5.9 ROLLE'S THEOREM

It is geometrically evident that a sufficiently "smooth" curve which crosses the x-axis at both endpoints of an interval $[a, b]$ must have a "turning point" somewhere between a and b. The precise statement of this fact is known as Rolle's theorem.

Theorem 5.10 (Rolle). *Assume f has a derivative (finite or infinite) at each point of an open interval (a, b), and assume that f is continuous at both endpoints a and b. If $f(a) = f(b)$ there is at least one interior point c at which $f'(c) = 0$.*

Proof. We assume f' is never 0 in (a, b) and obtain a contradiction. Since f is continuous on a compact set, it attains its maximum M and its minimum m somewhere in $[a, b]$. Neither extreme value is attained at an interior point (otherwise f' would vanish there) so both are attained at the endpoints. Since $f(a) = f(b)$, then $m = M$, and hence f is constant on $[a, b]$. This contradicts the assumption that f' is never 0 on (a, b). Therefore $f'(c) = 0$ for some c in (a, b).

5.10 THE MEAN-VALUE THEOREM FOR DERIVATIVES

Theorem 5.11 (Mean-Value Theorem). *Assume that f has a derivative (finite or infinite) at each point of an open interval (a, b), and assume also that f is continuous at both endpoints a and b. Then there is a point c in (a, b) such that*

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

Geometrically, this states that a sufficiently smooth curve joining two points A and B has a tangent line with the same slope as the chord AB. We will deduce Theorem 5.11 from a more general version which involves two functions f and g in a symmetric fashion.

Theorem 5.12 (Generalized Mean-Value Theorem). *Let f and g be two functions, each having a derivative (finite or infinite) at each point of an open interval (a, b) and each continuous at the endpoints a and b. Assume also that there is no interior point x at which both $f'(x)$ and $g'(x)$ are infinite. Then for some interior point c we have*

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

NOTE. When $g(x) = x$, this gives Theorem 5.11.

Proof. Let $h(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)]$. Then $h'(x)$ is finite if both $f'(x)$ and $g'(x)$ are finite, and $h'(x)$ is infinite if exactly one of $f'(x)$ or $g'(x)$ is infinite. (The hypothesis excludes the case of *both* $f'(x)$ and $g'(x)$ being infinite.) Also, h is continuous at the endpoints, and $h(a) = h(b) = f(a)g(b) - g(a)f(b)$. By Rolle's theorem we have $h'(c) = 0$ for some interior point, and this proves the assertion.

NOTE. The reader should interpret Theorem 5.12 geometrically by referring to the curve in the xy-plane described by the parametric equations $x = g(t)$, $y = f(t)$, $a \le t \le b$.

There is also an extension which does not require continuity at the endpoints.

Theorem 5.13. *Let f and g be two functions, each having a derivative (finite or infinite) at each point of (a, b). At the endpoints assume that the limits $f(a+)$, $g(a+)$, $f(b-)$ and $g(b-)$ exist as finite values. Assume further that there is no interior point x at which both $f'(x)$ and $g'(x)$ are infinite. Then for some interior point c we have*

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

Proof. Define new functions F and G on $[a, b]$ as follows:

$$F(x) = f(x) \quad \text{and} \quad G(x) = g(x) \qquad \text{if } x \in (a, b);$$

$$F(a) = f(a+), \quad G(a) = g(a+), \quad F(b) = f(b-), \quad G(b) = g(b-).$$

Then F and G are continuous on $[a, b]$ and we can apply Theorem 5.12 to F and G to obtain the desired conclusion.

The next result is an immediate consequence of the Mean-Value Theorem.

Theorem 5.14. *Assume f has a derivative (finite or infinite) at each point of an open interval (a, b) and that f is continuous at the endpoints a and b.*

a) *If f' takes only positive values (finite or infinite) in (a, b), then f is strictly increasing on $[a, b]$.*
b) *If f' takes only negative values (finite or infinite) in (a, b), then f is strictly decreasing on $[a, b]$.*
c) *If f' is zero everywhere in (a, b) then f is constant on $[a, b]$.*

Proof. Choose $x < y$ and apply the Mean-Value Theorem to the subinterval $[x, y]$ of $[a, b]$ to obtain

$$f(y) - f(x) = f'(c)(y - x) \qquad \text{where } c \in (x, y).$$

All the statements of the theorem follow at once from this equation.

By applying Theorem 5.14 (c) to the difference $f - g$ we obtain:

Corollary 5.15. *If f and g are continuous on $[a, b]$ and have equal finite derivatives in (a, b), then $f - g$ is constant on $[a, b]$.*

5.11 INTERMEDIATE-VALUE THEOREM FOR DERIVATIVES

In Theorem 4.33 we proved that a function f which is continuous on a compact interval $[a, b]$ assumes every value between its maximum and its minimum on the interval. In particular, f assumes every value between $f(a)$ and $f(b)$. A similar result will now be proved for functions which are derivatives.

Theorem 5.16 (Intermediate-value theorem for derivatives). *Assume that f is defined on a compact interval $[a, b]$ and that f has a derivative (finite or infinite) at each interior point. Assume also that f has finite one-sided derivatives $f'_+(a)$ and $f'_-(b)$ at the endpoints, with $f'_+(a) \neq f'_-(b)$. Then, if c is a real number between $f'_+(a)$ and $f'_-(b)$, there exists at least one interior point x such that $f'(x) = c$.*

Proof. Define a new function g as follows:

$$g(x) = \frac{f(x) - f(a)}{x - a} \quad \text{if } x \neq a, \quad g(a) = f'_+(a).$$

Then g is continuous on the closed interval $[a, b]$. By the intermediate-value theorem for continuous functions, g takes on every value between $f'_+(a)$ and $[f(b) - f(a)]/(b - a)$ in the interior (a, b). By the Mean-Value Theorem, we have $g(x) = f'(k)$ for some k in (a, x) whenever $x \in (a, b)$. Therefore f' takes on every value between $f'_+(a)$ and $[f(b) - f(a)]/(b - a)$ in the interior (a, b). A similar argument applied to the function h, defined by

$$h(x) = \frac{f(x) - f(b)}{x - b} \quad \text{if } x \neq b, \quad h(b) = f'_-(b),$$

shows that f' takes on every value between $[f(b) - f(a)]/(b - a)$ and $f'_-(b)$ in the interior (a, b). Combining these results, we see that f' takes on every value between $f'_+(a)$ and $f'_-(b)$ in the interior (a, b), and this proves the theorem.

NOTE. Theorem 5.16 is still valid if one or both of the one-sided derivatives $f'_+(a)$, $f'_-(b)$, is infinite. The proof in this case can be given by considering the auxiliary function g defined by the equation $g(x) = f(x) - cx$, if $x \in [a, b]$. Details are left to the reader.

The intermediate-value theorem shows that a derivative cannot change sign in an interval without taking the value 0. Therefore, we have the following strengthening of Theorem 5.14(a) and (b).

Theorem 5.17. *Assume f has a derivative (finite or infinite) on (a, b) and is continuous at the endpoints a and b. If $f'(x) \neq 0$ for all x in (a, b) then f is strictly monotonic on $[a, b]$.*

The intermediate-value theorem also shows that monotonic derivatives are necessarily continuous.

Theorem 5.18. *Assume f' exists and is monotonic on an open interval (a, b). Then f' is continuous on (a, b).*

Proof. We assume f' has a discontinuity at some point c in (a, b) and arrive at a contradiction. Choose a closed subinterval $[\alpha, \beta]$ of (a, b) which contains c in its interior. Since f' is monotonic on $[\alpha, \beta]$ the discontinuity at c must be a jump discontinuity (by Theorem 4.51). Hence f' omits some value between $f'(\alpha)$ and $f'(\beta)$, contradicting the intermediate-value theorem.

5.12 TAYLOR'S FORMULA WITH REMAINDER

As noted earlier, if f is differentiable at c, then f is approximately a linear function near c. That is, the equation

$$f(x) = f(c) + f'(c)(x - c),$$

is approximately correct when $x - c$ is small. Taylor's theorem tells us that, more generally, f can be approximated by a polynomial of degree $n - 1$ if f has a derivative of order n. Moreover, Taylor's theorem gives a useful expression for the error made by this approximation.

Theorem 5.19 (Taylor). *Let f be a function having finite nth derivative $f^{(n)}$ everywhere in an open interval (a, b) and assume that $f^{(n-1)}$ is continuous on the closed interval $[a, b]$. Assume that $c \in [a, b]$. Then, for every x in $[a, b]$, $x \neq c$, there exists a point x_1 interior to the interval joining x and c such that*

$$f(x) = f(c) + \sum_{k=1}^{n-1} \frac{f^{(k)}(c)}{k!}(x - c)^k + \frac{f^{(n)}(x_1)}{n!}(x - c)^n.$$

Taylor's theorem will be obtained as a consequence of a more general result that is a direct extension of the generalized Mean-Value Theorem.

Theorem 5.20. *Let f and g be two functions having finite nth derivatives $f^{(n)}$ and $g^{(n)}$ in an open interval (a, b) and continuous $(n - 1)$st derivatives in the closed interval $[a, b]$. Assume that $c \in [a, b]$. Then, for every x in $[a, b]$, $x \neq c$, there exists a point x_1 interior to the interval joining x and c such that*

$$\left[f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(c)}{k!}(x - c)^k\right] g^{(n)}(x_1) = f^{(n)}(x_1)\left[g(x) - \sum_{k=0}^{n-1} \frac{g^{(k)}(c)}{k!}(x - c)^k\right].$$

NOTE. For the special case in which $g(x) = (x - c)^n$, we have $g^{(k)}(c) = 0$ for $0 \leq k \leq n - 1$ and $g^{(n)}(x) = n!$. This theorem then reduces to Taylor's theorem.

Proof. For simplicity, assume that $c < b$ and that $x > c$. Keep x fixed and define new functions F and G as follows:

$$F(t) = f(t) + \sum_{k=1}^{n-1} \frac{f^{(k)}(t)}{k!}(x - t)^k,$$

$$G(t) = g(t) + \sum_{k=1}^{n-1} \frac{g^{(k)}(t)}{k!}(x - t)^k,$$

for each t in $[c, x]$. Then F and G are continuous on the closed interval $[c, x]$ and have finite derivatives in the open interval (c, x). Therefore, Theorem 5.12 is applicable and we can write

$$F'(x_1)[G(x) - G(c)] = G'(x_1)[F(x) - F(c)], \qquad \text{where } x_1 \in (c, x).$$

This reduces to the equation

$$F'(x_1)[g(x) - G(c)] = G'(x_1)[f(x) - F(c)], \tag{a}$$

since $G(x) = g(x)$ and $F(x) = f(x)$. If, now, we compute the derivative of the sum defining $F(t)$, keeping in mind that each term of the sum is a product, we find that all terms cancel but one, and we are left with

$$F'(t) = \frac{(x - t)^{n-1}}{(n - 1)!} f^{(n)}(t).$$

Similarly, we obtain

$$G'(t) = \frac{(x - t)^{n-1}}{(n - 1)!} g^{(n)}(t).$$

If we put $t = x_1$ and substitute into (a), we obtain the formula of the theorem.

5.13 DERIVATIVES OF VECTOR-VALUED FUNCTIONS

Let $\mathbf{f} : (a, b) \to \mathbf{R}^n$ be a vector-valued function defined on an open interval (a, b) in $\mathbf{R}$. Then $\mathbf{f} = (f_1, \ldots, f_n)$ where each component f_k is a real-valued function defined on (a, b). We say that $\mathbf{f}$ is differentiable at a point c in (a, b) if each component f_k is differentiable at c and we define

$$\mathbf{f}'(c) = (f_1'(c), \ldots, f_n'(c)).$$

In other words, the derivative $\mathbf{f}'(c)$ is obtained by differentiating each component of $\mathbf{f}$ at c. In view of this definition, it is not surprising to find that many of the theorems on differentiation are also valid for vector-valued functions. For example, if $\mathbf{f}$ and $\mathbf{g}$ are vector-valued functions differentiable at c and if λ is a real-valued function differentiable at c, then the sum $\mathbf{f} + \mathbf{g}$, the product $\lambda\mathbf{f}$, and the dot product $\mathbf{f} \cdot \mathbf{g}$ are differentiable at c and we have

$$(\mathbf{f} + \mathbf{g})'(c) = \mathbf{f}'(c) + \mathbf{g}'(c),$$
$$(\lambda\mathbf{f})'(c) = \lambda'(c)\mathbf{f}(c) + \lambda(c)\mathbf{f}'(c),$$
$$(\mathbf{f} \cdot \mathbf{g})'(c) = \mathbf{f}'(c) \cdot \mathbf{g}(c) + \mathbf{f}(c) \cdot \mathbf{g}'(c).$$

The proofs follow easily by considering components. There is also a chain rule for differentiating composite functions which is proved in the same way. If $\mathbf{f}$ is vector-valued and if u is real-valued, then the composite function $\mathbf{g}$ given by $\mathbf{g}(x) = \mathbf{f}[u(x)]$ is vector-valued. The chain rule states that

$$\mathbf{g}'(c) = \mathbf{f}'[u(c)]u'(c),$$

if the domain of $\mathbf{f}$ contains a neighborhood of $u(c)$ and if $u'(c)$ and $\mathbf{f}'[u(c)]$ both exist.

The Mean-Value Theorem, as stated in Theorem 5.11, does not hold for vector-valued functions. For example, if $\mathbf{f}(t) = (\cos t, \sin t)$ for all real t, then

$$\mathbf{f}(2\pi) - \mathbf{f}(0) = \mathbf{0},$$

but $\mathbf{f}'(t)$ is never zero. In fact, $\|\mathbf{f}'(t)\| = 1$ for all t. A modified version of the Mean-Value Theorem for vector-valued functions is given in Chapter 12 (Theorem 12.8).

5.14 PARTIAL DERIVATIVES

Let S be an open set in Euclidean space $\mathbf{R}^n$, and let $f : S \to \mathbf{R}$ be a real-valued function defined on S. If $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{c} = (c_1, \ldots, c_n)$ are two points of S having corresponding coordinates equal except for the kth, that is, if $x_i = c_i$ for $i \neq k$ and if $x_k \neq c_k$, then we can consider the limit

$$\lim_{x_k \to c_k} \frac{f(\mathbf{x}) - f(\mathbf{c})}{x_k - c_k}.$$

When this limit exists, it is called the *partial derivative* of f with respect to the kth coordinate and is denoted by

$$D_k f(\mathbf{c}), \qquad f_k(\mathbf{c}), \qquad \frac{\partial f}{\partial x_k}(\mathbf{c}),$$

or by a similar expression. We shall adhere to the notation $D_k f(\mathbf{c})$.

This process produces n further functions $D_1 f, D_2 f, \ldots, D_n f$ defined at those points in S where the corresponding limits exist.

Partial differentiation is not really a new concept. We are merely treating $f(x_1, \ldots, x_n)$ as a function of one variable at a time, holding the others fixed. That is, if we introduce a function g defined by

$$g(x_k) = f(c_1, \ldots, c_{k-1}, x_k, c_{k+1}, \ldots, c_n),$$

then the partial derivative $D_k f(\mathbf{c})$ is exactly the same as the ordinary derivative $g'(c_k)$. This is usually described by saying that we differentiate f with respect to the kth variable, holding the others fixed.

In generalizing a concept from $\mathbf{R}^1$ to $\mathbf{R}^n$, we seek to preserve the important properties in the one-dimensional case. For example, in the one-dimensional case, the existence of the derivative at c implies continuity at c. Therefore it seems desirable to have a concept of derivative for functions of several variables which will imply continuity. Partial derivatives do *not* do this. A function of n variables can have partial derivatives at a point with respect to each of the variables and yet not be continuous at the point. We illustrate with the following example of a function of two variables:

$$f(x, y) = \begin{cases} x + y, & \text{if } x = 0 \text{ or } y = 0, \\ 1, & \text{otherwise.} \end{cases}$$

The partial derivatives $D_1 f(0, 0)$ and $D_2 f(0, 0)$ both exist. In fact,

$$D_1 f(0, 0) = \lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x - 0} = \lim_{x \to 0} \frac{x}{x} = 1,$$

and, similarly, $D_2 f(0, 0) = 1$. On the other hand, it is clear that this function is not continuous at (0, 0).

The existence of the partial derivatives with respect to each variable separately implies continuity in each variable separately; but, as we have just seen, this does not necessarily imply continuity in all the variables simultaneously. The difficulty with partial derivatives is that by their very definition we are forced to consider only one variable at a time. Partial derivatives give us the rate of change of a function in the direction of each coordinate axis. There is a more general concept of derivative which does not restrict our considerations to the special directions of the coordinate axes. This will be studied in detail in Chapter 12.

The purpose of this section is merely to introduce the notation for partial derivatives, since we shall use them occasionally before we reach Chapter 12.

If f has partial derivatives $D_1 f, \ldots, D_n f$ on an open set S, then we can also consider *their* partial derivatives. These are called *second-order* partial derivatives. We write $D_{r,k} f$ for the partial derivative of $D_k f$ with respect to the rth variable. Thus,

$$D_{r,k} f = D_r(D_k f).$$

Higher-order partial derivatives are similarly defined. Other notations are

$$D_{r,k} f = \frac{\partial^2 f}{\partial x_r \, \partial x_k}, \qquad D_{p,q,r} f = \frac{\partial^3 f}{\partial x_p \, \partial x_q \, \partial x_r}.$$