(3.2) **Definition.** A sequence (x_n) of real numbers is a *Cauchy sequence* if for each k in $\mathbb{Z}^+$ there exists M_k in $\mathbb{Z}^+$ such that

(3.2.1) $$|x_m - x_n| \leqq k^{-1} \qquad (m, n \geqq M_k).$$

(3.3) **Theorem.** *A sequence (x_n) of real numbers converges if and only if it is a Cauchy sequence.*

Proof: Assume that (x_n) converges to a real number x_0. Let the sequence (N_k) satisfy (3.1.1). Write $M_k \equiv N_{2k}$. Then

$$|x_m - x_n| \leqq |x_m - x_0| + |x_n - x_0| \leqq (2k)^{-1} + (2k)^{-1} = k^{-1}$$

for m, $n \geqq M_k$. Therefore (x_n) is a Cauchy sequence.

Assume conversely that (x_n) is a Cauchy sequence. Let the sequence (M_k) satisfy (3.2.1). Write $N_k \equiv \max\{3k, M_{2k}\}$. Then

$$|x_m - x_n| \leqq (2k)^{-1} \qquad (m, n \geqq N_k).$$

Let y_k be the $(2k)^{\text{th}}$ rational approximation to x_{N_k}. For $m \geq n$,

$$\begin{aligned} |y_m - y_n| &\leq |y_m - x_{N_m}| + |x_{N_m} - x_{N_n}| + |x_{N_n} - y_n| \\ &\leq (2m)^{-1} + (2m)^{-1} + (2n)^{-1} + (2n)^{-1} = m^{-1} + n^{-1}. \end{aligned}$$

Therefore $y \equiv (y_n)$ is a real number. To see that (x_n) converges to y, we consider $n \geqq N_k$ and compute

$$\begin{aligned} |y - x_n| &\leqq |y - y_n| + |y_n - x_{N_n}| + |x_{N_n} - x_n| \\ &\leqq n^{-1} + (2n)^{-1} + (2k)^{-1} \leqq (3k)^{-1} + (6k)^{-1} + (2k)^{-1} = k^{-1}. \quad \square \end{aligned}$$

A subsequence of a convergent sequence converges to the same limit. If a sequence converges, then any sequence obtained from it by modifications (including, perhaps, insertions or deletions) which involve only a finite number of terms converges to the same limit.

If $x \equiv (x_n)$ is a regular sequence of rational numbers, then (x_n^*) converges to x, by (2.14).

A sequence (x_n) is *increasing* (respectively, *strictly increasing*) if $x_{n+1} \geqq x_n$ (respectively, $x_{n+1} > x_n$) for each n. *Decreasing* and *strictly decreasing* sequences are defined analogously, in the obvious way. A theorem of classical mathematics states that every bounded increasing sequence of real numbers converges. A counterexample to this statement is given by any increasing sequence (x_n) such that $x_n = 0$ or $x_n = 1$ for each n, but it is not known whether $x_n = 0$ for all n.

It is useful to supplement Definition (3.1) by writing

$$\lim_{n \to \infty} x_n = \infty$$

or

$$x_n \to \infty \quad \text{as} \quad n \to \infty$$

to express the fact that for each k in $\mathbb{Z}^+$ there exists N_k in $\mathbb{Z}^+$ with $x_n > k$ for all $n \geqq N_k$. We also define

$$\lim_{n\to\infty} x_n = -\infty$$

or

$$x_n \to -\infty \quad \text{as} \quad n \to \infty$$

to mean that $\lim_{n\to\infty} -x_n = \infty$.

The next proposition shows that we may work with real numbers constructed as limits by working with their approximations.

(3.4) **Proposition.** *Assume that $x_n \to x_0$ as $n \to \infty$, and $y_n \to y_0$ as $n \to \infty$, where x_0 and y_0 are real numbers. Then*

(a) $x_n + y_n \to x_0 + y_0$ *as* $n \to \infty$
(b) $x_n y_n \to x_0 y_0$ *as* $n \to \infty$
(c) $\max\{x_n, y_n\} \to \max\{x_0, y_0\}$ *as* $n \to \infty$
(d) $x_0 = c$ *whenever* $x_n = c$ *for all* n
(e) *if* $x_0 \neq 0$ *and* $x_n \neq 0$ *for all* n, *then* $x_n^{-1} \to x_0^{-1}$ *as* $n \to \infty$
(f) *if* $x_n \leqq y_n$ *for all* n, *then* $x_0 \leqq y_0$.

Proof: (a) For each k in $\mathbb{Z}^+$ there exists N_k in $\mathbb{Z}^+$ such that

$$|x_n - x_0| \leqq (2k)^{-1}, \quad |y_n - y_0| \leqq (2k)^{-1} \qquad (n \geqq N_k).$$

Then

$$|x_n + y_n - (x_0 + y_0)| \leqq (2k)^{-1} + (2k)^{-1} = k^{-1} \qquad (n \geqq N_k).$$

Therefore $x_n + y_n \to x_0 + y_0$ as $n \to \infty$.

(b) Choose m in $\mathbb{Z}^+$ such that $|y_0| \leqq m$ and $|x_n| \leqq m$ for all n. For each k in $\mathbb{Z}^+$ choose N_k in $\mathbb{Z}^+$ with

$$|x_n - x_0| \leqq (2mk)^{-1}, \quad |y_n - y_0| \leqq (2mk)^{-1} \quad (n \geqq N_k).$$

Then for $n \geqq N_k$,

$$\begin{aligned} |x_n y_n - x_0 y_0| &\leqq |x_n(y_n - y_0)| + |y_0(x_n - x_0)| \\ &\leqq m(|y_n - y_0| + |x_n - x_0|) \leqq k^{-1}. \end{aligned}$$

Therefore $x_n y_n \to x_0 y_0$ as $n \to \infty$.

(c) Since

$$|\max\{x_n, y_n\} - \max\{x_0, y_0\}| \leqq \max\{|x_n - x_0|, |y_n - y_0|\},$$

it follows that

$$\max\{x_n, y_n\} \to \max\{x_0, y_0\} \quad \text{as } n \to \infty.$$

(d) If $x_n = c$ for all n, then (x_n) converges to c. Therefore $x_0 = c$.

(e) Since $|x_0| > 0$,

$$|x_n| \geqq |x_0| - |x_n - x_0| > \tfrac{1}{2}|x_0|$$

whenever n is large enough, say for $n \geqq n_0$. Let k and n be positive integers such that $n \geqq n_0$ and $|x_n - x_0| < (2k)^{-1}|x_0|^2$. Then

$$|x_n^{-1} - x_0^{-1}| = |x_n|^{-1}|x_0|^{-1}|x_n - x_0| \leqq 2|x_0|^{-2}(2k)^{-1}|x_0|^2 = k^{-1}.$$

Therefore $x_n^{-1} \to x_0^{-1}$ as $n \to \infty$.

(f) We compute

$$\begin{aligned} y_0 - x_0 &= \lim_{n\to\infty} y_n - \lim_{n\to\infty} x_n = \lim_{n\to\infty}(y_n - x_n) = \lim_{n\to\infty}|y_n - x_n| \\ &= \lim_{n\to\infty} \max\{y_n - x_n, x_n - y_n\} = \max\{y_0 - x_0, x_0 - y_0\} \geqq 0, \end{aligned}$$

by (a), (b), (c), and (d). □

For each sequence (x_n) of real numbers the number

$$s_n \equiv \sum_{k=1}^{n} x_k$$

is called the n^{th} *partial sum* of (x_n), and (s_n) is called the *sequence of partial sums* of the sequence (x_n). A *sum* s_0 of (x_n) is a limit of the sequence (s_n) of partial sums. We write

$$s_0 = \sum_{n=1}^{\infty} x_n$$

to indicate that s_0 is a sum of (x_n). A sequence which is meant to be summed is called a *series*. A series is said to *converge* to its sum. Thus the sequence $(2^{-n})_{n=1}^{\infty}$ converges to 0 as a sequence, but as a series it converges to $\sum_{n=1}^{\infty} 2^{-n} = 1$.

A convergent series remains convergent, but not necessarily to the same sum, after modification of finitely many of its terms.

The series (x_n) is often loosely referred to as *the series* $\sum_{n=1}^{\infty} x_n$.

If the series $\sum_{n=1}^{\infty} x_n$ converges, then $x_n \to 0$ as $n \to \infty$.

A series $\sum_{n=1}^{\infty} x_n$ is said to *converge absolutely* when the series $\sum_{n=1}^{\infty} |x_n|$ converges.

In classical analysis a series of nonnegative terms converges if the partial sums are bounded. This is not true in constructive analysis. However, we have the following result.

(3.5) **Proposition.** *If* $\sum_{n=1}^{\infty} y_n$ *is a convergent series of nonnegative terms, and if* $|x_n| \leqq y_n$ *for each n, then* $\sum_{n=1}^{\infty} x_n$ *converges.*

Proof: Since $\sum_{n=1}^{\infty} y_n$ is convergent, the sequence of partial sums is a Cauchy sequence. Therefore for each k in $\mathbb{Z}^+$ there exists an N_k in $\mathbb{Z}^+$ with

$$\sum_{j=n+1}^{m} y_j \leqq k^{-1} \qquad (m > n \geqq N_k).$$

Then

$$\left|\sum_{j=n+1}^{m} x_j\right| \leqq \sum_{j=n+1}^{m} y_j \leqq k^{-1} \qquad (m > n \geqq N_k).$$

Therefore the sequence of partial sums of the series $\sum_{n=1}^{\infty} x_n$ is a Cauchy sequence. By (3.3), the series converges. □

The criterion of Proposition (3.5) is known as the *comparison test.* It follows from the comparison test that every absolutely convergent series is convergent.

The terms of an absolutely convergent series $\sum_{n=1}^{\infty} x_n$ may be reordered without affecting the sum s_0 of the series. More precisely, if λ: $\mathbb{Z}^+ \to \mathbb{Z}^+$ is a bijection, then $\sum_{n=1}^{\infty} x_{\lambda(n)}$ exists and equals s_0. This may not be true if the series $\sum_{n=1}^{\infty} x_n$ is merely convergent.

A sequence (x_n) is said to *diverge* if there exists ε in $\mathbb{R}^+$ such that for each k in $\mathbb{Z}^+$ there exist m and n in $\mathbb{Z}^+$ with $m, n \geqq k$ and $|x_m - x_n| \geqq \varepsilon$. The motivation for this definition is, of course, that a sequence cannot be both convergent and divergent. A series is said to *diverge* if the sequence of its partial sums diverges.

The series $\sum_{n=1}^{\infty} n^{-1}$ diverges, because

$$\left|\sum_{n=1}^{2^{k+1}} n^{-1} - \sum_{n=1}^{2^k} n^{-1}\right| > \tfrac{1}{2} \qquad (k \in \mathbb{Z}^+).$$

The series $\sum_{n=1}^{\infty} x_n$ diverges whenever there exists r in $\mathbb{R}^+$ such that $|x_n| \geqq r$ for infinitely many values of n.

Let $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ be series of nonnegative terms. The *comparison test* for divergence is that $\sum_{n=1}^{\infty} x_n$ diverges whenever $\sum_{n=1}^{\infty} y_n$ diverges and there is a positive integer N with $x_n \geqq y_n$ for all $n \geqq N$.

The following very useful test for convergence and divergence is called the *ratio test.*

(3.6) **Proposition.** *Let* $\sum_{n=1}^{\infty} x_n$ *be a series, c a positive number, and N a positive integer. Then* $\sum_{n=1}^{\infty} x_n$ *converges if* $c<1$ *and*

(3.6.1) $$|x_{n+1}| \leqq c|x_n| \quad (n \geqq N),$$

and diverges if $c>1$ *and*

(3.6.2) $$|x_{n+1}| > c|x_n| \quad (n \geqq N).$$

Proof: Assume that $c<1$ and that (3.6.1) is valid. Then $|x_n| \leq c^{n-N}|x_N|$ for $n \geqq N$. By the comparison test, $\sum_{n=1}^{\infty} x_n$ converges.

Next, assume that $c>1$ and that (3.6.2) holds. Then

$$|x_n| \geqq c^{n-N-1}|x_{N+1}| \geqq |x_{N+1}| \quad (n \geqq N+1)$$

and

$$|x_{N+1}| > c|x_N| \geqq 0.$$

Therefore $\sum_{n=1}^{\infty} x_n$ diverges. □

A corollary of the ratio test is that if the limit

$$L \equiv \lim_{n \to \infty} |x_{n+1} x_n^{-1}|$$

exists, then $\sum_{n=1}^{\infty} x_n$ converges whenever $L<1$ and diverges whenever $L>1$.

The ratio test says nothing in case $L=1$. To handle this case, we introduce stronger tests based on *Kummer's criterion.*

(3.7) **Lemma.** *Let* (a_n) *and* (x_n) *be sequences of positive numbers, c a positive number, and N a positive integer. Then* $\sum_{n=1}^{\infty} x_n$ *converges if* $a_n x_n \to 0$ *as* $n \to \infty$ *and*

(3.7.1) $$a_n x_n x_{n+1}^{-1} - a_{n+1} \geqq c \quad (n \geqq N),$$

while $\sum_{n=1}^{\infty} x_n$ *diverges if* $\sum_{n=1}^{\infty} a_n^{-1}$ *diverges and*

(3.7.2) $$a_n x_n x_{n+1}^{-1} - a_{n+1} \leqq 0 \quad (n \geqq N).$$

Proof: Assume that $a_n x_n \to 0$ and that (3.7.1) is valid. Let ε be an arbitrary positive number, and choose an integer $\nu \geqq N$ so that $a_k x_k - a_j x_j \leqq c\,\varepsilon$ whenever $j > k \geqq \nu$. For such j and k we have

$$\sum_{n=k+1}^{j} x_n \leqq c^{-1} \sum_{n=k+1}^{j} x_n(a_{n-1} x_{n-1} x_n^{-1} - a_n)$$
$$= c^{-1}(a_k x_k - a_j x_j) \leqq \varepsilon.$$

Thus $\left(\sum_{n=1}^{j} x_n\right)_{j=1}^{\infty}$ is a Cauchy sequence, and so $\sum_{n=1}^{\infty} x_n$ converges.

Next assume that $\sum_{n=1}^{\infty} a_n^{-1}$ diverges and that (3.7.2) holds. Then for each $n \geqq N$, $x_n \geqq a_N x_N a_n^{-1}$. Thus $\sum_{n=1}^{\infty} x_n$ diverges, by comparison with $\sum_{n=1}^{\infty} a_n^{-1}$. □

(3.8) **Lemma.** *Let (y_n) be a sequence of positive numbers, c a positive number, and N a positive integer such that*

$$n(y_n y_{n+1}^{-1} - 1) \geqq c \qquad (n \geqq N).$$

Then $\lim_{n\to\infty} y_n = 0$.

Proof: For each $n > N$,

$$y_N y_n^{-1} = (y_N y_{N+1}^{-1})(y_{N+1} y_{N+2}^{-1}) \cdots (y_{n-1} y_n^{-1})$$
$$\geqq (1 + cN^{-1}) \ldots (1 + c(n-1)^{-1})$$
$$\geqq 1 + c \sum_{k=N}^{n-1} k^{-1}.$$

Given $\varepsilon > 0$, choose an integer $\nu > N$ so that $\sum_{k=N}^{n-1} k^{-1} > c^{-1}(\varepsilon^{-1} y_N - 1)$ for all $n \geqq \nu$. Then for such n we have $y_n < \varepsilon$. Hence $y_n \to 0$ as $n \to \infty$. □

The next convergence test is known as *Raabe's test.*

(3.9) **Proposition.** *Let $\sum_{n=1}^{\infty} x_n$ be a series of positive numbers such that $n(x_n x_{n+1}^{-1} - 1)$ converges to a limit L. Then $\sum_{n=1}^{\infty} x_n$ converges if $L > 1$, and diverges if $L < 1$.*

Proof: First note that

$$n(n x_n/(n+1) x_{n+1} - 1) = n(n+1)^{-1}(n(x_n x_{n+1}^{-1} - 1) - 1)$$
$$\to L - 1 \qquad \text{as } n \to \infty.$$

If $L>1$, it follows from (3.8) that $nx_n \to 0$ as $n \to \infty$. We then obtain the convergence of $\sum_{n=1}^{\infty} x_n$ by taking $a_n \equiv n$ $(n \in \mathbb{Z}^+)$ in Kummer's criterion. The same choice of a_n yields divergence of $\sum_{n=1}^{\infty} x_n$ in case $L<1$. □

Important real numbers represented by series are

$$e = 1 + \sum_{n=1}^{\infty} (n!)^{-1}$$

and

$$\pi = 4 \sum_{n=0}^{\infty} (-1)^n (2n+1)^{-1}.$$

The series for e converges by the ratio test. The convergence of the series for π is a consequence of the general result that a series $\sum_{n=1}^{\infty} (-1)^n x_n$ converges whenever (i) $x_n \geqq 0$ for all n and (ii) the sequence (x_n) is decreasing and converges to 0. To see this, consider positive integers m and n with $m \geqq n$. Then

$$\begin{aligned} 0 &\leqq (x_n - x_{n+1}) + (x_{n+2} - x_{n+3}) + \ldots + (-1)^{m+n} x_m \\ &= (-1)^n \sum_{k=n}^{m} (-1)^k x_k \\ &= x_n - (x_{n+1} - x_{n+2}) - \ldots + (-1)^{m+n} x_m \leqq x_n. \end{aligned}$$

It follows that the sequence of partial sums of the series is a Cauchy sequence. Therefore the series converges.

4. Continuous Functions

A property P which is applicable to the elements of a set S is defined by a statement of the requirements that an element of S must satisfy in order to have property P. To construct an element of S with property P we must construct an element of S, perform certain additional constructions which depend on the property P, and prove that the entities constructed satisfy certain requirements that are characteristic of the property P. Each property P applicable to elements of a set S determines a subset of S which is denoted by

$$\{x\colon\ x \in S,\ x \text{ has property } P\}$$

or

$$\{x \in S\colon\ x \text{ has property } P\}.$$

When the context makes it clear which set S is under discussion, we also write simply

$$\{x\colon\ x \text{ has property } P\}.$$

Properties applicable to elements of a set S, and subsets of S, are essentially the same things regarded from different points of view.

Among the most important subsets of $\mathbb{R}$ are the intervals.

(4.1) **Definition.** For all real numbers a and b we define

$$(a,b) \equiv \{x\colon\ x \in \mathbb{R},\ a < x < b\},$$
$$(a,b] \equiv \{x\colon\ x \in \mathbb{R},\ a < x \leqq b\},$$
$$[a,b) \equiv \{x\colon\ x \in \mathbb{R},\ a \leqq x < b\},$$
$$[a,b] \equiv \{x\colon\ x \in \mathbb{R},\ a \leqq x \leqq b\}.$$

Each of these sets in an *interval*, whose left and right *end points* are a and b, respectively. The interval (a, b) is *open*, $[a, b]$ is *closed*, $(a, b]$ is *half-open on the left*, and $[a, b)$ is *half-open* on the right. If $a<b$, then the intervals are said to be *proper*.

The above intervals are called *finite intervals*. We also introduce *infinite intervals*, by the following definitions:

$$(-\infty, a) \equiv \{x\colon\ x \in \mathbb{R},\ x < a\},$$
$$(-\infty, a] \equiv \{x\colon\ x \in \mathbb{R},\ x \leqq a\},$$
$$(a, \infty) \equiv \{x\colon\ x \in \mathbb{R},\ a < x\},$$
$$[a, \infty) \equiv \{x\colon\ x \in \mathbb{R},\ a \leqq x\}.$$

An interval I is *nonvoid* if we can construct a real number belonging to I. A nonvoid, closed, finite interval is called a *compact interval.* A nonvoid finite interval I with left and right end points a and b has *length* $|I| \equiv b-a$.

If an interval I is a subset of an interval J (that is, if every element of I also belongs to J), then we say that I is a *subinterval* of J.

The rules for manipulating intervals, which we use in the sequel without mention or proof, are implicit in Proposition (2.11).

(4.2) **Definition.** A nonvoid set A of real numbers is *bounded above* if there exists a real number b, called an *upper bound* of A, such that $x \leqq b$ for all x in A. A real number b is called a *supremum*, or *least upper bound*, of A if it is an upper bound of A, and if for each $\varepsilon > 0$ there exists x in A with $x > b - \varepsilon$.

We say that A is *bounded below* if there exists a real number b, called a *lower bound* of A, such that $b \leqq x$ for all x in A. A real number b is called an *infimum*, or *greatest lower bound*, of A if it is a lower bound of A, and if for each $\varepsilon>0$ there exists x in A with $x<b+\varepsilon$.

The supremum (respectively, infimum) of A is unique, if it exists, and is written $\sup A$ (respectively, $\inf A$).

A classical theorem asserts that every nonvoid set of real numbers that is bounded above has a supremum. A counterexample to this is provided by the set $\{x_n: n\in\mathbb{Z}^+\}$ where $x_n=0$ or $x_n=1$ for each n, but it is not known whether $x_n=0$ for all n.

We now prove the constructive *least-upper-bound principle.*

(4.3) **Proposition.** *Let A be a nonvoid set of real numbers that is bounded above. Then* $\sup A$ *exists if and only if for all x, y in* $\mathbb{R}$ *with $x<y$, either y is an upper bound of A or there exists a in A with $x<a$.*

Proof: If $\sup A$ exists and $x<y$, then either $\sup A<y$ or $x<\sup A$; in the latter case we can find a in A with

$$\sup A-(\sup A-x)<a$$

and hence $x<a$. Thus the stated condition is necessary.

Conversely, assume that the stated condition holds. Let a_1 be an element of A, and choose an upper bound b_1 of A with $b_1>a_1$. We construct recursively a sequence (a_n) in A and a sequence (b_n) of upper bounds of A such that for each n in $\mathbb{Z}^+$,

(i) $a_n\leqq a_{n+1}<b_{n+1}\leqq b_n$

and

(ii) $b_{n+1}-a_{n+1}\leqq\frac{3}{4}(b_n-a_n)$.

Having found $a_1,\ldots,a_n$ and $b_1,\ldots,b_n$, if $a_n+\frac{3}{4}(b_n-a_n)$ is an upper bound of A, we set $b_{n+1}\equiv a_n+\frac{3}{4}(b_n-a_n)$ and $a_{n+1}\equiv a_n$; while if there exists a in A with $a>a_n+\frac{1}{4}(b_n-a_n)$, we set $a_{n+1}\equiv a$ and $b_{n+1}\equiv b_n$. This completes the recursive construction.

By (i) and (ii), we have

$$0\leqq b_n-a_n\leqq(3/4)^{n-1}(b_1-a_1)\qquad(n\in\mathbb{Z}^+).$$

Hence the sequences (a_n) and (b_n) converge to a common limit ℓ with $a_n\leqq\ell\leqq b_n$ for each n in $\mathbb{Z}^+$. Since each b_n is an upper bound of A, so is ℓ. On the other hand, given $\varepsilon>0$, we can choose n so that $\ell\geqq a_n>\ell-\varepsilon$, where $a_n\in A$. Hence $\ell=\sup A$. □

In Proposition (4.3), if A is contained in some interval I, then in order to prove that $\sup A$ exists it is sufficient to consider arbitrary points x and y in I with $x<y$.

(4.4) **Corollary.** *Let the subset A of $\mathbb{R}$ have the property that for each $\varepsilon>0$ there exists a subfinite set $\{y_1, \dots, y_n\}$ of points of A such that for each x in A at least one of the numbers $|x-y_1|, \dots, |x-y_n|$ is less than ε. (Such a set A is called* totally bounded.) *Then* $\sup A$ *and* $\inf A$ *exist.*

Proof: It will suffice to prove that $\sup A$ exists. To this end, let x and y be real numbers with $x<y$, and set $\alpha \equiv \frac{1}{4}(y-x)$. Choose points $a_1, \dots, a_N$ in A such that for each a in A at least one of the numbers $|a-a_1|, \dots, |a-a_N|$ is less than α. For some n with $1 \leq n \leq N$ we have

$$a_n > \max\{a_1, \dots, a_N\} - \alpha.$$

Either $x<a_n$ or $a_n<x+2\alpha$. In the latter case, if $a \in A$ and we choose k with $|a-a_k|<\alpha$, we have

$$a \leqq a_k + |a-a_k| < a_n + \alpha + \alpha < x + 4\alpha = y,$$

so that y is an upper bound of A. Thus $\sup A$ exists, by (4.3). □

Often when one real number depends on another the dependence is smooth, or continuous. An exact description of what this means is given in the following definition.

(4.5) **Definition.** A real-valued function f defined on a compact interval I is *continuous* on I if for each $\varepsilon>0$ there exists $\omega(\varepsilon)>0$ such that $|f(x)-f(y)| \leqq \varepsilon$ whenever $x, y \in I$ and $|x-y| \leqq \omega(\varepsilon)$. The operation $\varepsilon \mapsto \omega(\varepsilon)$ is called a *modulus of continuity* for f.

A real-valued function f on an arbitrary interval J is *continuous* on J if it is continuous on every compact subinterval I of J.

For example, when a and b are real numbers with $a<b$, then f is continuous on (a, b) if and only if it is continuous on $[a+\delta, b-\delta]$ for each δ with $0<\delta<\frac{1}{2}(b-a)$.

A modulus of continuity ω is an indispensable part of the definition of a continuous function on a compact interval, although sometimes it is not mentioned explicitly. In the same way, moduli of continuity of the restrictions of f to each compact subinterval are indispensable parts of the definition of a continuous function f on a general interval.

Constant functions, and the identity function $x \mapsto x$, are continuous on every interval.

(4.6) **Proposition.** *If f: $[a, b] \to \mathbb{R}$ is a continuous function on a compact interval, then the quantities*

$$\sup f \equiv \sup\{f(x) \colon x \in [a, b]\}$$

and

$$\inf f \equiv \inf\{f(x):\ x \in [a, b]\}$$

(called, respectively, the supremum *and the* infimum *of f on the interval* $[a, b]$*) exist.*

Proof: Consider any $\varepsilon > 0$. Choose real numbers $a = a_0 \leqq a_1 \leqq \ldots \leqq a_n = b$ such that $a_{i+1} - a_i \leqq \omega(\varepsilon)$ $(0 \leqq i \leqq n-1)$, where ω is a modulus of continuity for f. Then for each x in $[a, b]$ we have $|x - a_i| \leqq \omega(\varepsilon)$, and therefore $|f(x) - f(a_i)| \leqq \varepsilon$, for some i. Since ε is arbitrary, it follows that the set $\{f(x):\ x \in [a, b]\}$ is totally bounded. Therefore $\sup f$ and $\inf f$ exist, by (4.4). □

Let f: $A \to \mathbb{R}$ and g: $A \to \mathbb{R}$ be two functions defined on the same set A. We define the sum function $f + g$: $A \to \mathbb{R}$ by

$$(f+g)(x) \equiv f(x) + g(x) \qquad (x \in A).$$

Such functions as fg, $|f|$, and $\max\{f, g\}$ are defined similarly. If $g(x) \neq 0$ for each x in A, then g^{-1}: $A \to \mathbb{R}$ is defined by

$$g^{-1}(x) \equiv (g(x))^{-1} \qquad (x \in A).$$

We also write $1/g$ for g^{-1}, and f/g for the quotient function fg^{-1}.

The proof of the following proposition, which resembles the proof of Proposition (3.4), is left to the reader.

(4.7) **Proposition.** *Let f and g be continuous real-valued functions defined on an interval I. Then the functions $f+g$, fg, and* $\max\{f, g\}$ *are continuous on I. If f is bounded away from 0 on every compact subinterval J of I – that is, if $|f(x)| \geqq c$ for all x in J and some $c > 0$ (depending on J) – then f^{-1} is continuous on I.*

Proposition (4.7) implies that the quotient of continuous functions is continuous, provided that the denominator is bounded away from 0 on every compact subinterval. It also implies that a polynomial function

$$x \mapsto c_0 x^n + c_1 x^{n-1} + \ldots + c_n$$

is continuous on every interval, and that $|f|$ is continuous on each interval where f is continuous.

The composition of continuous functions is continuous, in the sense that if f: $I \to J$ and g: $J \to \mathbb{R}$ are continuous, then $g \circ f$ is continuous, provided that f maps every compact subinterval of I into a compact subinterval of J. To prove this, it is sufficient to consider the case in which I and J are both compact. Let ω be a modulus of continuity for f, and σ a modulus of continuity for g. Then if $x, y \in I$, $\varepsilon > 0$, and $|x-y| \leqq \omega(\sigma(\varepsilon))$, we have $|f(x)-f(y)| \leqq \sigma(\varepsilon)$. Therefore

$$|g(f(x)) - g(f(y))| \leqq \varepsilon.$$

It follows that $g \circ f$ is continuous, with modulus of continuity $\varepsilon \mapsto \omega(\sigma(\varepsilon))$.

Classically, a continuous function maps an interval onto an interval. We now prove a weak version of this result, known as the *intermediate value theorem.*

(4.8) **Theorem.** *Let f be a continuous map defined on an interval I, and let a, b be points of I with $f(a) < f(b)$. Then for each y in $[f(a), f(b)]$ and each $\varepsilon > 0$, there exists x in $[\min\{a, b\}, \max\{a, b\}]$ such that $|f(x) - y| < \varepsilon$.*

Proof: Since f is continuous, we must have $a \neq b$. We may assume that $a < b$. Consider y in $[f(a), f(b)]$ and $\varepsilon > 0$. Let

$$m \equiv \inf\{|f(x) - y| : a \leqq x \leqq b\},$$

which exists by (4.6). Suppose that $m > 0$. Then $f(a) - y \leqq -m$ and $f(b) - y \geqq m$. Let ω be a modulus of continuity for f on $[a, b]$, and choose points $a = x_0 \leqq x_1 \leqq \ldots \leqq x_n = b$ such that $x_{k+1} - x_k \leqq \omega(m)$ for $0 \leqq k \leqq n - 1$. Then for such k we have

$$|f(x_{k+1}) - y - (f(x_k) - y)| = |f(x_{k+1}) - f(x_k)| \leqq m.$$

Since $|f(x) - y| \geqq m$ for all x in $[a, b]$, it follows that the quantities $f(x_k) - y$ and $f(x_{k+1}) - y$ are either both positive or both negative. Therefore the quantities $f(x_i) - y$ $(0 \leqq i \leqq n)$ are either all positive or all negative. Hence $f(a) - y$ and $f(b) - y$ are either both positive or both negative. This contradiction ensures that the possibility $m > 0$ is ruled out; so that $m < \varepsilon$, and the desired conclusion follows. □

Under additional hypotheses satisfied by many of the common elementary functions of analysis, Theorem (4.8) can be strengthened to yield the conclusion that $f(x) = y$ for some x in $[\min\{a, b\}, \max\{a, b\}]$. For example, this strong conclusion obtains whenever f is *strictly increasing*, in the sense that $f(x) > f(x')$ for any two points x, x' of its domain with $x > x'$. For in that case, taking $a < b$ we can construct sequences (a_n), (b_n) in $[a, b]$ such that for each n in $\mathbb{Z}^+$,

(i) $a = a_1 \leqq a_2 \leqq \ldots \leqq a_n \leqq b_n \leqq \ldots \leqq b_2 \leqq b_1 = b$

(ii) $f(a_n) \leqq y \leqq f(b_n)$

(iii) $b_{n+1} - a_{n+1} \leqq (2/3)(b_n - a_n)$.

The sequences (a_n), (b_n) then converge to a common limit x in $[a, b]$ with $f(x) = y$.

Just as sequences of real numbers can converge to real numbers, sequences of continuous functions can converge to continuous functions. In fact, most of the important functions of analysis are defined as limits of sequences of continuous functions.

(4.9) **Definition.** A sequence (f_n) of continuous functions on a compact interval I *converges* on I to a continuous function f if for each $\varepsilon > 0$ there exists N_ε in $\mathbb{Z}^+$ such that

(4.9.1) $$|f_n(x) - f(x)| \leqq \varepsilon \quad (x \in I,\ n \geqq N_\varepsilon).$$

A sequence (f_n) of continuous functions on an arbitrary interval J *converges* on J to a continuous function f if it converges to f on every compact subinterval I of J; in that case, f is called the *limit* of the sequence (f_n).

Definition (4.9) can be recast to bear a closer resemblance to Definition (3.1). To this end, we define the *norm* $\|f\|_I$ of a continuous function f on a compact interval I to be the supremum of $|f|$ on I. Then (f_n) converges to f on I if and only if for each k in $\mathbb{Z}^+$ there exists N_k in $\mathbb{Z}^+$ with

$$\|f_n - f\|_I \leqq k^{-1} \quad (n \geqq N_k).$$

(4.10) **Definition.** A sequence (f_n) of continuous functions on a compact interval I is a *Cauchy sequence* on I if for each $\varepsilon > 0$ there exists M_ε in $\mathbb{Z}^+$ such that

(4.10.1) $$|f_m(x) - f_n(x)| \leqq \varepsilon \quad (x \in I;\ m, n \geqq M_\varepsilon).$$

A sequence of continuous functions on an arbitrary interval J is a *Cauchy sequence* on J if it is a Cauchy sequence on every compact subinterval of J.

The sequence (f_n) is a Cauchy sequence on the compact interval I if and only if for every k in $\mathbb{Z}^+$ there exists M_k in $\mathbb{Z}^+$ such that

$$\|f_m - f_n\|_I \leqq k^{-1} \quad (n, n \geqq M_k).$$

Notice that a sequence (c_n) of real numbers converges if and only if the corresponding sequence of constant functions, which we also denote by (c_n), converges on a given nonvoid interval I, and that a sequence of real numbers is a Cauchy sequence if and only if the corresponding sequence of constant functions is a Cauchy sequence on I. Because of these remarks, the following theorem is a generalization of Theorem (3.3).

(4.11) **Theorem.** *A sequence (f_n) of continuous functions on an interval J converges on J if and only if it is a Cauchy sequence on J.*

Proof: Assume that (f_n) converges to f on J. Let I be any compact subinterval of J. For each $\varepsilon>0$ choose N_ε in $\mathbb{Z}^+$ satisfying (4.9.1), and write $M_\varepsilon \equiv N_{\varepsilon/2}$. Then whenever $m, n \geqq M_\varepsilon$ and $x \in I$, we have

$$|f_m(x)-f_n(x)| \leqq |f_m(x)-f(x)|+|f_n(x)-f(x)|$$
$$\leqq \varepsilon/2+\varepsilon/2=\varepsilon.$$

Therefore (f_n) is a Cauchy sequence on I. It follows that (f_n) is a Cauchy sequence on J.

Assume conversely that (f_n) is a Cauchy sequence on J. Then for each x in J, $(f_n(x))$ is a Cauchy sequence of real numbers, whose limit we denote by $f(x)$. We shall show that $f\colon J \to \mathbb{R}$ is a continuous function and that (f_n) converges to f on J. It is enough to show that f is continuous on each compact subinterval I of J, and that (f_n) converges to f on I. To this end, choose the positive integers M_ε such that (4.10.1) is valid, and for each n in $\mathbb{Z}^+$ let ω_n be a modulus of continuity for f_n on I. For each $\varepsilon>0$ write

$$\omega(\varepsilon) \equiv \omega_M(\varepsilon/3),$$

where $M \equiv M_{\varepsilon/3}$. Then whenever $x, y \in I$ and $|x-y| \leqq \omega(\varepsilon)$, we have

$$|f(x)-f(y)| \leqq |f(x)-f_M(x)|+|f_M(x)-f_M(y)|+|f_M(y)-f(y)|$$
$$=\lim_{n\to\infty}|f_n(x)-f_M(x)|+|f_M(x)-f_M(y)|+\lim_{n\to\infty}|f_M(y)-f_n(y)|$$
$$\leqq \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.$$

Therefore f is continuous, with modulus of continuity ω. Finally, if $x \in I$, $\varepsilon>0$, and $n \geqq M_\varepsilon$, then

$$|f_n(x)-f(x)|=\lim_{m\to\infty}|f_n(x)-f_m(x)| \leqq \varepsilon.$$

Hence (f_n) converges to f on I. □

Notations to express the fact that (f_n) converges to f are

$$\lim_{n\to\infty} f_n = f$$

and

$$f_n \to f \quad \text{as } n \to \infty.$$

We also write simply $f_n \to f$.

To each sequence (f_n) of continuous functions on an interval I corresponds a sequence (g_n) of *partial sums*, defined by

$$g_n \equiv \sum_{k=1}^{n} f_k.$$

If (g_n) converges to a continuous function g on I, then g is the *sum* of the *series* $\sum_{n=1}^{\infty} f_n$,

$$g \equiv \sum_{n=1}^{\infty} f_n,$$

and the series is said to *converge* to g on I. If $\sum_{n=1}^{\infty} |f_n|$ converges on I, then $\sum_{n=1}^{\infty} f_n$ is said to *converge absolutely* on I. An absolutely convergent series of functions converges.

The comparison test and the ratio test for convergence carry over to series of functions. The *comparison test* states that if $\sum_{n=1}^{\infty} g_n$ is a convergent series of nonnegative continuous functions on an interval I, then the series $\sum_{n=1}^{\infty} f_n$ of continuous functions on I converges on I whenever $|f_n(x)| \leqq g_n(x)$ for all n in $\mathbb{Z}^+$ and all x in I.

The *ratio test* states that if $\sum_{n=1}^{\infty} f_n$ is a series of continuous functions on an interval J such that for each compact subinterval I of J there exist a constant c_I, $0 < c_I < 1$, and a positive integer N_I with

$$|f_{n+1}(x)| \leqq c_I |f_n(x)| \qquad (x \in I,\ n \geqq N_I),$$

then $\sum_{n=1}^{\infty} f_n$ converges absolutely on J.

A *power series* is a series of the form $\sum_{n=0}^{\infty} a_n(x-x_0)^n$, where $a_n(x-x_0)^n$ represents the function $x \mapsto a_n(x-x_0)^n$ and where $a_0(x-x_0)^0 \equiv a_0$ for all x. The ratio test has the following corollary.

(4.12) **Proposition.** *Let the power series* $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ *have the property that there exist* $r > 0$ *and* N *in* $\mathbb{Z}^+$ *such that* $|a_{n+1}| \leqq r^{-1}|a_n|$ *for all* $n \geqq N$. *Then the series converges absolutely on the interval* $(x_0 - r, x_0 + r)$.

Proof: If I is a compact subinterval of $(x_0 - r, x_0 + r)$, then there exists r_0 with $0 < r_0 < r$ such that $|x - x_0| \leqq r_0$ for all x in I. Then

$$|a_{n+1}(x-x_0)^{n+1}| \leqq r^{-1} r_0 |a_n(x-x_0)^n| \qquad (n \geqq N,\ x \in I).$$

By the ratio test, the series therefore converges absolutely on I. □

5. Differentiation

The rate at which a function is changing is a fundamental property of the function. Here is the precise definition of this concept.

(5.1) **Definition.** Let f and g be continuous functions on a proper compact interval I, and let δ be an operation from $\mathbb{R}^+$ to $\mathbb{R}^+$ such that

$$|f(y)-f(x)-g(x)(y-x)| \leqq \varepsilon |y-x|$$

whenever $\varepsilon>0$, $x, y \in I$, and $|y-x| \leqq \delta(\varepsilon)$. Then f is said to be *differentiable* on I, g is called a *derivative* of f on I, and δ is called a *modulus of differentiability* for f on I.

If f and g are continuous functions on a proper interval J, then g is a *derivative* of f on J if it is a derivative of f on every proper compact subinterval of J; f is then said to be *differentiable* on J.

To express that g is a derivative of f we write

$$g=f', \quad \text{or} \quad g=Df, \quad \text{or} \quad g(x)=\frac{df(x)}{dx}.$$

One way to interpret Definition (5.1) is that the *difference quotient*

$$(f(y)-f(x))(y-x)^{-1}$$

approaches $g(x)$ as y approaches x. In other words, g is the rate of change of f.

If f has two derivatives on I, then they are equal functions.

(5.2) **Theorem.** *Let f_1 and f_2 be differentiable functions on an interval I. Then f_1+f_2 and $f_1 f_2$ are differentiable on I. In case f_1 is bounded away from 0 on every compact subinterval of I, then f_1^{-1} is differentiable on I. The function $x \mapsto x$ is differentiable on $\mathbb{R}$. For each c in $\mathbb{R}$ the function $x \mapsto c$ is differentiable on $\mathbb{R}$. The derivatives in question are given by the following relations:*

(a) $D(f_1+f_2)=Df_1+Df_2$

(b) $D(f_1 f_2)=f_1 Df_2+f_2 Df_1$

(c) $Df_1^{-1}=-f_1^{-2} Df_1$

(d) $\dfrac{dx}{dx}=1$

(e) $\dfrac{dc}{dx}=0.$

Proof: It is enough to consider the case in which I is compact. Let δ_1 and δ_2 be moduli of differentiability for f_1 and f_2, respectively, on I, and ω_1 a modulus of continuity for f_1 on I.

(a) Whenever $x, y \in I$ and $|y-x| \leqq \delta(\varepsilon) \equiv \min\{\delta_1(\varepsilon/2),\ \delta_2(\varepsilon/2)\}$, we have

$$\begin{aligned}
&|f_1(y)+f_2(y)-(f_1(x)+f_2(x))-(f_1'(x)+f_2'(x))(y-x)| \\
&\quad \leqq |f_1(y)-f_1(x)-f_1'(x)(y-x)|+|f_2(y)-f_2(x)-f_2'(x)(y-x)| \\
&\quad \leqq \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.
\end{aligned}$$

Thus f_1+f_2 is differentable on I, with derivative $f_1'+f_2'$ and modulus of differentiability δ.

(b) Let M be a common bound for $|f_1|$, $|f_2|$, and $|f_2'|$ on the interval I. (For instance, define $M \equiv \max\{\|f_1\|_I, \|f_2\|_I, \|f_2'\|_I\}$.) Then whenever $x, y \in I$ and

$$|y-x| \leqq \delta(\varepsilon) \equiv \min\{\delta_1((3M)^{-1}\varepsilon), \delta_2((3M)^{-1}\varepsilon), \omega_1((3M)^{-1}\varepsilon)\},$$

we have

$$\begin{aligned}
&|f_1(y)\,f_2(y)-f_1(x)\,f_2(x)-(f_1(x)\,f_2'(x)+f_2(x)\,f_1'(x))(y-x)| \\
&\quad \leqq |f_1(y)|\,|f_2(y)-f_2(x)-f_2'(x)(y-x)| \\
&\qquad +|f_1(y)-f_1(x)|\,|f_2'(x)|\,|y-x| \\
&\qquad +|f_2(x)|\,|f_1(y)-f_1(x)-f_1'(x)(y-x)| \\
&\quad \leqq 3M(3M)^{-1}\varepsilon|y-x|=\varepsilon|y-x|.
\end{aligned}$$

Therefore $f_1 f_2$ is differentiable on I, with derivative $f_1 f_2'+f_2 f_1'$ and modulus of differentiability δ.

(c) For each $\varepsilon>0$ write

$$\delta(\varepsilon) \equiv \min\{\delta_1(\tfrac{1}{2}M^{-2}\varepsilon), \omega_1(\tfrac{1}{2}M^{-4}\varepsilon)\}$$

where $M \equiv \max\{\|f_1^{-1}\|_I, \|f_1'\|_I\}$. Then whenever $x, y \in I$ and $|y-x| \leqq \delta(\varepsilon)$, we have

$$\begin{aligned}
&|f_1^{-1}(y)-f_1^{-1}(x)+f_1^{-2}(x)\,f_1'(x)(y-x)| \\
&\quad =|f_1^{-1}(x)\,f_1^{-1}(y)|\,|f_1(y)-f_1(x)-f_1(y)\,f_1^{-1}(x)\,f_1'(x)(y-x)| \\
&\quad \leqq M^2|f_1(y)-f_1(x)-f_1'(x)(y-x)| \\
&\qquad +M^2|f_1'(x)\,f_1(x)^{-1}|\,|f_1(y)-f_1(x)|\,|y-x| \\
&\quad \leqq M^2(\tfrac{1}{2}M^{-2}\varepsilon)|y-x|+M^4(\tfrac{1}{2}M^{-4}\varepsilon)|y-x|=\varepsilon|y-x|.
\end{aligned}$$

Therefore f_1^{-1} is differentiable on I, with derivative $-f_1^{-2}f_1'$ and modulus of differentiability δ.

(d) This is obvious.

(e) This is obvious too. □

(5.3) **Corollary.** *For all positive integers n,*

(5.3.1) $$\frac{dx^n}{dx} = n x^{n-1}.$$

Proof: The proof is by induction on n. When $n=1$, (5.3.1) is just (d) of Theorem (5.2). If (5.3.1) is true for a given value of n, then

$$\frac{dx^{n+1}}{dx} = \frac{d(x \cdot x^n)}{dx} = x^n + x(n x^{n-1}) = (n+1) x^n,$$

by (b) of Theorem (5.2). Therefore (5.3.1) is true for all n. □

Theorem (5.2) and its corollary imply the formula

$$D(f_1 f_2^{-1}) = f_2^{-2}(f_2 D f_1 - f_1 D f_2)$$

for the derivative of a quotient, and the formula

$$D\left(\sum_{k=0}^{n} a_{n-k} x^k\right) = \sum_{k=1}^{n} k\, a_{n-k} x^{k-1}$$

for the derivative of a polynomial.

The next theorem is the so-called *chain rule* for the derivative of a composite function. Its intuitive meaning is that the rate of change of quantity C with respect to quantity A is the product of the rate of change of C with respect to some third quantity B by the rate of change of B with respect to A.

(5.4) **Theorem.** *Let $f\colon I \to \mathbb{R}$ and $g\colon J \to \mathbb{R}$ be differentiable functions such that f maps each compact subinterval of I into a compact subinterval of J. Then $g \circ f$ is differentiable, and*

(5.4.1) $$(g \circ f)' = (g' \circ f) f'.$$

Proof: It is no loss of generality to assume that I and J are compact. Let δ_f be a modulus of differentiability and ω_f a modulus of continuity for f on I. Let δ_g be a modulus of differentiability for g on J. For each $\varepsilon > 0$ write

$$\delta(\varepsilon) \equiv \min\{\omega_f(\delta_g(\alpha)), \delta_f(\beta)\},$$

where

$$\alpha \equiv (1 + \|f'\|_I)^{-1} \frac{\varepsilon}{2} \quad \text{and} \quad \beta \equiv (\alpha + \|g'\|_J)^{-1} \frac{\varepsilon}{2}.$$

Then for x, y in I and $|y-x| \leqq \delta(\varepsilon)$ we have $|f(y)-f(x)| \leqq \delta_g(\alpha)$, so that

$$|g(f(y))-g(f(x))-g'(f(x))(f(y)-f(x))| \leqq \alpha |f(y)-f(x)|.$$

Also

$$|f(y)-f(x)| \leqq \|f'\|_I |y-x| + |f(y)-f(x)-f'(x)(y-x)|$$

and

$$|f(y)-f(x)-f'(x)(y-x)| \leqq \beta |y-x|.$$

Using these inequalities and noting that $\alpha \|f'\|_I < \varepsilon/2$, we compute

$$\begin{aligned}
&|g(f(y))-g(f(x))-g'(f(x))f'(x)(y-x)| \\
&\quad \leqq |g(f(y))-g(f(x))-g'(f(x))(f(y)-f(x))| \\
&\qquad + |g'(f(x))|\,|f(y)-f(x)-f'(x)(y-x)| \\
&\quad \leqq \alpha |f(y)-f(x)| + \|g'\|_J |f(y)-f(x)-f'(x)(y-x)| \\
&\quad \leqq \alpha \|f'\|_I |y-x| + (\alpha + \|g'\|_J)|f(y)-f(x)-f'(x)(y-x)| \\
&\quad < \frac{\varepsilon}{2}|y-x| + \frac{\varepsilon}{2}|y-x| = \varepsilon |y-x|.
\end{aligned}$$

It follows that $g \circ f$ is differentiable on I, with derivative $(g' \circ f) f'$ and modulus of differentiability δ. □

The next lemma is known as *Rolle's theorem.*

(5.5) **Lemma.** *Let f be differentiable on the interval $[a,b]$, and let $f(a) = f(b)$. Then for each $\varepsilon > 0$ there exists x in $[a,b]$ with $|f'(x)| \leqq \varepsilon$.*

Proof: Let δ be a modulus of differentiability for f on $[a,b]$. Let

$$m \equiv \inf\{|f'(x)|: x \in [a,b]\},$$

which exists, by (4.6). Suppose that $m > 0$. We may assume that $f'(a) \geqq m$. For each x in $[a,b]$ we have $f'(x) \geqq m$. For if $f'(x) < m$, then $f'(x) \leqq -m$, so that, by the intermediate value theorem (4.8), there exists ξ in $[a,b]$ with $|f'(\xi)| < m$; this contradicts the definition of m. Now choose points $a = x_0 \leqq x_1 \leqq \ldots \leqq x_n = b$ so that $x_{k+1} - x_k \leqq \delta(\frac{1}{2}m)$ $(0 \leqq k \leqq n-1)$. Then

$$\begin{aligned}
0 &= f(b) - f(a) \\
&= \sum_{k=0}^{n-1} (f(x_{k+1}) - f(x_k)) \\
&= \sum_{k=0}^{n-1} f'(x_k)(x_{k+1}-x_k) + \sum_{k=0}^{n-1} (f(x_{k+1}) - f(x_k) - f'(x_k)(x_{k+1}-x_k)) \\
&\geqq \sum_{k=0}^{n-1} m(x_{k+1}-x_k) - \sum_{k=0}^{n-1} \tfrac{1}{2}m(x_{k+1}-x_k) \\
&= \tfrac{1}{2}m(b-a) > 0.
\end{aligned}$$

This contradiction ensures that $m=0$. The desired conclusion follows immediately. □

Rolle's theorem implies the *mean value theorem*, which gives a basic estimate for the difference of two values of a differentiable function.

(5.6) **Theorem.** *Let f be differentiable on the interval $[a,b]$. Then for each $\varepsilon>0$ there exists x in $[a,b]$ with*

$$|f(b)-f(a)-f'(x)(b-a)|\leqq\varepsilon.$$

Proof: Define the function h on $[a,b]$ by

$$h(x)\equiv(x-a)(f(b)-f(a))-f(x)(b-a) \qquad (x\in[a,b]).$$

Then $h(b)=h(a)=-f(a)(b-a)$. By (5.5), there exists x in $[a,b]$ with

$$\varepsilon\geqq|h'(x)|=|f(b)-f(a)-f'(x)(b-a)|. \quad \square$$

A function f on a proper interval I is *increasing* (respectively, *strictly increasing*) if $f(x)\geqq f(y)$ (respectively, $f(x)>f(y)$) whenever $x,y\in I$ and $x>y$. We say that f is *decreasing* (respectively, *strictly decreasing*) if $-f$ is increasing (respectively, strictly increasing). It follows from Theorem (5.6) that if $f:I\rightarrow\mathbb{R}$ is differentiable on I and $f'(x)\geqq 0$ (respectively, $f'(x)\leqq 0$) for all x in I, then f is increasing (respectively, decreasing) on I.

(5.7) **Definition.** Let $f, f^{(1)}, f^{(2)}, \ldots, f^{(n-1)}$ be differentiable functions on a proper interval I such that

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

and set $f^{(n)}\equiv Df^{(n-1)}$. Then $f^{(n)}$ is called the n^{th} *derivative* of f on I, and is also written $D^n f$; f is then said to be *n times differentiable* on I. The function f itself may be written $f^{(0)}$ or $D^0 f$.

A natural way to simplify a continuous function and set it up for computation is to replace it by a polynomial approximation. The basic result on polynomial approximation of differentiable functions is *Taylor's theorem* ((5.10) below). To see the motivation for Taylor's theorem, consider an n times differentiable function f on an interval I, and a point a in I. It is natural to approximate f by a polynomial of degree n whose derivatives of orders $0,1,\ldots,n$ at a have the same values as the corresponding derivatives of f at a. The unique such