $\sum n! z^n$ diverges for all $z \neq 0$. We have seen in Example 2 that for any complex number c, $\lim\limits_{n\to\infty} c^n/n! = 0$, or, replacing c by z^{-1}, $\lim\limits_{n\to\infty} 1/n!z^n = 0$. This precludes the possibility that $\lim n!z^n = 0$, so the given series cannot converge.

26. $\sum 1/n^2$ converges. In a later section we shall give another proof of this, at present we rely on a tricky observation.

$$\sum_{n=1}^{N}\left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{N+1}$$

Thus, the series

$$\sum \left(\frac{1}{n} - \frac{1}{n+1}\right)$$

converges to 1. But

$$\frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}$$

thus

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$$

Now, $2n^2 \geq n^2 + n = n(n+1)$, thus

$$\frac{1}{n^2} \leq \frac{1}{n(n+1)}$$

so by comparison $\sum 1/n^2$ also converges.

27. $\sum 1/n^{(1+\varepsilon)}$ converges for any $\varepsilon > 0$. Let k be an integer so large that $k\varepsilon > 2$. Then, for any n; if $m \geq n^k$,

$$\frac{1}{m^{(1+\varepsilon)}} \leq \frac{1}{n^{k(1+\varepsilon)}} \leq \frac{1}{n^{k+2}}$$

Between n^k and $(n+1)^k$ there are $(n+1)^k - n^k$ integers. Since

$$\left(\frac{n+1}{n}\right)^k \to 1$$

there is an n_0 such that for $n \geq n_0$, $(n+1)^k \leq 2n^k$, or $(n+1)^k - n^k \leq n^k$. Thus,

$$\sum_{m=n^k+1}^{(n+1)^k} \frac{1}{m^{(1+\varepsilon)}} \leq \frac{n^k}{n^{k+2}} = \frac{1}{n^2}$$

Well, now we can show that the sequence of partial sums

$$\left\{\sum_{n=1}^{N} \frac{1}{n^{(1+\varepsilon)}}\right\}$$

is bounded, for

$$\sum_{n=1}^{N^k} \frac{1}{n^{(1+\varepsilon)}} \leq \sum_{n=1}^{n_0{}^k} \frac{1}{n^{(1+\varepsilon)}} + \sum_{n=n_0{}^k+1}^{N^k} \frac{1}{n^{(1+\varepsilon)}}$$

$$\leq n_0{}^k + \sum_{n=n_0}^{N} \sum_{m=n^k+1}^{(n+1)^k} \frac{1}{m^{(1+\varepsilon)}}$$

$$\leq n_0{}^k + \sum_{n=n_0}^{N} \frac{1}{n^2} \leq n_0{}^k + \sum \frac{1}{n^2} < \infty$$

Now a special kind of a series is a power series: the geometric series, and the series in Examples 24 and 25 are such series. A **power series** is a series of the form

$$\sum_{n=0}^{\infty} a_n z^n$$

Such a series has the property that if it converges for some z_0, then it converges for all z such that $|z| < |z_0|$, and if it diverges for some z_1, then it diverges for all z such that $|z| > |z_1|$. Thus, the geometric series diverges for $|z| \geq 1$ and converges for $|z| < 1$; the series $\sum z^n/n!$ converges for all z, and $\sum n! z^n$ converges for no z. This general property of power series is easily deduced from the comparison test. We make the following somewhat stronger statement.

Proposition 8. *Let $\{c_n\}$ be a sequence of complex numbers.*

(i) *If $\{|c_n| t^n\}$ is bounded for some positive number t, then $\sum c_n z^n$ converges absolutely for all z, $|z| < t$.*
(ii) *If $\{|c_n| t^n\}$ is unbounded, then $\sum c_n z^n$ diverges for all z, $|z| > t$.*

Proof.
(i) Suppose $M \geq |c_n| t^n$ for all n. Let z be such that $|z| < t$. Then

$$|c_n z^n| \leq |c_n| t^n \left(\frac{|z|}{t}\right)^n \leq M\left(\frac{|z|}{t}\right)^n \qquad \text{for all } n$$

and since $|z|/t < 1$, $\sum (|z|/t)^n < \infty$, so by the comparison test the series $\sum c_n z^n$ converges absolutely.
(ii) If $\{|c_n| t^n\}$ is unbounded so is $\{c_n z^n\}$ for all z, $|z| > t$. Thus, we cannot have $\lim c_n z^n = 0$, so $\sum c_n z^n$ cannot converge.

Definition 4. Let $\{c_n\}$ be a sequence of complex numbers. The **power series** associated to $\{c_n\}$ is the series $\sum_{n=0}^{\infty} a_n z^n$. The **radius of convergence** of the power series is the least upper bound R of all real numbers t such that the sequence $\{|c_n| t^n\}$ is bounded.

According to Proposition 8 the series $\sum_{n=0}^{\infty} a_n z^n$ converges for z inside the disk of radius $R(|z| < R)$, and diverges for z outside that disk (see Problem 12).

Examples

28. $\sum_{n=0}^{\infty} z^n/n$ has radius of convergence one. For if $t > 1$, then $\{t^n/n\}$ is unbounded, and if $t < 1$, $t^n/n \to 0$. Notice that we can make no clear assertion for z on the unit circle, since $\sum_{n=0}^{\infty} (1)^n/n$ diverges, but $\sum_{n=0}^{\infty} (-1)^n/n$ converges.

29. If $\{c_n\}$ is bounded, but does not tend to zero, $\sum_{n=0}^{\infty} c_n z^n$ has radius of convergence one. For clearly $\{c_n t^n\}$ is bounded for $t < 1$, and unbounded for $t > 1$.

There are two final tests of some importance. These are as follows:

Root test. If eventually

$$(|c_n|)^{1/n} < r \qquad \text{for some } r < 1$$

then $\sum c_n$ converges absolutely. If there are infinitely many n such that

$$(|c_n|)^{1/n} > R \qquad \text{for some } R > 1$$

then $\sum c_n$ diverges.

Ratio test. If there is an $r < 1$ such that eventually

$$\left|\frac{c_{n+1}}{c_n}\right| < r < 1$$

then $\sum c_n$ converges absolutely. If

$$\left|\frac{c_{n+1}}{c_n}\right| > R > 1 \qquad \text{for infinitely many } n$$

then $\sum c_n$ diverges.

These are both derived by comparison with the geometric series. We leave it to the student to derive these tests (Problem 13). Let us here indicate why the convergence assertions are true. Suppose $(|c_n|)^{1/n} < r < 1$, for n large enough (say $n \geq N$). Then $|c_n| < r^n$ eventually, so the partial sums $\sum |c_n|$ are bounded by

$$\sum_{n=0}^{N} |c_n| + \frac{1}{1-r}$$

by comparison with the geometric series. As for the ratio test, suppose

$$\left|\frac{c_{n+1}}{c_n}\right| < r \qquad \text{for } n \geq N$$

Then we have

$$
\begin{aligned}
&|c_{N+1}| < r|c_N| \\
&|c_{N+2}| < r|c_{N+1}| < r^2|c_N| \\
&|c_{N+3}| < r^3|c_N| \\
&\quad \cdots \\
&|c_{N+k}| < r^k|c_N|
\end{aligned}
$$

by induction. Thus, $\sum_{n=0}^{m} |c_n| \leq \sum_{n=0}^{N} |c_n| + |c_N| \sum r^k < \infty$ since $r < 1$.

• EXERCISES

11. Which of the following series converge?

(a) $\sum \sin\left(\frac{1}{n}\right)$.

(b) $\sum \sin\left(\frac{1}{n^2}\right)$.

(c) $\sum \tan\left(\frac{1}{n}\right)$.

(d) $\sum \tan\left(\frac{1}{n}\right) - \sin\left(\frac{1}{n}\right)$.

(e) $\sum \frac{n^5 + 8}{4n^6 + n^4}$.

(f) $\sum \frac{n^3 + n^2 + n + 1}{n^4 + n^5 + n^6 + n^7}$.

(g) $\sum \frac{n}{2^n}$.

(h) $\sum \frac{n!}{(2n)!} x^n$, $x > 0$.

(i) $\sum \frac{n!}{n^k} x^n$, k a positive integer, $0 < x < 1$.

(j) $\sum (-1)^n \sin \frac{1}{n}$.

(k) $\sum (-1)^n \frac{n}{n+1}$.

(l) $\sum (-1)^n \frac{n}{(n+1)^2}$.

(m) $\sum \left(\frac{1}{n^2} + \frac{1}{(n+1)^2} + \frac{1}{(n+2)^2}\right)$.

(n) $\sum \left(\frac{1}{n} - \frac{1}{n+1} + \frac{1}{n+2}\right)$.

12. Verify directly that $\lim z^n/n! = 0$ for every z.
13. Suppose $\lim c_n = c$. Then for any integer k, $\lim c_n^k = c^k$.

14. Find the disk of convergence of the following power series.

(a) $\sum_{n=0}^{\infty} z^n$.

(b) $\sum_{n=0}^{\infty} \frac{z^n}{n}$.

(c) $\sum_{n=0}^{\infty} \frac{z^n}{(n!)^2}$.

(d) $\sum_{n=0}^{\infty} \frac{n!}{(2n)!} z^n$.

(e) $\sum_{n=0}^{\infty} \left(\frac{z}{2n}\right)^n$.

(f) $\sum_{n=0}^{\infty} n!\, z^n$.

(g) $\sum \frac{z^n}{(2n)^2}$.

(h) $\sum (1 + nh)z^n$.

(i) $\sum z^{n^2}$.

(j) $\sum (1 + z)^n$.

• PROBLEMS

12. Let $\{c_n\}$ be a sequence of complex numbers, and let R be the radius of convergence of the power series $\sum c_n z^n$. Show that $\sum c_n z^n$ converges absolutely for $|z| < R$, $\sum c_n z^n$ diverges for $|z| > R$.

13. Derive the convergence and divergence assertions of the root and ratio tests.

2.4 Convergence in R^n

The notion of convergence of a sequence of vectors is easy to conceive, since a vector in R^n is just an n-tuple of real numbers. Thus, a sequence of vectors is an n-tuple of real sequences, and the question of convergence of the vector sequence is just that of the *simultaneous* convergence of those n real sequences. We might also directly paraphrase Definition 2 of convergence, using the notion of distance in R^n discussed in Chapter 1. These two possible notions are in fact the same.

Definition 5. Let $\{\mathbf{v}_k\}$ be a sequence of vectors in R^n. The sequence **converges** if there is a vector $\mathbf{v} \in R^n$ such that to every positive number $\varepsilon > 0$ there corresponds an integer K such that $\|\mathbf{v}_k - \mathbf{v}\| < \varepsilon$ for $k \geq K$. We write $\lim_{k \to \infty} \mathbf{v}_k = \mathbf{v}$ if $\{\mathbf{v}_k\}$ converges to $\mathbf{v}$.

Thus, $\lim \mathbf{v}_k = \mathbf{v}$ means precisely that $\lim \|\mathbf{v}_k - \mathbf{v}\| = 0$; that is, the distance between the general term $\mathbf{v}_k$ and $\mathbf{v}$ tends to zero as k becomes infinite. When put this way it sounds like just the notion we have in mind. Recalling that in Section 2.1 we said that a complex sequence $\{c_k\}$ converges to c precisely when $|c_k - c| \to 0$, we see that this coincides with the above definition when $n = 2$. Now, if we write out the sequence $\mathbf{v}_k$ of vectors in R^n as an n-tuple

$$\mathbf{v}_k = (v_k{}^1, \ldots, v_k{}^n) \tag{2.8}$$

we can view the given sequence as the n real sequences $\{v_k{}^j\}$, where $j = 1, \ldots, n$. We now verify the fact mentioned above, that $\mathbf{v}_k \to \mathbf{v}$ precisely when $v_k{}^j \to v^j$ for all j. Notice that Proposition 2 is that fact in the case of R^2.

Proposition 9. *The sequence* (2.8) *converges to the vector* $\mathbf{v} = (v^1, \ldots, v^n)$ *if and only if* $\lim\limits_{k\to\infty} v_k{}^j = v^j$ *for all* j.

Proof. If $\mathbf{w} = (w^1, \ldots, w^n)$ is a vector in R^n, then by definition

$$\|\mathbf{w}\| = (\textstyle\sum (w^i)^2)^{1/2}$$

Then, in particular

$$|v_k{}^j - v^j| \le \|\mathbf{v}_k - \mathbf{v}\| \qquad j = 1, \ldots, n \tag{2.9}$$

Suppose now that $\mathbf{v}_k \to \mathbf{v}$. Then, given $\varepsilon > 0$, there is a K such that $\|\mathbf{v}_k - \mathbf{v}\| < \varepsilon$ for $k \ge K$. Thus, by Equation (2.9) for each j, $|v_k{}^j - v^j| < \varepsilon$ for $k \ge K$. But this means precisely that $\lim\limits_{k\to\infty} v_k{}^j = v^j$.

Conversely, if $v_k{}^j \to v^j$ for all j, then $(v_k{}^j - v^j)^2 \to 0$ for all j, so $[\sum (v_k{}^j - v^j)^2]^{1/2} = \|\mathbf{v}_k - \mathbf{v}\| \to 0$ as $k \to \infty$. But then, by Definition 5, $\mathbf{v}_k \to \mathbf{v}$.

In precisely the same way we can verify that if the sequence of vectors (2.8) satisfies a Cauchy criterion so do each of the real sequences $\{v_k{}^j\}$, and thus are convergent. Hence, by Proposition 9 the sequence of vectors $\{\mathbf{v}_k\}$ also converges, so we have a Cauchy criterion for vector sequences also. This fact, as well as some basic algebraic properties of convergence of vectors is easily verifiable. Accordingly, we make these assertions, leaving the proofs to the reader.

Proposition 10. **(Cauchy Criterion)** *Let* $\{\mathbf{v}_k\}$ *be a sequence of vectors in* R^n. *Suppose to every* $\varepsilon > 0$ *there corresponds a* K *such that* $\|\mathbf{v}_r - \mathbf{v}_s\| < \varepsilon$ *whenever both* $r, s \ge K$. *Then the sequence* $\{\mathbf{v}_k\}$ *is convergent.*

Proposition 11. *Suppose* $\lim \mathbf{v}_k = \mathbf{v}$, $\lim \mathbf{w}_k = \mathbf{w}$, $\lim c_k = c$, *where* $\{\mathbf{v}_k\}$, $\{\mathbf{w}_k\}$ *are sequences of vectors in* R^n, *and* $\{C_k\}$ *is a sequence of real numbers. Then*

(i) $\lim(\mathbf{v}_k + \mathbf{w}_k) = \mathbf{v} + \mathbf{w}$,
(ii) $\lim\langle \mathbf{v}_k, \mathbf{w}_k\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$,
(iii) $\lim c_k\mathbf{v}_k = c\mathbf{v}$.

Example

30. Let us find a point of a given plane in R^3 which is closest to the origin. A plane is given by the equation $\langle \mathbf{x}, \mathbf{a}\rangle = c$ for fixed $\mathbf{a}, c$. Let $m = \text{g.l.b. } \{\|\mathbf{x}\|;\ \langle \mathbf{x}, \mathbf{a}\rangle = c\}$. Choose a sequence $\{\mathbf{x}_n\}$ on the plane such that $\|\mathbf{x}_n\| \to m$. We shall show that $\{\mathbf{x}_n\}$ actually converges. Now,

$$\|\mathbf{x}_n - \mathbf{x}_m\|^2 = \|\mathbf{x}_n\|^2 + \|\mathbf{x}_m\|^2 - 2\langle \mathbf{x}_n, \mathbf{x}_m\rangle \tag{2.10}$$

We can estimate the last term by using the fact that the midpoint $\frac{1}{2}(\mathbf{x}_n + \mathbf{x}_m)$ between $\mathbf{x}_n$ and $\mathbf{x}_m$ must also be on the given plane.

$$m^2 \le \left\|\frac{1}{2}(\mathbf{x}_n + \mathbf{x}_m)\right\|^2 = \frac{\|\mathbf{x}_n\|^2}{4} + \frac{\|\mathbf{x}_m\|^2}{4} + \frac{1}{2}\langle \mathbf{x}_n, \mathbf{x}_m\rangle$$

Thus,

$$-2\langle \mathbf{x}_n, \mathbf{x}_m\rangle \le \|\mathbf{x}_n\|^2 + \|\mathbf{x}_m\|^2 - 4m^2 \tag{2.11}$$

Combining (2.10) and (2.11), we find that

$$\|\mathbf{x}_n - \mathbf{x}_m\|^2 \le 2(\|\mathbf{x}_n\|^2 + \|\mathbf{x}_m\|^2 - 2m^2) \tag{2.12}$$

Now, since $\|\mathbf{x}_n\| \to m$, if $\varepsilon > 0$ is given, there is an n_0 such that for $n, m \ge n_0$, we have $\|\mathbf{x}_n\| < m + \varepsilon$, $\|\mathbf{x}_m\| < m + \varepsilon$. Inequality (2.12) then gives

$$\|\mathbf{x}_n - \mathbf{x}_m\|^2 \le 2((m+\varepsilon)^2 + (m+\varepsilon)^2 - 2m^2) \le 4m\varepsilon + 2\varepsilon^2 = \varepsilon(4m + 2\varepsilon)$$

This can be made as small as we please by choosing ε small. Thus if n, m are large enough, $\|\mathbf{x}_n - \mathbf{x}_m\|$ is small, so the sequence $\{x_n\}$ is Cauchy, and thus convergent. If $\mathbf{x} = \lim \mathbf{x}_n$, then $\|\mathbf{x}\| = \lim\|\mathbf{x}_n\| = m$, so $\mathbf{x}$ is the closest point on the plane to the origin.

Let us pause for a moment to consider the reasons, as illustrated by the above example, for studying the convergence of vectors. The central problem of calculus is to find an object, usually considered as a point in a given collection of points, which has certain specified properties (i.e., the maximum of a given function, or a zero of a function). At least, the theoretical aspect of the problem is to *prove the existence* of a point with such and such properties. Our technique for doing this is to use the desired properties to develop a sequence of approximations; our hope is that the approximations will converge; and that the limit will have the desired properties. It is thus essential to be able to discuss the question of convergence without already knowing the limit. Hence, for example, we have the Cauchy criterion. Further, we will need techniques, or criteria, to apply to the given properties in order to be able to extract the desired Cauchy sequence of approximation. For example, we will want to know: (a) If we have a convergent sequence of points having a property, does the limit have that property? (b) If we have a sequence of points having a property, does the sequence converge? or, at least does it have a convergent subsequence? These questions lead us to the reconsideration of the closed sets introduced in Section 1.11.

Recall that a closed set in R^n is a set whose complement is open. More precisely, S is closed if and only if corresponding to every $\mathbf{v} \notin S$, there is an $\varepsilon > 0$ such that any vector within ε of $\mathbf{v}$ is also not in S. In particular, if S is a closed set, and $\mathbf{v} \notin S$, then $\mathbf{v}$ cannot be the limit of a sequence of vectors in S. To put it positively, a closed set contains the limits of all convergent sequences it contains. This is in fact a defining criterion for closedness:

Proposition 12. *Let S be a set in R^n. The following assertions are equivalent:*

(i) *S is closed.*
(ii) *If $\{\mathbf{v}_k\}$ is a convergent sequence contained in S, then* $\lim \mathbf{v}_k \in S$.

Proof. Suppose S is closed. Let $\{\mathbf{v}_k\}$ be a sequence contained in S and suppose it converges to $\mathbf{v}$. If $\mathbf{v} \notin S$, since S is closed, there is an $\varepsilon > 0$ such that no vector in S gets within ε of $\mathbf{v}$. This is nonsense since $\mathbf{v}$ is the limit of a sequence in S. Thus, we must have $\mathbf{v} \in S$.

Suppose now S is not closed. Then there is a $\mathbf{v} \notin S$ such that for every $\varepsilon > 0$ there is a vector in S which is within ε of $\mathbf{v}$. in particular, for each n, taking $\varepsilon = 1/n$ there is a $\mathbf{v}_n$ such that $\|\mathbf{v}_n - \mathbf{v}\| \leq 1/n$ and $\mathbf{v}_n \in S$. Thus, $\mathbf{v}_n \to \mathbf{v}$ so (ii) does not hold for S.

We are now in a position to state our last basic consequence of the fundamental existence axiom for the real number system. This is that every bounded sequence in R^n has a convergent subsequence. It is easy to derive this from the Cauchy criterion, itself an assertion of existence. Let us illustrate the situation in R^2. Suppose $\{c_k\}$ is a sequence of complex numbers which is bounded; that is, it remains in some fixed square S_0 of side length K. Cut that square into four equal squares. At least one of these new squares has infinitely many of the $\{c_k\}$; let S_1 be one such square. Cut S_1 into four equal pieces and let S_2 be one of these new squares which has infinitely many of the $\{c_k\}$; now do the same with S_2 and so on (see Figure 2.4). In this way we obtain a sequence of squares $\{S_n\}$ with the properties:

(i) $S_n \supset S_{n+1}$,
(ii) side length of S_n is $K/2^n$,
(iii) S_n has infinitely many of the $\{c_k\}$.

Now that this is done, we can, for each integer n, select a $k(n)$ such that $c_{k(n)} \in S_n$, and $\{c_{k(n)}\}$ forms a subsequence of $\{c_k\}$. (For this we need to know that S_n contains infinitely many $\{c_k\}$, so that we can choose $k(n)$ greater than any previously chosen index.) Now, $\{c_{k(n)}\}$ is a Cauchy sequence. For let $\varepsilon > 0$, and choose N so that $\varepsilon > K\sqrt{2}/2^N$. Then, if $n, m \geq N$, we have $c_{k(n)}, c_{k(m)} \in S_N$, so

$$|c_{k(n)} - c_{k(m)}| < \sqrt{\left(\frac{K}{2^N}\right)^2 + \left(\frac{K}{2^N}\right)^2} = \frac{K\sqrt{2}}{2^N} < \varepsilon$$

Since the sequence $\{c_{k(n)}\}$ is a Cauchy sequence, by Proposition 10 it converges, and the argument for R^2 is concluded. This is the basic idea of the verification of

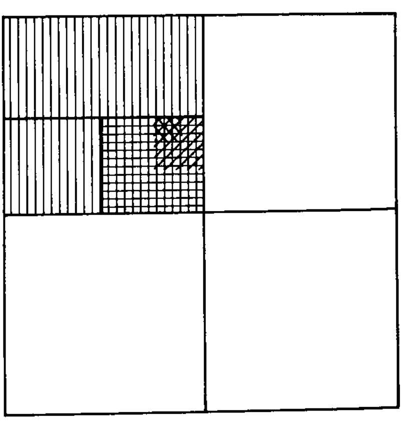

Figure 2.4

Theorem 2.4. *Every sequence in a closed and bounded set S in R^n has a subsequence which converges to a point of S.*

Proof. Suppose that S is closed and bounded and $\{\mathbf{v}_k\}$ is a sequence in S. We shall find a Cauchy subsequence. Since the sequence is bounded, it is contained in some ball $B(O, R)$. This ball can be covered by finitely many balls of radius 1. Since the $\{\mathbf{v}_k\}$ are infinite, there is one such ball which contains infinitely many. Call it B_1, and let $\mathbf{v}_{k(1)} \in B_1$. B_1 can be covered by finitely many balls of radius $\frac{1}{2}$. Let B_2 be one such which contains infinitely many of the $\{\mathbf{v}_k\}$ and let $\mathbf{v}_{k(2)} \in B_2$ with $k(2) > k(1)$.

Continuing in this way we obtain a sequence $\{B_n\}$ of balls, a subsequence $\{\mathbf{v}_{k(n)}\}$ of $\{\mathbf{v}_k\}$ such that (i) B_n has radius $1/n$, (ii) $\mathbf{v}_{k(n)} \in B_n$, (iii) $B_n \supset B_{n+1}$. Then $\{\mathbf{v}_{k(n)}\}$ is a Cauchy sequence, for if $n, m \geq N$, $\mathbf{v}_{k(n)}$ and $\mathbf{v}_{k(m)} \in B_N$ which has radius $1/N$, so

$$\|\mathbf{v}_{k(n)} - \mathbf{v}_{k(m)}\| < \frac{2}{N} \qquad \text{for all } n, m \geq N$$

By Proposition 10 there is a $\mathbf{v}$ such that $\mathbf{v}_{k(n)} \to \mathbf{v}$ as $n \to \infty$. Since S is a closed set, and $\{\mathbf{v}_{k(n)}\} \in S$, we also have $\mathbf{v} \in S$, so the theorem is proven.

Example

31. The unit sphere $S = \{\mathbf{x} \in R^n\colon \|\mathbf{x}\| = 1\}$ is closed. For if $\mathbf{x}_n \to \mathbf{x}$, then certainly $\|\mathbf{x}_n\| \to \|\mathbf{x}\|$, so if $\mathbf{x}_n \in S$, so is $\mathbf{x}$. Now suppose T is a linear transformation of R^n to R^n. We want to know if there is an $\mathbf{x} \in S$ at which $\|T\mathbf{x}\|$ is a maximum. First of all, the set of numbers of the form $\|T\mathbf{x}\|$ with $\mathbf{x} \in S$ is bounded. Let $A = (a_j{}^i)$ be the matrix representing T, and $M = \max |a_j{}^i|$. Then

$$T\mathbf{x} = T(x^1, \ldots, x^n) = (\textstyle\sum a_j{}^1 x^j, \ldots, \sum a_j{}^n x^j)$$

so

$$\begin{aligned} \|T\mathbf{x}\| &= [(\textstyle\sum a_j{}^1 x^j)^2 + \cdots + (\sum a_j{}^n x^j)^2]^{1/2} \qquad (2.13) \\ &\leq [nM^2 \|x\|^2 + \cdots + nM^2 \|\mathbf{x}\|^2]^{1/2} \leq nM \|\mathbf{x}\| \end{aligned}$$

Thus, nM is the desired bound. By the least upper bound axiom then, $m = \sup\{\|T\mathbf{x}\|: x \in S\}$ exists, and there is a sequence $\{\mathbf{x}_n\} \subset S$ such that $\|T\mathbf{x}_n\| \to m$. According to the above theorem there is a subsequence $\{\mathbf{y}_n\}$ which converges, say to $\mathbf{y}$. Since $\|T\mathbf{x}_n\| \to m$, we also have $\|T\mathbf{y}_n\| \to m$, and by (2.13), in fact $\|T\mathbf{y}\| = \lim \|T\mathbf{y}_n\| = m$.

• PROBLEMS

14. Prove Proposition 10.
15. Prove Proposition 11.
16. Let Π be a plane in R^3, and suppose $\mathbf{x}_0$ is the point on Π which is closest to the origin. Show that if $x \in \Pi$, then $\mathbf{x}_0$ is orthogonal to $\mathbf{x} - \mathbf{x}_0$. (*Hint:* If not, then one of $\mathbf{x} - \mathbf{x}_0$, $\mathbf{x} + \mathbf{x}_0$ is closer to the origin than $\mathbf{x}_0$.)
17. Find the point on the plane given by the equation $\langle \mathbf{x}, (1, 1, 1)\rangle = 3$ which is closest to the origin.
18. Find the point on the plane $\langle \mathbf{x}, (1, 0, 1)\rangle = 2$ which is closest to $-(1, 1, 1)$.
19. Let L be a linear function from R^n to R^m. Show that the kernel and range of L are both closed.
20. Let $L\colon R^n \to R$ be a linear function. Show that if $\lim \mathbf{x}_n = \mathbf{x}$, then also $\lim L(\mathbf{x}_n) = L(\mathbf{x})$.
21. Let $\mathbf{v}_0$ be a vector in R^n, and Π the set of $\mathbf{x}$ such that $\langle \mathbf{x}, \mathbf{v}_0\rangle = c$. Show that Π is closed.
22. Show that for any $\mathbf{v}_0 \in R^n$ and $r > 0$,

$$\{\mathbf{v} \in R^n\colon \|\mathbf{v} - \mathbf{v}_0\| \le r\}$$

is closed.

23. Show that $\mathbf{v}_k \to \mathbf{v}$ in R^n if and only if

$$\max_{1 \le i \le n} |v_k{}^i - v^i| \to 0$$

2.5 Continuity

We turn now to the consideration of functions from subsets of R^n to R^m. The basic notion of analysis being that of convergence, the fundamental class of functions will consist of those which respect convergence; that is, those which take convergent sequences into convergent sequences. These are continuous functions.

Definition 6. Let S be a set in R^n, and f a function defined on S, taking values in R^m. f **is continuous** on S if whenever $\mathbf{v}_k \to \mathbf{v}$ with $\mathbf{v}_k \in S$, all k, $\mathbf{v} \in S$, then $f(\mathbf{v}_k) \to f(\mathbf{v})$.

We shall be concerned most usually with the local study of a function near a given point. For this purpose we make this additional definition.

Definition 7. A function f from a set in R^n, taking values in R^m, will be said to be **continuous at** $\mathbf{v}_0 \in R^n$ if f is defined in a neighborhood of $\mathbf{v}_0$ and $\mathbf{v} \to \mathbf{v}_0$ implies $f(\mathbf{v}) \to f(\mathbf{v}_0)$.

Examples

32. $f\colon R^n \to R, f(\mathbf{v}) = \|\mathbf{v}\|$ is continuous. For if $\mathbf{v}_n \to \mathbf{v}$, then $\|\mathbf{v}_n - \mathbf{v}\| \to \mathbf{0}$ so that $\|\mathbf{v}_n\| \to \|\mathbf{v}\|$ since

$$| \, \| \mathbf{v}_n \| - \| \mathbf{v} \| \, | \leq \| \mathbf{v}_n - \mathbf{v} \|$$

33. $f\colon C \to C$, $f(z) = \bar{z}$ is continuous: $z_n \to z$ implies $|z_n - z| = |\bar{z}_n - \bar{z}| \to 0$, so that also $\bar{z}_n \to \bar{z}$.

34. A linear function on R^n is continuous. Let

$$f(v^1, \ldots, v^n) = \sum_{i=1}^{n} a_i v^i \tag{2.14}$$

Then, if $\mathbf{v}_k \to \mathbf{v}$ we have $v_k{}^1 \to v^1, \ldots, v_k{}^n \to v^n$, so that $\sum_{i=1}^n a_i v_k{}^i \to \sum_{i=1}^n a_i v^i$ since the limit of a sum is the sum of the limits. Thus, $f(\mathbf{v}_k) \to f(\mathbf{v})$.

Roughly, the idea of continuity of a function f is this: as a moving point $\mathbf{p}$ gets close to $\mathbf{p}_0$, the value $f(\mathbf{p})$ of f at $\mathbf{p}$ gets close to $f(\mathbf{p}_0)$. That is, we can ensure that $f(\mathbf{p})$ is as close as we please to $f(\mathbf{p}_0)$ by choosing $\mathbf{p}$ sufficiently close to $\mathbf{p}_0$. This leads to the so-called "$\varepsilon - \delta$" criterion for continuity, which we now give.

Proposition 13. *Let S be a subset of R^n, and let f be an R^m valued function defined on S.*

(i) *Let $\mathbf{x}_0 \in X$. f is continuous at $\mathbf{x}_0$ if and only if, to every $\varepsilon > 0$, there corresponds a $\delta > 0$ such that $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ implies $\|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \varepsilon$.*

(ii) *If S is open, f is continuous on S if and only if f is continuous at every point of S.*

Proof. (i) Supposing first that the $\varepsilon - \delta$ criterion is true, we shall show that f is continuous at $\mathbf{x}_0$. Let $\mathbf{x}_n \to \mathbf{x}_0$. We have to show $f(\mathbf{x}_n) \to f(\mathbf{x}_0)$. Given $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $\mathbf{x}$ is within δ of $\mathbf{x}_0$ we have $\|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \varepsilon$. Since $\mathbf{x}_n \to \mathbf{x}_0$, there is an N such that $n \geq N$ implies $\|\mathbf{x}_n - \mathbf{x}_0\| < \delta$. Thus, for $n \geq N$, $\|f(\mathbf{x}_n) - f(\mathbf{x}_0)\| < \varepsilon$, as desired.

Conversely, if the $\varepsilon - \delta$ criterion is false, then there is an ε_0 such that for every $\delta > 0$ there is an x_δ for which $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ but $\|f(\mathbf{x}) - f(\mathbf{x}_0)\| \geq \varepsilon_0$. Selecting

$$\delta = 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}$$

we obtain the corresponding sequence $\mathbf{x}_1, \mathbf{x}_{1/2}, \ldots, \mathbf{x}_{1/n}$, which converges to $\mathbf{x}_0$. But $f(\mathbf{x}_{1/n}) \nrightarrow f(\mathbf{x}_0)$ since the $f(\mathbf{x}_{1/n})$ are always outside the ball of radius ε_0 centered at $\mathbf{x}_0$.

Part (ii) is left as an exercise.

Examples

35. $f\colon R^2 \to R$ defined by

$$f(x, y) = \frac{5x}{1 + y^2}$$

is continuous at (0, 0). For

$$\left|\frac{5x}{1 + y^2}\right| \leq 5\,|x| \leq 5\,\|(x, y)\|$$

Thus, if ε is given we can choose $\delta = \varepsilon/5$. Then $\|(x, y)\| < \delta$ implies

$$\left|\frac{5x}{1 + y^2}\right| < 5\delta = \varepsilon$$

36.

$$f(x, y, z) = \frac{y^3 z}{1 + x^2 + z^2}$$

is continuous at (0, 0, 0). We have

$$|f(x, y, z) - f(0, 0, 0)| = \left|\frac{y^3 z}{1 + x^2 + z^2}\right| \leq |y^3 z| \leq \|(x, y, z)\|^4$$

Thus for each $\varepsilon > 0$ choose $\delta = \varepsilon^4 = \varepsilon$. Then $\|(x, y, z)\| < \delta$ implies $|f(x, y, z)| < \delta^4 = \varepsilon$.

37.

$$f(x, y) = \frac{(x+y)^2}{x^2+y^2} \qquad (x, y) \neq (0, 0) \qquad f(0, 0) = 0$$

This function is not continuous, since

$$f(x, x) = \frac{4x^2}{2x^2} = 2 \nrightarrow 0$$

If we redefine $f(0, 0) = 2$, this new function is still not continuous, since

$$f(0, y) = \frac{y^2}{y^2} = 1 \nrightarrow 2$$

38. We can easily verify the continuity of the linear function (2.14) by the $\varepsilon - \delta$ criterion. For

$$|f(\mathbf{v}) - f(\mathbf{w})| = |\sum a_i(v^i - w^i)| \leq \|(a^1, \ldots, a^n)\| \; \|\mathbf{v} - \mathbf{w}\|$$

by Schwarz's inequality. Thus, if $\varepsilon > 0$ is given, we can take $\delta = \|(a^1, \ldots, a^n)\|^{-1}\varepsilon$. Then $\|\mathbf{v} - \mathbf{v}_0\| < \delta$ implies $|f(\mathbf{v}) - f(\mathbf{v}_0)| < \varepsilon$.

The facts concerning convergence discussed in previous sections have application to the study of continuity, as might be expected. In particular, the assertion that every sequence in a closed bounded set has a convergent subsequence has profound significance for the behavior of continuous functions. Here is an important illustration.

Proposition 14. (Intermediate Value Theorem) *Let f be a continuous function on the interval $\{x \in R: a \leq x \leq b\}$, and suppose that $f(a) \leq \gamma \leq f(b)$. Then there is a c in the interval such that $f(c) = \gamma$.*

Proof. We seek (as in Figure 2.5) not just a point at which the value of f is γ, but more precisely the first such point c. We must find a way to describe this point which permits us to use the existence theorem. If $x < c$ we must have $f(x) < \gamma$, otherwise the graph of f crosses the line $y = \gamma$ between a and c. Thus, c is a lower bound for the set of x such that $f(x) \geq \gamma$. Since c is in that set, it must be the *greatest* such lower bound. So if there exists a first c at which $f(c) = \gamma$, it is the greatest lower bound of $\{x \in R: a \leq x \leq b, f(x) \geq \gamma\}$. We now show that this point (which exists by the least upper bound property) is the desired c.

Let $c = \text{g.l.b.}\{x: a \le x \le b, f(x) \ge \gamma\}$. Then c is a limit of a sequence $\{x_n\}$ in this set. Since $\gamma \le f(x_n)$ we must also have $\gamma \le \lim f(x_n) = f(c)$ since f is continuous. Now, if $f(c) \ne \gamma$, we must have $f(c) > \gamma$. Again, by continuity, there is a δ such that if $|x - c| < \delta$, then

$$|f(x) - f(c)| < \frac{f(c) - \gamma}{2}$$

from which it follows that for all x between c and $c - \delta$, $f(x) > \gamma$. Thus, $f(c - \delta) \ge \gamma$, contradicting the definition of c as a lower bound for the set of x with $f(x) \ge \gamma$. Hence $f(c) > \gamma$ is impossible, so we must have $f(c) = \gamma$.

Now, the most important fact about continuous real-valued functions is that they are bounded on closed and bounded sets. This follows easily from Theorem 2.4. If, say, f is continuous and not bounded above on the set S, then, for every positive integer n, there is an $\mathbf{x}_n \in S$ such that $f(\mathbf{x}_n) > n$. If S is closed and bounded, $\{\mathbf{x}_n\}$ has a convergent sequence $\{\mathbf{x}_{n(k)}\}$. Let $\lim \mathbf{x}_{n(k)} = \mathbf{x}_0$. Since f is continuous, $f(\mathbf{x}_0) = \lim_{k\to\infty} f(\mathbf{x}_{n(k)}) > \lim_{k\to\infty} n(k)$. But $n(k) \to \infty$ as $k \to \infty$, so this is impossible. Thus f is bounded on S. What is more it attains its least upper bound. For if m is this least upper bound, but is *not* a value of f, then $g(\mathbf{x}) = (f(\mathbf{x}) - m)^{-1}$ is an unbounded function on S, again a contradiction. To conclude: if f is a continuous real-valued function on a closed and bounded set S in R^n, then there are $\mathbf{x}_1, \mathbf{x}_2 \in S$ such that

$$f(\mathbf{x}_1) = \sup\{f(\mathbf{x}): \mathbf{x} \in S\}$$
$$f(\mathbf{x}_2) = \inf\{f(\mathbf{x}): \mathbf{x} \in S\}$$

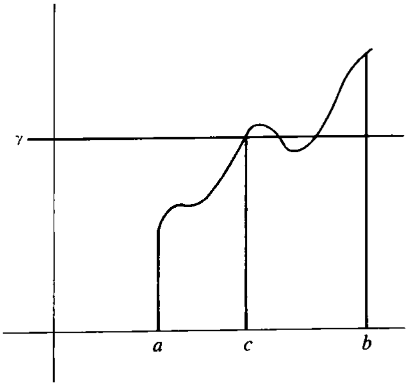

Figure 2.5

Here are the proofs in a slightly more general context.

Theorem 2.5. *Let f be a continuous R^m-valued function on the closed and bounded set S in R^n. Then the set of values of f on S,*

$$f(S) = \{f(\mathbf{x}) : \mathbf{x} \in S\}$$

is closed and bounded.

Proof. First, $f(S)$ is closed. Suppose $\mathbf{y}_n \in f(S)$ and $\mathbf{y}_n \to \mathbf{y} \in R^m$. We must show that $\mathbf{y} \in f(S)$. But this is easy. Since $\mathbf{y}_n \in f(S)$, there is for each n, $\mathbf{x}_n \in S$ such that $f(\mathbf{x}_n) = \mathbf{y}_n$. Since S is closed and bounded there is a subsequence $\{z_k\}$ of $\{x_n\}$ which converges, $\mathbf{z}_k \to \mathbf{z} \in S$. Since f is continuous, $f(\mathbf{z}_k) \to f(\mathbf{z})$. On the other hand, $\{f(\mathbf{z}_k)\}$ is a subsequence of $\{\mathbf{y}_n\}$, so $f(\mathbf{z}_k) \to \mathbf{y}$. Thus $f(\mathbf{z}) = \lim f(\mathbf{z}_k) = \mathbf{y}$ and $\mathbf{y} \in f(S)$.

If $f(S)$ is not bounded, there is for each n an $\mathbf{x}_n \in S$ such that $\|f(\mathbf{x}_n)\| \geq n$. But $\{\mathbf{x}_n\}$ has a convergent subsequence $\{\mathbf{z}_n\}$. Let $\lim \mathbf{z}_k = \mathbf{z}$. Then $\lim f(\mathbf{z}_k) = f(\mathbf{z})$. But $\{f(\mathbf{z}_k)\}$ is a subsequence of $\{f(\mathbf{x}_n)\}$, so $\|f(\mathbf{z}_k)\| \to \infty$, which is impossible since $\{f(\mathbf{z}_k)\}$ is convergent. Thus, $f(S)$ must be bounded.

In particular, suppose f is a real-valued function defined on the closed and bounded set S. Then $f(S)$ is bounded, so $M = \sup\{t : t \in f(S)\}$ exists, and since $f(S)$ is closed, $M \in f(S)$. Thus there is an $\mathbf{x}_1 \in S$ such that

$$f(\mathbf{x}_1) = \sup\{f(\mathbf{x}) : x \in S\}$$

Similarly, there is an $\mathbf{x}_2$ such that $f(\mathbf{x}_2) = \inf\{f(\mathbf{x}) : \mathbf{x} \in S\}$. This basic fact we state as

Theorem 2.6. *A continuous function attains its maximum and minimum on a closed bounded set.*

• PROBLEMS

24. Let $\mathbf{x}_0 \in R^n$. Show that $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{x}_0 \rangle$ is continuous on R^n.
25. Show that a linear function $L: R^n \to R^m$ is continuous.
26. Prove part (ii) of Proposition 13.
27. Show that if f is a continuous real-valued function on a closed and bounded set S, there is an $\mathbf{x}_2$ such that $f(\mathbf{x}_2) = \text{g.l.b.}\{f(\mathbf{x}) : \mathbf{x} \in S\}$.
28. Suppose that f, g are R^m-valued functions continuous at $\mathbf{p}_0 \in R^n$. Show that $f + g$ and $\langle f, g \rangle$ are also continuous at $\mathbf{p}_0$. If $c \in R$, then also cf is continuous at $\mathbf{p}_0$.

2.6 Calculus of One Variable

Theorem 2.6, which asserts that a continuous function attains its maximum and minimum on a closed and bounded set, is the fundamental theoretical tool of the calculus. We shall now give a brief review of the fundamentals of calculus, leaving the recollection of techniques to the student's memory. We shall give brief justifications of some of the more basic or special facts.

First of all, we studied in the calculus a limit concept which was more general than the sequential limit we have been studying. We recall the definition.

Definition 8. Suppose f is a real-valued function defined in a set

$$\{x: 0 < |x - x_0| < \delta_0\}$$

We say $\lim f(x) = L$ if and only if, for all $\varepsilon > 0$, there is $\delta > 0$ such that $|x - x_0| < \delta$ implies $|f(x) - L| < \varepsilon$.

First of all, the relationship between the two concepts of limit is an easy one: $\lim_{x \to x_0} f(x) = L$ if and only if for *every* sequence x_n converging to x_0, we have $\lim_{n \to \infty} f(x_n) = L$. We can thus rephrase the notion of continuity using Definition 8. f is continuous at x_0 if and only if $\lim_{x \to x_0} f(x) = f(x_0)$.

Proposition 15.

(i) *Suppose f is defined in $I = \{x: 0 < |x - x_0| < \delta\}$. Then $\lim_{x \to x_0} f(x) = L$ if and only if, for every sequence $\{x_n\}$ in I such that $x_n \to x_0$ we have*

$$\lim f(x_n) = L$$

(ii) *If f is also defined at x_0, f is continuous at x_0 if and only if*

$$\lim_{x \to x_0} f(x) = f(x_0)$$

Proof. We will prove only (i). The proof of (ii) is the same and is left as a problem. Suppose first that $\lim_{x \to x_0} f(x) = L$. Let $\{x_n\}$ be a sequence such that $x_n \to x_0$. Given $\varepsilon > 0$, there is a $\delta > 0$ such that $|f(x) - L| < \varepsilon$ for any x such that $|x - x_0| < \delta$. Now since $x_n \to x_0$, there is an N such that for $n \geq N$, $|x_n - x_0| < \delta$. Thus if $n \geq N$, $|f(x_n) - L| < \varepsilon$. Thus, $f(x_n) \to L$.

Now suppose $\lim_{x \to x_0} f(x) = L$ is false. Then there is an ε_0 such that for every δ we can find an x_δ such that $|x_\delta - x_0| < \delta$ but $|f(x) - L| \geq \varepsilon$. Consider the sequence $\{c_n\}$ of x's for $\delta = 1, \frac{1}{2}, \ldots, 1/n, \ldots$. Then $|c_n - x_0| < 1/n$, so certainly $c_n \to x_0$. But $f(c_n)$ is always outside the interval of width ε and center L, so it cannot converge to L.

Definition 9. Let f be a real-valued function defined in an interval about $x_0 \in R$. f is **differentiable at** x_0 if the limit

$$\lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t}$$

exists. If it does the limit is called the **derivative of** f at x_0 and is denoted

$$f'(x_0) \quad \text{or} \quad \frac{df}{dx}(x_0)$$

If f is differentiable in an interval I and the derivative f' is also differentiable there, then f is said to be **twice differentiable on** I and $(f')'$ is the second derivative of f and is denoted by

$$f'' \quad \text{or} \quad \frac{d^2f}{dx^2}$$

The higher derivatives $f''', \ldots, f^{(n)}, \ldots$ are defined successively in the obvious manner. A function which has derivatives of all orders on the interval will be said to be **infinitely differentiable** there. If f, g are n-times differentiable on I, so are $f + g, fg$, and cf for c a real number. If f is differentiable in an interval I it is continuous there. If f is differentiable at a point x_0 where it attains a local maximum (or minimum), then $f'(x_0) = 0$. This, together with Theorem 2.6 gives this basic existence theorem.

Theorem 2.7. **(Mean Value Theorem)** *Let f be differentiable on the closed interval $[a, b]$. There is a point $\xi \in (a, b)$ such that*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a} \tag{2.15}$$

Proof. This theorem has a nice geometric interpretation (Figure 2.6). There is a point $(\xi, f(\xi))$ on the graph $y = f(x)$ at which the tangent line is parallel to the line through $(a, f(a))$ and $(b, f(b))$. Clearly (see Problem 30), we need only verify this when the latter line is horizontal, that is, $f(b) = f(a)$. In this case, let $\xi_0 \in [a, b]$,

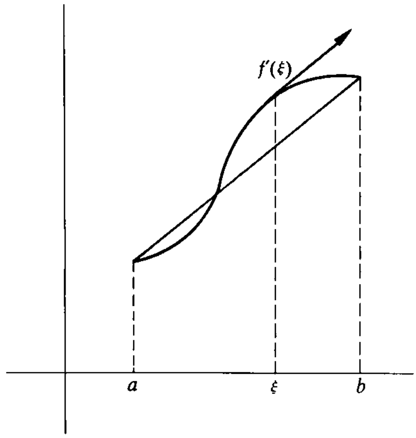

Figure 2.6

$\xi_1 \in [a, b]$ be the points at which f attains its maximum and minimum respectively on the interval (Figure 2.7). If either ξ_0 or ξ_1 is interior, then f has a local maximum there, so $f'(\xi) = 0$ for the appropriate ξ. If this is false, then $\{\xi_0, \xi_1\}$ are the points $\{a, b\}$, so $f(a) = f(b)$ is at once the maximum and minimum of f. Thus, f is constant on $[a, b]$, so f' is identically zero and we can choose any point for our ξ.

Now suppose that f is a differentiable function defined on the interval $[a, b]$, and g is a function defined on the range of f, and differentiable there. Then the composed function $h = g \circ f$, defined by

$$h(x) = g(f(x))$$

is also differentiable on a, b. For if $x_0 \in [a, b]$, then

$$\frac{h(x) - h(x_0)}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \cdot \frac{f(x) - f(x_0)}{x - x_0} \tag{2.16}$$

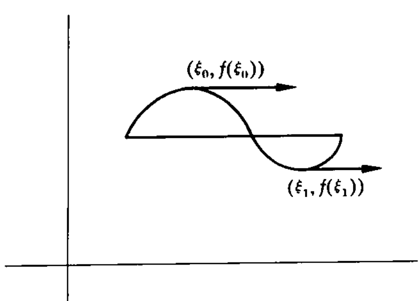

Figure 2.7

Taking the limit on both sides, we have (since $x \to x_0$ implies $f(x) \to f(x_0)$),

$$\lim_{x\to x_0} \frac{h(x)-h(x_0)}{x-x_0} = \lim_{f(x)\to f(x_0)} \frac{g(f(x))-g(f(x_0))}{f(x)-f(x_0)} \cdot \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$$

The limits on the right exist since f is differentiable at x_0, and g is differentiable at $f(x_0)$, so the limit on the left exists. Thus h is differentiable and we obtain the **chain rule**:

$$h'(x_0) = (g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$$

(Notice that if $f(x) = f(x_0)$, then (2.16) is invalid and the proof breaks down. However, that case can be treated separately.)

If f is a function from the interval $[a, b]$ to the interval $[\alpha, \beta]$ and there exists a function g: $[\alpha, \beta] \to [a, b]$ such that

$$\begin{aligned} g \circ f(x) &= x \qquad \text{for all } x \in [a, b] \\ f \circ g(y) &= y \qquad \text{for all } y \in [\alpha, \beta] \end{aligned}$$

we say that f is **invertible** and g is its **inverse**. The mean value theorem gives us a condition under which a differentiable function is invertible. If a function f has an inverse, it must be one-to-one. From (2.14) we see that this will be guaranteed if f' is never zero. This is the sufficient **condition for the invertibility of** f.

Theorem 2.8. *Suppose that f is a continuously differentiable function defined on the interval $[a, b]$, and f' is never zero. Let $f(a) = \alpha$ and $f(b) = \beta$. There is a continuously differentiable function g defined on the interval between α and β such that*

$$g(f(x)) = x \quad \textit{and} \quad g'(f(x)) = \frac{1}{f'(x)} \qquad \textit{for all } x$$

Proof. f is one-to-one. For if $a \le a_1 < b_1 \le b$, there is, by the mean value theorem a ξ between a_1 and b_1 such that

$$f(b_1) - f(a_1) = f'(\xi)(b_1 - a_1) \neq 0$$

by hypothesis. Thus $f(b_1) \neq f(a_1)$. By the intermediate value theorem every y between α and β is attained by f. Now we can define g as follows: let $g(y)$ be that

x such that $f(x)=y$. Clearly, $g(f(x))=x$ and $f(g(y))=y$. Now g is differentiable:

$$\lim_{y\to y_0}\frac{g(y)-g(y_0)}{y-y_0}=\lim_{y\to y_0}\frac{x-x_0}{f(x)-f(x_0)}=\lim_{x\to x_0}\frac{1}{[f(x)-f(x_0)]/x-x_0}=\frac{1}{f'(x_0)}$$

A further fundamental fact to be drawn from the mean value theorem is this: *A function is determined, up to a constant, by its derivative.*

Theorem 2.9. *Suppose that f, g are differentiable on the interval $[a, b]$ and that $f'(x) = g'(x)$ for all $x \in [a, b]$. Then there is a constant C such that*

$$f(x) = g(x) + C$$

Proof. Let $h=f-g$. By hypothesis $h'(x)=0$ for all $x\in[a, b]$. By the mean value theorem, for any $c\in[a, b]$, there is a ξ, $a\le\xi\le c$ such that

$$h'(\xi)=\frac{h(c)-h(a)}{c-a}$$

But $h'(\xi)=0$, so $h(c)=h(a)$. This for all $c\in[a, b]$, so h is constant and thus f differs from g by a constant, as desired.

Now, given any real-valued function f defined on interval I, we consider those differentiable functions F defined on I such that $F'=f$. By Theorem 2.9, any two such functions differ by a constant; thus by specifying the value of such an f at any point it is completely determined. We denote by $\int_a^x f=F(x)$ that function (if it exists) such that $F(a)=0$ and $F'(x)=f(x)$ for all $x\in[a, b]$. $\int_a^x f$ is called the **indefinite integral of** f. Every continuous function has an indefinite integral, which is given by the process of Riemann integration which we now describe.

Let f be a bounded function defined on the interval I. A partition P of I consists of an increasing sequence of points $a_0 < a_1 < \cdots < a_n$ such that $I = [a_0; a_n]$. We now construct two sums, corresponding to the approximations to the area under the graph of f given in Figure 2.8:

$$\Sigma(P,f)=\sum_{i=1}^{n} M_i(a_i-a_{i-1})$$

$$\sigma(P,f)=\sum_{i=1}^{n} m_i(a_i-a_{i-1})$$

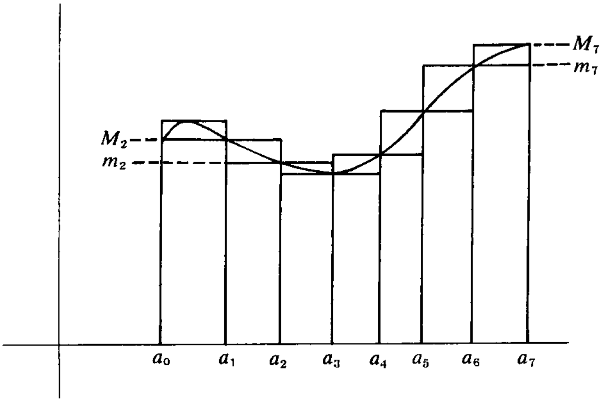

Figure 2.8

where M_i, m_i are the maximum and minimum values of f on the interval $[a_{i-1}, a_i]$.

Definition 10. Let f be a bounded real-valued function defined on the interval. f is **Riemann integrable** if

$$\inf_P \Sigma(P, f) = \sup_P \sigma(P, f) \tag{2.17}$$

(i.e., if we can find partitions for which the two sums Σ and σ are as close as we please). In this case the common value is called the **definite integral of f over the interval** I, and denoted $\int_I f$.

If f, g are integrable on the interval I, then so is $f + g$ and cf, $c \in R$. Further $\int_I (f+g) = \int_I f + \int_I g$, $\int_I cf = c \int_I f$. If f is integrable on the interval I, then f is integrable on every interval $J \subset I$. If f is integrable on the intervals $[a, b]$ and $[b, c]$ with $a < b < c$, then f is integrable on $[a, c]$ and

$$\int_{[a,c]} f = \int_{[a,b]} f + \int_{[b,c]} f$$

Furthermore, if $f \geq g$ and both functions are integrable, then $\int_I f \geq \int_I g$. Finally, if f is integrable on $[a, b]$, then

$$F(x) = \int_{[a,x]} f$$

is a continuous function of x.

The fundamental theorem of calculus says more: if f is continuous on $[a, b]$, then $\int_{[a,b]} f = \int_a^b f$; that is, the definite and the indefinite integrals of f coincide. The proof of this is actually quite easy to describe. Define these functions on the interval $[a, b]$, corresponding to the two sides of Equation (2.17);

$$\bar{F}(x) = \inf\{\Sigma(P, f) \colon P \text{ a partition of } [a, x]\}.$$

$$\underline{F}(x) = \sup\{\sigma(P, f) \colon P \text{ a partition of } [a, x]\}$$

To prove that f is Riemann integrable on $[a, b]$ is to prove $\bar{F}(b) = \underline{F}(b)$. We show, using Theorem 2.9, that in fact $\bar{F}(x) = \underline{F}(x)$ for all $x \in [a, b]$.

First of all $\bar{F}$ is differentiable in $[a, b]$. Let $x \in [a, b]$ and $h > 0$, then

$$\bar{F}(x + h) \le \bar{F}(x) + Mh \tag{2.18}$$

$$\bar{F}(x + h) \ge \bar{F}(x) + mh \tag{2.19}$$

where M, m are the maximum and minimum of f in the interval $[x, x + h]$. These inequalities can be routinely verified (see Problem 32); Figure 2.9 is convincing: $\bar{F}(x + h)$ is just $\bar{F}(x)$ plus the infimum of all $\Sigma(P, f)$ over partitions of $[x, x + h]$. Any such sum lies between Mh and mh. Now Equations (2.18) and (2.19) give

$$m \le \frac{\bar{F}(x + h) - \bar{F}(x)}{h} \le M$$

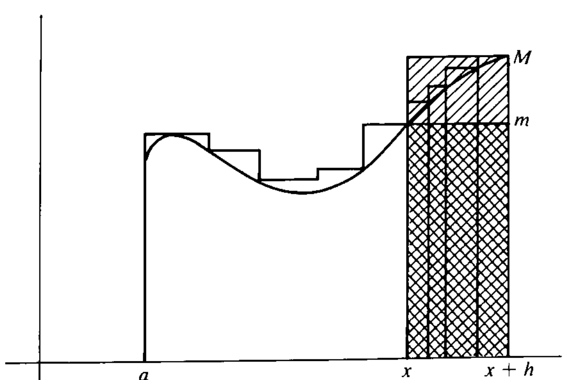

Figure 2.9

Letting $h \to 0$, since f is continuous, M and m both tend to $f(x)$. Thus $\bar{F}'(x)$ exists and is $f(x)$. Similarly, one verifies that $\underline{F}'(x)$ also exists for all x and has the same value. Thus, $\bar{F}$ and $\underline{F}$ differ by a constant. Since $\bar{F}(a) = \underline{F}(a) = 0$ is obvious, we have that $\bar{F}(x) = \underline{F}(x)$ for all x. Thus $\int_{[a,x]} f$ is defined for all x, is differentiable and has derivative f. This, then, is the proof of

Theorem 2.10. (Fundamental Theorem of Calculus) *Suppose f is continuous on the interval $[a, b]$. Then the integral $\int_a^x f$ exists for all $x \in [a, b]$. This is a differentiable function of f, and*

$$\frac{d}{dx}\int_a^x f = f(x)$$

● PROBLEMS

29. Prove Proposition 15(ii).

30. In the text the mean value theorem is proven in the case where $f(b) = f(a)$. The way to do the general case is to compare the graph of f with the line through $f(b)$ and $f(a)$. More precisely, let g be the function whose graph is that line, and consider $h = f - g$.

(a) Show that

$$h(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a) \tag{2.20}$$

(b) Show that $h(a) = h(b) = 0$.

(c) Now from the text there is a ξ between a and b such that $h'(\xi) = 0$. Differentiating (2.20), deduce that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

31. Suppose that f is differentiable on the interval $[a, b]$, and $f'(x) > 0$ for all x. Show that f is strictly increasing, that is, $f(x) < f(y)$ if $x < y$.

32. Verify inequalities (2.18) and (2.19).

33. Give an example of a continuous function of a real variable which is not differentiable. Give an example of an integrable function which is not continuous.

34. Find the real-valued function f, continuous on the interval $[0, 1]$ such that

$$\int_0^x f(t)\,dt = \int_x^1 f(t)\,dt \qquad \text{for all } x \in [0, 1]$$

35. Suppose f is k times differentiable on R, and $f^{(k)}(x) = 0$ for all x. Verify that f is a polynomial of degree at most $k - 1$.

2.7 Multiple Integration

The calculus of many variables results from the attempt to study functions of several variable quantities by generalizing to that context the calculus of a single variable. Some notions generalize easily, others require some ideas of linear algebra to be properly understood. The integration theory is much closer to that of one variable than is differentiation, hence we shall describe it first.

A **closed rectangle** in R^n is a set of the form

$$\{(x^1, \ldots, x^n) \in R^n : a^i \le x^i \le b^i\} = [(a^1, \ldots, a^n), (b^1, \ldots, b^n)]$$

for some fixed points $\mathbf{a} = (a^1, \ldots, a^n)$, $\mathbf{b} = (b^1, \ldots, b^n)$ in R^n. As in the case of intervals, we denote the corresponding open and half-open rectangles in the same way:

$$\begin{aligned}(\mathbf{a}, \mathbf{b}) &= \{\mathbf{x} \in R^n : a^i < x^i < b^i\}\\ [\mathbf{a}, \mathbf{b}) &= \{\mathbf{x} \in R^n : a^i \le x^i < b^i\}\\ (\mathbf{a}, \mathbf{b}] &= \{\mathbf{x} \in R^n : a^i < x^i \le b^i\}\end{aligned}$$

The term rectangle will refer to any of these possibilities. The **volume** of the rectangle R determined by the vectors **a** and **b** is

$$\mathrm{Vol}(R) = (b^1 - a^1) \cdots (b^n - a^n)$$

Notice that the volume of R is the same whether R is closed, open or half-open. Of course, this is as it should be since the faces contribute no volume.

Now let S be any set. The **characteristic function** of S, denoted by χ_S is the function which is one on S and identically zero off S. We should want to define integral so that the volume of S coincides with the integral of χ_S. In particular, for a rectangle R we shall have $\int \chi_R = \mathrm{Vol}(R)$. The notion of integral will be built up piece by piece so that things turn out that way. Now suppose that f is a finite linear combination of characteristic functions of rectangles: $f = \sum_{i=1}^{j} a_i\, \chi(R_i)$. Such a function is called **simple function**: It is constant on each of some finite collection of rectangles, and identically zero off their union.

Definition 11. Let f be a simple function. If $f = \sum_{i=1}^{k} a_i \chi_{Ri}$, we define

$$\int f = \sum_{i=1}^{k} a_i \operatorname{Vol}(R_i) \tag{2.21}$$

We immediately have a problem. It may be possible to also write the same function in another way, $f = \sum_{j=1}^{m} c_j \chi_{S_j}$ for some other collection of rectangles. For Definition 11 to make sense, we must be assured that the sum $\sum_{j=1}^{k} c_j \operatorname{Vol}(S_j)$ coincides with (2.21). In case the a_i and c_i are all one and the $\{R_i\}$ and $\{S_j\}$ are nonoverlapping (intersect only in faces), this amounts to the assertion that the volume of a set is the sum of the volumes of its rectangular pieces, *no matter how it is so partitioned.* The verification that (2.21) is the same for all expressions of the function f as a combination of characteristic functions is a long verification which is omitted. We now make this general definition of the integral.

Definition 12. Let f be a bounded real-valued function which is identically zero outside some rectangle R. The **upper integral** of f is

$$\overline{\int} f = \inf\{\int \sigma : \sigma \text{ a simple function on } R \text{ such that } \sigma \geq f\}$$

The **lower integral** of f is

$$\underline{\int} f = \sup\{\int \sigma : \sigma \text{ a simple function on } R \text{ such that } \sigma \leq f\}$$

f is **integrable** if

$$\overline{\int} f = \underline{\int} f; \qquad \text{the common value is } \textbf{the integral} \int f$$

This is the direct generalization of the definition of the Riemann integral given in Section 2.6. On the plane and in space it bears the same relation to area and volume as does the Riemann integral to length.

Definition 13. Let S be a set in R^n. If χ_S is integrable, we define the **volume of** S to be

$$\operatorname{Vol}(S) = \int \chi_S$$

Now there are sets for which χ_S is *not* integrable; these are highly pathological and shall not occur in this text. Notice that if $R_1, \ldots, R_n$ are nonoverlapping rectangles contained in the set S, then the sum of the volumes

$\sum \text{Vol}(R_i) = \int (\sum \chi_{R_i})$ is less than $\int \chi_S$, since $\chi_S > \sum \chi_{R_i}$. Thus the volume of S is greater than the sum of the volumes of any collection of nonoverlapping rectangles contained in S. Similarly, if now $R_1, \ldots, R_n$ are nonoverlapping rectangles containing S, $\bar{\int} \chi_S \le \sum \text{Vol}(R_i)$. Thus, the volume of S is trapped between the volume of any union of rectangles containing S and the volume of any union of rectangles contained in S. If we can make these two volumes as close as we please by proper choices of the rectangles, then $\int \chi_S$ is integrable (for then $\underline{\int} \chi_S = \bar{\int} \chi_S$), and its integral is the volume of S.

Theorem 2.11. *Let R be a closed rectangle in R^n. If f is continuous on R and zero off R, then f is integrable.*

Proof. Given $\varepsilon > 0$, we must find simple functions σ, τ such that $\sigma \ge f \ge \tau$ and $\int \sigma \le \int \tau + \varepsilon \text{Vol}(R)$; for then it will follow that

$$\bar{\int} f \le \int \sigma \le \int \tau + \varepsilon \text{Vol}(R) \le \underline{\int} f + \varepsilon \text{Vol}(R)$$

for any $\varepsilon > 0$. Thus, $\bar{\int} f \le \underline{\int} f$. In any case, since the inequality, $\underline{\int} f \le \bar{\int} f$ is obvious, f is integrable.

Such functions σ, τ are easily found using the basic property of uniform continuity (discussed in miscellaneous Problem 80). According to that theorem, given $\varepsilon > 0$, there is a $\delta > 0$ such that, if $|\mathbf{x} - \mathbf{y}| < \delta$ then $|f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon$. Now partition R into a finite set S of rectangles each of which has the property that any two points are within δ of each other. Thus, if for each such rectangle ρ, m_ρ, and M_ρ are respectively the maximum and minimum of f on ρ, we must have $M_\rho - m_\rho < \varepsilon$. Let

$$\sigma = \sum_{\rho \in S} M_\rho \chi_\rho \qquad \tau = \sum_{\rho \in S} m_\rho \chi_{\rho_0}$$

where ρ_0 is the open rectangle corresponding to ρ. Then $\sigma \ge f \ge \tau$ certainly, and

$$\int \sigma = \sum_{\rho \in S} M_\rho \text{Vol}(\rho) < \sum_{\rho \in S} (m_\rho + \varepsilon) \text{Vol}(\rho)$$

$$< \int \tau + \varepsilon \sum_{\rho \in S} \text{Vol}(\rho) < \int \tau + \varepsilon \text{Vol}(R)$$

since S is a partition of R into rectangles.

These following basic properties of the integral are easily derived.

Proposition 16. *The collection of integrable functions is a vector space and the integral is a linear function. That is:*

(i) *If f is integrable and $c \in R$, then cf is integrable and $\int cf = c \int f$.*
(ii) *If f, g are integrable, so is $f+g$ and $\int (f+g) = \int f + \int g$.*
(iii) *Furthermore, if $f \leq g$ then $\int f \leq \int g$.*

Proof. We leave the proof of (i) to the reader. (ii) is certainly true for simple functions. For if $f = \sum a_i \chi_{R_i}$, $g = \sum b_j \chi_{S_j}$, where R_i, S_j are rectangles, then $f+g = \sum a_i \chi_{R_i} + \sum b_j \chi_{S_j}$ is also simple, and thus integrable. By Definition 1,

$$\int (f+g) = \sum a_i \operatorname{Vol}(R_i) + \sum b_j \operatorname{Vol}(S_j) = \int f + \int g$$

More generally, now let f, g be any integrable functions. If $\varepsilon > 0$, there are simple functions σ_1, σ_2, τ_1, τ_2, such that

$$\sigma_1 \geq f \geq \sigma_2 \qquad \tau_1 \geq g \geq \tau_2$$

and

$$\int \sigma_1 \leq \int \sigma_2 + \varepsilon \qquad \int \tau_1 \leq \int \tau_2 + \varepsilon$$

Thus

$$\sigma_1 + \tau_1 \geq f + g \geq \sigma_2 + \tau_2$$

so

$$\overline{\int} (f+g) \leq \int \sigma_1 + \int \tau_1 \leq \int (\sigma_2 + \tau_2) + 2\varepsilon \leq \underline{\int} (f+g) + 2\varepsilon$$

Since $\varepsilon > 0$ was arbitrary, we obtain $\overline{\int} (f+g) \leq \underline{\int} (f+g)$, so $f+g$ is integrable. Finally,

$$\overline{\int} (f+g) \leq \int \sigma_2 + \int \tau_2 + 2\varepsilon \leq \int f + \int g + 2\varepsilon$$

so letting $\varepsilon \to 0$,

$$\int (f+g) \leq \int f + \int g$$

Similarly,

$$\underline{\int}(f+g)+2\varepsilon \geq \int \sigma_1 + \int \tau_2 \geq \int f + \int g$$

so again letting $\varepsilon \to 0$,

$$\underline{\int}(f+g) \geq \int f + \int g$$

(iii) Finally if $f \leq g$, then $g - f \geq 0$. But certainly the function which is identically zero is a simple function. Thus $\int (g-f) \geq \underline{\int}(g-f) \geq 0$. By (ii) it follows that $\int g - \int f \geq 0$, or $\int g \geq \int f$.

We shall now give the basic tool for computing integrals: **Fubini's theorem.** According to that result we can integrate by integrating one variable at a time. For the purpose of showing this, write the variable $(x^1, \ldots, x^n)$ of R^n as $(\mathbf{x}, y)$ where $\mathbf{x} \in R^{n-1}$ and $y \in R$: $\mathbf{x} = (x^1, \ldots, x^{n-1})$, $y = x^n$. Let f be a function defined on a rectangle R in R^n, and suppose for each y fixed, $f(\mathbf{x}, y)$ is an integrable function of $\mathbf{x}$. Define $F(y) = \int f(\mathbf{x}, y)\, d\mathbf{x}$. If F is an integrable function of y, its integral

$$\int F(y)\, dy = \int \left[\int f(\mathbf{x}, y)\, d\mathbf{x} \right] dy$$

is called the **iterated integral** of f. We shall now show that if f is integrable this is the same as $\int f$. More generally (after applying this principle n times) if all functions appearing in the following formula are integrable, then the formula is valid.

$$\int f(x^1, \ldots, x^n)\, dx^1 \cdots dx^n$$
$$= \int \left[\int \cdots \left[\int f(x^1, \ldots, x^n)\, dx^1 \right] dx^2 \cdots dx^n \right] \qquad (2.22)$$

This follows from Fubini's theorem.

Theorem 2.12. *Let f be an integrable function on a rectangle R in R^n. We refer to the coordinates of R^n as $(\mathbf{x}, \mathbf{y})$, where $\mathbf{x} \in R^k$, $\mathbf{y} \in R^{n-k}$*

(i) *These functions of* $\mathbf{y}$, $\underline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}$, $\overline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}$ *are integrable.*

(ii) *These functions of* $\mathbf{x}$, $\underline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{y}$, $\overline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{y}$ *are integrable.*

(iii) $\int f$ *is given by any iterated integral of* f; *for example,*

$$\int f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}\, d\mathbf{y} = \int \left[\overline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \right] d\mathbf{y} = \int \left[\underline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{y} \right] d\mathbf{x}$$

Proof. It is easily verified that the collection of functions for which the assertions (i), (ii), and (iii) are true is a vector space. Furthermore, these assertions are obvious for the characteristic function of a rectangle. Thus, Fubini's theorem holds for simple functions.

Now, suppose f is a bounded, real-valued function on the given rectangle R, and suppose that σ is a simple function, and $f \geq \sigma$. By definition of the lower integral with respect to the $\mathbf{x}$ coordinate,

$$\underline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \geq \int \sigma(\mathbf{x}, \mathbf{y})\, d\mathbf{x}$$

Now this inequality is maintained after taking the lower integrals with respect to y, thus

$$\underline{\int} \left[\underline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \right] d\mathbf{y} \geq \int \left[\int \sigma(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \right] d\mathbf{y} = \int \sigma(\mathbf{x}, \mathbf{y})\, d\mathbf{x}\, d\mathbf{y} \tag{2.23}$$

since Theorem 2.12 is true for simple functions. Equation (2.23) being true for any $\sigma \leq f$, we can take the least upper bound on the right, obtaining

$$\underline{\int} \left[\underline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \right] \mathrm{d}\mathbf{y} \geq \underline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}\, d\mathbf{y}$$

Now, by considering simple functions σ such that $\sigma \geq f$ and applying the same kind of reasoning we obtain this inequality

$$\overline{\int} \left[\overline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \right] d\mathbf{y} \leq \overline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}\, d\mathbf{y}$$

As a result, we obtain this string of inequalities, which is valid for any bounded, real-valued function on R:

$$\overline{\int} f \geq \overline{\int} \left[\overline{\int} f \right] \geq \left\{ \begin{matrix} \underline{\int} \left[\overline{\int} f \right] \\ \overline{\int} \left[\underline{\int} f \right] \end{matrix} \right\} \geq \underline{\int} \left[\underline{\int} f \right] \geq \underline{\int} f \tag{2.24}$$

(The second and third inequalities follow immediately from the fact that the upper integral always dominates the lower integral.) Now, if f is indeed integrable, the first and last terms of (2.24) are the same, so all are the same. That the second and top third are equal implies that $\bar{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}$ is integrable. That the bottom third and fourth are equal says that $\underline{\int} f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}$ is integrable. The equation

$$\int f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}\, d\mathbf{y} = \int \left[\int f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \right] d\mathbf{y}$$

now just states the equality of the end terms with the interior terms.

Now we shall illustrate the use of Fubini's theorem. Before doing that, we should remark that we rarely have the occasion to integrate functions defined on a rectangle; more often such a function is defined or considered only on a given measurable domain D. We make the following definition.

Definition 14. Let D be a domain contained in a rectangle R. Given a function f defined on D, we say f is **integrable** if this is so for the function $\tilde{f}$ defined on R by

$$\tilde{f}(\mathbf{x}) \begin{cases} = f(\mathbf{x}) & \mathbf{x} \in D \\ = 0 & \mathbf{x} \in R, \mathbf{x} \notin D \end{cases}$$

We define $\int_D f = \int \tilde{f}$.

If D is a subdomain of a rectangle R bounded by a surface which is the graph of a function, or has some other redeeming property, then the function f will be integrable if $\tilde{f}$ is. We shall not pursue this theoretical inquiry, but rather tacitly assume our domains are redeemable.

Example

39. $\{D = (x, y)\colon \ 0 \le y \le x^2, \ \ 0 \le x \le 1\}$, $f(x, y) = x^2 + y^2$. Define $\tilde{f}(x, y) = x^2 + y^2$ if $(x, y) \in D$, and $\tilde{f}(x, y) = 0$ otherwise. Then

$$\int_D f = \int_R \tilde{f} = \int_{-1}^{1} \left[\int_{-1}^{1} \tilde{f}(x, y)\, dy \right] dx = \int_0^1 \left[\int_0^{x^2} (x^2 + y^2)\, dy \right] dx$$

since, for fixed x, $\tilde{f}(x, y)$ is zero if $x < 0$ or $y > x^2$ and otherwise is $x^2 + y^2$. We thus obtain

$$\int_D f = \int_0^1 \left[x^2 y + \frac{y^3}{3} \right] \Big|_0^{x^2} dx = \int_0^1 \left(x^4 + \frac{x^6}{3} \right) dx = \frac{1}{5} + \frac{1}{21} = \frac{26}{105}$$

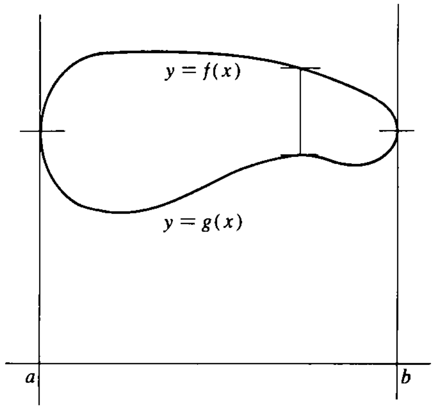

Figure 2.10

Let us do the same example, iterating this time in the other order.

$$\int_D f = \int_{-1}^{1}\left[\int_{-1}^{1} f(x, y)\,dx\right]dy = \int_0^1\left[\int_{\sqrt{y}}^{1}(x^2 + y^2)\,dx\right]dy$$

$$= \frac{1}{3} + \frac{1}{3} - \frac{2}{15} - \frac{2}{7} = \frac{26}{105}$$

The general technique can be described as follows: Try to write the domain in either of these forms (Figure 2.10).

$$D = \{(x, y)\colon a \le x \le b, g(x) \le y \le f(x)\}$$

or (Figure 2.11)

$$D = \{(x, y)\colon a \le y \le b, \phi(y) \le x \le \psi(y)\}$$

Then, given the function f defined on D, we can write

$$\int f = \int_a^b\left[\int_{f(x)}^{g(x)} f(x, y)\,dy\right]dx$$

in the first case; and in the second

$$\int f = \int_\alpha^\beta\left[\int_{\phi(x)}^{\psi(y)} f(x, y)\,dx\right]dy$$

Of course, if neither case can be obtained, then D might have to be broken up into pieces in each of which either representation is possible. The computation of integrals in more than two dimensions is done in pretty much the same way, but with a certain amount of additional care. For example, one should try to pick out one of the coordinates, say z, so that the given domain takes the form $g(y) \leq x \leq f(y)$, where y represents all the other coordinates and ranges through some domain D_0. Now one proceeds to break down D_0 in the same way.

Examples

40. $D = \{(x, y, z): \ x^2 + y^2 + z^2 \leq 1, \ \ x \geq 0, \ \ y \geq 0, \ \ z \geq 0\}$, $f(x, y, z) = xyz$.
Now z ranges between 0 and $(1 - (x^2 + y^2))^{1/2}$, so

$$D = \{(x, y, z): x^2 + y^2 \leq 1, 0 \leq x, 0 \leq y, 0 \leq z \leq [1 - (x^2 + y^2)]^{1/2}\}$$

Thus, continuing the analysis of

$$D_0 = \{(x, y): x^2 + y^2 \leq 1, \ 0 \leq x, 0 \leq y\}$$

$$D = \{(x, y, z): \ 0 \leq x \leq 1, \ 0 \leq y \leq (1 - x^2)^{1/2}, \\ 0 \leq z \leq [1 - (x^2 + y^2)]^{1/2}\}$$

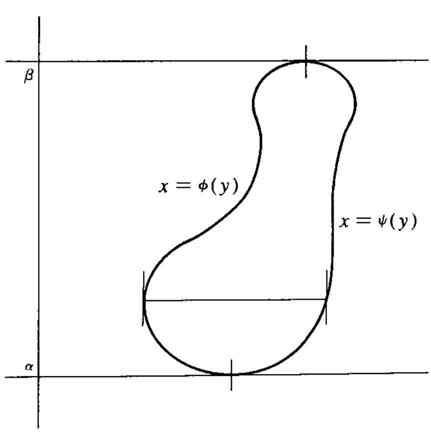

Figure 2.11

and

$$\int_D f = \int_0^1 x\left[\int_0^{(1-x^2)^{1/2}} y\left[\int_0^{[1-(x^2+y^2)]^{1/2}} z\,dz\right]dy\right]dx$$

$$= \frac{1}{2}\int_0^1 x\left[\int_0^{(1-x^2)^{1/2}} y(1-(x^2+y^2))\,dy\right]dx$$

$$= \frac{1}{2}\int_0^1 \left[\frac{x(1-x^2)^2}{2} - x\,\frac{(1-x^2)^2}{4}\right]dx = \frac{1}{24}$$

41. $D = \{(x, y, z)\colon x^2 + y^2 + z^2 \le 1,\ (x - \tfrac{1}{2})^2 + y^2 \le \tfrac{1}{4},$
$x \ge 0 \quad y \ge 0 \quad z \ge 0\} \qquad f(x, y, z) = 1$

(see Figure 2.12). We may rewrite this domain as

$$D = \{(x, y, z)\colon (x - \tfrac{1}{2})^2 + y^2 \le \tfrac{1}{4},\ x \ge 0,\ y \ge 0,$$
$$0 \le z \le [1 - (x^2 + y^2)]^{1/2}\}$$
$$= \{(x, y, z)\colon 0 \le x \le 1,\ 0 \le y \le [\tfrac{1}{4} - (x - \tfrac{1}{2})^2]^{1/2},$$
$$0 \le z \le [1 - (x^2 + y^2)]^{1/2}\}$$

Thus

$$\mathrm{Vol}(D) = \int_0^1\left[\int_0^{[\frac{1}{4}(x-\frac{1}{2})^2]^{1/2}}\left[\int_0^{[1-(x^2+y^2)]^{1/2}} dx\right]dy\right]dz$$

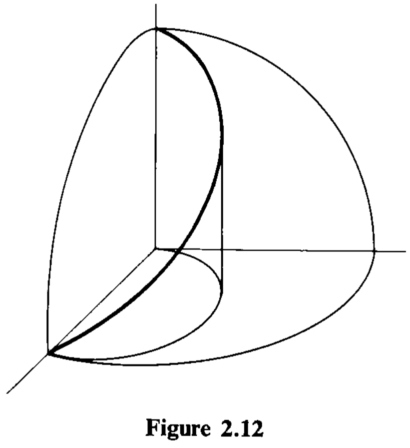

Figure 2.12

Integration is clearly of value in computing volumes; it also plays a role in the study of mass. Suppose E is a domain in R^3 filled with a certain fluid. If D is any subdomain in E, we shall let (D) be the mass of the fluid contained in D. What information do we need in order to compute mass (D), and how do we compute it? The answer is suggested by comparison of the properties of mass with those of volume. In fact, it is clear that the intuitive properties of mass are the same as the properties of volume; so we should also expect to be able to compute masses by integration. In fact, we introduce the notion of density: for $x_0 \in E$, the **density** $\sigma(x_0)$ **of the fluid at** x_0 is the limit

$$\lim_{R \to x_0} \frac{\text{mass}(R)}{\text{Vol}(R)}$$

where we mean by $R \to x_0$, that x_0 is in the rectangle R, and the lengths of the sides of R tend to zero (we might call mass (R)/Vol (R) the relative density of the fluid in the rectangle R). Now, the mass of the fluid in any domain is computable in terms of this density function σ. Suppose D is such a domain and $\{R_i\}$ is a collection of pairwise disjoint rectangles in D and almost filling D. Then

$$\sum \text{mass}(R_i) = \sum \frac{\text{mass}(R_i)}{\text{Vol}(R_i)} \text{Vol}(R_i)$$

is an approximation to mass (D) and as the size of the rectangles gets smaller and smaller, the approximation gets better. On the other hand, this sum is the integral of a simple function approximating σ, and thus approximates $\int_D \sigma$. Taking the limit we obtain mass $(D) = \int_D \sigma$.

• EXERCISES

15. Compute the volume of these domains:
 (a) $\{(x, y) \in R^2 \colon x^2 + y^2 \leq 1\}$.
 (b) $\{(x, y) \in R^2 \colon x^2 \leq y \leq 1\}$.
 (c) $\{(x, y, z) \in R^3 \colon 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq x^2 + y^2\}$.
 (d) $\{(x, y, z) \in R^3 \colon -1 \leq x \leq 1, 0 \leq y \leq 2, y \leq z \leq y + x^2\}$.

16. Verify that the volume of a right circular cylinder of radius r and height h is $\frac{1}{2}\pi r^2 h$.

17. Integrate the function f on the unit rectangle $[(0, 0), (1, 1)]$ in R^2.
 (a) $f(x, y) = x \cos 2\pi y$.
 (b) $f(x, y) = |(x - \frac{1}{2})(y - \frac{1}{2})|$.
 (c) $f(x, y) = xe^{xy} + ye^{-x}$.

(d) $f(x, y) = \begin{cases} x & \text{if } x \leq y \\ y & \text{if } y \leq x. \end{cases}$

(e) $f(x, y) = \begin{cases} x + y & \text{if } x + y \leq 1 \\ 1 & \text{if } x + y \geq 1. \end{cases}$

(f) $f(x, y) = (1 + x^2 + y^2)^{1/2}$.

18. Integrate the function f on the domain D in R^2.

(a) $D = \{(x, y)\colon 0 \leq x, 0 \leq y, x + y \leq 1\}, f(x, y) = x^2 + y^2$.

(b) $D = \{(x, y)\colon 0 \leq y \leq x \leq 1\}, f(x, y) = xy^2$.

(c) $D = \{(x, y)\colon 0 \leq y \leq x \leq 1\}, f(x, y) = x^2y$.

(d) $D = \{(x, y)\colon x^2 + y^2 \leq 1\}, f(x, y) = (x^2 - y^2)^2$.

19. Integrate the function f on the domain D in R^3.

(a) D is the intersection of the unit ball with the octant $\{x \geq 0, y \geq 0, z \geq 0\}$ and $f(x, y, z) = x + y + z$.

(b) D is as above and $f(x, y, z) = xyz$.

(c) D is the unit cube in the first octant and $f(x, y, z) = x^2 + y^2 + z^2$.

(d) D is the domain in the first octant bounded by the coordinate axes and the plane $x + y + z = 1$ and $f(x, y, z) = z$.

• PROBLEMS

36. Verify that the integral on R^n as defined in this section coincides, when $n = 1$, with the Riemann integral defined in the previous section.

37. Let f be a bounded, nonnegative, real-valued function defined on the interval I, and let $D = \{(x, y) \in R^2;\ x \in I, 0 \leq y \leq f(x)\}$. Verify this assertion: f is integrable if and only if D is measurable, and $\int_I f = \text{Vol}(D)$.

38. Use Problem 37 to verify this. Let D be a domain in R^2 and suppose that D is of the form

$$\{(x, y) \in R^2\colon a \leq x \leq b, g(x) \leq y \leq f(x)\}$$

Then, if D is measurable, $\text{Vol}\,(D) = \int_a^b [f(x) - g(x)]\,dx$.

39. Complete the proof of Fubini's theorem by verifying the second and third inequalities of Equation (2.24).

40. State and prove Fubini's theorem in three dimensions.

41. Suppose the unit ball is filled with a fluid whose density is proportional to the distance to the boundary. Find the radius of the ball centered at the origin which has precisely half the mass.

42. Suppose a cone of base radius r and height h is filled with mud (Figure 2.13). Suppose the density of the mud is equal to the distance from the base. What is the mass of the mud?

43. A beach B is shaped in the form of a crescent (see Figure 2.14)

$$B = \{(x, y)\colon 1 \leq x^2 + y^2;\ (x - \tfrac{1}{2})^2 + y^2 \leq 1\}$$

and the human density σ increases with the distance from the water. More precisely, $\sigma(x, y) = (x^2 + y^2)^{-1}$. What is the mass of humanity on that beach?

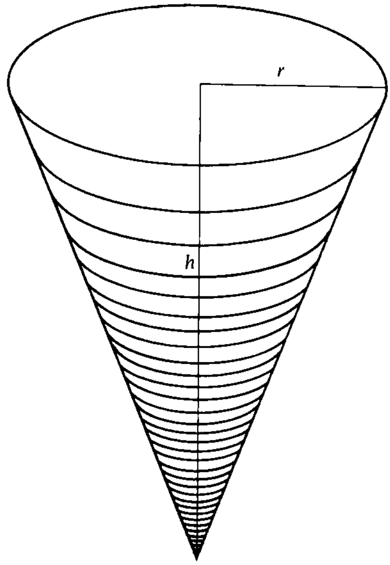

Figure 2.13

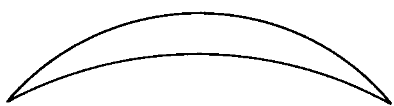

Figure 2.14

2.8 Partial Differentiation

Although the integral in R^n is defined without reference to the coordinates, it is computed by a succession of integrations, one coordinate at a time. The notion of differentiation is, to begin with, generalized to R^n one coordinate at a time. Later we shall see how to build out of this generalization an invariant notion of derivation.

Let $\mathbf{x}_0 \in R^n$, and suppose that f is a real-valued function defined in a neighborhood of $\mathbf{x}_0$. For each i consider the function of the single variable x^i given by

$$f(x_0{}^1, \ldots, x^i, \ldots, x_0{}^n)$$

If this function is differentiable, we denote the derivative by $\partial f/\partial x^i$, and call it the partial derivative of f in the x^i direction. More precisely,

Definition 15. Let f be a real-valued function defined in a neighborhood of $\mathbf{x}_0$ in R^n. The **partial derivative of f with respect to x^i** at $\mathbf{x}_0$ is the limit

$$\frac{\partial f}{\partial x^i}(\mathbf{x}_0) = \lim_{t\to 0} \frac{f(x_0^{\ 1}, \ldots, x_0^{\ i} + t, \ldots, x_0^{\ n}) - f(x_0^{\ 1}, \ldots, x_0^{\ n})}{t}$$

Another way of describing the partial derivative is this. Consider the function f only as a function on the line through $\mathbf{x}_0$ and in the $\mathbf{E}_i$ direction. This restriction is a function of one variable and $\partial f/\partial x^i$ is its derivative. These partial derivatives are computed merely by considering all but the relevant variable as constant.

Examples

42.

$$f(x, y) = xy$$

$$\frac{\partial f}{\partial x}(x, y) = y \qquad \frac{\partial f}{\partial y}(x, y) = x$$

43.

$$\frac{\partial}{\partial x}(x^2 y) = 2xy \qquad \frac{\partial}{\partial y}(x^2 y) = x^2$$

44.

$$f(x, y) = \cos[x(1 + y)]$$

$$\frac{\partial f}{\partial x}(x, y) = -(1 + y)\sin[x(1 + y)]$$

$$\frac{\partial f}{\partial y}(x, y) = -x\sin[x(1 + y)]$$

45.

$$f(x, y) = x^y$$

$$\frac{\partial f}{\partial x}(x, y) = yx^{y-1} \qquad \frac{\partial f}{\partial y}(x, y) = x^y \ln x$$

Of course, if the functions

$$\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}$$

are also defined in a neighborhood of $\mathbf{x}_0$, we may subject them to further partial differentiation, and keep going in this way as far as possible. We shall refer to any such operation as a **partial differentiation** and call its **order** the number of individual partial derivatives involved.

Thus, the order of

$$\frac{\partial}{\partial x^i}\left(\frac{\partial f}{\partial x^j}\right)$$

is 2; the order of

$$\frac{\partial^2}{\partial x^2}\left(\frac{\partial}{\partial y}\left(\frac{\partial^3 f}{\partial z^3}\right)\right)$$

is 6. We introduce a notational convention which deletes parentheses.

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$$

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

$$\frac{\partial^2 f}{\partial x^i\,\partial x^j} = \frac{\partial}{\partial x^i}\left(\frac{\partial f}{\partial x^j}\right)$$

$$\frac{\partial^3 f}{\partial x^i\,\partial x^j\,\partial x^k} = \frac{\partial}{\partial x^i}\left(\frac{\partial}{\partial x^j}\left(\frac{\partial f}{\partial x^k}\right)\right)$$

$$\frac{\partial^6 f}{\partial x^2\,\partial y\,\partial z^3} = \frac{\partial}{\partial x}\left(\frac{\partial^5 f}{\partial x\,\partial y\,\partial z^3}\right)$$

and so forth.

Suppose now that f is a function defined in an open set N in R^n and that $\partial f/\partial x^1, \ldots, \partial f/\partial x^n$ all exist in N. If we set all the variables constant except one, say x^i, then $\partial f/\partial x^i$ is just the derivative of f along this line. Thus, if $\partial f/\partial x^i = 0$, f is constant along the line on which only x^i varies. In such circumstances we say that f is independent of x^i, since f does not vary as x^i alone varies. If, moreover, $\partial f/\partial x^i$ is zero at all points of N for all i, then f depends on none of the variables, so is constant. As this is an important observation, we make it.

Proposition 17. *Suppose that f is a real-valued function defined in a neighborhood of $\mathbf{x}_0$ in R^n. f is constant near $\mathbf{x}_0$ if and only if all the derivatives $\partial f/\partial x^1, \ldots, \partial f/\partial x^n$ exist and are zero near $\mathbf{x}_0$.*

Proof. If f is constant, it is obvious that $\partial f/\partial x^i = 0$ for all i. On the other hand, suppose that these conditions are valid in a ball $B(\mathbf{x}_0, r)$ centered at $\mathbf{x}_0$. Let $\mathbf{y} = (y^1, \ldots, y^n) \in B(\mathbf{x}_0, r)$. We will show that $f(y) = f(\mathbf{x}_0)$. Figure 2.15 illustrates the proof. Consider the function of x^n:

$$f(x_0^1, \ldots, x_0^{n-1}, x^n)$$

This function has derivative zero by hypothesis, so is constant. Thus,

$$f(x_0^1, \ldots, x_0^{n-1}, x_0^n) = f(x_0^1, \ldots, x_0^{n-1}, y^n)$$

Now, the function of x^{n-1},

$$f(x_0^1, \ldots, x_0^{n-2}, x^{n-1}, y^n)$$

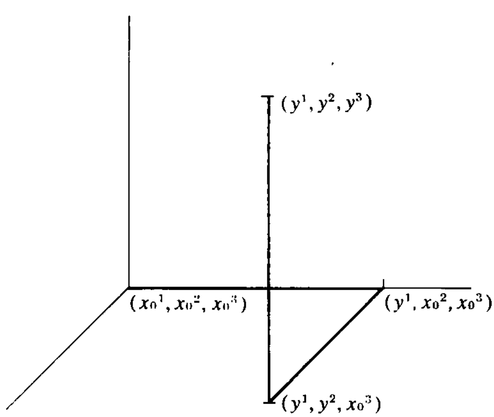

Figure 2.15

also has derivative zero, and thus must be constant, so

$$f(x_0^1, \ldots, x_0^{n-1}, y^n) = f(x_0^1, \ldots, y^{n-1}, y^n)$$

This together with the preceding equation gives

$$f(x_0^1, \ldots, x_0^{n-1}, x_0^n) = f(x_0^1, \ldots, x_0^{n-2}, y^{n-1}, y^n)$$

Continuing in this way, we can replace each x_0^j by the corresponding y^j one at a time, ending up with the desired equation $f(x_0) = f(y)$.

As far as the higher order differentiations are concerned, there is one basic fact we should now verify. This is that each partial differentiation depends only on the number of derivatives with respect to each coordinate, and not on the order in which they are performed. For example,

$$\frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial^2 f}{\partial y\, \partial x} \tag{2.25}$$

$$\frac{\partial^5 f}{\partial x^2\, \partial y\, \partial z^2} = \frac{\partial^5 f}{\partial z\, \partial x\, \partial y\, \partial x\, \partial z} = \frac{\partial^5 f}{\partial y\, \partial x\, \partial z\, \partial z\, \partial x} = \cdots$$

We shall verify only the first equation; it being clear that all others follow from a succession of applications of the first one. The verification of (2.25) amounts to an interesting application of Fubini's theorem.

Theorem 2.13. *Let f be a real-valued function defined in a neighborhood N of (x_0, y_0) in R^2 and suppose that all first- and second-order partial derivatives of f exist and are continuous on N. Then*

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

throughout N.

Proof. We apply Fubini's theorem to $\partial^2 f/\partial x\, \partial y$ in a sufficiently small rectangle $R = ((x_0, y_0), (s, t))$ contained in N (see Figure 2.16)

$$\int_{x_0}^{s} \left[\int_{y_0}^{t} \frac{\partial^2 f}{\partial x\, \partial y}\, dy\right] dx = \int_{y_0}^{t} \left[\int_{x_0}^{s} \frac{\partial^2 f}{\partial x\, \partial y}\, dx\right] dy \tag{2.26}$$

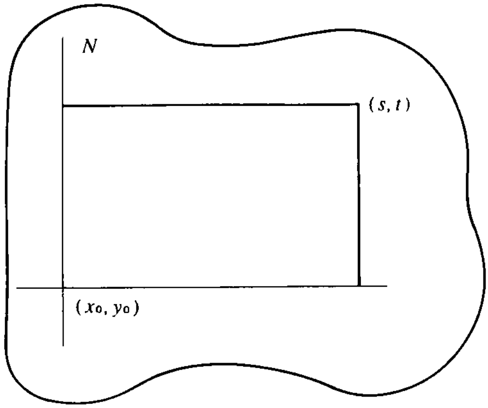

Figure 2.16

Now, we can easily evaluate the integral on the right-hand side. For fixed y,

$$\int_{x_0}^{s} \frac{\partial^2 f}{\partial x\, \partial y}\, dx = \int_{x_0}^{s} \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(x, y)\right) dx = \frac{\partial f}{\partial y}(s, y) - \frac{\partial f}{\partial y}(x_0, y) \tag{2.27}$$

Integrating once again (this time with respect to y) we obtain from Equations (2.26) and (2.27)

$$\int_{x_0}^{s}\left[\int_{y_0}^{t} \frac{\partial^2 f}{\partial x\, \partial y}\, dy\right] dx = \int_{y_0}^{t} \frac{\partial}{\partial y}\,[f(s, y) - f(x_0, y)]\, dy$$

$$= f(s, t) - f(x_0, t) - [f(s, y_0) - f(x_0, y_0)] \tag{2.28}$$

Now, we can differentiate this equation with respect to s first, and then t. By the fundamental theorem of calculus, we know how to differentiate the integral on the left with respect to the upper limit of integration:

$$\frac{\partial}{\partial s}\left\{\int_{x_0}^{s}\left[\int_{y_0}^{t} \frac{\partial^2 f}{\partial x\, \partial y}(x, y)\, dy\right] dx\right\} = \int_{y_0}^{t} \frac{\partial^2 f}{\partial x\, \partial y}(s, y)\, dy$$

Then, from (2.28)

$$\int_{y_0}^{t} \frac{\partial^2 f}{\partial x\, \partial y}(s, y)\, dy = \frac{\partial f}{\partial x}(s, t) - \frac{\partial f}{\partial x}(s, y_0)$$

Differentiating this equation now with respect to t, we obtain

$$\frac{\partial^2 f}{\partial x\,\partial y}(s, t) = \frac{\partial^2 f}{\partial y\,\partial x}(s, t)$$

as desired.

Another important application of Fubini's theorem is this result, which allows us to differentiate under the integral sign.

Proposition 18. *Suppose that f is a continuously differentiable function of two variables x and $\mathbf{y}$, $a \le x \le b$, and $\mathbf{y} \in D$, a domain in R^n. Define the function F on the interval $[a, b]$ by*

$$F(x) = \int_D f(x, \mathbf{y})\, d\mathbf{y}$$

Then F is differentiable and

$$\frac{dF}{dx}(x) = \int_D \frac{\partial f}{\partial x}(x, \mathbf{y})\, d\mathbf{y}$$

Proof. We shall show that F is the indefinite integral of the function

$$\int_D \frac{\partial f}{\partial x}(x, \mathbf{y})\, d\mathbf{y}$$

and thus by the fundamental theorem of calculus, the proposition follows. By Fubini's theorem

$$\int_a^t \left[\int_D \frac{\partial f}{\partial x}(x, \mathbf{y})\, d\mathbf{y}\right] dx = \int_D \left[\int_a^t \frac{\partial f}{\partial x}(x, \mathbf{y})\, dx\right] d\mathbf{y}$$

But by the fundamental theorem of calculus, the inner integral on the right is $f(t, \mathbf{y}) - f(a, \mathbf{y})$. Thus

$$\int_a^t \left[\int_D \frac{\partial f}{\partial x}(x, \mathbf{y})\, d\mathbf{y}\right] dx = \int_D [f(t, \mathbf{y}) - f(a, \mathbf{y})]\, d\mathbf{y} = F(t) - F(a)$$

Let us return now to the consideration of the first-order derivatives. These are obtained by differentiating *after* restricting the function to lines parallel to the coordinate axes. We generalize this notion to allow differentiation along *any* line. That is, we make this definition.

Definition 16. Let $\mathbf{x}_0 \in R^n$ and suppose f is a real-valued function defined in a neighborhood of $\mathbf{x}_0$. If $\mathbf{v}$ is a vector in R^n, we define the **directional derivative** $df(\mathbf{x}_0, \mathbf{v})$ to be

$$\frac{d}{dt} f(\mathbf{x}_0 + t\mathbf{v})\Big|_{t=0}$$

This is clearly the same as

$$\lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$$

We leave it as an exercise to verify that

$$\frac{\partial f}{\partial x^i}(\mathbf{x}_0) = df(\mathbf{x}_0, \mathbf{E}_i) \tag{2.29}$$

Now, in certain pathological cases the directional derivatives need not hang together in any nice way, but typically we need only know the partial derivatives in order to find any directional derivative.

Proposition 19. *Suppose f is defined in a neighborhood of $\mathbf{x}_0$ and the partial derivatives $\partial f/\partial x^1, \ldots, \partial f/\partial x^n$ all exist near $\mathbf{x}_0$. Then the directional derivatives $df(\mathbf{x}_0, \mathbf{v})$ vary linearly in $\mathbf{v}$.*

Proof. The argument consists in looking at the difference

$$f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)$$

one variable at a time. In order to expose the idea without encumbering the argument with a pile of indices, we consider the two-variable case. Write the difference

$$f(x_0 + th, y_0 + tk) - f(x_0, y_0)$$

as

$$\{f(x_0 + th, y_0 + tk) - f(x_0 + th, y_0)\} + \{f(x_0 + th, y_0) - f(x_0, y)\}$$

We can find a better expression for the term in the second set of braces by applying the mean value theorem to the function $f(s, y_0)$ of s. That is, there is a ξ_0 between x_0 and $x_0 + th$ such that

$$f(x_0 + th, y_0) - f(x_0, y_0) = \frac{\partial f}{\partial x}(\xi_0, y_0)th$$

Similarly, by applying the mean value theorem to the function $f(x_0 + th, s)$, we can rewrite the term in the first set of braces as

$$\frac{\partial f}{\partial y}(x_0 + th, \eta_0)tk$$

for some η_0 between y_0 and $y_0 + tk$. Thus, we have for suitable (ξ_0, η_0) in the rectangle $[(x_0, y_0), (x_0 + th, y_0 + tk)]$,

$$\frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} = \frac{\partial f}{\partial x}(\xi_0, y_0)h + \frac{\partial f}{\partial y}(x_0 + th, \eta_0)k$$

Letting $t \to 0$, we obtain by continuity that

$$d((x_0, y_0), (h, k)) = \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k \tag{2.30}$$

Thus the proposition is verified, at least in R^2.

This linear function, $df(\mathbf{x}_0, \mathbf{v})$ of the vector $\mathbf{v}$ in R^n is called the **differential of f at $\mathbf{x}_0$**. We will make a systematic study of this in a later chapter. The vector-valued function

$$\left(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}\right)$$

is called the **gradient** of f and is denoted by ∇f. It is clear from Proposition 19 that the generalization of (2.30) to n variables is

$$df(\mathbf{x}_0, \mathbf{v}) = \sum \frac{\partial f}{\partial x^i}(\mathbf{x}_0)v^i = \langle \mathbf{v}, \nabla f(\mathbf{x}_0)\rangle \tag{2.31}$$

The gradient behaves as a sort of "total derivative." It is not as powerful in the analysis of a function as the derivative in one variable and it is somewhat more cumbersome, but it does provide a similar kind of tool. For example,

Proposition 20. *The gradient of a function vanishes at any point at which it attains a maximum or minimum value.*

Proof. If $\mathbf{x}_0 = (x_0^1, \ldots, x_0^n)$ is (for instance) a maximum value of f, then $f(x_0^1, \ldots, x^i, \ldots, x_0^n)$, as a function of x^i, attains a maximum at x_0^i. Thus, $\partial f/\partial x^i$ vanishes at x_0^i. Since this is true for all i, $\nabla f(x_0) = 0$.

Examples

46. Consider $f(x, y, z) = x^2 + xy + y^2$.

$$\nabla f = (2x + y, x + 2y)$$

Thus ∇f is zero when

$$x = -\frac{y}{2} \qquad x = -2y$$

that is, only at the origin. This is the only critical point, and a minimum at that.

47. $f(x, y, z) = x \cos y + z$

$$\nabla f = (\cos y, \; -x \sin y, \; 1)$$

is never zero, so f has no critical values.

48. $f(x, y, z) = x \cos (yz)$

$$\nabla f = (\cos(yz), \; -xz \sin yz, \; -xy \sin yz)$$

∇f is zero only when $x = 0$ and $yz = \pi(n + \frac{1}{2})$ for any integer n. Clearly, f has both negative and positive values near any point on the line $\{x = 0\}$, so no such point is critical. Thus, f has no critical points.

• EXERCISES

20. Find the first partial derivatives of these functions.
(a) xyz (b) $\sin(xy)$ (c) x^{y^z} (d) $x^2y + y^2x$
21. Differentiate x^{x^x}. (*Hint:* This is the same as finding the directional derivative of x^{y^z} at a point (x, x, x) in the direction of $(1, 1, 1)$.)

22. If f is differentiable at $\mathbf{x}_0$, then

$$\frac{\partial f}{\partial x^i}(\mathbf{x}_0) = df(\mathbf{x}_0, \mathbf{E}_i)$$

for all i.

23. Suppose that f, g are differentiable at $\mathbf{x}_0$ in R^n. Show that fg is also differentiable and $\nabla(fg)(\mathbf{x}_0) = f(\mathbf{x}_0)\nabla g(\mathbf{x}_0) + g(\mathbf{x}_0)\nabla f(\mathbf{x}_0)$.

24. If f is differentiable at $\mathbf{x}_0$, and $f(\mathbf{x}_0) \neq 0$, then

$$\nabla\left(\frac{1}{f}\right)(\mathbf{x}_0) = \frac{-1}{f^2}\nabla f(\mathbf{x}_0)$$

25. What is the minimum of $x^2 + y^2 + (2y + 1)^2$?

26. What is the maximum of

$$\frac{x + 3y}{1 + x^2 + y^2}?$$

27. Compute the differentials of the functions in Exercise 20.

• PROBLEMS

44. Suppose f is a differentiable function of two variables and g_1, g_2 are differentiable functions of one variable so that the range of (g_1, g_2) is in the domain of f. Find the derivative of $h(t) = f(g_1(t), g_2(t))$.

45. Let f be a differentiable function of two variables. Show that f is a function of $x - y$ alone if and only if $\partial f/\partial x + \partial f/\partial y = 0$.

46. Suppose that $L\colon R^n \to R$ is a linear function. What is ∇L?

47. Let $T\colon R^n \to R^n$ be a linear transformation. Define the function on $R^n \times R^n$: $f(\mathbf{x}, \mathbf{y}) = \langle T\mathbf{x}, \mathbf{y}\rangle$. Show that f is differentiable, and $\nabla f(\mathbf{x}, \mathbf{y}) = \langle T^t\mathbf{y}, T\mathbf{x}\rangle$ (recall that T^t is the transpose of T: if T is represented by the matrix (a_j^i), then T^t is represented by (b_j^i) where $b_j^i = a_i^j$).

48. If $T\colon R^n \to R^n$ is a linear transformation, then the function $g(x) = \langle T\mathbf{x}, \mathbf{x}\rangle$ is differentiable, and $\nabla g(\mathbf{x}) = T^t\mathbf{x} + T\mathbf{x}$.

2.9 Improper Integrals

We return now to the study of functions of one variable; in fact, we will be considering functions defined on the whole real line. Our interest will focus on the "behavior at infinity" of such functions. For this purpose we introduce the notion of $\lim f(x)$ as $x \to \infty$.

Definition 17. If f is a real-valued function defined in an infinite interval $\{x: x > a\}$ we say that $f(x)$ **converges to** L **as** x **becomes infinite**, written $\lim_{x\to\infty} f(x) = L$ if, for every $\varepsilon > 0$ there is an $M > 0$ such that $x > M$ implies $|f(x) - L| < \varepsilon$. Similarly, if f is defined in $\{x: x<b\}$ we say $\lim_{x\to\infty} f(x) = L$ (the **limit of** $f(x)$ **is** L **as** x **becomes negatively infinite**) if, for every $\varepsilon > 0$ there is an $M > 0$ such that $x < -M$ implies $|f(x) - L| < \varepsilon$.

Examples

49. $\lim_{x\to\infty} 1/x = 0$. For given $\varepsilon > 0$, we can take $M = \varepsilon^{-1}$. Then $x > M$ implies $|1/x - 0| < \varepsilon$.

50.

$$\lim_{x\to\infty} \frac{4x^2 + 3x + 5}{8x^2 - 7} = \frac{1}{2}$$

For, so long as $x > 0$,

$$\frac{4x^2 + 3x + 5}{8x^2 - 7} = \frac{4 + 3/x + 5/x^2}{8 - 7/x^2} \tag{2.32}$$

Now, we can compute the desired limit by using the standard algebraic rules (the limit of a sum is the sum of the limits, etc.). (See Exercise 28.) Since $1/x$, $1/x^2$ tend to zero as $x \to \infty$, the limit of (2.32) as $x \to \infty$ is $4/8 = 1/2$.

51.

$$\lim_{x\to\infty} \frac{x\,|x|}{1 + x^2} = 1 \qquad \lim_{x\to-\infty} \frac{x\,|x|}{1 + x^2} = -1$$

If

$$x > 0, \frac{x\,|x|}{1 + x^2} = \frac{x^2}{1 + x^2} = \frac{1}{1 + 1/x^2}$$

if

$$x < 0, \frac{x\,|x|}{1 + x^2} = -\frac{x^2}{1 + x^2} = \frac{-1}{1 + 1/x^2}$$

52. $\lim_{x \to \infty} \arctan x = \pi/2.$

Definition 17 is the analog for functions defined on an infinite interval of the notion of convergence of a sequence (a function defined on the integers). Just as we pass from sequences to series we can pass from infinite limits of functions to infinite sums; that is, integrals over infinite intervals.

Definition 18. Let f be a continuous function on the interval $\{x: x \geq a\}$. We say f is **integrable** if $\lim_{x \to \infty} \int_a^x f$ exists, in which case we write the limit as $\int_a^\infty f$. f is **absolutely integrable** if $\lim_{x \to \infty} \int_a^x |f|$ exists.

Examples

53. x^{-2} is integrable on the interval $[1, \infty)$. For

$$\int_1^m x^{-2}\,dx = -x^{-1}\Big|_1^m = -\frac{1}{m} + 1$$

so

$$\int_1^\infty x^{-2}\,dx = \lim_{m \to \infty}\left(-\frac{1}{m} + 1\right) = 1$$

54. $x^{-1} \cos x$ is not absolutely integrable on the interval $[1, \infty)$. For

$$\int_1^\infty \left|\frac{\cos x}{x}\right| dx \geq \sum_{n=1}^\infty \int_{2\pi n - \pi/3}^{2\pi n + \pi/3} \frac{\cos x}{x}\,dx$$

Between $2\pi n - \pi/3$ and $2\pi n + \pi/3$, $x^{-1} \cos x \geq (2\pi n + \pi/3)^{-1} \cdot \frac{1}{2}$. Thus,

$$\int_1^\infty \left|\frac{\cos x}{x}\right| dx \geq \sum_{n=1}^\infty \frac{1}{2} \cdot \frac{1}{(2\pi n + \pi/3)} \cdot \frac{2\pi}{3} = \infty$$

The theory of integration on infinite intervals is entirely analogous to the theory of infinite series. We have the following facts (whose counterparts in the theory of series are easily recognized).

Proposition 21. *Let f be continuous on the interval $\{x : x \geq a\}$.*

(i) *f is absolutely integrable if and only if the set $\{\int_a^x |f|\}$ is bounded.*
(ii) *If f is absolutely integrable, then f is integrable.*
(iii) **(Comparison Test).** *If there exists a $b \geq a$ and a constant K and an integrable positive function g defined on $\{x\colon x > b\}$ such that $Kg \geq |f|$, then f is absolutely integrable.*

Proof.
(i) If $|f|$ is integrable, clearly $\{\int_a^x |f|\}$ is bounded. On the other hand, if $\{\int_a^x |f|\}$ is bounded, let $L = \sup\{\int_a^x |f|\}$. Then for $\varepsilon > 0$, $L - \varepsilon$ is not an upper bound, so there exists an x_0 such that $\int_a^{x_0} |f| \geq L - \varepsilon$. Then for all $x \geq x_0$,

$$L \geq \int_a^x |f| \geq \int_a^{x_0} |f| \geq L - \varepsilon$$

so

$$\left| L - \int_a^x |f| \right| < \varepsilon$$

(ii) Suppose $\int_a^\infty |f| = L$. Let $c_n = \int_a^n f$. We show that $\{c_n\}$ is a Cauchy sequence. Let $\varepsilon > 0$. Then there is an x_0 such that for $x \geq x_0$,

$$\left| \int_a^x |f| - L \right| < \frac{\varepsilon}{2}$$

Then for $n, m \geq x_0$,

$$|c_n - c_m| = \left| \int_n^m f \right| \leq \int_n^m |f| \leq \left| \int_a^n |f| - \int_a^m |f| \right|$$

$$\leq \left| \int_a^n |f| - L \right| + \left| \int_a^m |f| - L \right| < \varepsilon$$

Thus $\{c_n\}$ is Cauchy, so converges, say to c. We shall show that in fact $\int_a^\infty f = c$. Let $\varepsilon > 0$, and find N so that $|c_n - c| < \varepsilon/2$ for $n \geq N$. Then for $x \geq \max(x_0, N)$,

$$\left| \int_a^x f - c \right| \leq \left| \int_a^N f - c \right| + \int_N^x |f| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

as in the previous computation.

(iii) Under the given hypothesis, if $x \geq 1$, then

$$\int_{a}^{x} |f| \leq \int_{a}^{b} |f| + K \int_{b}^{\infty} g < \infty$$

Thus by (i), f is absolutely integrable.

Here is an easily derived relationship between the absolute convergence of series and integrals which provides yet another test for the convergence of series.

Proposition 22. (Integral Test) *Let f be a positive, decreasing function defined on R^+. Then $\int_1^\infty f$ exists if and only if $\sum_{n=1}^{\infty} f(n) < \infty$.*

Proof. For $x, n \leq x \leq n+1$ we have $f(n) \geq x \geq n+1$. Thus $f(n) \geq \int_n^{n+1} f \geq f(n+1)$. Thus, by comparison the series $\sum \int_n^{n+1} f$ and $\sum f(n)$ converge or diverge together. But the convergence of the first series is the same as the existence of $\int_1^\infty f$, and conversely.

This proposition gives an easy proof that $\sum 1/n^{(1+\varepsilon)} < \infty$ for $\varepsilon > 0$. (Compare to the work of Example 18.) For if we consider the integral $\int_1^\infty dt/t^{1+\varepsilon}$, we have

$$\int_1^x \frac{dt}{t^{1+\varepsilon}} = \frac{-1}{\varepsilon t^\varepsilon}\bigg|_1^x = \frac{1}{\varepsilon} - \frac{1}{\varepsilon x^\varepsilon} \to \frac{1}{\varepsilon}$$

as $x \to \infty$.

Example

55.

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^2} < \infty$$

For

$$\int_2^x \frac{dt}{t(\log t)^2} = \int_{\log 2}^{\log x} \frac{du}{u^2} = -u^{-1}\bigg|_{\log 2}^{\log x} = \frac{1}{\log 2} - \frac{1}{\log x}$$

Thus

$$\int_2^\infty \frac{dt}{t(\log t)^2} = \lim_{x\to\infty}\left(\frac{1}{\log 2} - \frac{1}{\log x}\right) = \frac{1}{\log 2} < \infty$$

• EXERCISES

28. Verify these algebraic properties of $\lim\limits_{x\to\infty}$. Suppose $\lim\limits_{x\to\infty} f(x)$, $\lim\limits_{x\to\infty} g(x)$ exist.

(a) $\lim\limits_{x\to\infty} f(x)+g(x)=\lim\limits_{x\to\infty} f(x)+\lim\limits_{x\to\infty} g(x)$.

(b) $\lim\limits_{x\to\infty} f(x)g(x)=\lim\limits_{x\to\infty} f(x)\lim\limits_{x\to\infty} g(x)$.

(c) $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)}=\dfrac{\lim\limits_{x\to\infty} f(x)}{\lim\limits_{x\to\infty} g(x)}$, if $\lim\limits_{x\to\infty} g(x)\neq 0$.

29. Compute these limits as $x\to\infty$.

(a) $\dfrac{\sin x}{x}$.

(b) $\dfrac{x^2+3x+1}{x^4+1}$.

(c) $\dfrac{x^2-1}{x^2+1}$.

(d) $\tan \dfrac{1}{x}$.

(e) $x \sin \dfrac{1}{x}$.

30. Which of these series converge:

(a) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n\log n}$

(b) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\log n)^2}$

(c) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\log\log n)^2}$

(d) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{(\log n)^2(\log\log n)^2}$

(e) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{\tan\left(\frac{\pi}{2}-\frac{1}{n}\right)}$

(f) $\displaystyle\sum_{n=2}^{\infty} \frac{\log n}{n^{3/2}}$

(g) $\displaystyle\sum_{n=2}^{\infty} \frac{(-1)^n}{(\log n)^2}$

(h) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{(\log n)^2}$

(i) $\displaystyle\sum_{n=2}^{\infty} \frac{1}{(n\sin n)^2}$

(j) $\displaystyle\sum_{n=3}^{\infty} \frac{1}{n\log(\log n)^{1+\varepsilon}}$

2.10 The Space of Continuous Functions

The mathematician attacks his problems with a certain store of techniques. Occasionally a problem will require the development of a new technique; more often the problem is solved by viewing it in one way, and then another and then again another until a viewpoint is obtained which allows for the application of one of those techniques. Sometimes if the viewpoint is clever enough, or profound enough—or naive enough—the applicable technique is quite elementary and surprising and leads to further deep discoveries. This is the case with the contraction lemma (a fixed point theorem) which we shall apply several times in this text to obtain some of the basic facts of calculus. First, in this section, we shall develop the particular viewpoint in the relevant context. It is simple enough—instead of looking at continuous functions one at a time, we consider them all.

Let us illustrate this with a particular problem. Suppose we are interested in finding a differentiable function with these properties:

$$f'(x) = f(x) \quad \text{for all } x \qquad \text{and} \qquad f(0) = 1 \tag{2.33}$$

To find such a function means first of all to verify that a solution to our problem exists, and secondly to establish some technique for computing it. We already have enough experience with calculus to know that this second objective will be hard to fulfill. What we in fact seek is a means of effectively approximating our solution. This provides a clue: let us look for a sequence of functions $\{f_n\}$ which converges to a function with the properties (2.33). Such a sequence would be a sequence of differentiable function $\{f_n\}$ such that the sequence $\{f_n(x)\}$ converges for all x, and $f'_n(x) = f_{n-1}(x)$. If we had such a sequence, we could take the limit and deduce that

$$\lim f'_n(x) = \lim f_{n-1}(x)$$

so $f(x) = \lim f_n(x)$ will solve our problem.

Now this is a good idea, because Equation (2.33) itself provides the technique for generating such a sequence. Let f_0 be any function, and define $f_1 = f'_0$. Then let $f_2 = f'_1$, $f_3 = f'_2$, and so forth. Will the sequence $\{f_n\}$ converge? Well, that is a problem. Notice that $f_2 = f'_1 = f''_0$, $f_3 = f'_2 = f'''_0$, and more generally $f_n = f_0^{(n)}$. Thus, we must be very careful to choose an infinitely differentiable function for f_0. Suppose f_0 is chosen as a polynomial of degree n. Then $f_{n+1} = f_0^{(n+1)} = 0$, and so all the rest of our functions are zero. Thus, the sequence certainly converges, but hardly to a solution, since the condition $f(0) = 1$ is not verified. In fact, this present approach has obviously petered out fruitlessly and it may be because we have not incorporated the initial condition $f(0) = 1$ in our approach. Can we put all of (2.33) in one statement, and then proceed with this technique of generating an approximating sequence? The fundamental theorem of calculus says yes; in fact, (2.33) can be rewritten as

$$f(x) = \int_0^x f(t)\,dt + 1 \tag{2.34}$$

This now is an operation involving integration rather than differentiation, and so we have the added advantage of not having to choose a very well-behaved function for the first approximant. Let us try again, with (2.34) rather than (2.33). Letting $f_0 = 1$, we find

$$f_1(x) = \int_0^x 1\,dt + 1 = x + 1$$

$$f_2(x) = \int_0^x (t + 1)\,dt + 1 = \frac{x^2}{2} + x + 1$$

$$f_3(x) = \int_0^x \left(\frac{t^2}{2} + t + 1\right) dt + 1 = \frac{x^3}{3!} + \frac{x^2}{2!} + x + 1$$

$$f_n(x) = \int_0^x f_{n-1}(t)\,dt + 1 = \frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!} + \cdots + \frac{x^2}{2!} + x + 1 \tag{2.35}$$

Now we're getting somewhere. We have already seen that the series (2.35) converges for any x. Thus, letting

$$f(x) = \lim f_n(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

this must be the sought after function. (Of course the reader has long since recognized the solution of our problem as being the exponential function. Thus he should be reassured to see that it did in fact turn out that way.) What we need now is the theoretical mathematics that *will allow us* to take the limit in (2.35) and correctly deduce

$$f(x) = \int_0^x f(t)\,dt + 1 = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Thus we are led to the question of convergence in the *space of continuous functions*. We now proceed to that theory.

Let X be a closed bounded set in R^n, and let $C(X)$ denote the **space of all continuous complex-valued functions on** X. We know that if f and g are two functions in $C(X)$, then so are $f+g$ and fg and cf, for c, a complex number. In particular, $C(X)$ is a vector space on which multiplication is defined. The vector space $C(X)$ is quite different from the vector spaces C^n, R^n: $C(X)$ is usually infinite dimensional (see Problem 49). $C(X)$ does not have any obvious "standard basis"—in fact, we wouldn't know how to choose one. In other particulars, however, $C(X)$ is not very different. There is in this space a reasonable notion of closeness. Two functions are close if their values are *everywhere* close; that is, if the maximum of their difference is small. This leads to a notion of length and distance in $C(X)$.

Definition 19. Let X be a closed and bounded set in R^n, and $C(X)$ the space of continuous functions on X. If $f \in C(X)$, the **length of** f is

$$\|f\| = \max\{|f(\mathbf{x})| : \mathbf{x} \in X\}$$

If f, g are in $C(X)$, the **distance between** f **and** g is $\|f - g\|$.

The properties of length and distance are just those of the corresponding notions in R^n:

$$\|cf\| = |c| \, \|f\|$$

$$\|f+g\| \leq \|f\| + \|g\|$$

If $\|f\| = 0$, then $f = 0$. What is important is what we can consider the notion of convergence of a sequence of continuous functions. We say that $f_n \to f$ if $\|f_n - f\| \to 0$, that is, if the distance between the general term of the sequence and f becomes arbitrarily small. This is the same as saying that the values of f_n at points of X converge to the values of f **in a uniform manner**. The value of these notions lies not only in their naturality, but in the now realizable possibility of finding specific functions satisfying given properties by techniques of approximation. Let us make this precise.

Definition 20. Let X be a closed bounded set, and $\{f_n\}$ a sequence in $C(X)$. We say that $\{f_n\}$ is **uniformly convergent** if there is an $f \in C(X)$ such that

$$\lim_{n \to \infty} \|f_n - f\| = 0$$

We say that the sequence is **uniformly Cauchy** if, for every $\varepsilon > 0$ there is an N such that

$$\|f_n - f_m\| < \varepsilon \qquad \text{whenever } n, m \geq N$$

Examples

56. Let x be the interval $[0, 1]$, $f_n(x) = (1 - x)x^n$. This sequence converges uniformly to zero. Let us compute $\max|f_n(x)| = \|f_n\|$.

$$f_n'(x) = n(1 - x)x^{n-1} - x^n$$

so $f_n'(x) = 0$ has the solutions $x = 0$, $x = n/(n + 1)$. Thus

$$\|f_n\| = \left(1 - \frac{n}{n+1}\right)\left(\frac{n}{n+1}\right)^n = \frac{1}{n+1}\left(\frac{n}{n+1}\right)^n$$

which tends to zero.

57. On the same interval the sequence $f_n(x) = \sin x/n$ tends to zero, for

$$\|f_n\| = \sin\frac{1}{n} \to 0 \qquad \text{as } n \to \infty$$

58. Consider the convergence of the sequence $\{nx \sin x/n\}$ on the interval $[0, 1]$. Now we know that $\sin x/n \to 0$ as $n \to \infty$, but $nx \to \infty$, so we cannot make any deduction about the product. We have to refine our information about $\sin x/n$. For large values of n, it is very close to x/n. Thus

$$nx \sin\frac{x}{n} \sim nx \cdot \frac{x}{n} = x^2 \tag{2.36}$$

so we guess that $nx \sin x/n \to x^2$. Let us prove it by computing

$$\left\|nx \sin\frac{x}{n} - x^2\right\| \tag{2.37}$$

In order to do that, let us provide an estimate to our guess (2.36).

$$\left\|\sin\frac{x}{n} - \frac{x}{n}\right\| \leq \frac{1}{n^2} \qquad \text{in the interval } [0, 1] \tag{2.38}$$

Then (2.37) becomes

$$\left\| nx \sin \frac{x}{n} - x^2 \right\| = \left\| nx\left(\sin \frac{x}{n} - \frac{x}{n}\right) + nx \cdot \frac{x}{n} - x^2 \right\|$$

$$= \left\| nx\left(\sin \frac{x}{n} - \frac{x}{n}\right) \right\| \tag{2.39}$$

$$\leq \|nx\| \left\| \sin \frac{x}{n} - \frac{x}{n} \right\| \tag{2.40}$$

$$\leq n \cdot \frac{1}{n^2} = n^{-1}$$

and since $n^{-1} \to 0$ as $n \to \infty$, we are through.

59. On the interval [0, 1] the sequence $\{\sin nx\}$ is not convergent. It is not even a Cauchy sequence. The distance $\|\sin nx - \sin mx\|$ does not become arbitrarily small as $n, m \to \infty$. In particular, if $m = 2n$, we have

$$\|\sin(nx) - \sin(2nx)\| \geq \left| \sin\left(n \cdot \frac{\pi}{2n}\right) - \sin\left(2n \cdot \frac{\pi}{2n}\right) \right| = 1$$

The basic theorem about convergence of continuous functions is the following, which plays the same role in $C(X)$ as the least upper bound axiom does for R. It provides the assertion of existence of functions with prescribed properties. In order to verify that a sequence of functions has a continuous limit, we need only verify that it is a uniformly Cauchy sequence.

Theorem 2.14. *A uniformly Cauchy sequence of continuous functions is uniformly convergent.*

Proof. Suppose $\{f_n\}$ is a uniformly Cauchy sequence of continuous functions on X. This means: for every $\varepsilon > 0$, there is an $N > 0$ such that $\|f_n - f_m\| < \varepsilon$ for $n, m \geq N$. This means precisely

$$|f_n(\mathbf{x}) - f_m(\mathbf{x})| < \varepsilon \qquad \text{for all } \mathbf{x} \in X \tag{2.41}$$

Thus, for each x, $\{f_n(x)\}$ is a uniformly Cauchy sequence of real numbers, and thus converges. Denote the limit, $\lim f_n(\mathbf{x})$ by $f(\mathbf{x})$. We must show that this function $\mathbf{x} \to f(\mathbf{x})$ is continuous, and that f_n converges uniformly to f.

First of all, if $\varepsilon > 0$, choose N as above, and let $m \to \infty$ in (2.41). We obtain, for $n > N$,

$$\lim_{m\to\infty} |f_n(\mathbf{x}) - f_m(\mathbf{x})| = |f_n(\mathbf{x}) - f(\mathbf{x})| \leq \varepsilon \qquad \text{for all } \mathbf{x} \in X$$

Thus, if $n \geq N$, $\|f_n - f\| \geq \varepsilon$. This implies that $\lim_{n\to\infty} \|f_n - f\| = 0$, as desired.

Now f is continuous. Fix $\mathbf{x}_0 \in X$. Let $\varepsilon > 0$ and choose N so large that $\|f_N - f\| < \varepsilon/3$. Since f_N is continuous, there is a $\delta > 0$ such that $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ implies $|f_N(\mathbf{x}) - f_N(\mathbf{x}_0)| < \varepsilon/3$. Then if $|\mathbf{x} - \mathbf{x}_0| < \delta$,

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{x}_0)| &\leq |f(\mathbf{x}) - f_N(\mathbf{x})| + |f_N(\mathbf{x}) - f_N(\mathbf{x}_0)| + |f_N(\mathbf{x}) - f(\mathbf{x}_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

as desired.

Having seen one vector space of functions, we can easily see them everywhere. The collection of bounded real-valued functions on a set X is a vector space over the reals. The collection of all bounded functions on X taking values in R^n is also a vector space; similarly, the space of continuous functions taking values in R^n. All the spaces here are endowed with the same concept of length:

$$\|f\| = \sup\{\|f(\mathbf{x})\| : \mathbf{x} \in X\}$$

Of even more interest are the spaces of functions on which is defined some analytic operations. For example, if I is an interval, the space of all real-valued functions which are differentiable on I is a vector space. The space $C^1(I)$ of all functions whose derivative is continuous is also a vector space, as is the space $C^{(n)}(I)$ of all functions which have continuous nth derivatives. The space $R(I)$ of functions which are integrable on I is a vector space. These (and other) examples are further elaborated in the exercises. Suffice it to say here that the mathematical theory which follows this point of view (called functional analysis) is a recent (20th-century) development which has had profound impact, not only in foundations of mathematics, but in the practical application of mathematics in all branches of science.

Let us return to the space $C(X)$ of continuous functions on a closed bounded set X in R^n. Once we begin thinking of these functions as points in a space, on which are defined such notions as distance and convergence, we are easily led to consider functions on that space. Naturally, such a function is **continuous** if it takes convergent sequences into convergent sequences.

Examples

60. Let $g \in C(X)$ and define $\phi(f) = fg$. ϕ is continuous, for if $f_n \to f$, that is, $\|f_n - f\| \to 0$, then

$$\|f_n g - fg\| \leq \|f_n - f\| \; \|g\| \to 0$$

61. Define $\psi\colon C(X) \to C(X)$, $\psi(f) = f^2$. ψ is also continuous, for

$$\|f_n^2 - f^2\| = \|(f_n - f)(f_n + f)\| \leq \|f_n - f\| \cdot \|f_n + f\| \tag{2.42}$$

If $f_n \to f$, the term $\|f_n + f\|$ remains bounded while $\|f_n - f\| \to 0$, thus also $\|f_n^2 - f^2\| \to 0$.

62. If P is any polynomial, $\psi_P(f) = P(f)$ is continuous on $C(X)$ (Problem 55).

63. Define $M\colon C(X) \to R$, $M(f) = \|f\|$.
This is continuous, since

$$\|M(f) - M(g)\| = |\; \|f\| - \|g\| \;| \leq \|f - g\|$$

64. Let $x_0 \in X$ and define $F_0\colon C(X) \to R$, $F_0(f) = f(x_0)$. Certainly F_0 is continuous: for if $f_n \to f$ in $C(X)$, then the maximum over X of $|f_n(x) - f(x)|$ tends to zero; in particular, $|f_n(x_0) - f(x_0)| \to 0$, so $F_0(f_n) \to F_0(f)$.

65. The definite integral is a continuous function on $C(I)$, where $I = [a, b] \subset R$. For

$$\left| \int_I f_n - \int_I f \right| \leq \left| \int_I (f_n - f) \right| \leq \|f_n - f\|(b - a)$$

so if $f_n \to f$, also $\int_I f_I \to \int_I f$. A stronger and more important statement than that of Example 65 is that the indefinite integral, as a function from $C(I)$ to $C(I)$ is continuous. This is contained in the next proposition.

Proposition 23. *Let $I = \{x \in R\colon a \leq x \leq b\}$. Suppose f_n is a sequence of continuous functions on I converging uniformly to f. Let $F_n(x) = \int_a^x f_n$, $F(x) = \int_a^x f$. Then $F_n \to F$ uniformly.*

Proof.

$$|F_n(x) - F(x)| = \left|\int_a^x (f_n - f)\right| \le \|f_n - f\|(x-a) \le \|f_n - f\|(b-a)$$

Thus, taking the maximum on the left,

$$\|F_n - F\| \le \|f_n - f\|(b-a)$$

so if $f_n \to f$ uniformly so also $F_n \to F$.

Problem 56 is intended to demonstrate that on the other hand, differentiation is *not* a continuous function on $C(I)$. (It isn't even everywhere defined; i.e., there are continuous functions that do not have a derivative.) Nevertheless, Proposition 23 has this consequence for differentiation.

Proposition 24. *Let $\{f_n\}$ be a sequence of continuously differentiable functions on the interval $[a, b]$ and suppose that* (i) *$\{f'_n\}$ is uniformly Cauchy,* (ii) *$f_n(a) = 0$ for all n. Then $\{f_n\}$ is uniformly convergent to a differentiable function f and $f' = \lim f'_n$.*

Proof. The proof of this proposition consists in a rereading of Proposition 23 via the fundamental theorem of calculus. By that theorem

$$f_n(x) = \int_a^x f'_n$$

so by Proposition 23, f_n is also convergent. If we let $g = \lim f'_n$, then $\lim f_n = \int_a^x g$. Thus, $\lim f_n$ is indeed differentiable and its derivative is $g = \lim f'_n$.

Let us return now to the consideration of our original problem. In fact, let us generalize it slightly. Let c be a complex number, and let us seek a differentiable complex-valued function f such that

$$f'(x) = cf(x) \quad \text{for all } x \qquad \text{and} \qquad f(0) = 1 \tag{2.43}$$

This is, by the fundamental theorem of calculus the same as seeking a continuous function f such that

$$f(x) = c\int_0^x f(t)\,dt + 1 \tag{2.44}$$

Now that we have the necessary theory and point of view available, we may follow a more sophisticated approach. Let I be the interval $I = [-R, R]$, and define the function T on $C(I)$:

$$Tf(x) = c\int_0^x f(t)\,dt + 1 \tag{2.45}$$

We seek a function f such that $f = Tf$, that is, a **fixed point** of the transformation. Our technique is that of successive approximation. Let f_0 be any continuous function, and define $f_1 = Tf_0$, $f_2 = Tf_1 = T^2 f_0$, and in general $f_n = Tf_{n-1} = T^n f_0$. We must show that the sequence $\{f_n\}$ converges. If we choose $f_0 = 1$ we can compute the sequence explicitly, and we find that

$$f_n(x) = \frac{(cx)^n}{n!} + \frac{(cx)^{n-1}}{(n-1)!} + \cdots + cx + 1$$

Then if $m > n$,

$$f_m(x) - f_n(x) = \frac{(cx)^m}{m!} + \frac{(cx)^{m-1}}{(m-1)!} + \cdots + \frac{(cx)^{n+1}}{(n+1)!}$$

On the interval $[-R, R]$ the maximum of this expression is dominated by replacing c by $|c|$, and x by R. Thus,

$$\|f_m - f_n\| = \frac{(|c|\,R)^m}{m!} + \frac{(|c|\,R)^{m-1}}{(m-1)!} + \cdots + \frac{(|c|\,R)^{n+1}}{(n+1)!}$$

$$\|f_m - f_n\| = \sum_{k=0}^{m} \frac{(|c|\,R)^k}{k!} - \sum_{k=0}^{n} \frac{(|c|\,R)^k}{k!} \tag{2.46}$$

Since the series

$$\left\{\sum_{k=0}^{\infty} \frac{(|c|\,R)^k}{k!}\right\}$$

converges, its sequence of partial sums is a Cauchy sequence, so by (2.46), $\{f_n\}$ is a Cauchy sequence and is thus uniformly convergent. Since T is continuous on $C(I)$, we have

$$\lim f_n = \lim T(f_{n-1}) = T(\lim f_{n-1}) = T(\lim f_n)$$

so $\lim f_n$ solves the given problem. This function is important enough for us to spend a few more paragraphs discussing it.

Definition 21. The **exponential function**, denoted $\exp(cx)$, or e^{cx}, for any complex number c is the solution of the differential equation

$$f'(x) = cf(x) \qquad f(0) = 1$$

First of all, this definition makes sense, because there *is only one* solution. If g also solves, then

$$\frac{d}{dx}\left[\frac{e^{cx}}{g}\right] = \frac{ce^{cx}g - e^{cx}g'}{g^2} = \frac{ce^{cx}g - ce^{cx}g}{g^2} = 0$$

since $g' = cg$. Thus $e^{cx}g^{-1}$ is constant. Since its value at 0 is 1, $e^{cx}g^{-1} \equiv 1$, or $e^{cx} \equiv g$. From these discussions we have these additional properties of the exponential function

Proposition 25.

(i) $e^{cx} = \sum_{n=0}^{\infty} \frac{(cx)^n}{n!}$

(ii) $e^{x+y} = e^x e^y$.

(iii) e^{cx} is *never zero*.

Proof. Part (i) follows directly from the argument above. Part (ii) follows from the uniqueness. Fix y, and define $h(x) = e^{x+y}/e^y$. Then

$$h'(x) = \frac{e^{x+y}}{e^y} = h(x) \quad \text{and} \quad h'(0) = \frac{e^{0+y}}{e^y} = 1$$

Thus we must have $h(x) = e^x$, so (ii) is verified. Part (iii) follows immediately from (ii):

$$e^{cx}e^{-cx} = e^{cx-cx} = e^0 = 1$$

so $(e^{cx})^{-1} = e^{-cx}$.

• PROBLEMS

49. Let I be a nonempty interval in R. Show that $C(I)$ is infinite dimensional.

50. Show that the sequence of functions on the closed unit disk in C defined by

$$f_n(x) = \sum_{k=1}^{n} \frac{x^k}{k^2}$$

converges.

51. Does the sequence $\left\{\sum_{k=1}^{n} z^k/k\right\}$ converge on the closed unit disk?

52. Let $\{a_n\}$ be a sequence of complex numbers such that $\sum |a_n| < \infty$. Verify these facts:

(a) For every z, $|z| \leq 1$, $f(z) = \sum_{n=1}^{\infty} a_n z^n$ converges, and

$$|f(z)| \leq \sum_{n=1}^{\infty} |a_n|$$

(b) f is continuous on $\{z \in C\colon |z| \leq 1\}$. This is true because f is the uniform limit of the polynomials $f_N(z) = \sum_{n=1}^{N} a_n z^n$, since $\|f - f_N\| \leq \sum_{n=N+1}^{\infty} |a_n| \to 0$ as $N \to \infty$.

53. Let f, g be continuous functions on the closed and bounded set X. Show that $\|fg\| \leq \|f\| \cdot \|g\|$. Is $\|fg\| < \|f\| \cdot \|g\|$ possible?

54. Show that on the interval $[0, 1]$,

$$\left\| \sin \frac{x}{n} - \frac{x}{n} \right\| \leq \frac{1}{n^2} \qquad \text{for all } n$$

55. Let $\mathbf{x}_1, \ldots, \mathbf{x}_k \in X$ and p be any polynomial in k variables. Define $\Psi\colon C(X) \to C$

$$\Psi(f) = p(f(\mathbf{x}_1), \ldots, f(\mathbf{x}_k))$$

Show that Ψ is continuous.

56. Find a sequence $\{f_n\}$ of differentiable functions which is uniformly convergent, but such that $\{f'_n(\frac{1}{2})\}$ is *not* convergent.

2.11 The Fixed Point Theorem

The fixed point theorem is a generalization of the technique of successive approximations described above in the discussion of the exponential function. This technique was first used by Newton as a technique for finding roots of polynomial equations. Simply stated, Newton's method is this. First, a technique is described by means of which one can transform a given approximation to a root into a better approximation. One then chooses a reasonable approximation, applies this technique to it to find a better one. Having this, one again applies the technique: if it's a good one, the result is an even better approximation. Continuing in this way, one obtains a sequence of approximations which should converge to the root. Now, having described the procedure, let us turn to Newton's specific technique for bettering approximations.

Let f be a given real polynomial. We want to find a point x_0 such that $f(x_0) = 0$. Choose a p_1 so that $f(p_1)$ is small. Now, replace the function by its linear approximation at p_1: $L(x) = f(p_1) + f'(p_1)(x - p_1)$, and let p_2 be the root of $L(x) = 0$. In other words, replace the graph of f by its tangent line and let p_2 be the x intercept of that line (see Figure 2.17). Now apply this procedure to p_2. Let p_3 be the root of the linear approximation to f at p_2, and so forth. We can describe Newton's technique abstractly as follows: For any point p, let $T(p)$ be the zero of the linear approximation of f at p: $T(p)$ solves the equation $f(p) + f'(p)(T(p) - p) = 0$. (We must

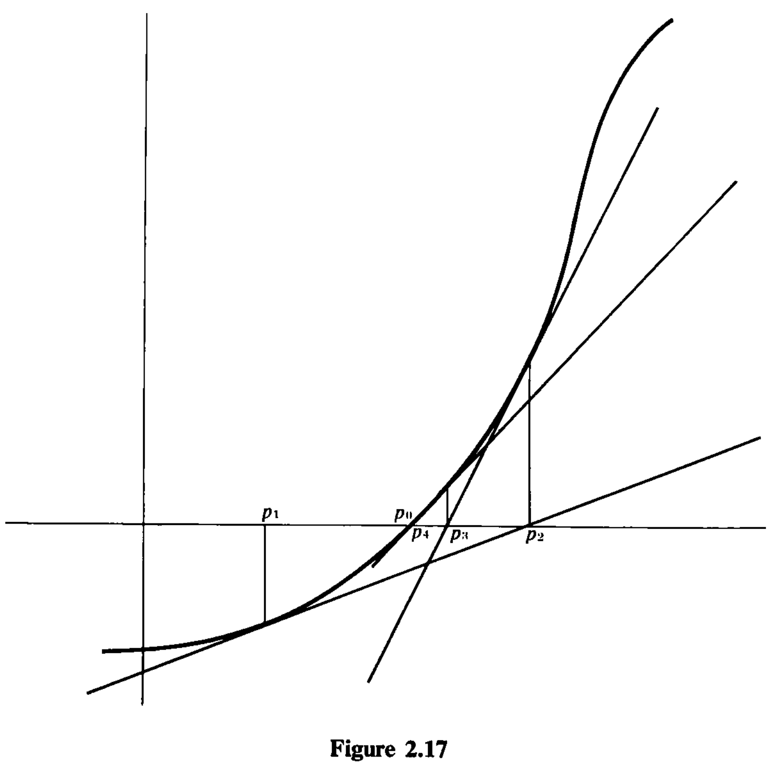

Figure 2.17

assume that $f' \neq 0$ for T to be a well-defined function.) Clearly, if $f(p) = 0$, we have $T(p) = p$, and conversely, thus we are in reality seeking a fixed point of T!

Suppose T has the property of **contraction** on some interval I. There is a $c < 1$ such that $|Tx - Ty| < c|x - y|$, all $x, y \in I$. Then Newton's method works. There is a root of $f(x) = 0$ (or $f'(x) = 0$) on the interval I, and it is the limit of the sequence x_0, Tx_0, $T^2x_0, \ldots,$ where x_0 is any point of I. This is the content of the fixed point theorem.

We now state and prove it explicitly for subsets of $C(X)$. It will be clear that the theorem is true for subsets of R^n, by virtue of the same argument.

Theorem 2.15. *Suppose S is a closed set of functions in $C(X)$: that S contains all limits of sequences in S. Suppose T is a mapping of S onto S which is a contraction, that is, there is a $c < 1$ such that*

$$\|T(f) - T(g)\| \leq c\|f - g\| \qquad \text{for all } f, g \in S$$

Then there is a unique continuous function f_0 such that $T(f_0) = f_0$.

Proof. Certainly the fixed point is unique. For if $T(f_0) = f_0$ and $T(f_1) = f_1$, then $\|f_0 - f_1\| = \|T(f_0) - T(f_1)\| \leq c\,\|f_0 - f_1\| < \|f_0 - f_1\|$ unless $\|f_0 - f_1\| = 0$, that is, $f_0 = f_1$.

Now let $f \in C(X)$. Let the sequence $\{f_n\}$ be defined as follows: $f_1 = f$, $f_2 = Tf_1$, $f_3 = Tf_2, \ldots, f_n = Tf_{n-1}$. $\{f_n\}$ is a Cauchy sequence. For

$$\|f_{n+1} - f_n\| = \|Tf_n - Tf_{n-1}\| \leq c\,\|f_n - f_{n-1}\|$$

so we can verify by induction that

$$\|f_{n+1} - f_n\| < c^n\,\|f_1 - f_0\|$$

Thus, for $m > n$ we have

$$\|f_m - f_n\| \leq \left\|\sum_{j=n}^{m-1}(f_{j+1} - f_j)\right\| \leq \sum_{j=n}^{m-1} \|f_{j+1} - f_j\|$$
$$\leq \left(\sum_{j=n}^{m-1} c^j\right) \|f_1 - f_0\| < \|f_1 - f_0\|\frac{c^n}{1 - c}$$

Since $c < 1$, $\{f_n\}$ is Cauchy, so has a limit $f_0 \in C(X)$. Since T is continuous, $Tf_0 = \lim_{n\to\infty} Tf_n = \lim_{n\to\infty} f_{n+1} = f_0$, and thus f_0 is the desired fixed function.

As an illustration on the real numbers let us prove that if $a > 0$, there is an $x_0 > 0$ such that $x_0^2 = a$, by Newton's method. First, we describe the map T. Let $p > 0$, the linear approximation to $x^2 - a$ at p is $p^2 - a + 2p(x - p)$. Thus, the zero of this linear polynomial is

$$Tp = \frac{a - p^2}{2p} + p = \frac{1}{2}\left(p + \frac{a}{p}\right)$$

Clearly, if T has a fixed point x_0, we must have $x_0{}^2 = a$. Thus, we must show that T is a contraction on some closed interval:

$$\begin{aligned} |Tx - Ty| &= \frac{1}{2}\left|x - y + \frac{a}{x} - \frac{a}{y}\right| = \frac{1}{2}\left|x - y + \frac{a}{xy}(y - x)\right| \\ &= \frac{1}{2}|x - y|\left|1 - \frac{a}{xy}\right| \end{aligned}$$

Since a, x, y are all positive, $1 - (a/xy) \leq 1$, so we need only ensure that $1 - (a/xy) \geq -1$, for T to be a contraction with $c = \frac{1}{2}$. Let $I = \{x: x^2 \geq a/2\}$. Then for $x, y \in I$, $xy > a/2$, so $a/xy \leq 2$, which is the desired inequality. Thus, by the fixed point theorem there is an x_0 with $x_0{}^2 \geq a/2$ such that $x_0{}^2 = a$.

We shall now give a somewhat more subtle application of the fixed point theorem. Sometimes a relation between two real variables determines one as a function of the other. For example, the relation $x + y = 0$ determines y as a function of x: $y = -x$; $x^2 + y^2 = 1$ gives $y = (1 - x^2)^{1/2}$ near the value $(0, 1)$, and near $(1, 0)$ we should write $x = (1 - y^2)^{1/2}$ as a function of y. The relations

$$e^{xy} = 1 \qquad \sin(x(\log y)) = 0$$

are somewhat less transparent, nevertheless we can ask whether or not they do determine y as a function of x.

Suppose now, in general we have an equation (see Figure 2.18)

$$F(x, y) = 0 \tag{2.47}$$

defined in the plane. We ask: does there exist a function g of x such that (2.47) amounts to saying $y = g(x)$? More precisely, is there a function g such that

$$F(x, y) = 0 \quad \text{if and only if} \quad y = g(x)$$

It is not hard to find a necessary condition. For there to be such a function

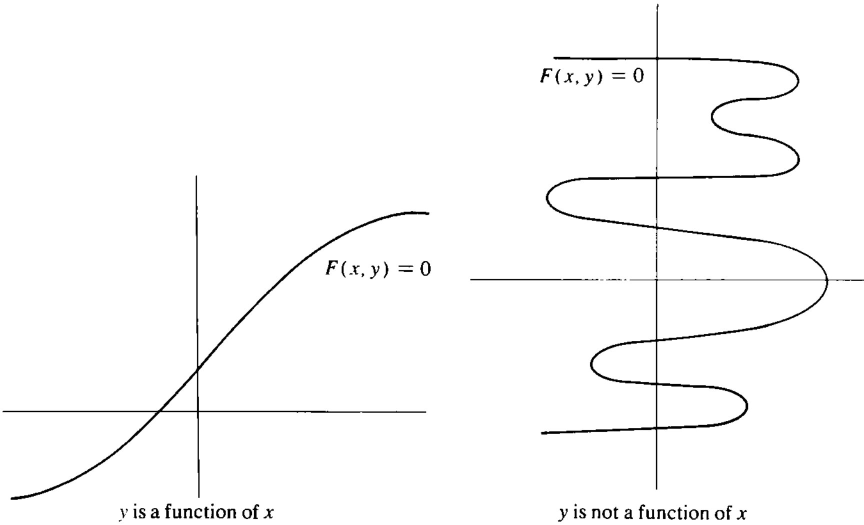

Figure 2.18

it must be the case that each line $x = \text{constant}$ intersects the set $F(x, y) = 0$ in only one point (see Figure 2.19). Thus the function $F(x, y)$, as a function of y on lines $x = \text{constant}$ must take the value 0 only once. The root of $F(x, y) = 0$ is then the value $g(x)$. Now we recall from one-variable theory that a function $H(y)$ will take all values once if $H'(y) \neq 0$. Thus the reasonable condition to impose on F is that it has a continuous partial derivative with respect to y, and $\partial F/\partial y \neq 0$. This condition turns out to be enough.

More precisely, suppose that F is defined and has continuous partial derivatives in the neighborhood of the origin in R^2, and $\partial F/\partial y(0, 0) \neq 0$.

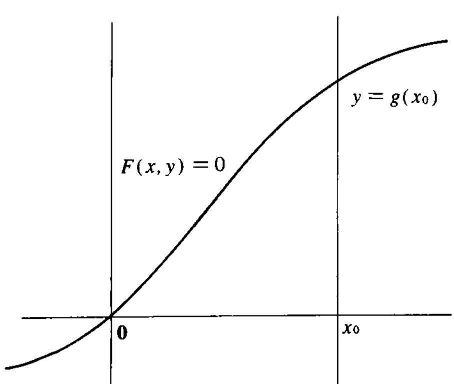

Figure 2.19

We seek a function g defined in a neighborhood of $x = 0$ such that $g(0) = 0$ and $F(x, g(x)) = 0$. If we fix $x = x_0$ near 0, then we seek a root of $F(x_0, y) = 0$. This brings us right back to Newton's method. Define T as a function of y as Newton did: $T(y)$ is the zero of the linear approximation of $F(x_0, y)$ at y; that is,

$$F(x_0, y) + \frac{\partial F}{\partial y}(x_0, y)(Ty - y) = 0$$

or

$$Ty = y - \left[\frac{\partial F}{\partial y}(x_0, y)\right]^{-1} F(x_0, y) \tag{2.48}$$

Just as in Newton's case the solution of $F(x_0, y) = 0$ is the fixed point of T. Thus, we need only verify that T is a contraction in some interval of values of y for x_0 near x so that it will have a fixed point; and we define $g(x_0)$ to be that fixed point. This application of the fixed point theorem really works, as we now shall prove.

Theorem 2.16. *Suppose that F has continuous partial derivatives in a neighborhood of* $(0, 0)$, *and that* $F(0, 0) = 0$, $\partial F/\partial y(0, 0) \neq 0$. *Then there is a function g defined for x in some interval* $(-\varepsilon, \varepsilon)$ *such that*

$$F(x, y) = 0 \quad \textit{if and only if} \quad y = g(x)$$

Proof. Instead of (2.48) we consider something slightly simpler. For x near 0, define

$$T_x(y) = y - \left[\frac{\partial F}{\partial y}(0, 0)\right]^{-1} F(x, y) \tag{2.49}$$

We want to find the fixed point, if it exists, of (2.49). Thus we seek suitable intervals, $-\varepsilon < x < \varepsilon$, $-\eta < y < \eta$ in which T_x is a contraction

$$T_x(y_1) - T_x(y_2) = y_1 - y_2 - \left[\frac{\partial F}{\partial y}(0, 0)\right]^{-1} [F(x, y_1) - F(x, y_2)] \tag{2.50}$$

By the mean value theorem there is a ξ between y_1 and y_2 such that

$$F(x, y_1) - F(x, y_2) = \frac{\partial F}{\partial y}(x, \xi)(y_1 - y_2)$$

Equation (2.50) becomes, upon substitution,

$$T_x(y_1) - T_x(y_2) = (y_1 - y_2)\left[1 - \frac{\partial F}{\partial y}(0, 0)^{-1}\frac{\partial F}{\partial y}(x, \xi)\right] \tag{2.51}$$

Now the term in brackets is continuous in (x, ξ) and has the value 0 at $(0, 0)$. Thus we may choose ε so that that term is less than $\frac{1}{2}$ if $-\varepsilon < x < \varepsilon$, $-\varepsilon < y_1 < \varepsilon$, $-\varepsilon < y_2 < \varepsilon$ and ξ is between y_1 and y_2. With this choice of ε, (2.51) gives

$$|T_x(y_1) - T_x(y_2)| < \tfrac{1}{2}|y_1 - y_2|$$

so T_x is indeed a contraction. Define $g(x)$ as the fixed point of T_x. Then, if $F(x, y) = 0$, then by (2.49) $T_x(y) = y$, so we must have $y = g(x)$. On the other hand, if $y = g(x)$, then $T_x(y) = y$, so again by (2.49) we must have $F(x, y) = 0$. The theorem is proved.

To say that the function g exists is already good enough, but much more is true: g is a continuously differentiable function. We will leave the verification of this fact to the interested reader (see Problem 58). In Section 7.2 we shall reconsider this theorem (known as the **implicit function theorem**) in many more variables. The beauty of the fixed point theorem is that the general context does not at all complicate the ideas, nor the verifications.

● EXERCISES

31. Find, by Newton's method, a sequence of numbers converging to the square root of a, for any $a > 0$. Now, do the cube root.

32. Find a sequence converging to a root of these polynomials:
(a) $x^3 + x^2 + x + 1$ (c) $x^3 - 2x^2 - 3x + 2$
(b) $x^2 - x + 1$ (d) $x^5 - x - 1$

33. (a) Let $F(x, y) = x\sin(xy)$. For what values of (x, y) such that $F(x, y) = 0$ is it true that nearby the equation $F(x, y) = 0$ defines y as a function of x?
(b) Same problem for
(i) $F(x, y) = xy^2 + 2xy + 1$, (ii) $F(x, y) = x^y - y$,
(iii) $F(x, y) = x^2 + y^2$.

34. Let $F(x, y)$ be differentiable in a domain D, and $(x_0, y_0) \in D$ such that $F(x_0, y_0) = 0$. Suppose g is differentiable and has the property $g(x_0) = y_0$, $F(x, g(x)) = 0$. Show that

$$g'(x_0) = -\frac{\partial F/\partial x(x_0, y_0)}{\partial F/\partial y(x_0, y_0)}$$

35. Find g' where g is defined implicitly by
(a) $x \sin(xy) = 0$ (c) $e^{xy} = 1$
(b) $\cos(x + y) = y$ (d) $e^{xy} = y$

• PROBLEMS

57. Prove the fixed point theorem in R^n:

Theorem *If S is a subset of R^n and T is defined on $\bar{S}$ and is a contraction on S, then there is a unique $\mathbf{y}_0 \in \bar{S}$ such that $T(\mathbf{y}_0) = \mathbf{y}_0$.*

58. Let F have continuous partial derivatives near (x_0, y_0) and suppose $F(x_0, y_0) = 0$, $\partial F/\partial y(x_0, y_0) \neq 0$. Let g be the function described in Theorem 2.15 ($F(x, y(x)) = 0$ and $g(x_0) = y_0$). We can prove that g is differentiable as follows.

(a) First of all, by the mean value theorem, for any (x, y), there is a (ξ, η) on the line between (x_0, y_0) and (x, y) such that

$$F(x, y) - F(x_0, y_0) = \frac{\partial F}{\partial x}(\xi, \eta)(x - x_0) + \frac{\partial F}{\partial y}(\xi, \eta)(y - y_0)$$

Why is the mean value theorem applicable?

(b) Now, if we substitute $y = g(x)$, $y_0 = g(x_0)$, we have

$$0 = \frac{\partial F}{\partial x}(\xi, \eta)(x - x_0) + \frac{\partial F}{\partial y}(\xi, \eta)(g(x) - g(x_0))$$

Thus

$$\frac{g(x) - g(x_0)}{x - x_0} = \frac{-\partial F/\partial x(\xi, \eta)}{\partial F/\partial y(\xi, \eta)}$$

Conclude that g is differentiable and

$$g'(x_0) = -\frac{\partial F/\partial x(x_0, g(x_0))}{\partial F/\partial y(x_0, g(x_0))}$$

2.12 Summary

A *sequence* $z_1, \ldots, z_n, \ldots$ of complex numbers is a function from the positive integers to C. The sequence $\{z_n\}$ *converges* to z if, for every $\varepsilon > 0$ there is an N such that $|z_n - z| < \varepsilon$ for $n \geq N$.

A convergent sequence is bounded, but not conversely. A monotonic bounded sequence of real numbers *is* convergent. *Cauchy criterion*: a sequence $\{z_n\}$ converges if, for every $\varepsilon > 0$, there is an N such that $|z_n - z_m| < \varepsilon$ for both $n, m \geq N$.

The *series* formed of a sequence $\{z_n\}$ is the sequence of sums $\{\sum_{i=1}^{n} z_i\}$. If the sequence of sums converges, we say that the *series converges* and denote the limit by $\sum_{n=1}^{\infty} z_n$. If $\sum z_n$ converges, then $z_n \to 0$, but not conversely. If $\{c_k\}$ is a sequence of nonnegative numbers, $\sum c_k$ converges if and only if the sequence $\sum_{k=1}^{n} c_k$ is bounded. A series $\sum z_n$ *converges absolutely* if $\sum |z_n| < \infty$. Absolutely convergent series may be summed in any convenient way.

Tests for Convergence

COMPARISON TEST. Suppose $|z_n| \leq |w_n|$ for all but finitely many n. Then (i) if $\sum |w_n|$ converges, $\sum z_n$ is absolutely convergent, (ii) if $\sum |z_n|$ diverges, so does $\sum |w_n|$.

ROOT TEST. If $|c_n|^{1/n} \leq r$ for some $r < 1$ and all but finitely many n, $\sum c_n$ is absolutely convergent.

RATIO TEST. If $|c_{n+1}/c_n| \leq r$ for some $r < 1$ and all but finitely many n, $\sum c_n$ is absolutely convergent.

The sequence $\{\mathbf{v}_k\}$ of vectors in R^n is said to *converge to* $\mathbf{v}$ if, for every $\varepsilon > 0$, there is an N such that $\|\mathbf{v}_k - \mathbf{v}\| < \varepsilon$ for $k > N$. A sequence of vectors converges if and only if it does so in each coordinate.

A set S is closed if and only if $\mathbf{v}_k \in S$, $\lim \mathbf{v}_k = \mathbf{v}$ implies $\mathbf{v} \in S$ also. Every sequence contained in a closed and bounded set has a convergent subsequence.

An R^m-valued function defined in R^n is said to be *continuous* at $\mathbf{v}_0$ if f is defined in a neighborhood of $\mathbf{v}_0$ and $\mathbf{v}_k \to \mathbf{v}_0$ implies $f(\mathbf{v}_k) \leq f(\mathbf{v}_0)$. A function is *continuous* on a set S if it is continuous at every point of S. If S is a closed and bounded set, and f is a continuous real-valued function defined on S, then f is bounded and attains its maximum and minimum.

Sections 2.6 and 2.7 are mainly about integration. We shall not recollect the definitions here; only the major results.

FUNDAMENTAL THEOREM OF CALCULUS. Suppose f is continuous on the interval $[a, b]$. Then the integral

$$F(x) = \int_a^x f$$

exists for all $x \in [a, b]$. F is differentiable on (a, b) and $F' = f$.

FUBINI'S THEOREM. Let f be an integrable function defined on a rectangle $R = I_1 \times \cdots \times I_n$ in R^n. $\int f$ can be computed by iteration:

$$\int_R f = \int_{I_1}\left[\cdots\left[\int_{I_n} f(x^1, \ldots, x^n)\, dx^n\right] dx^{n-1} \cdots\right] dx^1$$

Let f be a real-valued function defined in a neighborhood of $\mathbf{x}_0$ in R^n. If $\mathbf{v}$ is a vector in R^n, the directional derivative $df(\mathbf{x}_0, \mathbf{v})$ of f at x_0 in the direction $\mathbf{v}$ is defined by

$$\lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}$$

(if it exists). The partial derivative of f with respect to x^i at $\mathbf{x}_0$ is

$$\frac{\partial f}{\partial x^i}(\mathbf{x}_0) = df(\mathbf{x}_0, \mathbf{E}_i)$$

If these partial derivatives are all defined and continuous near $\mathbf{x}_0$, then $df(\mathbf{x}_0, \mathbf{v})$ is linear in $\mathbf{v}$. We can write

$$df = \sum \frac{\partial f}{\partial x^i}\, dx^i$$

If the partial derivatives $\partial f/\partial x^i$ all exist in an open set we may be able to compute the derivatives $\partial(\partial f/\partial x^i)/\partial x^j$. These are the second-order partial derivatives. If all first and second derivatives of f exist and are continuous in an open set N, then

$$\frac{\partial^2 f}{\partial x^i\, \partial x^j} = \frac{\partial^2 f}{\partial x^j\, \partial x^i}$$

throughout N.

Suppose that f has continuous partial derivatives in the domain $I \times D$, where I is an interval of reals, and D is a domain in R^n. Let

$$F(x) = \int_D f(x, \mathbf{y})\, d\mathbf{y}$$

Then F is differentiable and

$$\frac{dF}{dx}(x) = \int_D \frac{\partial f}{\partial x}(x, \mathbf{y})\, d\mathbf{y}$$

Suppose f is a real-valued function defined on R. We say that $f(x)$ converges to L as $x \to \infty$ written $\lim_{x \to \infty} f(x) = L$ if $|f(x) - L|$ can be made arbitrarily small by taking x sufficiently large. If now f is a continuous function on R such that

$$\lim_{x \to \infty} \int_a^x f$$

exists, we say that f is integrable on R. If $\lim_{x \to \infty} \int_a^x |f|$ exists, f is absolutely integrable. *Integral test*: If f is a positive, decreasing continuous function defined on R, then $\int_1^\infty f$ exists if and only if $\sum_{n=1}^{\infty} f(n) < \infty$.

Let X be a closed and bounded set in R^n. We denote by $C(X)$ the collection of all complex-valued continuous functions on X. $C(X)$ is a vector space. If f is in $C(X)$, the *length* of f is

$$\|f\| = \max\{|f(\mathbf{x})| : \mathbf{x} \in X\}$$

For f, g in $C(X)$ the distance between f and g is $\|f - g\|$. If $\{f_n\}$ is a sequence in $C(X)$, and $\|f_n - f\| \to 0$ as $n \to \infty$ for some $f \in C(X)$, we say that $\{f_n\}$ converges uniformly to f. *Cauchy criterion*. Suppose $\{f_n\}$ is a sequence in $C(X)$ satisfying the following condition: for each $\varepsilon > 0$, there is an N such that $\|f_n - f_m\| < \varepsilon$ whenever $n, m \geq N$. Then there is an $f \in C(X)$ such that $f_n \to f$ uniformly.

INTEGRATION. If X is an interval in R, and $f_n \to f$ uniformly in $C(X)$ then also $\int_a^x f_n \to \int_a^x f$ uniformly.

The exponential function, denoted $\exp(cx)$, or e^{cx} for any complex number c is the solution of the differential equation $y' = cy$, $y(0) = 1$. It has these properties:

$$e^{cx} = \sum_{n=0}^{\infty} \frac{(cx)^n}{n!}$$

$$e^{c(x+y)} = e^{cx}e^{cy}$$

e^{cx} is never zero.

FIXED POINT THEOREM. *Let S be a closed set of functions in $C(X)$ and T a mapping of S onto S which is a contraction; that is, there is $c < 1$ such that*

$$\|T(f) - T(g)\| \leq c\|f - g\| \qquad \text{for all } f, g \in S$$

Then there is a unique continuous function f_0 such that $T(f_0) = f_0$.

IMPLICIT FUNCTION THEOREM. Suppose that F has continuous partial derivatives in a neighborhood of $(0, 0)$, and that $F(0, 0) = 0$, $\partial F/\partial y(0, 0) \neq 0$. Then there is a function g defined for x in some interval $(-\varepsilon, \varepsilon)$ such that

$$F(x, y) = 0 \quad \text{if and only if} \quad y = g(x)$$

● FURTHER READING

M. Spivak, *Calculus*, Benjamin, New York, 1967. This is an eloquent text in the one-variable calculus. It is an excellent reference for a full treatment of the material in this chapter.

T. A. Bak and J. Lichtenberg, *Mathematics for Scientists*, Benjamin, New York, 1966. This is a review of the theory of calculus from the point of view of the physical scientist. It includes a chapter on numerical analysis.

C. W. Burrill and J. R. Knudsen, *Real Variables*, Holt, Rinehart and Winston, New York, 1969. An advanced text, going thoroughly through the material of this chapter and beyond to the theory of Lebesque integration.

● MISCELLANEOUS PROBLEMS

59. Let $\{x_n\}$ and $\{y_n\}$ be sequences. Then $\{x_n + y_n\}$ is also a sequence. So also is $\{rx_n\}$ for any real number r; thus the collection S of all real sequences is a vector space. Show that it is *not* finite dimensional.

60. Show that the collection B of bounded sequences is a linear subspace of the vector space S of all sequences (Problem 59).

61. Show that the collection C of convergent sequences is a linear subspace of B. Also C_0, the collection of all sequences converging to zero is a linear subspace of B. These spaces are all infinite dimensional.

62. Define the function "lim" on convergent sequences in the obvious way: $\lim: C \to R$: $\lim\{x_n\} = \lim x_n$. Show that lim is a linear function.

63. What is the dimension of the space of linear functionals on C which annihilate C_0?

64. Let $x_1 = 4$, $x_2 = \frac{1}{2}(4 + \frac{3}{4})$, and once $x_2, \ldots, x_n$ are defined, let $x_{n+1} = \frac{1}{2}(x_n + 3/x_n)$. Prove that $\{x_n\}$ converges. Assuming that, find the limit.

65. (a) Show that for every integer k,

$$\lim n^k/(n+1)^k = 1$$

$$\lim n^k/(n+1)^{k+1} = 0$$

$$\lim n^{k+1}/(n+1)^k \text{ does not exist}$$

(b) Let k be an integer, and $1 > h > 0$. Show that $\lim n^k h^n = 0$.
(c) Show that $\lim n/h^n$ does not exist.

66. Let $x_1 = 1$, and in general

$$x_{n+1} = 3 \cdot \frac{1 + x_n}{3 + x_n}$$

Find $\lim x_n$.

67. Suppose $\lim z_n = z$.

(a) Let $y_n = \frac{1}{2}(z_{n-1} + z_n)$. Then $\lim y_n = z$.

(b) Let k be a positive integer. Now let $\{y_n\}$ be defined by

$$y_n = \frac{1}{k+1}(z_n + z_{n+1} + \cdots + z_{n+k})$$

Then $\lim y_n = z$ also.

(c) This time take

$$y_n = \frac{1}{n}(z_1 + \cdots + z_n)$$

Once again $\lim y_n = z$.

68. Suppose that f is continuous at c, and $\lim c_n = c$. Then $\lim f(c_n) = f(c)$.

69. Let $\{c_n\}$ be a sequence of complex numbers, and suppose $(|c_n|)^{1/n} = R$. Show that R^{-1} is the radius of convergence of $\sum c_n z^n$.

70. Let $\{s_n\}$, $\{t_n\}$ be two sequences of positive numbers such that $\lim s_n t_n^{-1}$ exists and is nonzero. Then $\sum s_n$ converges if and only if $\sum t_n$ converges.

71. Let $\{c_n\}$ be a sequence of positive numbers. Suppose that for every sequence of positive numbers $\{p_n\}$ such that $\sum p_n < \infty$ we have also $\sum c_n p_n < \infty$. Prove that $\{c_n\}$ is bounded.

72. Verify Schwarz's inequality:

$$\left(\sum_{n=1}^{\infty} |a_n b_n|\right)^2 \leq \sum_{n=1}^{\infty} |a_n|^2 \cdot \sum_{n=1}^{\infty} |b_n|^2$$

(*Hint:* It is true by virtue of the same fact for finite sums, which was discussed in Problem 74 of Chapter 1.)

73. Prove that if $\sum |a_n|^2 < \infty$, then $\sum (1/n)|a_n| < \infty$. Is the reverse implication true?

74. Let S be a subset of R^n. Show that $\perp(S) = \{\mathbf{v} \in R^n\colon \langle \mathbf{v}, \mathbf{s}\rangle = 0$ for all $\mathbf{s} \in S\}$ is a closed set.

75. Suppose that f is a continuous positive real-valued function defined on a set S in R^n. Show that $\log f$ is also continuous.

76. Suppose that f is a continuous real-valued function defined on all of R^n. Let $\mathbf{x}_0, \mathbf{x}_1 \in R^n$ and $c \in R$ be such that $f(\mathbf{x}_0) \leq c \leq f(\mathbf{x}_1)$. Show that there is an $\mathbf{x}_2 \in R^n$ such that $f(\mathbf{x}_2) = c$.

77. Show that if f is a continuous function on the interval I taking only rational values, then f must be constant.

78. A set S in R^n is called *connected* if every continuous real-valued function has the intermediate value property. Show that this is equivalent to the following definition:

A set S is *not connected* if there is a continuous real-valued function f defined on S which takes precisely two values.

79. Verify the following assertions:

(a) A ball in R^n is connected.

(b) The set of integers is not connected.

(c) The sphere $\{\mathbf{x} \in R^3: \|\mathbf{x}\| = 1\}$ is connected.

(d) The union of two balls in R^n is connected if and only if they intersect.

(e) An open set is not connected if and only if it can be written as the disjoint union of two nonempty open subsets.

(f) A closed set is not connected if and only if it can be written as the disjoint union of two nonempty closed sets.

80. Let f be a continuous function on the closed and bounded set X. Then f is *uniformly continuous*; that is, given $\varepsilon > 0$, there is a $\delta > 0$ such that for all $x, y \in X$ such that $|x - y| < \delta$ we have $|f(x) - f(y)| < \varepsilon$. Supposing not, we can derive a contradiction as follows. There is an ε_0 such that for every δ, " $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon_0$ " is not true. Taking $\delta = 1/n$, there are x_n, y_n with $|x_n - y_n| < 1/n$ but $|f(x_n) - f(y_n)| \geq \varepsilon_0$. Since X is closed and bounded, these sequences have convergent subsequences: $\{x'_n\}$, $\{y_n'\}$. Show that $\lim x'_n = \lim y'_n$ but $|f(\lim x'_n) - f(\lim y'_n)| \geq \varepsilon_0$, a contradiction.

81. Let L be a linear functional on R^n and choose v_0 such that $\|v_0\| = 1$ and

$$L(v_0) = \max\{L(v): \|v\| = 1\}$$

Show that for every $v \in R^n$, $L(v) = L(v_0) \langle v, v_0 \rangle$.

82. Let f be an integrable function on the rectangle $[\mathbf{a}, \mathbf{b}]$. Let R_t be rectangle $[\mathbf{a}, \mathbf{b} + t(\mathbf{b} - \mathbf{a})]$, for $0 \leq t \leq 1$. Verify that f is integrable on each rectangle R_t, and define $F(t) = \int_{R_t} f$. Show that f is continuous. Is f differentiable?

83. Let $Q = \{p/q: p, q \text{ integers with } 0 \leq p \leq q\}$. Q is a subset of the unit interval $[0, 1]$ which is not measurable. For surely $\int \chi_Q = 0$, and if $R_1 \cup \cdots \cup R_n \supset Q$, then also $R_1 \cup \cdots \cup R_n \supset [0, 1]$, so $\int \sum \chi_{R_i} \geq 1$, and thus $\overline{\int} \chi_Q = 1$.

84. Let f be an integrable nonnegative function defined on the domain $B \subset R^2$ and consider $D = \{(x, y, z) \in R^3;\ 0 \leq z \leq f(x, y);\ (x, y) \in B\}$. Verify that $\mathrm{Vol}(D) = \int_B f$.

85. Suppose that f is a continuous decreasing real-valued function of a real variable and $\lim_{x\to\infty} f(x) = 0$. Then $\int_0^\infty f(x) \sin x\, dx$ converges (compare this with Leibniz's theorem for series).

86. Suppose that f is a real-valued function defined on R^n. We say that

$$f(\mathbf{x}) \to +\infty \quad \text{as} \quad \|\mathbf{x}\| \to \infty$$

if, for every M there is a K such that $f(\mathbf{x}) \geq M$ whenever $\|\mathbf{x}\| \geq K$. Show that if f is a real-valued continuous function on R^n such that $f(\mathbf{x}) \to +\infty$ as $\|\mathbf{x}\| \to \infty$, then f attains a minimum at some point.

87. Define

$$f(\mathbf{x}) \to 0 \quad \text{as} \quad \|\mathbf{x}\| \to \infty$$

in a way suggested by the definition in the above problem. Show that if a continuous function on R^n has this property, then it attains both a maximum and a minimum on R^n.

88. Suppose f is a real-valued function which has continuous partial derivatives in the ball $\{\mathbf{x} \in R^n\colon \|\mathbf{x}\| < 1\}$. Show that the function

$$g(\mathbf{x}) = \int_0^1 f(t\mathbf{x})\, dt$$

has the same properties, and find ∇g.

89. Let l^2 be the space of sequences $\{c_n\}$ of real numbers such that

$$\sum_{n=1}^{\infty} |c_n|^2 < \infty$$

Because of the result in Problem 72 (Schwarz's inequality), if $\{c_n\}$ and $\{d_n\}$ are in l^2, then

$$\langle\{c_n\}, \{d_n\}\rangle = \sum_{n=1}^{\infty} c_n d_n$$

converges. Show that l^2 is a Euclidean vector space with that inner product.

90. The space of continuous functions on the unit interval can be made into a Euclidean vector space in this way:

$$\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt$$

Corresponding to this inner product is a notion of length which we denote by $\|\cdot\|_2$ so as to distinguish it from the modulus $\|\cdot\|_\infty$ introduced in the text. Show that this length is deficient in these respects:

(a) We can have $\|f_n\|_2 \to 0$ without having $\|f_n\|_\infty \to 0$.

(b) We can have a sequence $\{f_n\}$ of continuous functions which is a Cauchy sequence in the sense of the length $\|\cdot\|_2$, but which does not converge to a continuous function. On the other hand, show that

(c) if $\|f_n\|_\infty \to 0$, then $\|f_n\|_2 \to 0$ also.

91. Suppose L: $C[0, 1] \to R$ is a linear function. Show that L is continuous if and only if there is an $M > 0$ such that

$$|L(f)| \leq M \|f\|_\infty$$

92. Show that there is a unique differentiable function f_0 such that

$$f'_0(x) = (f_0(x))^2 \qquad \text{for all } x \qquad \text{and} \qquad f_0(0) = \tfrac{1}{2}$$

Do it by applying the fixed point theorem to the function T defined below on the set $\{f \in C[\frac{1}{2}, \frac{1}{2}]: \|f\| \leq \frac{3}{4}\}$:

$$Tf(x) = \int_0^x f^2(t)\, dt + \tfrac{1}{2}$$

93. We can talk of open and closed sets, and convergence in the space M^n of $(n \times n)$ matrices, merely by considering them as vectors in R^{n^2}. Doing so verify these statements:

(a) The set G of invertible $(n \times n)$ matrices is open.

(b) The set of triangular matrices is closed.

(c) The function $A \to A^2$ is continuous.

(d) If p is any polynomial in one variable the function

$$T \to p(T)$$

is continuous.

(e) $\lim\limits_{n \to \infty} (x/n!) \sum_{n=0}^{N} (1/n!)T^n$ exists for all $T \in L(R^n, R^m)$.

94. Suppose g is a continuous real-valued function on the interval $[-a, a]$. Show that the implication

$$\int_{-a}^{a} g(t)f(t)\, dt = 0$$

for all $f \in F$ implies $g = 0$ holds whenever F is any one of these classes:

(a) $F = C([-a, a])$.

(b) $F = C^1([-a, a])$.

(c) F is the collection of all polynomials.

(d) F is the collection $\{\chi_I: I \text{ a subinterval of } [-a, a]\}$.

(e) F is the collection of all continuously differentiable functions such that $f(-a) = f(a) = 0$.