5
Differential calculus

We now wish to give a rigorous treatment of differential calculus. We shall give a precise definition of what it means to differentiate a function, and build a theory of differentiation on this definition. The most important theorem of the chapter is the so-called mean value theorem, which we shall establish rigorously and then put to a variety of applications.

Differentiating a function means finding its instantaneous rate of change. Graphically, it means finding the slope of the tangent to the curve $y = f(x)$ at a given point P (see Fig. 5.1).

To obtain a numerical value for the rate of change of $f(x)$ at the instant x, we consider the average rate of change over an interval $[x, x+h]$, i.e.,

$$\frac{f(x+h)-f(x)}{h},$$

and see what happens when $h \to 0$. This is equivalent to drawing a chord PQ on the graph and considering the limit of the slope of PQ as Q→P (see Fig. 5.2).

We are led to define the *derivative* $f'(x)$ of the function $f(x)$ as the limit

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

Of course we cannot assume that the limit will always exist, so we say $f(x)$ is *differentiable* whenever the limit does exist. We shall exclude the case where

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \pm\infty,$$

so that, if $f(x)$ is differentiable over an interval I, then $f'(x)$ is another function defined on I.

The notation $f'(x)$ will be used exclusively in preference to the Leibnitz notation dy/dx, though it will sometimes be useful to rephrase things the Leibnitz way in order to see what is going on in

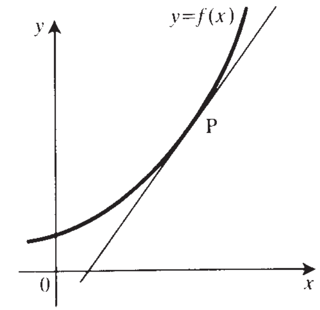

Fig. 5.1

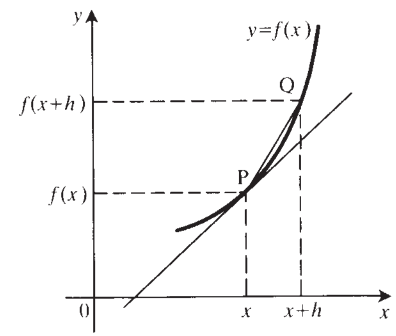

Fig. 5.2

a situation before attempting a rigorous justification. The expression dy/dx of course means the *limiting value* of the ratio between the change dy in y and the change dx in x as both dy and dx tend to zero.

The formal definition is as follows.

5.1 Definition

$f(x)$ is *differentiable* at x if

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

exists (finite), and the value of the limit wherever it exists is called the *derivative* of $f(x)$ at x, and is denoted by $f'(x)$. □

5.2 Examples

(i) $f(x) \equiv C$ (constant).
In this case we have

$$\frac{f(x+h)-f(x)}{h}=\frac{C-C}{h}=0$$

for all x and for all $h \neq 0$, so $f(x)$ is differentiable at every x with derivative $f'(x) \equiv 0$.

(ii) $f(x)=x$.
Here we have

$$\frac{f(x+h)-f(x)}{h}=\frac{(x+h)-x}{h}=\frac{h}{h}=1$$

for all x and for all $h \neq 0$, so $f(x)$ is differentiable everywhere with derivative $f'(x) \equiv 1$.

(iii) $f(x)=1/x \quad (x \neq 0),$
$\quad\quad =0 \quad (x=0).$

If $x \neq 0$, then

$$\frac{f(x+h)-f(x)}{h}=-\frac{1}{x(x+h)}$$

for all h satisfying $0<|h|<|x|$. Therefore $f(x)$ is differentiable with $f'(x)=-1/x^2$ for all $x \neq 0$.

If $x=0$, then

$$\frac{f(x+h)-f(x)}{h}=\frac{f(h)-f(0)}{h}=\frac{1}{h^2}$$

which diverges to infinity as $h \to 0$, so $f(x)$ is not differentiable at $x=0$. □

The following theorem describes the fundamental relationship between differentiable functions and continuous functions.

5.3 Theorem

If a function is differentiable at any point, then it must be continuous there. □

Proof Suppose $f(x)$ is differentiable at x. Then

$$\begin{aligned} f(x+h)-f(x) &= h\frac{f(x+h)-f(x)}{h} \\ &\to 0 \,.\, f'(x) \\ &= 0 \end{aligned}$$

as $h \to 0$. Hence $f(x)$ is continuous at x. □

Observe that $1/x$ is discontinuous at $x=0$ whatever (finite) value we give it at $x=0$, so cannot be differentiable there, by 5.3, confirming our findings of 5.2 (iii).

Observe also that the converse of 5.3 is false. The standard counter-example is the following.

5.4 Counter-example

$f(x)=|x|$.
$f(x)$ is continuous at $x=0$, but is not differentiable there since

$$\frac{f(h)-f(0)}{h}=\frac{|h|}{h}=\operatorname{sgn} h \qquad (h \neq 0)$$

which has no limit as $h \to 0$. □

The above example leads us to introduce the notion of one-sided differentiability, which we define as follows.

5.5 Definition

$f(x)$ is differentiable *on the right* at x if

$$\lim_{h \to 0_+} \frac{f(x+h)-f(x)}{h}$$

exists (finite). The limit is called the *right-hand* derivative and is denoted by $f'_+(x)$. □

Left-hand differentiability and derivative are defined similarly.

For example, $f(x)=|x|$ has $f'_+(0)=1$, $f'_-(0)=-1$.

Our next task is to decide which functions are differentiable, and to find their derivatives. We shall adopt a similar procedure to that used in Chapter 4 of constructing new functions from old ones by various methods. Here we not only have to establish differentiability of the new functions, we also have to determine their derivatives in terms of the derivatives of the old functions.

5.6 Theorem: Arithmetic of differentiable functions

If $f(x)$, $g(x)$ are differentiable at x, then so are
(i) $s(x)=f(x)+g(x)$,
(ii) $p(x)=f(x)g(x)$,
(iii) $q(x)=f(x)/g(x)$,
provided, in case (iii), $g(x)\neq 0$, and their derivatives are
(i) $s'(x)=f'(x)+g'(x)$,
(ii) $p'(x)=f'(x)g(x)+f(x)g'(x)$,
(iii) $q'(x)=(f'(x)g(x)-f(x)g'(x))/(g(x))^2$.

Proofs (i) We have

$$\frac{s(x+h)-s(x)}{h}=\frac{f(x+h)-f(x)}{h}+\frac{g(x+h)-g(x)}{h}$$
$$\to f'(x)+g'(x)$$

as $h\to 0$. Hence $s(x)$ is differentiable at x with derivative

$$s'(x)=f'(x)+g'(x).$$

(ii) We have

$$\frac{p(x+h)-p(x)}{h}=\frac{f(x+h)g(x+h)-f(x)g(x)}{h}$$
$$=\frac{f(x+h)-f(x)}{h}g(x+h)+f(x)\frac{g(x+h)-g(x)}{h}$$
$$\to f'(x)g(x)+f(x)g'(x)$$

as $h\to 0$ (by 5.3).

(iii) If $g(x)\neq 0$, then also $g(x+h)\neq 0$ for all $|h|<\delta$ for some $\delta>0$ by continuity (5.3). Therefore, if h satisfies $0<|h|<\delta$ for this δ, then we have

$$\frac{q(x+h)-q(x)}{h}=\frac{f(x+h)g(x)-f(x)g(x+h)}{hg(x)g(x+h)}$$
$$=\frac{1}{g(x)g(x+h)}$$
$$\times\left\{\frac{f(x+h)-f(x)}{h}g(x)-f(x)\frac{g(x+h)-g(x)}{h}\right\}$$
$$\to\frac{f'(x)g(x)-f(x)g'(x)}{(g(x))^2}$$

as $h\to 0$, by 5.3.

5.7 Corollaries

(i) If $f(x)$ is differentiable at x, then so is $Cf(x)$ for any constant C, and the derivative is $Cf'(x)$.

(ii) If $f(x)$, $g(x)$ are differentiable at x, then so is $f(x)-g(x)$, with derivative $f'(x)-g'(x)$.

Proofs (i) follows from 5.2 (i) and 5.6 (ii).
(ii) follows from (i) and 5.6 (i). □

5.8 Exercises

Prove that for any positive integer n,
(i) x^n is differentiable with derivative nx^{n-1},
(ii) $1/x^n$ is differentiable for all $x\neq 0$ with derivative $-n/x^{n+1}$.
Hint. Use 5.2, 5.6 and induction.

5.9 Applications

(i) Any polynomial $p(x)$ is differentiable for all x.
(ii) Any rational function $r(x)=p(x)/q(x)$, where $p(x)$, $q(x)$ are polynomials, is differentiable except where $q(x)=0$.

5.10 Theorem: Composition of differentiable functions

If $g(x)$ is differentiable at x, and if $f(y)$ is differentiable at $y=g(x)$, then $c(x)=f(g(x))$ is differentiable at x with derivative $c'(x)=f'(g(x))g'(x)$. □

Observe that Leibnitz notation is suggestive here. If we write $z=f(y)$, we have $z=f(g(x))=c(x)$, and so

$$c'(x)=\frac{dz}{dx}=\frac{dz}{dy}\frac{dy}{dx}=f'(y)g'(x)=f'(g(x))g'(x)$$

on the assumption that dy can be cancelled. Two questions arise: firstly, can dy be cancelled in the limit? and secondly, how do we proceed if $dy=0$?

The formal proof is as follows.

Proof of 5.10 We distinguish two cases.

Case 1
$g'(x) \neq 0$

In this case there must exist $\delta > 0$ such that

$$\frac{g(x+h)-g(x)}{h} \neq 0,$$

and therefore $g(x+h) \neq g(x)$, for all $0 < |h| < \delta$. Hence for h in this range we have

$$\begin{aligned}\frac{c(x+h)-c(x)}{h} &= \frac{f(g(x+h))-f(g(x))}{h}\\ &= \frac{f(g(x+h))-f(g(x))}{g(x+h)-g(x)}\frac{g(x+h)-g(x)}{h}\\ &= \frac{f(y+k)-f(y)}{k}\frac{g(x+h)-g(x)}{h},\end{aligned}$$

where $k = g(x+h) - g(x) \neq 0$,

$$\begin{aligned}&\to f'(y)g'(x)\\ &= f'(g(x))g'(x)\end{aligned}$$

as $h \to 0$, since $k \to 0$ as $h \to 0$ by the continuity of $g(x)$.

Case 2
$g'(x) = 0$

The argument of case 1 shows that

$$\frac{c(x+h)-c(x)}{h} \to f'(g(x))g'(x)$$

as $h \to 0$ via values such that $g(x+h) \neq g(x)$. If $g(x+h) = g(x)$ for any $h \neq 0$, then we have

$$\frac{c(x+h)-c(x)}{h} = \frac{f(g(x+h))-f(g(x))}{h} = 0,$$

and hence

$$\frac{c(x+h)-c(x)}{h} \to 0 = f'(g(x))g'(x)$$

as $h \to 0$ via these values also. □

5.11 Examples

(i) $f(x) = \sin\frac{1}{x}$ $\quad(x \neq 0),$

$\qquad = 0 \quad (x = 0).$

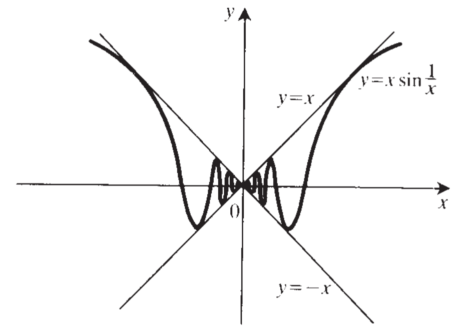

Fig. 5.3

$f(x)$ is differentiable for all x except $x = 0$, where $f(x)$ is not differentiable since $f(x)$ is not continuous at $x = 0$.

Of course we are assuming $\sin x$ is differentiable for all x. This will be justified shortly. (See 5.18.)

(ii) $f(x) = x \sin\frac{1}{x}$ $\quad(x \neq 0),$

$\qquad = 0 \quad (x = 0)$ (see Fig. 5.3).

$f(x)$ is continuous everywhere and differentiable everywhere except $x = 0$. In fact

$$\frac{f(h)-f(0)}{h} = \sin\frac{1}{h}$$

which has no limit as $h \to 0$.

(iii) $f(x) = x^2 \sin\frac{1}{x}$ $\quad(x \neq 0),$

$\qquad = 0 \quad (x = 0)$ (see Fig. 5.4).

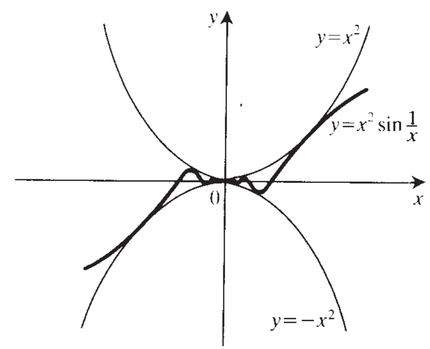

Fig. 5.4

$f(x)$ is differentiable everywhere including $x=0$. In fact,

$$\frac{f(h)-f(0)}{h}=h\sin\frac{1}{h}\to 0$$

as $h\to 0$, so $f'(0)$ exists and $=0$.

5.12 Exercise

Find the derivative $f'(x)$ of the function

$$f(x)=x^2\sin\frac{1}{x}\qquad (x\neq 0)$$
$$=0\qquad (x=0)$$

at $x\neq 0$. (Assume the derivative of $\sin x$ is $\cos x$.) Show $f'(x)$ is discontinuous at $x=0$. Explain why this doesn't contradict 5.3. □

5.13 Inverse function theorem

If $y=f(x)$ is differentiable and strictly increasing over an interval $[a, b]$, and if $f'(x)\neq 0$ for all $x\in[a, b]$, then the inverse function $x=\phi(y)$ is differentiable over $[f(a), f(b)]$ with derivative $\phi'(y)=1/f'(x)=1/f'(\phi(y))$. □

Again we are in a situation where it is illuminating to use Leibnitz notation. Here we have

$$\phi'(y)=\frac{dx}{dy}=1\bigg/\frac{dy}{dx}=1/f'(x)$$

and the question is whether we can invert dy/dx in the limit. The justification is as follows.

Proof of 5.13 For any $k\neq 0$ we have

$$\frac{\phi(y+k)-\phi(y)}{k}=\frac{h}{f(x+h)-f(x)}$$

where $x=\phi(y)$, $x+h=\phi(y+k)$. Now $k\to 0$ implies

$$h=\phi(y+k)-\phi(y)\to 0$$

by the continuity of $\phi(y)$. (See 4.26.) Therefore

$$\frac{\phi(y+k)-\phi(y)}{k}\to\frac{1}{f'(x)}$$

as $k\to 0$, which is the required result. □

5.14 Application: Differentiability of $x^{1/n}$ $(n\geq 2)$

Theorem 5.13 shows that $f(x)=x^{1/n}$ is differentiable for all $x>0$ with derivative

$$f'(x)=\frac{1}{ny^{n-1}},$$

where $y=x^n$,

$$=\frac{1}{nx^{(n-1)/n}}$$

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

If n is odd, $x^{1/n}$ is differentiable for all $x<0$ also. At $x=0$, $x^{1/n}$ has an infinite derivative, one-sided if n is even, two-sided if n is odd, which corresponds to the fact that the tangent to the graph is vertical at the origin. This indicates what happens when the condition $f'(x)\neq 0$ is left out of Theorem 5.13. □

The final method we shall consider for constructing differentiable functions is by means of infinite series. There is a Weierstrassian theorem (cf. 4.30) which enables differentiation of an infinite series of functions to be justified, but its proof relies on the mean value theorem, so it will have to wait until that theorem is covered.

We shall therefore proceed forthwith to the mean value theorem, and among its applications the first will be to prove Weierstrass' theorem for differentiable functions. We shall attack the mean value theorem via a special case of it known as Rolle's theorem, which is as follows.

5.15 Rolle's theorem

If $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b), and if $f(a)=f(b)$, then there exists $c\in(a, b)$ such that $f'(c)=0$ (see Fig. 5.5).

Proof If $f(x)$ is constant on $[a, b]$, then $f'(x)=0$ for all $x\in[a, b]$. Otherwise $f(x)$ attains a maximum or minimum at an interior point $c\in(a, b)$ (see 4.58).

If $f(c)\geq f(x)$ for all $x\in[a, b]$, then

$$\frac{f(c+h)-f(c)}{h}\leq 0$$

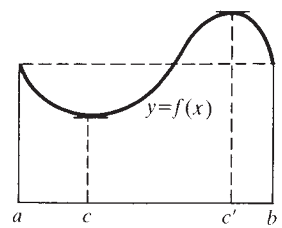

Fig. 5.5

if $h > 0$, and

$$\frac{f(c+h)-f(c)}{h} \geqslant 0$$

if $h < 0$, so letting $h \to 0$ we obtain $f'(c) \leqslant 0$, $f'(c) \geqslant 0$ simultaneously, and hence $f'(c) = 0$.

If $f(c) \leqslant f(x)$ for all $x \in [a, b]$, then $f'(c) = 0$ similarly. □

5.16 Mean value theorem

If $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b), then

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

for some $c \in (a, b)$ (see Fig. 5.6) □

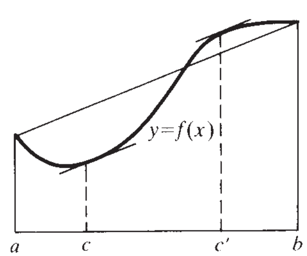

Fig. 5.6

Observe that the expression

$$\frac{f(b)-f(a)}{b-a}$$

represents the 'mean value' of $f'(x)$ over the interval $[a, b]$, and the mean value theorem is saying that $f'(x)$ must attain this mean value somewhere in the open interval (a, b).

Proof Let

$$g(x) = f(x) - Cx$$

where the constant C is chosen in such a way that $g(x)$ satisfies the conditions of Rolle's theorem. In fact, we require

$$f(a) - Ca = f(b) - Cb,$$

which gives

$$C = \frac{f(b)-f(a)}{b-a}.$$

Then, by Rolle's theorem, there must exist $c \in (a, b)$ such that $g'(c) = 0$. Hence $f'(c) = C$ as required. □

An equivalent way of stating the mean value theorem is to say that if $f(x)$ is continuous over the interval $[x, x+h]$ and differentiable over the interval $(x, x+h)$, then

$$\frac{f(x+h)-f(x)}{h} = f'(x+\theta h)$$

for some θ satisfying $0 < \theta < 1$. Observe that this statement is equally valid for h positive or negative.

As promised, our first application of the mean value theorem is to the differentiation of infinite series.

5.17 Weierstrass' theorem

If $\sum_1^\infty f_n(x)$ is a convergent series of differentiable functions on the interval $[a, b]$, and if there exists a convergent series of positive constants $\sum_1^\infty M_n'$ such that

$$|f_n'(x)| \leqslant M_n'$$

for all $x \in [a, b]$ and for all $n \geqslant 1$ (*W′-condition*), then $s(x) =$

$\sum_1^\infty f_n(x)$ is differentiable on $[a, b]$ and its derivative is $s'(x) = \sum_1^\infty f_n'(x)$. □

Observe that Abel's series $\sum_1^\infty (\sin nx)/n$ (see 4.29) formally differentiates to $\sum_1^\infty \cos nx$, which doesn't satisfy the W′-condition. In fact, $\sum_1^\infty \cos nx$ diverges for every x.

Proof of 5.17 The W′-condition clearly implies immediately that $\sum_1^\infty f_n'(x)$ converges for each $x \in [a, b]$. Suppose $x \in [a, b]$ and $\varepsilon > 0$ are given. Then we can choose N such that

$$\sum_{N+1}^{\infty} M_n' < \tfrac{1}{3}\varepsilon,$$

and for each $n = 1, 2, \ldots, N$ we can choose $\delta_n > 0$ such that

$$\left|\frac{f_n(x+h) - f_n(x)}{h} - f_n'(x)\right| < \frac{\varepsilon}{3N}$$

for all $0 < |h| < \delta_n$. Let $\delta = \min_{1 \le n \le N} \delta_n$. Then $\delta > 0$ and, for any $0 < |h| < \delta$, we have

$$\begin{aligned}
&\left|\frac{s(x+h) - s(x)}{h} - \sum_1^\infty f_n'(x)\right| \\
&= \left|\sum_1^\infty \left(\frac{f_n(x+h) - f_n(x)}{h} - f_n'(x)\right)\right| \\
&\le \sum_1^N \left|\frac{f_n(x+h) - f_n(x)}{h} - f_n'(x)\right| + \sum_{N+1}^\infty \left|\frac{f_n(x+h) - f_n(x)}{h} - f_n'(x)\right| \\
&< \sum_1^N \frac{\varepsilon}{3N} + \sum_{N+1}^\infty |f_n'(x + \theta_n h) - f_n'(x)|,
\end{aligned}$$

for some $0 < \theta_n < 1$ by the mean value theorem,

$$\begin{aligned}
&\le \tfrac{1}{3}\varepsilon + \sum_{N+1}^\infty (|f_n'(x+\theta_n h)| + |f_n'(x)|) \\
&< \tfrac{1}{3}\varepsilon + 2\sum_{N+1}^\infty M_n' \\
&< \tfrac{1}{3}\varepsilon + \tfrac{2}{3}\varepsilon \\
&= \varepsilon.
\end{aligned}$$

5.18 Application to power series

If the power series $\sum_0^\infty a_n x^n$ has radius of convergence R, then $f(x) = \sum_0^\infty a_n x^n$ is differentiable over $|x| < R$ and its derivative there is

$$f'(x) = \sum_1^\infty n a_n x^{n-1}.$$

Proof The differentiated series $\sum_1^\infty n a_n x^{n-1}$ also has radius of convergence R (see 3.38, question 17). Therefore, for any r satisfying $0 < r < R$, the series $\sum_1^\infty n a_n r^{n-1}$ is absolutely convergent, and so $\sum_1^\infty a_n x^n$ satisfies the W′-condition over $[-r, r]$. (Take $M_n' = n\,|a_n|\,r^{n-1}$.) The result follows. □

We can now differentiate powers and logarithms. In fact, by definition,

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

which has infinite radius of convergence, so e^x is differentiable for all x with derivative obtained by differentiating the series term by term, i.e., e^x itself.

$\log_e x$ is the inverse function of e^x, so is differentiable for all $x > 0$ by 5.13 with derivative $1/x$.

By definition, $x^a = e^{a \log x}$, so is differentiable for all $x > 0$ by 5.10 with derivative ax^{a-1}.

We can also differentiate $\sin x$ and $\cos x$, and hence all the trigonometric functions. In fact, by definition,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

which has infinite radius of convergence, so is differentiable for all x with derivative

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots,$$

i.e. $\cos x$. Similarly $\cos x$ is differentiable for all x with derivative $-\sin x$.

Hence, e.g., $\tan x$ is differentiable for all $x \ne (n + \frac{1}{2})\pi$ with derivative $\sec^2 x$, $\cot x$ is differentiable for all $x \ne n\pi$ with derivative $-\operatorname{cosec}^2 x$, etc.

We must now consider the other applications of the mean value theorem. The first of these concerns uniqueness of primitives.

5.19 Definition

$F(x)$ is a *primitive* (or an *indefinite integral*) of $f(x)$ if $F'(x) = f(x)$. □

5.20 Theorem

If $f(x)$ is differentiable over an interval I, and if $f'(x) = 0$ for all $x \in I$, then $f(x)$ must be constant on I.

Proof For any $a, b \in I$ we have

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

for some c between a and b. Therefore $f'(c) = 0$ and hence $f(a) = f(b)$. □

5.21 Corollary

Any two primitives of a given function $f(x)$ must differ by a constant. □

5.22 Application

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

for all $|x| < 1$.

Proof For all $|x| < 1$ we have

$$\frac{1}{1 + x} = 1 - x + x^2 - x^3 + \cdots$$

since the right-hand side is a geometric series. Now $\log(1 + x)$ is a primitive for $1/(1 + x)$ and

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

is a primitive for

$$1 - x + x^2 - x^3 + \cdots$$

by 5.18. Therefore

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + C$$

where C is a constant. Putting $x = 0$ clearly gives $C = 0$. □

5.23 Exercise

Prove that

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$$

for all $|x| < 1$. *Hint*: Consider primitives for both sides of the expansion

$$\frac{1}{1 + x^2} = 1 - x^2 + x^4 - x^6 + \cdots.$$ □

As we remarked at the beginning of the chapter, $f'(x)$ represents the rate of change of $f(x)$. In particular, $f'(x) > 0$ implies $f(x)$ increases, and $f'(x) < 0$ implies $f(x)$ decreases. This apparently self-evident fact is actually quite hard to prove rigorously, at least it would be if we did not have the mean value theorem. With the mean value theorem the proof is spectacularly simple as we now demonstrate.

5.24 Theorem

If $f(x)$ is differentiable over an interval I, and if $f'(x) > 0$ for all $x \in I$, then $f(x)$ is strictly increasing over I. □

Proof Suppose that $a < b$ both $\in I$. Then, by the mean value theorem,

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

for some $c \in (a, b)$. Therefore $c \in I$ and so $f'(c) > 0$. Hence

$$f(b) - f(a) > 0$$

i.e. $f(a) < f(b)$ as required. □

Similarly, one can prove that, if $f'(x) < 0$ over the interval I, then $f(x)$ decreases over I.

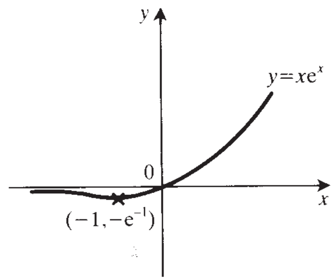

Fig. 5.7

The sceptical reader should attempt to prove 5.24 without the aid of the mean value theorem. It should convince him that the mean value theorem is deeper than it looks, which is really not too surprising when one recalls that it is ultimately based on the Bolzano–Weierstrass theorem (2.38) via the minimax theorem (4.58) and Rolle's theorem (5.15).

Theorem 5.24 can often be used to find maxima and minima of functions. The following example illustrates the method.

5.25 Example

$f(x) = xe^x$.

We have

$$f'(x) = (x+1)e^x$$

so $f'(x) < 0$ if $x < -1$, and $f'(x) > 0$ if $x > -1$. Therefore $f(x)$ decreases strictly over $x < -1$, and increases strictly over $x > -1$. Hence $f(x)$ takes a minimum at $x = -1$ (see Fig. 5.7).

5.26 Exercise

Find maxima and minima of $f(x) = x^2e^x$ and hence sketch the graph. □

Another application of 5.24 is to the proof of inequalities. By way of illustration we shall give a new proof of Bernoulli's inequality. (See 1.12.)

5.27 Example

For all real $x > -1$, and for all integers $n \geqslant 1$, we have

$$(1+x)^n \geqslant 1 + nx.$$

In fact, let

$$f(x) = (1+x)^n - 1 - nx.$$

Then

$$f'(x) = n(1+x)^{n-1} - n \begin{cases} <0 & \text{if } -1 < x < 0 \\ >0 & \text{if } \quad x > 0. \end{cases}$$

Therefore $f(x)$ decreases strictly over the interval $-1 < x < 0$, and increases strictly over the interval $x > 0$. It follows that $f(x)$ takes its minimum value at $x = 0$, and hence $f(x) \geqslant f(0)$, i.e.,

$$(1+x)^n - 1 - nx \geqslant 0,$$

for all $x > -1$, which gives the required result. □

Observe that this method also gives conditions for equality. In fact, for $n \geqslant 2$, we have equality only at $x = 0$.

5.28 Exercises

Prove the following inequalities.

(i) $e^x \geqslant 1 + x$ for all x.

(ii) $\log x \leqslant x - 1$ for all $x > 0$.

(iii) $\sin x < x < \tan x$ for all $0 < x < \frac{1}{2}\pi$.

Investigate conditions for equality in each case. □

In geometrical terms, the mean value theorem says that, between any two points A, B on the graph of a differentiable function, there is a third point C where the tangent is parallel to the chord AB (see Fig. 5.6). This property is enjoyed by plane curves in general, not just those arising as graphs of functions (see Fig. 5.8).

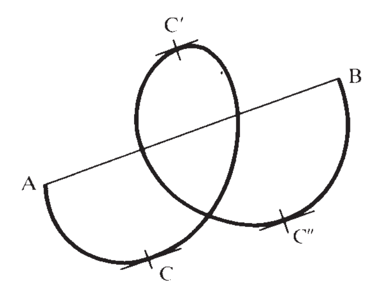

Fig. 5.8

To see what this means arithmetically we represent the curve as the locus of the points having co-ordinates (x, y) where $x = f(t)$, $y = g(t)$ are two functions of a parameter t. Existence of a tangent at every point is ensured by requiring that $f'(t)$, $g'(t)$ both exist and never vanish simultaneously. The slope at the point $(f(t), g(t))$ is then

$$\frac{dy}{dx} = \frac{g'(t)}{f'(t)}$$

which is always well defined, though possibly infinite which simply means that the tangent is vertical. The slope of the chord joining the points corresponding to $t = a$, $t = b$ is of course

$$\frac{g(b) - g(a)}{f(b) - f(a)},$$

which again may be infinite with the obvious interpretation.

We arrive at the following theorem.

5.29 Cauchy's mean value theorem

If $f(t)$, $g(t)$ are continuous for all $t \in [a, b]$ and differentiable for all $t \in (a, b)$, and if $f'(t)$, $g'(t)$ do not vanish simultaneously on (a, b), then

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

for some $c \in (a, b)$.

Proof Consider the function

$$h(t) = (g(b) - g(a))f(t) - (f(b) - f(a))g(t).$$

Clearly, $h(t)$ is continuous on $[a, b]$, differentiable on (a, b) and $h(a) = h(b)$. Therefore, by Rolle's theorem, there exists $c \in (a, b)$ such that $h'(c) = 0$, i.e.,

$$(g(b) - g(a))f'(c) = (f(b) - f(a))g'(c).$$

The condition of $f'(c)$, $g'(c)$ not both being zero ensures that the equation

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

has a meaning. □

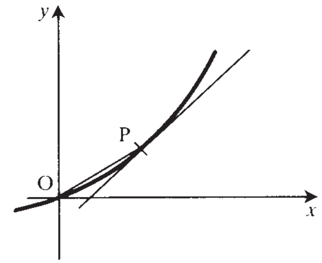

Fig. 5.9

5.30 Application: L'Hôpital's rule

If $f(t)$, $g(t)$ both $\to 0$ as $t \to c$, then

$$\lim_{t \to c} \frac{f(t)}{g(t)} = \lim_{t \to c} \frac{f'(t)}{g'(t)}$$

whenever the second limit exists.

Proof The existence of $\lim_{t \to c} f'(t)/g'(t)$ presupposes that $f'(t)$, $g'(t)$ both exist and $g'(t) \neq 0$ for all t satisfying $0 < |t - c| < \delta$ for some $\delta > 0$. The condition that $\lim_{t \to c} f(t) = \lim_{t \to c} g(t) = 0$ implies that, if we assume $f(c) = g(c) = 0$, then $f(t)$ and $g(t)$ are continuous at c. Therefore, for any t satisfying $0 < |t - c| < \delta$, we can apply Cauchy's mean value theorem to obtain

$$\frac{f(t)}{g(t)} = \frac{f(t) - f(c)}{g(t) - g(c)} = \frac{f'(u)}{g'(u)}$$

for some u between t and c. Now $u \to c$ as $t \to c$ so

$$\lim_{t \to c} \frac{f(t)}{g(t)} = \lim_{u \to c} \frac{f'(u)}{g'(u)} = \lim_{t \to c} \frac{f'(t)}{g'(t)}$$

as required. □

The geometrical interpretation of l'Hôpital's rule is that if a curve passes through the origin O, then the limiting slope of the chord OP is equal to the limiting slope of the tangent at P as $P \to$ O along the curve (see Fig. 5.9).

5.31 Example

$$\lim_{x \to 0} \frac{\log(1 + x)}{x} = \lim_{x \to 0} \frac{1}{1 + x} = 1.$$ □

5.32 Exercises

Use L'Hôpital's rule to evaluate the following limits.

$$\text{(i) } \lim_{x\to 0}\frac{\sin x}{x} \qquad \text{(ii) } \lim_{x\to\pi}\frac{\sin x}{x-\pi} \qquad \text{(iii) } \lim_{x\to 0}\frac{1-\cos x}{x^2}$$ □

Our final application of the mean value theorem is concerned with the representation of functions by power series.

Suppose we are given a function $f(x)$ and we wish to represent it in the form

$$f(x)=\sum_0^\infty a_n x^n.$$

If we assume the power series has a positive radius of convergence R then, by 5.18, $f(x)$ must be differentiable and

$$f'(x)=\sum_0^\infty n a_n x^{n-1}$$

for all $|x|<R$. By 5.18 again, $f'(x)$ must be differentiable, i.e. $f(x)$ is *twice* differentiable, and the derivative of $f'(x)$, which we denote by $f''(x)$ and call the *second* derivative of $f(x)$, must be given by

$$f''(x)=\sum_0^\infty n(n-1)a_n x^{n-2}$$

for all $|x|<R$. In fact, $f(x)$ must be differentiable any number of times, and if we denote the kth derivative by $f^{(k)}(x)$, we have

$$f^{(k)}(x)=\sum_0^\infty n(n-1)\cdots(n-k+1)a_n x^{n-k}$$

for all $|x|<R$.

The above power series representations of the successive derivatives of $f(x)$ enable us to determine the coefficients $(a_n)_{n\geq 0}$ in terms of $f(x)$. In fact, putting $x=0$ in the formula for $f^{(k)}(x)$ gives

$$a_k=\frac{1}{k!}f^{(k)}(0)$$

for each $k\geq 0$.

5.33 Definition

We say $f(x)$ is *infinitely* differentiable over an interval I if derivatives $f'(x)$, $f''(x)$, $f'''(x)$, . . . , $f^{(k)}(x)$, . . . of all orders exist at every $x\in I$. □

5.34 Examples

Any polynomial $p(x)$ is infinitely differentiable for all x. Any rational function $p(x)/q(x)$ is infinitely differentiable except at zeros of the denominator $q(x)$. The functions e^x, $\sin x$, $\cos x$ are infinitely differentiable for all x, and the function $\log x$ is infinitely differentiable for all $x>0$. □

5.35 Definition

If $f(x)$ is infinitely differentiable at $x=0$, then

$$a_n=\frac{1}{n!}f^{(n)}(0)$$

is called the nth *Maclaurin coefficient* of $f(x)$, and the series

$$\sum_0^\infty a_n x^n=\sum_0^\infty \frac{1}{n!}f^{(n)}(0)x^n$$

is called the *Maclaurin series* of $f(x)$. □

5.36 Exercises

Find the Maclaurin series of the following functions.

(i) e^x (ii) $\sin x$ (iii) $\cos x$

(iv) $\dfrac{1}{1-x}$ (v) $\log(1+x)$ □

It is natural to expect, or at least hope, that the Maclaurin series of $f(x)$ will converge to $f(x)$ on its interval of convergence. Unfortunately, this is in general not so as the following example demonstrates.

5.37 Example

$$\begin{aligned} f(x)&=e^{-1/x^2} \qquad (x\neq 0),\\ &=0 \qquad (x=0). \end{aligned}$$

$f(x)$ is easily shown to be infinitely differentiable for all x with $f^{(n)}(0)=0$ for all x. Therefore the Maclaurin coefficients of $f(x)$ are all zero, and hence the Maclaurin series of $f(x)$ is trivial, so doesn't converge to $f(x)$ anywhere except at $x=0$. □

It follows that validity of a Maclaurin expansion has to be verified in each particular case. For many of the important functions of analysis this can be done without too much difficulty. For example, the Maclaurin series of e^x, $\sin x$, $\cos x$ are their defining series, and so validity for all x is immediate. The Maclaurin series of $(1-x)^{-1}$ is the geometric series $\sum_0^\infty x^n$ which we already know converges to $(1-x)^{-1}$ for all $|x|<1$ (see 3.2). The function $\log(1+x)$ has Maclaurin series

$$\sum_1^\infty \frac{(-1)^{n-1}}{n} x^n$$

which converges to $\log(1+x)$ for all $|x|<1$ (see 5.22). We shall show shortly that this expansion is also valid at $x=1$, i.e.

$$\sum_1^\infty \frac{(-1)^{n-1}}{n} = \log 2$$

(see 5.42).

Partial sums of Maclaurin series give polynomial approximations to the functions in question, which can be expected to be fairly accurate, at least for small x. If one requires polynomial approximations for other x, it is more appropriate to use power series of the form

$$\sum_0^\infty a_n(x-a)^n$$

where a is some fixed point $\neq 0$. If

$$f(x) = \sum_0^\infty a_n(x-a)^n,$$

then, as above, the coefficients a_n are given by

$$a_n = \frac{1}{n!} f^{(n)}(a)$$

for all $n \geqslant 0$.

5.38 Definition

If $f(x)$ is infinitely differentiable at $x=a$, then

$$a_n = \frac{1}{n!} f^{(n)}(a)$$

is called the nth *Taylor* coefficient of $f(x)$ at $x=a$, and the series

$$\sum_0^\infty a_n(x-a)^n = \sum_0^\infty \frac{1}{n!} f^{(n)}(a)(x-a)^n$$

is called the *Taylor* series of $f(x)$ at $x=a$. □

5.39 Exercises

Find the Taylor series of the following functions at the points given.

(i) e^x at $x=2$.

(ii) $\sin x$ at $x=\frac{1}{2}\pi$.

(iii) $\dfrac{1}{1-x}$ at $x=-1$. □

As for Maclaurin series, we cannot expect Taylor series to converge to the function given, though in many important cases they do. (See 5.39.)

5.40 Definition

The Nth *remainder* $R_N(x)$ for the Taylor series of $f(x)$ at $x=a$ is the difference between $f(x)$ and the Nth partial sum of its Taylor series at $x=a$, i.e.,

$$R_N(x) = f(x) - \sum_0^{N-1} a_n(x-a)^n$$

where $a_n = f^{(n)}(a)/n!$. □

Clearly the Taylor series converges to $f(x)$ if and only if $R_N(x) \to 0$ as $N \to \infty$. The following theorem gives a formula for $R_N(x)$ which enables the question of its convergence to zero to be decided in many cases.

5.41 Taylor's theorem with remainder

If $f(x)$ is infinitely differentiable over $[a, x]$, then

$$f(x) = \sum_0^{N-1} \frac{(x-a)^n}{n!} f^{(n)}(a) + \frac{(x-a)^N}{N!} f^{(N)}(c)$$

for some $c \in (a, x)$. □

Observe that case $N=1$ is the mean value theorem. The general case is sometimes called the Nth mean value theorem. An alternative formulation useful in many applications is

$$f(x+h)=\sum_{0}^{N-1}\frac{h^n}{n!}f^{(n)}(x)+\frac{h^N}{N!}f^{(N)}(x+\theta h)$$

for some $0<\theta<1$, if $f(x)$ is infinitely differentiable over the interval $[x, x+h]$.

Proof of 5.41 Let $g(t)$ be defined by

$$g(t)=\sum_{0}^{N-1}\frac{(x-t)^n}{n!}f^{(n)}(t)+A\frac{(x-t)^N}{N!}$$

where A is a constant chosen so as to make $g(t)$ satisfy the conditions of Rolle's theorem over $[a, x]$. In fact, A must satisfy the equation

$$f(x)=\sum_{0}^{N-1}\frac{(x-a)^n}{n!}f^{(n)}(a)+\frac{A(x-a)^N}{N!}.$$

Now the derivative of $g(t)$ is

$$g'(t)=\frac{(x-t)^{N-1}}{(N-1)!}(f^{(N)}(t)-A),$$

so Rolle's theorem gives $A=f^{(N)}(c)$ for some $c\in(a, x)$. The result follows immediately. □

We content ourselves with two applications.

5.42 Theorem

$$\sum_{1}^{\infty}(-1)^{n-1}/n=\log 2.$$

Proof If $f(x)=\log(1+x)$ and $a=0$, then the Nth remainder $R_N(x)$ is

$$\begin{aligned}R_N(x)&=\frac{x^N}{N!}f^{(N)}(c)\qquad(\text{for some } c\in(0, x))\\&=\frac{x^N}{N!}\frac{(-1)^{N-1}(N-1)!}{(1+c)^N}\\&=\frac{x^N}{N}\frac{(-1)^{N-1}}{(1+c)^N}.\end{aligned}$$

Therefore

$$\begin{aligned}|R_N(1)|&=\frac{1}{N(1+c)^N}\qquad(\text{for some } 0<c<1)\\&<\frac{1}{N},\end{aligned}$$

and hence $R_N(1)\to 0$ as $N\to\infty$. □

5.43 Theorem

If $f(x)$ is infinitely differentiable at x and $f'(x)=0$, then $f(x)$ has a maximum at x if $f''(x)<0$, and a minimum if $f''(x)>0$.

Proof The second mean value theorem (5.41 case $N=2$) gives

$$f(x+h)=f(x)+hf'(x)+\tfrac{1}{2}h^2f''(x+\theta h)$$

for some $0<\theta<1$. If $f''(x)>0$, then there exists $\delta>0$ such that $f''(x+h)>0$ for all $|h|<\delta$, since $f''(x)$ is continuous. (See 5.3.) Therefore for all $|h|<\delta$ we have

$$\begin{aligned}f(x+h)&=f(x)+\tfrac{1}{2}h^2f''(x+\theta h)\\&\geqslant f(x)\end{aligned}$$

showing $f(x)$ has a minimum at x. The argument for a maximum is similar.

5.44 Miscellaneous exercises

1. Discuss the differentiability and continuity of the following functions at $x=0$.

(i) $\log|x|$ (ii) $x\log|x|$ (iii) $x^2\log|x|$

(Give each function the value 0 at $x=0$.)

2. Sketch the graphs of the following functions.

(i) e^{-x^2} (ii) xe^{-x^2} (iii) $x^2e^{-x^2}$

3. Prove the following inequalities.

(i) $\sin x>\dfrac{2}{\pi}x$ $(0<x<\tfrac{1}{2}\pi)$

(ii) $\dfrac{\sin x}{x}>\dfrac{\pi-x}{\pi}$ $(0<x<\pi)$

(iii) $\dfrac{\sin x}{x} \geqslant \dfrac{\pi^2 - x^2}{\pi^2 + x^2}$ (all real x)

4. Use L'Hôpital's rule (5.30) to find the following limits.

(i) $\lim_{x\to 0} \dfrac{\tan x - x}{x - \sin x}$ (ii) $\lim_{x\to 1} \left(\dfrac{1}{\log x} - \dfrac{1}{x-1}\right)$

(iii) $\lim_{n\to\infty} n(2^{1/n} - 1)$ (iv) $\lim_{n\to\infty} \left(1 + \dfrac{3}{n}\right)^n$

Hint: (for (iii), (iv)) put $x = 1/n$ and let $x \to 0$.

5. Prove that, if $f(x)$ is twice differentiable at x, then

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

Find the value of

$$\lim_{h\to 0} \frac{f(x) - 3f(x+h) + 3f(x+2h) - f(x+3h)}{h^3}$$

under suitable assumptions on $f(x)$.

Generalize.

6. Given that $f(x)$ is twice differentiable at $x = c$, and that $f'(c) = 0$, $f''(c) > 0$, show that $f(x)$ has a (local) minimum at $x = c$ (cf. 5.43).

What can you say if $f''(c) < 0$? $f''(c) = 0$?

7. Prove that $(x - \alpha)^2$ is a factor of the polynomial $p(x)$ if and only if $p(\alpha) = p'(\alpha) = 0$.

8. Discuss the convergence or otherwise of the series

$$\sum_1^\infty (-1)^{n-1} \frac{\sin \log n}{n}.$$

Hint: Group the terms in pairs and use the mean value theorem. (Observe that the alternating series test is inapplicable here.)

9. Given that $f(0) = 0$, $f''(x) \leqslant 0$ for all $x \geqslant 0$, prove that

$$f(a+b) \leqslant f(a) + f(b)$$

for all $a \geqslant 0$, $b \geqslant 0$. *Hint:* Given $a \leqslant b$, apply the mean value theorem to $f(x)$ over the intervals $[0, a]$, $[b, a+b]$.

Prove the inequality

$$\frac{a+b}{1+a+b} \leqslant \frac{a}{1+a} + \frac{b}{1+b}$$

for all $a \geqslant 0$, $b \geqslant 0$.

10. Prove that, if $f''(x) \geqslant 0$ for all x, then

$$f\left(\frac{a+b}{2}\right) \leqslant \frac{f(a) + f(b)}{2}$$

for all a, b.

Prove conversely that, if the above inequality holds for all a, b, then we must have $f''(x) \geqslant 0$ for all x. *Hint*: (for second part): Use question 5.

11. Show that the Maclaurin series of $(1+x)^\alpha$ is

$$\sum_0^\infty \binom{\alpha}{n} x^n$$

where

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}.$$

Show that it has radius of convergence 1.

Prove that the function

$$f(x) = \sum_0^\infty \binom{\alpha}{n} x^n \qquad (|x| < 1)$$

satisfies the differential equation

$$(1+x)f'(x) = \alpha f(x).$$

Deduce that $f(x) = (1+x)^\alpha$ for $|x| < 1$.

12. Prove that, for all $x > -1$,

$$(1+x)^\alpha \geqslant 1 + \alpha x,$$

if $\alpha < 0$ or $\alpha > 1$, but

$$(1+x)^\alpha \leqslant 1 + \alpha x$$

if $0 < \alpha < 1$. *Hint* Use 5.41, case $N = 2$.

Investigate conditions for equality.

13. Use 5.41 to prove the following inequalities.

(i) $x - \dfrac{x^2}{2} < \log(1+x) < x - \dfrac{x^2}{2} + \dfrac{x^3}{3}$ $(x > 0)$

(ii) $x - \dfrac{x^3}{6} \leqslant \sin x \leqslant x$ $(x \geqslant 0)$

(iii) $1 - \dfrac{x^2}{2} \leqslant \cos x \leqslant 1 - \dfrac{x^2}{2} + \dfrac{x^4}{24}$ (all real x)