Chapter 9
Derivatives and Differentiability

Abstract Derivatives are usually introduced by fully exploiting the possibility of *dividing* real numbers. We propose an approach that can be extended almost literally to function defined on general normed spaces.

Definition 9.1 Let $f : (a, b) \to \mathbb{R}$ be a function, and let $x_0 \in (a, b)$ be a distinguished point. We say that f is differentiable at x_0 if a real number A exists with the property that

$$f(x) = f(x_0) + A(x - x_0) + o(|x - x_0|) \quad \text{as } x \to x_0. \tag{9.1}$$

The number A is called the *derivative* of f at x_0, and is denoted by any of the symbols

$$f'(x_0), \quad Df(x_0), \quad \mathrm{d}f(x_0), \quad \dot{f}(x_0).$$

The reader should recall that (9.1) is an equivalent way of requiring

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

Remark 9.1 The assumption that f be defined on an open interval (a, b) is essentially for definiteness. Equation (9.1) shows that x_0 must be an accumulation point of the domain of f, but it should also belong to it. It would be possible to differentiate functions defined on a closed interval $[a, b]$, for instance, but at the end-points the derivative would lose several properties. For this reason we define the derivative of a function only at interior points of its domain.

Exercise 9.1 Suppose that f is differentiable at x_0. We want to prove that the number A in (9.1) is uniquely determined. For the sake of contradiction, assume

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
S. Secchi, *A Circle-Line Study of Mathematical Analysis*,
La Matematica per il 3+2 141, https://doi.org/10.1007/978-3-031-19738-3_9

that

$$f(x) = f(x_0) + A(x - x_0) + o(|x - x_0|)$$
$$f(x) = f(x_0) + B(x - x_0) + o(|x - x_0|)$$

as $x \to x_0$. Deduce that $A - B = o(1)$ as $x \to x_0$, and conclude that $A = B$.

If (9.1) holds, then

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = A. \tag{9.2}$$

On the other hand, if

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = A,$$

then

$$\frac{f(x) - f(x_0)}{x - x_0} = A + o(1) \quad \text{as } x \to x_0,$$

or $f(x) - f(x_0) = A(x - x_0) + o(|x - x_0|)$ as $x \to x_0$. We have proved

Theorem 9.1 *For a function $f : (a, b) \to \mathbb{R}$ the following conditions are equivalent:*

(i) f is differentiable at $x_0 \in (a, b)$ and $f'(x_0) = A$;
(ii) the limit $\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$ exists as a real number and is equal to A.

Remark 9.2 Calculus books usually propose the derivative as the limit of the incremental ratio, namely (9.2). Our Definition 9.1 can be formally generalized to the case in which the function f is defined on a normed vector space, like $\mathbb{R}^n$ for $n \geq 2$. Equivalent definitions may be used as they are needed: we will see that (9.1) is the most convenient characterization of the derivative for proving the chain rule.

We record a third definition of the derivative in terms of continuous functions.

Theorem 9.2 *A function $f : (a, b) \to \mathbb{R}$ is differentiable at $x_0 \in (a, b)$ if and only if there exists a continuous function $\omega : (a, b) \to \mathbb{R}$ such that*

$$f(x) = f(x_0) + \omega(x)(x - x_0) \quad \textit{for every } x \in (a, b). \tag{9.3}$$

In this case, $f'(x_0) = \omega(x_0)$.

Proof Condition (9.3) simply means that the function

$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0}, \quad x \neq x_0$$

can be extended at $x = x_0$ continuously. Of course this is true if and only if the limit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists as a real number. This means that f is differentiable at x_0, and by (9.3) we must have $\omega(x_0) = \lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$. □

Corollary 9.1 *If a function is differentiable at a point, then it is continuous at that point.*

Proof This is immediate from (9.3). □

Exercise 9.2 Prove the previous Corollary by using each of the equivalent definitions of the derivative.

9.1 Rules of Differentiation, or the Algebra of Calculus

If two functions f and g are defined on a neighborhood of a point x_0, we can define pointwise the functions $f+g$ and $f \cdot g$: indeed $x \mapsto f(x)+g(x)$ and $x \mapsto f(x)g(x)$ are well defined in a neighborhood of x_0. If $g \neq 0$ in a neighborhood of x_0, then the quotient $x \mapsto f(x)/g(x)$ is also defined.

Theorem 9.3 (Differentiation Rules) *Suppose that f and g are defined on a neighborhood (a, b) of the point x_0. Then*

(i) the function $f + g$ is differentiable at x_0, and $(f + g)'(x_0) = f'(x_0) + g'(x_0)$;
(ii) the function $f \cdot g$ is differentiable at x_0, and

$$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0);$$

(iii) if $g(x_0) \neq 0$, then the function f/g is differentiable at x_0, and

$$(f/g)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

Proof The proof of (i) is left as an easy exercise. A standard proof of (ii) is as follows:

$$\begin{aligned}
\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\
&= \frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x_0) \frac{g(x) - g(x_0)}{x - x_0} \\
&\to f'(x_0)g(x_0) + f(x_0)g'(x_0).
\end{aligned}$$

Another proof, based instead on Definition 9.1, starts from the assumptions $f(x) = f(x_0) + f'(x_0)(x - x_0) + o(1)$, $g(x) = g(x_0) + g'(x_0)(x - x_0) + o(1)$ and then

$$\begin{aligned} f(x)g(x) &= \left[f(x_0) + f'(x_0)(x - x_0) + o(1)\right]\left[g(x_0) + g'(x_0)(x - x_0) + o(1)\right] \\ &= f(x_0)g(x_0) + f'(x_0)g(x_0)(x - x_0) \\ &\quad + f(x_0)g'(x_0)(x - x_0) + [\dots]\, o(1), \end{aligned}$$

where $[\dots]$ contains all the terms that are multiplied by some $o(1)$ in the algebraic expansion. It follows that the linearization of fg at x_0 (exists and) is $f'(x_0)g(x_0) + f(x_0)g'(x_0)$.

The proof of (iii) is more traditional. First of all, since differentiability implies continuity, the condition $g(x_0) \neq 0$ implies that $g \neq 0$ in a neighborhood of x_0. Then we construct

$$\begin{aligned} \frac{\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}}{x - x_0} &= \frac{1}{x - x_0} \frac{f(x)g(x_0) - f(x_0)g(x)}{g(x)g(x_0)} \\ &= \frac{1}{x - x_0} \frac{f(x)g(x_0) - f(x_0)g(x_0) + f(x_0)g(x_0) - f(x_0)g(x)}{g(x)g(x_0)} \\ &\to \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}. \end{aligned}$$

□

Remark 9.3 Formula (iii) is not easily proved by means of Definition 9.1. The trouble is that it is not trivial to extract a linearization formula from the quotient

$$\frac{f(x_0) + f'(x_0)(x - x_0) + o(1)}{g(x_0) + g'(x_0)(x - x_0) + o(1)}.$$

Exercise 9.3 Try to deduce formula (iii) from

$$\frac{f(x_0) + f'(x_0)(x - x_0) + o(1)}{g(x_0) + g'(x_0)(x - x_0) + o(1)}.$$

As a hint, you may begin with the identity

$$\frac{1}{1 - z} = 1 + z + z^2 + z^3 + \cdots = 1 + z + O(z^2).$$

Theorem 9.4 (Chain Rule) *Let $f : (a, b) \to \mathbb{R}$ be differentiable at a point $x_0 \in (a, b)$, and let g be a function defined on a neighborhood of the range $f((a, b))$. If g is differentiable at the point $f(x_0)$, then $g \circ f$ is differentiable at x_0, and*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0). \tag{9.4}$$

Proof From Definition 9.1, we know that there exists function σ and τ such that $\sigma(x) = o(1)$ as $x \to x_0$, $\tau(y) = o(1)$ as $y \to f(x_0)$, and

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \sigma(x)|(-x_0)$$
$$g(y) = g(f(x_0)) + g'(f(x_0))(y - f(x_0) + \tau(y)(y - f(x_0)).$$

Then

$$\begin{aligned} g \circ f(x) &= g(f(x_0)) + g'(f(x_0))(f(x) - f(x_0)) + \tau(f(x)(f(x) - f(x_0)) \\ &= g(f(x_0)) + g'(f(x_0))(f'(x_0) + \sigma(x)|x - x_0|) \\ &\quad +\tau(f(x)(f(x) - f(x_0)) \\ &= g(f(x_0)) + g'(f(x_0))(f'(x_0) + \sigma(x)|x - x_0|) \\ &\quad +\tau(f(x))(f'(x_0)(x - x_0) + \sigma(x)(x - x_0)) \\ &= g(f(x_0)) + g'(f(x_0)f'(x_0)(x - x_0) + \\ &\quad + ([\ldots]\sigma(x) + [\ldots]\tau(f(x)))\,(x - x_0). \end{aligned}$$

As $x \to x_0$, it is immediate to check that $[\ldots]\sigma(x) + [\ldots]\tau(f(x)) \to 0$, and the conclusion follows. □

Important: Warning

The Calculus "proof" of the Chain Rule goes as follows:

$$\begin{aligned} \frac{g(f(x)) - g(f(x_0))}{x - x_0} &= \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \frac{f(x) - f(x_0)}{x - x_0} \\ &\to g'(f(x_0))f'(x_0). \end{aligned}$$

There is a subtle flaw in this computation, since division by $f(x) - f(x_0)$ is legitimate only if $f(x) \neq f(x_0)$ in a neighborhood of x_0. Unfortunately the assumptions of the Theorem do not ensure that this additional condition is satisfied by f. There is a way out, but we do not emphasize this approach, since it cannot be generalized to higher dimension.

Exercise 9.4 Provide a proof of the Chain Rule according to Theorem 9.2.

Theorem 9.5 (Differentiation of the Inverse Function) *Suppose that f is an invertible function on an interval (a, b). If f is differentiable at a point $x_0 \in (a, b)$ and $f'(x_0) \neq 0$, then f^{-1} is differentiable at $y_0 = f(x_0)$. Moreover,*

$$Df^{-1}(y_0) = \frac{1}{f'(x_0)}.$$

Fig. 9.1 Differentiating the inverse function

Proof Since f is continuous and invertible on an interval, its inverse function f^{-1} is continuous on the range of f. Then

$$\lim_{y\to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{x\to x_0} \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}}$$
$$= \frac{1}{f'(x_0)}.$$

□

The necessity of all the assumptions should be clear from Fig. 9.1.

Example 9.1

1. The function f defined by

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

is differentiable at any $x \neq 0$: indeed

$$f'(x) = \sin \frac{1}{x} - \frac{1}{x} \cos \frac{1}{x}.$$

However

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

does not converge as $x \to 0$. Hence f is not differentiable at $x = 0$.

2. The function f defined by

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

is differentiable at $x = 0$, since

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

and $0 \le |x \sin(1/x)| \le x$ for every x. Therefore $f'(0) = 0$.

Exercise 9.5 Suppose that $f : (a, b) \to \mathbb{R}$ is differentiable at a point a. Show that

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

Provide an example of a function such that the limit on the right-hand side exists, but the function is not differentiable at a.

9.2 Mean Value Theorems

The derivative is a local object that can provide *global* properties of functions. This is essentially the basis of mean value theorems.

Theorem 9.6 (Fermat) *Let f be a function defined on the interval $[a, b]$. If f has a local maximum or a local minimum at a point $x_0 \in (a, b)$, and if $f'(x_0)$ exists, then $f'(x_0) = 0$.*

Proof Suppose that x_0 is a local maximum of f, so that there exists $\delta > 0$ such that $a < x_0 - \delta < x_0 < x_0 + \delta < b$. If $x_0 - \delta < x, x_0$, then

$$\frac{f(x) - f(x_0)}{x - x_0} \ge 0,$$

since f attains a local maximum at x_0. Letting $x \to x_0$ in the last inequality, we get $f'(x_0) \ge 0$. Similarly, if $x_0 < x < x_0 + \delta$, then

$$\frac{f(x) - f(x_0)}{x - x_0} \le 0,$$

and letting $x \to x_0$ we get $f'(x_0) \le 0$. Necessarily $f'(x_0) = 0$. The proof for a local minimum reduces to the previous one by considering $-f$ instead of f. □

Theorem 9.7 (Cauchy) *Let f and g be two functions defined on $[a.b]$, which are differentiable on (a, b) and continuous on $[a, b]$. Then there exists a point $c \in (a, b)$ such that*

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

Proof Let us introduce the function h defined on $[a, b]$ by

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

obviously h is continuous and differentiable on (a, b), and $h(a) = h(b)$. It remains to prove that the derivative of h vanishes somewhere inside (a, b). If h turns out to be constant on $[a, b]$, then the proof is complete.

If $h(x) > h(a)$ for some $x \in (a, b)$, then h must attain a global maximum inside (a, b), and at this point h' vanishes by Theorem 9.6.

If $h(x) < h(a)$ for some $x \in (a, b)$, then h must attain a global minimum inside (a, b), and at this point h' vanishes again by Theorem 9.6. □

The simple choice $g = \mathrm{id}$ is surprisingly important: see Fig. 9.2.

Theorem 9.8 (Lagrange) *Let f be a function defined on $[a.b]$, which is differentiable on (a, b) and continuous on $[a, b]$. Then there exists a point $c \in (a, b)$ such that*

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

Corollary 9.2 (Monotonicity) *Suppose f is differentiable on (a, b).*

(a) If $f' \geq 0$ on (a, b), then f is monotonically increasing on (a, b).
(b) If $f' \leq 0$ on (a, b), then f is monotonically decreasing on (a, b).
(c) if $f' = 0$ identically on (a, b), then f is constant on (a, b).

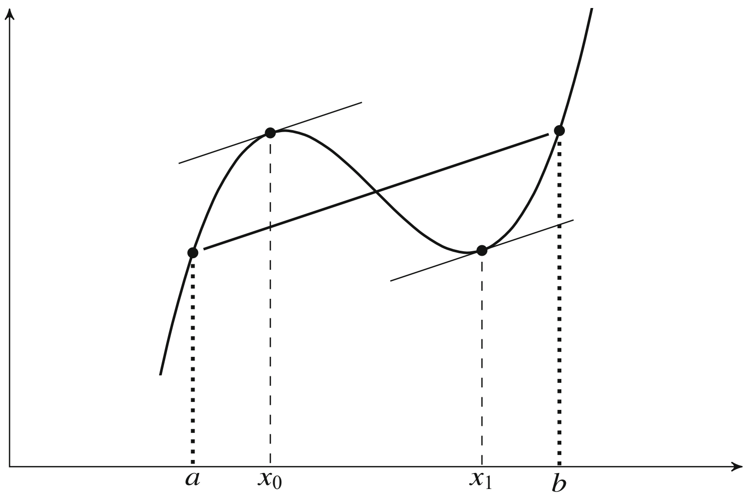

Fig. 9.2 Lagrange's theorem

Proof For any points x_1 and x_2 in (a, b), Theorem 9.8 yields $f(x_2) - f(x_1) = (x_2 - x_1) f'(c)$ for some c between x_1 and x_2. It is now immediate to conclude according to the sign of f'. □

Exercise 9.6 Suppose that f is a differentiable function such that $f'(x) = f(x)$ for each $x \in \mathbb{R}$. If $f(0) = 1$, prove that $f(x) = e^x$ for each $x \in \mathbb{R}$.

Exercise 9.7 Let f be a differentiable function on $\mathbb{R}$ with

$$L = \sup \left\{ |f'(x)| \;\middle|\; x \in \mathbb{R} \right\} < 1.$$

(a) Fix any $s_0 \in \mathbb{R}$, and define $s_n = f(s_{n-1})$ for each $n = 1, 2, \ldots$ Prove that the sequence $\{s_n\}_n$ is convergent. *Hint:* show that $\{s_n\}_n$ is a Cauchy sequence.
(b) Prove the Banach-Caccioppoli Fixed Point Theorem: there exists a point $x \in \mathbb{R}$ such that $f(x) = x$.

Mean value theorems are typically used in Calculus to derive criteria for the existence of limits. The well-known result which goes under the name of De l'Hospital is the most celebrated one.[1] We follow [2] for the proof.

Theorem 9.9 (De l'Hospital) *Suppose f and g are differentiable on (a, b), and $g'(x) \neq 0$ for all $x \in (a, b)$, where $-\infty \leq a < b \leq +\infty$. Suppose*

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = A, \tag{9.5}$$

where $A \in \tilde{\mathbb{R}}$. If either

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0 \tag{9.6}$$

or

$$\lim_{x \to a} g(x) = +\infty, \tag{9.7}$$

then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = A. \tag{9.8}$$

An analogous statement holds as $x \to b$.

We remark that A may be infinite.

[1] We write De l'Hospital since this is the ancient and original name. Nowadays it is customary to write De l'Hôpital.

Proof Let us start with the case $-\infty \leq A < +\infty$. Let $q > A$ be any real number, and choose r such that $A < r < q$. By (9.5) there exists a point $c \in (a, b)$ such that $a < x < c$ implies

$$\frac{f'(x)}{g'(x)} < r. \tag{9.9}$$

If $a < x < y < b$, Theorem 9.7 yields a point $t \in (x, y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r. \tag{9.10}$$

Suppose that (9.6) holds. When $x \to a$ in (9.10) we see that $a < y < c$ implies

$$\frac{f(y)}{g(y)} \leq r < q \tag{9.11}$$

Suppose now that (9.7) holds. We fix y in (9.10) and select $c_1 \in (a, y)$ such that $g(x) > g(y)$ and $g(x) > 0$ for every $x \in (a, c_1)$. Then it follows from (9.10) that

$$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)} \tag{9.12}$$

for every $x \in (a, c_1)$. Letting $x \to a$ in (9.12) we see that the right-hand side of (9.12) converges to r, and therefore there exists a point $c_2 \in (a, c_1)$ such that

$$\frac{f(x)}{g(x)} \leq r < q \tag{9.13}$$

for every $x \in (a, c_2)$. Equations (9.11) and (9.13) imply that $f(x)/g(x) < q$ for every $x \in (a, c_2)$.

If $-\infty < A \leq +\infty$, a completely similar argument shows that, given any $p < A$, there exists a point c_3 such that $p < f(x)/g(x)$ for every $x \in (a, c_3)$. Since p and q are arbitrary, we have proved that $f(x)/g(x) \to A$ as $x \to a$. □

Exercise 9.8 For every $x \in \mathbb{R}$, consider the functions

$$f(x) = x + \sin x \cos x$$
$$g(x) = e^{\sin x} f(x).$$

(a) Show that $\lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} g(x) = +\infty$.
(b) Show that $f'(x) = 2\cos^2 x$ and $g'(x) = e^{\sin x} \cos x\,(2\cos x + f(x))$.
(c) Show that $f'(x)/g'(x) = \frac{2e^{-\sin x} \cos x}{2\cos x + f(x)}$ if $\cos x \neq 0$ and $x > 3$.

(c) Show that $\lim_{x \to +\infty} \frac{2e^{-\sin x} \cos x}{2\cos x + f(x)} = 0$ and yet $\lim_{x \to +\infty} f(x)/g(x)$ does not exist.

This exercise shows that the assumption "$g'(x) \neq 0$" is necessary in Theorem 9.9.

Remark 9.4 Calculus books often write that De l'Hospital's theorem is a tool for the analysis of indeterminate forms $[0/0]$ and $[\infty/\infty]$. As we have seen, no condition on f is needed when $g(x) \to +\infty$ as $x \to a$.

Exercise 9.9 Suppose that $f(x) + f'(x) \to L$ as $x \to +\infty$. Prove that $f(x) \to L$ as $x \to a$ and $f'(x) \to 0$ as $x \to +\infty$. *Hint:* write $f(x) = e^x f(x)/e^x$, and remark that $e^x \to +\infty$ as $x \to +\infty$. De l'Hospital's theorem applies even if we have no information about $e^x f(x)$ as $x \to +\infty$.

9.3 The Intermediate Property for Derivatives

Although a differentiable function may have a discontinuous derivative, it is interesting that derivatives always have the intermediate value property.

Theorem 9.10 (Darboux) *Suppose that f is a real differentiable function on $[a, b]$, and suppose that $f'(a) < \lambda < f'(b)$. Then there exists a point $\xi \in (a, b)$ such that $f'(\xi) = \lambda$.*

Proof We define $g(x) = f(x) - \lambda x$. By assumption $g'(a+) < 0$ and $g'(b-) > 0$. It follows that $g(t_1) < g(a)$ and $g(t_2) < g(b)$ for some t_1, t_2 in (a, b). As a consequence, the function g must attain its minimum at some point $\xi \in (a, b)$. We already know that $g'(\xi) = 0$, and thus $f'(\xi) = \lambda$. □

Example 9.2 Define the polynomial $P(x) = (x^2 - 1)^2$. Let $f : [0, 1] \to \mathbb{R}$ be the function such that $f(0) = 0$ and

$$f(x) = \frac{1}{n^{3/2}} P\left(2n(n+1)x - 2n - 1\right) \quad \text{if } \frac{1}{n+1} \leq x \leq \frac{1}{n}.$$

Clearly f is a differentiable function, but f' is not continuous, and event not bounded on $[0, 1]$. Indeed $f'_+(0) = 0$, but $f(b_n) \to +\infty$ at the points

$$b_n = \frac{4n+1}{4n(n+1)} \to 0.$$

Nevertheless, Theorem 9.10 applies, and f' attains every positive value γ on every interval $[0, b_n]$.

9.4 Derivatives at End-Points

The idea of linearization is fruitful at *inner* points of the domain of definition: the function can be identified, at an infinitesimal scale, with a linear function. It is nonetheless convenient, from time to time, to extend the definition of derivative at end-points.

Definition 9.2 Let $f\colon [a,b] \to \mathbb{R}$ be a function. We say that f is differentiable at a if the limits

$$f'(a+) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}$$

exists as a real number. In this case, we call $f'(a+)$ the right-derivative of f at a. Similarly we define the left derivative of f at b.

Unfortunately, several fundamental results of differential calculus do not extend to end-point derivatives. As an example, we propose the function $f\colon [a,b] \to \mathbb{R}$ such that $f(x) = mx + q$, where $m \neq 0$ and q are real numbers. It is easy to check that a and b are global extremum points of f (their nature depends on the sign of m), but $f'(a+) = m = f'(b-)$ are different than zero. In other words, Fermat's Theorem does not hold at end-points.

9.5 Derivatives of Derivatives

A nice feature of derivatives in one variable is that we can easily differentiate derivatives. We will see that this requires much more attention in higher dimension, since the derivative is no longer a real number.

Definition 9.3 Suppose that a function f is defined on an interval (a,b), and that the derivative f' of f exists at every point of (a,b). Hence the function $f'\colon (a,b) \to \mathbb{R}$ is defined in such a way that $f'\colon x \mapsto f'(x)$ for every $x \in (a,b)$. We say that f is twice differentiable at $x \in (a,b)$ if f' is differentiable at x. In this case we denote by $f''(x)$ or $D^2 f(x)$ or $\mathrm{d}^2 f(x)$ the derivative of f' at x, and call it the second derivative of f at x.

More generally, if f is differentiable n times at every point of (a,b), we say that f is differentiable $(n+1)$-times at $x \in (a,b)$ if the function $f^{(n)}$ is differentiable at x. In this case we denote by $f^{(n+1)}(x)$, or $D^{n+1} f(x)$ or $\mathrm{d}^{n+1} f(x)$ the derivative of $f^{(n)}$ at x, and call it the derivative of f of order $n+1$ at x.

Definition 9.4 (Regularity Classes) Let $f\colon (a,b) \to \mathbb{R}$ be a function. We write $f \in C^0(a,b)$ if f is continuous on (a,b). For $n \in \mathbb{N}$, we write $f \in C^n(a,b)$ if the derivatives f', f'', …, $f^{(n-1)}$ exist on (a,b), and if f^n exists and is a continuous function on (a,b). We formally write $f \in C^\infty(a,b)$ to mean that

$f \in \bigcap_{n=0}^{\infty} C^n(a,b)$. Hence, a function is of class C^∞ if and only if it can be differentiated as many times as we please.

We now attach a very specific polynomial to every function with a high degree of differentiability.

Definition 9.5 (Taylor Polynomial) Let $f : (a,b) \to \mathbb{R}$ be n-times differentiable at a point $x_0 \in (a,b)$. The Taylor polynomial of degree n at x_0 is defined to be

$$\begin{aligned} P(n, x_0; x) &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \\ &= f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots \\ &\quad \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n. \end{aligned} \tag{9.14}$$

Taylor polynomials express the local behavior of functions, and generalize the concept of linear approximation which was introduced in the definition of the first derivative.

Theorem 9.11 (Local Polynomial Approximation) *Let $f : (a,b) \to \mathbb{R}$ be n-times differentiable at a point $x_0 \in (a,b)$, and let $P(n, x_0; \cdot)$ be its Taylor polynomial of degree n. Then*

$$f(x) = P(n, x_0; x) + (x-x_0)^n o(1) \quad \text{as } x \to x_0. \tag{9.15}$$

If, in addition, $f^{(n+1)}(x_0)$ exists, then

$$\lim_{x \to x_0} \frac{\zeta(x)}{x-x_0} = \frac{f^{(n+1)}(x_0)}{(n+1)!}.$$

Proof We consider the function

$$\zeta(x) = \frac{f(x) - P(n, x_0; x)}{(x-x_0)^n}, \tag{9.16}$$

defined for $x \neq x_0$. It is easy to check that all the derivatives of order $1 \leq j \leq n-1$ of the numerator of ζ vanish at x_0. We apply Theorem 9.9 $n-1$ times to (9.16), to get

$$\lim_{x \to x_0} \zeta(x) = \lim_{x \to x_0} \frac{f^{(n-1)}(x) - f^{(n-1)}(x_0) - (x-x_0) f^{(n)}(x_0)}{n!(x-x_0)}, \tag{9.17}$$

provided the last limit exists. By definition,

$$\lim_{x \to x_0} \frac{f^{(n-1)}(x) - f^{(n-1)}(x_0)}{x - x_0} = f^{(n)}(x_0),$$

thus from (9.17) we deduce $\lim_{x \to x_0} \zeta(x) = 0$. We can define $\zeta(x_0) = 0$, so that ζ becomes a continuous function on (a, b). By applying again Theorem 9.9 n times, we finally see that

$$\lim_{x \to x_0} \frac{\zeta(x)}{x - x_0} = \frac{1}{(n+1)!} \lim_{x \to x_0} \frac{f^{(n)}(x) - f^{(n)}(x_0)}{(x - x_0)} = \frac{f^{(n+1)}(x_0)}{(n+1)!}.$$

□

Theorem 9.12 (Lagrange Remainder) *Let* $f : (a, b) \to \mathbb{R}$ *be* $n + 1$ *times differentiable on* (a, b)*, and let* $x_0 \in (a, b)$*. For each* $x \in (a, b)$*,* $x \neq x_0$*, there exists a point* ξ *between* x_0 *and* x *such that*

$$f(x) = P(n, x_0; x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}.$$

Proof Suppose without loss of generality that $x_0 < x$. We define $F : [x_0, x] \to \mathbb{R}$,

$$F(t) = f(x) - f(y) - f'(t)(x - t) - \frac{f''(t)}{2!}(x - t)^2 - \cdots - \frac{f^{(n)}(t)}{n!}(x - t)^n.$$

Then

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

Next we introduce the function $G : [x_0, x] \to \mathbb{R}$,

$$G(t) = \frac{(x - t)^{n+1}}{(n+1)!}.$$

We have $F(x) = G(x) = 0$, and $F'(t)/G'(t) = f^{(n+1)}(t)$. We now apply Theorem 9.7 to F and G on $[x_0, x]$, and find a point ξ between x_0 and x such that

$$\frac{F(x_0)}{G(x_0)} = \frac{F(x) - F(x_0)}{G(x) - G(x_0)} = \frac{F'(\xi)}{G'(\xi)} = f^{(n+1)}(\xi).$$

□

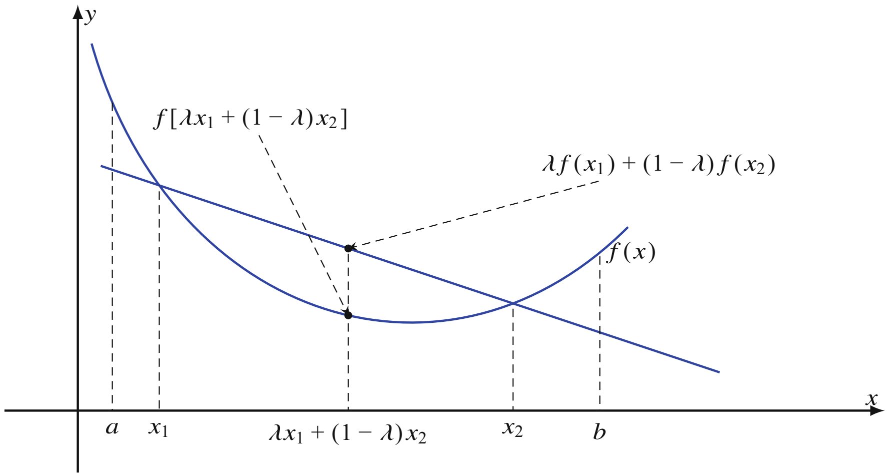

Fig. 9.3 A convex function

9.6 Convexity

Convexity is a fundamental property in mathematical analysis. Most Calculus books propose the definition of convexity as a property of the second derivative. We are going to see that this is just the top of the iceberg.

Definition 9.6 Let I be an interval[2] (open, closed, bounded or unbounded), and let $f: I \to \mathbb{R}$ be a function. We say that f is a convex function on I, if for each $x_1 \in I$, $x_2 \in I$, and for all real numbers $\lambda \geq 0$, $\mu \geq 0$ such that $\lambda + \mu = 1$, there results

$$f(\lambda x_1 + \mu x_2) \leq \lambda f(x_1) + \mu f(x_2). \tag{9.18}$$

The function f is concave on I if and only if the function $-f$ is convex on I.

Exercise 9.10 Prove that a function f is convex on an interval I if and only if for every $x_1 \in I$, $x_2 \in I$ and $\lambda \in [0, 1]$ there results

$$f\,(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda) f(x_2).$$

See Fig. 9.3.

Remark 9.5 The crucial fact is that $\lambda x_1 + \mu x_2$ must be an element of I, as soon as x_1 and x_2 belong to I and $\lambda + \mu = 1$. It is easy to check that this is actually correct, since I is an interval.

[2] Hence I is characterized by the following property: if $x \in I$, $y \in I$ and $x < z < y$, then $z \in I$.

Let us manipulate the *convexity inequality* (9.18). Let x be any point between x_1 and x_2. We set $x = \lambda x_1 + \mu x_2$. Since $\lambda + \mu = 1$, we see that

$$\lambda = \frac{x_2 - x}{x_2 - x_1}, \quad \mu = \frac{x - x_1}{x_2 - x_1}.$$

Hence (9.18) is equivalent to

$$f(x) \le \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2). \tag{9.19}$$

By symmetry, we can always suppose that $x_2 > x_1$. Then we get

$$(x_2 - x_1) f(x) \le (x_2 - x) f(x_1) + (x - x_1) f(x_2). \tag{9.20}$$

Writing $x_2 - x = (x_2 - x_1) - (x - x_1)$ in (9.20), we find

$$\frac{f(x) - f(x_1)}{x - x_1} \le \frac{f(x_2) - f(x_1)}{x_2 - x_1}. \tag{9.21}$$

Writing $x_2 - x_1 = (x_2 - x) + (x - x_1)$ in (9.20) we find

$$\frac{f(x_1) - f(x)}{x_1 - x} \le \frac{f(x_2) - f(x)}{x_2 - x}. \tag{9.22}$$

Comparing (9.18), (9.21) and (9.22), we have proved

Theorem 9.13 *The function f is convex on the interval I if and only if, for every $x_0 \in I$, the map*

$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$$

is (defined for $x \neq x_0$ and) monotonically increasing.

We thus see that convexity is just another way of stating that the incremental ratio is an increasing function. Since monotone functions always have one-sided limits, we deduce

Corollary 9.3 *A convex function f defined on an interval (a, b) is left- and right-differentiable at every $x_0 \in (a, b)$. Moreover $f'(x_0-) \le f'(x_0+)$.*

If we remember (9.22) and let $x \to x_1$ and then $x \to x_2$, we see that

$$f'(x_1+) \le \frac{f(x_2) - f(x_1)}{x_2 - x_1} \le f'(x_2-). \tag{9.23}$$

We can finally relate convexity to derivatives.

Theorem 9.14 *Let f be a differentiable function in the interval $[a, b]$. A necessary and sufficient condition for f to be convex is that f' be monotonically increasing.*

Proof If f is convex, then (9.23) implies $f'(x_1) \leq f'(x_2)$, so that f' is increasing. In the other direction, we remark that convexity is equivalent to (9.22) for all points x_1, x and x_2 such that $x_1 < x < x_2$. If f' is increasing, then there exists points ξ_1 and ξ_2 such that $x_1 < \xi_1 < x < \xi_2 < x_2$ and

$$\frac{f(x_1) - f(x)}{x_1 - x} = f'(\xi_1), \qquad \frac{f(x_2) - f(x)}{x_2 - x} = f'(\xi_2).$$

Since $f'(x_1) \leq f'(x_2)$, the conclusion follows. □

In particular, it is indeed true that convex functions are those functions whose second derivative is positive, but there is no need to restrict our definitions to twice differentiable functions.

Corollary 9.4 *Let f be twice differentiable in $[a, b]$. The function f is convex if and only if $f''(x) \geq 0$ for every x.*

Proof This follows immediately from the characterization of increasing functions in terms of the first derivative, see Corollary 9.2. □

Example 9.3 Prove that the function $x \mapsto |x|$ is convex on $I = \mathbb{R}$. Of course this conclusion would be meaningless if we had defined convex functions through the sign of the second derivative.

9.7 Problems

9.1 Suppose that f is differentiable at the point x_0. Prove that

$$\lim_{n \to +\infty} n \left[f\left(x_0 + \frac{\alpha}{n}\right) - f\left(x_0 - \frac{\beta}{n}\right) \right] = (\alpha + \beta) f'(x_0).$$

Give an example to show that the existence of the previous limit does not imply the differentiability of f at x_0.

9.2 Suppose that f is differentiable at the point x_0. let $\{h_n\}_n$ and $\{k_n\}_n$ be two nonincreasing sequences which converge to x_0. Prove that

$$\lim_{n \to +\infty} \frac{f(x_0 + h_n) - f(x_0 - k_n)}{h_n + k_n} = f'(x_0).$$

Give an example to show that the existence of the previous limit does not imply the differentiability of f at x_0.

9.3 Suppose that f' is continuous on an interval $[a, b]$. Prove that for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| < \varepsilon$$

whenever $0 < |t - x| < \delta$ and $x \in [a, b]$, $t \in [a, b]$.

9.4 Let f be differentiable on $[a, b]$. Suppose that $0 < m \leq f'(x) \leq M$ for each $x \in [a, b]$, and that $f(a) < 0 < f(b)$. Given $x_1 \in [a, b]$, define a sequence $\{x_n\}_n$ by

$$x_{n+1} = x_n - \frac{f(x_n)}{M}$$

for $n = 1, 2, 3, \ldots$ Prove that $\{x_n\}_n$ converges to a limit x_0 such that $f(x_0) = 0$. Furthermore, prove that

$$|x_{n+1} - x_n| \leq \frac{f(x_1)}{m} \left(1 - \frac{m}{M}\right)^n .$$

9.5 Suppose f is a real-valued function defined on the half-line $(a, +\infty)$. Suppose that f is twice differentiable on $(a, +\infty)$, and define

$$M_0 = \sup_{x>a} |f(x)|$$

$$M_1 = \sup_{x>a} |f'(x)|$$

$$M_2 = \sup_{x>a} |f''(x)|.$$

Prove that $M_1^2 \leq M_0 M_2$ as follows: for each $h > 0$ deduce from Taylor's expansion that there exists $\xi \in (x, x + 2h)$ such that

$$f'(x) = \frac{f(x + 2h) - f(x)}{2h} - hf'(\xi).$$

Therefore

$$|f'(x)| \leq hM_1 + \frac{M_0}{h}.$$

Now optimize the right-hand side with respect to $h > 0$.

9.6 If f is twice differentiable on $(0, +\infty)$, f'' is bounded and $\lim_{x \to +\infty} f(x) = 0$, prove that $\lim_{x \to +\infty} f'(x) = 0$. *Hint:* consider the limit $a \to +\infty$ in the previous problem.

9.7 (a) Prove that for each $x > 0$, $x \neq 1$, we have

$$\frac{x-1}{x} < \log x < x - 1.$$

(b) For each $j \in \mathbb{N}$, $j > 1$, prove that

$$\log \frac{j+1}{j} < \frac{1}{j} < \log \frac{j}{j-1}.$$

(c) For each $n \in \mathbb{N}$, $k \in \mathbb{N}$, $n > 1$, prove that

$$\log\left(k + \frac{1}{n}\right) < \sum_{j=n}^{kn} \frac{1}{j} < \log\left(k + \frac{k}{n-1}\right)$$

and

$$\lim_{n\to+\infty} \sum_{j=n}^{kn} \frac{i}{j} = \log k.$$

(d) Deduce from (c) and from the identity

$$\sum_{j=1}^{2n} \frac{(-1)^{j+1}}{j} = \sum_{j=1}^{2n} \frac{1}{j} - 2\sum_{k=1}^{n} \frac{1}{2k}$$

that

$$\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} = \log 2.$$

9.8 If $x > 1$, $x \neq e$, prove that there exists one and only one number $f(x) > 0$ such that $f(x) \neq x$ and

$$x^{f(x)} = (f(x))^x.$$

Hint: $x^y = y^x$ if and only if $\frac{\log x}{x} = \frac{\log y}{y}$.

9.8 Comments

The standard definition of derivative as the limit of the incremental ratio should not be considered as the one used by mathematicians from the beginning of Calculus. They would rather use a principle of disappearing quantities which roughly correspond to an expansion of functions at first order as in

$$f(t+\Delta t) = f(t) + f'(t)\Delta t + f''(t)(\Delta t)^2 \approx f(t) + f'(t)\Delta t.$$

In this sense, the one-dimensional derivative has progressively lost its definition as a linearization procedure in favor of an iconic limit:

$$f'(t) = \lim_{\Delta t \to 0} \frac{f(t+\Delta t) - f(t)}{\Delta t}.$$

There are several good reasons to define the derivative as a linearization, and the most important one is that the derivative of a function of several variables is not a number.

The theory of convex functions is a long but elementary exercise, in the case of functions of a single real variable. The topic becomes much more exciting in higher dimensions, where intervals must be replaced by convex sets and a new fact comes into play: it is possible to draw conclusion about a function on a convex set by assuming a property of that function on every straight line. The interested reader may start from [1].

References

1. R.T. Rockafellar, *Convex Analysis* (Princeton University Press, 1970)
2. W. Rudin, *Principles of Mathematical Analysis*. International Series in Pure and Applied Mathematics, 3rd edn. (McGraw-Hill Book Co., New York, 1976)