Chapter 12
Differentiation

12.1 The Definition of Differentiability

Consider a point that is moving on a line, and let $s(t)$ denote the location of the point on the line at time t. Back when we talked about real-life problems that could lead to the definition of limits (see Chapter 9, p. 121), we saw that the definition of instantaneous velocity required taking the limit of the fraction $\big(s(t)-s(t_0)\big)/(t-t_0)$ in t_0. Having precisely defined what a limit is, we can now *define the instantaneous velocity of the point at a t_0 to be the limit*

$$\lim_{t\to t_0}\frac{s(t)-s(t_0)}{t-t_0}$$

(assuming, of course, that this limit exists and is finite).

We also saw that if we want to define the tangent line to the graph of a function f at $\big(a,f(a)\big)$, then the slope of that line is exactly the limit of $\big(f(x)-f(a)\big)/(x-a)$ in a. *We agree to define the tangent line as the line that contains the point $\big(a,f(a)\big)$ and has slope*

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

again assuming that this limit exists and is finite.

Other than the two examples above, many problems in mathematics, physics, and other fields can be grasped in the same form as above. This is the case when we have to find the rate of some change (not necessarily happening in space). If, for example, the temperature of an object at time t is $H(t)$, then we can ask how fast the temperature is changing at time t_0. The average change over the interval $[t_0,t]$ is $\big(H(t)-H(t_0)\big)/(t-t_0)$. Clearly, the instantaneous change of temperature will be defined as the limit

$$\lim_{t\to t_0}\frac{H(t)-H(t_0)}{t-t_0}$$

(assuming that it exists and is finite).

© Springer New York 2015
M. Laczkovich, V.T. Sós, *Real Analysis*, Undergraduate Texts in Mathematics, DOI 10.1007/978-1-4939-2766-1_12

We use the following names for the quotients appearing above. If f is defined at the points a and b, then the quotient $\big(f(b)-f(a)\big)/(b-a)$ is called the **difference quotient** of f between a and b. It is clear that the difference quotient $\big(f(b)-f(a)\big)/(b-a)$ agrees with the slope of the line passing through the points $\big(a,f(a)\big)$ and $\big(b,f(b)\big)$.

In many cases, using the notation $b-a=h$, the difference quotient between a and $x=a+h$ is written as

$$\frac{f(a+h)-f(a)}{h}.$$

Definition 12.1. Let f be defined on a neighborhood of the point a. We say that the function f is *differentiable at the point* a if the finite limit

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a} \tag{12.1}$$

exists. The limit (12.1) is called the *derivative* of f at a.

The derivative of f at a is most often denoted by $f'(a)$. Sometimes, other notations,

$$\dot{f}(a)\,,\qquad \left.\frac{df}{dx}\right|_{x=a},\qquad \left.\frac{df(x)}{dx}\right|_{x=a},\qquad y'(a)\,,\qquad \left.\frac{dy}{dx}\right|_{x=a}$$

are used (the last two in accordance with the notation $y=f(x)$).

Equipped with the above definition, we can say that if the function $s(t)$ describes the location of a moving point, then the instantaneous velocity at time t_0 is $s'(t_0)$. Similarly, if the temperature of an object at time t is given by $H(t)$, then the instantaneous temperature change at time t_0 is $H'(t_0)$. The definition of a tangent should also be updated with our new definition. The tangent of f at a is the line that passes through the point $\big(a,f(a)\big)$ and has slope

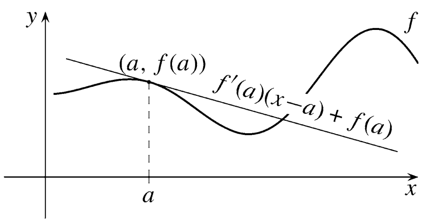

Fig. 12.1

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}=f'(a).$$

Since the equation of this line is $y=f'(a)\cdot(x-a)+f(a)$, we accept the following definition.

Definition 12.2. Let f be differentiable at the point a. The *tangent line of f at* $\big(a,f(a)\big)$ is the line with equation $y=f'(a)\cdot(x-a)+f(a)$.

The visual meaning of the derivative $f'(a)$ is then the slope of the tangent of graph f at $(a, f(a))$ (Figure 12.1).

Examples 12.3. **1.** The constant function $f(x) = c$ is differentiable at all values a, and its derivative is zero. This is because

$$\frac{f(x)-f(a)}{x-a} = \frac{c-c}{x-a} = 0$$

for all $x \neq a$.
2. The function $f(x) = x$ is differentiable at all values a, and $f'(a) = 1$. This is because

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

for all $x \neq a$.
3. The function $f(x) = x^2$ is differentiable at all values a, and $f'(a) = 2a$. This is because

$$\frac{f(x)-f(a)}{x-a} = \frac{x^2-a^2}{x-a} = x+a,$$

and so

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

Thus by the definition of tangent lines, the tangent of the parabola $y = x^2$ at the point (a, a^2) is the line with equation $y = 2a(x-a) + a^2 = 2ax - a^2$. Since this passes through the point $(a/2, 0)$, we can construct the tangent by drawing a line connecting the point $(a/2, 0)$ to the point (a, a^2) (Figure 12.2).[1]

Differentiability is a stronger condition than continuity. This is shown by the theorem below and the remarks following it.

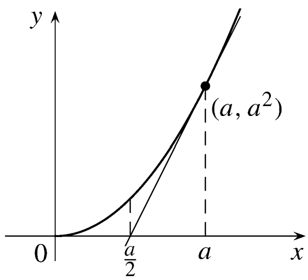

Fig. 12.2

Theorem 12.4. *If f is differentiable at a, then f is continuous at a.*

Proof. If f is differentiable at a, then

$$\lim_{x\to a}(f(x)-f(a)) = \lim_{x\to a}\left[\frac{f(x)-f(a)}{x-a}\cdot(x-a)\right] = f'(a)\cdot 0 = 0.$$

This means exactly that f is continuous at a. □

Remarks 12.5. **1.** Continuity is a necessary but not sufficient condition for differentiability. There exist functions that are continuous at a point a but are not differentiable there. An easy example is the function $f(x) = |x|$ at $a = 0$. Clearly,

[1] So the calculus was correct; see page 5.

$$f(x) = \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0, \end{cases}$$

so

$$\frac{f(x) - f(0)}{x - 0} = \begin{cases} 1, & \text{if } x > 0, \\ -1, & \text{if } x < 0. \end{cases}$$

Then

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

and so f is not differentiable at 0.
2. There even exist functions that are continuous everywhere but not differentiable anywhere. Using the theory of series of functions, one can show that for suitable $a, b > 0$, the function

$$f(x) = \sum_{n=1}^{\infty} a^n \cdot \sin(b^n x)$$

has this property (for example with the choice $a = 1/2,\ b = 10$). Similarly, one can show that the function

$$g(x) = \sum_{n=1}^{\infty} \frac{\langle 2^n x \rangle}{2^n}$$

is everywhere continuous but nowhere differentiable, where $\langle x \rangle$ denotes the smallest distance from x to an integer.
3. There even exists a function that is differentiable at a point a but is not continuous at any other point. For example, the function

$$f(x) = \begin{cases} x^2, & \text{if } x \text{ is rational}, \\ -x^2, & \text{if } x \text{ is irrational} \end{cases}$$

is differentiable at 0. This is because

$$\left| \frac{f(x) - f(0)}{x - 0} \right| = \left| \frac{x^2}{x} \right| = |x| \to 0, \qquad \text{if } x \to 0,$$

so

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

At the same time, it can easily be seen—see function 6 in Example 10.2—that the function $f(x)$ is not continuous at any $x \neq 0$.

As we saw, the function $|x|$ is not differentiable at 0, and accordingly, the graph does not have a tangent at 0 (here the graph has a cusp). If, however, we look only on the right-hand side of the point, then the difference quotient has a limit from that side. Accordingly, in the graph of $|x|$, the "right-hand chords" of $(0,0)$ have a limit (which is none other than the line $y = x$). The situation is similar on the left-hand side of 0.

As the above example illustrates, it is reasonable to introduce the one-sided variants of derivatives (and hence one-sided differentiability).

Definition 12.6. If the finite limit

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

exists, then we call this limit the *right-hand derivative of f at a.* We can similarly define the left-hand derivative as well.

We denote the right-hand derivative at a by $f'_+(a)$, and the left-hand derivative at a by $f'_-(a)$.

Remark 12.7. It is clear that f is differentiable at a if and only if both the right- and left-hand derivatives of f exist at a, and $f'_+(a) = f'_-(a)$. (Then this shared value is $f'(a)$.)

As we have seen, differentiability does not follow from continuity. We will now show that convexity—which is a stronger property than continuity—implies one-sided differentiability.

Theorem 12.8. *If f is convex on the interval (a,b), then f is differentiable from the left and from the right at all points $c \in (a,b)$.*

Proof. By Theorem 9.20, the function $x \mapsto (f(x)-f(c))/(x-c)$ is monotone increasing on the set $(a,b)\setminus\{c\}$. Fix $d \in (a,c)$. Then

$$\frac{f(d)-f(c)}{d-c} \le \frac{f(x)-f(c)}{x-c}$$

for all $x \in (c,b)$, so the function $(f(x)-f(c))/(x-c)$ is monotone increasing and bounded from below on the interval (c,b).

By statement (ii) of Theorem 10.68, it then follows that the limit

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

exists and is finite, which means that f is differentiable from the right at c. It can be shown similarly that f has a left-hand derivative at c. □

Linear Approximation. In working with a function that arises in a problem, we can frequently get a simpler and more intuitive result if instead of the function, we use another simpler one that "approximates the original one well." One of the simplest classes of functions comprises the linear functions ($y = mx+b$). We show that the differentiability of a function f at a point a means exactly that the function can be "well approximated" by a linear function. As we will soon see, the best linear approximation for f at a point a is the function $y = f'(a)(x-a)+f(a)$.

If f is continuous at a, then for all c,

$$f(x)-[c\cdot(x-a)+f(a)] \to 0, \quad \text{if} \quad x \to a.$$

Thus every linear function $\ell(x)$ such that $\ell(a) = f(a)$ "approximates f well" in the sense that $f(x) - \ell(x) \to 0$ as $x \to a$. The differentiability of f at a, by the theorem below, means exactly that the function

$$t(x) = f'(a) \cdot (x-a) + f(a) \tag{12.2}$$

approximate significantly better: not only does the difference $f - t$ tend to zero as $x \to a$, but it tends to zero faster than $(x-a)$.

Theorem 12.9. *The function f is differentiable at a if and only if at a, it can be "well approximated" locally by a linear polynomial in the following sense: there exists a number α (independent of x) such that*

$$f(x) = \alpha \cdot (x-a) + f(a) + \varepsilon(x) \cdot (x-a),$$

where $\varepsilon(x) \to 0$ as $x \to a$. The number α is the derivative of f at a.

Proof. Suppose first that f is differentiable at a, and let

$$\varepsilon(x) = \frac{f(x) - f(a)}{x-a} - f'(a).$$

Since

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x-a},$$

we have $\varepsilon(x) \to 0$ as $x \to a$. Thus

$$f(x) = f(a) + f'(a)(x-a) + \varepsilon(x)(x-a),$$

where $\varepsilon(x) \to 0$ as $x \to a$.

Now suppose that

$$f(x) = \alpha \cdot (x-a) + f(a) + \varepsilon(x)(x-a),$$

where $\varepsilon(x) \to 0$ as $x \to a$. Then

$$\frac{f(x) - f(a)}{x-a} = \alpha + \varepsilon(x) \to \alpha, \quad \text{if} \quad x \to a.$$

Thus f is differentiable at a, and $f'(a) = \alpha$. □

The following theorem expresses that $t(x) = f'(a) \cdot (x-a) + f(a)$ is the "best" linear function approximating f.

Theorem 12.10. *If the function f is differentiable at a, then for all $c \neq f'(a)$,*

$$\lim_{x \to a} \frac{f(x) - [f'(a)(x-a) + f(a)]}{f(x) - [c(x-a) + f(a)]} = 0.$$

Proof. As $x \to a$,

$$\frac{f(x) - [f'(a)(x-a) + f(a)]}{f(x) - [c(x-a) + f(a)]} = \frac{\frac{f(x)-f(a)}{x-a} - f'(a)}{\frac{f(x)-f(a)}{x-a} - c} \to \frac{f'(a) - f'(a)}{f'(a) - c} = 0.$$

□

Theorem 12.9 gives a necessary and sufficient condition for f to be differentiable at a. With the help of this, we can give another (equivalent to (12.1)) definition of differentiability.

Definition 12.11. The function f is *differentiable* at a if there exists a number α (independent of x) such that

$$f(x) = \alpha \cdot (x-a) + f(a) + \varepsilon(x)(x-a),$$

where $\varepsilon(x) \to 0$ if $x \to a$.

The significance of this equivalent definition lies in the fact that if we want to extend the concept of differentiability to other functions—not necessarily real-valued or of a real variable—then we cannot always find a definition analogous to Definition 12.1, but the generalization of Definition 12.11 is generally feasible.

Derivative Function. We will see that the derivative is the most useful tool in investigating the properties of a function. This is true locally and globally. The existence of the derivative $f'(a)$ and its value describes local properties of f: from the value of $f'(a)$, we can deduce the behavior of f at a.[2]

If, however, f is differentiable at every point of an interval, then from the values of $f'(x)$ we get global properties of f. In applications, we mostly come upon functions that are differentiable in some interval. The definition of this is the following.

Definition 12.12. Let $a < b$. We say that f is *differentiable on the interval* (a,b) if it is differentiable at every point of (a,b). We say that f is *differentiable on* $[a,b]$ if it is differentiable on (a,b) and differentiable from the right at a and from the left at b.

Generally, we can think of differentiation as an operation that maps functions to functions.

Definition 12.13. The *derivative function* of the function f is the function that is defined at every x where f is differentiable and has the value $f'(x)$ there. It is denoted by f'.

The basic task of differentiation is to find relations between functions and their derivatives, and to apply them. For the applications, however, first we need to decide where a function is differentiable, and we have to find its derivative. We begin with this latter problem, and only thereafter do we inspect what derivatives tell us about functions and how to use that information.

[2] One such link was already outlined when we saw that differentiability implies continuity.

Exercises

12.1. Where is the function $(\{x\}-1/2)^2$ differentiable?

12.2. Let $f(x)=x^2$ if $x\le 1$, and $f(x)=ax+b$ if $x>1$. For what values of a and b will f be differentiable everywhere?

12.3. Let $f(x)=|x|^{\alpha}\cdot\sin|x|^{\beta}$ if $x\neq 0$, and let $f(0)=0$. For what values of α and β will f be continuous at 0? When will it be differentiable at 0?

12.4. Prove that the graph of the function x^2 has the tangent line $y=mx+b$ if the line and the graph of the function intersect at exactly one point.

12.5. Where are the tangent lines to the function $2x^3-3x^2+8$ horizontal?

12.6. When is the x-axis tangent to the graph of x^3+px+q?

12.7. Let $f(2^{-n})=3^{-n}$ for all positive integers n, and let $f(x)=0$ otherwise. Where is f differentiable?

12.8. Are there any points where the Riemann function is differentiable?

12.9. Are there any points where the square of the Riemann function is differentiable? (H)

12.10. At what angle does the graph of x^2 intersect the line $y=2x$? (That is, what is the angle between the tangent of the function and the line?)

12.11. Prove that the function $f(x)=\sqrt{x}$ is differentiable for all $a>0$, and that $f'(a)=1/(2\sqrt{a})$.

12.12. Prove that the function $1/x$ is differentiable at all points $a>0$, and find its derivative. Show that every tangent of the function $1/x$ forms a triangle with the two axes whose area does not depend on which point the tangent is taken at.

12.13. Prove that if f is differentiable at 0, then the function $f(|x|)$ is differentiable at 0 if and only if $f'(0)=0$.

12.14. Prove that if f is differentiable at a, then

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

Show that the converse of this statement does not hold.

12.15. Let f be differentiable at the point a. Prove that if $x_n<a<y_n$ for all n and if $y_n-x_n\to 0$, then

$$\lim_{n\to\infty}\frac{f(y_n)-f(x_n)}{y_n-x_n}=f'(a).\ \text{(H S)}$$

12.16. Suppose that f is differentiable everywhere on $(-\infty,\infty)$. Prove that if f is even (odd), then f' is odd (even).

12.17. Let
$B = \{f\colon f \text{ is bounded on } [a,b]\}$,
$C = \{f\colon f \text{ is continuous on } [a,b]\}$,
$M = \{f\colon f \text{ is monotone on } [a,b]\}$,
$X = \{f\colon f \text{ is convex on } [a,b]\}$,
$D = \{f\colon f \text{ is differentiable on } [a,b]\}$,
$I = \{f\colon f \text{ has an inverse on } [a,b]\}$.
From the point of containment, how are the sets B, C, M, X, D, and I related?

12.2 Differentiation Rules and Derivatives of the Elementary Functions

The differentiability of some basic functions can be deduced from the properties we have seen. These include the polynomials, trigonometric functions, and logarithmic functions. On the other hand, it is easy to see that all of the elementary functions can be expressed in terms of polynomials, trigonometric functions, and logarithmic functions using the four basic arithmetic operations, taking inverses, and composition. Thus to determine the differentiability of the remaining elementary functions, we need theorems that help us deduce their differentiability and to calculate their derivatives based on the differentiability and derivatives of the component functions in terms of which they can be expressed. These are called the **differentiation rules.** Below, we will determine the derivatives of the power functions with integer exponent, trigonometric functions, and logarithmic functions; we will introduce the differentiation rules, and using all this information, we will determine the derivatives of the rest of the elementary functions.

Theorem 12.14. *For an arbitrary positive integer n, the function x^n is differentiable everywhere on $(-\infty,\infty)$, and $(x^n)' = n\cdot x^{n-1}$ for all x.*

Proof. For all a,

$$\lim_{x\to a}\frac{x^n-a^n}{x-a} = \lim_{x\to a}\left(x^{n-1}+x^{n-2}\cdot a+\cdots+x\cdot a^{n-2}+a^{n-1}\right) = na^{n-1}$$

by the continuity of x^k. □

Theorem 12.15.

(i) *The functions* $\sin x$ *and* $\cos x$ *are differentiable everywhere on* $(-\infty,\infty)$. *Moreover,*

$$(\sin x)' = \cos x \ \ \textit{and} \ \ (\cos x)' = -\sin x$$

for all x.

(ii) *The function* $\operatorname{tg} x$ *is differentiable at all points* $x \neq \frac{\pi}{2} + k\pi$ $(k \in \mathbb{Z})$*, and at those points,*

$$(\operatorname{tg} x)' = \frac{1}{\cos^2 x}.$$

(iii) *The function* $\operatorname{ctg} x$ *is differentiable for all* $x \neq k\pi$ $(k \in \mathbb{Z})$*, and at those points,*

$$(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x}.$$

Proof. (i) For arbitrary $a \in \mathbb{R}$ and $x \neq a$,

$$\frac{\sin x - \sin a}{x-a} = \frac{2\sin(x-a)/2 \cdot \cos(x+a)/2}{x-a} = \frac{\sin((x-a)/2)}{(x-a)/2} \cdot \cos\frac{x+a}{2}$$

by the second identity of (11.37). Now $\lim_{x \to a} \sin((x-a)/2)/((x-a)/2) = 1$ and $\lim_{x \to a} \cos(x+a)/2 = \cos a$ by (11.45), by the continuity of the cos function and the theorem about limits of compositions of functions. Therefore,

$$\lim_{x \to a} \frac{\sin x - \sin a}{x-a} = \cos a.$$

Similarly, for arbitrary $a \in \mathbb{R}$ and $x \neq a$,

$$\frac{\cos x - \cos a}{x-a} = -\frac{2\sin(x-a)/2 \cdot \sin(x+a)/2}{x-a} = -\frac{\sin((x-a)/2)}{(x-a)/2} \cdot \sin\frac{x+a}{2}$$

by the fourth identity of (11.37). Then using (11.45) and the continuity of the sin function, we get that

$$\lim_{x \to a} \frac{\cos x - \cos a}{x-a} = -\sin a.$$

(ii) For arbitrary $a \neq (\pi/2) + k\pi$ and $x \neq a$,

$$\frac{\operatorname{tg} x - \operatorname{tg} a}{x-a} = \left(\frac{\sin x}{\cos x} - \frac{\sin a}{\cos a}\right) \cdot \frac{1}{x-a} = \frac{\sin x \cos a - \sin a \cos x}{\cos x \cos a} \cdot \frac{1}{x-a} =$$
$$= \frac{\sin(x-a)}{x-a} \cdot \frac{1}{\cos x \cos a}.$$

Then using (11.45) and the continuity of cos, we get that

$$\lim_{x \to a} \frac{\operatorname{tg} x - \operatorname{tg} a}{x-a} = \frac{1}{\cos^2 a}.$$

(iii) For arbitrary $a \neq k\pi$ and $x \neq a$,

$$\frac{\operatorname{ctg} x - \operatorname{ctg} a}{x-a} = \left(\frac{\cos x}{\sin x} - \frac{\cos a}{\sin a}\right) \cdot \frac{1}{x-a} = \frac{\cos x \sin a - \cos a \sin x}{\sin x \sin a} \cdot \frac{1}{x-a} =$$
$$= -\frac{\sin(x-a)}{x-a} \cdot \frac{1}{\sin x \sin a},$$

which gives us

$$\lim_{x \to a} \frac{\operatorname{ctg} x - \operatorname{ctg} a}{x-a} = -\frac{1}{\sin^2 a}.$$

□

Theorem 12.16. *If $a > 0$ and $a \neq 1$, then the function $\log_a x$ is differentiable at every point $x > 0$, and*

$$(\log_a x)' = \frac{1}{\log a} \cdot \frac{1}{x}. \tag{12.3}$$

Proof. By Corollary 11.11,

$$\lim_{h \to 0} \left(1 + \frac{h}{x}\right)^{1/h} = e^{1/x}$$

for all $x > 0$. Thus

$$\lim_{h \to 0} \frac{\log_a(x+h) - \log_a x}{h} = \lim_{h \to 0} \log_a \left(1 + \frac{h}{x}\right)^{1/h} = \log_a e^{1/x} =$$
$$= \frac{1}{x} \cdot \log_a e = \frac{1}{\log a} \cdot \frac{1}{x}.$$

□

It is clear that if $a > 0$ and $a \neq 1$, then the function $\log_a |x|$ is differentiable on the set $\mathbb{R} \setminus \{0\}$, and $(\log_a |x|)' = 1/(x \cdot \log a)$ for all $x \neq 0$.

We now turn our attention to introducing the **differentiation rules.** As we will see, a large portion of these are consequences of the definition of the derivative and theorems about limits.

Theorem 12.17. *If the functions f and g are differentiable at a, then cf ($c \in \mathbb{R}$), $f+g$, and $f \cdot g$ are also differentiable at a, and*

(i) $(cf)'(a) = cf'(a)$,
(ii) $(f+g)'(a) = f'(a) + g'(a)$,
(iii) $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.

If $g(a) \neq 0$, then $1/g$ and f/g are differentiable at a. Moreover,

(iv) $\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g^2(a)}$,
(v) $\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}$.

Proof. The shared idea of the proofs of these is to express the difference quotients of each function with the help of the difference quotients $(f(x)-f(a))/(x-a)$ and $(g(x)-g(a))/(x-a)$:

(i) The difference quotient of the function $F = cf$ is

$$\frac{F(x)-F(a)}{x-a} = \frac{cf(x)-cf(a)}{x-a} = c\cdot\frac{f(x)-f(a)}{x-a} \to c\cdot f'(a).$$

(ii) The difference quotient of $F = f+g$ is

$$\frac{F(x)-F(a)}{x-a} = \frac{(f(x)+g(x))-(f(a)+g(a))}{x-a} = \frac{f(x)-f(a)}{x-a} + \frac{g(x)-g(a)}{x-a}.$$

Thus

$$\lim_{x\to a}\frac{F(x)-F(a)}{x-a} = f'(a)+g'(a).$$

(iii) The difference quotient of $F = f\cdot g$ is

$$\frac{F(x)-F(a)}{x-a} = \frac{f(x)\cdot g(x)-f(a)\cdot g(a)}{x-a} = \\ = \frac{f(x)-f(a)}{x-a}\cdot g(x) + f(a)\cdot\frac{g(x)-g(a)}{x-a}.$$

Since $g(x)$ is differentiable at a, it is continuous there (by Theorem 12.4), and so

$$\lim_{x\to a}\frac{F(x)-F(a)}{x-a} = \lim_{x\to a}\frac{f(x)-f(a)}{x-a}\cdot\lim_{x\to a} g(x) + f(a)\cdot\lim_{x\to a}\frac{g(x)-g(a)}{x-a} = \\ = f'(a)g(a)+f(a)g'(a).$$

If $g(a)\neq 0$, then by the continuity of g, it follows that $g(x)\neq 0$ on a neighborhood of a; that is, the functions $1/g(x)$ and $f(x)/g(x)$ are defined here.

(iv) The difference quotient of $F = 1/g$ is

$$\frac{F(x)-F(a)}{x-a} = \frac{1/g(x)-1/g(a)}{x-a} = \frac{g(a)-g(x)}{g(a)g(x)}\cdot\frac{1}{x-a} = \\ = -\frac{1}{g(x)g(a)}\cdot\frac{g(x)-g(a)}{x-a} \to -\frac{1}{g^2(a)}\cdot g'(a).$$

(v) The difference quotient of $F = f/g$ is

$$\frac{F(x)-F(a)}{x-a} = \frac{f(x)/g(x)-f(a)/g(a)}{x-a} = \frac{1}{g(a)g(x)}\,\frac{f(x)g(a)-f(a)g(x)}{x-a} = \\ = \frac{1}{g(x)g(a)}\left(\frac{f(x)-f(a)}{x-a}\cdot g(x) - \frac{g(x)-g(a)}{x-a}\cdot f(x)\right),$$

which implies

$$\lim_{x\to a} \frac{F(x)-F(a)}{x-a} = \frac{f'(a)g(a)-f(a)g'(a)}{g^2(a)}.$$

□

Remarks 12.18. **1.** The statements of the theorem hold for right- and left-hand derivatives as well.
2. Let I be an interval, and suppose that the functions $f\colon I\to\mathbb{R}$ and $g\colon I\to\mathbb{R}$ are differentiable on I as stated in Definition 12.12. By the above theorem, it follows that cf, $f+g$, and $f\cdot g$ are also differentiable on I, and the equalities

$$(cf)' = cf',\ (f+g)' = f'+g',\ (fg)' = f'g+fg'$$

hold. If we also suppose that $g\neq 0$ on I, then $1/g$ and f/g are also differentiable on I, and

$$\left(\frac{1}{g}\right)' = -\frac{g'}{g^2},\quad \text{moreover}\quad \left(\frac{f}{g}\right)' = \frac{f'g-fg'}{g^2}.$$

We emphasize that here we are talking about equality of *functions,* not simply numbers.
3. From the above theorem, we can use induction to easily prove the following statements.

If the functions $f_1,\dots,f_n$ are differentiable at a, then

(i) $f_1+\cdots+f_n$ is differentiable at a, and

$$(f_1+\cdots+f_n)'(a) = f_1'(a)+\cdots+f_n'(a);$$

moreover,

(ii) $f_1\cdot\ldots\cdot f_n$ is differentiable at a, and

$$(f_1\cdot\ldots\cdot f_n)'(a) = \left(f_1'\cdot f_2\cdots f_n + f_1\cdot f_2'\cdot f_3\cdots f_n+\cdots+f_1\cdots f_{n-1}\cdot f_n'\right)(a).$$

4. By (ii) above, it follows that if $f_1(a)\cdot\ldots\cdot f_n(a)\neq 0$, then

$$\left(\frac{(f_1\cdot\cdots\cdot f_n)'}{f_1\cdot\cdots\cdot f_n}\right)(a) = \left(\frac{f_1'}{f_1}+\cdots+\frac{f_n'}{f_n}\right)(a).$$

Thus if $f_1,\dots,f_n$ are defined on the interval I, are nowhere zero, and are differentiable on the interval I, then

$$\frac{(f_1\cdot\cdots\cdot f_n)'}{f_1\cdot\cdots\cdot f_n} = \frac{f_1'}{f_1}+\cdots+\frac{f_n'}{f_n}. \tag{12.4}$$

The next theorem is known as the **chain rule.**

Theorem 12.19. *If the function g is differentiable at a and the function f is differentiable at $g(a)$, then the function $h = f \circ g$ is differentiable at a, and*

$$h'(a) = f'(g(a)) \cdot g'(a).$$

With the notation $y = g(x)$ and $z = f(y)$, the statement of the theorem can easily be remembered in the form

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

This formula is where the name "chain rule" comes from.

Proof. By the assumptions, it follows that the function h is defined on a neighborhood of the point a. Indeed, f is defined on a neighborhood V of $g(a)$. Since g is differentiable at a, it is continuous at a, so there exists a neighborhood U of a such that $g(x) \in V$ for all $x \in U$. Thus h is defined on U. After these preparations, we give two proofs of the theorem.

I. Following the proof of the previous theorem, let us express the difference quotient of h using the difference quotients of f and g. Suppose first that $g(x) \neq g(a)$ on a punctured neighborhood of a. Then

$$\frac{h(x)-h(a)}{x-a} = \frac{f(g(x))-f(g(a))}{x-a} = \frac{f(g(x))-f(g(a))}{g(x)-g(a)} \cdot \frac{g(x)-g(a)}{x-a}. \tag{12.5}$$

Since g is continuous at a, if $x \to a$, then $g(x) \to g(a)$. Thus by the theorem on the limits of compositions of functions,

$$\lim_{x\to a} \frac{h(x)-h(a)}{x-a} = \lim_{t\to g(a)} \frac{f(t)-f(g(a))}{t-g(a)} \cdot \lim_{x\to a} \frac{g(x)-g(a)}{x-a} = f'(g(a)) \cdot g'(a).$$

The proof of the special case above used twice the fact that $g(x) \neq g(a)$ on a punctured neighborhood of a: first, when we divided by $g(x) - g(a)$, and second, when we applied Theorem 10.41 for the limit of the composition. Recall that the assumption of Theorem 10.41 requires that the inner function shouldn't take on its limit in a punctured neighborhood of the place, unless the outer function is continuous. This, however, does not hold for us, since the outer function is the difference quotient $(f(t) - f(g(a)))/(t - g(a))$, which is not even defined at $g(a)$.

Exactly this circumstance gives an idea for the proof in the general case. Define the function $F(t)$ as follows: let

$$F(t) = \frac{f(t)-f(g(a))}{t-g(a)}$$

if $t \in V$ and $t \neq g(a)$, and let $F(t) = f'(g(a))$ if $t = g(a)$. Then F is continuous at $g(a)$, so by Theorem 10.41, $\lim_{x\to a} F(g(x)) = f'(g(a))$. To finish the proof, it suffices to show that

$$\frac{h(x)-h(a)}{x-a} = F(g(x)) \cdot \frac{g(x)-g(a)}{x-a} \tag{12.6}$$

for all $x \in U$. We distinguish two cases. If $g(x) \neq g(a)$, then (12.6) is clear from (12.5). If, however, $g(x) = g(a)$, then $h(x) = f(g(x)) = f(g(a)) = h(a)$, so both sides of (12.6) are zero. Thus the proof is complete.

II. This proof is based on the definition of differentiability given in 12.11. According to this, the differentiability of f at $g(a)$ means that

$$f(t) - f(g(a)) = f'(g(a))(t - g(a)) + \varepsilon_1(t)(t - g(a)) \tag{12.7}$$

for all $t \in V$, where $\varepsilon_1(t) \to 0$ if $t \to g(a)$.

Let $\varepsilon_1(g(a)) = 0$. Similarly, by the differentiability of the function g, it follows that

$$g(x) - g(a) = g'(a)(x - a) + \varepsilon_2(x)(x - a) \tag{12.8}$$

for all $x \in U$, where $\varepsilon_2(x) \to 0$ as $x \to a$. If we substitute $t = g(x)$ in (12.7), and then apply (12.8), we get that

$$\begin{aligned} h(x) - h(a) = f(g(x)) - f(g(a)) &= \\ &= f'(g(a))(g(x) - g(a)) + \varepsilon_1(g(x))(g(x) - g(a)) = \\ &= f'(g(a))g'(a)(x - a) + \varepsilon(x)(x - a), \end{aligned}$$

where

$$\varepsilon(x) = f'(g(a))\varepsilon_2(x) + \varepsilon_1(g(x))(g'(a) + \varepsilon_2(x)).$$

Since $g(x) \to g(a)$ as $x \to a$, we have $\varepsilon_1(g(x)) \to 0$ as $x \to a$ (since by $\varepsilon_1(g(a)) = 0$, ε_1 is continuous at $g(a)$). Then from $\varepsilon_2(x) \to 0$, we get that $\varepsilon(x) \to 0$ if $x \to a$. Now this, by Definition 12.11, means exactly that the function h is differentiable at a, and $h'(a) = f'(g(a))g'(a)$. □

The following theorem gives the differentiation rule for inverse functions.

Theorem 12.20. *Let f be strictly monotone and continuous on the interval (a,b), and let φ denote the inverse of f. If f is differentiable at the point $c \in (a,b)$ and $f'(c) \neq 0$, then φ is differentiable at $f(c)$, and*

$$\varphi'(f(c)) = \frac{1}{f'(c)}.$$

Proof. The function φ is defined on the interval $J = f((a,b))$. By the definition of the inverse function, $\varphi(f(c)) = c$ and $f(\varphi(y)) = y$ for all $y \in J$. Let $F(x)$ denote the difference quotient $(f(x) - f(c))/(x - c)$. If $y \neq f(c)$, then

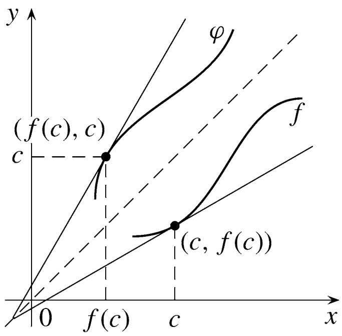

Fig. 12.3

$$\frac{\varphi(y)-\varphi(f(c))}{y-f(c)} = \frac{\varphi(y)-c}{f(\varphi(y))-f(c)} = \\ = \frac{1}{F(\varphi(y))}. \tag{12.9}$$

Since φ is strictly monotone, if $y \neq f(c)$, then $\varphi(y) \neq c$. Thus we can apply Theorem 10.41 on the limits of compositions of functions. We get that

$$\varphi'(f(c)) = \lim_{y\to f(c)} \frac{\varphi(y)-\varphi(f(c))}{y-f(c)} = \lim_{y\to f(c)} \frac{1}{F(\varphi(y))} = \lim_{x\to c} \frac{1}{F(x)} = \frac{1}{f'(c)}.$$

□

Remarks 12.21. **1.** The statement of the theorem can be illustrated with the following geometric argument. The graphs of the functions f and φ are the mirror images of each other in the line $y = x$. The mirror image of the tangent to graph f at the point $(c, f(c))$ gives us the tangent to graph φ at $(f(c), c)$. The slopes of these are the reciprocals of each other, that is, $\varphi'(f(c)) = 1/f'(c)$ (Figure 12.3).
2. Let f be strictly monotone and continuous on the interval (a,b), and suppose that f is differentiable everywhere on (a,b). If f' is nonvanishing everywhere, then by the above theorem, φ is everywhere differentiable on the interval $J = f((a,b))$, and $\varphi'(f(x)) = 1/f'(x)$ for all $x \in (a,b)$. If $y \in J$, then $\varphi(y) \in I$ and $f(\varphi(y)) = y$. Thus $\varphi'(y) = 1/f'(\varphi(y))$. Since this holds for all $y \in J$, we have that

$$\varphi' = \frac{1}{f' \circ \varphi}. \tag{12.10}$$

3. If $f'(c) = 0$ (that is, if the tangent line to graph f is parallel to the x-axis at the point $(c, f(c))$), then by (12.9), it is easy to see that the difference quotient $(\varphi(y) - \varphi(f(c)))/(y - f(c))$ does not have a finite limit at $f(c)$. Indeed, in this case, the limit of the difference quotient $F(x) = (f(x) - f(c))/(x - c)$ at c is zero, so $\lim_{y\to f(c)} F(\varphi(y)) = 0$. If, however, f is strictly monotone increasing, then the difference quotient $F(x)$ is positive everywhere, and so $\lim_{y\to f(c)} 1/F(\varphi(y)) = \infty$ (see Remark 10.39). Then

$$\lim_{y\to f(c)} \frac{\varphi(y)-\varphi(f(c))}{y-f(c)} = \infty,$$

and we can similarly get that if f is strictly monotone decreasing, then the value of the above limit is $-\infty$. (These observations agree with the fact that if $f'(c) = 0$, then the tangent line to graph φ at $(f(c), c)$ is parallel to the y-axis.) This remark motivates the following extension of the definition of the derivative.

Definition 12.22. Let f be defined on a neighborhood of a point a. We say that the *derivative of f at a is infinite* if

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \infty,$$

and we denote this by $f'(a) = \infty$. We define $f'(a) = -\infty$ similarly.

Remark 12.23. Definitions 12.1 and 12.22 can be stated jointly as follows: if the limit (12.1) exists and has the value β (which can be finite or infinite), then we say that **the derivative of** f **at** a **exists,** and we use the notation $f'(a) = \beta$. We emphasize that the "differentiable" property is reserved for the cases in which the derivative is finite. Thus a function f is differentiable at a if and only if its derivative there exists and is *finite.*

We extend the definition of the one-sided derivatives (Definition 12.6) for the cases in which the one-sided limits of the difference quotient are infinite. We use the notation $f'_+(a) = \infty$, $f'_-(a) = \infty$, $f'_+(a) = -\infty$, and $f'_-(a) = -\infty$; their meaning is straightforward.

Using the concepts above, we can extend Theorem 12.20 as follows.

Theorem 12.24. *Let f be strictly monotone and continuous on the interval (a,b). Let φ denote the inverse function of f. If $f'(c) = 0$, then φ has a derivative at $f(c)$, and in fact, $\varphi'(f(c)) = \infty$ if f is strictly monotone increasing, and $\varphi'(f(c)) = -\infty$ if f is strictly monotone decreasing.*

Now let us return to the elementary functions. With the help of the differentiation rules, we can now determine the derivatives of all of them.

Theorem 12.25. *If $a > 0$, then the function a^x is differentiable everywhere, and*

$$(a^x)' = \log a \cdot a^x \tag{12.11}$$

for all x.

Proof. The statement is clear for $a = 1$, so we can suppose that $a \neq 1$. Since $(\log_a x)' = 1/(x \cdot \log a)$, by the differentiation rule for inverse functions, we get that

$$(a^x)' = a^x \cdot \log a.$$

$\square$

We note that the differentiability of the function a^x can also be deduced easily from Theorem 12.8.

Applying Theorems 12.16 and 12.25 for $a = e$, we get the following.

Theorem 12.26.

(i) *For all x,*

$$(e^x)' = e^x. \tag{12.12}$$

(ii) *For all $x > 0$,*

$$(\log x)' = \frac{1}{x}. \tag{12.13}$$

According to (12.11), the function e^x *is the only exponential function that is the derivative of itself.* This fact motivates us to consider e to be one of the most important constants[3] of analysis (and of mathematics, more generally). By equality (12.13), out of all the logarithmic functions, the derivative of the logarithm with base e is the simplest. This is why we chose the logarithm with base e from among all the other logarithm functions (see Remark 11.17).

Using the derivatives of the exponential and logarithmic functions, we can easily find the derivatives of the power functions.

Theorem 12.27. *For arbitrary $b \in \mathbb{R}$, the function x^b is differentiable at all points $x > 0$, and*

$$(x^b)' = b \cdot x^{b-1}. \tag{12.14}$$

Proof. Since $x^b = e^{b \log x}$ for all $x > 0$, we can apply the differentiation rule for compositions of functions, Theorem 12.19. □

The derivatives of the inverse trigonometric functions can easily be found by the differentiation rule for inverse functions.

Theorem 12.28.

(i) *The function* $\arcsin x$ *is differentiable on the interval* $(-1,1)$*, and*

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} \tag{12.15}$$

for all $x \in (-1,1)$.

(ii) *The function* $\arccos x$ *is differentiable on the interval* $(-1,1)$*, and*

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}} \tag{12.16}$$

for all $x \in (-1,1)$.

(iii) *The function* $\operatorname{arc\,tg} x$ *is differentiable everywhere, and*

$$(\operatorname{arc\,tg} x)' = \frac{1}{1+x^2} \tag{12.17}$$

for all x.

(iv) *The function* $\operatorname{arcctg} x$ *is differentiable everywhere, and*

$$(\operatorname{arcctg} x)' = -\frac{1}{1+x^2} \tag{12.18}$$

for all x.

Proof. The function $\sin x$ is strictly increasing and differentiable on $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, and its derivative is $\cos x$. Since $\cos x \neq 0$ if $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, by Theorem 12.20, if $x \in (-1,1)$, then

$$(\arcsin x)' = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-\sin^2(\arcsin x)}} = \frac{1}{\sqrt{1-x^2}},$$

which proves (i). Statement (ii) follows quite simply from (i) and (11.50).

[3] The other central constant is π. The relation between these two constants is given by $e^{i\pi} = -1$, which is a special case of the identity (11.64).

The function $\operatorname{tg} x$ is strictly increasing on the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, and its derivative is $1/\cos^2 x \neq 0$ there. Thus

$$(\operatorname{arc\,tg} x)' = \cos^2(\operatorname{arc\,tg} x).$$

However,

$$\cos^2 x = \frac{1}{1+\operatorname{tg}^2 x},$$

so

$$\cos^2(\operatorname{arc\,tg} x) = \frac{1}{1+x^2},$$

which establishes (iii). Statement (iv) is clear from (iii) and (11.52). □

By the definition of the hyperbolic functions and (12.12), the assertions of the following theorem, which strengthen their link to the trigonometric functions, are clear once again.

Theorem 12.29.

(i) *The functions* $\operatorname{sh} x$ *and* $\operatorname{ch} x$ *are differentiable everywhere on* $(-\infty, \infty)$, *and*

$$(\operatorname{sh} x)' = \operatorname{ch} x \ \textit{ and } \ (\operatorname{ch} x)' = \operatorname{sh} x$$

for all x.

(ii) *The function* $\operatorname{th} x$ *is differentiable everywhere on* $(-\infty, \infty)$. *Moreover,*

$$(\operatorname{th} x)' = \frac{1}{\operatorname{ch}^2 x}$$

for all x.

(iii) *The function* $\operatorname{cth} x$ *is differentiable at all points* $x \neq 0$, *and there,*

$$(\operatorname{cth} x)' = -\frac{1}{\operatorname{sh}^2 x}.$$

Finally, consider the inverse hyperbolic functions.

Theorem 12.30.

(i) *The function* $\operatorname{arsh} x$ *is differentiable everywhere, and*

$$(\operatorname{arsh} x)' = \frac{1}{\sqrt{x^2+1}} \tag{12.19}$$

for all x.

(ii) *The function* $\operatorname{arch} x$ *is differentiable on the interval* $(1, \infty)$, *and*

$$(\operatorname{arch} x)' = \frac{1}{\sqrt{x^2-1}} \tag{12.20}$$

for all $x > 1$.

(iii) *The function* $\operatorname{arth} x$ *is differentiable everywhere on* $(-1,1)$*, and*

$$(\operatorname{arth} x)' = \frac{1}{1-x^2} \tag{12.21}$$

for all $x \in (-1,1)$.

Proof. (i) Since the function $\operatorname{sh} x$ is strictly increasing on $\mathbb{R}$ and its derivative there is $\operatorname{ch} x \neq 0$, the function $\operatorname{arsh} x$ is differentiable everywhere. The derivative can be computed either from Theorem 12.20 or by the identity (11.59).
(ii) Since the function $\operatorname{ch} x$ is strictly increasing on $(0,\infty)$ and its derivative there is $\operatorname{sh} x \neq 0$, the function $\operatorname{arch} x$ is differentiable on the interval $(1,\infty)$.

We leave the proofs of (12.20) and statement (iii) to the reader. □

Remark 12.31. It is worth noting that the derivatives of the functions $\log x$, $\operatorname{arc\,tg} x$, and $\operatorname{arth} x$ are rational functions, and the derivatives of $\arcsin x$, $\arccos x$, $\operatorname{arsh} x$, and $\operatorname{arch} x$ are algebraic functions. (The definition of an algebraic function can be found in Exercise 11.45.)

Exercises

12.18. Suppose $f+g$ is differentiable at a, and g is not differentiable at a. Can f be differentiable at a?

12.19. Let $f(x) = x^2 \cdot \sin(1/x)$, $f(0) = 0$. Prove that f is differentiable everywhere. (S)

12.20. Prove that if $0 < c < 1$, then the right-hand derivative of x^c at 0 is infinity.

12.21. Prove that if n is a positive odd integer, then the derivative of $\sqrt[n]{x}$ at 0 is infinity.

12.22. Where is the tangent line of the function $\sqrt[3]{\sin x}$ vertical?

12.23. Prove that the graphs of the functions $\sqrt{4a(a-x)}$ and $\sqrt{4b(b+x)}$ cross each other at right angles, that is, the tangent lines at the intersection point are perpendicular. (S)

12.24. Prove that the curves $x^2 - y^2 = a$ and $xy = b$ cross each other at right angles. That is, the graphs of the functions $\pm\sqrt{x^2-a}$ and b/x cross each other perpendicularly.

12.25. At what angle do the graphs of the functions 2^x and $(\pi - e)^x$ cross each other? (S)

12.26. Give a closed form for $x+2x^2+\cdots+nx^n$. (Hint: differentiate the function $1+x+\cdots+x^n$.) Use this to compute the sums

$$\frac{1}{2}+\frac{2}{4}+\frac{3}{8}+\cdots+\frac{n}{2^n} \qquad \text{and} \qquad \frac{1}{3}+\frac{2}{9}+\frac{3}{27}+\cdots+\frac{n}{3^n}.$$

12.27. Let $f(x)=x\cdot(x+1)\cdot\ldots\cdot(x+100)$, and let $g=f\circ f\circ f$. Compute the value of $g'(0)$.

12.28. Prove that the function x^x is differentiable for all $x>0$, and compute its derivative.

12.29. The function x^x is strictly monotone on $[1,\infty)$. What is the value of the derivative of its inverse at the point 27?

12.30. The function x^5+x^2 is strictly monotone on $[0,\infty)$. What is the value of the derivative of its inverse at the point 2?

12.31. Prove that the function $x+\sin x$ is strictly monotone increasing. What is the value of the derivative of its inverse at the point $1+(\pi/2)$?

12.32. Let $f(x)=\log_x 3$ $(x>0,\ x\neq 1)$. Compute the derivatives of f and f^{-1}.

12.33. Let us apply differentiation to find limits. The method consists in changing the function being considered into a difference quotient and finding its limit through differentiation. For example, instead of

$$\lim_{x\to 0}(x+e^x)^{1/x},$$

we can take its logarithm to get the quotient

$$\frac{\log(x+e^x)}{x},$$

which is the difference quotient of the numerator at 0. The limit of the quotient is thus the derivative of the numerator at 0. If this limit is A, then the original limit is e^A. Finish this computation.

12.34. Apply the method above, or one of its variants, to find the following limits:

(a) $\lim_{x\to 0}(\cos x)^{1/\sin x}$,

(b) $\lim_{x\to 0}\left(\frac{e^x+1}{2}\right)^{1/\operatorname{sh} x}$,

(c) $\lim_{x\to 0}\frac{\operatorname{sh}^2 x}{\log\cos 3x}$,

(d) $\lim_{x\to 1}(2-x)^{1/\cos(\pi/(2x))}$,

(e) $\lim\limits_{x\to\infty}\left(x^{1/x}-1\right)\cdot\frac{x}{\log x}$.

12.35. Prove, using the method above, that if $a_1, \ldots, a_n > 0$, then

$$\lim_{x \to 0} \sqrt[x]{\frac{a_1^x + \cdots + a_n^x}{n}} = \sqrt[n]{a_1 \cdot \ldots \cdot a_n}.$$

12.36. Let T_n denote the nth Chebyshev polynomial (see Exercise 11.32). Prove that if $T_n(a) = 0$, then $|T_n'(a)| = n/\sqrt{1-a^2}$.

12.37. Let f be convex on the open interval I.

(a) Prove that the function $f'_+(x)$ is monotone increasing on I.
(b) Prove that if the function $f'_+(x)$ is continuous at a point x_0, then f is differentiable at x_0.
(c) Prove that the set $\{x \in I : f \text{ is not differentiable at } x\}$ is countable.

12.3 Higher-Order Derivatives

Definition 12.32. Let the function f be differentiable in a neighborhood of a point a. If the derivative function f' has a derivative at a, then we call the derivative of f' at a the *second derivative* of f. We denote this by $f''(a)$. Thus

$$f''(a) = \lim_{x \to a} \frac{f'(x) - f'(a)}{x - a}.$$

If $f''(a)$ exists and is finite, then we say that f is twice differentiable at a. The *second derivative function* of f, denoted by f'', is the function defined for the points x at which f is twice differentiable, and its value there is $f''(x)$.

We can define the kth-order derivatives by induction:

Definition 12.33. Let the function f be $k-1$ times differentiable in a neighborhood of the point a. Let the $(k-1)$th derivative function of f be denoted by $f^{(k-1)}$. The derivative of $f^{(k-1)}$ at a, if it exists, is called the *kth (order) derivative* of f. The kth derivative function is denoted by $f^{(k)}$; this is defined where f is k times differentiable.

The kth derivative at a can be denoted by the symbols

$$\left.\frac{d^k f}{dx^k}\right|_{x=a}, \qquad \left.\frac{d^k f(x)}{dx^k}\right|_{x=a}, \qquad y^{(k)}(a), \qquad \left.\frac{d^k y}{dx^k}\right|_{x=a}$$

as well. To keep our notation consistent, we will sometimes use the notation

$$f^{(0)} = f, \quad f^{(1)} = f', \quad f^{(2)} = f''$$

too.

If $f^{(k)}$ exists for all $k \in \mathbb{N}^+$ at a, then we say that f is *infinitely differentiable at a.*

It is easy to see that if p is an nth-degree polynomial, then its kth derivative is an $(n-k)$th-degree polynomial for all $k \le n$. Thus the nth derivative of the polynomial p is constant, and the kth derivative is identically zero for $k > n$. It follows that *every polynomial is infinitely differentiable.* With the help of higher-order derivatives, we can easily determine the multiplicity of a root of a polynomial.

Theorem 12.34. *The number a is a root of the polynomial p with multiplicity k if and only if*

$$p(a) = p'(a) = \ldots = p^{(k-1)}(a) = 0 \ \text{ and } \ p^{(k)}(a) \neq 0. \tag{12.22}$$

Proof. Clearly, it is enough to show that if a is a root of multiplicity k, then (12.22) holds (since, for different k's the statements (12.22) exclude each other). We prove this by induction. If $k=1$, then $p(x) = (x-a)\cdot q(x)$, where $q(a) \neq 0$. Then $p'(x) = q(x) + (x-a)\cdot q'(x)$, which gives $p'(a) = q(a) \neq 0$, so (12.22) holds for $k=1$.

Let $k > 1$, and suppose that the statement holds for $k-1$. Since $p(x) = (x-a)^k \cdot q(x)$, where $q(a) \neq 0$, we have

$$p'(x) = k\cdot(x-a)^{k-1}\cdot q(x) + (x-a)^k \cdot q'(x) = (x-a)^{k-1}\cdot r(x),$$

where $r(a) \neq 0$. Then the number a is a root of the polynomial p' with multiplicity $k-1$, so by the induction hypothesis,

$$p'(a) = p''(a) = \ldots = p^{(k-1)}(a) = 0 \ \text{ and } \ p^{(k)}(a) \neq 0.$$

Since $p(a) = 0$ is also true, we have proved (12.22). □

Some of the differentiation rules apply for higher-order derivatives as well. Out of these, we will give those for addition and multiplication. To define the rule for multiplication, we need to introduce **binomial coefficients.**

Definition 12.35. If $0 \le k \le n$ are integers, then the number $\dfrac{n!}{k!(n-k)!}$ is denoted by $\binom{n}{k}$, where $0!$ is defined to be 1.

By the definition, it is clear that $\binom{n}{0} = \binom{n}{n} = 1$ for all n. It is also easy to check that

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \tag{12.23}$$

for all $n \ge 2$ and $k = 1, \ldots, n-1$.

Theorem 12.36 (Binomial Theorem). *The following identity holds:*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k. \tag{12.24}$$

The name of the theorem comes from the fact that a *binomial* (that is, a polynomial with two terms) appears on the left-hand side of (12.24).

Proof. We prove this by induction. The statement is clear for $n = 1$. If it holds for n, then

$$(a+b)^{n+1} = (a+b)^n \cdot (a+b) = \\ = \left[a^n + \binom{n}{1} a^{n-1} \cdot b + \cdots + \binom{n}{n-1} a \cdot b^{n-1} + b^n\right] \cdot (a+b).$$

If we multiply this out, then in the resulting sum, the terms a^{n+1} and b^{n+1} appear, and moreover, for all $1 \le k \le n$, the terms $\binom{n}{k-1} a^{n-k+1} \cdot b^k$ and $\binom{n}{k} a^{n-k+1} \cdot b^k$ also appear. The sum of these two terms, according to (12.23), is exactly $\binom{n+1}{k} a^{n-k+1} \cdot b^k$. Thus we get the identity (12.24) (with $n+1$ replacing n). □

Theorem 12.37. *If f and g are n times differentiable at a, then $f+g$ and $f \cdot g$ are also n times differentiable there, and*

$$(f+g)^{(n)}(a) = f^{(n)}(a) + g^{(n)}(a), \tag{12.25}$$

as well as

$$(f \cdot g)^{(n)}(a) = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)}(a) \cdot g^{(k)}(a). \tag{12.26}$$

The identity (12.26) is called the **Leibniz rule.**

Proof. (12.25) is straightforward by induction. We also use induction to prove (12.26). If $n = 1$, then the statement is clear by the differentiation rule for products.

Suppose that the statement holds for n. If f and g are $n+1$ times differentiable at a, then they are n times differentiable in a neighborhood U of a, so by the induction hypothesis,

$$(f \cdot g)^{(n)} = f^{(n)} \cdot g + \binom{n}{1} f^{(n-1)} \cdot g' + \cdots + \binom{n}{n-1} f' \cdot g^{(n-1)} + f \cdot g^{(n)} \tag{12.27}$$

in U. Using the differentiation rules for sums and products, we get that $(f \cdot g)^{(n+1)}(a)$ is the sum of the terms $f^{(n+1)}(a) \cdot g(a)$ and $f(a) \cdot g^{(n+1)}(a)$, as well as terms of the form $\binom{n}{k-1} f^{(n-k+1)}(a) \cdot g^{(k)}(a)$ and $\binom{n}{k} f^{(n-k+1)}(a) \cdot g^{(k)}(a)$ for all $k = 1, \ldots, n$. These sum to $\binom{n+1}{k} f^{(n-k+1)}(a) \cdot g^{(k)}(a)$, which shows that (12.26) holds for $n+1$. □

The higher-order derivatives of some elementary functions are easy to compute.

Examples 12.38. **1.** It is easy to see that the exponential function a^x is infinitely differentiable, and that

$$(a^x)^{(n)} = (\log a)^n \cdot a^x \tag{12.28}$$

for all n.

2. It is also easy to check that the power function x^b is infinitely differentiable on the interval $(0, \infty)$, and that

$$(x^b)^{(n)} = b(b-1) \cdot \ldots \cdot (b-n+1) \cdot x^{b-n} \tag{12.29}$$

for all n and $x > 0$.

3. The functions $\sin x$ and $\cos x$ are also infinitely differentiable, and their higher-order derivatives are

$$(\sin x)^{(2n)} = (-1)^n \cdot \sin x, \qquad (\sin x)^{(2n+1)} = (-1)^n \cdot \cos x,$$
$$(\cos x)^{(2n)} = (-1)^n \cdot \cos x, \qquad (\cos x)^{(2n+1)} = (-1)^{n+1} \cdot \sin x \tag{12.30}$$

for all n and x.

Remark 12.39. The equalities $(\sin x)'' = -\sin x$ and $(\cos x)'' = -\cos x$ can be expressed by saying that the functions $\sin x$ and $\cos x$ satisfy the relation

$$y'' + y = 0;$$

this means that if we write $\sin x$ or $\cos x$ in place of y, then we get an equality. Such a relation that links the derivatives of a function to the function itself (possibly using other known functions) is called a **differential equation.** A differential equation is said to have order n if the highest-order derivative that appears in the differential equation is the nth. So we can say that the functions $\sin x$ and $\cos x$ satisfy the second-order differential equation $y'' + y = 0$. More specifically, we will say that this differential equation is an **algebraic differential equation**, since only the basic operations are applied to the function and its derivatives.

It is clear that the exponential function a^x satisfies the first-order differential equation $y' - \log a \cdot y = 0$. But it is not immediately clear that every exponential function satisfies the *same* differential equation. Indeed, if $y = a^x$, then $y'/y = \log a$, which is constant. Thus $(y'/y)' = 0$, that is, a^x satisfies the second-order algebraic differential equation

$$y'' \cdot y - (y')^2 = 0.$$

We can similarly show that every power function satisfies the same second-order algebraic differential equation. The function x will appear in this differential equation, but can be removed by increasing the order (see Exercise 12.43).

The function $\log_a x$ satisfies the equation $y' - (\log a \cdot x)^{-1} = 0$. From this, we get that $x \cdot y'$ is a constant, that is,

$$x \cdot y'' + y' = 0.$$

It is easy to see that the logarithmic functions satisfy a single third-order algebraic differential equation, in which x does not appear. The inverse trigonometric and hyperbolic functions satisfy similar equations.

One can show that if two functions both satisfy an algebraic differential equation, then their sum, product, quotient, and composition also satisfy an algebraic differential equation (a different one, generally more complicated). It follows that *every elementary function satisfies an algebraic differential equation* (which, of course, depends on the function).

In the next chapter, we will discuss differential equations in more detail.

Exercises

12.38. Prove that if f is twice differentiable at a, then f is continuous in a neighborhood of a.

12.39. How many times is the function $|x|^3$ differentiable at 0?

12.40. Give a function that is k times differentiable at 0 but is not $k+1$ times differentiable there.

12.41. Prove that for the Chebyshev polynomial $T_n(x)$ (see Exercise 11.31),

$$(1-x^2)T_n''(x) - x\,T_n'(x) + n^2 T_n(x) = 0$$

for all x.

12.42. Prove that the **Legendre polynomial**

$$P_n(x) = \frac{1}{2^n\, n!}\left((x^2-1)^n\right)^{(n)}$$

satisfies

$$(1-x^2)P_n''(x) - 2x\,P_n'(x) + n(n+1)P_n(x) = 0$$

for all x.

12.43. (a) Prove that every power function satisfies a second-order algebraic differential equation.
(b) Prove that every power function satisfies a single third-order algebraic differential equation that does not contain x.

12.44. Prove that the logarithmic functions satisfy a single third-order algebraic differential equation that does not contain x. (S)

12.45. Prove that each of the functions $\arcsin x$, $\arccos x$, $\operatorname{arctg} x$, $\operatorname{arsh} x$, $\operatorname{arch} x$, and $\operatorname{arth} x$ (individually) satisfies a third-degree algebraic differential equation that does not contain x.

12.46. Prove that the function $e^x + \log x$ satisfies an algebraic differential equation.

12.47. Prove that the function $e^x \cdot \sin x$ satisfies an algebraic differential equation.

12.4 Linking the Derivative and Local Properties

Definition 12.40. Let the function f be defined on a neighborhood of the point a. We say that f is *locally increasing at a* if there exists a $\delta > 0$ such that $f(x) \le f(a)$ for all $a - \delta < x < a$, and $f(x) \ge f(a)$ for all $a < x < a + \delta$.

Let f be defined on a right-hand neighborhood of a. We say that f is *locally increasing on the right at a* if there exists a $\delta > 0$ such that $f(x) \ge f(a)$ for all $a < x < a + \delta$ (Figure 12.4).

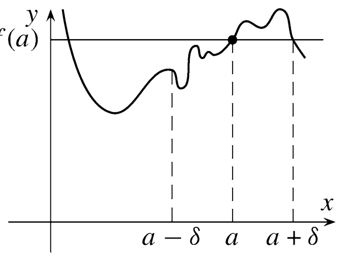

Fig. 12.4

We similarly define the concepts of *strictly locally increasing, locally decreasing,* and *strictly locally decreasing at a,* as well as *(strictly) locally increasing and decreasing from the left.*

Remark 12.41. We have to take care in distinguishing the concepts of local and monotone increasing (or decreasing) functions. The precise link between the two is the following.

On the one hand, it is clear that if f is monotone increasing on (a,b), then f is locally increasing for all points in (a,b).

On the other hand, it can be shown that if f is locally increasing at every point in (a,b), then it is monotone increasing in (a,b) (but since we will not need this fact, we leave the proof of it as an exercise; see Exercise 12.54).

It is possible, however, for a function f to be locally increasing at a point a but not be monotone increasing on any neighborhood $U(a)$. Consider the following examples.

1. The function
$$f(x) = \begin{cases} x \cdot \sin^2(1/x), & \text{if } x \ne 0, \\ 0, & \text{if } x = 0 \end{cases}$$
is locally increasing at 0, but it does not have a neighborhood of 0 in which f is monotone (Figure 12.5).
2. The function
$$f(x) = \begin{cases} 1/x, & \text{if } x \ne 0, \\ 0, & \text{if } x = 0 \end{cases}$$
is strictly locally increasing at 0, but it is not monotone increasing in any interval. In fact, in the *right-hand* punctured neighborhoods of 0, f is strictly monotone decreasing.

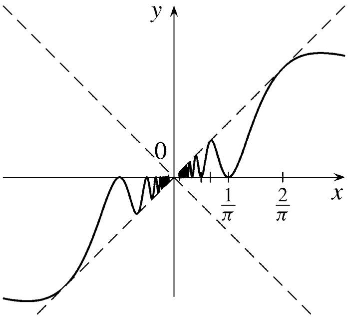

Fig. 12.5

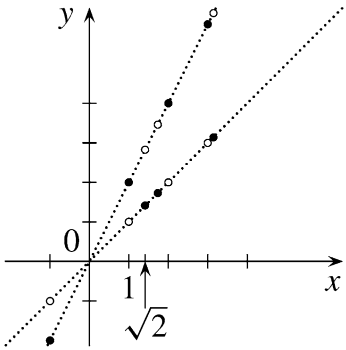

Fig. 12.6

3. Similarly, the function

$$f(x) = \begin{cases} \operatorname{tg} x, & \text{if } x \in (0,\pi) \setminus \left\{\frac{\pi}{2}\right\}, \\ 0, & \text{if } x = \frac{\pi}{2} \end{cases}$$

is strictly locally decreasing at $\frac{\pi}{2}$ but is strictly monotone increasing on the intervals $(0,\pi/2)$ and $(\pi/2,\pi)$.

4. The function

$$f(x) = \begin{cases} x, & \text{if } x \text{ is irrational}, \\ 2x, & \text{if } x \text{ is rational} \end{cases}$$

is strictly locally increasing at 0, but there is no interval on which it is monotone (Figure 12.6).

Definition 12.42. We say that the function f has a *local maximum* (or *minimum*) at a if a has a neighborhood U in which f is defined and for all $x \in U$, $f(x) \le f(a)$ (or $f(x) \ge f(a)$). We often refer to the point a itself as the *local maximum (or minimum) of the function.*

If for all $x \in U \setminus \{a\}$, $f(x) < f(a)$ (or $f(x) > f(a)$), then we say that a is a *strict local maximum (or minimum).*

Local maxima and local minima are collectively called *local extrema*.

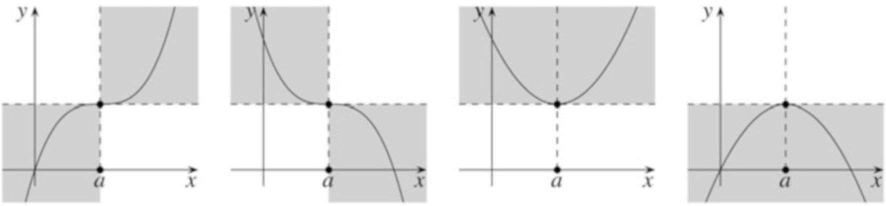

Fig. 12.7

Remark 12.43. We defined absolute (global) extrema in Definition 10.54. The following connections exist between absolute and local extrema.

An absolute extremum is not necessarily a local extremum, since a condition for a point to be a local extremum is that the function be defined in a neighborhood of the point. So for example, the function x on the interval $[0,1]$ has an absolute minimum at 0, but this is not a local minimum. However, if the function $f: A \to \mathbb{R}$ has an absolute extremum at $a \in A$, and A contains a neighborhood of a, then a is a local extremum.

A local extremum is not necessarily an absolute extremum, since the fact that f does not have a value larger than $f(a)$ in a neighborhood of a does not prevent it from having a larger value outside the neighborhood.

Consider the following three properties:

I. The function f is locally increasing at a.
II. The function f is locally decreasing at a.
III. a is a local extremum of the function f.

The function f satisfies one of the properties I, II, and III if and only if there exists a $\delta > 0$ such that the graph of f over the interval $(a-\delta, a)$ lies entirely in one section of the plane separated by the horizontal line $y = f(a)$ and the graph of f over $(a, a+\delta)$ lies in one of the sections separated by $y = f(a)$. The four possibilities correspond to the function at a being locally increasing, locally decreasing, a local maximum, and a local minimum (Figure 12.7).

It is clear that properties I and II can hold simultaneously only if f is constant in a neighborhood of a.

In the strict variants of the properties above, however, only one can hold at one time.

Of course, it is possible that none of I, II, and III holds. This is the case, for example, with the function

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

at the point $x = 0$.

Let us now see the connection between the sign of the derivative and the properties above.

Theorem 12.44. *Suppose that f is differentiable at a.*

(i) *If $f'(a) > 0$, then f is strictly locally increasing at a.*
(ii) *If $f'(a) < 0$, then f is strictly locally decreasing at a.*
(iii) *If f is locally increasing at a, then $f'(a) \geq 0$.*
(iv) *If f is locally decreasing at a, then $f'(a) \leq 0$.*
(v) *If f has a local extremum at a, then $f'(a) = 0$.*

Proof. (i) If

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} > 0,$$

then by Theorem 10.30, there exists a $\delta > 0$ such that

$$\frac{f(x) - f(a)}{x - a} > 0$$

for all $0 < |x - a| < \delta$. Thus $f(x) > f(a)$ if $a < x < a + \delta$, and $f(x) < f(a)$ if $a - \delta < x < a$. But this means precisely that the function f is strictly locally increasing at a (Figure 12.8). Statement (ii) can be proved similarly.
(iii) If f is locally increasing at a, then there exists a $\delta > 0$ such that

$$\frac{f(x) - f(a)}{x - a} \geq 0$$

if $0 < |x - a| < \delta$. But then

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

Statement (iv) can be proved similarly.
(v) If $f'(a) \neq 0$, then by (i) and (ii), f is either locally strictly increasing or locally strictly decreasing at a, so a cannot be a local extremum. Thus if f has a local extremum at a, then necessarily $f'(a) = 0$. □

Remarks 12.45. **1.** The one-sided variants of the statements (i)–(iv) above also hold (and can be proved the same way). That is, if $f'_+(a) > 0$, then f is strictly locally increasing on the right at a; if f is locally increasing on the right at a, then $f'_+(a) \geq 0$, assuming that the right-hand derivative exists.

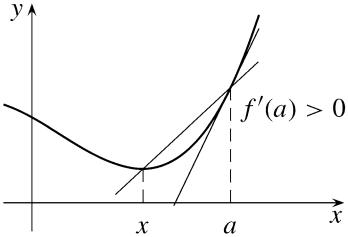

Fig. 12.8

2. *None of the converses of the statements (i)–(v) is true.* If we know only that f is strictly locally increasing at a, cannot deduce that $f'(a) > 0$. For example, the function $f(x) = x^3$ is strictly locally increasing at 0 (and in fact strictly monotone increasing on the whole real line), but $f'(0) = 0$.

If we know only that $f'(a) \geq 0$, we cannot deduce that the function f is locally increasing at a. For example, for the function $f(x) = -x^3$, $f'(0) = 0 \geq 0$, but f is not locally increasing at 0 (and in fact, it is strictly locally decreasing there).

Similarly, if we know only that $f'(a) = 0$, cannot deduce that the function f has a local extremum at a. For example, for the function $f(x) = x^3$, $f'(0) = 0$, but f does not have a local extremum at 0 (since x^3 is strictly monotone increasing on the entire real line). We can also express this by saying that if f is differentiable at a, then the assumption $f'(a) = 0$ is a necessary *but not sufficient* condition for f to have a local extremum at a.

3. If in statement (iii), we assume f to be strictly locally increasing instead, then we still cannot say more than $f'(a) \geq 0$ generally (since the converse of statement (i) does not hold).

4. From $f'(a) > 0$, we can deduce only that f is locally increasing at a, and not that it is monotone increasing. Consider the following example. Let f be a function such that $x - x^2 \leq f(x) \leq x + x^2$ for all x (Figure 12.9). Then $f(0) = 0$, and so if $x > 0$, then

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

while if $x < 0$, the reverse inequalities hold. Thus by the squeeze theorem,

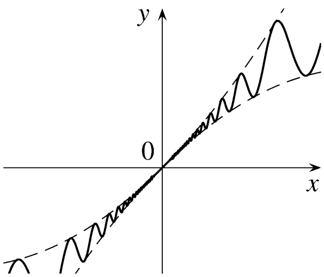

Fig. 12.9

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = 1 > 0.$$

On the other hand, it is clear that we can choose the function f such that it is not monotone increasing in any neighborhood of 0. For this, if we choose $\delta > 0$, we need $-\delta < x < y < \delta$ such that $f(x) > f(y)$. If, for example, $f(x) = x - x^2$ for all rational x and $f(x) = x + x^2$ for all irrational x, then this holds for sure. We can even construct f to be differentiable everywhere: draw a "smooth" (that is, differentiable) wave between the graphs of the functions $x - x^2$ and $x + x^2$ (one such function can be seen in Exercise 12.53).

Even though the condition $f'(a) = 0$ is not sufficient for f to have a local extremum at a, in certain important cases, statement (v) of Theorem 12.44 is still applicable for finding the extrema.

Example 12.46. Find the (absolute) maximum of the function $f(x) = x \cdot (1 - x)$ in the interval $[0, 1]$. Since the function is continuous, by Weierstrass's theorem (Theorem 10.55) f has an absolute maximum in $[0, 1]$. Suppose that f takes on its largest value at the point $a \in [0, 1]$. Then either $a = 0$, $a = 1$, or $a \in (0, 1)$. In the last case f has a local maximum at a, and since f is everywhere differentiable, by statement (v) of Theorem 12.44 we have $f'(a) = 0$. Now $f'(x) = 1 - 2x$, so the condition $f'(a) = 0$ is satisfied only by $a = 1/2$. We get that the function attains its maximum at one of the points 0, 1, 1/2. However, $f(0) = f(1) = 0$ and $f(1/2) = 1/4 > 0$, so 0 and 1 cannot be maxima of the function. Thus only $a = 1/2$ is possible. Thus we have shown that the function $f(x) = x \cdot (1 - x)$ over the interval $[0, 1]$ attains its maximum at the point 1/2; that is, $a = 1/2$ is its absolute maximum.

Remark 12.47. This argument can be applied in the cases in which we are dealing with a function f that is continuous on a closed and bounded interval $[a, b]$ and is differentiable inside, on (a, b). Then f has a largest value by Weierstrass's theorem. If this is attained at a point c, then either $c = a$, $c = b$, or $c \in (a, b)$. In this last case, we are talking about a local extremum as well, so $f'(c) = 0$. Thus if we find all points $c \in (a, b)$ where f' vanishes, then the absolute maximum points of f must be among these, a, and b. We can then locate the absolute maxima by computing f at all of these values (not forgetting about a and b), and determining those at which the value of f is the largest. (We should note that in some cases, we have to compute the value of f at infinitely many points. It can happen that f' has infinitely many roots in (a, b); see Exercise 12.52.)

Example 12.48. As another application of the argument above, we deduce **Snell's**[4] **law.** By what is called **Fermat's**[5] **principle,** light traveling between two points takes the path that can be traversed in the least amount of time. In Figure 12.10, the x-axis separates two fluids in which the speed of light is respectively v_1 and v_2. Looking from the point P_1, we will see point P_2 in the direction that a light ray arrives at P_1 if it starts at P_2. The light ray—by Fermat's principle—"chooses" the path that takes the shortest time to traverse. To determine the path of the light ray, we thus need to solve the following problem.

[4] Willebrord Snellius (1580–1626), Dutch mathematician.

[5] Pierre de Fermat (1601–1665), French mathematician.

Let a line e be given in the plane, and in the two half-planes determined by this line, let there be given the points P_1 and P_2. If a moving point travels with velocity v_1 in the half-plane in which P_1 is located, and with velocity v_2 if it is in the half-plane of P_2, what path must it take to get from P_1 to P_2 in the shortest amount of time?

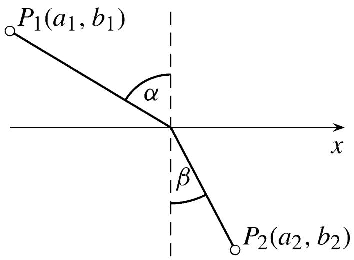

Fig. 12.10

Let the line e be the x-axis, let the coordinates of P_1 be (a_1, b_1), and let the coordinates of P_2 be (a_2, b_2). We may assume that $a_1 < a_2$ (Figure 12.10). Clearly, the point needs to travel in a straight line in both half-planes, so the problem is simply to find where the point crosses the x-axis, that is, where the path bends (is refracted).

If the path intersects the x-axis at the point x, the time necessary for the point to traverse the entire path is

$$f(x) = \frac{1}{v_1} \cdot \sqrt{(x-a_1)^2 + b_1^2} + \frac{1}{v_2} \cdot \sqrt{(x-a_2)^2 + b_2^2},$$

and so

$$f'(x) = \frac{1}{v_1} \cdot \frac{x-a_1}{\sqrt{(x-a_1)^2 + b_1^2}} + \frac{1}{v_2} \cdot \frac{x-a_2}{\sqrt{(x-a_2)^2 + b_2^2}}. \qquad (12.31)$$

Our task is to find the absolute minimum of f. Since if $x < a_1$, then $f(x) > f(a_1)$, and if $x > a_2$, then $f(x) > f(a_2)$, it suffices to find the minimum of f in the interval $[a_1, a_2]$. Since f is continuous, Weierstrass's theorem applies, and f attains its minimum on $[a_1, a_2]$. Since f is also differentiable, the minima can be only at the endpoints of the interval and at the points where the derivative is zero.

Now by (12.31),

$$f'(a_1) = \frac{(a_1 - a_2)}{v_2 \cdot \sqrt{(a_1 - a_2)^2 + b_2^2}} < 0,$$

and so by Theorem 12.44, f is strictly locally decreasing at a_1. Thus in a suitable right-hand neighborhood of a_1, every value of f is smaller that $f(a_1)$, so a_1 cannot be a minimum. Similarly,

$$f'(a_2) = \frac{(a_2 - a_1)}{v_1 \cdot \sqrt{(a_2 - a_1)^2 + b_1^2}} > 0,$$

and so by Theorem 12.44, f is strictly locally increasing at a_2. Thus in a suitable left-hand neighborhood of a_2, every value of f is smaller that $f(a_2)$, so a_2 cannot be a minimum either. Thus the minimum of the function f is at a point $x \in (a_1, a_2)$ where $f'(x) = 0$. By (12.31), this is equivalent to saying that

$$\frac{x-a_1}{\sqrt{(x-a_1)^2+b_1^2}} : \frac{a_2-x}{\sqrt{(x-a_2)^2+b_2^2}} = \frac{v_1}{v_2}.$$

We can see in the figure that

$$\frac{x-a_1}{\sqrt{(x-a_1)^2+b_1^2}} = \sin\alpha \qquad \text{and} \qquad \frac{a_2-x}{\sqrt{(x-a_2)^2+b_2^2}} = \sin\beta,$$

where α and β are called the angle of incidence and angle of refraction respectively. Thus the path taking the least time will intersect the line separating the two fluids at the point where

$$\frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2}.$$

This is **Snell's law.**

Exercises

12.48. We want to create a box with no lid out of a rectangle with sides a and b by cutting out a square of size x at each corner of the rectangle. How should we chose x to maximize the volume of the box? (S)

12.49. Which cylinder inscribed into a given sphere has the largest volume?

12.50. Which right circular cone inscribed into a given sphere has the largest volume?

12.51. Which right circular cone inscribed into a given sphere has the largest surface area? (The surface area of the cone includes the base circle.)

12.52. Let

$$f(x) = \begin{cases} x^2 \cdot \sin 1/x, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Prove that the derivative of f has infinitely many roots in $(0,1)$.

12.53. Let

$$f(x) = \begin{cases} x + 2x^2 \cdot \sin 1/x, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Show that $f'(0) > 0$, but f is not monotone increasing in any neighborhood of 0. (S)

12.54. Prove that if f is locally increasing at all points in (a,b), then it is monotone increasing in (a,b). (H)

12.55. Determine the absolute extrema of the functions below in the given intervals.

(a) $x^2 - x^4$, $[-2,2]$;
(b) $x - \operatorname{arc\,tg} x$, $[-1,1]$;
(c) $x + e^{-x}$, $[-1,1]$;
(d) $x + x^{-2}$, $[1/10, 10]$;
(e) $\operatorname{arc\,tg}(1/x)$, $[1/10, 10]$;
(f) $\cos x^2$, $[0,\pi]$;
(g) $\sin(\sin x)$, $[-\pi/2, \pi/2]$;
(h) $x \cdot e^{-x}$, $[-2,2]$;
(i) $x^n \cdot e^{-x}$, $[-2n, 2n]$;
(j) $x - \log x$, $[1/2, 2]$;
(k) $1/(1+\sin^2 x)$, $(0,\pi)$;
(l) $\sqrt{1-e^{-x^2}}$, $[-2,2]$;
(m) $x \cdot \sin(\log x)$, $[1,100]$;
(n) x^x, $(0,\infty)$;
(o) $\sqrt[x]{x}$, $(0,\infty)$;
(p) $(\log x)/x$, $(0,\infty)$;
(q) $x \cdot \log x$, $(0,\infty)$;
(r) $x^x \cdot (1-x)^{1-x}$, $(0,1)$.

12.5 Intermediate Value Theorems

The following three theorems—each a generalization of the one that it follows—are some of the most frequently used theorems in differentiation. When we are looking for a link between properties of a function and its derivative, most often we use one of these intermediate value theorems.

Theorem 12.49 (Rolle's Theorem[6]). *Suppose that the function f is continuous on $[a,b]$ and differentiable on (a,b). If $f(a) = f(b)$, then there exists a $c \in (a,b)$ such that $f'(c) = 0$.*

Proof. If $f(x) = f(a)$ for all $x \in (a,b)$, then f is constant in (a,b), so $f'(x) = 0$ for all $x \in (a,b)$. Then we can choose c to be any number in (a,b).

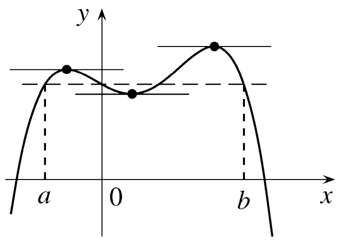

Fig. 12.11

We can thus suppose that there exists an $x_0 \in (a,b)$ for which $f(x_0) \neq f(a)$. Consider first the case $f(x_0) > f(a)$. By Weierstrass's theorem, f has an absolute maximum in $[a,b]$. Since $f(x_0) > f(a) = f(b)$, neither a nor b can be its absolute maximum. Thus if c is an absolute maximum, then $c \in (a,b)$, and so c is also a local maximum too. By statement (v) of Theorem 12.44, it then follows that $f'(c) = 0$.

If $f(x_0) < f(a)$, then we argue similarly, considering the absolute minimum of f instead (Figure 12.11). □

An important generalization of Rolle's theorem is the following theorem.

[6] Michel Rolle (1652–1719), French mathematician.

Theorem 12.50 (Mean Value Theorem). *If the function f is continuous on $[a,b]$ and differentiable on (a,b), then there exists a $c \in (a,b)$ such that*

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

Proof. The equation for the chord between the points $(a, f(a))$ and $(b, f(b))$ is given by

$$y = h_{a,b}(x) = \frac{f(b)-f(a)}{b-a}(x-a)+f(a).$$

The function

$$F(x) = f(x) - h_{a,b}(x)$$

satisfies the conditions of Rolle's theorem. Indeed, since f and $h_{a,b}$ are both continuous in $[a,b]$ and differentiable on (a,b), their difference also has these properties. Since $F(b) = F(a) = 0$, we can apply Rolle's theorem to F. We get that there exists a $c \in (a,b)$ such that $F'(c) = 0$. But this means that

$$0 = F'(c) = f'(c) - h'_{a,b}(c) = f'(c) - \frac{f(b)-f(a)}{b-a},$$

so

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

which concludes the proof of the theorem (Figure 12.12). □

The geometric meaning of the mean value theorem is the following: if the function f is continuous on $[a,b]$ and differentiable on (a,b), then the graph of f has a point in (a,b) where the tangent line is parallel to the chord $h_{a,b}$.

The following theorem is a generalization of the previous one.

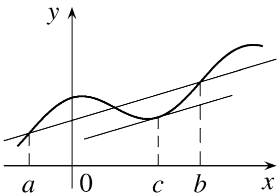

Fig. 12.12

Theorem 12.51 (Cauchy's Mean Value Theorem). *If the functions f and g are continuous on $[a,b]$, differentiable on (a,b), and for $x \in (a,b)$ we have $g'(x) \neq 0$, then there exists a $c \in (a,b)$ such that*

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

Proof. By Rolle's theorem, we know that $g(a) \neq g(b)$. Indeed, if $g(a) = g(b)$ held, then the derivative of g would be zero at at least one point of the interval (a,b), which we did not allow. Let

$$F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)}(g(x)-g(a)).$$

The function F is continuous on $[a,b]$, differentiable on (a,b), and $F(a) = F(b) = 0$. Thus by Rolle's theorem, there exists a $c \in (a,b)$ such that $F'(c) = 0$. Then

$$0 = F'(c) = f'(c) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(c).$$

Since by the assumptions, $g'(c) \neq 0$, we get that

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

which concludes the proof. □

It is clear that the mean value theorem is a special case of Cauchy's mean value theorem if we apply the latter with $g(x) = x$.

A simple but important corollary of the mean value theorem is the following.

Theorem 12.52. *If f is continuous on $[a,b]$, differentiable on (a,b), and $f'(x)=0$ for all $x \in (a,b)$, then the function f is constant on $[a,b]$.*

Proof. By the mean value theorem, for every x in $(a,b]$, there exists a $c \in (a,b)$ such that

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

So by $f'(c) = 0$, we have $f(x) = f(a)$. □

The following corollary is sometimes called the *fundamental theorem of integration;* we will later see why.

Corollary 12.53. *If f and g are continuous on $[a,b]$, differentiable on (a,b), and moreover, $f'(x) = g'(x)$ for all $x \in (a,b)$, then with a suitable constant c, we have $f(x) = g(x) + c$ for all $x \in [a,b]$.*

Proof. Apply Theroem 12.52 to the function $f - g$. □

Exercises

12.56. Give an example of a differentiable function $f \colon \mathbb{R} \to \mathbb{R}$ that has a point c such that $f'(c)$ is not equal to the difference quotient $(f(b) - f(a))/(b - a)$ for any $a < b$. Why does this not contradict the mean value theorem?

12.57. Prove that if f is twice differentiable on $[a,b]$, and for a $c \in (a,b)$ we have $f''(c) \neq 0$, then there exist $a \leq x_1 < c < x_2 \leq b$ such that

$$f'(c) = \frac{f(x_1) - f(x_2)}{x_1 - x_2}. \tag{H}$$

12.58. Prove that

$$(\alpha-\beta)\cdot\cos\alpha \le \sin\alpha-\sin\beta \le (\alpha-\beta)\cdot\cos\beta$$

for all $0<\beta<\alpha<\pi/2$.

12.59. Prove that $|\operatorname{arc\,tg} x-\operatorname{arc\,tg} y|\le|x-y|$ for all x,y.

12.60. Let f be differentiable on the interval I, and suppose that the function f' is bounded on I. Prove that f is Lipschitz on I.

12.61. Prove that if $f'(x)=x^2$ for all x, then there exists a constant c such that $f(x)=(x^3/3)+c$.

12.62. Prove that if $f'(x)=f(x)$ for all x, then there exists a constant c such that $f(x)=c\cdot e^x$ for all x.

12.63. Let $f\colon\mathbb{R}\to(0,\infty)$ be differentiable and strictly monotone increasing. Suppose that the tangent line of the graph of f at every point $(x,f(x))$ intersects the x-axis at the point $x-a$, where $a>0$ is a constant. Prove that f is an exponential function.

12.64. Let $f\colon(0,\infty)\to(0,\infty)$ be differentiable and strictly monotone increasing. Suppose that the tangent line to the graph of f at every point $(x,f(x))$ intersects the x-axis at the point $c\cdot x$, where $c>0$ is a constant. Prove that f is a power function.

12.65. Prove that if f and g are differentiable everywhere, $f(0)=0$, $g(0)=1$, $f'=g$, and $g'=-f$, then $f(x)=\sin x$ and $g(x)=\cos x$ for all x. (H)

12.66. Prove that the function x^5-5x+2 has three real roots.

12.67. Prove that the function $x^7+8x^2+5x-23$ has at most three real roots.

12.68. At most how many real roots can the function $x^{16}+ax+b$ have?

12.69. For what values of k does the function x^3-6x^2+9x+k have exactly one real root?

12.70. Prove that if p is an nth-degree polynomial, then the function $e^x-p(x)$ has at most $n+1$ real roots.

12.71. Let f be n times differentiable on (a,b). Prove that if f has n distinct roots in (a,b), then $f^{(n-k)}$ has at least k roots in (a,b) for all $k=1,\dots,n-1$. (H)

12.72. Prove that if p is an nth-degree polynomial and every root of p is real, then every root of p' is also real.

12.73. Prove that every root of the Legendre polynomial

$$P_n(x)=\frac{1}{2^n n!}\left((x^2-1)^n\right)^{(n)}$$

is real.

12.74. Let f and g be n times differentiable functions on $[a,b]$, and suppose that they have n common roots in $[a,b]$. Prove that if the functions $f^{(n)}$ and $g^{(n)}$ have no common roots in $[a,b]$, then for all $x \in [a,b]$ such that $g(x) \neq 0$, there exists a $c \in (a,b)$ such that

$$\frac{f(x)}{g(x)} = \frac{f^{(n)}(c)}{g^{(n)}(c)}.$$

12.75. Let f be continuous on (a,b) and differentiable on $(a,b) \setminus \{c\}$, where $a < c < b$. Prove that if $\lim_{x \to c} f'(x) = A$, where A is finite, then f is differentiable at c and $f'(c) = A$.

12.76. Prove that if f is twice differentiable at a, then

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

12.77. Let f be differentiable on $(0,\infty)$. Prove that if there exists a sequence $x_n \to \infty$ such that $f(x_n) \to 0$, then there also exists a sequence $y_n \to \infty$ such that $f'(y_n) \to 0$.

12.78. Let f be differentiable on $(0,\infty)$. Prove that if $\lim_{x \to \infty} f'(x) = 0$, then $\lim_{x \to \infty} f(x)/x = 0$.

12.6 Investigation of Differentiable Functions

We begin with monotonicity criteria.

Theorem 12.54. *Let f be continuous on $[a,b]$ and differentiable on (a,b).*

(i) *f is monotone increasing (decreasing) on $[a,b]$ if and only if $f'(x) \geq 0$ ($f'(x) \leq 0$) for all $x \in (a,b)$.*
(ii) *f is strictly monotone increasing (decreasing) on $[a,b]$ if and only if $f'(x) \geq 0$ ($f'(x) \leq 0$) for all $x \in (a,b)$, and $[a,b]$ does not have a nondegenerate subinterval on which f' is identically zero.*

Proof. (i) Suppose that $f'(x) \geq 0$ for all $x \in (a,b)$. By the mean value theorem, for arbitrary $a \leq x_1 < x_2 \leq b$ there exists a $c \in (x_1, x_2)$ such that

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(c).$$

Since $f'(c) \geq 0$, we have $f(x_1) \leq f(x_2)$, which means exactly that f is monotone increasing on $[a,b]$.

Conversely, if f is monotone increasing on $[a,b]$, then it is locally increasing at every x in (a,b). Thus by statement (iii) of Theorem 12.44, we see that $f'(x) \geq 0$.

The proof is similar for the monotone decreasing case.

(ii) It is easy to see that a function f is strictly monotone on $[a,b]$ if and only if it is monotone in $[a,b]$, and if $[a,b]$ does not have a subinterval on which f is constant. Then the statement can be proved easily by Theorem 12.52 □

As an application of the theorem above, we introduce a simple but useful method for proving inequalities.

Corollary 12.55. *Let f and g be continuous on $[a,b]$ and differentiable on (a,b). If $f(a) = g(a)$ and $a < x \le b$ implies $f'(x) \ge g'(x)$, then $f(x) \ge g(x)$ for all $x \in [a,b]$. Also, if $f(a) = g(a)$ and $a < x \le b$ imply $f'(x) > g'(x)$, then $f(x) > g(x)$ for all $x \in (a,b]$.*

Proof. Let $h = f - g$. If $a < x \le b$ implies $f'(x) \ge g'(x)$, then $h'(x) \ge 0$ for all $x \in (a,b)$, and so by statement (i) of Theorem 12.54, h is monotone increasing on $[a,b]$. If $f(a) = g(a)$, then $h(a) = 0$, and so $h(x) \ge h(a) = 0$, and thus $f(x) \ge g(x)$ for all $x \in [a,b]$. The second statement follows similarly, using statement (ii) of Theorem 12.54. □

Example 12.56. To illustrate the method, let us show that

$$\log(1+x) > \frac{2x}{x+2} \tag{12.32}$$

for all $x > 0$. Since for $x = 0$ we have equality, by Corollary 12.55 it suffices to show that the derivatives of the given functions satisfy the inequality for all $x > 0$. It is easy to check that $(2x/(x+2))' = 4/(x+2)^2$. Thus we have only to show that for $x > 0$, $1/(1+x) > 4/(x+2)^2$, which can be checked by multiplying through.

More applications of Corollary 12.55 can be found among the exercises.

Remark 12.57. With the help of Theorem 12.54, we can find the local and absolute extrema of an arbitrary differentiable function even if it is not defined on a closed and bounded interval. This is because by the sign of the derivative, we can determine on which intervals the function is increasing and on which it is decreasing, and this generally gives us enough information to find the extrema.

Consider the function $f(x) = x \cdot e^{-x}$, for example. Since $f'(x) = e^{-x} - x \cdot e^{-x}$, we have $f'(x) > 0$ if $x < 1$, and $f'(x) < 0$ if $x > 1$. Thus f is strictly monotone increasing on $(-\infty, 1]$, and strictly monotone decreasing on $[1, \infty)$. It follows that f has an absolute maximum at 1 (which is also a local maximum), and that f does not have any local or absolute minima.

In Theorem 12.44 we saw that if f is differentiable at a, then for f to have a local extremum at a, it is necessary (but generally not sufficient) for $f'(a) = 0$ to hold. The following theorems give *sufficient* conditions for the existence of local extrema.

Theorem 12.58. *Let f be differentiable in a neighborhood of the point a.*

(i) If $f'(a) = 0$ and f' is locally increasing (decreasing) at a,[7] then a is a local minimum (maximum) of f.

(ii) If $f'(a) = 0$ and f' is strictly locally increasing (decreasing) at a, then the point a is a strict local minimum (maximum) of f.

Proof. (i) Consider the case that f' is locally increasing at a. Then there exists a $\delta > 0$ such that $f'(x) \le 0$ if $a-\delta < x < a$, and $f'(x) \ge 0$ if $a < x < a+\delta$. By Theorem 12.54, it then follows that f is monotone decreasing on $[a-\delta, a]$, and monotone increasing on $[a, a+\delta]$. Thus if $a-\delta < x < a$, then $f(x) \ge f(a)$. Moreover, if $a < x < a+\delta$, then again $f(x) \ge f(a)$. This means exactly that f has a local minimum at a. The statement for the local maximum can be proved similarly.

(ii) If f' is strictly locally increasing at a, then there exists a $\delta > 0$ such that $f'(x) < 0$ if $a-\delta < x < a$, and $f'(x) > 0$ if $a < x < a+\delta$.

It then follows that f is strictly monotone decreasing on $[a-\delta, a]$, and strictly monotone increasing on $[a, a+\delta]$. Thus if $a-\delta < x < a$, then $f(x) > f(a)$. Moreover, if $a < x < a+\delta$, then again $f(x) > f(a)$. This means exactly that f has a strict local minimum at a. One can argue similarly for the case of a strict local maximum. □

Remark 12.59. The sign change of f' at a is not necessary for f to have a local extremum at the point a. Let f be a function such that $x^2 \le f(x) \le 2x^2$ for all x. Then the point 0 is a strict local (and absolute) minimum of f. On the other hand, it is possible that f is not differentiable (or even continuous) at any point other than 0; this is the case, for example, if $f(x) = x^2$ for all rational x, and $f(x) = 2x^2$ for all irrational x.

We can also construct such an f to be differentiable; for this we need to place a differentiable function between the graphs of the functions x^2 and $2x^2$ each of whose one-sided neighborhoods of 0 contains a section that is monotone decreasing and a section that is monotone increasing as well. We give such a function in Exercise 12.95.

Theorem 12.60. *Let f be twice differentiable at a. If $f'(a) = 0$ and $f''(a) > 0$, then f has a strict local minimum at a. If $f'(a) = 0$ and $f''(a) < 0$, then f has a strict local maximum at a.*

Proof. Suppose that $f''(a) > 0$. By Theorem 12.44, it follows that f' is strictly locally increasing at a. Now apply the previous theorem (Figure 12.13). The proof for the case $f''(a) < 0$ is similar. □

Remark 12.61. If $f'(a) = 0$ and $f''(a) = 0$, then we cannot deduce whether f has a local extremum at a. The different possibilities are illustrated by the functions

[7] That is, if f' changes signs at the point a, meaning that it is nonpositive on a left-sided neighborhood of a and nonnegative on a right-sided neighborhood, or vice versa.

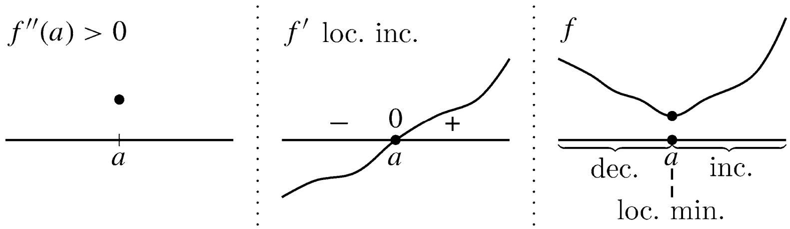

Fig. 12.13

$f(x) = x^3$, $f(x) = x^4$, and $f(x) = -x^4$ at $a = 0$. In this case, we can get sufficient conditions for f to have a local extremum at a by the value of higher-order derivatives.

Theorem 12.62.

(i) *Let the function f be $2k$ times differentiable at the point a, where $k \geq 1$. If*

$$f'(a) = \ldots = f^{(2k-1)}(a) = 0 \ \textit{and} \ f^{(2k)}(a) > 0, \tag{12.33}$$

then f has a strict local minimum at a. If

$$f'(a) = \ldots = f^{(2k-1)}(a) = 0 \ \textit{and} \ f^{(2k)}(a) < 0,$$

then f has a strict local maximum at a.

(ii) *Let the function f be $2k+1$ times differentiable at a, where $k \geq 1$. If*

$$f'(a) = \ldots = f^{(2k)}(a) = 0 \ \textit{and} \ f^{(2k+1)}(a) \neq 0, \tag{12.34}$$

then f is strictly monotone in a neighborhood of a, that is, f does not have a local extremum there.

Proof. (i) We prove only the first statement, using induction. We already saw the $k = 1$ case in Theorem 12.60. Let $k > 1$, and suppose that the statement holds for $k-1$. If (12.33) holds for f, then for the function $g = f''$, we have

$$g'(a) = \ldots = g^{(2k-3)}(a) = 0 \ \text{and} \ g^{(2k-2)}(a) > 0.$$

Thus by the induction hypothesis, f'' has a strict local minimum at a. Since by $k > 1$ we have $f''(a) = 0$, there must exist a $\delta > 0$ such that $f''(x) > 0$ at all points $x \in (a-\delta, a+\delta) \setminus \{a\}$. Then by Theorem 12.54, it follows that f' is strictly monotone increasing on $(a-\delta, a+\delta)$. Thus f' is strictly locally increasing at a, so we can apply Theorem 12.58.

(ii) Suppose (12.34) holds. Then by the already proved statement (i), a is a strict local extremum of f'. Since $f'(a) = 0$, either $f'(x) > 0$ for all $x \in (a-\delta, a+\delta) \setminus \{a\}$, or $f'(x) < 0$ for all $x \in (a-\delta, a+\delta) \setminus \{a\}$. By Theorem 12.54, it then follows that f is strictly monotone on $(a-\delta, a+\delta)$, so it does not have a local extremum at the point a (Figure 12.14). □

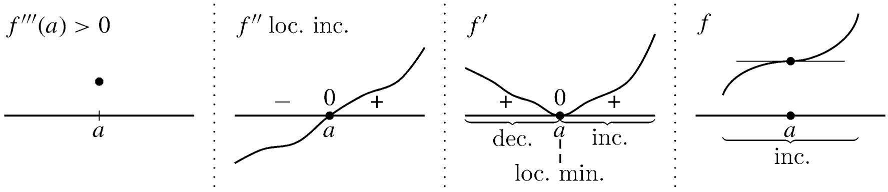

Fig. 12.14

We now turn to the conditions for convexity.

Theorem 12.63. *Let f be differentiable on the interval I.*

(i) *The function f is convex (concave) on I if and only if f' is monotone increasing (decreasing) on I.*
(ii) *The function f is strictly convex (concave) on I if and only if f' is strictly monotone increasing (decreasing) on I.*

Proof. (i) Suppose that f' is monotone increasing on I. Let $a, b \in I$, $a < b$, and let $a < x < b$ be arbitrary. By the mean value theorem, there exist points $u \in (a, x)$ and $v \in (x, b)$ such that

$$f'(u) = \frac{f(x) - f(a)}{x - a} \text{ and } f'(v) = \frac{f(b) - f(x)}{b - x}.$$

Since $u < v$ and f' is monotone increasing, $f'(u) \le f'(v)$, so

$$\frac{f(x) - f(a)}{x - a} \le \frac{f(b) - f(x)}{b - x}.$$

Then by a simple rearrangement, we get that

$$f(x) \le \frac{f(b) - f(a)}{b - a} \cdot (x - a) + f(a),$$

which shows that f is convex.

Now suppose that f is convex on I, and let $a, b \in I$, $a < b$, be arbitrary. By Theorem 9.20, the function $F(x) = (f(x) - f(a))/(x - a)$ is monotone increasing on the set $I \setminus \{a\}$. Thus $F(x) \le ((f(b) - f(a))/(b - a)$ for all $x < b$, $x \ne a$. Since $f'(a) = \lim_{x \to a} F(x)$, we have that (Figures 12.15)

$$f'(a) \le \frac{f(b) - f(a)}{b - a}. \tag{12.35}$$

Similarly, the function $G(x) = (f(x) - f(b))/(x - b)$ is monotone increasing on the set $I \setminus \{b\}$, so $G(x) \ge ((f(b) - f(a))/(b - a)$ for all $x > a$, $x \ne b$.

Since $f'(b) = \lim_{x \to b} G(x)$, we have that

$$f'(b) \ge \frac{f(b) - f(a)}{b - a}. \tag{12.36}$$

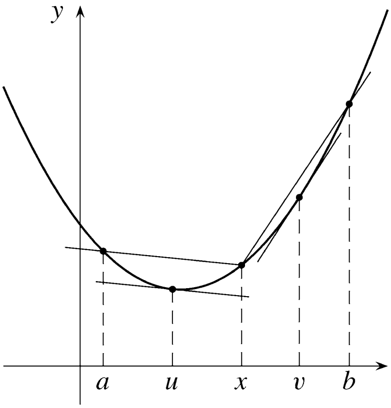

Fig. 12.15

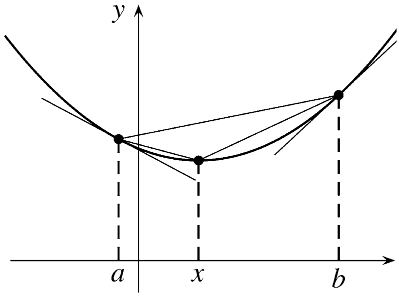

Fig. 12.16

If we combine (12.35) and (12.36), we get that $f'(a) \le f'(b)$. Since this is true for all $a, b \in I$ if $a < b$, f' is monotone increasing on I (Figure 12.16).

The statement for concavity can be proved in the same way. Statement (ii) follows by a straightforward change of the argument above. □

Rearranging equation (12.35), we get that if $a < b$, then $f(b) \ge f'(a) \cdot (b-a) + f(a)$. Equation (12.36) states that if $a < b$, then $f(a) \ge f'(b) \cdot (a-b) + f(b)$. This means that for arbitrary $a, x \in I$,

$$f(x) \ge f'(a) \cdot (x-a) + f(a), \tag{12.37}$$

that is, a tangent line drawn at any point on the graph of f lies below the graph itself. Thus we have proved the "only if" part of the following theorem.

Theorem 12.64. *Let f be differentiable on the interval I. The function f is convex on I if and only if for every $a \in I$, the graph of the function f lies above the tangent of the graph at the point a, that is, if and only if* (12.37) *holds for all $a, x \in I$.*

Proof. We now have to prove only the "if" part of the statement. Suppose that (12.37) holds for all $a, x \in I$. If $a, b \in I$ and $a < b$, then it follows that both (12.35) and (12.36) are true, so $f'(a) \le f'(b)$. Thus f' is monotone increasing on I, so by Theorem 12.63, f is convex. □

Theorem 12.65. *Let f be twice differentiable on I. The function f is convex (concave) on I if and only if $f''(x) \ge 0$ ($f''(x) \le 0$) for all $x \in I$.*

Proof. The statement of the theorem is a simple corollary of Theorems 12.63 and 12.54. □

Definition 12.66. We say that a point a is an *inflection point* of the function f if f is continuous at a, f has a (finite or infinite) derivative at a, and there exists a $\delta > 0$ such that f is convex on $(a-\delta, a]$ and concave on $[a, a+\delta)$, or vice versa.

So for example, 0 is an inflection point of the functions x^3 and $\sqrt[3]{x}$.

Theorem 12.67. *If f is twice differentiable at a, and f has an inflection point at a, then $f''(a) = 0$.*

Proof. If f is convex on $(a-\delta, a]$, then f' is monotone increasing there; if it is concave on $[a, a+\delta)$, then f' is monotone decreasing there. Thus f' has a local maximum at a, and so $f''(a) = 0$.

The proof is similar in the case that f is concave on $(a-\delta, a]$ and convex on $[a, a+\delta)$. □

Remark 12.68. Let f be differentiable on a neighborhood of the point a. By Theorem 12.63, a is an inflection point of f if and only if a is a local extremum of f' such that f' is increasing in a left-hand neighborhood of a and is decreasing in a right-hand neighborhood of a, or the other way around. From this observation and by Theorem 12.67, we get the following theorem.

Theorem 12.69. *Let f be twice differentiable on a neighborhood of the point a. Then a is an inflection point of f if and only if f'' changes sign at the point a, that is, if $f''(a) = 0$ and f'' is locally increasing or decreasing at a.*

Corollary 12.70. *Let f be three times differentiable at a. If $f''(a) = 0$ and $f'''(a) \neq 0$, then f has an inflection point at a.*

Remark 12.71. In the case $f''(a) = f'''(a) = 0$, it is possible that f has an inflection point at a, but it is also possible that it does not. The different cases are illustrated by the functions $f(x) = x^4$ and $f(x) = x^5$ at the point $a = 0$. As in the case of extrema, we can refer to the values of higher-order derivatives to help determine when a point of inflection exists.

Theorem 12.72.

(i) Let the function f be $2k+1$ times differentiable at a, where $k \geq 1$. If

$$f''(a) = \ldots = f^{(2k)}(a) = 0 \ \ \textit{and} \ \ f^{(2k+1)}(a) \neq 0, \tag{12.38}$$

then f has an inflection point at a.

(ii) If

$$f''(a) = \ldots = f^{(2k-1)}(a) = 0 \ \ \textit{and} \ \ f^{(2k)}(a) \neq 0, \tag{12.39}$$

then f is strictly convex or concave in a neighborhood of a, and so a is not a point of inflection.

Proof. (i) We already saw the case for $k = 1$ in the previous theorem, so we may assume that $k > 1$. If (12.38) holds, then by statement (ii) of Theorem 12.62, f'' is strictly monotone in a neighborhood of a. Thus f'' is locally increasing or decreasing at the point a, and we can apply Theorem 12.69.

(ii) Assume (12.39). Then necessarily $k > 1$. By statement (i) of Theorem 12.62, f'' has a strict local extremum at a. Since $f''(a) = 0$, this means that for a suitable neighborhood U of a, the sign of f'' does not change in $U \setminus \{a\}$. Thus f' is strictly monotone on U, so f is either strictly convex or strictly concave on U, and so a cannot be a point of inflection. □

Remark 12.73. If the function f is infinitely differentiable at a and $f^{(n)}(a) \neq 0$ for at least one $n \geq 2$, then with the help of Theorems 12.62 and 12.72, we can determine whether f has a local extremum or a point of inflection at a. This is because if $f'(a) \neq 0$, then a cannot be a local extremum. Now suppose that $f'(a) = 0$, and let n be the smallest positive integer for which $f^{(n)}(a) \neq 0$. In this case, the function f has a local extremum at a if and only if n is even.

If $f''(a) \neq 0$, then a is not an inflection point of f. If, however, $f''(a) = 0$, and n is the smallest integer greater than 2 for which $f^{(n)}(a) \neq 0$, then a is a point of inflection of f if and only if n is odd.

It can occur, however, that f is infinitely differentiable at a, and that $f^{(n)}(a) = 0$ for all n, while f is not zero in any neighborhood of a. In the following chapter we will see that among such functions we will find some that have local extrema at a, but we will also find those that do not; we will find those that have points of inflection at a, and those that do not (see Remark 13.17).

A **complete investigation of the function** f is accomplished by finding the following pieces of information about the function:

1. the (one-sided) limits of the accumulation points of the domain;
2. the intervals on which f is monotone increasing or decreasing;
3. the local and absolute extrema of f;
4. the points where f is continuous or differentiable;
5. the intervals on which f is convex or concave;
6. the inflection points of f.

Example 12.74. Carry out a complete investigation of the function $x^{2n}e^{-x^2}$ for all $n \in \mathbb{N}$.

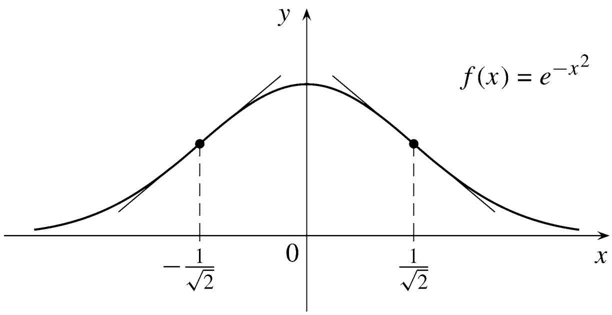

Fig. 12.17

1. Consider first the case $n = 0$. If $f(x) = e^{-x^2}$, then $f'(x) = -2x \cdot e^{-x^2}$, so the sign of $f'(x)$ agrees with the sign of $-x$. Then f is strictly increasing on the interval $(-\infty, 0]$, and strictly decreasing on the interval $[0, \infty)$, so f must have a strict local and global maximum at $x = 0$.

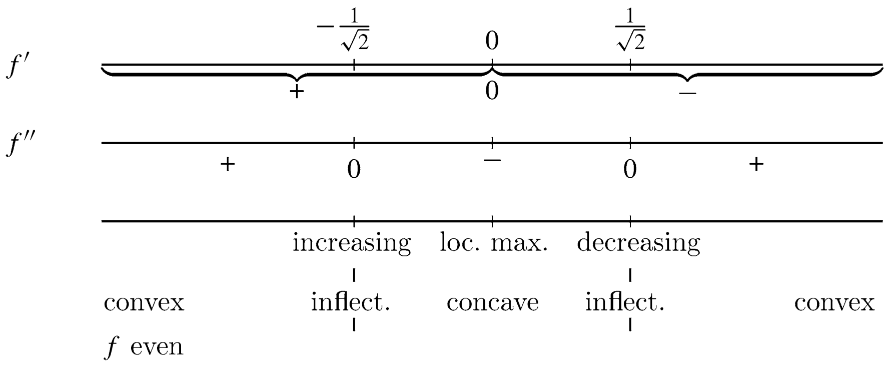

Fig. 12.18

On the other hand, $f''(x) = e^{-x^2}(4x^2 - 2)$, so $f''(x) > 0$ if $|x| > 1/\sqrt{2}$, and $f''(x) < 0$ if $|x| < 1/\sqrt{2}$. So f is strictly concave on $[-1/\sqrt{2}, 1/\sqrt{2}]$, and strictly convex on the intervals $(-\infty, -1/\sqrt{2}]$ and $[1/\sqrt{2}, \infty)$, so the points $\pm 1/\sqrt{2}$ are points of inflection. All these properties are summarized in Figure (12.18).

If we also consider that $\lim_{x\to\pm\infty} f(x) = 0$, then we can sketch the graph of the function as seen in Figure 12.17. Inspired by the shape of its graph, the function e^{-x^2}—which appears many times in probability—is called a **bell curve**.

2. Now let $f(x) = x^{2n}e^{-x^2}$, where $n \in \mathbb{N}^+$. Then

$$f'(x) = \left(2n \cdot x^{2n-1} - 2x^{2n+1}\right) e^{-x^2} = 2 \cdot x^{2n-1} e^{-x^2} \left(n - x^2\right).$$

Thus $f'(x)$ is positive on the intervals $(-\infty, -\sqrt{n})$ and $(0, \sqrt{n})$, and negative on the intervals $(-\sqrt{n}, 0)$ and $(\sqrt{n}, \infty)$. On the other hand,

$$\begin{aligned}
\tfrac{1}{2} \cdot f''(x) &= \left[\left(n \cdot x^{2n-1} - x^{2n+1}\right) \cdot e^{-x^2}\right]' = \\
&= \left(n(2n-1)x^{2n-2} - (2n+1)x^{2n} - 2nx^{2n} + 2x^{2n+2}\right) \cdot e^{-x^2} = \\
&= x^{2n-2} \cdot \left(2x^4 - (4n+1)x^2 + n(2n-1)\right) \cdot e^{-x^2}.
\end{aligned}$$

It is easy to see that the roots of f'' are the numbers

$$\pm\sqrt{\frac{4n+1 \pm \sqrt{16n+1}}{4}}.$$

If we denote these by x_i $(1 \le i \le 4)$, then $x_1 < x_2 < 0 < x_3 < x_4$. It is also easy to check that f'' is positive if $|x| < x_3$ or $|x| > x_4$, and negative if $x_3 < |x| < x_4$ (Figure 12.19). The behavior of the function can be summarized by the Figure 12.18.

If we also consider that $f(0) = 0$ and $\lim_{x\to\pm\infty} f(x) = 0$, then we can sketch the graph as in Figure 12.20.

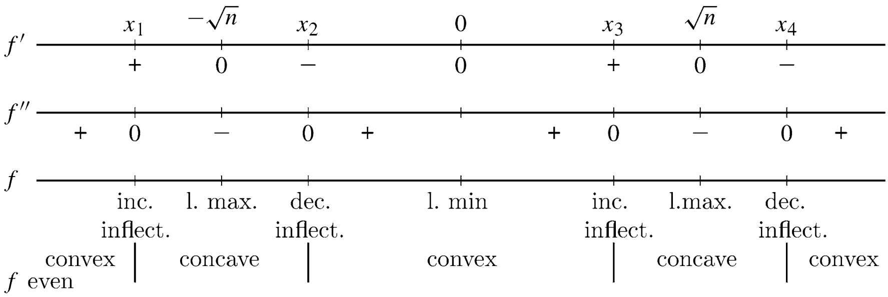

Fig. 12.19

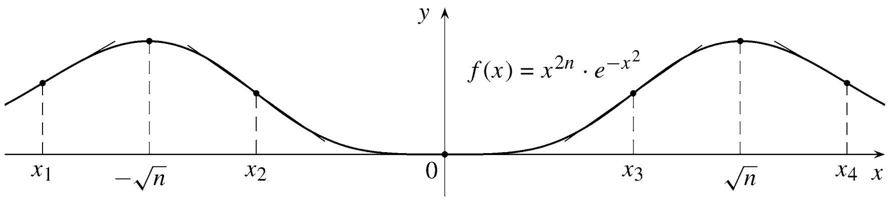

Fig. 12.20

Exercises

12.79. Prove that if $x \in [0,1]$, then $2^x \leq 1+x \leq e^x$ and $2x/\pi \leq \sin x \leq x$.

12.80. Prove that if f is a rational function, then there exists an a such that f is monotone and either concave or convex on $(-\infty, a)$. A similar statement is true on a suitable half-line (b, ∞).

12.81. For which $a > 0$, does the equation $a^x = x$ have roots? (H)

12.82. For which $a > 0$ is the sequence defined by the recurrence $a_1 = a$, $a_{n+1} = a^{a_n}$ convergent? (H)

12.83. Prove that $\left(1+\frac{1}{x}\right)^{x+(1/2)} > e$ for all $x > 0$. (H)

12.84. Prove that for all $x \geq 0$ and $n \geq 1$,

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

12.85. Prove that for all $0 \leq x \leq K$ and $n \geq 1$,

$$e^x \leq 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} \cdot e^K.$$

12.86. Prove that for all $x \geq 0$,

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots . \qquad (12.40)$$

As a special case,

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots .$$

12.87. Prove that e is irrational. (S)

12.88. Prove that for all $x \geq 0$ and $n \geq 1$,

$$1 - x + \frac{x^2}{2!} - \ldots - \frac{x^{2n-1}}{(2n-1)!} \leq e^{-x} \leq 1 - x + \frac{x^2}{2!} - \ldots + \frac{x^{2n}}{(2n)!}.$$

12.89. Prove that (12.40) holds for all x.

12.90. Prove that for all $x \geq 0$ and $n \geq 1$,

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots - \frac{x^{4n-2}}{(4n-2)!} \leq \cos x \leq 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + \frac{x^{4n}}{(4n)!}$$

and

$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots - \frac{x^{4n-1}}{(4n-1)!} \leq \sin x \leq x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + \frac{x^{4n+1}}{(4n+1)!}.$$

12.91. Prove that for all $x \geq 0$ and $n \geq 1$,

$$x - \frac{x^2}{2} + \cdots - \frac{x^{2n}}{2n} \leq \log(1+x) \leq x - \frac{x^2}{2} + \cdots + \frac{x^{2n+1}}{2n+1}.$$

12.92. Prove that for all $x \in [0,1]$,

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

As a special case,

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \ldots . \tag{12.41}$$

12.93. Prove that

$$x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots - \frac{x^{4n-1}}{4n-1} \leq \operatorname{arc\,tg} x \leq x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + \frac{x^{4n+1}}{4n+1}$$

for every $x \geq 0$ and $n \geq 1$.

12.94. Prove that

$$\operatorname{arc\,tg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots .$$

for every $|x| \leq 1$.

As a special case,

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \dots. \tag{12.42}$$

12.95. Let

$$f(x) = \begin{cases} x^4 \cdot (2 + \sin(1/x)), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Show that f has a strict local minimum at 0, but that f' does not change sign at 0. (S)

12.96. Let

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

Prove that

(a) f is continuous on $[0, \infty)$,
(b) f is differentiable on $[0, \infty)$,
(c) f is strictly monotone increasing on $[0, \infty)$,
(d) $f'(1/(2\pi k)) = 0$ if $k \in \mathbb{N}^+$, that is, f' is 0 at infinitely many places in the interval $[0, 1]$.

12.97. Give a function that is monotone decreasing and differentiable on $(0, \infty)$, satisfies $\lim_{x\to\infty} f(x) = 0$, but $\lim_{x\to\infty} f'(x) \neq 0$.

12.98. Let f be differentiable on a punctured neighborhood of a, and let

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} = \infty.$$

Does it then follow that $\lim_{x\to a} f'(x) = \infty$?

12.99. Let f be convex and differentiable on the open interval I. Prove that f has a minimum at the point $a \in I$ if and only if $f'(a) = 0$.

12.100. Let f be convex and differentiable on $(0, 1)$. Prove that if $\lim_{x\to 0+0} f(x) = \infty$, then $\lim_{x\to 0+0} f'(x) = -\infty$. Show that the statement no longer holds if we drop the convexity assumption.

12.101. Carry out a complete investigation of each of the functions below.

$x^3 - 3x$,	$x^2 - x^4$,	$x - \operatorname{arc\,tg} x$,	$x + e^{-x}$,
$x + x^{-2}$,	$\operatorname{arc\,tg}(1/x)$,	$\cos x^2$,	$\sin(\sin x)$,
$\sin(1/x)$,	$x \cdot e^{-x}$,	$x^n \cdot e^{-x}$,	$x - \log x$,
$1/(1 + \sin^2 x)$,	$\left(1 + \frac{1}{x}\right)^x$,	$\left(1 + \frac{1}{x}\right)^{x+1}$,	$\sqrt{1 - e^{-x^2}}$,
$x \cdot \sin(\log x)$,	x^x,	$\sqrt[x]{x}$,	$(\log x)/x$,

$x \cdot \log x$, $\quad x^x \cdot (1-x)^{1-x}$, $\quad \operatorname{arc\,tg} x - \frac{1}{2}\log(1+x^2)$,

$\operatorname{arc\,tg} x - \frac{x}{x+1}$, $\quad x^4/(1+x)^3$, $\quad e^x/(1+x)$, $\quad e^x/\operatorname{sh} x$,

$\sqrt[\log x]{x}$, $\quad e^{-x} \cdot \left[\frac{1-x^2}{2}\sin x - \frac{(1+x)^2}{2}\cos x\right]$,

$x^{2n+1}e^{-x^2} \quad (n \in \mathbb{N})$.