Table of contents

Integers and exponents

Definition. A set of real numbers is called an *inductive set* if

(a) The number 1 is in the set.

(b) For every x in the set, the number $x + 1$ is in the set also.

The set R^+ of positive real numbers is an example of an inductive set. [The number 1 is in R^+ because $1 > 0$. And if x is in R^+ (so that $x > 0$), then $x + 1$ is in R^+ (since $x + 1 > 1 > 0$).]

Definition. A real number that belongs to *every* inductive set is called a *positive integer*; such a number is necessarily positive because R^+ is an inductive set.

Let P denote the set of positive integers. We prove some basic properties of this set.

Theorem 1. Every element of P *is greater than or equal to* 1.

Proof. We shall show that the set A of all real numbers greater than or equal to 1 is inductive. It then follows that every positive integer belongs to this set.

The number 1 belongs to the set A, since $1 \geq 1$. Suppose x belongs to the set A. Then $x \geq 1$; it follows that $x + 1 \geq 1 + 1 > 1$, so that $x + 1$ belongs to the set A. Thus A is inductive. □

Theorem 2. 1 is in P.

Proof. 1 belongs to every inductive set (by definition of "inductive.") Hence 1 belongs to P (by definition of P). □

Theorem 3. If x is in P, so is x + 1.

Proof. Suppose that x is a given element of P. Let I be an arbitrary inductive set. Then x is in I (by definition of P). Hence x + 1 is in I (by definition of "inductive"). Since I is arbitrary, x + 1 is in I for every inductive set I. We conclude that x + 1 is in P (by definition of P). □

Theorem 4 (Principle of induction). Let S be a set of positive integers. If 1 is in S, and if for every x in S, x + 1 is also in S, then necessarily S contains every positive integer.

Proof. S is inductive, by hypothesis. Therefore every positive integer is in S, by definition of P. □

Now we show that P is closed under addition and multiplication.

Theorem 5. If a and b are in P, so is a + b.

Proof. Let a be a fixed positive integer. Then we let S be the set of all positive integers b for which a + b is a positive integer. We shall show that S contains all positive integers; then the theorem is proved. We use the principle of induction.

The number 1 is in S, because a + 1 is a positive integer (by Theorem 3). Given an element b in S, we show that b + 1 is in S. Now a + b is a positive integer by hypothesis; hence (a+b) + 1 is a positive integer by Theorem 3. Thus a + (b+1) is a positive integer, so b + 1 belongs to S, by definition of S. Thus S is inductive. □

<u>Theorem 6</u>. <u>If</u> a <u>and</u> b <u>are in</u> P, <u>so is</u> a · b.

The proof is left as an exercise.

<u>Definition</u>. A number x is called an <u>integer</u> if it is 0, or is a positive integer, or is the negative of a positive integer. It is easy to see that the negative of any integer is an integer, since −(−a) = a and −0 = 0.

Let Z denote the set of integers. We now show that Z is closed under addition, multiplication, <u>and</u> subtraction. Closure under multiplication is easy, so we leave the proof as an exercise:

<u>Theorem 7</u>. <u>If</u> a <u>and</u> b <u>are in</u> Z, <u>so is</u> a · b. □

Closure under addition and subtraction are more difficult:

<u>Theorem 8</u>. <u>If</u> a <u>and</u> b <u>are in</u> Z, <u>so are</u> a + b <u>and</u> a − b.

<u>Proof</u>. We proceed in several steps.

Step 1. We show that the theorem is true in the case where a is a positive integer and $b = 1$. That is, if a is a positive integer, we show that $a + 1$ and $a - 1$ are integers. That $a + 1$ is an integer (in fact, a positive integer) has already been proved. We prove that $a - 1$ is an integer, by induction on a. It is true if $a = 1$, since $a - 1 = 0$ if $a = 1$. Supposing it true for a, we prove it true for $a + 1$. That is, we show $(a+1) - 1$ is an integer. But that is trivial, since $(a+1) - 1 = a$, which is an integer by hypothesis (in fact, a positive integer).

Step 2. We show the theorem is true if a is any integer and $b = 1$.

We consider three cases. If a is a positive integer, this result follows from Step 1. If $a = 0$, the result is immediate, since

$$0 + 1 = 1 \quad \text{and} \quad 0 - 1 = -1.$$

Finally, suppose $a = -c$, where c is a positive integer. Then

$$a + 1 = -c + 1 = -(c-1),$$

$$a - 1 = -c - 1 = -(c+1).$$

Both $c - 1$ and $c + 1$ are integers, by Step 1; then $a + 1$ and $a - 1$ are also integers.

<u>Step 3</u>. We show the theorem is true if a is any integer and b is a positive integer.

We proceed by induction on b, holding a fixed. We know the theorem holds if $b = 1$, by Step 2. Supposing it holds for b, we show it holds for $b + 1$. That is, we show that $a + (b+1)$ and $a - (b+1)$ are integers. Now

$$a + (b+1) = (a+b) + 1,$$

$$a - (b+1) = (a-b) - 1.$$

Both $a + b$ and $a - b$ are integers, by the induction hypothesis; then Step 2 applies to show that $(a+b) + 1$ and $(a-b) - 1$ are integers.

<u>Step 4</u>. The theorem is true in general. Let a be any integer. The case where b is a positive integer was treated in Step 3, and the case where $b = 0$ is trivial. Consider finally the case where $b = -d$, where d is a positive integer. Then

$$a + b = a - d \quad \text{and} \quad a - b = a + d;$$

Step 3 applies to show that both $a - d$ and $a + d$ are integers. □

Now we prove the "obvious" fact that if n is an integer, then $n + 1$ is the "next" integer after n:

<u>Theorem 9</u>. <u>If</u> n <u>is in</u> Z <u>and</u> $n < a < n+1$, <u>then</u> a <u>is not in</u> Z.

<u>Proof</u>. From the hypothesis of the theorem, it follows that

$$0 < a-n < 1.$$

If a were in Z, then $a - n$ would be an integer, by the preceding theorem. But 1 is the smallest positive integer, by Theorem 1. Therefore a is not in Z. □

Now we define integral exponents.

<u>Definition</u>. Let a be any real number. We define a^n, when n is a positive integer, by induction, as follows. We define

$$a^1 = a,$$

and supposing a^n is defined, we define

$$a^{n+1} = a^n \cdot a.$$

Then a^n is defined for every positive integer n. The number n in this expression is called the <u>exponent</u>, and the number a is called the <u>base</u>.

Exponents satisfy three basic laws, which are stated in the following three theorems. They are called the <u>laws of exponents</u>.

<u>Theorem 10</u>. $a^n \cdot a^m = a^{n+m}$.

<u>Proof</u>. Let a and n be fixed. We prove the theorem "by induction on m." The theorem is true for $m = 1$, since $a^n \cdot a^1 = a^n \cdot a = a^{n+1}$ by definition. Suppose it is true for m; we show it is true for $m + 1$. It follows that it holds for all m. We have

$$a^n \cdot a^{m+1} = a^n \cdot (a^m \cdot a) \quad \text{by definition,}$$

$$= (a^n \cdot a^m) \cdot a \quad \text{by associativity of multiplication,}$$

$$= (a^{n+m}) \cdot a \quad \text{by the induction hypothesis,}$$

$$= a^{(n+m)+1} \quad \text{by definition,}$$

$$= a^{n+(m+1)} \quad \text{by associativity of addition.}$$

Thus the theorem is proved for $m + 1$, as desired. □

Similar proofs hold for the following two theorems, whose proofs are left as exercises:

Theorem 11. $(a^n)^m = a^{nm}$. □

Theorem 12. $a^n \cdot b^n = (a \cdot b)^n$. □

Now we define negative exponents.

Definition. Let a be a real number different from zero. We define zero and negative exponents by the rules:

$$a^0 = 1,$$

$$a^{-n} = 1/(a^n) \quad \text{if } n \text{ is a positive integer.}$$

Theorem 13. The "laws of exponents" hold when n and m are arbitrary integers, provided a and b are non-zero.

The proof is left as an exercise.

Later on, (in Section G) we shall extend this definition to define "rational exponents"; that is, we shall define a^r when a is positive and r is rational. Still later (in Section M), we shall extend the definition still further to define a^x when a is positive and x is an arbitrary real number. In each of these cases, the same three laws of exponents will hold.

Exercises

1. Prove Theorems 6 and 7.
2. Prove Theorems 11 and 12.
3. Show that if a set A of integers is bounded above, then A has a largest element. [Hint: Use the least upper bound axiom.]
4. Let F be the set of all real numbers of the form $a + b\sqrt{2}$, where a and b are rational. Show that F is closed under addition, subtraction, multiplication, and division. Conclude that F is an "ordered field", that is, that F satisfies Axioms 1 - 9. Show that F does not contain $\sqrt{3}$.
5. Let n and m be positive integers; let a and b be non-zero real numbers. Let p be any integer. Given that the laws of exponents hold for positive integral exponents, prove them for arbitrary integral exponents as follows:

   (a) Show $a^n a^{-m} = a^{n-m}$ in the three cases $n - m > 0$ and $n - m = 0$ and $n - m < 0$.

   (b) Show $a^{-n} a^{-m} = a^{-n-m}$; and $a^o a^p = a^p$.

   (c) Show $(a^n)^{-m} = a^{-nm} = (a^{-n})^m$.

   (d) Show $(a^{-n})^{-m} = a^{nm}$, and $(a^o)^p = (a^p)^o = a^o$.

   (e) Show $a^{-n} b^{-n} = (ab)^{-n}$, and $a^o b^o = (ab)^o$.

6. Let a and h be real numbers; let m be a positive integer. Show by induction that if a and $a + h$ are positive, then

$$(a+h)^m \geq a^m + ma^{m-1}h.$$

[Note: Be explicit about where you use the fact that a and $a + h$ are positive. Note that h is not assumed to be positive.]
We shall use this result later on.

<u>Square roots, and the existence of irrational numbers.</u>

Definition. If $b^2 = a$, then we say that b is a <u>square</u> <u>root</u> of a.

A negative number has no square root (see Theorem I.20), and the number 0 has only one square root, namely 0. We shall show that a positive real number has exactly two square roots, one positive and one negative.

Theorem. Let $a > 0$. *Then there is a number* $b > 0$ *such that* $b^2 = a$.

Proof. Step 1. Let x *and* y *be positive numbers. Then* $x < y$ *if and only if* $x^2 < y^2$.

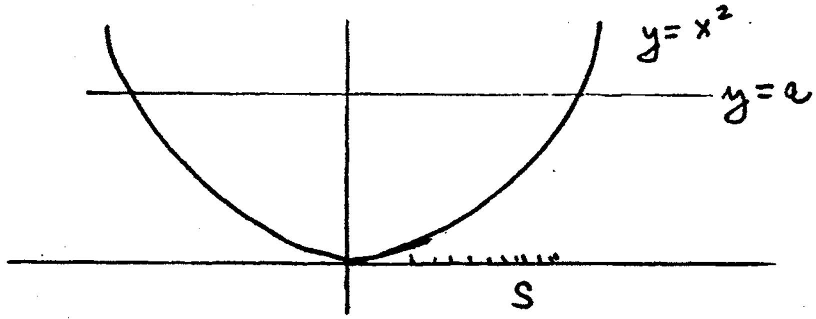
If $x < y$, we multiply both sides, first by x and then by y, to obtain the inequalities

$$x \cdot x < y \cdot x \quad \text{and} \quad y \cdot x < y \cdot y.$$

Thus $x^2 < y^2$. Conversely if $x^2 < y^2$, then it cannot be true that $x = y$ (for that would imply $x^2 = y^2$), or that $y < x$ (for that would imply, by what we just proved, that $y^2 < x^2$). Hence we must have $x < y$.

Step 2. We construct b as follows: Consider the set

$$S = \{x \mid x > 0 \quad \text{and} \quad x^2 < a\}.$$



The set S is nonempty; indeed if x is a number such that $0 < x \leq 1$ and $x < a$, then

$$x^2 < ax \leq a \cdot 1 = a,$$

so that x is in S. Furthermore, S is bounded above; indeed, $1 + a$ is an upper bound on S:

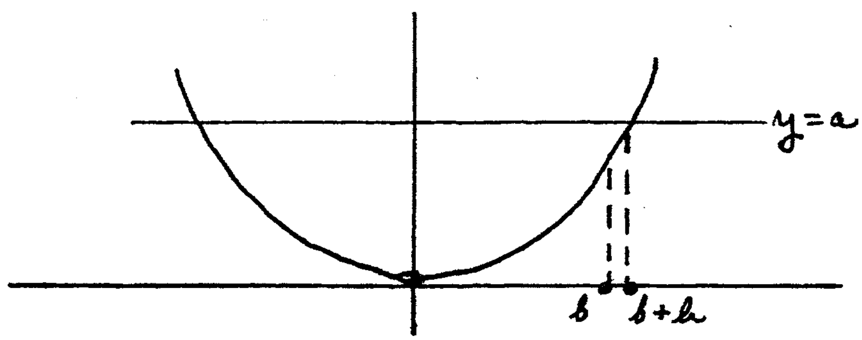
For if x is in S, then $x^2 < a$; since

$$a < 1 + 2a + a^2 = (1+a)^2,$$

it follows from Step 1 that $x < 1 + a$.

Let b denote the supremum of S; we show that $b^2 = a$. We verify this fact by showing that neither inequality $b^2 < a$ or $b^2 > a$ can hold.

Step 3. Assume first that $b^2 < a$. We shall show that there is a positive number h such that $(b+h)^2 < a$. It then follows that $b + h$ belongs to S (by definition of S), contradicting the fact that b is an upper bound for S.



To find h, we proceed as follows: The inequality $(b+h)^2 < a$ is equivalent to the inequality

$$h(2b+h) < a-b^2.$$

Now $a - b^2$ is positive; it seems reasonable that if we take h to be sufficiently small, this inequality will hold. Specifically, we first specify that $h \leq 1$; then we have

$$h(2b+h) \leq h(2b+1).$$

It is then easy to see how small h should be; if we choose $h < (a-b^2)/(2b+1)$, then

$$h(2b+1) < a - b^2$$

and we are finished.

Step 4. Now assume that $b^2 > a$. We shall show that there is a number h such that $0 < h < b$ and $(b-h)^2 > a$. It follows that $b - h$ is an upper bound for S: For if x is in S, then $a > x^2$, so that $(b-h)^2 > x^2$, whence by Step 1, $b - h > x$. This contradicts the fact that b is the least upper bound for S.

To find h, we proceed as follows: The inequality $(b-h)^2 > a$ is equivalent to the inequality

$$h(2b-h) < b^2 - a.$$

Now $b^2 - a$ is positive; it seems reasonable that if h is sufficiently small, this inequality will hold. Our first requirement is that $0 < h < b$. Then we note that $h(2b-h) = 2hb - h^2 < 2hb$. It is now easy to see how small h should be; if we choose $h < (b^2-a)/2b$, then

$$2hb < b^2 - a$$

and we are finished. □

Corollary. If $a > 0$, *then* a *has exactly two square roots.*

We denote the positive square root of a by $\sqrt{a}$.

Proof. Let $b > 0$ and $b^2 = a$. Then $(-b)^2 = a$. Thus a has at least two square roots, b and $-b$. Conversely, if c is any square root of a, then $c^2 = a$, whence

$$(b+c)(b-c) = b^2 - c^2 = 0.$$

It follows that $c = -b$ or $c = b$. □

We now demonstrate the existence of irrational numbers.

Theorem. Let a *be a positive integer; let* $b = \sqrt{a}$. *Then either* b *is a positive integer or* b *is irrational.*

Proof. Suppose that $b = \sqrt{a}$ and b is a rational number that is not an integer. We derive a contradiction.

Let us write $b = m/n$, where m and n are positive integers and n is as small as possible. (I.e., we choose n to be the smallest positive integer such that nb is an integer, and we set $m = nb$.)

Choose q to be the unique integer such that

$$q < m/n < q+1.$$

Then

$$qn < m < qn + n, \text{ or}$$

(*) $$0 < m - qn < n.$$

We compute as follows:

$$(m/n)^2 = b^2 = a,$$
$$m^2 = n^2 a$$
$$m(m-qn) = n(na-qm).$$

Then using (*), we can write

$$b = \frac{m}{n} = \frac{na-qm}{m-qn}.$$

This equation expresses b as a ratio of positive integers; and by (*) the denominator is less than n. Thus we reach a contradiction. □

Corollary. $\sqrt{2}$ is irrational.

Proof. Let $b = \sqrt{2}$. Then b cannot be an integer, for the square of 1 equals 1 while the square of any integer greater than 1 is at least 4. It follows that b is irrational. □

The same proof shows that the number $\sqrt{n}$ is irrational whenever n is a positive integer less than 100 that is not one of the integers 1, 4, 9, 16, 25, 36, 49, 64, or 81.

1.17 The Riemann condition.

The most useful criterion for determining whether f is integrable on $[a,b]$ is given in the following theorem. It is called the Riemann condition for existence of the integral.

Theorem 1. Suppose f is defined on $[a,b]$. Then f is integrable on $[a,b]$ if and only if given any $\varepsilon > 0$, there exist, correspondingly, step functions s and t, with $s \leqslant f \leqslant t$ on $[a,b]$, such that

$$\int_a^b t - \int_a^b s < \varepsilon.$$

Proof. We know from Theorem 1.9 of the text (p. 74) that

$$\int_a^b s \leqslant \underline{I} \leqslant \overline{I} \leqslant \int_a^b t,$$

provided s and t are step functions such that $s \leqslant f \leqslant t$. Suppose that the condition of the theorem holds. Given $\varepsilon > 0$, choose $s \leqslant f \leqslant t$ so that $\int_a^b t - \int_a^b s < \varepsilon$. It follows that $\overline{I} - \underline{I} < \varepsilon$. Because this latter inequality holds for every positive ε, it must be true that $\overline{I} = \underline{I}$. Hence f is integrable.

Conversely, suppose f is integrable. Given $\varepsilon > 0$, choose a step function s with $s \leqslant f$, such that $\int_a^b s$ is within $\varepsilon/2$ of $\underline{I}$. This we can do because $\underline{I}$ is the supremum of the set of such numbers $\int_a^b s$. Similarly choose $t \geqslant f$ so that $\int_a^b t$ is within $\varepsilon/2$ of $\overline{I}$. Now because f is integrable, $\underline{I} = \overline{I}$. Therefore, $\int_a^b t$ and $\int_a^b s$ are within ε of each other. □

We now obtain a slight strengthening of the preceding theorem:

Theorem 2. Suppose f is defined on $[a,b]$. Let A be a number. Suppose that given $\varepsilon > 0$, there exist step functions s and t with $s \leq f \leq t$, such that

$$\int_a^b t - \int_a^b s < \varepsilon, \quad \text{and}$$

$$\int_a^b s \leq A \leq \int_a^b t.$$

Then $\int_a^b f$ exists and equals A.

Proof. We know that $\int_a^b f$ exists, by the preceding theorem. Given $\varepsilon > 0$, choose s and t satisfying the hypotheses of the corollary. Then we have

$$\int_a^b s \leq A \leq \int_a^b t$$

by hypothesis, and

$$\int_a^b s \leq \int_a^b f \leq \int_a^b t$$

by definition of the integral. Because $\int_a^b t - \int_a^b s < \varepsilon$, it follows that

$$|A - \int_a^b f| < \varepsilon.$$

Since the latter inequality holds for every $\varepsilon > 0$, we must have $A = \int_a^b f$. □

<u>The calculation of</u> $\int x^p$.

Let p be a positive integer. We seek to show that the function $f(x) = x^p$ is integrable, and to obtain a formula for it, using the definition of the integral. The proof proceeds by applying the Riemann condition. It involves some hard work, as we shall see.

We need the following lemma, which is easily proved by induction. (See the Exercises of Section A.)

<u>Lemma 3.</u> <u>Let</u> a <u>and</u> $a + h$ <u>be positive real numbers.</u> <u>If</u> m <u>is a positive integer, we have</u>

$$(a+h)^m \geq a^m + ma^{m-1}h.$$

Now we prove our desired integration formula. First we consider a special case:

<u>Theorem 4</u>. <u>Let</u> b <u>be</u> <u>a</u> <u>positive</u> <u>real</u> <u>number</u>. <u>Then</u>

$$\int_0^b x^p \, dx = \frac{b^{p+1}}{p+1} .$$

<u>Proof</u>. The theorem is proved by applying the Riemann condition, as expressed in Theorem 2 preceding. We show that given $\varepsilon > 0$, there are step functions s and t such that

(i) $s \leq f \leq t$ on $[0,b]$,

(ii) $\int_0^b t - \int_0^b s < \varepsilon$, and

(iii) $\int_0^b s \leq b^{p+1}/(p+1) \leq \int_0^b t$.

This will prove our result.

Let $f(x) = x^p$. To define s and t, we begin by partitioning $[0,b]$ into n equal intervals. That is, we consider the partition

$$x_0 = 0,\ x_1 = b/n, \ldots, x_k = kb/n, \ldots, x_n = nb/n = b.$$

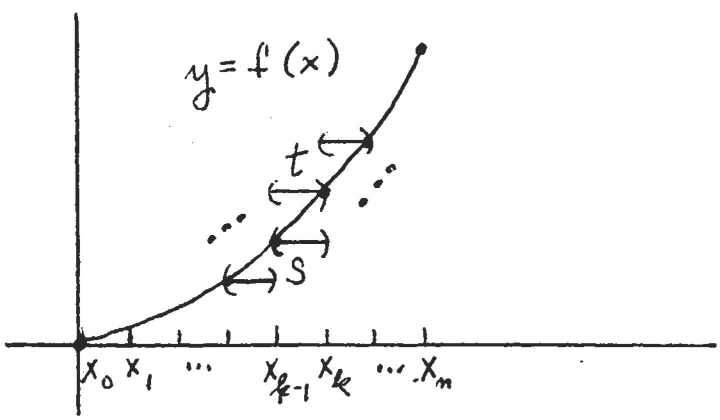
It is not obvious how n should be chosen. In fact, given $\varepsilon > 0$, we shall choose $n > b^{p+1}/\varepsilon$; we will see later that this is the "right" choice.

Now $f(x)$ is strictly increasing on $[0,b]$; that is, if $0 \leq x_1 < x_2$, then $f(x_1) < f(x_2)$. (This result is easily proved by induction on p.) Hence it is easy to define s and t. We let

$$s(x) = f(x_{k-1}) \quad \text{for} \quad x_{k-1} < x < x_k, \quad \text{and}$$

$$t(x) = f(x_k) \quad \text{for} \quad x_{k-1} < x < x_k.$$

Then $s \leq f \leq t$ on the interval (x_{k-1}, x_k).



These equations define s and t except at the partition points. It doesn't matter how we define them at the partition points, so long as $s \leq f \leq t$ holds. Suppose we let $s(x) = 0$ and $t(x) = b^p$ at the partition points. Then $s \leq f \leq t$ on the entire interval $[0,b]$. Thus (i) holds.

Let us compute $\int_0^b t - \int_0^b s$ and show (ii) holds. Now the value of $s(x)$ on (x_{k-1}, x_k) is given by

$$s_k = f(x_{k-1}) = (x_{k-1})^p = ((k-1)b/n)^p.$$

Similarly, the value of $t(x)$ on this same interval is

$$t_k = f(x_k) = (x_k)^p = (kb/n)^p.$$

Furthermore, $x_k - x_{k-1} = b/n$. We compute

$$\int_0^b s = \sum_{k=1}^{n} s_k \cdot (x_k - x_{k-1})$$

$$= \sum_{k=1}^{n} s_k \cdot (b/n)$$

$$= \left[0^p + \left(\frac{b}{n}\right)^p + \left(\frac{2b}{n}\right)^p + \cdots + \left(\frac{(n-1)b}{n}\right)^p\right]\left(\frac{b}{n}\right).$$

Similarly,

$$\int_0^b t = \sum_{k=1}^{n} t_k \cdot (b/n)$$

$$= \left[\left(\frac{b}{n}\right)^p + \left(\frac{2b}{n}\right)^p + \cdots + \left(\frac{nb}{n}\right)^p\right]\left(\frac{b}{n}\right).$$

Subtracting these equations, we obtain

$$\int_0^b t - \int_0^b s = \left(\frac{nb}{n}\right)^p \left(\frac{b}{n}\right) = \frac{b^{p+1}}{n} .$$

Since we (cleverly!) chose n so that $n > b^{p+1}/\varepsilon$, it follows that, as desired.

$$\int_0^b t - \int_0^b s < \varepsilon.$$

To check (iii) requires some work. Plugging the preceding computations into the desired inequalities

$$\int_0^b s \leqslant b^{p+1}/(p+1) \leqslant \int_0^b t,$$

we obtain the inequalities

$$b^{p+1}\ \frac{[0^p+1^p+2^p+\cdots+(n-1)^p]}{n^{p+1}} \leqslant \frac{b^{p+1}}{p+1} \leqslant b^{p+1}\ \frac{[1^p+2^p+\cdots+n^p]}{n^{p+1}} .$$

Simplifying further, we have

$$(*) \qquad 0^p + 1^p + \cdots + (n-1)^p \leqslant \frac{n^{p+1}}{p+1} \leqslant 1^p + 2^p + \cdots + n^p.$$

These are the inequalities we must prove.

We proceed by induction on n, holding p fixed. Both inequalities are trivial when $n = 1$. We assume them true for n, and verify their correctness for $n + 1$.

Let us begin with the left inequality in (*). We add n^p to both sides to obtain

$$0^p + 1^p + \cdots + (n-1)^p + n^p \leqslant \frac{n^{p+1}}{p+1} + n^p.$$

If we can show that

$$\frac{n^{p+1}}{p+1} + n^p \leqslant \frac{(n+1)^{p+1}}{p+1} ,$$

we are through; for the left inequality in (*) then holds for $n + 1$ and the induction step is verified. But this latter inequality follows at once from Lemma 3. If we set $a = n$ and $h = 1$ and $m = p+1$, this lemma takes the form

$$(n+1)^{p+1} \geqslant n^{p+1} + (p+1)n^p,$$

which (if we divide through by $p+1$) is exactly what we want.

Now we consider the right inequality in (*). Adding $(n+1)^p$ to both sides, we obtain the inequality

$$\frac{n^{p+1}}{p+1} + (n+1)^p \leqslant 1^p + 2^p + \cdots + n^p + (n+1)^p.$$

If we can show that

$$\frac{(n+1)^{p+1}}{p+1} \leqslant \frac{n^{p+1}}{p+1} + (n+1)^p,$$

then we are through; the right inequality of (*) then holds for $n + 1$, and the induction step is verified. Once again, we use Lemma 3. If we set $a = n+1$ and $h = -1$ and $m = p+1$ in that lemma, we obtain

$$n^{p+1} \geqslant (n+1)^{p+1} + (p+1)(n+1)^p(-1),$$

which gives our desired inequality. □

Using the basic properties of the integral (proved in D of these notes), we now derive the general integration formula for x^p.

Theorem 5. *Let* p *be a positive integer. Then for all* a *and* b,

$$\int_a^b x^p \, dx = \frac{b^{p+1} - a^{p+1}}{p+1} .$$

Proof. Step 1. We first verify that the given formula holds when $a = 0$. That is, we show that for all b,

$$\int_0^b x^p \, dx = b^{p+1}/(p+1).$$

The case $b > 0$ was proved in Theorem 4. The case $b = 0$ is trivial. Consider the case $b < 0$; let $b = -c$, where $c > 0$. Applying basic properties of the integral, we compute

$$\begin{aligned}
\int_0^c x^p \, dx &= \int_{-c}^0 (-x)^p \, dx && \text{by the reflection property,} \\
&= \int_b^0 (-1)^p x^p \, dx && \text{by laws of exponents,} \\
&= (-1)^p \int_b^0 x^p \, dx && \text{by the linearity property,} \\
&= (-1)^{p+1} \int_0^b x^p \, dx && \text{by our convention.}
\end{aligned}$$

On the other hand, Theorem 4 implies that

$$\begin{aligned}
\int_0^c x^p \, dx &= c^{p+1}/(p+1) = (-b)^{p+1}/(p+1) \\
&= (-1)^{p+1} b^{p+1}/(p+1).
\end{aligned}$$

Comparing these two computations gives us our desired formula.

<u>Step 2</u>. The theorem now follows. We have

$$\int_a^b x^p\,dx = \int_a^0 x^p\,dx + \int_0^b x^p\,dx \quad \text{by the additivity property,}$$

$$= \int_0^b x^p\,dx - \int_0^a x^p\,dx \quad \text{by our convention,}$$

$$= \frac{b^{p+1}}{p+1} - \frac{a^{p+1}}{p+1} \quad \text{by Step 1 of this proof.} \ \square$$

NOTE: Let us introduce the notation

$$f(x)\Big|_a^b = f(b) - f(a).$$

With this notation, the preceding theorem can be written in the form

$$\int_a^b x^p\,dx = \frac{x^{p+1}}{p+1}\Big|_a^b .$$

NOTE: The preceding theorem, along with the linearity property of the integral, now enables us to compute the integral of any polynomial function. We simply "integrate term-by-term."

For example, one has the following computation:

$$\int_1^2 (x^3-3x+5)\,dx = \int_1^2 x^3\,dx - 3\int_1^2 x\,dx + 5\int_1^2 1\,dx$$

$$= \frac{x^4}{4}\Big|_1^2 - 3\,\frac{x^2}{2}\Big|_1^2 + 5\,x\Big|_1^2$$

$$= (4 - \tfrac{1}{4}) - 3(2 - \tfrac{1}{2}) + 5(2 - 1)$$

$$= \frac{15}{4} - \frac{9}{2} + 5 = \frac{17}{4} .$$

Properties of integrals

In this section, we prove the four basic properties of the integral that we shall need.

Theorem. (Properties of the integral)

(1) (Linearity property.) If f and g are integrable on $[a,b]$, then so is $cf + dg$ (here c and d are constants), and furthermore

$$\int_a^b (cf+dg) = c \int_a^b f + d \int_a^b g.$$

(2) (Additivity property.) Suppose f is defined on $[a,c]$ and $a < b < c$. Then

$$\int_a^c f = \int_a^b f + \int_b^c f;$$

the two integrals on the right exist if and only if the integral on the left exists.

(3) (Comparison property.) If $f(x) \leqslant g(x)$ for all x in $[a,b]$, then

$$\int_a^b f \leqslant \int_a^b g,$$

provided both integrals exist.

(4) (Reflection property.) If f is integrable on $[a,b]$, then $f(-x)$ is integrable on $[-b,-a]$, and

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

We use the first three of these properties repeatedly. Property (4) is used only in deriving the formula for $\int_a^b x^p\,dx$.

Let us note that once we make the convention that

$$\int_a^a f = 0 \quad \text{and that} \quad \int_b^a f = -\int_a^b f \quad \text{if} \quad a < b,$$

then the formula

$$\int_a^c f = \int_a^b f + \int_b^c f$$

holds without regard to the requirement that $a < b < c$. The proof is left as an exercise.

<u>Proof.</u> First, one verifies these properties for step functions. This is quite straightforward. Property (3) has already been proved; properties (1) and (2) will be assigned as exercises; and property (4) is proved as follows:

Let s be a step function on $[a,b]$ relative to the partition $x_0,\ldots,x_n$. Let $s(x) = s_k$ for x in (x_{k-1},x_k). The function

$$u(x) = s(-x)$$

is then a step function relative to the partition $-x_k,\ldots,-x_1,-x_0$ of the interval $[-b,-a]$. Indeed, if x is in the interval $(-x_k,-x_{k-1})$, then $-x$ is in the interval (x_{k-1},x_k), so that

$$u(x) = s(-x) = s_k.$$

Then by definition,

$$\int_{-b}^{-a} u(x)\ dx = \sum_{k=n}^{1} s_k \cdot ((-x_{k-1}) - (-x_k)).$$

But

$$\int_a^b s(x)\ dx = \sum_{i=1}^{n} s_k \cdot (x_k - x_{k-1});$$

and these two expressions are equal. Thus (4) holds for step functions.

Step 2. We first prove property (1) in the case where c and d are non-negative. Suppose that f and g are integrable on $[a,b]$. Choose step functions s_i and t_i such that

$$s_1 \leq f \leq t_1 \qquad \text{and} \qquad s_2 \leq g \leq t_2$$

and

$$\int_a^b t_1 - \int_a^b s_1 < \frac{\varepsilon}{2(c+1)} \qquad \text{and} \qquad \int_a^b t_2 - \int_a^b s_2 < \frac{\varepsilon}{2(d+1)}.$$

Then let $s = cs_1 + ds_2$ and let $t = ct_1 + dt_2$. Now s and t are step functions, and (since c and d are non-negative)

$$s \leq cf + dg \leq t .$$

Furthermore, by property (1) for step functions,

$$\int_a^b t - \int_a^b s = [c\int_a^b t_1 + d\int_a^b t_2] - [c\int_a^b s_1 + d\int_a^b s_2]$$

$$= c[\int_a^b t_1 - \int_a^b s_1] + d[\int_a^b t_2 - \int_a^b s_2]$$

$$\leq \frac{c\varepsilon}{2(c+1)} + \frac{d\varepsilon}{2(d+1)} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

Hence the integral of $cf + dg$ exists by the Riemann condition.

Now, by definition of the integral, we have

$$\int_a^b s_1 \leq \int_a^b f \leq \int_a^b t_1 \quad \text{and} \quad \int_a^b s_2 \leq \int_a^b g \leq \int_a^b t_2.$$

We multiply the first set of inequalities by c, and the second by d, and add, obtaining the inequalities:

$$\int_a^b s = c\int_a^b s_1 + d\int_a^b s_2 \leq \boxed{c\int_a^b f + d\int_a^b g} \leq c\int_a^b t_1 + d\int_a^b t_2 = \int_a^b t.$$

Here we use property (1) for step functions again. Since the expression in the box lies between the integrals of s and t, by the Riemann condition it must equal the integral of $cf + dg$.

Step 3. To complete the proof of property (1), it suffices to show that

$$\int_a^b (-f) = -\int_a^b f.$$

This is easy. Given $\varepsilon > 0$, choose step functions s and t such that $s \leq f \leq t$ on $[a,b]$, and

$$\int_a^b t - \int_a^b s < \varepsilon \quad .$$

Then $-s$ and $-t$ are step functions on $[a,b]$, and $-t \leq -f \leq -s$ on $[a,b]$. Furthermore,

$$\left[\int_a^b -s\right] - \left[\int_a^b -t\right] = -\int_a^b s + \int_a^b t < \varepsilon.$$

Here we use property (1) for step functions. Thus the integral of $-f$ exists, by the Riemann condition.

Now by definition of the integral

$$\int_a^b s \leq \int_a^b f \leq \int_a^b t.$$

Multiplying these inequalities by -1, we conclude that

$$\int_a^b (-t) = -\int_a^b t \leq \boxed{-\int_a^b f} \leq -\int_a^b s = \int_a^b (-s) \ .$$

Here we use property (1) for step functions, again. Since the expression in the box lies between the integrals of $-t$ and $-s$, by the Riemann condition it must equal the integral of $-f$.

<u>Step 4.</u> Now we prove property (2). We consider first the "existence" part of the statement. Suppose the integrals

$$\int_a^b f \quad \text{and} \quad \int_b^c f$$

exist. Choose step functions s_1 and t_1 with $s_1 \leq f \leq t_1$ on $[a,b]$, and choose step functions s_2 and t_2 with $s_2 \leq f \leq t_2$ on $[b,c]$, such that

$$\int_a^b t_1 - \int_a^b s_1 < \varepsilon/2 \quad \text{and} \quad \int_b^c t_2 - \int_b^c s_2 < \varepsilon/2.$$

The values of these functions at the partition points do not matter, so we can assume that t_1 and t_2 are equal at c, and s_1 and s_2 are equal at c. Then t_1 and t_2 combine to define a step function t such that $f \leqslant t$ on $[a,c]$, and s_1 and s_2 combine to form a step function s such that $s \leqslant f$ on $[a,c]$. Furthermore, using property (2) for step functions,

$$\int_a^c t - \int_a^c s = \left(\int_a^b t + \int_b^c t\right) - \left(\int_a^b s + \int_b^c s\right)$$

$$= \left(\int_a^b t_1 + \int_b^c t_2\right) - \left(\int_a^b s_1 + \int_b^c s_2\right)$$

(by the way s and t were constructed)

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

Hence $\int_a^c f$ exists, by the Riemann condition.

Conversely, suppose $\int_a^c f$ exists. Then given $\varepsilon > 0$, we can choose step functions s and t with $s \leqslant f \leqslant t$ on $[a,c]$, such that

$$\int_a^c t - \int_a^c s < \varepsilon.$$

Let s_1 and t_1 be the restrictions of s and t, respectively, to $[a,b]$, and let s_2 and t_2 be their restrictions to $[b,c]$. As before, using property (2) for step fucntions, we have

$$\left(\int_a^b t_1 + \int_b^c t_2\right) - \left(\int_a^b s_1 + \int_b^c s_2\right) < \varepsilon,$$

or

$$\left(\int_a^b t_1 - \int_a^b s_1\right) + \left(\int_b^c t_2 - \int_b^c s_2\right) < \varepsilon.$$

Since each expression in parentheses is nonnegative, each is less than ε. Hence $\int_a^b f$ and $\int_b^c f$ exist.

Now in either of these cases, we have

$$\int_a^b s_1 \leqslant \int_a^b f \leqslant \int_a^b t_1 \quad \text{and} \quad \int_b^c s_2 \leqslant \int_b^c f \leqslant \int_b^c t_2,$$

by definition. Adding, we obtain

$$\int_a^c s = \int_a^b s_1 + \int_b^c s_2 \leqslant \boxed{\int_a^b f + \int_b^c f} \leqslant \int_a^b t_1 + \int_b^c t_2 = \int_a^c t.$$

Since the expression in the box lies between $\int_a^c s$ and $\int_a^c t$, the Riemann condition implies that it equals $\int_a^c f$.

Step 5. We prove the comparison property (3).

Consider the set of all step functions s such that $s \leq f$ on $[a,b]$; also consider the set of all step functions t such that $g \leq t$ on $[a,b]$. Because $f \leq g$ on $[a,b]$, we conclude that $s \leq t$ on $[a,b]$, whence

$$\int_a^b s \leq \int_a^b t,$$

<u>because</u> (3) holds for step functions. Holding t fixed and letting s vary, we conclude that

$$\sup \left\{ \int_a^b s \right\} \leq \int_a^b t,$$

for any fixed $t \geq g$. <u>Now</u> letting t vary, we see that

$$\sup \left\{ \int_a^b s \right\} \leq \inf \left\{ \int_a^b t \right\}.$$

That is,

$$\underline{I}(f) \leq \overline{I}(g).$$

Since both f and g are integrable, we have $\underline{I}(f) = \int_a^b f$ and $\overline{I}(g) = \int_a^b g$, so our result is proved.

<u>Step 6.</u> Finally, we prove the reflection property (4).

Given $\varepsilon > 0$, choose step functions s and t so that $s \leq f \leq t$ on $[a,b]$ and $\int_a^b t - \int_a^b s < \varepsilon$. Then $s(-x)$ and $t(-x)$ are step functions on $[-b,-a]$, and

$$s(-x) \leq f(-x) \leq t(-x)$$

on $[-b,-a]$. Now

$$\int_{-b}^{-a} t(-x) - \int_{-b}^{-a} s(-x) = \int_a^b t - \int_a^b s < \varepsilon;$$

here we use the fact that (4) holds for step functions. Thus

$$\int_{-b}^{-a} f(-x)$$

exists., by the Riemann condition. Using (4) for step functions again,

$$\int_a^b s = \int_{-b}^{-a} s(-x) \leq \boxed{\int_{-b}^{-a} f(-x)} \leq \int_{-b}^{-a} t(-x) = \int_a^b t.$$

Since the expression in the box lies between the integrals of s and t, it must by the Riemann condition equal the integral of f.

Exercises

1. Prove property (1) for step functions. [Hint: If s and t are step functions, the first thing to do is to choose a partition P that is compatible with both s and t. Then show $cs + dt$ is a step function compatible with P.]

2. Prove property (2) for step functions. [Hint: If P_1 is a partition of $[a,b]$ and P_2 is a partition of $[b,c]$, then $P_1 \cup P_2$ is a partition of $[a,c]$.]

3. We know (2) holds if $a < b < c$. Show that with our convention, it holds in all cases:

$a = b$,	$a < c < b$,	$c < a < b$,
$a = c$,	$b < a < c$,	$c < b < a$.
$b = c$,	$b < c < a$,	

4. Let $x_0 < \cdots < x_n$ be a partition of $[a,b]$. Let s be a step function on $[a,b]$ such that $s(x) = s_k$ for $x_{k-1} < x < x_k$. Let h be an increasing function on $[a,b]$. Suppose we define

$$\int_a^b s\, dh = \sum_{k=1}^{n} s_k \cdot (h(x_k) - h(x_{k-1})).$$

(a) Show that this integral is well-defined.

(b) Show that this integral satisfies the linearity, additivity, and comparison properties. You need to use the fact that h is increasing in order to prove one of these properties; which one?

[This definition is actually an important one in mathematics. It leads to a generalization of the integral called the Riemann–Stieltjes integral; one defines $\int_a^b f\, dh$ by using upper and lower integrals, just as before. This integral is important in probability theory.]

1.21 Integrability of bounded piecewise-monotonic functions.

The definition of "piecewise-monotonic" is given on p. 77 of the text.

Lemma. If f is bounded on $[a,b]$ and monotonic on (a,b), then f is integrable on $[a,b]$.

(Note that we need to assume f is bounded in the hypothesis of this lemma. The function

$$f(x) = \begin{cases} 1/x & \text{for} \quad 0 < x \leq 1 \\ 0 & \text{for} \quad x = 0 \end{cases}$$

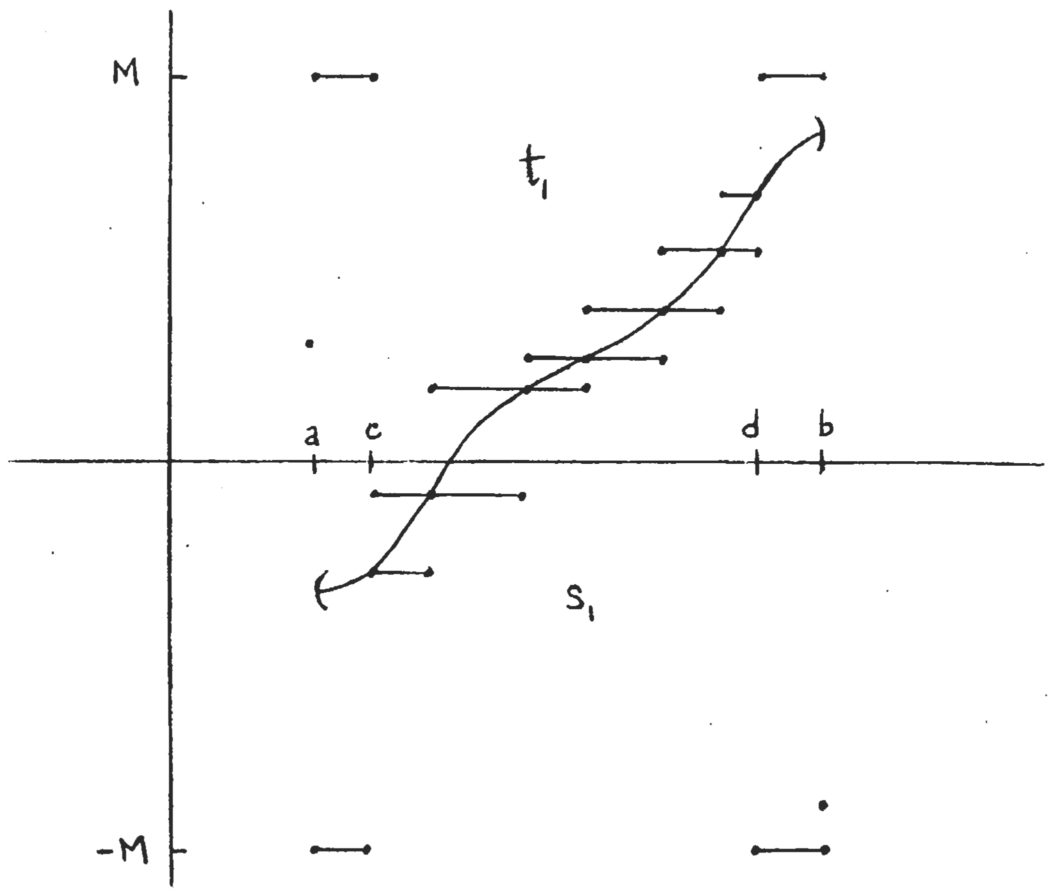
is monotonic on $(0,1)$, but it is not bounded.)

Proof. Choose M so that $-M \leq f(x) \leq M$. We apply the Riemann condition.

Given $\varepsilon > 0$, let us choose numbers c and d (close to a and b respectively), such that

$$a < c < d < b$$

and such that $c - a < \varepsilon/M$ and $b - d < \varepsilon/M$.



Now f is monotonic on [c,d] so it is integrable on [c,d]. Therefore we can find step functions s and t defined on [c,d] such that $s \leq f \leq t$ on [c,d], and such that $\int_c^d t - \int_c^d s < \varepsilon$. <u>Extend</u> t to a step function t_1 defined on [a,b] by setting

$$t_1(x) = \begin{cases} M & \text{for} \quad a \leq x \leq c, \\ t(x) & \text{for} \quad c \leq x \leq d, \\ M & \text{for} \quad d < x \leq b. \end{cases}$$

Similarly, extend s to a step function s_1 defined on [a,b] by setting

$$s_1(x) = \begin{cases} -M & \text{for} \quad a \leqslant x < c, \\ t(x) & \text{for} \quad c \leqslant x \leqslant d, \\ -M & \text{for} \quad d < x \leqslant b. \end{cases}$$

Then $s_1 \leqslant f \leqslant t_1$ on all of $[a,b]$. Furthermore,

$$\int_a^b t_1 - \int_a^b s_1 = \int_a^c (t_1-s_1) + \int_c^d (t_1-s_1) + \int_d^b (t_1-s_1)$$

$$= 2M(c-a) + \int_c^d (t_1-s_1) + 2M(d-b)$$

$$< 2\varepsilon + \varepsilon + 2\varepsilon = 5\varepsilon.$$

Since ε is arbitrary, the Riemann condition is satisfied. □

<u>Theorem</u>. <u>If</u> f <u>is bounded and piecewise-monotonic on</u> $[a,b]$, <u>then</u> f <u>is integrable on</u> $[a,b]$.

<u>Proof</u>. By hypothesis, there is a partition $x_0 < x_1 < \ldots < x_n$ of $[a,b]$ such that f is monotonic on each open interval (x_{i-1}, x_i). By the preceding lemma, f is integrable on $[x_{i-1}, x_i]$ for each i. By the additivity property of integrals (Theorem on p. D.1), it follows that f is integrable on $[a,b]$. □

<u>Exercise</u>

1. Suppose f is bounded on $[a,b]$. Suppose also that f is integrable on <u>every</u> closed interval $[c,d]$ contained in the open interval (a,b). Show that f is integrable on $[a,b]$.

3.3 Continuity of the square root function.

The following theorem shows that the square-root function is continuous for $x \geq 0$. We will give a different proof, based on the intermediate-value theorem, shortly.

Theorem. (i) $\lim_{x \to 0+} \sqrt{x} = 0$.

(ii) If $a > 0$, $\lim_{x \to a} \sqrt{x} = \sqrt{a}$.

Proof. (i) Given $\varepsilon > 0$, we wish to ensure that $|\sqrt{x}-0| < \varepsilon$. This will occur if $x < \varepsilon^2$. So the choice $\delta = \varepsilon^2$ will work; if $0 < x < \varepsilon^2$, then $\sqrt{x} < \varepsilon$.

(ii) Given $\varepsilon > 0$, we wish to ensure that

$$|\sqrt{x} - \sqrt{a}| < \varepsilon.$$

But

$$|\sqrt{x} - \sqrt{a}| = \frac{|x-a|}{\sqrt{x}+\sqrt{a}} \leq \frac{|x-a|}{\sqrt{a}} .$$

So we need merely choose $\delta = \varepsilon\sqrt{a}$; if $|x-a| < \varepsilon\sqrt{a}$, then $|\sqrt{x}-\sqrt{a}| < \varepsilon$. □

Exercises on continuity

1. Show <u>directly</u> <u>from</u> <u>the</u> <u>definition</u> that $f(x) = 1/x$ is continuous at $x = 3$. (That is, given $\epsilon > 0$, define a $\delta > 0$ and show it will work.)

2. Let $f(x)$ be defined for all x, and continuous except for $x = -1$ and $x = 3$. Let

$$g(x) = \begin{cases} x^2 + 1 & \text{for } x > 0, \\ x - 3 & \text{for } x \leq 0. \end{cases}$$

For what values of x can you be sure that $f(g(x))$ is continuous? Explain.

Rational exponents - an application of the intermediate-value theorem.

It is a consequence of the intermediate-value theorem, that, given a positive integer n and a real number $a \geq 0$, there is exactly one real number $b \geq 0$ such that

$$b^n = a.$$

We denote b by $\sqrt[n]{a}$, and call it the $n^{\underline{th}}$ <u>root</u> of a. (See Theorem 3.9, p. 145 of Apostol.)

It follows from the general theorem about continuity of inverses that the $n^{\underline{th}}$ <u>root function</u>, defined by the rule

$$f(x) = \sqrt[n]{x} \quad \text{for} \quad x \geq 0,$$

is continuous. (See Theorem 3.10, p. 147 of Apostol.)

Now (finally!) we can introduce rational exponents. We do so only when the base is a <u>positive</u> real number.

<u>Definition</u>. Let r be a rational number; let a be a <u>positive</u> real number. We can write $r = m/n$, where m and n are integers <u>and</u> n <u>is positive</u>. We then define

$$a^r = (\sqrt[n]{a})^m.$$

(Here we use the fact that $\sqrt[n]{a}$ is non-zero, so m can be negative.)

We must show that this definition makes sense. A problem might arise from the fact that the number r can be represented as a ratio of integers in many different ways. We must show that the value of a^r does not depend on how we represent r. This is the substance of the following lemma.

Lemma 1. *Suppose* $m/n = p/q$, *where* m, n, p, q *are integers, and* n *and* q *are positive. Then* $(\sqrt[n]{a})^m = (\sqrt[q]{a})^p$.

Proof. Let $c = \sqrt[n]{a}$ and $d = \sqrt[q]{a}$. Then $a = c^n$ and $a = d^q$ by definition. Because $m/n = p/q$, we have $mq = np$. Using these facts, we compute

$$a^p = (c^n)^p = c^{np} = c^{mq} = (c^m)^q, \quad \text{and}$$

$$a^p = (d^q)^p = d^{qp} = (d^p)^q, \quad \text{so that}$$

$$(c^m)^q = (d^p)^q.$$

(We use here the laws of integral exponents.) We conclude (by uniqueness of the q^{th} roots) that

$$c^m = d^p, \quad \text{or}$$

$$(\sqrt[n]{a})^m = (\sqrt[q]{a})^p. \quad \square$$

On the basis of Lemma 1, we know that a^r is well-defined if r is a rational number and a is positive. In particular, we have the equation

$$a^{1/n} = \sqrt[n]{a},$$

by definition. The definition of $a^{m/n}$ can then be written in the form

$$a^{m/n} = (a^{1/n})^m.$$

Consider now the three basic laws of exponents. We already know that these laws hold in the following cases:

(i) positive integral exponents; arbitrary bases.

(ii) integral exponents; non-zero bases.

We now comment that these laws also hold in the following case:

(iii) rational exponents; positive bases.

The proof is not difficult, but it is tedious. It is given in Theorem 2 following.

Later on, we shall extend our definition to arbitrary <u>real</u> exponents; that is, we shall define a^x when x is an arbitrary real number (and a is a positive real number). Furthermore, we shall verify that the laws of exponents also holds in this new situation; i.e., in the case:

(iv) real exponents; positive bases.

So you can skip the proof of Theorem 2 if you wish, for we are going to prove the more general result involving real exponents later on.

Before proving Theorem 2, we make the following remark about negative bases: If a is negative, one can still define $\sqrt[n]{a}$ provided n is odd. For in that case there exists exactly one real number b such that $b^n = a$. We shall define $\sqrt[n]{a} = b$ in this case. It is tempting to use exponent notation in this situation, defining $a^{m/n} = (\sqrt[n]{a})^m$ if n is odd and a is negative. However, this practice is dangerous! For the laws of exponents do not always hold in these circumstances. For example, if we used this definition, we would have

$$((-8)^2)^{1/6} = 2, \quad \text{while} \quad (-8)^{1/3} = -2.$$

Thus the second law of exponents would not hold in this situation. For this reason, we make the following convention:

We shall use rational exponent notation only when the base is positive.

Now we verify the laws of exponents for rational exponents and positive bases.

Theorem 2. If r and s are rational numbers, and if a and b are positive real numbers, then

(i) $a^r a^s = a^{r+s}$,

(ii) $(a^r)^s = a^{rs}$,

(iii) $a^r b^r = (ab)^r$.

Proof. Let $r = m/n$ and $s = p/q$, where m, n, p, q are integers, and where n and q are positive.

To prove (i), we note that

$$a^r a^s = a^{m/n}\, a^{p/q}$$

$$= a^{mq/nq}\, a^{np/nq}$$

$$= (\sqrt[nq]{a})^{mq}\, (\sqrt[nq]{a})^{np} \quad \text{by definition,}$$

$$= (\sqrt[nq]{a})^{mq+np} \quad \text{by (iii) for integral exponents,}$$

$$= a^{(mq+np)/nq} \quad \text{by definition,}$$

$$= a^{r+s}.$$

To prove (ii), we verify first that

$$(\sqrt[n]{a})^m = \sqrt[n]{a^m}.$$

Let $c = \sqrt[n]{a}$; then $c^n = a$ by definition. We compute

$$a^m = (c^n)^m = c^{nm} = (c^m)^n$$

by (ii) for integral exponents. By uniqueness of $n^{\underline{th}}$ roots, we have

$$\sqrt[n]{a^m} = c^m = (\sqrt[n]{a})^m,$$

as desired.

It now follows that

$$(*) \qquad a^{m/n} = (a^{1/n})^m = (a^m)^{1/n}.$$

The first equation follows from the definition of $a^{m/n}$, and the second from what we just proved. The formula (*) is of course a special case of our desired formula (ii).

Now we prove (ii) in general: Let

$$c = (a^r)^s = (a^{m/n})^{p/q}.$$

Then

$$c = (((a^m)^{1/n})^p)^{1/q} \quad \text{by (*) (applied twice)}$$

$$= (((a^m)^p)^{1/n})^{1/q} \quad \text{by (*).}$$

It follows that

$$c^q = (((a^m)^p)^{1/n}, \qquad \text{and}$$

$$(c^q)^n = (a^m)^p, \quad \text{by definition, so that}$$

$$c^{qn} = a^{mp} \quad \text{by (ii) for integral exponents.}$$

Then

$$c = \sqrt[qn]{a^{mp}} \qquad \text{by definition,}$$

$$= (a^{mp})^{1/nq} \qquad \text{by definition,}$$

$$= a^{mp/nq} \qquad \text{by } (*),$$

$$= a^{rs}.$$

To check (iii), let $c = \sqrt[n]{a}$ and $d = \sqrt[n]{b}$. We first note that

$$(cd)^n = c^n d^n \quad \text{by (iii) for integral exponents,}$$

$$= ab \qquad \text{by definition.}$$

It follows that

$$cd = \sqrt[n]{ab}.$$

We then prove (iii) as follows:

$$a^{m/n}b^{m/n} = (\sqrt[n]{a})^m(\sqrt[n]{b})^m = c^m d^m \quad \text{by definition,}$$

$$= (cd)^m \quad \text{by (iii) for integral exponents,}$$

$$= (\sqrt[n]{ab})^m \quad \text{by (*),}$$

$$= (ab)^{m/n} \quad \text{by definition.}$$

Thus the three laws hold for rational exponents. □

The small span theorem and the extreme-value theorem.

There are three fundamental theorems concerning a function that is continuous on a closed interval $[a,b]$. The first is the Intermediate-Value Theorem, which is stated and proved on p. 144 of Apostol. We consider the other two here.

We begin with a definition.

Definition. If the function f is bounded on the interval $[c,d]$, we define the span of f on this interval as follows: Let

$$M(f) = \sup \{f(x);\ x \in [c,d]\},$$

$$m(f) = \inf \{f(x);\ x \in [c,d]\}.$$

Then we define the span of f by the equation

$$\text{span}(f) = M(f) - m(f).$$

Theorem. (The small-span theorem). Let f be continuous on the closed interval $[a,b]$. Given $\varepsilon > 0$, there is a partition

$$x_0 < x_1 < \dots < x_n$$

of the interval $[a,b]$ such that f is bounded on each closed subinterval $[x_{i-1}, x_i]$, and such that the span of f on each closed subinterval is at most ε.

Proof. The proof of this theorem is a bit tricky, but the theorem is so useful that the effort is justified.

One proceeds by what is sometimes called "the method of successive bisections," or less elegantly, "chopping the interval in half repeatedly"!

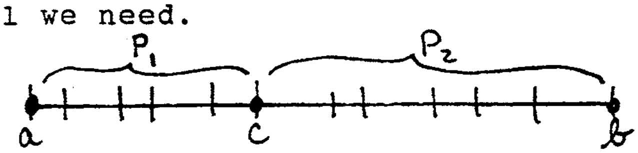
For purposes of this proof, let us make up some terminology. If f is defined on the interval [c,d], we shall say that

f <u>is</u> ε<u>-pleasant</u> <u>on</u> <u>the</u> <u>interval</u> [c,d]

if there is some partition of [c,d] such that the span of f on each closed subinterval of the partition is at most ε. If there is no such partition, we shall say that f is ε-<u>unpleasant</u> on [c,d]!

Our object then is to prove that if f is continuous on [a,b], then f is ε-pleasant on [a,b].

We make the following remark: Let c be any number with $a < c < b$. If f is ε-pleasant on [a,c], and if f is also ε-pleasant on [c,b], then f is ε-pleasant on all of [a,b]. The proof is easy. One merely takes the appropriate partitions of [a,c] and [c,b] and puts them together to get a partition of [a,b]. This simple fact is all we need.



We now prove the theorem. Assume that f is continuous on [a,b], and that f is ε-unpleasant on [a,b]. We shall derive a contradiction.

<u>First step.</u> Let c be the midpoint of [a,b]. Since f is ε-unpleasant on [a,b], it must be true that f is ε-unpleasant on either [a,c], or on [c,b], or on

both. Let $[a_1,b_1]$ denote the left half $[a,c]$ of our interval if f is ε-unpleasant on $[a,c]$. Otherwise, let $[a_1,b_1]$ denote the right half $[c,b]$ of our interval. In either case, f is ε-unpleasant on $[a_1,b_1]$.

<u>General</u> <u>step.</u> Assume that $[a_n,b_n]$ is an interval contained in $[a,b]$ and that f is ε-unpleasant on $[a_n,b_n]$. Let c_n be the midpoint of $[a_n,b_n]$. As before, let $[a_{n+1},b_{n+1}]$ denote the left half $[a_n,c_n]$ of $[a_n,b_n]$ if f is ε-unpleasant on this half; otherwise, let $[a_{n+1},b_{n+1}]$ denote the right half. In either case, f is ε-unpleasant on $[a_{n+1},b_{n+1}]$.

We now have defined a sequence of intervals

$$[a_1,b_1],\ [a_2,b_2],\ldots$$

that is "nested" in the sense that each interval contains all its successors. Furthermore, each interval has half the length of the preceding one, and f is ε-unpleasant on each of them. It follows by an easy induction proof that the length of the n^{th} interval is

$$b_n - a_n = (b-a)/2^n.$$

Because the intervals are nested, we have

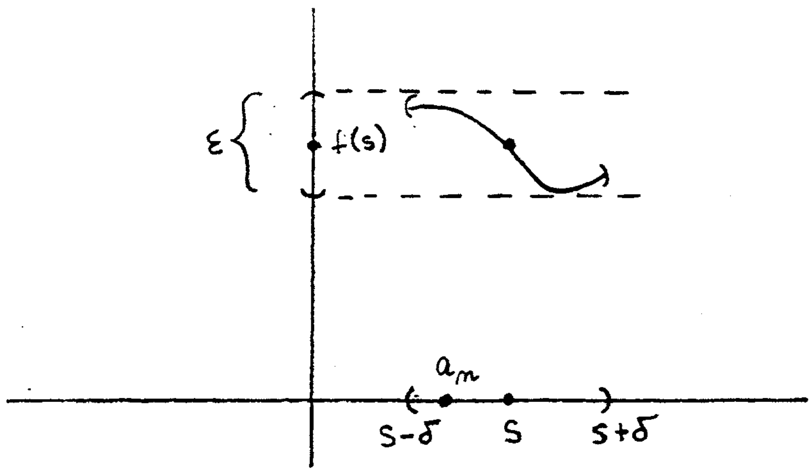
$$a \leq a_1 \leq a_2 \leq \ \ldots \ \leq b_2 \leq b_1 \leq b\ .$$

Let s be the least upper bound of the numbers a_i. Since all the numbers a_i belong to the interval $[a,b]$, so does s. Now we derive a contradiction. We have three cases, according as $a < s < b$, or $s = a$, or $s = b$.

Consider first the case where $a < s < b$. Since f is continuous at s, we may choose a neighborhood $(s-\delta, s+\delta)$ of s such that

$$|f(x) - f(s)| < \varepsilon/2$$

for all x in this neighborhood of s.



Because s is the least upper bound of the numbers a_i, there must be an n such that

$$s - \delta < a_n \leq s.$$

Because the a_i are increasing, we have

$$s - \delta < a_n \leq a_{n+1} \leq a_{n+2} \cdots \leq s.$$

Now let us choose m so large that $m > n$, and so that

$$(b - a)/2^m < \delta.$$

Then $b_m - a_m < \delta$, so that $b_m < a_m + \delta \leq s + \delta$. Then the interval $[a_m, b_m]$ is contained in the interval $(s - \delta, s + \delta)$. Therefore the span of f in the interval $[a_m, b_m]$ is at most ε.

It follows that f is ε-pleasant on $[a_m, b_m]$. Indeed, for <u>any</u> partition of $[a_m, b_m]$, the span of f in each subinterval of the partition will be at most ε.

The other two cases proceed similarly. For example, suppose $s = a$. In this case, we have $a_n = a$ for all n. (This means merely that we chose the left half of the interval at each step of the construction.) Since f is continuous at a, there must be a δ such that

$$|f(x) - f(a)| < \varepsilon/2$$

for $a \leq x < a + \delta$. Choose m so large that $(b-a)/2^m < \delta$. Then $b_m - a_m < \delta$, so that $b_m < a_m + \delta = a + \delta$. Then the interval $[a_m, b_m]$ is contained in the interval $[a, a + \delta)$, so the span of f in the interval $[a_m, b_m]$ is at most ε. It follows that f is ε-pleasant in $[a_m, b_m]$, as before.

The proof when $s = b$ is similar. □

Here is an important application of this theorem:

<u>Theorem.</u> <u>If</u> f <u>is continuous on</u> $[a,b]$, <u>then</u> f <u>is bounded on</u> $[a,b]$ <u>and integrable on</u> $[a,b]$.

<u>Proof.</u> Given $\varepsilon > 0$, choose a partition

$$x_0 < x_1 < \ldots < x_n$$

of $[a,b]$ such that the span of f on each closed subinterval of the partition is at most ε. Define

$$s_k = \inf \{f(x) \quad \text{for} \quad x_{k-1} \leq x \leq x_k\},$$

$$t_k = \sup \{f(x) \quad \text{for} \quad x_{k-1} \leq x \leq x_k\}.$$

Now f is bounded on $[a,b]$; indeed, $f(x)$ is at most the largest of the numbers $t_1,\ldots,t_n$ and at least the smallest of the numbers $s_1,\ldots,s_n$.

We define step functions s and t by letting their values equal s_k and t_k, respectively, on (x_{k-1},x_k); at the partition points, we set $s(x_k) = t(x_k) = f(x_k)$. Then $s(x) \leq f(x) \leq t(x)$ for all x. Now $t_k - s_k \leq \varepsilon$ because the span of f on $[x_{k-1},x_k]$ is at most ε. Therefore

$$\int_a^b t - \int_a^b s = \int_a^b (t-s) \leq \varepsilon(b-a).$$

The Riemann condition applies to show that f is integrable on $[a,b]$. □

Finally, we prove the third big theorem about continuous functions.

<u>Theorem. (Extreme-value theorem). Let</u> f <u>be continuous on the closed interval</u> $[a,b]$. <u>Then there are points</u> x_0 <u>and</u> x_1 <u>of</u> $[a,b]$ <u>such that for every</u> x <u>in</u> $[a,b]$, <u>we have</u>

$$f(x_0) \leq f(x) \leq f(x_1).$$

The number $M = f(x_1)$ is called the <u>maximum-value</u> of f on $[a,b]$, and the number $m = f(x_0)$ is called the <u>minimum-value</u> of $f(x)$ on $[a,b]$. Both are called <u>extreme values</u> of f on $[a,b]$.

<u>Proof.</u> We show that the point x_1 exists; the proof that x_0 exists is similar.

We know that f is bounded on $[a,b]$, by the previous theorem; define

$$M = \sup \{f(x) \text{ for } x \text{ in } [a,b]\} .$$

We wish to show that $M = f(x_1)$ for some point x_1 of $[a,b]$. Suppose this is not true. We derive a contradiction.

Then by assumption, we have $f(x) < M$ for all x in $[a,b]$. Consider the function

$$g(x) = \frac{1}{M - f(x)} .$$

Since the denominator does not vanish, g is well-defined; because f is continuous, so is g. Therefore g is bounded on $[a,b]$, by the preceding theorem. Choose C so that $g(x) \leq C$ for x in $[a,b]$. Then

$$0 < \frac{1}{M - f(x)} \leq C,$$

so that $1/C \leq M - f(x)$, or $f(x) \leq M - 1/C$, for every x in $[a,b]$. This contradicts the fact that M is the least upper bound of the values of $f(x)$ on $[a,b]$. □

We shall use the extreme-value theorem shortly, when we prove the fundamental theorems of calculus.

Exercises on the intermediate-value, extreme-value, and small-span theorems.

1. Let $f(x) = x + [x]$ for $0 \leq x \leq 3$.

   (a) Draw the graph of f; show f is strictly increasing.

   (b) Define a function g by the following rule:
   If $y = f(x)$ for some x in $[0,3]$, let $g(y)$ equal that x. Because f is strictly increasing, g is well-defined. What is the domain of g?

2. Let $f(x) = x^4 + 2x^2 + 1$ for $0 \leq x \leq 10$.

   (a) Show f is strictly increasing; what is the domain of its inverse function g?

   (b) Find an expression for g, using radicals.
   ("Radicals" are the symbols $\sqrt{\ }$, $\sqrt[3]{\ }$, $\sqrt[4]{\ }$, etc.)

3. Let $f(x) = 2x^5 - 5x^4 + 5$ for $x \geq 2$. We will show later that f is strictly increasing (since its derivative is positive for $x > 2$).

   (a) Show that f is unbounded.
   (Hint: $f(x) > x^4(2x-5) > 2x - 5$.)

   (b) What is the domain of its inverse function g?

   [Note: A famous theorem of Modern Algebra states that it is not possible to express g in terms of algebraic operations and radicals.]

4. Let $f(x)$ be defined and continuous and strictly increasing for $x \geq 0$; suppose that $f(0) = a$.

   (a) Show that if f is unbounded, then f takes on every value greater than a.

   (b) If f is bounded, let M be the least upper bound of the values of f. Show that f takes on every value between a and M, but does not take on the value M.

5. Show by example that the conclusion of the intermediate value theorem can fail if f is only continuous on $[a,b)$ and bounded on $[a,b]$.

6. Show by example that the conclusion of the extreme value theorem can fail if f is only continuous on $[a,b)$ and bounded on $[a,b]$.

7. Let $f(x) = x$ for $0 \leq x < 1$; let $f(1) = 5$. Show that the conclusion of the small span theorem fails for the function $f(x)$.

8. A function f defined on $[a,b]$ is said to be <u>piecewise-continuous</u> if it is continuous except at finitely many points of $[a,b]$. Said differently, f is piecewise-continuous if there is some partition of $[a,b]$ such that f is continuous on each open subinterval determined by the partition.

   (a) Show that if f is bounded on $[a,b]$ and piecewise-continuous on $[a,b]$, then f is integrable on $[a,b]$. [Hint: Examine the proof we gave for piecewise-monotonic functions.]

   (b) Show that f can be piecewise-continuous on $[a,b]$ without being bounded on $[a,b]$.

9. Consider the function

$$f(x) = \begin{cases} (-1)^{[1/x]} & \text{if } x > 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Show that f is integrable on $[0,1]$. Show that f is neither piecewise-monotonic nor piecewise-continuous on $[0,1]$.

10. Challenge exercise. Define a function f on the interval $[0,1]$ by setting $f(x) = 0$ if x is irrational ; and $f(x) = 1/n$ if x is a rational number of the form $x = m/n$, where m and n are positive integers having no common factors other than 1 ; and $f(0) = 1$.

(a) Show that f is integrable on $[0,1]$.

(b) Show that f is continuous at each irrational and discontinuous at each rational.

Theorem. *Let* m *and* n *be integers; let* $n > 0$. *Let*

$$h(y) = (\sqrt[n]{y})^m \text{ for } y > 0.$$

Then h *is differentiable, and*

$$h'(y) = \frac{m}{n}(\sqrt[n]{y})^m \cdot y^{-1}.$$

Proof. Step 1. We first prove the theorem in the case $m = 1$. Let $f(x) = x^n$ for $x > 0$. Then the inverse function to f, denoted $g(y)$, is the n^{th} root function. By the theorem on the derivative of an inverse function, $g'(y)$ exists and

$$g'(y) = \frac{1}{f'(x)},$$

where $x = g(y)$. Now $f'(x) = n\,x^{n-1}$. Therefore

$$g'(y) = \frac{1}{n\,x^{n-1}} = \frac{1}{n(\sqrt[n]{y})^{n-1}}$$

$$= \frac{1}{n(\sqrt[n]{y})^n\,(\sqrt[n]{y})^{-1}} = \frac{\sqrt[n]{y}}{n\,y}$$

$$= \frac{1}{n}\sqrt[n]{y}\cdot y^{-1}.$$

Step 2. We prove the theorem in general. If $m = 0$, it is trivial. Otherwise, we apply the chain rule. We have

$$h(y) = (\sqrt[n]{y})^m;$$

then

$$h'(y) = m(\sqrt[n]{y})^{m-1}\,[\tfrac{1}{n}\sqrt[n]{y}\cdot y^{-1}]$$

$$= \frac{m}{n}(\sqrt[n]{y})^m \cdot y^{-1}. \quad \square$$

Once one has checked that the laws of exponents hold for rational exponents (Notes G), one can write this formula in a manner that is much easier to remember:

Theorem. Let r be a rational constant; let $h(x) = x^r$ for $x > 0$. Then h is differentiable and

$$h'(x) = rx^{r-1}.$$

We will give a different proof of this theorem later on, one which holds when r is an arbitrary real constant.

Exercises on derivatives

1. Define a new derivative by the formula

$$D^{\#}f(x) = \lim_{h\to 0} \frac{(f(x+h))^3 - (f(x))^3}{h} .$$

Assuming that f and g are continuous, and that $D^{\#}f(x)$ and $D^{\#}g(x)$ exist, derive formulas for $D^{\#}(f(x)g(x))$ and $D^{\#}(1/f(x))$ in terms of $D^{\#}f(x)$ and $D^{\#}g(x)$.

2. Define a new derivative by the formula

$$D^{*}f(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h^2} .$$

If $f(x) = x^2 + 3$, show that $D^{*}f(x)$ exists only at the point $x = 0$, and compute $D^{*}f(0)$.

3. Assume the usual properties of the sine and cosine functions. Define

$$f(x) = \begin{cases} x\sin(1/x) & \text{for} \quad x \neq 0, \\ 0 & \text{for} \quad x = 0. \end{cases}$$

$$g(x) = \begin{cases} x^2\sin(1/x) & \text{for} \quad x \neq 0, \\ 0 & \text{for} \quad x = 0. \end{cases}$$

(a) Apply the definition of derivative to determine whether $f'(0)$ and $g'(0)$ exist. Compute them if they do exist.

(b) Show that $f'(x)$ and $g'(x)$ are not continuous at $x = 0$. Explain which part of the definition of continuity is violated in each case.

4. If $f(x) = u(v(x))$, write down a formula for $f''(x)$, assuming u', u'', v', and v'' exist at the points in question.

5. Suppose $f(x)$ is continuous and strictly monotonic on the interval $[a,b]$; let $g(y)$ be its inverse function. Show that if f' and f'' exist on $[a,b]$, then g'' exists at each point y for which $f'(g(y)) \neq 0$, and

$$g''(y) = - \frac{f''(g(y))}{[f'(g(y))]^3} .$$

6. Let $f(x) = 2x^5 - 5x^4 + 5$ for $x \geq 2$; let $g(y)$ be the inverse function to f. Let c be the number for which $f(c) = 0$. (See Exercise 3 of Section G.)

(a) Note that $g(0) = c$; show that $g(-11) = 2$ and $g(86) = 3$.

(b) Show that

$$g'(0) = \frac{1}{10c^3(c-2)}.$$

(c) Compute $g'(-11)$ and $g'(86)$.

7. Suppose f is a function defined for all x such that:

$f(1) = 2$ and $f(2) = 3$ and $f(3) = 4$;

$f'(1) = 6$ and $f'(2) = 10$ and $f'(3) = 7$;

$f''(1) = 3$ and $f''(2) = 2$ and $f''(3) = 1$.

(a) Let $h(x) = f(f(x))$; compute $h(1)$, $h'(1)$, and $h''(1)$. (Answers: 3, 60, 102.)

(b) Suppose f is strictly increasing. Let $g(y)$ be its inverse function, and compute $g(3)$, $g'(3)$, and $g''(3)$. (Answers: 2, 1/10, -1/500.)

8. Derive a formula for the derivative of $\sqrt{x}$ directly from the definition.

9. Using the fact that $f(x) = \sqrt[3]{x}$ is defined and continuous for all x, derive a formula for $f'(x)$, when $x \neq 0$, directly from the definition of the derivative.

[Hint: $a^3 - b^3 = (a-b)(a^2+ab+b^2)$. Let $a = \sqrt[3]{x+h}$ and $b = \sqrt[3]{x}$.]

The fundamental theorems of calculus.

Here are the two basic theorems relating integrals and derivatives. You should know the proofs of these theorems.

First, we need to discuss "one-sided" derivatives.

If a function is defined on an interval [a,b], we know what it means for f to be continuous on [a,b]. It means that f is continuous in the ordinary sense at each point of the open interval (a,b), and that f satisfies the appropriate version of one-sided continuity at each of the end points a and b.

What shall it mean for f to be differentiable on [a,b]? It will mean that f is differentiable in the ordinary sense at each point of (a,b), and that the appropriate one-sided derivatives of f exist at the end points. More specifically, the one-sided derivative of f at a is the one-sided limit

$$f'(a) = \lim_{h \to 0+} \frac{f(a+h) - f(a)}{h} .$$

Similarly , the one-sided derivative of f at b is the one-sided limit

$$f'(b) = \lim_{h \to 0-} \frac{f(a+h) - f(a)}{h} .$$

Of course, if it happens that f is defined and differentiable in some open interval that contains [a,b], then it is automatically true that f is differentiable on [a,b], in the sense just defined. This is the situation that usually occurs in practice.

Now we prove a lemma:

Lemma 1. Suppose f *is integrable on the closed interval having* c *and* d *as end points and that* $|f(x)| \le M$ *on this interval. Then*

$$\left|\int_c^d f\right| \le M|d - c|.$$

Proof. Assume first that $c < d$. Now

$$-M \le f(x) \le M$$

for all x in $[c,d]$. The comparison theorem for integrals tells us that

$$-M(d-c) \le \int_c^d f \le M(d-c).$$

On the other hand, if $d < c$, the comparison theorem tells us that

$$-M(c-d) \le \int_d^c f \le M(c-d).$$

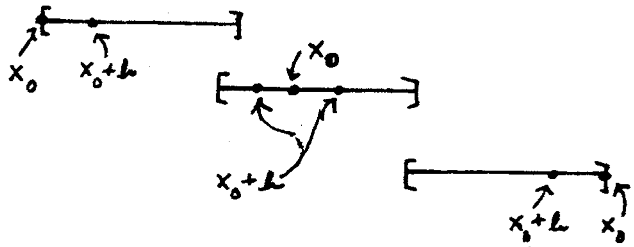
In either case, we conclude that $\left|\int_c^d f\right| \le M|d-c|$. □

Theorem 2. Suppose f *is integrable on* $[a,b]$. *Let* c *be a point of* $[a,b]$. *Let*

$$A(x) = \int_c^x f(t)\,dt$$

for x *in* $[a,b]$. *Then* $A(x)$ *is continuous on* $[a,b]$.

Proof. Throughout this proof, let h denote a number such that $h \neq 0$ and $x_0 + h$ is in $[a,b]$. This means that h is small, and that h is positive if $x_0 = a$, and h is negative if $x_0 = b$.



We know f is bounded on $[a,b]$; choose M so that $|f(x)| \leq M$ for x in $[a,b]$. Then we compute

$$A(x_0+h) - A(x_0) = \int_c^{x_0+h} f - \int_c^{x_0} f$$

$$= \int_{x_0}^{x_0+h} f(x)\,dx.$$

By the preceding lemma, we have

$$|A(x_0+h)-A(x_0)| = \left|\int_{x_0}^{x_0+h} f(x)\,dx\right| \leq M|h|.$$

We use this inequality to show that $A(x)$ is continuous at x_0. Given $\varepsilon > 0$, let $\delta = \varepsilon/M$. Then if $|h| < \delta$, the above inequality shows that

$$|A(x_0+h) - A(x_0)| \leq M|h| < M(\varepsilon/M) = \varepsilon. \quad \square$$

Theorem 3. (First fundamental theorem of calculus.) *Let f be integrable on [a,b]; let c be a point of [a,b]. Let*

$$A(x) = \int_c^x f(t)\,dt.$$

If f is continuous at the point x_0 of [a,b], then $A'(x_0)$ exists and $A'(x_0) = f(x_0)$.

Proof. Let h be as in the preceding proof. As before, we compute

$$A(x_0+h) - A(x_0) = \int_{x_0}^{x_0+h} f(t)\,dt.$$

Now since $f(x_0)$ is a constant, we have the equation

$$f(x_0)\cdot h = \int_{x_0}^{x_0+h} f(x_0)\,dt.$$

Subtracting and using linearity, we see that

$$(*) \qquad \frac{A(x_0+h) - A(x_0)}{h} - f(x_0) = \frac{1}{h}\int_{x_0}^{x_0+h} (f(t)-f(x_0))\,dt.$$

To prove that $A'(x_0)$ exists and equals $f(x_0)$ is equivalent to showing that

$$\lim_{h \to 0} \frac{A(x_0+h) - A(x_0)}{h} = f(x_0).$$

(The limit is a one-sided limit if x_0 equals a or b).
To prove this statement, it suffices to show that the right side of (*) approaches zero.

We use the continuity of f at x_o. Given $\varepsilon > 0$, choose $\delta > 0$ so that

$$|f(x) - f(x_o)| < \varepsilon$$

whenever $|x-x_0| < \delta$ and x is in $[a,b]$. Then if $0 < |h| < \delta$, the inequality

$$|f(x) - f(x_0)| < \varepsilon$$

holds for all x in the interval having end points x_0 and $x_0 + h$. It follows from the preceding lemma that

$$\left| \int_{x_0}^{x_0+h} (f(x) - f(x_0))dx \right| \leq \varepsilon |h|.$$

We conclude that for $0 < |h| < \delta$,

$$\left| \frac{A(x_0+h) - A(x_0)}{h} - f(x_0) \right| \leq \varepsilon,$$

as desired. □

Theorem 4. (Second fundamental theorem of calculus.) Suppose $P(x)$ is defined on $[a,b]$ and that $P'(x)$ exists and is continuous on $[a,b]$. Let c be a point of $[a,b]$. Then for all x in $[a,b]$,

$$\int_c^x P'(t)dt = P(x) - P(c).$$

Proof. Since $P'(x)$ is continuous on $[a,b]$, it is integrable. Furthermore, if

$$A(x) = \int_c^x P',$$

then by the first fundamental theorem, $A'(x)$ exists and equals $P'(x)$. We conclude that the function $P(x) - A(x)$ is continuous on $[a,b]$ (in fact, differentiable on $[a,b]$) and that its derivative vanishes on $[a,b]$.

It follows from the mean-value theorem (see p. 187 of the text) that $P(x) - A(x)$ is constant on $[a,b]$. Let

$$P(x) - A(x) = K$$

for all x in $[a,b]$. Setting $x = c$, we see that

$$P(c) - 0 = K.$$

Therefore,

$$A(x) = P(x) - K = P(x) - P(c),$$

<u>Definition</u>. If $f(x)$ is a function defined on $[a,b]$, a <u>primitive</u> of f is a function $P(x)$ defined on $[a,b]$ such that $P'(x) = f(x)$. (Such a function P does not always exist, of course.) We also call $P(x)$ an <u>antiderivative</u> of f, and we write

$$\int f(x)\,dx = P(x) + C.$$

The second fundamental theorem says that if f is continuous, one can compute $\int_a^b f$ provided one can find a primitive P of f; for then $\int_a^b f = P(b) - P(a)$.

<u>Remark</u>. These two theorems may be summarized as follows:

(1) $$\frac{d}{dx}\int_c^x f = f(x) \quad \underline{\text{if}}\ f\ \underline{\text{is continuous at}}\ x.$$

(2) $$\int_c^x \frac{d}{dx} P = P(x) - P(c) \quad \underline{\text{if}}\ \frac{dP}{dx}\ \underline{\text{is continuous on the interval having end points}}\ c\ \underline{\text{and}}\ x.$$

These theorems say, in essence, that integration and differentiation are inverse operations. But in each case, there is a continuity requirement that the integrand must satisfy in order for the theorem to hold.

<u>Corollary 5</u>. <u>Let</u> r <u>be a rational constant with</u> $r \neq -1$. <u>If</u> a <u>and</u> b <u>are positive real numbers, then</u>

$$\int_a^b x^r\,dx = \frac{b^{r+1} - a^{r+1}}{r+1}.$$

Proof. Let $P(x) = x^{r+1}/(r+1)$ for all $x > 0$. Then we have shown (see notes I) that $P'(x) = x^r$ for all $x > 0$. Since the function x^r is continuous for all $x > 0$, it is continuous on $[a,b]$, so the second fundamental theorem applies to give our formula. □

Exercises

1. If $b > 0$, show that

$$\int_0^b [t]\,dt = \tfrac{1}{2}[b]\,(2b - [b] - 1).$$

[Hint: Let $n = [b]$. Evaluate $\int_0^n [t]\,dt$ and $\int_n^b [t]\,dt$.]

2. Let $A(x) = \int_0^x [t]\,dt$.

(a) Use the first fundamental theorem of calculus to show that $A'(x) = [x]$ when x is not an integer, and that $A'(x)$ does not exist when x is an integer. See the figure on p. 127 of Apostol.

(b) Use the formula of Exercise 1 to verify the same result.

3. Use the chain rule to evaluate:

(a) $\frac{d}{dx}\int_1^{x^2} \frac{dt}{1+t^5}$.   (b) $\frac{d}{dx}\int_{x^3}^{1} \frac{dt}{1+t^5}$.   (c) $\frac{d}{dx}\int_{x^3}^{x^2} \frac{dt}{1+t^5}$.

4. Suppose $F(t)$ is continuous for $a \leq t \leq b$. Let

$$A(x) = \int_a^x F(t)\,dt$$

for x in $[a,b]$.

(a) Suppose $g(u)$ is a function whose values lie in the interval $[a,b]$, with g differentiable. Consider the function

$$B(u) = A(g(u)) = \int_a^{g(u)} F(t)\,dt.$$

Use the chain rule to show that

$$B'(u) = F(g(u))g'(u).$$

We express this fact in words as follows: The derivative of

$$\int_a^{g(u)} F(t)\,dt$$

with respect to u equals the integrand, evaluated at the upper limit, times the derivative of the upper limit.

(b) If $g(u)$ and $h(u)$ are two functions whose values lie in $[a,b]$, and if g and h are differentiable, derive a formula for the derivative with respect to u of

$$\int_{h(u)}^{g(u)} F(t)\,dt.$$

[Hint: Write

$$\int_h^g F = \int_a^g F - \int_a^h F.\ ]$$

5. Suppose f is integrable on $[a,b]$. Let

$$A(x) = \int_a^x f(t)\,dt$$

for x in $[a,b]$. Let x_0 be a point of (a,b).

(a) If f is continuous at x_0, what can you say about the function $A(x)$?

(b) If f is continuous on $[a,b]$, what can you say about $A(x)$?

(c) If f is continuous from the right at x_0, what can you say about $A(x)$? [Hint: Examine the proof of the first fundamental theorem.]

(d) If f' exists on $[a,b]$ what can you say about $A(x)$?

Justify your answers, using the theorems we have proved.

The trigonometric functions.

For the present, we shall assume the following theorem concerning existence of the sine and cosine functions. Later on, when we study power series, we shall prove this theorem.

Theorem 1. *There exist two functions* $\sin x$ *and* $\cos x$, *defined for all real numbers* x, *satisfying the following conditions:*

(i) $\sin 0 = 0$; $\cos 0 = 1$.

(ii) $D \sin x = \cos x$; $D \cos x = -\sin x$.

From the properties listed in this theorem, one can derive all the other familiar properties of the trigonometric functions, as we shall see.

Theorem 2. *Conditions (i) and (ii) specify the functions* $\sin x$ *and* $\cos x$ *uniquely.*

Proof. Step 1. We first note the following fact: If $u(x)$ and $v(x)$ are functions satisfying the equations $u'(x) = v(x)$ and $v'(x) = -u(x)$ for all x, then $u^2 + v^2$ is constant. This result follows from the fact that the derivative of $u^2 + v^2$ is $2uu' + 2vv' = 2uv - 2vu = 0$.

Step 2. We prove the theorem. Suppose $\operatorname{Sin} x$ and $\operatorname{Cos} x$ are two other functions satisfying these conditions. Let

$$u(x) = \sin x - \operatorname{Sin} x \quad \text{and} \quad v(x) = \cos x - \operatorname{Cos} x .$$

Direct computation shows that $u' = v$ and $v' = -u$. Then $u^2 + v^2 = K$ for some constant K. Substituting $x = 0$, we see that $K = 0$. It follows that $\sin x - \operatorname{Sin} x = 0$ for all x and $\cos x - \operatorname{Cos} x = 0$ for all x. □

Theorem 3. The functions $\sin x$ and $\cos x$ have the following properties:

(a) $\sin^2 x + \cos^2 x = 1$.

(b) $\sin x$ and $\cos x$ are continuous for all x and have values in the interval $[-1,1]$.

(c) $\int \sin x\, dx = -\cos x + C$; $\int \cos x\, dx = \sin x + C$.

(d) $\sin(x+y) = \sin x \cos y + \cos x \sin y$,

$\cos(x+y) = \cos x \cos y - \sin x \sin y$.

(e) $\sin 2x = 2 \sin x \cos x$

$\cos 2x = \cos^2 x - \sin^2 x$

$= 2\cos^2 x - 1 = 1 - 2\sin^2 x$.

(f) $\sin(-x) = -\sin x$,

$\cos(-x) = \cos x$.

(g) There is at least one positive number x such that $\cos x = 0$.

(h) There is a unique number $a > 0$ such that $\cos a = 0$ and such that $\cos x$ is positive for $0 < x < a$. We commonly denote the number $2a$ by π.

(i) The number π satisfies the inequalities

$$3 \le \pi \le 3.6.$$

(j) $\sin(\pi/2) = 1$; $\sin x$ is strictly increasing on $[0,\pi/2]$. $\cos(\pi/2) = 0$; $\cos x$ is strictly decreasing on $[0,\pi/2]$.

(k) $\sin x$ is strictly decreasing from 1 to 0 on $[\pi/2,\pi]$. $\cos x$ is strictly decreasing from 0 to -1 on $[\pi/2,\pi]$.

(ℓ) $\sin(x+\pi) = -\sin x$; $\cos(x+\pi) = -\cos x$.

(m) $\sin(x+2\pi) = \sin x$; $\cos(x+2\pi) = \cos x$.

<u>Proof</u>. The functions $\sin x$ and $\cos x$ are continuous because they are differentiable. Applying Step 1 of the preceding proof to the functions $\sin x$ and $\cos x$, we see that $\sin^2 x + \cos^2 x$ is K, a constant. Substituting $x = 0$, we have $K = 1$. Parts (a), (b), (c) follow.

To prove (d), define

$$u(x) = \sin(x+a) - \sin x \cos a - \cos x \sin a;$$

$$v(x) = \cos(x+a) - \cos x \cos a + \sin x \sin a.$$

Then $u'(x) = v(x)$ and $v'(x) = -u(x)$, as you can check. It follows that $u^2 + v^2 = K$, a constant. Substituting $x = 0$, one sees that $K = 0$. Then $u(x) = 0$ and $v(x) = 0$ for all x.

(e) These formulas follow at once from (d).

To prove (f) we set

$$u(x) = \cos(-x) - \cos x$$
$$v(x) = \sin(-x) + \sin x .$$

Then $u' = v$ and $v' = -u$, as you can check. It follows that $u^2 + v^2 = K$, a ccnstant. Setting $x = 0$, we see that $K = 0$. Then $u = 0$ and $v = 0$ for all x..

(g) We suppose first that $\cos x > 0$ for all $x > 0$ and derive a contradiction. If $\cos x > 0$ for all $x > 0$, then since $D \sin x = \cos x$, the function $\sin x$ is strictly increasing for all $x \geq 0$. Therefore

$$0 = \sin 0 < \sin x < \sin 2x = 2 \sin x \cos x$$

for all $x > 0$. We can rewrite this in the form

$$0 < (2 \cos x - 1) \sin x$$

for $x > 0$. Since $\sin x$ is positive, then $2 \cos x - 1$ is positive, so $\cos x > 1/2$ for $x > 0$. The comparison theorem for integrals implies that

$$\int_0^b \cos x \, dx \geq \int_0^b 1/2 \, dx, \quad \text{or}$$

$$\sin b \geq 1/2\, b,$$

for all $b > 0$. This is impossible if $b > 2$.

Therefore $\cos b \leq 0$ for at least one $b > 0$. Since $\cos 0 = 1$, the intermediate-value theorem applied to the interval $[0,b]$ gives us a point x such that $0 < x \leq b$ and $\cos x = 0$.

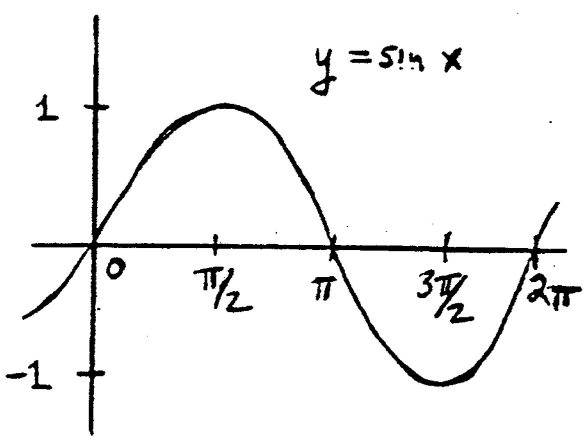
(h) Let a be the inf of the set S consisting of those positive values of x for which $\cos x = 0$. We show that $\cos a = 0$. If $\cos a \neq 0$, then by continuity there is an open interval I about a on which $\cos x \neq 0$. This fact implies that the right hand end point of I is a lower bound for S, a contradiction. Therefore $\cos a = 0$. By choice of a, we know that $\cos x$ is nonzero on the interval $[0,a)$. Because $\cos 0$ is positive, the intermediate-value theorem implies that $\cos x$ must be positive for $0 \leq x < a$.

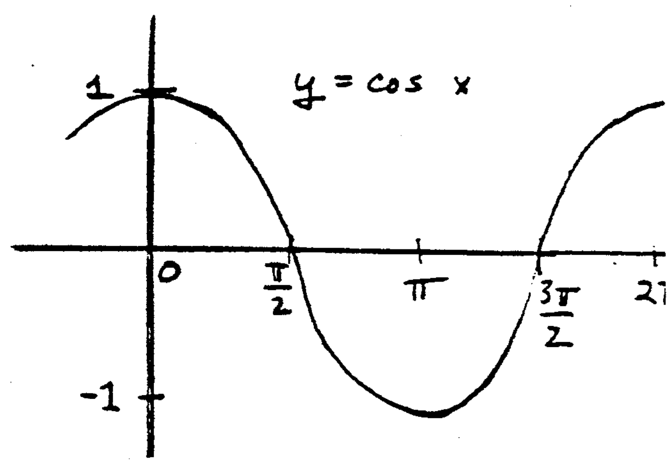
We leave (i) as an exercise.

(j) Because $\cos(\pi/2) = 0$, we must have $\sin(\pi/2) = \pm 1$. Because $\cos x > 0$ on $(0,\pi/2)$, the function $\sin x$ is strictly increasing on $[0,\pi/2]$; therefore $\sin(\pi/2) = +1$. We know $\cos(\pi/2) = 0$; because $D \cos x = -\sin x$ and $\sin x$ is positive on $(0,\pi/2)$, $\cos x$ is strictly decreasing on $[0,\pi/2]$.

Condition (k) follows by computing $\sin(x+\pi/2)$ and $\cos(x+\pi/2)$; conditions (ℓ) and (m) follow similarly. □

<u>Remark</u>. Conditions (j) - (m) suffice to show that the graphs of $y = \sin x$ and $y = \cos x$ are the familiar wave-shaped curves, as you can check.





<u>Definition</u>. We define $\tan x = (\sin x)/(\cos x)$ and $\sec x = 1/\cos x$. This definition makes sense whenever $\cos x \neq 0$; i.e., whenever $x \neq k\pi + \pi/2$, where k is an integer.

<u>Theorem</u> 4. (a) $D \tan x = \sec^2 x$; <u>therefore</u> $\tan x$ <u>is strictly increasing on any interval on which it is defined</u>.

(b) $\tan 0 = 0$; $\tan(-x) = -\tan x$; $\tan x$ <u>is unbounded above and below on the interval</u> $(-\pi/2,\pi/2)$.

(c) $\tan(x+\pi) = \tan x$.

(d) $\tan(x+y) = (\tan x + \tan y)/(1 - \tan x \tan y)$ *if* $\tan x$ *and* $\tan y$ *and* $\tan(x+y)$ *are defined*.

(e) $D \sec x = \sec x \tan x$.

(f) $1 + \tan^2 x = \sec^2 x$.

Proof. The proof is left as an exercise. □

Exercises.

1. Prove the half-angle formulas

$$\cos^2 x = \tfrac{1}{2}(1 + \cos 2x), \qquad \sin^2 x = \tfrac{1}{2}(1 - \cos 2x).$$

2. Show that $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$.

3. (a) Show that

$$\cos 3x = 4\cos^3 x - 3\cos x.$$

Compute $\cos \pi/6$ and $\sin \pi/6$.

(b) Compute $\cos \pi/3$ and $\sin \pi/3$.

4. Prove Theorem 4.

5. (a) Show that $\sqrt{3}/2 \leq \cos x \leq 1$ for $0 \leq x \leq \pi/6$.

(b) Apply the comparison theorem for integrals to the inequalities of (a) to conclude that $3 \leq \pi \leq 3.6$.

6. Use the half-angle formulas and the substitution rule to compute

$$\int_0^{\pi/3} \sin^2 x \, dx.$$

Remark on motivation

The general pattern of the development of calculus, after the basic theorems have been established, is to introduce and study various functions that arise in the applications.

The most elementary such functions are of course those involving only algebraic operations: the integral powers of x, the polynomial functions, and the rational functions. The next functions one studies arise naturally from certain real-life situations, which are important enough to study in detail. Both the trigonometric functions and the exponential function arise in this way, as we now describe.

In a first course in trigonometry, the sine and cosine functions are usually introduced as functions of an angle, and their study is motivated by their usefulness in "solving triangles", a problem of importance to navigators and surveyors. This approach is of course misleading, for one would never include them in a beginning course in calculus if that were their primary use.

In fact, their importance comes instead from a physical situation called "simple harmonic motion" or "one-dimensional vibration". It arises frequently, and is characterized by the equation

$$f''(x) = -k^2 f(x).$$

This equation is called a differential equation because it is an equation involving an unknown function f(x) and one or more of its derivatives. "Solving" such an equation means finding a function satisfying the equation. This particular equation arises, for example, in describing the motion of a particle acted on by a force that is proportional to the displacement of the particle from its "rest" position. The sine and cosine functions arise in solving this equation; one checks readily that the function sin kx and cos kx do satisfy this equation. But more generally, it is a fact that every solution of this equation can be expressed in terms of these two functions. This we now prove:

Theorem 6. Suppose f(x) is defined for all x, and satisfies the equation

$$f''(x) = -k^2 f(x),$$

where $k \neq 0$. Let $a = f(0)$ and $b = f'(0)$. Then

$$f(x) = a \cos(kx) + b \sin(kx)$$

for all x.

Proof. We show that given k and a and b, there is at most one function f(x) satisfying the conditions:

$$f''(x) = -k^2 f(x),$$
$$f(0) = a,$$
$$f'(0) = b.$$

The theorem follows, since the function $f(x) = a \cos(kx) + b \sin(kx)$ does satisfy these conditions, by direct computation.

So suppose f and g are two functions satisfying the given conditions. Set

$$u(x) = f(x/k) - g(x/k),$$
$$v(x) = \frac{1}{k}[f'(x/k) - g'(x/k)].$$

Then

$$u'(x) = v(x), \text{ and}$$
$$v'(x) = \frac{1}{k^2}[f''(x/k) - g''(x/k)] = -u(x).$$

It follows from the proof of Theorem 2 that $u^2 + v^2$ is constant. Setting $x = 0$, we see that this constant is zero. Thus u and v are identically zero. □

The exponential and logarithm functions

In this section, we study the exponential and logarithm functions and derive their properties.

We also define a^b for $a > 0$ and b arbitrary, and we verify the laws of exponents.

As we did for the trig functions, we shall assume a theorem concerning the existence of the exponential function, postponing the proof until after we have studied power series. Thus we assume the following:

Theorem 1. There exists a function E(x), defined for all real numbers x, satisfying the following conditions

$$E'(x) = E(x);\ \ E(0) = 1.$$

We call E the exponential function, for reasons to be seen. It is sometimes denoted exp(x).

Theorem 2. (i) The equation

$$E(a+b) = E(a)E(b)$$

holds for all a and b. In particular, $E(a)E(-a) = 1$ for all a.

(ii) E(x) is continuous, positive, and strictly increasing.

(iii) The conditions $E'(x) = E(x)$ and $E(0) = 1$ determine E(x) uniquely.

(iv) If n is an integer and a is a real number, then

$$E(na) = E(a)^n.$$

In particular, if e is defined by the equation $e = E(1)$, then

$$E(n) = e^n.$$

This equation shows why E is called the "exponential function".

(v) The number e satisfies the inequalities

$$2 \leq e \leq 4.$$

(vi) E(x) takes on every positive real value exactly once.

Proof. (i) For fixed b, let us set

$$f(x) = E(x+b)E(-x).$$

Then

$$\begin{aligned} f'(x) &= E'(x+b)E(-x) - E(x+b)E'(-x) \\ &= E(x+b)E(-x) - E(x+b)E(-x) \\ &= 0. \end{aligned}$$

Hence f equals a constant K. Setting $x = 0$, we see that $K = E(b)$. Thus

$$(*) \qquad E(x+b)E(-x) = E(b)$$

for all x and b.

If we set $b = 0$ in equation (*), we obtain the equation

$$E(x)E(-x) = 1,$$

which holds for all x. If then we multiply both sides of equation (*) by E(x), we obtain the equation

$$E(x)E(x+b)E(-x) = E(x)E(b), \quad \text{or}$$

$$E(x+b) = E(x)E(b).$$

Setting $x = a$ gives our desired equation.

(ii) E(x) is continuous because it is differentiable. The equation

$$E(x)E(-x) = 1$$

implies that $E(x) \neq 0$ for all x. The intermediate-value theorem then applies to show that, since E(x) is positive for $x = 0$, it is positive for all x. It follows that $E'(x) = E(x)$ is positive for all x, so that E is strictly increasing.

(iii) Let $\bar{E}(x)$ be another function satisfying the given conditions. Set

$$g(x) = \bar{E}(x)E(-x).$$

One checks readily that $g'(x) = 0$, so that $g(x)$ is constant. Setting $x = 0$, we see this constant is 1. Hence $\bar{E}(x)E(-x) = 1$, or $\bar{E}(x) = E(x)$.

(iv) One proves the result for positive integers by induction: The equation

$$E(na) = E(a)^n$$

holds for $n = 1$, trivially. If it holds for n, compute

$$\begin{aligned} E((n+1)a) &= E(na+a) \\ &= E(na)E(a) \quad \text{by (i)} \\ &= E(a)^n E(a) \quad \text{by the induction hypothesis,} \\ &= E(a)^{n+1}. \end{aligned}$$

The equation holds when $n = 0$ by definition (both sides equal 1), and it holds for negative integers because $E(na)\,E(-na) = 1$, so that

$$E(-na) = 1/E(na) = 1/E(a)^n.$$

(v) Because E is increasing, $E(x) \geq 1$ for $x \geq 0$. The comparison theorem implies that

$$1 \leq \int_0^1 E(x)dx = \int_0^1 E'(x)dx = E(1) - E(0) = e - 1.$$

Hence $e \geq 2$. We leave the other inequality as an exercise.

(vi) It follows from what we have proved that

$$E(n) = e^n \geq 2^n, \text{ and}$$

$$E(-n) = 1/E(n) \leq 1/2^n.$$

Given any positive real number r, we may choose a positive integer n such that

$1/2^n \leq r \leq 2^n$. The intermediate-value theorem then implies that E(x) takes on the value r for some x in the interval [−n,n]. □

Remark. Since $e^n = E(n)$ for all integers n, it seems reasonable to define e^x for arbitrary real x by the equation $e^x = E(x)$. Theorems 1 and 2 can then be restated in this new notation, which is standard, as follows:

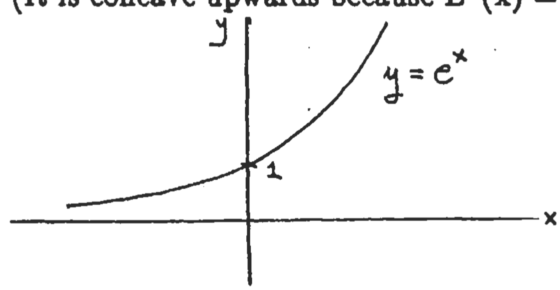
$$\frac{d}{dx}(e^x) = e^x \quad \text{and} \int e^x\,dx = e^x + C.$$

$$e^{a+b} = e^a \cdot e^b,$$

$$e^{na} = (e^a)^n.$$

The latter two equations are special cases of the laws of exponents, which shall prove shortly in full generality.

Remark. The preceding theorem implies that the function $E(x) = e^x$ has the following familiar graph. (It is concave upwards because $E'(x) = E(x) > 0$.)



Exercises 1. Show that

$$\frac{1}{2} \leq \int_0^{1/2} E(x)dx \leq \frac{1}{2}\sqrt{e}.$$

Show the integral equals $\sqrt{e} - 1$; conclude that $2.25 \leq e \leq 4$.

2. Show more generally, by integrating E(x) over the interval [0,1/n], that

$$1/n \leq \sqrt[n]{e} - 1 \leq \sqrt[n]{e}/n.$$

Conclude that

$$(1+\tfrac{1}{n})^n \leq e \leq (1+\tfrac{1}{n})^{n+1}.$$

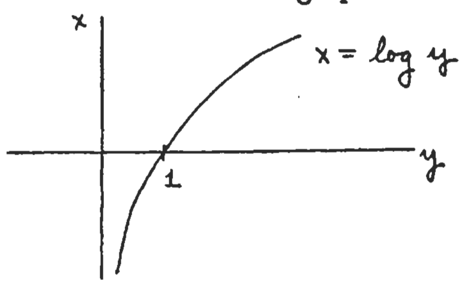
These inequalities give a (not very useful) way of computing e. Try $n = 2$ and ~~$n = 3$~~ $n=4$ in these formulas, using your calculator.

The logarithm function.

Definition. The function e^x is strictly increasing and takes on each positive value exactly once. We define the (natural) logarithm function to be its inverse. That is, if y is any positive number, we define

$$\log y = x \quad \text{if and only if} \quad y = e^x.$$

The logarithm function thus has the graph



It is strictly increasing and continuous. It is defined only for $y > 0$, and it takes on every real value.

The fact that these functions are inverses of each other implies that:

$$e^{\log y} = y \quad \text{if } y > 0,$$
$$\log(e^x) = x \quad \text{for all } x.$$

Theorem 3. The logarithm function has the following properties:

(i) $\frac{d}{dx}(\log x) = \frac{1}{x}$ and

$$\int \frac{1}{x}\,dx = \log|x| + C.$$

(ii) $\log(ab) = \log a + \log b$ if a and b are positive.

(iii) $\log(a^n) = n \log a$ if n is an integer and a is positive.

Proof (i) Let $f(x) = e^x$ and $g(y) = \log y$. Then g is the inverse function to f. We use the formula for the derivative of an inverse function:

$$g'(y) = \frac{1}{f'(g(y))}.$$

Now $f'(x) = f(x)$ for all x, so

$$g'(y) = \frac{1}{f'(g(y))} = \frac{1}{e^{\log y}} = \frac{1}{y}.$$

since $f(g(y)) = e^{\log y} = y$.

If x is positive, the derivative of $\log|x| = \log x$ is $1/x$. If x is negative, the derivative of $\log|x| = \log(-x)$ is $(-1)/(-x) = 1/x$. Thus $\int dx/x = \log|x| + C$.

(ii) Given a and b, let

$$x = \log a \quad \text{and} \quad y = \log b.$$

We have the equation

$$e^{x+y} = e^x \cdot e^y.$$

Since both sides of this equation are positive, we can take their logs to conclude that

$$x + y = \log(e^x \cdot e^y),$$

so that

$$\log a + \log b = \log(a \cdot b),$$

as desired.

(iv) Given $a > 0$, let $x = \log a$. Then $a = e^x$, so

$$a^n = (e^x)^n = e^{nx} \quad \text{by (iv) of Theorem 2,}$$

so that

$$\log a^n = nx = n(\log a). \quad \square$$

Exponents

Theorem 4. There is one and only one function a^x, defined for all positive a and all real x, such that the following four conditions hold:

(i) The function a^x is positive and continuous.

(ii) $a^1 = a$.

(iii) $a^{x+y} = a^x a^y$.

(iv) $(a^x)^y = a^{xy}$.

This function satisfies the following additional conditions:

(v) $a^x b^x = (ab)^x$.

(vi) $\log a^x = x \log a$.

Proof. Uniqueness. Suppose a^x is defined and satisfies conditions (i) – (iv). Conditions (ii) and (iii) imply that

$$a^1 = a \quad \text{and} \quad a^{n+1} = a^n \cdot a$$

for every positive integer n. The equation

$$a^1 \cdot a^0 = a^{1+0} = a^1$$

implies (since $a > 0$) that $a^0 = 1$. Finally, the equation

$$a^n a^{-n} = a^0 = 1$$

implies that

$$a^{-n} = 1/a^n.$$

Hence integral powers of a must be defined as we have defined them earlier.

Now if n is a positive integer and m is any integer, (iv) implies that

$$(a^{1/n})^n = a \quad \text{or} \quad a^{1/n} = \sqrt[n]{a} .$$

Then, using (iv) again, we see that

$$(a^{m/n}) = (a^{1/n})^m = (\sqrt[n]{a})^m .$$

Thus for x rational, a^x is completely determined by conditions (ii) – (iv) and positivity.

Continuity now implies that a^x is determined for all x: Suppose f(x) and g(x) are two functions that satisfy (i)–(iv). Let x_0 be arbitrary. Given $\epsilon > 0$, choose δ so that

$|f(x)-f(x_0)| < \epsilon$ and $|g(x)-g(x_0)| < \epsilon$ for $|x-x_0| < \delta$. Then choose x_1 rational with $|x_1-x_0| < \delta$. It follows from what we have already showed that $f(x_1) = g(x_1)$. Then $|f(x_0)-g(x_0)| < 2\epsilon$. Since ϵ is arbitrary, we must have $f(x_0) = g(x_0)$.

Existence. We motivate the definition as follows: If n is a positive integer, then

$$\log a = \log(\sqrt[n]{a})^n = n \log \sqrt[n]{a},$$

so that

$$\log (\sqrt[n]{a})^m = m \log \sqrt[n]{a} = \frac{m}{n} \log a,$$

or

$$(\sqrt[n]{a})^m = E(\frac{m}{n} \log a).$$

This equation suggests the following definition.

We define, for arbitrary x,

$$a^x = E(x \log a).$$

Then condition (vi) holds trivially, for $\log a^x = x \log a$ by definition.

We show that the other conditions of the theorem are satisfied:

(i) a^x is positive and continuous.

(ii) $a^1 = E(\log a) = a$.

(iii)
$$\begin{aligned} a^{x+y} &= E((x+y)\log a) && \text{by definition,} \\ &= E(x \log a + y \log a) && \text{by distributivity,} \\ &= E(x \log a) \cdot E(y \log a) && \text{by Theorem 2,} \\ &= a^x \cdot a^y. && \text{by definition.} \end{aligned}$$

(iv)	$(a^x)^y = E(y \log a^x)$	by definition,
	$= E(y(x \log a))$	by (vi),
	$= E((xy) \log a)$	by associativity,
	$= a^{xy}.$	by definition.
(v)	$(ab)^x = E(x \log (ab))$	by definition,
	$= E(x(\log a + \log b))$	by Theorem 3,
	$= E(x \log a + x \log b)$	by distributivity,
	$= E(x \log a) \cdot E(x \log b)$	by Theorem 2,
	$= a^x \cdot b^x.$	by definition. □

<u>Theorem 5</u>. <u>Let</u> c <u>be</u> <u>a</u> <u>real</u> <u>constant</u>. <u>Then</u>

$$D(x^c) = cx^{c-1} \quad \text{if } x > 0,$$

$$\int x^c \, dx = \frac{x^{c+1}}{c+1} + C \quad \text{if } c \neq -1 \text{ and } x > 0.$$

<u>Proof</u>. Since $x^c = E(c \log x)$, we can use the chain rule. We have

$$\begin{aligned} D(x^c) &= E(c \log x)D(c \log x) \\ &= E(c \log x)c/x \\ &= x^c(c/x) = cx^{c-1}. \end{aligned}$$

The integration formula follows at once. □

<u>Theorem 6</u>. <u>Let</u> a <u>be</u> <u>a</u> <u>real</u> <u>constant</u>. $a > 0$. <u>Then</u>

$$D(a^x) = a^x \log a,$$

$$\int a^x \, dx = \frac{a^x}{\log a} + C \quad \text{if } a \neq 1.$$

The proof is left as an exercise

For other differentiation and integration formulas involving logarithms and exponentials, see 6.7 and 6.16 of Apostol.

Remark concerning common logarithms.

The logarithm function we have defined is sometimes called the "natural logarithm". A different version of the logarithm was once useful. It can be obtained as follows: Consider the function

$$f(x) = 10^x = E(x \log 10).$$

It is strictly increasing, since by the chain rule,

$$f'(x) = E(x \log 10) \cdot \log 10,$$

which is positive. Furthermore, since $x \log 10$ takes on all real values, $E(x \log 10) = f(x)$ takes on all positive values. The inverse of f is called the "common logarithm" or the "logarithm to the base 10", and denoted by $\log_{10} y$. That is, if $y > 0$, we define

$$\log_{10} y = x \text{ if and only if } y = 10^x.$$

This function was at one time useful for computational purposes, but it has long since fallen into oblivion.

A similar remark applies to obtain logarithms to other bases. If b is any positive number with $b \neq 1$, one defines

$$\log_b y = x \quad \text{if and only if } y = b^x.$$

Remark on motivation

Just as the sine and cosine functions arise most naturally as the fundamental solutions of the differential equation for simple harmonic motion, so the exponential function arises most naturally from consideration of the important differential equation

$$f'(x) = kf(x),$$

called the "equation of population growth (or decay)." If k is, for instance, the difference of the birth and death rate (per thousand, say, of a population) in a given time period, then this is the equation for the actual population (in thousands), as a function of time.

One checks at once that the function e^{kx} satisfies this equation. More generally, <u>every</u> solution of this equation can be expressed in terms of e^{kx}:

<u>Theorem 7</u>. Suppose f(x) is defined for all x and satisfies the equation

$$f'(x) = kf(x).$$

Let f(0) = a. Then

$$f(x) = ae^{kx}.$$

<u>Proof</u>. Let us set

$$g(x) = f(x)E(-kx).$$

Then we compute

$$g'(x) = f'(x)E(-kx) - kf(x)E'(-kx) = 0,$$

so g is constant. Since $g(0) = f(0)\cdot E(0) = a$,

$$g(x) = a.$$

Multiplying by E(kx), we have f(x) = aE(kx), as desired. □

Exercises

1. If a is constant, show that in general

$$D(a^x) \neq xa^{x-1}, \text{ and}$$

$$\int a^x\,dx \neq \frac{a^{x+1}}{x+1}.$$

[Anyone who makes the mistake, on a quiz, of thinking these are equalities gets clobbered!]

2. Evaluate $\int_0^1 \pi^x\,dx$ and $\int_0^1 x^\pi\,dx$.

Integration

The substitution rule

Apostol proves only one version of the substitution rule, the one given in Theorem 1 following. Sometimes the converse is needed; we prove this result in Theorem 2.

Theorem 1. Assume that f(u) and g(x) and g′(x) are continuous, and that f(g(x)) is defined for all x in the domain of g. If

$$\int f(u)du = P(u) + C, \text{ then}$$

$$\int f(g(x))g'(x)dx = P(g(x)) + C.$$

Proof. We are given that $P'(u) = f(u)$. The chain rule implies that the derivative of P(g(x)) equals

$$P'(g(x))g'(x) = f(g(x))g'(x).$$

This is just the desired result. □

Theorem 2. (Partial converse) Assume that f(u) and g(x) and g′(x) are continuous, and that f(g(x)) is defined for all x in the domain of g. Assume also that $u = g(x)$ has the differentiable inverse function $x = h(u)$. If

(*) $$\int f(g(x))g'(x)dx = Q(x) + C, \quad \text{then}$$

(**) $$\int f(u)du = Q(h(u)) + C.$$

Proof. Applying Theorem 1, we can substitute $x = h(u)$ in the given formula (*) to obtain the equation

$$\int \left[f(g(h(u)))g'(h(u)\right]h'(u)du = Q(h(u)) + C.$$

Because h is the inverse function to g, we know that $g(h(u)) = u$ and

$$h'(u) = 1/g'(h(u)).$$

This formula thus takes the form

$$\int f(u)du = Q(h(u)) + C.$$

which is the equation we wished to prove. □

Example 1. The usual application of the substitution rule uses Theorem 1. One begins with the given integrand, and tries to write it in the form $f(g(x))g'(x)$ for some suitable function f and g, where f is a function we know how to integrate. For example, suppose we wish to compute the integral

$$\int x^2 \cos(x^3)dx.$$

We see this is almost of the form $\int \cos u\, du$ if we set $u = x^2$. That is, we "group x^2 with dx and supply a factor of 3", writing the integral in the form

$$\frac{1}{3}\int \cos (x^3)[3x^2dx].$$

Then because we know that $\int \cos u\, du = \sin u + C$, we conclude from Theorem 1 that our given integral equals

$$\frac{1}{3} \sin(x^3) + C.$$

Example 2. On the other hand, sometimes Theorem 2 is the one that is useful. It often applies when there is nothing obvious to "group with the dx" to simplify the integrand. Trigonometric substitutions are of this type.

For example, consider the integral $\int f(u)du$, where

$$f(u) = 1/\sqrt{1+u^2}.$$

This is not something we know how to integrate. However, the substitution $u = \tan x$ will simplify the expression $\sqrt{1+u^2}$ at least. It is an acceptable substitution, since it has the differentiable inverse function $x = \arctan u$

Using this substitution, we have

$$1 + u^2 = 1 + \tan^2 x = \sec^2 x.$$

Then

$$\sqrt{1+u^2} = \sec x;$$

the sign is +, because x lies between $-\pi/2$ and $+\pi/2$, so sec x = 1/cos x is positive. And of course we have

$$1 + u^2 = \sec^2 x.$$

Hence the integral

$$\int (1/\sqrt{1+u^2})du$$

takes the form

$$\int (1/\sec x)\sec^2 x \, dx$$

which we know how to integrate. Indeed,

$$\int \sec x \, dx = \log|\sec x + \tan x| + C.$$

Then we can apply Theorem 2 to conclude that

$$\int \left[1/\sqrt{1+u^2}\right] du = \log|\sec(\arctan u) + \tan(\arctan u)| + C.$$

This answer can be written more simply. For if x = arctan u, then sec x = $\sqrt{1+u^2}$, as noted earlier, and tan x = u. Hence we have the formula

$$\int \left[1/\sqrt{1+u^2}\right] du = \log|\sqrt{1+u^2} + u| + C.$$

A strategy for integration.

Step 1. Determine whether you can simplify the integrand easily, using algebraic manipulations (such as completing the square), trig identities, or a simplifying substitution (especially for the "inside function" in $f(g(x))$.)

Step 2. Examine the form of the integrand to determine the appropriate method.

(a) A product of two dissimilar functions suggests integration by parts. [Examples: $x^2 \sin x$, xe^x.] The same holds for a function whose derivative is a more familiar function than the function itself. [Examples: $\log x$, $\arctan x$.]

(b) Rational functions of x can always be integrated by the method of partial fractions.

(c) Powers of trig functions can be integrated using the half-angle formulas, various substitutions or (if necessary) reduction formulas.

(d) For integrands involving $\sqrt{a^2-x^2}$, $\sqrt{a^2+x^2}$, $\sqrt{x^2-a^2}$, a trig substitution is often helpful.

(e) [Optional: Any rational function of $\sin x$ and $\cos x$ can be reduced to a rational function of u by means of the substitution $u = \tan(x/2)$.]

Exercises

Evaluate the following:

1. $\int \cos^2 x \, dx.$

(Use either a half-angle formula derived from the identity

$$\cos 2x = 1 - 2\sin^2 x = 2\cos^2 x - 1$$

or the reduction formula of p. 221.)

2. $\int_0^{1/2} \sqrt{1-\mu^2}\, d\mu.$

3. Integrate $\int \frac{d\mu}{1-\mu^2}$ by expressing $\frac{1}{1-\mu^2}$ in the form $\frac{a}{1-\mu} + \frac{b}{1+\mu}$. Use this formula to evaluate

$$\int \sec x \, dx = \int \frac{\cos x \, dx}{\cos^2 x} = \int \frac{\cos x \, dx}{1-\sin^2 x}.$$

4. $\int \frac{d\mu}{\sqrt{\mu^2-1}}.$

5. $\int_0^1 \frac{d\mu}{(\sqrt{\mu^2+1})^3}.$

6. Compute

$$\int_0^1 x f''(2x)\, dx,$$

given that f'' is continuous for all x, and

$f(0) = 1$, $f'(0) = 3$.
$f(1) = 5$, $f'(1) = 2$.
$f(2) = 7$, $f'(2) = 4$.

7. Evaluate $\int \frac{x^4 + 2}{x^4 + x^3 + x^2}\, dx$ completely, including the constants A, B, ... in the partial fraction decomposition.

Chapter 7 <u>Taylor's formula</u>.

If $f(x)$ has derivatives of orders $1,\ldots,n$ at the point $x = a$, then the polynomial function

$$T_n(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n,$$

where

$$a_m = \frac{f^{(m)}(a)}{m!}$$

for each m, is called the <u>n^{th} order Taylor approximation to f at</u> a. It has the crucial property that it agrees with f at a, and that also its first n derivatives agree with those of f at a. (See 7.1 and 7.2 of Apostol.)

In order to use this approximation for either practical or theoretical purposes, we need to find a way to obtain an upper bound on the difference between f and its Taylor approximation. Let us define $E_n(x) = f(x) - T_n(x)$. Then we have the equation

$$f(x) = T_n(x) + E_n(x),$$

which is called <u>Taylor's formula</u>. Here $E_n(x)$ is called the <u>remainder term</u>, or <u>error term</u>, in Taylor's formula.

There are a number of different formulas involving $E_n(x)$; all involve $(x-a)$ and $f^{(n+1)}(x)$ in some way. (See 7.7 of Apostol, where four different ways of expressing $E_n(x)$ are given.) Each of these expressions is useful in situations where the others are not much help; the study of such formulas leads to a general subject called Numerical Analysis. Here we are going to derive and use just one of these expressions, the one called the Lagrange form of the remainder.

First, we need a lemma.

Lemma 1. Suppose $h^{(n+1)}(x)$ exists on an open interval. I. (This implies that $h(x), h'(x), \ldots, h^{(n)}(x)$ exist and are continuous on this interval.) Let a and b be points of this interval; suppose that

$$h(a) = h'(a) = \cdots = h^{(n)}(a) = 0,$$

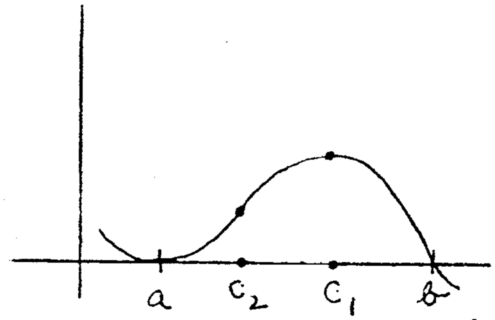
and

$$h(b) = 0.$$

Then there is some point c between a and b for which

$$h^{(n+1)}(c) = 0.$$

Proof. We suppose for convenience that $a < b$. (The same proof works if $b < a$). Because $h(a) = h(b) = 0$, the mean value theorem tells us there is some point c_1 with $a < c_1 < b$ such that $h'(c_1) = 0$. Now because $h'(a) = h'(c_1) = 0$, the mean-value theorem, applied to $h'(x)$, tells us there is some point c_2 with $a < c_2 < c_1$ such that $h''(c_2) = 0$.



Similarly continue. At the n^{th} stage, we have a point $c_n > a$ such that $h^{(n)}(c_n) = 0$. Since $h^{(n)}(a) = 0$, we can apply the mean-value theorem to $h^{(n)}(x)$ to find a point c with $a < c < c_n$ such that $h^{(n+1)}(c) = 0$. □

Theorem 2. Let $f^{(n+1)}(x)$ exist in an open interval I. Let a be a point of I. Let $T_n(x)$ be the n^{th} order Taylor approximation to f at a; let $E_n(x) = f(x) - T_n(x)$. Then given x in I, there is a point c between a and x such that

$$E_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

This expression is called the Lagrange form of the remainder.

Proof. Let b be a fixed point of the interval I different from a. We show there is a point c between a and b such that

$$E_n(b) = \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}.$$

Since b is arbitrary, this will suffice.

Consider the function $E_n(x) = f(x) - T_n(x)$. Because the functions f and T_n and their first n derivatives agree at a, we have

$$(*) \qquad E_n(a) = E_n'(a) = \cdots = E_n^{(n)}(a) = 0.$$

Thus part of the hypothesis of the preceding lemma is satisfied. Of course, $E_n(b)$ is not zero, in general. We seek to modify E_n, by adding a suitably chosen function, so as to get a function $h(x)$ that does vanish at b, but which has the same derivatives as E_n does at a (so that condition (*) still holds). A multiple of $(x-a)^{n+1}$ will do the job, for the first n derivatives of $(x-a)^{n+1}$ all vanish at a. The $(n+1)^{\underline{st}}$ derivative of $(x-a)^{n+1}$ does not vanish at a, of course. In fact, the $(n+1)^{\underline{st}}$ derivative is just the constant $(n+1)!$.

So let us define

$$h(x) = E_n(x) - A(x-a)^{n+1},$$

where we choose the constant A so that $h(b) = 0$. That is, we let $A = E_n(b)/(b-a)^{n+1}$. The hypotheses of the preceding lemma are then satisfied. We conclude there is some point c between a and b such that $h^{(n+1)}(c) = 0$.

Now we compute $h^{(n+1)}(x)$. Recall that

$$h(x) = E_n(x) - A(x-a)^{n+1} = f(x) - T_n(x) - A(x-a)^{n+1}.$$

Because $T_n(x)$ is a polynomial of degree n, its $(n+1)^{\underline{st}}$ derivative vanishes. Therefore we have

$$h^{(n+1)}(x) = f^{(n+1)}(x) - 0 - (n+1)!A$$

$$= f^{(n+1)}(x) - (n+1)!E_n(b)/(b-a)^{n+1}.$$

We substitute c for x, and recall that $h^{(n+1)}(c) = 0$. Solving for $E_n(b)$, we have

$$E_n(b) = \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1},$$

as desired. □

Theorem 3. *If the* $(n+1)^{\text{st}}$ *derivative of* f *satisfies the inequalities*

$$m \leq f^{(n+1)}(x) \leq M$$

for all x *in some interval about* a, *then for all* x *in this interval, we have*

$$\frac{m(x-a)^{n+1}}{(n+1)!} \leq E_n(x) \leq M\frac{(x-a)^{n+1}}{(n+1)!}$$

if $(x-a)^{n+1}$ *is positive; otherwise, the reverse inequalities hold.*

Proof. We apply the preceding theorem to calculate $E_n(x)$. Multiplying through

$$m \leq f^{(n+1)}(c) \leq M$$

by the number $(x-a)^{n+1}/(n+1)!$ gives us the desired inequalities. □

Application: Taylor's formula applied to indeterminate forms.

We can calculate limits for most familiar functions by using our basic theorems on limits, along with the continuity properties of the elementary functions. One situation where these theorems fail us is when we consider a limit of a quotient,

$$\lim_{x \to a} \frac{f(x)}{g(x)},$$

where the denominator approaches 0 as $x \longrightarrow a$. In this case, anything can happen. If $f(x)$ approaches a limit A different from zero and $g(x)$ approaches 0, then the limit of $f(x)/g(x)$ does not exist. If, however, $f(x)$ and $g(x)$ both approach zero, then the quotient may approach a finite limit. For example:

$$\lim_{x \to 2} \frac{x^2-4}{x-2} = \lim_{x \to 2} (x+2) = 4$$

$$\lim_{x \to 2} \frac{(x-2)^2}{x^2-4} = \lim_{x \to 2} \frac{x-2}{x+2} = 0$$

$$\lim_{x \to 2} \frac{x^2-4}{(x-2)^2} = \lim_{x \to 2} \frac{x+2}{x-2}, \quad \text{which does not exist.}$$

Taylor's formula can sometimes be of help in computing limits, if one knows the Taylor polynomials of the functions involved. In general, if $f^{(n+1)}$ is contiruous on an interval containing a, we have the formula

$$f(x) = f(a) + f'(a)(x-a) + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + B(x)(x-a)^{n+1}$$

where $B(x) = f^{(n+1)}(c)/(n+1)!$. We do not know exactly what $B(x)$ is, but we do know it is <u>bounded</u> on the interval in question, because $f^{(n+1)}$ is continuous. We use the letter B to remind us it is bounded.

We have for example the formulas

$$e^x = 1 + x + x^2/2! + \ldots + x^n/n! + B(x)x^{n+1},$$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^{n-1}\frac{x^n}{n} + B(x)x^{n+1},$$

and so on for other elementary functions. Here is how these formulas can be used in calculating limits.

<u>Example 1</u>. Calculate the limit as $x \to 0$ of $(\sin x)/x$. Now

$$\sin x = x - x^3/3! + Bx^5, \quad \text{so}$$

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + Bx^4,$$

which approaches 1 as x approaches 0. (Since B is a bounded function of x, we have $Bx^4 \longrightarrow 0$ as $x \longrightarrow 0$.)

Example 2. Calculate $\lim(\cos x-1)/x \sin x$ as $x \longrightarrow 0$.

$$\cos x = 1 - x^2/2! + Bx^4$$

$$\sin x = x - x^3/3! + Cx^5$$

$$\frac{\cos x - 1}{x \sin x} = \frac{-\ x^2/2! + Bx^4}{x^2 - x^4/3! + Cx^6} = \frac{-\ 1/2 + Bx^2}{1 - x^2/3! + Cx^4},$$

which approaches $-1/2$ as $x \longrightarrow 0$.

Example 3. Calculate $\lim(x \cot x - 1)/x^2$ as $x \longrightarrow 0$.

$$\cot x = \frac{\cos x}{\sin x} = \frac{1 - x^2/2! + Bx^4}{x - x^3/3! + Cx^5}.$$

$$\frac{x \cot x - 1}{x^2} = \frac{(x-x^3/2!+Bx^5) - (x-x^3/3!+Cx^5)}{x^2(x-x^3/3!+Cx^5)}$$

$$= \frac{-\ x^3/3 + (B-C)x^5}{x^3 - x^5/3! + Cx^7}$$

$$= \frac{-\ 1/3 + (B-C)x^2}{1 - x^2/3! + Cx^4}$$

which approaches $-1/3$ as $x \longrightarrow 0$.

Example 4. Calculate $\lim(\cos(ax) - \cos x)/x^2$ as $x \longrightarrow 0$.

$$\cos(ax) = 1 - (ax)^2/2 + B(ax)(ax)^4.$$

$$\cos(ax) - \cos x = -(ax)^2/2 + x^2/2 + B(ax)a^4x^4 - B(x)x^4.$$

Hence $(\cos(ax) - \cos x)/x^2 \longrightarrow (1-a^2)/2$ as $x \longrightarrow 0$, since both $B(ax)$ and $B(x)$ are bounded.

Example 5. Calculate the limit of $(\log(1+ax))/x$ as $x \longrightarrow 0$.

$$\log(1+ax) = (ax) - (ax)^2/2 + B(ax)(ax)^3;$$

so $(\log(1+ax))/x \longrightarrow a$ as $x \longrightarrow 0$.

Example 6. Show that $(1+ax)^{1/x} \longrightarrow e^a$ as $x \longrightarrow 0$. Now $(1+ax)^{1/x} = E((1/x)\log(1+ax))$. By continuity of the exponential function, it suffices to show that $(1/x)\log(1+ax) \longrightarrow a$. This is done in Example 5.

Exercises

Throughout, we use Taylor's formula with the Lagrange form of the remainder: $f(x) = T_n(x) + E_n(x)$, where

$$E_n(x) = \frac{f^{n+1}(c)}{(n+1)!}(x-a)^{n+1}.$$

1. Derive the following from Taylor's theorem:

(a) Given x, there is a c between 0 and x such that

$$e^x = \left[1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}\right] + \frac{e^c}{(n+1)!}x^{n+1}.$$

(b) Given $x > -1$, there is a c between 0 and x such that

$$\log(1+x) = \left[x - \frac{x^2}{2} + \frac{x^3}{3} + \dots + (-1)^{n-1}\frac{x^n}{n}\right] + \frac{(-1)^n}{(1+c)^{n+1}}\frac{x^{n+1}}{n+1}.$$

(c) Given x, there is a c between 0 and x such that

$$\sin x = \left[x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^n\frac{x^{2n+1}}{(2n+1)!}\right] + (-1)^{n+1}\frac{\cos c}{(2n+3)!}x^{2n+3},$$

$$\cos x = \left[1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n\frac{x^{2n}}{(2n)!}\right] + (-1)^{n+1}\frac{\cos c}{(2n+2)!}x^{2n+2}.$$

<u>Learn the Taylor polynomials for</u> e^x, $\log(1+x)$, $\sin x$ <u>and</u> $\cos x$.

2. (a) Use the third order Taylor approximation for e^x near $x = 0$ and the fact that $e < 4$ to show that

$$\frac{8}{3} + \frac{1}{24} < e < \frac{8}{3} + \frac{1}{6},$$

whence it follows that $2.7 < e < 2.9$.

(b) Now use the same Taylor approximation and the fact that $e < 3$ to show that

$$\frac{8}{3} + \frac{1}{24} < e < \frac{8}{3} + \frac{1}{8},$$

whence $2.7 < e < 2.8$.

3. Use the first and second order Taylor approximations to $\sqrt{x}$ near $a = 4$ to compute $\sqrt{3.8}$. (Actual value is 1.949359...)

4. Use the third order Taylor approximation to $\sin x$ near $a = 0$ to compute $\sin(1/2)$. Obtain an upper bound on the error.

5. What order Taylor polynomial should one use if one wishes to compute e to two decimal places of accuracy (i.e., with an error less than .005)? Use the fact that $e < 3$. Obtain an upper bound for the error. What about computing log 2?

6. (a) Show that the inequalities

$$x - x^3/3! \leq \sin x \leq x - x^3/3! + x^5/5!$$

hold for $0 \leq x \leq \pi/2$. [Hint: Consider the sign of the error term.]

(b) Use these inequalities and your trusty pocket calculator to show that $\sin .523 < 1/2$ and $\sin .524 > 1/2$. (Be sure you allow for round-off error.)

(c) Since $\sin \pi/6 = 1/2$, conclude that

$$3.138 < \pi < 3.144.$$

These inequalities give the approximation $\pi \sim 3.14$ with two decimal places of accuracy.

7. Find the limit as $x \longrightarrow 0$ of the function

$$\frac{\sin x - xe^{(x^2)} + 7x^3/6}{(\sin^2 x)(\sin x^3)} .$$

8. For what range of values of x can you replace $\cos x$ by $1 - x^2/2 + x^4/24$ with an error no greater than 5×10^{-4}?

9. The approximation $(1+x)^{1/3} \sim 1 + x/3$ is often used when $|x|$ is small. Find an upper bound for the error if $0 < x < .01$.

10. (a) Show that if $|x| < 1$, then

$$|e^x - (1+x+x^2/2)| < |x^3/2|.$$

(b) Show that if $|t| < 1$, then the approximation

$$\int_0^t e^{x^2}\, dx \sim t + t^3/3 + t^5/10$$

involves an error in absolute value no more than $|t^7/14|$.

L'Hopital's rule for 0/0

Theorem. Suppose $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$. If

$$\frac{f'(x)}{g'(x)} \to L \quad \text{as} \quad x \to a,$$

then also $f(x)/g(x) \to L$ as $x \to a$.

This result holds whether a and L are finite or infinite, and it also holds if the limits are one-sided.

Proof. The proof when a is finite is that given on p. 295 of the text. The crucial step is to use Cauchy's mean-value theorem to prove that $g(x) \neq 0$ for x near a, and that

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}$$

for some c between a and x. It follows that if $f'(x)/g'(x)$ approaches L as $x \to a$, then $f(x)/g(x)$ must approach L also. In the text, it is assumed that L is finite. But it does really not matter whether L is finite or infinite; precisely the same proof applies.

The proof in the case $a = +\infty$ is given on p. 298 of the text. Again, it is assumed that L is finite, but that doesn't matter; if L is $\pm\infty$ precisely the same proof applies. □

Remark. L'Hopital's rule also works if $f(x)$ and $g(x)$ both approach ∞ instead of 0. But the proof is more complicated. We shall give a proof shortly. The only cases of interest to us concern the logarithm and the exponential. For these functions, a direct proof is given on p. 301 of the text. Alternatively, they may be treated by using L'Hopital's rule for the case ∞/∞, as we shall see.

The behavior of log and exp that we are concerned with is stated in the following theorem:

<u>Theorem</u>. <u>As</u> $x \to +\infty$, <u>both</u> $\log x$ <u>and</u> e^x <u>approach</u> $+\infty$. <u>But</u> $\log x$ <u>approaches</u> ∞ <u>more slowly than any positive power of</u> x, <u>and</u> e^x <u>approaches</u> ∞ <u>more rapidly than any positive power of</u> x; <u>the same holds for any positive powers of</u> $\log x$ <u>and</u> e^x. <u>More precisely, if</u> a <u>and</u> b <u>are positive real numbers, then</u>

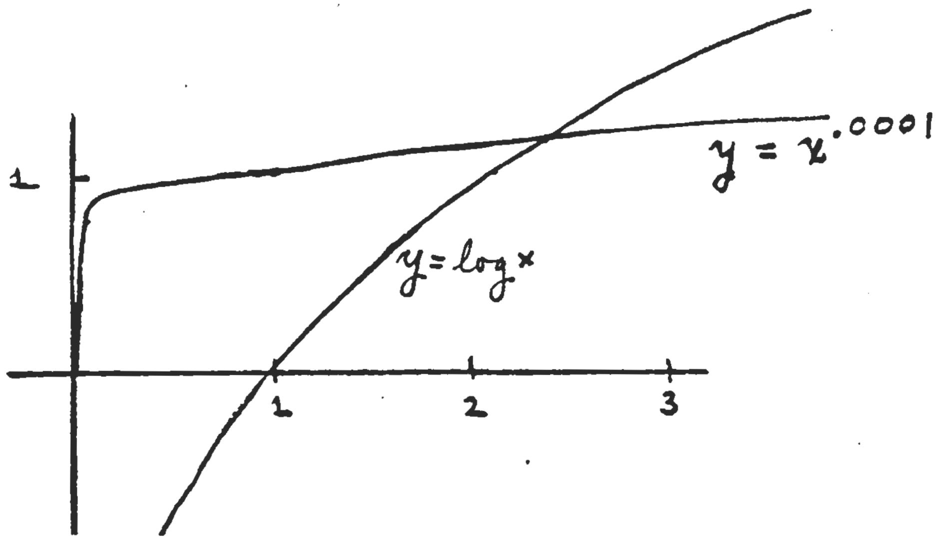
$$\lim_{x \to +\infty} \frac{(\log x)^b}{x^a} = 0 \quad \text{and} \quad \lim_{x \to +\infty} \frac{(e^x)^b}{x^a} = +\infty$$

<u>Corollary.</u> <u>The function</u> $\log x$ <u>goes to</u> $-\infty$ <u>very slowly as</u> x <u>goes to</u> 0. <u>More precisely, if</u> a <u>is a positive real number, then</u>

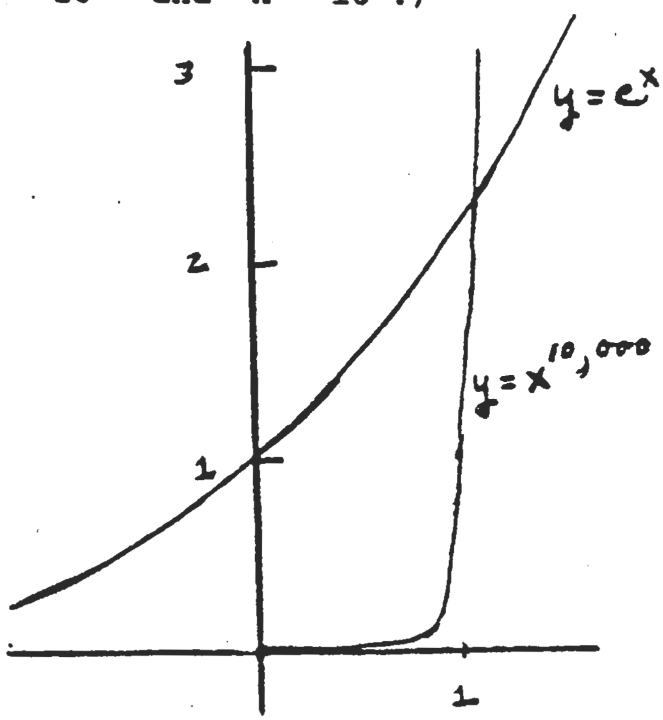
$$\lim_{x \to 0+} x^a \log x = 0 .$$

What does this theorem mean? Note that for any function $f(x)$ that goes to ∞ as x goes to ∞, a positive power of $f(x)$, say $[f(x)]^a$, goes to ∞ even more rapidly if the power a is large, and to goes to ∞ more slowly if a is small. This theorem says that no matter how high a power b you raise $\log x$ to, and how small a power a you raise x to, the power of $\log x$ will still go to ∞ more slowly than the power of x. Similarly, any power of e^x, no matter how small, will go to ∞ faster than any power of x, no matter how large.

For example, even though for small values of x, the graphs of the functions $\log x$ and $x^{.0001}$ appear as in the accompanying figure, it is still true that eventually the function $f(x) = x^{.0001}$ becomes much larger than $\log x$.



Similar graphs for the functions $x^{10,000}$ and e^x can be obtained by exchanging the axes in this figure. Although $x^{10,000}$ shoots up very rapidly to begin with, eventually e^x becomes much larger than $x^{10,000}$. (In fact, these curves cross again between $x = 10^5$ and $x = 10^6$.)



L'Hopital's rule for ∞/∞.

Theorem. Suppose $f(x) \to \infty$ and $g(x) \to \infty$ as $x \to a$. If

$$\frac{f'(x)}{g'(x)} \longrightarrow L \text{ as } x \longrightarrow a,$$

then also $f(x)/g(x) \longrightarrow L$ as $x \longrightarrow a$.

This result holds whether a and L are finite or infinite, and it also holds if the limits are one-sided.

Proof. Case 1. We prove the theorem first in the case where a is finite and $x \longrightarrow a+$.

The hypotheses of the theorem imply that f and g are defined and positive on some interval of the form (a,b], and that f′ and g′ exist and $g' \neq 0$ on some such interval.

Let x_0 be a fixed point of this interval. (We shall specify how to choose x_0 later.) Then let x be a point of this interval that is very close to a. Just how close will be determined later. For now we merely require that $a < x < x_0$ and that $f(x) > f(x_0)$ and $g(x) > g(x_0)$. (Since f and g go to ∞ as $x \to a+$, these inequalities hold if x is close enough to a.) Then we compute.

Let us apply the Cauchy mean-value theorem to the interval $[x,x_0]$. We conclude that there is a c with $x < c < x_0$ such that

$$f'(c)[g(x_0)-g(x)] = g'(c)[f(x_0)-f(x)]$$

or

$$f'(c)g(x)[g(x_0)/g(x)-1] = g'(c)f(x)[f(x_0)/f(x)-1]$$

or

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}\left[\frac{(g(x_0)/g(x)) - 1}{(f(x_0)/f(x)) - 1}\right].$$

For convenience, let $\lambda(x)$ denote the expression in brackets; then

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}\lambda(x).$$

Note that $\lambda(x) \to 1$ as $x \to a+$.

Now we verify the theorem in the case where L is finite. By choosing x_0 close to a, we can ensure that $f'(c)/g'(c)$ is close to L (since $a < c < x_0$); then we can make $\lambda(x)$ close to 1 by requiring that x be very close to a. Then $f(x)/g(x)$ will be close to L. The only question is: how close is "close enough"? Let us set

$$\epsilon_1 = |(f'(c)/g'(c)) - L| \quad \text{and} \quad \epsilon_2 = |\lambda(x) - 1|.$$

Then

$$(*) \qquad \left|\frac{f'(c)}{g'(c)}\cdot\lambda(x) - L\right| = |(L\pm\epsilon_1)(1\pm\epsilon_2) - L| \le |\epsilon_1| + |L\epsilon_2| + |\epsilon_1\epsilon_2|.$$

This inequality tells us how to proceed. Suppose $0 < \epsilon < 1$. First, we choose x_0 so that for all c with $a < c < x_0$, we have $\epsilon_1 < \epsilon/3$. Now x_0 is fixed. Then choose $\delta > 0$ so that for $a < x < a+\delta$, we have $g(x) > g(x_0)$ and $f(x) > f(x_0)$ and

$$|\lambda(x) - 1| = \epsilon_2 < \epsilon/3(1+|L|).$$

Then for $a < x < a+\delta$, inequality (*) tells us that

$$\left|\frac{f(x)}{g(x)} - L\right| = \left|\frac{f(c')}{g(c')}\lambda(x) - L\right| \le \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon^2}{9} < \epsilon, \text{ as desired.}$$

Finally, we consider the case where L is infinite. Given $M > 0$, we want to show that $f(x)/g(x) > M$ for x close to a. First, choose x_0 so that for all c with $a < c < x_0$, we have $f'(c)/g'(c) > 2M$. Then choose δ so that for $a < x < a+\delta$, we have $g(x) > g(x_0)$ and $f(x) > f(x_0)$ and $\lambda(x) > 1/2$. It follows that, for $a < x < a+\delta$,

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}\lambda(x) > 2M\cdot\frac{1}{2} = M.$$

We have now proved the rule in the case $x \to a+$. The case $x \to a-$ follows readily, as we now show. Note that as x approaches a from the left, $u = a-x$ approaches 0 from the right. Then

$$\begin{aligned}\lim_{x\to a-} (f(x)/g(x)) &= \lim_{u\to 0+} f(a-u)/g(a-u) \\ &= \lim_{u\to 0+} (-1)f'(a-u)/(-1)g'(a-u) \\ &= \lim_{x\to a-} f'(x)/g'(x),\end{aligned}$$

if the latter limit exists.

The case $x \to a$, with a finite, follows from the two cases $x \to a+$ and $x \to a-$.

Finally, the case $x \to \infty$ follows from the computation

$$\begin{aligned}\lim_{x\to\infty} f(x)/g(x) &= \lim_{t\to 0+} f(1/t)/g(1/t) \\ &= \lim_{t\to 0+} (-1/t^2)f'(1/t)/(-1/t^2)g'(1/t) \\ &= \lim_{x\to\infty} f'(x)/g'(x),\end{aligned}$$

if the latter limit exists. □

The behavior of log and exp

We now derive the theorem on p. P.2 from L'Hopital's rule. Consider first the log function. Given $c > 0$, we compute

$$\begin{aligned}\lim_{x\to\infty}(\log x)/x^c &= \lim_{x\to\infty} x^{-1}/cx^{c-1} \quad \text{by L'Hopital's rule} \\ &= \lim_{x\to\infty} 1/cx^c = 0.\end{aligned}$$

Then we set $c = b/a$ and compute

$$\lim_{x\to\infty} (\log x)^a/x^b = \lim_{x\to\infty} [\log x/x^c]^a = 0,$$

as desired.

Now we consider the exp function. Given $c > 0$, we compute

$$\begin{aligned}\lim_{x\to\infty} e^{cx}/x &= \lim_{x\to\infty} ce^{cx}/1 \quad \text{by L'Hopital's rule} \\ &= \infty.\end{aligned}$$

Then we set $c = a/b$ and compute

$$\lim_{x\to\infty} (e^x)^a/x^b = \lim_{x\to\infty}[e^{cx}/x]^b = \infty,$$

as desired.

Finally, we note that

$$\lim_{x\to 0+} x^a \log x = \lim_{t\to\infty} (1/t^a)\log(1/t)$$
$$= \lim_{t\to\infty} \frac{-\log t}{t^a} = 0,$$

as desired□

Example. Although

$$\lim_{x\to\infty} \frac{x+\sin x}{x}$$

assumes the indeterminate form ∞/∞, L'Hopital's rule does not apply, since the function $(1+\cos x)/1$ oscillates rather than approaches a limit as $x \to \infty$. However,

$$\frac{x+\sin x}{x} = 1 + \frac{\sin x}{x},$$

which approaches 1 because $|\sin x|/x \leq 1/x$ for $x > 0$.

This example shows that the converse of L'Hopital's rule is not true. For this is a case where $f(x)\to\infty$ and $g(x)\to\infty$ as $x\to\infty$, and $f(x)/g(x)$ approaches a limit, even though $f'(x)/g'(x)$ does not approach a limit.

10.13 Notes on error estimates.

Each of the standard convergence tools for series brings with it a method for estimating the error made in approximating the limit by taking only finitely many terms of the series. We treat this method for both the integral test and the ratio test.

Theorem 1. (Integral estimate.) Let f be a positive decreasing function defined for $x \geq 1$. Let $a_k = f(k)$. If the integral $\int_1^\infty f$ exists, then the series $\Sigma\, a_k$ converges, and

$$\int_{N+1}^{\infty} f(x)dx \leq \Big(\sum_{k=N+1}^{\infty} a_k\Big) \leq \int_{N}^{\infty} f(x)dx.$$

The expression in the middle is of course the error made in using the finite sum $a_1 + \ldots + a_N$ as an approximation to the sum of the series.

Theorem 2. (Ratio estimate.) Suppose that $\Sigma\, a_k$ is a series of non-zero terms. Suppose α is a number less than 1, and that

$$|a_{k+1}/a_k| \leq \alpha \quad \text{for all} \quad k > N.$$

Then

$$\left|\sum_{k=N+1}^{\infty} a_k\right| \leq \left|a_{N+1}\right| \left(\frac{1}{1-\alpha}\right).$$

The expression on the left is again the error made in approximating the series $\Sigma\, a_k$ by the finite sum $a_1 + \ldots + a_N$.

Exercises

1. Prove Theorem 1.

2. (a) Use Theorem 1 to estimate the error made in approximating the number

$$a = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

by the finite sum $1 + 1/4 + 1/9 + 1/16 + 1/25$. (In fact, $a = \pi^2/6$.)

(b) Given N, estimate the error made by using

$$\sum_{n=1}^{N} \frac{1}{n^2}$$

as an approximation to a.

(c) Estimate the error made by using

$$\sum_{n=1}^{N} \frac{1}{n^2} + \frac{1}{N+1}$$

as an approximation to a.

(d) Estimate the error made in approximating a by the sum $1 + 1/4 + 1/9 + 1/16 + 1/25 + 1/6$.

3. Prove Theorem 2.

4. Estimate the error made in approximating the number $\Sigma_{n=1}^{\infty} n/2^n$ by the finite sum $1/2 + 2/4 + 3/8 + 4/16 + 5/32$.

The basic theorems on power series.

Whenever we have a series $\sum \mu_n(x)$ of functions, there are three fundamental questions we ask:

(1) Given the series $\sum \mu_n(x)$, for what values of x does the series converge?

(2) Given $\sum \mu_n(x)$, if it converges to a function $f(x)$, what properties does $f(x)$ have? Specifically: Is f continuous? Can you calculate $\int_a^b f(x)$ by integrating the series term-by-term? Is f differentiable, and can you calculate $f'(x)$ by differentiating the series term-by-term?

(3) Given a function $f(x)$, under what conditions does it equal such a series, where the functions $\mu_n(x)$ are functions of a particular type?

We shall answer these questions for a power series. This is a series of the form:

$$a_0 + a_1x + a_2x^2 + \cdots = \sum_{n=0}^{\infty} a_nx^n.$$

Theorem 1. Given a power series $\sum a_nx^n$, exactly one of the following holds:

(a) The series converges only for $x = 0$.

(b) The series converges absolutely for all x.

(c) There is a number $r > 0$ such that the series converges absolutely if $|x| < r$ and diverges if $|x| > r$. (Nothing is said about what happens when $x = \pm r$.)

Proof. Step 1. We show that if the series $\sum a_nx^n$ converges for $x = x_0 \neq 0$, then it converges (absolutely) for any x with $|x| < |x_0|$.

For this purpose, we write

$$|a_n x^n| = |a_n x_0^n| \; |x/x_0|^n = c_n |x/x_0|^n ,$$

where $c_n = |a_n x_0^n|$. Now the series $\sum |x/x_0|^n$ converges, because it is a geometric series of the form $\sum y^n$ with $|y| < 1$. Furthermore, the sequence c_n approaches 0 as $n \to \infty$, because the series $\sum a_n x_0^n$ converges (by hypothesis). We can choose N so that $|a_n x_0^n| \leq 1$ for $n \geq N$. Then $|a_n x^n| \leq |x/x_0|^n$ for $n \geq N$. The comparison thest then implies that the series $\sum |a_n x^n|$ converges.

Step 2. Let S be the set of all numbers x for which the series $\sum a_n x^n$ converges. If S consists of 0 alone, then (a) holds. Otherwise, there is at least one number x_0 different from 0 belonging to S. It then follows that there is a positive number x_1 belonging to S; indeed, if x_1 is any positive number such that $x_1 < |x_0|$, then $\sum |a_n x_1^n|$ converges by Step 1, so that $\sum a_n x_1^n$ converges and x_1 belongs to S. We now consider two cases.

Case 1. The set S is bounded above. In this case, we set $r = \sup S$, and show that the series $\sum a_n x^n$ diverges if $|x| > r$ and converges (absolutely) if $|x| < r$.

Divergence if $|x| > r$ is clear. For suppose $|x| > r$ and the series $\sum a_n x^n$ converges. If we choose x_2 so that $r < x_2 < |x|$, then Step 1 implies that the series $\sum |a_n x_2^n|$ converges. Then the series $\sum a_n x_2^n$ converges, so that x_2 belongs to S, contradicting the fact that r is an upper bound for S.

Convergence if $|x| < r$ is also clear. If $|x| < r$, we can choose an element x_3 of S such that $|x| < x_3$ (otherwise $|x|$ would be a smaller upper bound on S than r). Step 1 then implies that $\sum |a_n x^n|$ converges.

Case 2. The set S is unbounded above. We show the series $\sum a_n x_n$ converges (absolutely) for all x. Given x, choose an element x_4 of S such that $|x| < x_4$. This we can do because S is unbounded above. Then $\sum |a_n x^n|$ coverges, by Step 1. □

Definition. The number r constructed in (c) of the preceding theorem is called the radius of convergence of the series. In case (a) we say that $r = 0$; and in case (b), we say that $r = \infty$.

Theorem 2. *Suppose* $\sum a_n x^n$ *has radius of convergence* $r > 0$. (*We allow* $r = \infty$.) *Let*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n.$$

(a) f *is continuous for* $|x| < r$.

(b) *For* $|x| < r$, *we have*

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

(c) *For* $|x| < r$, *we have*

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

(d) *The series in* (b) *and* (c) *have radii of convergence precisely* r.

Proof. In general, let

$$p_m(x) = a_0 + a_1 x + \cdots + a_m x^m.$$

It is a polynomial of degree m, and it is the $m^{\underline{th}}$ partial sum of our power series.

We are going to prove parts (a) and (b) for the fixed point b in $(-r,r)$. So as a preliminary, let us choose R so that $|b| < R < r$.

Step 1. *Given* $\varepsilon > 0$, *there is an integer* N *such that the inequality*

$$|f(x) - p_m(x)| < \varepsilon$$

holds for all $m \geqslant N$ *and all* x *with* $|x| \leqslant R$.

The proof is easy. Since $\sum |a_n R^n|$ converges, we can choose N sufficiently large that

$$\sum_{n=N+1}^{\infty} |a_n R^n| < \varepsilon.$$

It follows that if $|x| \leqslant R$ and $m \geqslant N$, we have

$$\sum_{n=m+1}^{\infty} |a_n x^n| \leqslant \sum_{n=m+1}^{\infty} |a_n R^N| \leqslant \sum_{n=N+1}^{\infty} |a_n R^N| < \varepsilon.$$

Then for $m \geqslant N$ and $|x| \leqslant R$,

$$|f(x)-p_m(x)| = |\sum_{n=m+1}^{\infty} a_n x^n| \leqslant \sum_{n=m+1}^{\infty} |a_n x^n| < \varepsilon.$$

Step 2. We show that f is continuous at b. This proves (a).

Given $\varepsilon > 0$, choose N as in Step 1. We have

$$|f(x) - p_N(x)| < \varepsilon,$$

for any x in the interval $[-R,R]$. In particular,

$$|f(b) - p_N(b)| < \varepsilon.$$

Now we use continuity of the polynomial $p_N(x)$ to choose δ so that whenever $|x-b| < \delta$, then x is in $[-R,R]$, and

$$|p_N(x) - p_N(b)| < \varepsilon.$$

Adding these three inequalities and using the triangle inequality, we see that whenever $|x-b| < \delta$, we have

$$|f(x) - f(b)| < 3\varepsilon.$$

Step 3. We show that $\sum a_n b^{n+1}/(n+1)$ converges to $\int_0^b f(x)\,dx$. This proves (b).

Given $\varepsilon > 0$, choose N as in Step 1. Then whenever $m \geq N$, the inequality

$$-\varepsilon < f(x) - p_m(x) < \varepsilon$$

holds <u>for all</u> x in the interval $[-R,R]$. The comparison property of integrals tells us that

$$\left|\int_0^b (f(x)-p_m(x))\,dx\right| \leqslant \varepsilon|b|.$$

This says that

$$\left|\int_0^b f(x)\,dx - (a_0b+a_1b^2/2+\cdots+a_mb^{m+1}/(m+1))\right| \leqslant \varepsilon|b|$$

for all $m \geqslant N$. It follows that $\sum a_nb^{n+1}/(n+1)$ converges to $\int_0^b f(x)\,dx$.

<u>Step 4</u>. We show that the power series

$$\sum_{n=1}^{\infty} na_nx^{n-1}$$

has radius of convergence at least r.

For this purpose, it suffices to show that if c is any number with $0 < c < r$, then $\sum na_nc^{n-1}$ converges. In fact, it suffices to show that $\sum na_nc^n$ converges, since multiplying the series by c does not affect convergence. This is what we shall show.

First, choose d such that $c < d < r$. Then write the general term of our series in the form

$$na_nc^n = na_n(\tfrac{c}{d})^n\cdot d^n .$$

We note that the series $\sum a_nd^n$ converges because $d < r$. It follows that the n^{th} term a_nd^n approaches 0 as n becomes large. Choose N

sufficiently large that $|a_n d^n| < 1$ for $n \geq N$. Then for $n \geq N$, we have

$$na_n c^n \leq n(\frac{c}{d})^n .$$

Now the series

$$\sum n(\frac{c}{d})^n$$

converges by the ratio test, since $0 < c/d < 1$. Therefore the series $\sum na_n c^n$ converges, by the comparison test.

<u>Step 5</u>. We prove part (c). Let

$$g(x) = \sum_{n=1}^{\infty} na_n x^{n-1}$$

for $|x| < r$. Part (b) of the theorem tells us that for $|x| < r$, we have

$$\int_0^x g(t)\, dt = \sum_{n=1}^{\infty} a_n x^n.$$

$$= f(x) - a_0.$$

Part (a) of the theorem tells us that $g(x)$ is continuous for $|x| < r$. Then the first fundamental theorem of calculus applies; we conclude that

$$g(x) = f'(x),$$

which is what we wanted to prove.

<u>Step 6</u>. We prove part (d). If the series $\sum a_n x^{n+1}/(n+1)$ had radius of convergence $q > r$, then so would the differentiated series $\sum a_n x^n$, by Step 4. But it does not. Similarly, if the series $\sum na_n x^{n-1}$ had radius of convergence $q > r$, then so would the integrated series $\sum a_n x^n$. But it does not. □

<u>Remark</u>. It follows readily that all the results of Theorem 2 hold for general power series of the form

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n.$$

There is a number r (which may be 0 or ∞) such that the series converges absolutely for $|x-a| < r$ and diverges for $|x-a| > r$. Furthermore for $|x-a| < r$, one has:

(a) $f(x)$ is continuous.

(b) $\int_a^x f(t)\,dt = \sum_{n=0}^{\infty} a_n \frac{(x-a)^{n+1}}{n+1}$.

(c) $f'(x) = \sum_{n=1}^{\infty} na_n (x-a)^{n-1}$.

The proof is immediate; one merely substitutes $(x-a)$ for x in the theorem.

Here is a theorem proving the uniqueness of a power series representation:

<u>Theorem 3</u>. <u>Suppose</u>

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n = \sum_{n=0}^{\infty} b_n (x-a)^n$$

<u>on some open interval</u> I <u>containing</u> a. <u>Then for all</u> k,

$$a_k = b_k = \frac{f^{(k)}(a)}{k!}.$$

Proof. We apply the preceding theorem. We write

$$f(x) = \sum a_n(x-a)^n.$$

Differentiating, we have

$$f'(x) = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}.$$

Applying the theorem once again, we have

$$f''(x) = \sum_{n=2}^{\infty} n(n-1)a_n(x-a)^{n-2}.$$

And so on. Differentiating k times, we have

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)a_n x^{n-k}.$$

When we evaluate at $x = 0$, all the terms vanish except for the first term. Thus

$$f^{(k)}(x) = k!a_k,$$

as desired. The same argument applies to compute b_k. □

Definition. If $f(x)$ equals a power series $\sum a_n(x-a)^n$ in some open interval containing a, we say f is analytic (or sometimes "real analytic") near a. By the preceding theorem, this power series is uniquely determined by f; its partial sums must be the Taylor polynomials of f at a. For this reason, the series is sometimes called the Taylor series of f at a.

Corollary 4. The function $f(x)$ is analytic near a if and only if both the following hold:

(1) All derivatives of f exist in an open interval about a.

(2) The error term $E_n(x)$ in Taylor's formula approaches 0 as $n \longrightarrow \infty$, for each x in that interval.

Remark. We know that it is possible for us to have

$$f(x) = \sum_{n=0}^{\infty} \mu_n(x)$$

for all x in an interval $[c,d]$, where each function $\mu_n(x)$ is continuous, without it following that $f(x)$ is continuous, or that its integral can be obtained by integrating the series term-by-term. However, this unpleasant situation does not occur if the analogue of the statement in Step 1 of the proof of Theorem 2 holds. This fact leads to the following definition.

Definition. The series $\sum \mu_n(x)$ is said to converge uniformly to $f(x)$ on the interval $[c,d]$ if given $\varepsilon > 0$, there is an N such that

$$\left| f(x) - \sum_{n=0}^{m} \mu_n(x) \right| < \varepsilon$$

for all $m \geq N$ and all x in $[c,d]$.

Theorem 5. The series $\sum \mu_n(x)$ converges uniformly on $[c,d]$ if there is a convergent series $\sum M_n$ of constants such that $|\mu_n(x)| < M_n$ for all x in $[c,d]$.

(The proof is just like that of Step 1. There the series of constants was the series $\sum |a_n R^n|$.)

Under the hypothesis of uniform convergence, the analogues of (a) and (b) of Theorem 2 hold:

Theorem 6. Suppose $\sum \mu_n(x)$ converges uniformly to $f(x)$ on $[c,d]$. If the functions $\mu_n(x)$ are continuous, so is $f(x)$, and furthermore the series

$$\sum_{n=0}^{\infty} \left(\int_c^x \mu_n(t)\,dt\right)$$

converges uniformly to $\int_c^x f(t)\,dt$ on $[c,d]$.

The proof is just like the ones given in Steps 2 and 3.

Remark. Part (c) of the theorem, about differentiating a power series term-by-term, does not carry over to more general uniformly convergent series. For instance, the series

$$\sum_{n=1}^{\infty} (\sin nx)/n^2$$

converges uniformly on any interval, by comparison with the series of constants $\sum 1/n^2$, but the differentiated series $\sum (\cos nx)/n$ does not even converge at $x = 0$. If however the differentiated series does converge uniformly on $[c,d]$, then $f'(x)$ does exist and equals this differentiated series. The proof is similar to that of (c).

Exercises

1. Prove Theorem 3.

2. Prove Theorem 4.

3. Prove the following theorem about term-by-term differentiation:

Suppose that the functions $\mu_n'(x)$ are continuous, that the series $\sum_{n=1}^{\infty} \mu_n'(x)$ converges uniformly on $[c,d]$, and that $\sum_{n=1}^{\infty} \mu_n(x)$ converges for at least one x in $[c,d]$. Then:

(a) $\sum_{n=1}^{\infty} \mu_n(x)$ converges uniformly on $[c,d]$, say to $f(x)$.

(b) $f'(x)$ exists and equals $\sum \mu_n'(x)$.

[Hint: Integrate the series $\sum \mu_n'(x)$.]

A family of non-analytic functions.

Let $m \geq 0$ be any nonnegative integer. Define

$$f_m(x) = \begin{cases} \dfrac{e^{-1/x^2}}{x^m} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

We will show that each of the functions $f_m(x)$ has continuous derivatives of all orders, for all x. We also show that none of them is analytic near 0; that is, none of them equals a power series of the form $\Sigma\, a_n x^n$ in an interval about 0.

Theorem 1. (a) *The function $f_m(x)$ is continuous for all x.*

(b) *Furthermore, $f_m'(x)$ exists for all x and satisfies the equation*

$$f_m'(x) = -mf_{m+1}(x) + 2f_{m+3}(x).$$

Proof. (a) The general theorem about composites of continuous functions shows that $f_m(x)$ is continuous when $x \neq 0$. To prove continuity at $x = 0$, we must show that

$$\lim_{x\to 0} \frac{e^{-1/x^2}}{x^m} = 0.$$

The substitution $\mu = 1/x^2$ simplifies the calculation. We have

$$\lim_{x\to 0} \frac{e^{-1/x^2}}{x^m} = \lim_{\mu\to\infty} \frac{e^{-\mu}}{1/\mu^{m/2}} = \lim_{\mu\to\infty} \frac{\mu^{m/2}}{e^{\mu}}\ .$$

This limit is zero because e^{μ} approaches infinity faster than any power of μ, as $\mu \longrightarrow \infty$.

(b) We check differentiability. If $x \neq 0$, we calculate directly:

$$f_m'(x) = D\left(\frac{1}{x^m} e^{-1/x^2}\right) = \frac{-m}{x^{m+1}} e^{-1/x^2} + \frac{1}{x^m} e^{-1/x^2}\left(\frac{2}{x^3}\right)$$

$$= -mf_{m+1}(x) + 2f_{m+3}(x).$$

To show the derivative exists at $x = 0$, we apply the definition of the derivative:

$$f_m'(0) = \lim_{h\to 0} \frac{f_m(0+h) - f_m(0)}{h}$$

$$= \lim_{h\to 0} \frac{(e^{-1/h^2}/h^m) - 0}{h} = \lim_{h\to 0} \frac{e^{-1/h^2}}{h^{m+1}}$$

This limit is zero, by part (a). Therefore, the derivative exists at $x = 0$ and equals 0. Thus the formula

$$f_m'(x) = -mf_{m+1}(x) + 2f_{m+3}(x)$$

holds when $x = 0$.

Theorem 2. *The function $f_m(x)$ has continuous derivatives of all orders, for all x, but $f_m(x)$ does not equal a power series $\Sigma\, a_n x^n$ on any interval about 0.*

Proof. We know that each function $f_m(x)$ is differentiable, for all x. The equation

$$f_m'(x) = -mf_{m+1}(x) + 2f_{m+3}(x)$$

shows us that $f_m'(x)$ is differentiable, for each x. This is the same as saying that derivative $f_m''(x)$ exists for all x.

In general, we proceed by induction. Suppose we are given that the n^{th} derivative of each function $f_m(x)$ exists, for all x. Then the preceding equation shows that the n^{th} derivative of the function $f_m'(x)$ also exists, for all x. This is the same as saying that the $(n+1)^{st}$ derivative of $f_m(x)$ exists.

It follows that the n^{th} derivative of $f_m(x)$ exists, for all x and all n. And of course it is continuous because the $(n+1)^{st}$ derivative exists.

Now we suppose $f_m(x) = \Sigma a_n x^n$ on some non-trivial interval about $x = 0$, and derive a contradiction. If $f_m(x)$ equals this power series, then the coefficients a_n must satisfy the equations

$$a_n = \frac{f_m^{(n)}(0)}{n!}$$

for all n. We know that $f_m(x)$ vanishes when $x = 0$. Using the equation

$$f_m'(x) = -mf_{m+1}(x) + 2f_{m+3}(x)$$

repeatedly, we see that all the derivatives of $f_m(x)$ also vanish at $x = 0$. Therefore $a_n = 0$ for all n, so $f_m(x)$

is identically zero in some interval about $x = 0$. But this is not true; indeed the function $f_m(x)$ vanishes <u>only</u> for $x = 0$.

FOURIER SERIES

Let us summarize what we know about power series:

I CONVERGENCE Given a power series

$$\sum_{k=0}^{\infty} a_k x^k,$$

there is a number r with $0 \leq r \leq \infty$, such that the series converges absolutely for $|x| < r$ and diverges for $|x| > r$; we call r the radius of convergence.

II UNIQUENESS If

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

for x in some non-trivial interval about 0, then

$$a_k = f^{(k)}(0)/k! \quad .$$

III TAYLOR SERIES If $f^{(k)}(0)$ exists for all k, then we can write down the series

$$\sum_{k=0}^{\infty} a_k x^k, \quad \text{where} \quad a_k = f^{(k)}(0)/k! \ .$$

This series is called the Taylor series of f. It may not converge to f, however; it will do so only if the error term $E_n(x)$ goes to 0 as n approaches ∞. In this case, f is said to be analytic.

IV DIFFERENTIATION AND INTEGRATION If $f(x)$ equals a power series in some non-trivial interval about 0, then $f'(x)$ and $\int_0^x f(t)\,dt$ can be computed by differentiating and integrating the series term-by-term. These new series have the same radius of convergence as the original series,

V APPROXIMATION Among all polynomials of degree n, the Taylor polynomial of f is the one that equals f at 0, and whose first n derivatives equal those of f at 0. If f is analytic, it approximates f very well for x near 0 (and less well as x becomes large).

Now we consider series whose terms are not powers of x, but are of the form $\sin nx$ and $\cos mx$, for n and m positive integers. One motivation for considering such functions is that they are periodic, so they would be natural functions to consider if one wished to represent a periodic function by an infinite series. (Such functions are often called wave functions, and are important in the applications.)

So let us consider a general series of the form

$$\tfrac{1}{2}a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) .$$

Such a series is called a trigonometric series. (The factor $\frac{1}{2}$ is inserted for later convenience.) We will consider the analogues of statements I - V for this new series.

I CONVERGENCE

About this there is little to say. Trigonometric series have no particularly nice convergence properties. For instance, the series $\sum (\cos nx)/n$ converges at $x = \pi$ and fails to converge at $x = 0$. What happens in between is anybody's guess!

II UNIQUENESS

Here there is a theorem. But since we don't know the series converges on a non-trivial interval, we must assume it.

Theorem 1. If the trigonometric series

$$\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

converges uniformly to a function $f(x)$ on the interval $[-\pi, \pi]$, then

$$a_0 = (1/\pi) \int_{-\pi}^{\pi} f(x)\, dx ,$$
$$a_n = (1/\pi) \int_{-\pi}^{\pi} f(x) \cos nx\, dx, \quad \text{and}$$
$$b_n = (1/\pi) \int_{-\pi}^{\pi} f(x) \sin nx\, dx .$$

(since $\cos nx \equiv 1$ if $n = 0$, the first of these equations is redundant.)

Proof. Since the series converges uniformly on $[-\pi, \pi]$, it will still converge uniformly if we multiply through by $\cos mx$ or $\sin mx$. Then we can compute the integrals of $f(x)$ or $f(x)\cos mx$ or $f(x)\sin mx$ by integrating the appropriate series term-by-term. It happens that if we integrate from $-\pi$ to π, all but one of the terms equal 0! This follows from the integration formulas

$$\int_{-\pi}^{\pi} \cos nx \cos mx \, dx = 0 \quad \text{if } n \neq m$$

$$\int_{-\pi}^{\pi} \sin nx \sin mx \, dx = 0 \quad \text{if } n \neq m$$

$$\int_{-\pi}^{\pi} \sin nx \cos mx \, dx = 0 \quad \text{always}$$

$$\int_{-\pi}^{\pi} \sin nx \, dx = 0$$

$$\int_{-\pi}^{\pi} \cos nx \, dx = 0.$$

Finally, the fact that the integrals from $-\pi$ to π of $\cos^2 nx$ and $\sin^2 nx$ equal π gives us the factor of $(1/\pi)$ in the above equations. □

III FOURIER SERIES

Suppose f is an integrable function defined on $[-\pi, \pi]$. Then we can write down the trigonometric series

$$\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

where the coefficients are given by the integral formulas in the statement of Theorem 1. This series is then called the Fourier series of f. Just as was the case with the Taylor series of a function, however, the series may not converge to f.

The remarkable fact about Fourier series is that they converge under very weak assumptions about the function f, in contrast to the situation for Taylor series, where the function must have derivatives of all orders and, in addition, be analytic. We shall state without proof several theorems concerning the convergence of Fourier series. In order to do so, we must make the following definition:

Suppose that f is continuous on an open interval about p, except possibly at the point p itself. If both the limits

$$\lim_{x \to p+} f(x) \quad \text{and} \quad \lim_{x \to p-} f(x)$$

exist (and are finite), we say that f has at most a jump discontinuity at p.

Theorem 2. Suppose $f(x)$ is continuous on the interval $[-\pi, \pi]$, and that f' is continuous on this interval except for finitely many jump discontinuities. Suppose also that $f(-\pi) = f(\pi)$. Then the Fourier series of f converges uniformly to f on the interval $[-\pi, \pi]$.

Remark. Note that if the Fourier series of f is to converge to f on $[-\pi, \pi]$, then it is necessary that $f(\pi) = f(-\pi)$, since all the functions involved have this property. Note further that if the convergence is to be uniform, it is necessary that f be continuous, since the limit of a uniformly convergent series of continuous functions is continuous.

What is remarkable is the fact that you need to assume very little more than these two necessary conditions in order to ensure that the series converges uniformly. The situation is very different from that for Taylor series!

Now of course if f is not continuous, then there is no hope of getting the Fourier series of f to converge uniformly. Even in this case, however, the series will try as hard as it can to converge! That is the substance of the following theorem:

Theorem 3. Suppose that f and f' are continuous on $[-\pi, \pi]$ except for finitely many jump discontinuities. Then the Fourier series of f converges to f at each point of $(-\pi, \pi)$ at which f is continuous; and the convergence is uniform on any closed interval in $(-\pi, \pi)$ on which f is continuous.

At every point p of $(-\pi, \pi)$, the series converges to the number

$$\tfrac{1}{2}\left(\lim_{x \to p+} f(x) + \lim_{x \to p-} f(x)\right).$$

And at π and $-\pi$, it converges to

$$\tfrac{1}{2}\left(\lim_{x \to -\pi+} f(x) + \lim_{x \to \pi-} f(x)\right).$$

We can understand better what happens at π and $-\pi$ if we note the following: Given $f(x)$, let us look at its values on the half-open interval $[-\pi, \pi)$, and extend f to the entire real line by defining

$$g(x + 2n\pi) = f(x)$$

for all x and all n. The function g is called the periodic extension of f.

Now if the Fourier series of f converges to f for some x in $[-\pi,\pi)$, it will automatically converge to the periodic extension g of f at any point of the form $x + 2n\pi$. In some sense, then, it is more natural to deal with functions $g(x)$ that are of period 2π and defined on the entire real line. It now becomes clear why the Fourier series of f may not converge to f at π or $-\pi$ (even if f, which is only defined on $[-\pi,\pi]$, is continuous there). For the periodic extension g of f will not be continuous at π unless the right and left hand limits of g at π equal the value of g at π.

Restated in these terms, Theorem 3 becomes the following:

<u>Theorem 4.</u> Let $g(x)$ be a function of period 2π, defined for all x. Suppose g and g' are continuous on $[-\pi,\pi]$ except for finitely many jump discontinuities. Then the Fourier series of g has the following properties:

(i) It converges to $g(x)$ whenever g is continuous at x.

(ii) It converges uniformly to g on each closed interval on which g is continuous.

(iii) It converges to the average of the right and left hand limits of g at each point where g is discontinuous.

To illustrate these theorems, we compute some examples. Before doing so, let us recall that we call $f(x)$ an <u>even</u> function if $f(x) = f(-x)$ for all x, and we call it an <u>odd</u> function if $f(x) = -f(-x)$ for all x. The integral of an odd function from $-a$ to a is always 0, because cancellation occurs. The following is an immediate consequence:

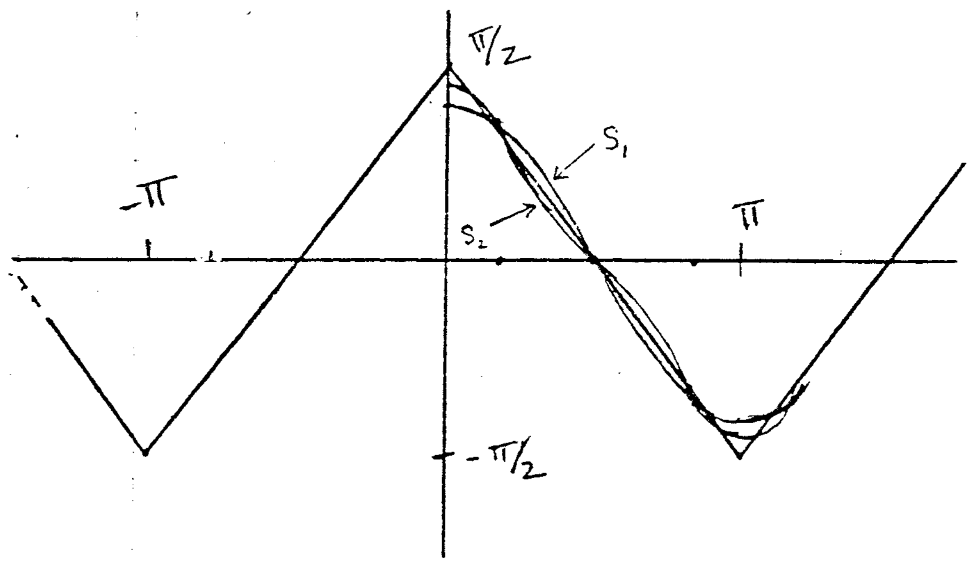
<u>Theorem 5</u>. If $f(x)$ is an even function, then all the terms $b_n \sin nx$ are missing from its Fourier series. If $f(x)$ is odd, then all the terms $\frac{1}{2}a_0$ and $a_n \cos nx$ are missing from its Fourier series.

<u>Proof</u>. If f is an even function, then $f(x)\sin nx$ is odd, while if f is odd, then $f(x)\cos nx$ is odd. $\square$

Example 1. Consider the function

$$f(x) = \tfrac{1}{2}\pi + x \quad \text{for} \quad -\pi \le x \le 0,$$
$$f(x) = \tfrac{1}{2}\pi - x \quad \text{for} \quad 0 \le x \le \pi.$$

Its graph is pictured; it is called a triangular wave. (We have actually pictured the periodic extension of f.)



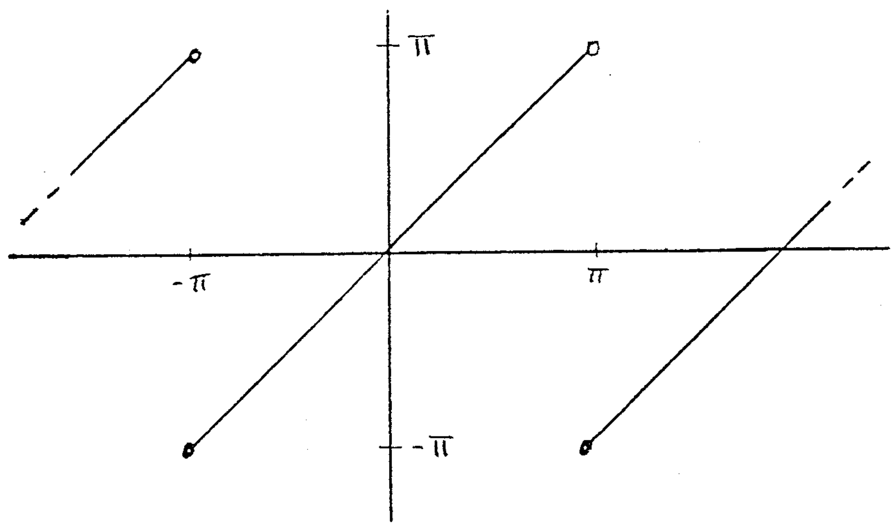
This function is even, so only cosine terms appear in its Fourier series. Direct computation of the coefficients a_n, by integration, gives us the series

$$\frac{4}{\pi}\left[\cos x + (1/9)\cos 3x + (1/25)\cos 5x + \dots\right].$$

The first two partial sums of this series, s_1 and s_2, are pictured above. Note how closely they approximate the function.

This function satisfies the hypotheses of Theorem 2; this theorem predicts that the series will converge uniformly (since f is continuous and $f(-\pi) = f(\pi)$ and f' has only jump discontinuities). And indeed, it does converge uniformly, by comparison with the series $\sum 1/n^2$.

Example 2. Consider the function $f(x) = x$ for $-\pi \le x < \pi$. Let g be its periodic extension. The graph of g is pictured below.

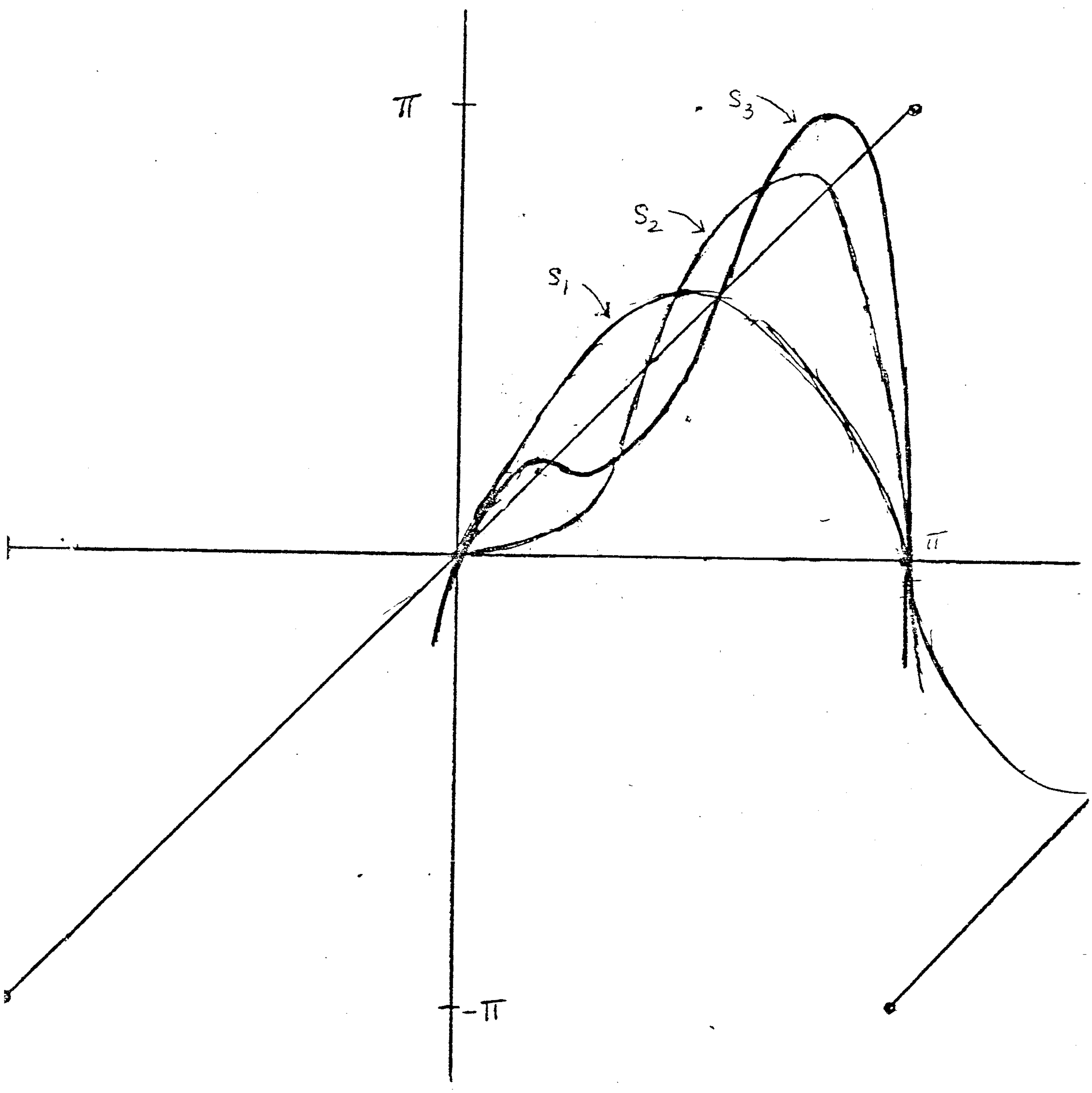


This is a case where the function f is continuous on $[-\pi, \pi)$, but its periodic extension has a jump at π. Thus we do not expect the Fourier series of f to converge to f at π or $-\pi$, but rather to the average of the left and right hand limits, which is 0. And this is exactly what happens.

Since f is an odd function, no cosine terms appear in its Fourier series. Direct computation gives us the series

$$2\left[\sin x - (1/2)\sin 2x + (1/3)\sin 3x - (1/4)\sin 4x + \dots\right] .$$

The first three partial sums s_1, s_2, and s_3 are pictured in the following figure. Note that the convergence is not nearly as rapid as in the preceding example, and that it gets much worse as one approaches π, where g fails to be continuous.



π
S_3
S_2
S_1
π
-π

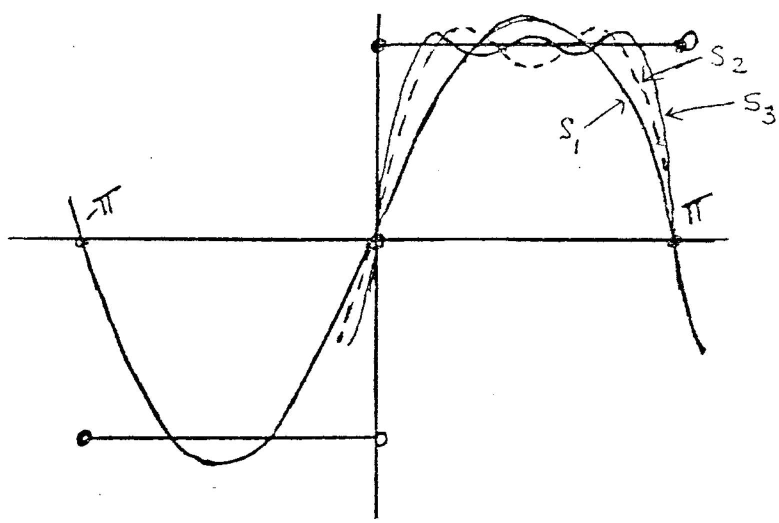
Example 3. Finally, let us consider the following function, which is called the square wave function:

$$f(x) = 1 \text{ if } 0 \le x < \pi$$
$$f(x) = -1 \text{ if } -\pi \le x < 0 .$$

This function is also odd; its Fourier series is the series

$$(4/\pi)\left[\sin x + (1/3)\sin 3x + (1/5)\sin 5x + \dots\right].$$

The first three partial sums are sketched. Note that the convergence becomes worse as one approaches the discontinuities (of the periodic extension of f).



IV DIFFERENTIATION AND INTEGRATION

We know that in general a uniformly convergent series can be integrated term by term. A much stronger result holds for Fourier series; in fact, one does not even need to assume that the series converges!

Theorem 6. Suppose that f is continuous on $[-\pi,\pi]$ except for finitely many jump discontinuities. Although the Fourier series of f need not converge to f, it is still true that if you integrate each term of the series from a to b (where a and b are points of $[-\pi,\pi]$), then the resulting series will converge to the number

$$\int_a^b f(x)\,dx.$$

There is no similar theorem about differentiating a Fourier series.

V APPROXIMATION

Just as was the case for an analytic function and its Taylor series, the Fourier series of a function f will, under quite weak conditions, approximate the function. The difference lies in how one measures the closeness of the approximation. Rather than measuring the actual difference between the values of the function and of the partial sums of the series, we measure the <u>average value</u>, in some sense, of this difference. Specifically, we make the following definition:

Suppose that $f(x)$ is a given function on the interval $[a,b]$. And suppose we seek to approximate f by another function $g(x)$ on this interval. In this case, we call the number

$$E(f,g) = \int_a^b (f(x) - g(x))^2 \, dx$$

the <u>mean square error</u> in this approximation.

One has the following theorem:

<u>Theorem 7</u>. Let $f(x)$ be continuous, except for finitely many jump discontinuities, on $[-\pi,\pi]$. Among all "trigonometric polynomials" of the form

$$h_n(x) = \tfrac{1}{2}a_0 + \sum_{i=1}^{n} (a_i \cos ix + b_i \sin ix) ,$$

the one for which the mean square error $E(f,h_n)$ is a minimum is the one for which the coefficients are the Fourier coefficients of f.

Furthermore, in this situation, the mean square error goes to zero as n approaches ∞.

What this last sentence says is that even though the Fourier series of f may not converge to g in the ordinary sense, it will converge "in the mean."

GENERALIZATIONS

It is remarkable how different the theorems concerning the convergence of power series and the convergence of Fourier series are. It is then natural to ask why the functions $\sin nx$ and $\cos nx$ play such a special role. Perhaps it is because they are periodic. But that is not the case; their periodicity is important only if one wishes to represent periodic functions. If the functions one wishes to represent are not periodic, there are many other systems of functions that will do as well.

The crucial property we needed was that when we multiplied two of the functions $\sin nx$ and $\cos nx$ together and integrated from $-\pi$ to π, we got zero!

For example, it is not at all hard to find a sequence of functions

$P_0(x),\ P_1(x),\ P_2(x),\ldots,P_n(x),\ldots$

such that $P_n(x)$ is a polynomial of degree n, for each n, and such that

$$\int_{-1}^{1} P_n(x)\, P_m(x)\, dx = 0$$

whenever $n \neq m$. These polynomials are uniquely determined up to a constant factor; it is traditional to multiply each by an appropriate constant so that

$$\int_{-1}^{1} P_n(x)\, P_n(x)\, dx = 1.$$

These polynomials are called the <u>Legendre polynomials</u>. It follows just as in the proof of Theorem 1 that if a series of the form

$$\sum_{n=0}^{\infty} a_n P_n(x)$$

converges uniformly to a function $f(x)$ on the interval $[-1,1]$, then the coefficients a_n are given by the equation

$$a_n = \int_{-1}^{1} f(x)\, P_n(x)\, dx.$$

For any integrable function $f(x)$, the series for which the coefficients are given by this formula is called the <u>Fourier-Legendre series</u> of f. There are theorems about these series that are directly analogous to those about Fourier series mentioned above. Their uses in the applications of mathematics are abundant.