Chapter 11
Various Important Classes of Functions (Elementary Functions)

In this chapter, we will familiarize ourselves with the most commonly occurring functions in mathematics and in applications of mathematics to the sciences. These are the polynomials, rational functions, exponential, power, and logarithm functions, trigonometric functions, hyperbolic functions, and their inverses. We call the functions that we can get from the above functions using basic operations and composition **elementary functions**.

11.1 Polynomials and Rational Functions

We call the function $p : \mathbb{R} \to \mathbb{R}$ a **polynomial function** (a polynomial, for short) if there exist real numbers $a_0, a_1, \ldots, a_n$ such that

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \tag{11.1}$$

for all x. Suppose that in the above description, $a_n \neq 0$. If x_1 is a root of p (that is, if $p(x_1) = 0$), then

$$p(x) = p(x) - p(x_1) = a_n(x^n - x_1^n) + \cdots + a_1(x - x_1).$$

Here using the equality

$$x^k - x_1^k = (x - x_1)(x^{k-1} + x_1 x^{k-2} + \cdots + x_1^{k-2} x + x_1^{k-1}),$$

and then taking out the common factor $x - x_1$ we get that $p(x) = (x - x_1) \cdot q(x)$, where $q(x) = b_{n-1} x^{n-1} + \cdots + b_1 x + b_0$ and $b_{n-1} = a_n \neq 0$.

If x_2 is a root of q, then by repeating this process with q, we obtain that $p(x) = (x - x_1)(x - x_2) \cdot r(x)$, where $r(x) = c_{n-2} x^{n-2} + \cdots + c_1 x + c_0$ and $c_{n-2} = a_n \neq 0$.

It is clear that this process ends in at most n steps, and in the last step, we get the following.

M. Laczkovich, V.T. Sós, *Real Analysis*, Undergraduate Texts in Mathematics, DOI 10.1007/978-1-4939-2766-1_11

Lemma 11.1. *Suppose that in (11.1), $a_n \neq 0$. If p has a root, then there exist not necessarily distinct real numbers $x_1, \dots, x_k$ and a polynomial p_1 such that $k \leq n$, the polynomial p_1 has no roots, and*

$$p(x) = (x - x_1) \cdot \ldots \cdot (x - x_k) \cdot p_1(x) \tag{11.2}$$

for all x. It then follows that p can have at most n roots.

The above lemma has several important consequences.

1. If a polynomial is not identically zero, then it has only finitely many roots. Clearly, in the expression (11.1), not all coefficients are zero. If a_m is the nonzero coefficient with largest index, then we can omit the terms with larger indices. Then by the lemma, p can have at most m roots.
2. If two polynomials agree in infinitely many points, then they are equal everywhere. (Apply the previous point to the difference of the two polynomials.)
3. The identically zero function can be expressed as (11.1) only if $a_0 = \ldots = a_n = 0$ (since the identically zero function has infinitely many roots).
4. If in an expression (11.1), $a_n \neq 0$ and the polynomial p defined by (11.1) has an expression of the form

$$p(x) = b_k x^k + b_{k-1} x^{k-1} \cdots + b_1 x + b_0,$$

where $b_k \neq 0$, then necessarily $k = n$ and $b_i = a_i$ for all $i = 0, \dots, n$. We see this by noting that the difference is the identically zero function, so this statement follows from the previous one.

The final corollary means that a not identically zero polynomial has a unique expression of the form (11.1) in which a_n is nonzero.

In this presentation of a polynomial, we call the coefficient a_n the **leading coefficient** of p, and the number n the **degree** of the polynomial. We denote the degree of p by $\operatorname{gr} p$.[1] The zero-degree polynomials are thus the constant functions different from zero. The identically zero function does not have a degree.

If a polynomial p is not identically zero, then its presentation of the form (11.2) is unique. Clearly, if $p(x) = (x - y_1) \cdot \ldots \cdot (x - y_m) \cdot p_2(x)$ is another presentation, then x_1 is also a root of this, whence one of $y_1, \dots, y_m$ must be equal to x_1 (since p_2 has no roots). We can suppose that $y_1 = x_1$. Then

$$(x - x_2) \cdot \ldots \cdot (x - x_k) \cdot p_1(x) = (x - y_2) \cdot \ldots \cdot (x - y_m) \cdot p_2(x)$$

for all $x \neq x_1$. Then the two sides agree at infinitely many points, so they are equal everywhere. Since x_2 is a root of the right-hand side, it must be equal to one of $y_2, \dots, y_m$. We can assume that $y_2 = x_2$. Repeating this argument, we run out of $x - x_i$ terms on the left-hand side, and at the kth step, we get that

$$p_1(x) = (x - y_{k+1}) \cdot \ldots \cdot (x - y_m) \cdot p_2(x).$$

[1] The notation is based on the Latin *gradus* = degree.

Since p_1 has no roots, necessarily $m = k$ and $p_1 = p_2$.

If in the presentation (11.2), an $x - \alpha$ term appears ℓ times, then we say that α is a **root of multiplicity** ℓ. So, for example, the polynomial $p(x) = x^5 - x^4 - x + 1$ has 1 as a root of multiplicity two, and -1 is a root of multiplicity one (often called a simple root), since $p(x) = (x-1)^2(x+1)(x^2+1)$, and x^2+1 has no roots.[2]

As for the analytic properties of polynomials, first of all, we should note that *a polynomial is continuous everywhere.* This follows from Theorem 10.44, taking into account the fact that constant functions and the function x are continuous everywhere. We now show that if in the presentation (11.1), $n > 0$ and $a_n \neq 0$, then

$$\lim_{x\to\infty} p(x) = \begin{cases} \infty, & \text{if } a_n > 0, \\ -\infty, & \text{if } a_n < 0. \end{cases} \tag{11.3}$$

This is clear from the rearrangement

$$p(x) = x^n \left(a_n + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n} \right),$$

using that $\lim_{x\to\infty} x^n = \infty$ and

$$\lim_{x\to\infty} \left(a_n + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n} \right) = a_n.$$

Rational functions are functions of the form p/q, where p and q are polynomials, and q is not identically zero. The rational function p/q is defined where the denominator is nonzero, so everywhere except for a finite number of points. By 10.44, it again follows that a rational function is continuous at every point where it is defined.

The following theorem is analogous to the limit relation (11.3).

Theorem 11.2. *Let*

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

and

$$q(x) = b_k x^k + b_{k-1} x^{k-1} \cdots + b_1 x + b_0,$$

where $a_n \neq 0$ and $b_k \neq 0$. Then

$$\lim_{x\to\infty} \frac{p(x)}{q(x)} = \begin{cases} \infty, & \text{ha } a_n/b_k > 0 \text{ and } n > k, \\ -\infty, & \text{ha } a_n/b_k < 0 \text{ and } n > k, \\ a_n/b_k, & \text{if } n = k, \\ 0, & \text{if } n < k. \end{cases}$$

[2] Since we defined polynomials on $\mathbb{R}$, we have been talking about only real roots the entire time. Among complex numbers, every nonconstant polynomial has a root; see the second appendix of the chapter.

Exercises

11.1. Show that if p and q are polynomials, then so are $p+q$, $p\cdot q$, and $p\circ q$.

11.2. Let p and q be polynomials. Prove that

(a) if none of p, q, $p+q$ are identically zero, then $\operatorname{gr}(p+q)\le\max(\operatorname{gr}p,\operatorname{gr}q)$;
(b) if neither p nor q is identically zero, then $\operatorname{gr}(p\cdot q)=(\operatorname{gr}p)+(\operatorname{gr}q)$;
(c) if none of p, q, and $p\circ q$ are identically zero then $\operatorname{gr}(p\circ q)=(\operatorname{gr}p)\cdot(\operatorname{gr}q)$. Does it suffice to assume that only p and q are not identically zero?

11.3. Let $p(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, where $a_n>0$. Prove that p is monotone increasing on the half-line (K,∞) if K is sufficiently large.

11.4. Prove that if the polynomial p is not constant, then p takes on each of its values at most k times, where $k=\operatorname{gr}p$.

11.5. Prove that if the rational function p/q is not constant, then it takes on each of its values at most k times, where $k=\max(\operatorname{gr}p,\operatorname{gr}q)$.

11.6. Prove that every polynomial is Lipschitz on every bounded interval.

11.7. Prove that every rational function is Lipschitz on every closed and bounded interval on which it is defined.

11.2 Exponential and Power Functions

Before we define the two important classes of the exponential and power functions, we fulfill our old promise (which we made after Theorem 3.27) and show that the identities regarding taking powers still apply when we take arbitrary real powers.

The simple proof is made possible by our newly gained knowledge of limits of sequences and their properties. We will use the following lemma in the proof of all three identities.

Lemma 11.3. *If $a>0$ and $x_n\to x$, then $a^{x_n}\to a^x$.*

Proof. Suppose first that $a>1$. In Theorem 3.25, we saw that

$$\sup\{a^r\colon r\in\mathbb{Q},\ r<x\}=\inf\{a^s\colon s\in\mathbb{Q},\ s>x\},$$

and by definition, the shared value is a^x. Let $\varepsilon>0$ be given. Then there exist rational numbers $r<x$ and $s>x$ such that

$$a^x-\varepsilon<a^r \text{ and } a^s<a^x+\varepsilon.$$

Since $x_n\to x$, for suitable n_0 we have $r<x_n<s$ if $n>n_0$. Now according to Theorem 3.27, for every $u<v$, $a^u<a^v$. Thus for every $n>n_0$,

$$a^x - \varepsilon < a^r < a^{x_n} < a^s < a^x + \varepsilon.$$

Since ε was arbitrary, we have shown that $a^{x_n} \to a^x$.

The statement can be proved similarly if $0 < a < 1$, while the case $a = 1$ is trivial. □

Theorem 11.4. *For arbitrary $a, b > 0$ and real exponents x, y,*

$$(ab)^x = a^x \cdot b^x, \qquad a^{x+y} = a^x \cdot a^y \qquad \textit{and} \qquad (a^x)^y = a^{xy}. \tag{11.4}$$

Proof. We have already seen these inequalities for *rational* exponents in Theorem 3.23.

We begin by showing that the first two equalities in (11.4) hold for all positive a, b and real numbers x, y. Choose two sequences (r_n) and (s_n) of rational numbers that tend to x and y respectively. (For example, if $r_n \in \big(x-(1/n), x+(1/n)\big) \cap \mathbb{Q}$ and $s_n \in (y-(1/n), y+(1/n)) \cap \mathbb{Q}$, then these sequences work.) Then by Lemma 11.3,

$$(ab)^x = \lim_{n\to\infty} (ab)^{r_n} = \lim_{n\to\infty} a^{r_n} \cdot b^{r_n} = a^x \cdot b^x$$

and

$$a^{x+y} = \lim_{n\to\infty} a^{r_n+s_n} = \lim_{n\to\infty} a^{r_n} \cdot a^{s_n} = a^x \cdot a^y.$$

We only outline the proof of the third identity for the case $a > 1$ and $x, y > 0$. (The remaining cases cant be proven similarly, or can be reduced to our case by considering reciprocals.)

Let now $r_n \to x$ and $s_n \to y$ be sequences consisting of rational numbers that satisfy $0 < r_n < x$ and $0 < s_n < y$ for all n. Then

$$a^{r_n s_n} = (a^{r_n})^{s_n} < (a^x)^{s_n} < (a^x)^y. \tag{11.5}$$

Here, other than using Theorem 3.27, in the middle inequality we used the fact that if $0 < u < v$ and $s > 0$ is rational, then $u^s < v^s$. This follows from $v^s/u^s = (v/u)^s > (v/u)^0 = 1$, since $v/u > 1$, and then we can apply Theorem 3.27 again. Now from (11.5), we get that

$$a^{xy} = \lim_{n\to\infty} a^{r_n s_n} \le (a^x)^y.$$

The inequality $a^{xy} \ge (a^x)^y$ can be proven similarly if we take sequences $r_n \to x$ and $s_n \to y$ consisting of rational numbers such that $r_n > x$ and $s_n > y$ for all n. □

We note that by the second identity of (11.4),

$$a^x \cdot a^{-x} = a^{x+(-x)} = a^0 = 1,$$

and so $a^{-x} = 1/a^x$ holds for all $a > 0$ and real numbers x.

Now we can continue and define exponential and power functions. If in the power a^b we consider the base to be fixed and let the exponent vary, then we get the exponential functions; if we consider the exponent to be fixed and let the base be a variable, then we get the power functions. The precise definition is the following.

Definition 11.5. Given arbitrary $a > 0$, the function $x \mapsto a^x$ $(x \in \mathbb{R})$ is called the *exponential function with base a*.

Given arbitrary $b \in \mathbb{R}$, the function $x \mapsto x^b$ $(x > 0)$ is called the *power function with exponent b*.

Since $1^x = 1$ for all x, the function that is identically 1 is one of the exponential functions. Similarly, by $x^0 = 1$ $(x > 0)$ and $x^1 = x$, the functions 1 and x are power functions over the interval $(0, \infty)$.

The most important properties of exponential functions are summarized by the following theorem.

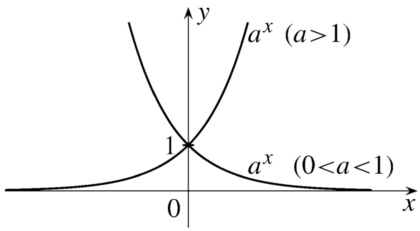

Fig. 11.1

Theorem 11.6.

(i) *If* $a > 1$, *then the exponential function* a^x *is positive everywhere, strictly monotone increasing, and continuous on* $\mathbb{R}$. *Moreover,*

$$\lim_{x\to\infty} a^x = \infty \quad \textit{and} \quad \lim_{x\to-\infty} a^x = 0. \tag{11.6}$$

(ii) *If* $0 < a < 1$, *then the exponential function* a^x *is positive everywhere, strictly monotone decreasing, and continuous on* $\mathbb{R}$. *Moreover,*

$$\lim_{x\to\infty} a^x = 0 \quad \textit{and} \quad \lim_{x\to-\infty} a^x = \infty.$$

(iii) *Given arbitrary* $a > 0$, *the function* a^x *is convex on* $\mathbb{R}$ *(Figure 11.1).*

Proof. We have already seen in Theorem 3.27 that if $a > 1$, then the function a^x is positive and strictly monotone increasing. Thus by Theorem 10.68, the limits $\lim_{x\to\infty} a^x$ and $\lim_{x\to-\infty} a^x$ exist. And since $a^n \to \infty$ and $a^{-n} \to 0$ if $n \to \infty$, (11.6) holds. The analogous statements when $0 < a < 1$ can be proven similarly.

The continuity of exponential functions is clear by Lemma 11.3, using Theorem 10.19. Thus we have proved statements (i) and (ii).

Let $a > 0$ and $x, y \in \mathbb{R}$. If we apply the inequality between the arithmetic and geometric means to the numbers a^x and a^y, then we get that

$$a^{(x+y)/2} = \sqrt{a^x \cdot a^y} \le \frac{a^x + a^y}{2}.$$

This means that the function a^x is weakly convex. Since it is continuous, it is convex by Theorem 10.76. □

The corresponding properties of power functions are given by the following theorem (Figure 11.2).

Theorem 11.7.

(i) *If* $b > 0$*, then the power function* x^b *is positive, strictly monotone increasing, and continuous on the interval* $(0,\infty)$*, and moreover,*

$$\lim_{x\to 0+0} x^b = 0 \text{ and } \lim_{x\to\infty} x^b = \infty.$$

(ii) *If* $b < 0$*, then the power function* x^b *is positive, strictly monotone decreasing, and continuous on the interval* $(0,\infty)$*, and moreover,*

$$\lim_{x\to 0+0} x^b = \infty \text{ and } \lim_{x\to\infty} x^b = 0.$$

(iii) *If* $b \geq 1$ *or* $b \leq 0$*, then the function* x^b *is convex on* $(0,\infty)$*. If* $0 \leq b \leq 1$*, then* x^b *is concave on* $(0,\infty)$*.*

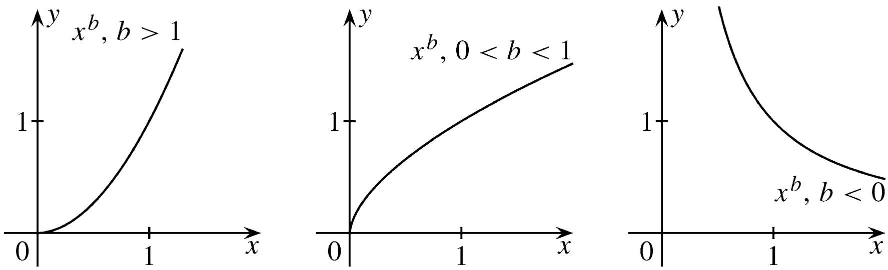

Fig. 11.2

To prove this theorem, we will require a generalization of Bernoulli's inequality (Theorem 2.5).

Theorem 11.8. *Let* $x > -1$.

(i) *If* $b \geq 1$ *or* $b \leq 0$*, then* $(1+x)^b \geq 1+bx$*.*
(ii) *If* $0 \leq b \leq 1$*, then* $(1+x)^b \leq 1+bx$*.*

Proof. We already proved the statement for nonnegative exponents in Exercise 3.33. The following simple proof is based on the convexity of the exponential function.

Let us change our notation: write a instead of x, and x instead of b. We have to show that if $a > -1$, then $x \in [0,1]$ implies $(1+a)^x \leq ax+1$, while $x \notin (0,1)$ implies $(1+a)^x \geq ax+1$. Both statements follow from the fact that the function $(1+a)^x$ is convex. We see this by noting that the chord connecting the points 0 and 1 is exactly $y = ax+1$; in other words, $h_{0,1}(x) = ax+1$. Thus if $x \in [0,1]$, then the inequality $(1+a)^x \leq h_{0,1}(x)$ follows from the definition of convexity, while for $x \notin (0,1)$, we have $(1+a)^x \geq h_{0,1}(x)$ by Lemma 10.74. □

Proof (Theorem 11.7). (i) Let $b > 0$. If $t > 1$, then $t^b > t^0 = 1$ by Theorem 3.27. Now if $0 < x < y$, then

$$y^b = \left(\frac{y}{x} \cdot x\right)^b = \left(\frac{y}{x}\right)^b \cdot x^b > 1 \cdot x^b = x^b,$$

which shows that x^b is strictly monotone increasing. Since for arbitrary $K > 0$, we have $x^b > K$ if $x > K^{1/b}$, it follows that $\lim_{x\to\infty} x^b = \infty$. Similarly, for arbitrary $\varepsilon > 0$, we have $x^b < \varepsilon$ if $x < \varepsilon^{1/b}$, and so $\lim_{x\to 0+0} x^b = 0$. (In both arguments, we used that $\left(a^{1/b}\right)^b = a^1 = a$ for all $a > 0$.)

Let $x_0 > 0$ be given; we will see that the function x^b is continuous at x_0. If $0 < \varepsilon < x_0^b$, then by the monotonicity of the power function with exponent $1/b$, we have

$$\left(x_0^b - \varepsilon\right)^{1/b} < x_0 < \left(x_0^b + \varepsilon\right)^{1/b}.$$

Now if

$$\left(x_0^b - \varepsilon\right)^{1/b} < x < \left(x_0^b + \varepsilon\right)^{1/b},$$

then

$$x_0^b - \varepsilon < x^b < x_0^b + \varepsilon.$$

This proves the continuity of x^b.
Statement (ii) can be proved the same way.
(iii) Since the function x^b is continuous, it suffices to show that for $b \ge 1$ and $b \le 0$, it is weakly convex, and for $0 \le b \le 1$, it is weakly concave.

Consider first the case $b \ge 1$ or $b \le 0$. We have to show that

$$\left(\frac{x+y}{2}\right)^b \le \frac{x^b + y^b}{2} \tag{11.7}$$

for all $x, y > 0$. Let us introduce the notation $(x+y)/2 = t$, $x/t = u$, $y/t = v$. Then $u + v = 2$. By statement (i) of Theorem 11.8, $u^b \ge 1 + b\cdot(u-1)$ and $v^b \ge 1 + b\cdot(v-1)$. Thus

$$\frac{u^b + v^b}{2} \ge 1 + b \cdot \frac{u+v-2}{2} = 1.$$

If we multiply this inequality through by t^b, then we get (11.7).

Now suppose that $0 \le b \le 1$; we have to show that $\left((x+y)/2\right)^b \ge (x^b + y^b)/2$. We get this result by repeating the argument we used above, except that we apply statement (ii) of Theorem 11.8 instead of (i). □

As another application of Theorem 11.8, we inspect the function $\left(1 + 1/x\right)^x$.

Theorem 11.9. *The function $f(x) = \left(1 + 1/x\right)^x$ is monotone increasing on the intervals $(-\infty, -1)$ and $(0, \infty)$, and*

$$\lim_{x\to-\infty}\left(1 + \frac{1}{x}\right)^x = \lim_{x\to\infty}\left(1 + \frac{1}{x}\right)^x = e. \tag{11.8}$$

Proof. If $0 < x < y$, then $y/x > 1$, so by Theorem 11.8,

$$\left(1+\frac{1}{y}\right)^{y/x} \geq 1+\frac{y}{x}\cdot\frac{1}{y} = 1+\frac{1}{x},$$

and so by the monotonicity of the power function,

$$\left(1+\frac{1}{y}\right)^{y} \geq \left(1+\frac{1}{x}\right)^{x}. \tag{11.9}$$

If, however, $x < y < -1$, then $0 < y/x < 1$, so again by Theorem 11.8,

$$\left(1+\frac{1}{y}\right)^{y/x} \leq 1+\frac{y}{x}\cdot\frac{1}{y} = 1+\frac{1}{x},$$

and then by the monotone decreasing property of the power function with exponent x, we get (11.9) again. Thus we have shown that f is monotone increasing on the given intervals.

By Theorem 10.68, it then follows that f has limits at infinity in both directions. Since $(1+1/n)^n \to e$ (since this was the definition of the number e), the limit at infinity can only be e. On the other hand,

$$\left(1-\frac{1}{n}\right)^{-n} = \left(\frac{n-1}{n}\right)^{-n} = \left(\frac{n}{n-1}\right)^{n} = \left(1+\frac{1}{n-1}\right)^{n-1}\cdot\left(1+\frac{1}{n-1}\right),$$

so

$$\lim_{n\to-\infty}\left(1+\frac{1}{n}\right)^{n} = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{-n} = e.$$

This gives us the first equality of (11.8). □

We can generalize Theorem 11.9 in the following way.

Theorem 11.10. *For an arbitrary real number b,*

$$\lim_{x\to-\infty}\left(1+\frac{b}{x}\right)^{x} = \lim_{x\to\infty}\left(1+\frac{b}{x}\right)^{x} = e^{b}. \tag{11.10}$$

Proof. The statement is clear for $b = 0$. If $b > 0$, then using Theorem 10.41 for limits of compositions yields

$$\lim_{x\to\infty}\left(1+\frac{b}{x}\right)^{x/b} = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^{x} = e,$$

and so by the continuity of power functions with exponent b,

$$\lim_{x\to\infty}\left(1+\frac{b}{x}\right)^{x} = \lim_{x\to\infty}\left[\left(1+\frac{b}{x}\right)^{x/b}\right]^{b} = e^{b}.$$

We can find the limit at $-\infty$ similarly, as well as in the $b < 0$ case. □

Corollary 11.11. *For an arbitrary real number b,* $\lim_{h\to 0}(1+bh)^{1/h} = e^b$.

Proof. Applying Theorem 10.41 twice, we get

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

□

By Theorem 11.10,

$$\lim_{n\to\infty}\left(1+\frac{b}{n}\right)^n = e^b \tag{11.11}$$

for all real numbers b. This fact has several important applications.

Examples 11.12. **1.** Suppose that a bank pays p percent yearly interest on a bond. A 1-dollar bond would then grow to $a+q$ dollars after a year, where $q = p/100$. If, however, after half a year, the bank pays interest of $p/2$ percent, and from this point on, the new interest applies to this increased value, then at the end of the year, our bond will be worth $1+(q/2)+[1+(q/2)]\cdot(q/2) = [1+(q/2)]^2$ dollars. If now we divide the year into n equal parts, and the interest is added to our bond after every $1/n$ year, which is then included in calculating interest from that point on, then by the end of the year, the bond will be worth $[1+(q/n)]^n$ dollars. This sequence is monotone increasing (why?), and as we have seen, its limit is $e^q = e^{p/100}$. This means that no matter how often we add the interest to our bond during the year, its value cannot exceed $e^{p/100}$, but it can get arbitrarily close to it.
2. In this application, we inspect how much a certain material (say a window pane) absorbs a certain radiation (say a fixed wavelength of light). The amount of absorption is a function of the thickness of the material. This function is not linear, but experience shows that for thin slices of the material, a linear approximation works well. This means that there exists a positive constant α (called the absorption coefficient) such that a slice of the material with thickness h absorbs about $\alpha\cdot h$ of entering radiation if h is sufficiently small.

Consider a slab of the material with thickness h, where h is now arbitrary. Subdivide this slab into n equal slices. If n is sufficiently large, then each slice with thickness h/n will absorb $\alpha\cdot(h/n)$ of the radiation that has made it that far. Thus after the ith slice, $(1-(\alpha h/n))^i$ of the radiation remains, so after the light leaves the whole slab, $1-(1-(\alpha h/n))^n$ of it is absorbed. If we let n go to infinity, we get that a slab of thickness h will absorb $1-e^{-\alpha h}$ of the radiation.

The following theorem gives an interesting characterization of exponential functions.

Theorem 11.13. *The function* $f\colon \mathbb{R}\to\mathbb{R}$ *is an exponential function if and only if it is continuous, not identically zero, and satisfies the identity*

$$f(x_1+x_2) = f(x_1)\cdot f(x_2) \tag{11.12}$$

for all $x_1, x_2 \in \mathbb{R}$.

Proof. We already know that the conditions are satisfied for an exponential function. Suppose now that f satisfies these conditions, and let $a = f(1)$. We will show that $a > 0$ and $f(x) = a^x$ for all x.

Since $f(x) = f\big((x/2)+(x/2)\big) = f(x/2)^2$ for all x, f is nonnegative everywhere. If f vanishes at a point x_0, then by the identity $f(x) = f\big((x-x_0)+x_0\big) = f(x-x_0)\cdot f(x_0)$, it would follow that f is identically zero, which we have excluded. Thus f is positive everywhere, so specifically $a = f(1)$ is positive. Since $f(0) = f(0+0) = f(0)^2$ and $f(0) > 0$, we have $f(0) = 1$. Thus we get that $1 = f(0) = f(x+(-x)) = f(x)\cdot f(-x)$, so $f(-x) = 1/f(x)$ for all x.

From assumption (11.12), using induction we know that

$$f(x_1+\cdots+x_n) = f(x_1)\cdot\ldots\cdot f(x_n)$$

holds for all n and $x_1,\ldots,x_n$. Then by the choice $x_1 = \cdots = x_n = x$, we get $f(nx) = f(x)^n$. Thus if p and q are positive integers, then

$$\left(f\left(\frac{p}{q}\right)\right)^q = f\left(\frac{p}{q}\cdot q\right) = f(p\cdot 1) = f(1)^p = a^p,$$

that is, $f(p/q) = a^{p/q}$. Since $f(-p/q) = 1/f(p/q) = 1/a^{p/q} = a^{-p/q}$, we have shown that $f(r) = a^r$ for all rational numbers r.

If x is an arbitrary real number, then let (r_n) be a sequence of rational numbers that tends to x. Since f is continuous on x,

$$f(x) = \lim_{n\to\infty} f(r_n) = \lim_{n\to\infty} a^{r_n} = a^x.$$

□

Remark 11.14. Theorem 11.13 characterizes exponential functions with the help of a **functional equation**. We ran into a similar functional equation in Chapter 8, when we mentioned functions that satisfy

$$f(x_1+x_2) = f(x_1)+f(x_2) \tag{11.13}$$

while talking about weakly convex functions; this functional equation is called **Cauchy's functional equation**. We also mentioned that there exist solutions of (11.13) that are not continuous. These solutions cannot be bounded from above on any interval, by Exercise 10.94. If f is such a function, then the function $e^{f(x)}$ satisfies the functional equation (11.12), and it is not bounded in any interval. This remark shows that in Theorem 11.13, the continuity condition cannot be dropped (although it can be weakened).

We will run into two relatives of the functional equations above, whose continuous solutions are exactly the power functions and logarithmic functions (see Exercises11.15 and 11.26). Equally noteworthy is **d'Alembert's**[3] **functional equation:**

$$f(x_1+x_2)+f(x_1-x_2)=2f(x_1)f(x_2).$$

See Exercises 11.35 and 11.44 regarding the continuous solutions of this functional equation.

Generalized mean. If $a>0$ and $b\neq 0$, then we will also denote the power $a^{1/b}$ by $\sqrt[b]{a}$ (which is strongly connected to the fact that $a^{1/k}=\sqrt[k]{a}$ for positive integers k by definition). Let $a_1,\ldots,a_n$ be positive numbers, and let $b\neq 0$. The quantity

$$G(b;a_1,\ldots,a_n)=\sqrt[b]{\frac{a_1^b+\cdots+a_n^b}{n}}$$

is called the **generalized mean with exponent** b of the numbers a_i. It is clear that $G(-1;a_1,\ldots,a_n)$, $G(1;a_1,\ldots,a_n)$, and $G(2;a_1,\ldots,a_n)$ are exactly the harmonic, arithmetic, and quadratic means of the numbers a_i. We also consider the geometric mean as a generalized mean, since we define the generalized mean with exponent 0 by

$$G(0;a_1,\ldots,a_n)=\sqrt[n]{a_1\cdot\ldots\cdot a_n}.$$

(To see the motivation behind this definition, see Exercise 12.35.)

Let $b\geq 1$. Apply Jensen's inequality (Theorem 9.19) to x^b. We get that

$$\left(\frac{a_1+\cdots+a_n}{n}\right)^b\leq\frac{a_1^b+\cdots+a_n^b}{n}.$$

Raising both sides to the power $1/b$, we get the inequality

$$\frac{a_1+\cdots+a_n}{n}\leq G(b;a_1,\ldots,a_n)=\sqrt[b]{\frac{a_1^b+\cdots+a_n^b}{n}}.$$

This is called the **generalized mean inequality** (which implies the inequality of arithmetic and quadratic means as a special case). However, this inequality is also just a special case of the following theorem.

Theorem 11.15. *Let $a_1,\ldots,a_n$ be fixed positive numbers. Then the function*

$$b\mapsto G(b;a_1,\ldots,a_n)\ \ (b\in\mathbb{R})$$

is monotone increasing on $\mathbb{R}$.

Proof. Suppose first that $0<b<c$. Then $c/b>1$. Apply Jensen's inequality to the function $x^{c/b}$ and the numbers a_i^b $(i=1,\ldots,n)$. We get that

[3] Jean le Rond d'Alembert (1717–1783), French mathematician.

$$\left(\frac{a_1^b + \cdots + a_n^b}{n}\right)^{c/b} \le \frac{a_1^c + \cdots + a_n^c}{n}.$$

If we raise both sides to the power $1/c$, then we get that

$$G(b; a_1, \ldots, a_n) \le H(c; a_1, \ldots, a_n). \tag{11.14}$$

Now let $b = 0$ and $c > 0$. Apply the inequality of arithmetic and geometric means for the numbers a_i^c $(i = 1, \ldots, n)$. We get that

$$G(0; a_1, \ldots, a_n)^c = \sqrt[n]{a_1^c \cdot \ldots \cdot a_n^c} \le \frac{a_1^c + \cdots + a_n^c}{n},$$

and if here we raise both sides to the power $1/c$, then we get (11.14) again.

It is easy to check that

$$G(-b; a_1, \ldots, a_n) = \frac{1}{G\left(b; \frac{1}{a_1}, \ldots, \frac{1}{a_n}\right)}$$

for all b. So using the inequalities we just proved, we get that for $b < c \le 0$, we have

$$G(b; a_1, \ldots, a_n) = \frac{1}{G\left(-b; \frac{1}{a_1}, \ldots, \frac{1}{a_n}\right)} \le \frac{1}{G\left(-c; \frac{1}{a_1}, \ldots, \frac{1}{a_n}\right)} = \\ = G(c; a_1, \ldots, a_n),$$

which proves the theorem. □

As we saw in Theorem 11.7, $\lim_{x \to 0+0} x^b = 0$ if $b > 0$; then it is in our best interest for all positive powers of zero to be zero. Thus we define $0^b = 0$ for all $b > 0$ (but we still do not define the nonpositive powers of zero). Regarding this change, for $b > 0$ we can extend the power function x^b to be defined at zero as well, where its value is zero. The new extended function x^b is then continuous from the right at zero if $b > 0$.

Exercises

11.8. Prove that for the numbers $0 < a < b$, $a^b = b^a$ holds if and only if there exists a positive number x such that

$$a = \left(1 + \frac{1}{x}\right)^x \quad \text{and} \quad b = \left(1 + \frac{1}{x}\right)^{x+1}.$$

11.9. Prove that if $a > 0$ and $a \ne 1$, then the function a^x is strictly convex.

11.10. Prove that if $b > 1$ or $b < 0$, then the function x^b is strictly convex.

11.11. Prove that if $0 < b < 1$, then the function x^b is strictly concave.

11.12. Let $x > -1$ and $b \in \mathbb{R}$. Prove that $(1+x)^b = 1 + bx$ holds if and only if at least one of $x = 0$, $b = 0$, $b = 1$ holds.

11.13. $\lim_{x\to -1-0}\left(1+\frac{1}{x}\right)^x = ?$

11.14. Let $G(b; a_1, \ldots, a_n)$ be the generalized mean with exponent b of the positive numbers $a_1, \ldots, a_n$. Prove that

$$\lim_{b\to\infty} G(b; a_1, \ldots, a_n) = \max(a_1, \ldots, a_n)$$

and

$$\lim_{b\to -\infty} G(b; a_1, \ldots, a_n) = \min(a_1, \ldots, a_n).$$

11.15. Prove that the function $f : (0,\infty) \to \mathbb{R}$ is a power function if and only if it is continuous, not identically zero, and satisfies the identity

$$f(x_1 \cdot x_2) = f(x_1) \cdot f(x_2)$$

for all positive x_1, x_2.

11.3 Logarithmic Functions

If $a > 0$ and $a \neq 1$, then the function a^x is strictly monotone and continuous on $\mathbb{R}$ by Theorem 11.6. Thus if $a > 0$ and $a \neq 1$, then a^x has an inverse function, which we call the **logarithm to base** a, and we denote it by $\log_a x$. Since the image of a^x is $(0,\infty)$, the function $\log_a x$ is defined on the open interval $(0,\infty)$, but its image is $\mathbb{R}$. Having the definition of the inverse function in mind, we come to the conclusion that if $a > 0$, $a \neq 1$, and $x > 0$, then $\log_a x$ is the only real number for which $a^{\log_a x} = x$ holds. Specifically, $\log_a 1 = 0$ and $\log_a a = 1$ (Figure 11.3).

Theorem 11.16.

(i) *If* $a > 1$, *then the function* $\log_a x$ *is strictly monotone increasing, continuous, and strictly concave on* $(0,\infty)$. *Moreover,*

$$\lim_{x\to 0+0} \log_a x = -\infty \quad \textit{and} \quad \lim_{x\to\infty} \log_a x = \infty. \qquad (11.15)$$

(ii) *If* $0 < a < 1$, *then the function* $\log_a x$ *is strictly monotone decreasing, continuous, and strictly convex on* $(0,\infty)$. *Moreover,*

$$\lim_{x\to 0+0} \log_a x = \infty \quad \textit{and} \quad \lim_{x\to\infty} \log_a x = -\infty. \qquad (11.16)$$

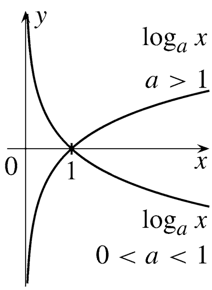

Fig. 11.3

(iii) *For all* $a > 0,\ a \neq 1$, *and* $x, y > 0$, *the identities*

$$\begin{aligned}\log_a(xy) &= \log_a x + \log_a y,\\ \log_a(x/y) &= \log_a x - \log_a y,\\ \log_a(1/y) &= -\log_a y,\end{aligned} \tag{11.17}$$

as well as

$$\log_a(x^y) = y \cdot \log_a x \quad \textit{and} \quad \log_a \sqrt[y]{x} = \tfrac{1}{y} \cdot \log_a x \tag{11.18}$$

hold.

Proof. As an application of Theorems 11.6 and 10.72, we get that the function $\log_a x$ is continuous everywhere; if $a > 1$, it is strictly monotone increasing; and if $0<a<1$, it is strictly monotone decreasing. Since its image is $\mathbb{R}$, the limit relations (11.15) and (11.16) follow from this. Thus we have proved (i) and (ii) except for the statements about convexity and concavity.

Let $a > 0$, $a \neq 1$. By the second identity in (11.4), we obtain that for arbitrary $x, y > 0$ we have

$$a^{\log_a x + \log_a y} = a^{\log_a x} \cdot a^{\log_a x} = x \cdot y = a^{\log_a(xy)}.$$

Since the function a^x is strictly monotone, we get the first identity (11.17) from this. Thus

$$\log_a x = \log_a\left(\frac{x}{y} \cdot y\right) = \log_a\left(\frac{x}{y}\right) + \log_a y,$$

which is the second identity in (11.17). Applying this to $x = 1$ gives us the third one. If we now apply the third identity of (11.4), then we get that

$$a^{\log_a(x^y)} = x^y = \left(a^{\log_a x}\right)^y = a^{y \cdot \log_a x},$$

which implies the first identity of (11.18). Applying this to $1/y$ instead of y will yield the second identity of (11.18), where we note that $\sqrt[y]{x} = x^{1/y}$.

If $0 < x < y$, then by the inequality of arithmetic and geometric means, $\sqrt{xy} < (x+y)/2$. Applying (11.17) and (11.18), we get that if $a > 1$, then

$$\frac{\log_a x + \log_a y}{2} < \log_a\left(\frac{x+y}{2}\right),$$

where we note also that the function $\log_a$ is strictly monotone increasing for $a > 1$. This means that if $a > 1$, then $\log_a$ is strictly weakly concave. Similarly, we get that if $0 < a < 1$, then the function $\log_a$ is strictly weakly convex. Since we are talking about continuous functions, when $a > 1$, the function $\log_a x$ is strictly concave, while if $0 < a < 1$, then it is strictly convex. This concludes the proof of the theorem. □

Remark 11.17. Let a and b be positive numbers different from 1. Then by applying (11.18), we get

$$\log_a x = \log_a\left(b^{\log_b x}\right) = \log_b x \cdot \log_a b,$$

which gives us

$$\log_b x = \frac{\log_a x}{\log_a b} \tag{11.19}$$

for all $x > 0$. This means that *two logarithmic functions differ by only a constant multiple.* Thus it is useful to choose one logarithmic function and write the rest of them as multiples of this one. But which one should we choose among the infinitely many logarithmic functions? This is determined by necessity; clearly, it is most useful to choose the one that we use the most. Often in engineering, the base-10 logarithm is used, while in computer science, it is base 2. We will use the logarithmic function at base e, since this will make the equations that arise in differentiation the simplest. *From now on, we will write* $\log x$ *instead of* $\log_e x$. (Sometimes, the logarithm at base e is denoted by $\ln x$, which stands for *logaritmus naturalis*, the natural logarithm.)

If we apply (11.19) to $a = e$, we get that

$$\log_b x = \frac{\log x}{\log b} \tag{11.20}$$

whenever $b > 0$, $b \neq 1$, and $x > 0$. Just as every logarithmic function can be expressed using the natural logarithm, we can also express every exponential function using e^x. In fact, if $a > 0$, then by the definition of $\log a$ and by the third identity of (11.4),

$$a^x = \left(e^{\log a}\right)^x = e^{x\cdot\log a} = (e^x)^{\log a},$$

that is, a^x is a power of the function e^x.

Considering that $e > 1$, the function $\log x$ is concave. This fact makes proving an important inequality possible.

Theorem 11.18 (Hölder's Inequality[4]). *Let p and q be positive numbers such that $1/p + 1/q = 1$. Then for arbitrary real numbers $a_1, \dots, a_n$ and $b_1, \dots, b_n$,*

$$|a_1 b_1 + \cdots + a_n b_n| \le \sqrt[p]{|a_1|^p + \cdots + |a_n|^p} \cdot \sqrt[q]{|b_1|^q + \cdots + |b_n|^q}. \tag{11.21}$$

Proof. First we show that

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \tag{11.22}$$

for all $a, b \ge 0$. This is clear if $a = 0$ or $b = 0$, so we can assume that $a > 0$ and $b > 0$. Since $\log x$ is concave, by Lemma 9.17,

$$\log\left(ta^p + (1-t)b^q\right) \ge t\log a^p + (1-t)\log b^q \tag{11.23}$$

for all $0 < t < 1$. Now we apply this inequality with $t = 1/p$. Then $1 - t = 1/q$, and so the right-hand side of (11.23) becomes $\log a + \log b = \log(ab)$. Since $\log x$ is monotone increasing, we get that $(1/p)a^p + (1/q)b^q \ge ab$, which is exactly (11.22).

[4] Otto Ludwig Hölder(1859–1937), German mathematician.

Now to continue the proof of the theorem, let $A = \sqrt[p]{|a_1|^p + \cdots + |a_n|^p}$ and $B = \sqrt[q]{|b_1|^q + \cdots + |b_n|^q}$. If $A = 0$, then $a_1 = \ldots = a_n = 0$, and so (11.21) holds, since both sides are zero. The same is the case if $B = 0$, so we can assume that $A > 0$ and $B > 0$.

Let $\alpha_i = |a_i|/A$ and $\beta_i = |b_i|/B$ for all $i = 1, \ldots, n$. Then

$$\alpha_1^p + \cdots + \alpha_n^p = \beta_1^q + \cdots + \beta_n^q = 1. \tag{11.24}$$

Now by (11.22),

$$\alpha_i \beta_i \le \frac{1}{p}\alpha_i^p + \frac{1}{q}\beta_i^q$$

for all i. Summing these inequalities, we get that

$$\alpha_1\beta_1 + \cdots + \alpha_n\beta_n \le \frac{1}{p} \cdot (\alpha_1^p + \cdots + \alpha_n^p) + \frac{1}{q} \cdot (\beta_1^q + \cdots + \beta_n^q) = \frac{1}{p} \cdot 1 + \frac{1}{q} \cdot 1 = 1,$$

using (11.24) and the assumption on the numbers p and q. If now we write $|a_i|/A$ and $|b_i|/B$ in place of α_i and multiply both sides by AB, then we get that

$$|a_1 b_1| + \cdots + |a_n b_n| \le AB,$$

from which we immediately get (11.21) by the triangle inequality. □

Hölder's inequality for the special case $p = q = 2$ gives the following, also very notable, inequality.

Theorem 11.19 (Cauchy–Schwarz[5]–Bunyakovsky[6]Inequality). *For arbitrary real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$,*

$$|a_1 b_1 + \cdots + a_n b_n| \le \sqrt{a_1^2 + \cdots + a_n^2} \cdot \sqrt{b_1^2 + \cdots + b_n^2}.$$

We also give a direct proof.

Proof. For arbitrary $i, j = 1, \ldots, n$, the number

$$A_{i,j} = a_i^2 b_j^2 + a_j^2 b_i^2 - 2a_i a_j b_i b_j$$

is nonnegative, since $A_{i,j} = (a_i b_j - a_j b_i)^2$. If we add all the numbers $A_{i,j}$ for all $1 \le i < j \le n$, then we get the difference

$$(a_1^2 + \cdots + a_n^2) \cdot (b_1^2 + \cdots + b_n^2) - (a_1 b_1 + \cdots + a_n b_n)^2,$$

which is thus also nonnegative. □

[5] Hermann Amandus Schwarz (1843–1921), German mathematician.

[6] ViktorYakovlevich Bunyakovsky (1804–1889), Russian mathematician.

Exercises

11.16. Prove that $1+1/2+\cdots+1/n > \log n$ for all n. (H)

11.17. Prove that the sequence $1+1/2+\cdots+1/n-\log n$ is monotone decreasing and convergent.[7]

11.18. Let f be a strictly monotone increasing and convex (concave) function on the open interval I. Prove that the inverse of f is concave (convex).

11.19. Let f be strictly monotone decreasing and convex (concave) on the open interval I. Prove that the inverse of f is convex (concave). Check these statements for the cases that f is an exponential, power, or logarithmic function.

11.20. Prove that $\lim_{x\to 0+0} x^{\varepsilon}\cdot\log x = 0$ for all $\varepsilon > 0$.

11.21. Prove that $\lim_{x\to\infty} x^{-\varepsilon}\cdot\log x = 0$ for all $\varepsilon > 0$.

11.22. $\lim_{x\to 0+0} x^x = ?$ $\quad\lim_{x\to\infty}\sqrt[x]{x} = ?$ $\quad\lim_{x\to 0+0}(1+1/x)^x = ?$

11.23. Let $\lim_{n\to\infty} a_n = a$, $\lim_{n\to\infty} b_n = b$. When does $\lim_{n\to\infty} a_n^{b_n} = a^b$ hold?

11.24. Prove that equality holds in (11.22) only if $a^p = b^q$.

11.25. Suppose that the numbers $a_1,\ldots,a_n,b_1,\ldots,b_n$ are nonnegative. Prove that equality holds in (11.21) if and only if $a_1 = \ldots = a_n = 0$ or if there exists a number t such that $b_i^q = t\cdot a_i^p$ for all $i = 1,\ldots,n$.

11.26. Prove that the function $f\colon (0,\infty)\to\mathbb{R}$ is a logarithmic function if and only if it is continuous, not identically zero, and satisfies the identity

$$f(x_1\cdot x_2) = f(x_1)+f(x_2)$$

for all positive x_1, x_2.

11.4 Trigonometric Functions

Defining trigonometric functions. Since trigonometric functions are functions that deal with angles, or more precisely, map angles to real numbers, we first need to clarify how angles are measured.

Let O be a point in the plane, let h and k be rays starting at point O, and let $(h,k)\lessdot$ denote the section of the plane (one of them, that is) determined by h and k. Let D be a disk centered at O, and C a circle centered at O. It is clear by inspection that the angle determined by h and k is proportional to the area of the sector $D\cap(h,k)\lessdot$, as well as the length of the arc $C\cap(h,k)\lessdot$. It would be a good choice, then, to measure

[7] The limit of this sequence is called **Euler's constant**. *Whether this number is rational has been an open question for a long time.*

the angle by the area of the sector $D \cap (h,k)\sphericalangle$ or the length of the arc $C \cap (h,k)\sphericalangle$ (or any other quantity directly proportional to these). We will choose the length of the arc. We agree that an angle at the point O is measured by the length of the arc of a unit circle centered at O that subtends the angle. We call this number the **angular measure**, and the units for it are **radians**. Thus the angular measure of a straight angle is the length of the unit semicircle, that is, π radians; the angular measure of a right angle is half of this, that is, $\pi/2$ radians. From now on, we omit the word "radians," so measurements of angles (unless otherwise specified) will be given in radians.

In trigonometry, we define the cosine (or sine) of an angle x smaller than $\pi/2$ by considering the right triangle with acute angle x and taking the quotient of the length of the side adjacent (or opposite) to the angle x and the hypotenuse. Let $0 < u < 1$ and $v = c(u) = \sqrt{1-u^2}$. Then the points $O = (0,0)$, $P = (u,0)$, and $Q = (u,v)$ define a right triangle whose angle at O is defined by the rays $\overrightarrow{OP}$ and $\overrightarrow{OQ}$. The arc of the circle C that falls into this section agrees with the graph of c over the interval $[u,1]$, whose length is $s(c;[u,1])$ (see Definition 10.78). If $s(c;[u,1]) = x$, then the angular measure of the angle $POQ\sphericalangle$ is x, and so $\cos x = \overline{OP}/\overline{OQ} = u/1 = u$, and $\sin x = \overline{PQ}/\overline{OQ} = v/1 = v$. We can also formulate this by saying that *if starting from the point* $(1,0)$*, we measure out an arc of length x onto the circle C in the positive direction, then the point we end up at has coordinates* $(\cos x, \sin x)$. Here the "positive direction" means that we move in a counterclockwise direction, that is, we start measuring the arc in the upper half-plane.[8]

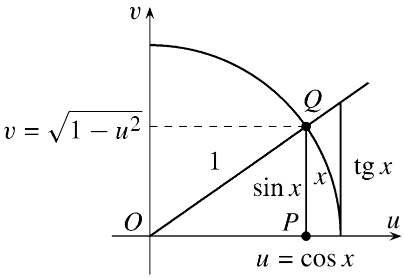

Fig. 11.4

We will accept the statement above as the definition of the functions $\cos x$ and $\sin x$.

Definition 11.20. Starting at the point $(1,0)$ on the unit circle centered at the origin, measure an arc of length x in the positive or negative direction according to whether $x > 0$ or $x < 0$. (If $|x| \geq 2\pi$, then we will circle around more than once as necessary.) The first coordinate of the point we end up at is denoted by $\cos x$, while the second coordinate is $\sin x$. Thus we have defined the functions cos and sin on the set of all real numbers.

Remark 11.21. In Remark 10.83, we saw that for $0 \leq x \leq \pi$, there exists a $u \in [-1,1]$ such that $S(u) = s(c;[u,1]) = x$. Thus the endpoint of an arc of length x measured on C is the point $(u,c(u)) = (u,\sqrt{1-u^2})$. It then follows that $\cos x = u$ and $\sin x = \sqrt{1-u^2} = \sqrt{1-\cos x^2}$. The relation $\cos x = u$ means that on the interval $[0,\pi]$, the function $\cos x$ is exactly the inverse of the function S.

[8] That is, in the part of the plane given by $\{(x,y) : y > 0\}$.

We often need to use the expressions $\sin x/\cos x$ and $\cos x/\sin x$, for which we use the notation $\operatorname{tg} x$ and $\operatorname{ctg} x$ respectively.

Properties of Trigonometric Functions. As we noted, measuring an arc of length π from any point $(\cos x, \sin x)$ of the circle C gets us into the antipodal point, whose coordinates are $(-\cos x, -\sin x)$. Thus

$$\cos(x+\pi) = -\cos x \;\text{ and }\; \sin(x+\pi) = -\sin x \tag{11.25}$$

for all x. Since $(\cos 0, \sin 0) = (1,0)$, we have $\cos 0 = 1$ and $\sin 0 = 0$. Thus by (11.25),

$$\cos(k\pi) = (-1)^k \;\text{ and }\; \sin(k\pi) = 0 \tag{11.26}$$

for all integers k. By the definition, we immediately get that

$$\cos(x+2\pi) = \cos x \;\text{ and }\; \sin(x+2\pi) = \sin x$$

for all x, that is, $\cos x$ and $\sin x$ are both periodic functions with period 2π. Since $(\cos x, \sin x)$ is a point on the circle C, we have

$$\sin^2 x + \cos^2 x = 1 \tag{11.27}$$

for all x. The circle C is symmetric with respect to the horizontal axis. So if we measure a segment of length $|x|$ in the positive or negative direction starting at the point $(1,0)$, then we arrive at a point that is symmetric with respect to the horizontal axis. This means that $(\cos(-x), \sin(-x)) = (\cos x, -\sin x)$, that is,

$$\cos(-x) = \cos x \;\text{ and }\; \sin(-x) = -\sin x \tag{11.28}$$

for all x. In other words, the function $\cos x$ is even, while the function $\sin x$ is odd. Connecting the identities (11.28) and (11.25), we obtain that

$$\cos(\pi - x) = -\cos x \;\text{ and }\; \sin(\pi - x) = \sin x \tag{11.29}$$

for all x. Substituting $x = \pi/2$ into this, we get $\cos(\pi/2) = 0$, so by (11.25), we see that

$$\cos\left(\frac{\pi}{2} + k\pi\right) = 0 \;\; (k \in \mathbb{Z}). \tag{11.30}$$

Since $\sin(\pi/2) = \sqrt{1 - \cos^2(\pi/2)} = 1$, then again by (11.25), we have

$$\sin\left(\frac{\pi}{2} + k\pi\right) = (-1)^k \;\; (k \in \mathbb{Z}). \tag{11.31}$$

The circle C is also symmetric with respect to the $45°$ line passing through the origin. If we measure x in the positive direction from the point $(1,0)$ and then reflect over this line, we get the point that we would get if we measured x in the negative direction starting at $(0,1)$. Since $(0,1) = (\cos(\pi/2), \sin(\pi/2))$, the endpoint of the arc we mirrored is the same as the endpoint of the arc of length $x - (\pi/2)$ measured from $(1,0)$ in the negative direction, which is the same as $(\pi/2) - x$ in the positive direction. Thus we see that the point $(\sin x, \cos x)$—the reflection of $(\cos x, \sin x)$ about the 45° line passing through the origin—agrees with the point $(\cos((\pi/2)-x), \sin((\pi/2)-x))$, so

$$\cos\left(\frac{\pi}{2}-x\right) = \sin x \text{ and } \sin\left(\frac{\pi}{2}-x\right) = \cos x \tag{11.32}$$

for all x. The following identities are the addition formulas for sin and cos.

$$\begin{aligned}\sin(x+y) &= \sin x\cos y + \cos x\sin y,\\ \sin(x-y) &= \sin x\cos y - \cos x\sin y,\\ \cos(x+y) &= \cos x\cos y - \sin x\sin y,\\ \cos(x-y) &= \cos x\cos y + \sin x\sin y.\end{aligned} \tag{11.33}$$

The proofs of the addition formulas are outlined in the first appendix of the chapter. The proof there is based on rotations about the origin. Later, using differentiation, we will give a proof that does not use geometric concepts and does not rely on geometric inspections (see the second appendix of Chapter 13).

The following identities are simple consequences of the addition formulas.

$$\begin{aligned}\sin 2x &= 2\sin x\cos x,\\ \cos 2x &= \cos^2 x - \sin^2 x = 1 - 2\sin^2 x = 2\cos^2 x - 1,\end{aligned} \tag{11.34}$$

$$\cos^2 x = \frac{1+\cos 2x}{2}, \quad \sin^2 x = \frac{1-\cos 2x}{2}, \tag{11.35}$$

$$\begin{aligned}\cos x\cos y &= \tfrac{1}{2}\left(\cos(x+y)+\cos(x-y)\right),\\ \sin x\sin y &= \tfrac{1}{2}\left(\cos(x-y)-\cos(x+y)\right),\\ \sin x\cos y &= \tfrac{1}{2}\left(\sin(x+y)+\sin(x-y)\right),\end{aligned} \tag{11.36}$$

$$\begin{aligned}\sin x + \sin y &= 2\sin\frac{x+y}{2}\cos\frac{x-y}{2},\\ \sin x - \sin y &= 2\sin\frac{x-y}{2}\cos\frac{x+y}{2},\\ \cos x + \cos y &= 2\cos\frac{x+y}{2}\cos\frac{x-y}{2},\\ \cos x - \cos y &= -2\sin\frac{x-y}{2}\sin\frac{x+y}{2}.\end{aligned} \tag{11.37}$$

Now we turn to the analytic properties of the functions sin and cos (Figure 11.5).

Theorem 11.22.

(i) *The function* $\cos x$ *is strictly monotone decreasing on the intervals* $[2k\pi, (2k+1)\pi]$ *and strictly monotone increasing on the intervals* $[(2k-1)\pi, 2k\pi]$, *($k \in \mathbb{Z}$). The only roots of the function* $\cos x$ *are the points* $(\pi/2)+k\pi$.

(ii) *The function* $\sin x$ *is strictly monotone increasing on the intervals* $[2k\pi-(\pi/2), 2k\pi+(\pi/2)]$ *and strictly monotone decreasing on the intervals* $[2k\pi+(\pi/2), 2k\pi+(3\pi/2)]$, *($k \in \mathbb{Z}$). The only roots of the function* $\sin x$ *are the points* $k\pi$.

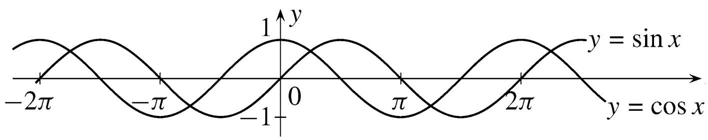

Fig. 11.5

Proof. (i) By remark 11.21, on the interval $[0,\pi]$ the function cos is none other than the inverse of the function S. Since S is strictly monotone decreasing on $[-1,1]$, the inverse function $\cos x$ is also strictly monotone decreasing on $[0,\pi]$. Thus by (11.25), it is clear that if $k \in \mathbb{Z}$ is even, then $\cos x$ is strictly monotone decreasing on $[2k\pi,(2k+1)\pi]$ and is strictly monotone increasing on $[(2k-1)\pi, 2k\pi]$. Now the statement about the roots of $\cos x$ is clear from this.

Statement (ii) follows from (i) via the identity (11.32). □

The following inequalities are especially important for applications.

Theorem 11.23. *For all x, the equations*

$$|\sin x| \le |x| \tag{11.38}$$

and

$$0 \le 1 - \cos x \le x^2 \tag{11.39}$$

hold.

Proof. It suffices to prove the inequality (11.38) for nonnegative x, since both sides are even. If $x > \pi/2$, then $|\sin x| \le 1 < \pi/2 < x$, and the statement holds. We can suppose that $0 \le x \le \pi/2$. Let $u = \cos x$ and $v = \sin x$. Then by the definition of $\cos x$ and $\sin x$, in the graph of the function $c(t) = \sqrt{1-t^2}$, the arc length over the interval $[u,1]$ is exactly x, since $(\cos x, \sin x) = (u,v)$ are the coordinates of the points that we get by measuring an arc of length x onto the circle C; see Figure 11.4. Then by Theorem 10.79,

$$0 \le \sin x = v \le \sqrt{(1-u)^2 + (v-0)^2} \le s(c;[u,1]) = x,$$

which proves (11.38).

For inequality (11.39), it again suffices to consider only the nonnegative x, since cos is an even function. If $x > \pi/2$, then

$$1 - \cos x \le 2 < \left(\frac{3}{2}\right)^2 < \left(\frac{\pi}{2}\right)^2 < x^2,$$

and so (11.39) is true. If, however, $0 \le x \le \pi/2$, then $\cos x \ge 0$, and so

$$1-\cos x = \frac{1-\cos^2 x}{1+\cos x} = \frac{\sin^2 x}{1+\cos x} \le \sin^2 x \le x^2,$$

which gives (11.39). □

Theorem 11.24. *For all $x, y \in \mathbb{R}$, the inequalities*

$$|\cos x - \cos y| \le |x-y| \tag{11.40}$$

and

$$|\sin x - \sin y| \le |x-y| \tag{11.41}$$

hold.

Proof. By Theorem 11.23 and the identity (11.37), we get that

$$|\cos x - \cos y| = 2\cdot\left|\sin\frac{x-y}{2}\right|\cdot\left|\sin\frac{x+y}{2}\right| \le 2\cdot\left|\frac{x-y}{2}\right|\cdot 1 = |x-y|$$

and

$$|\sin x - \sin y| = 2\cdot\left|\sin\frac{x-y}{2}\right|\cdot\left|\cos\frac{x+y}{2}\right| \le 2\cdot\left|\frac{x-y}{2}\right|\cdot 1 = |x-y|.$$

□

Theorem 11.25.

(i) *The functions* sin *and* cos *are continuous everywhere, and in fact, they are Lipschitz.*
(ii) *The function* $\sin x$ *is strictly concave on the intervals* $[2k\pi, (2k+1)\pi]$ *and strictly convex on the intervals* $[(2k-1)\pi, 2k\pi]$, *($k \in \mathbb{Z}$).*
(iii) *The function* $\cos x$ *is strictly concave on the intervals* $[2k\pi - (\pi/2), 2k\pi + (\pi/2)]$ *and strictly convex on the intervals* $[2k\pi + (\pi/2), 2k\pi + (2\pi/2)]$, *($k \in \mathbb{Z}$).*

Proof. Statement (i) is clear by the previous theorem.

If $0 \le x < y \le \pi$, then applying the first identity in (11.37) yields

$$\frac{\sin x + \sin y}{2} = \sin\frac{x+y}{2}\cos\frac{x-y}{2} < \sin\frac{x+y}{2}.$$

This shows that $\sin x$ is strictly weakly concave on the interval $[0,\pi]$. Since it is continuous, it is strictly concave there. Then statement (ii) follows immediately by the identity $\sin(x+\pi) = -\sin x$.

Finally, statement (iii) follows from (ii) using the identity (11.32). □

Theorem 11.26. *If* $|x| < \pi/2$ *and* $x \neq 0$, *then*

$$\cos x \leq \frac{\sin x}{x} \leq 1. \tag{11.42}$$

Proof. Since the function $(\sin x)/x$ is even, it suffices to consider the case $x > 0$. The inequality $(\sin x)/x \leq 1$ is clear by (11.38).

So all we need to show is that if $0 < x < \pi/2$, then

$$x \leq \operatorname{tg} x. \tag{11.43}$$

In Figure 11.6, we have $A = (\cos x, 0)$, $B = (\cos x, \sin x)$. For the circle K, the tangent at point B intersects the horizontal axis at point C. The reflection of the point B over the horizontal axis is D. Since OAB and OBC are similar triangles, $\overline{BC} = \sin x/\cos x = \operatorname{tg} x$.

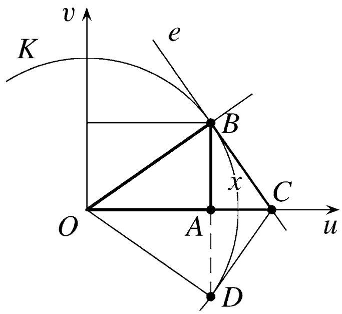

Fig. 11.6

Let us inscribe an arbitrary polygonal path in the arc DB. Extending this with segments OD and OB gives us a convex polygon T, which is part of the quadrilateral $ODCB$. By Lemma 10.82 it then follows that the perimeter of T is at most the perimeter of $ODCB$, which is $2 + 2\operatorname{tg} x$. Since the supremum of polygonal paths inscribed into DB is $2x$, we get that $2 + 2x \leq 2 + 2\operatorname{tg} x$, which proves (11.43). □

Theorem 11.27. *The limit relations*

$$\lim_{x \to 0} \frac{1 - \cos x}{x} = 0 \tag{11.44}$$

and

$$\lim_{x \to 0} \frac{\sin x}{x} = 1 \tag{11.45}$$

hold.

Proof. The two statements follow from inequalities (11.39) and (11.42) by applying the squeeze theorem. □

Now we summarize the properties of the functions $\operatorname{tg} x$ and $\operatorname{ctg} x$. The function $\operatorname{tg} x = \sin x / \cos x$ is defined where the denominator is nonzero, that is, at the points $x \neq (\pi/2) + k\pi$, where k is an arbitrary integer. By the addition formulas of sin and cos, it is easy to deduce the following identities:

$$\begin{aligned} \operatorname{tg}(x+y) &= \frac{\operatorname{tg} x + \operatorname{tg} y}{1 - \operatorname{tg} x \cdot \operatorname{tg} y}, \\ \operatorname{tg}(x-y) &= \frac{\operatorname{tg} x - \operatorname{tg} y}{1 + \operatorname{tg} x \cdot \operatorname{tg} y}, \\ \operatorname{tg} 2x &= \frac{2 \operatorname{tg} x}{1 - \operatorname{tg}^2 x}. \end{aligned} \tag{11.46}$$

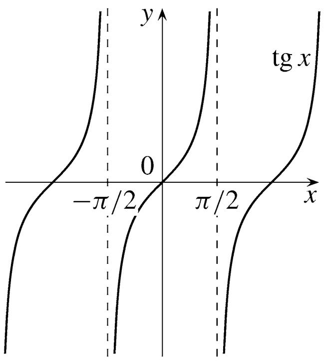

Fig. 11.7

The function $\operatorname{tg} x$ is continuous on its domain, since it is the quotient of two continuous functions. The function $\operatorname{tg} x$ is odd, and it is periodic with period π, since

$$\operatorname{tg}(x+\pi) = \frac{\sin(x+\pi)}{\cos(x+\pi)} = \frac{-\sin x}{-\cos x} = \frac{\sin x}{\cos x} = \operatorname{tg} x$$

for all $x \neq (\pi/2) + k\pi$. Since on the interval $[0, \pi/2)$, the function $\sin x$ is strictly monotone increasing, $\cos x$ is strictly monotone decreasing, and both are positive, so there $\operatorname{tg} x$ is strictly monotone increasing. Since $\operatorname{tg} 0 = 0$ and $\operatorname{tg} x$ is odd, $\operatorname{tg} x$ is strictly increasing on the whole interval $(-\pi/2, \pi/2)$.

The limit relations

$$\lim_{x \to -(\pi/2)+0} \operatorname{tg} x = -\infty \qquad \text{and} \qquad \lim_{x \to (\pi/2)-0} \operatorname{tg} x = \infty \tag{11.47}$$

hold, following from the facts

$$\lim_{x \to \pm\pi/2} \sin x = \sin(\pm\pi/2) = \pm 1, \qquad \lim_{x \to \pm\pi/2} \cos x = \cos(\pm\pi/2) = 0,$$

and that $\cos x$ is positive on $(-\pi/2, \pi/2)$ (Figure 11.7).

The function $\operatorname{ctg} x = \cos x / \sin x$ is defined where the denominator is not zero, that is, at the points $x \neq k\pi$, where k is an arbitrary integer. The function $\operatorname{ctg} x$ is continuous on its domain, is odd, and is periodic with period π. By (11.32) we get that

$$\operatorname{ctg} x = \operatorname{tg}\left(\frac{\pi}{2} - x\right) \tag{11.48}$$

for all $x \neq k\pi$. It then follows that the function $\operatorname{ctg} x$ is strictly monotone decreasing on the interval $(0, \pi)$, and that

$$\lim_{x \to 0+0} \operatorname{ctg} x = \infty \qquad \text{and} \qquad \lim_{x \to \pi-0} \operatorname{ctg} x = -\infty. \tag{11.49}$$

Exercises

11.27. Prove the following equalities:

(a) $\cos\frac{\pi}{6}=\frac{\sqrt{3}}{2}$, (b) $\cos\frac{\pi}{4}=\frac{\sqrt{2}}{2}$, (c) $\cos\frac{\pi}{3}=\frac{1}{2}$,

(d) $\cos\frac{2\pi}{3}=-\frac{1}{2}$, (e) $\cos\frac{3\pi}{4}=-\frac{\sqrt{2}}{2}$, (f) $\cos\frac{5\pi}{6}=-\frac{\sqrt{3}}{2}$,

(g) $\sin\frac{\pi}{6}=\frac{1}{2}$, (h) $\sin\frac{\pi}{4}=\frac{\sqrt{2}}{2}$, (i) $\sin\frac{\pi}{3}=\frac{\sqrt{3}}{2}$,

(j) $\sin\frac{2\pi}{3}=\frac{\sqrt{3}}{2}$, (k) $\sin\frac{3\pi}{4}=\frac{\sqrt{2}}{2}$, (l) $\sin\frac{5\pi}{6}=\frac{1}{2}$.

11.28. Prove that

$$\sin 3x = 4\cdot\sin x\cdot\sin\left(x+\frac{\pi}{3}\right)\cdot\sin\left(x+\frac{2\pi}{3}\right)$$

for all x.

11.29. Prove that

$$\sin 4x = 8\cdot\sin x\cdot\sin\left(x+\frac{\pi}{4}\right)\cdot\sin\left(x+\frac{2\pi}{4}\right)\cdot\sin\left(x+\frac{3\pi}{4}\right)$$

for all x. How can the statement be generalized?

11.30. Let (a_n) denote sequence (15) in Example 4.1, that is, let $a_1=0$ and $a_{n+1}=\sqrt{2+a_n}$ $(n\ge 1)$. Prove that $a_n = 2\cdot\cos(\pi/2^n)$.

11.31. Prove that if n is a positive integer, then $\cos nx$ can be written as an nth-degree polynomial in $\cos x$, that is, there exists a polynomial T_n of degree n such that $\cos nx = T_n(\cos x)$ for all x. (H)

11.32. Prove that if n is a positive integer, then $\sin nx/\sin x$ can be written as an nth-degree polynomial of $\cos x$, that is, there exists a polynomial U_n of degree n such that $\sin nx = \sin x\cdot U_n(\cos x)$ for all x.[9]

11.33. Can $\sin nx$ be written as a polynomial of $\sin x$ for all $n\in\mathbb{N}^+$?

11.34. Prove that the equation $x\cdot\sin x = 100$ has infinitely many solutions.

11.35. Let $f:\mathbb{R}\to\mathbb{R}$ be continuous, not identically zero, and suppose that $|f(x)|\le 1$ and $f(x+y)+f(x-y)=2f(x)f(y)$ for all x,y. Prove that for a suitable constant c, $f(x)=\cos cx$ for all x. ($*$H)

---

[9] The polynomials T_n and U_n defined as such are called the **Chebyshev polynomial**s.

11.36. (a) Let $A_n = \sum_{k=1}^{n} \sin^{-2}(k\pi/2n)$. Show that $A_1 = 1$ and $A_{2n} = 4 \cdot A_n - 1$ for all $n = 1, 2, \ldots$.
(b) Prove that $A_{2^n} = (2/3) \cdot 4^n + (1/3) \qquad (n = 0, 1, \ldots)$.
(c) Prove that $(\sin^{-2} x) - 1 < x^{-2} < \sin^{-2} x$ for all $0 < x < \pi/2$, and deduce from this that

$$A_n - n < (2n/\pi)^2 \cdot \sum_{k=1}^{n} 1/k^2 < A_n$$

for all n.
(d) Prove that $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$. (H S)

11.5 The Inverse Trigonometric Functions

Since the function $\cos x$ is strictly monotone on the interval $[0, \pi]$, it has an inverse there, which we call the **arccosine** and denote by $\arccos x$. We are already familiar with this function (Figures 11.8 and 11.9). In Remark 11.21, we noted that the function $\cos x$ on the interval $[0, \pi]$ agrees with the inverse of $S(u) = s(c; [u, 1])$. Thus the function arccos is none other than the function S. The function $\arccos x$ is thus defined on the interval $[-1, 1]$, and there it is strictly monotone decreasing and continuous.

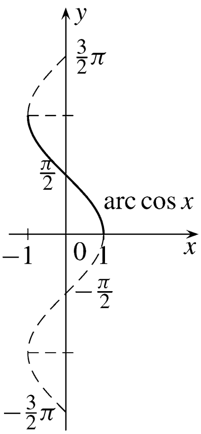

Fig. 11.8

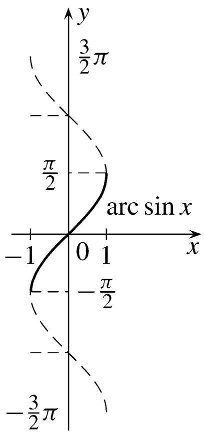

Fig. 11.9

The function $\sin x$ is strictly monotone increasing on the interval $[-\pi/2, \pi/2]$, so it has an inverse there, which we call the **arcsine** function, and we denote it by $\arcsin x$. This function is defined on the interval $[-1, 1]$, and is strictly monotone increasing and continuous there. From the identities (11.32), we see that for all $x \in [-1, 1]$,

$$\arccos x = \frac{\pi}{2} - \arcsin x. \tag{11.50}$$

The function $\operatorname{tg} x$ is strictly monotone increasing on the interval $(-\pi/2, \pi/2)$, so it has an inverse there, which we call the **arctangent**, and we denote it by $\operatorname{arc\,tg} x$. By the limit relation (11.47) and the continuity of $\operatorname{tg} x$, it follows that the function $\operatorname{tg} x$ takes on every real value over $(-\pi/2, \pi/2)$. Thus the function $\operatorname{arc\,tg} x$ is defined on all of the real number line, is continuous, and is strictly monotone increasing (Figure 11.10). It is also clear that

$$\lim_{x\to-\infty} \operatorname{arc\,tg} x = -\frac{\pi}{2} \quad \text{and} \quad \lim_{x\to\infty} \operatorname{arc\,tg} x = \frac{\pi}{2}. \tag{11.51}$$

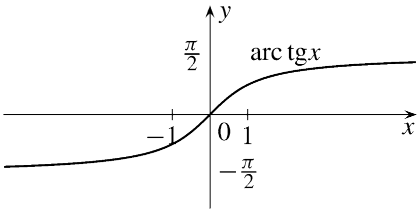

Fig. 11.10

The function $\operatorname{ctg} x$ is strictly monotone decreasing on the interval $(0, \pi)$, so it has an inverse there, which we denote by $\operatorname{arcctg} x$. Since $\operatorname{ctg} x$ takes on every real value over $(0, \pi)$, the function $\operatorname{arcctg} x$ is defined on the whole real number line, is continuous, and is strictly monotone decreasing. By (11.48), it follows that for all x,

$$\operatorname{arcctg} x = \frac{\pi}{2} - \operatorname{arc\,tg} x. \tag{11.52}$$

Exercises

11.37. Draw the graphs of the following functions:

(a) $\arcsin(\sin x)$, (b) $\arccos(\cos x)$, (c) $\operatorname{arc\,tg}(\operatorname{tg} x)$,

(d) $\operatorname{arcctg}(\operatorname{ctg} x)$, (e) $\operatorname{arc\,tg} x - \operatorname{arcctg}(1/x)$.

11.38. Prove the following identities:

(a) $\arcsin x = \arccos\sqrt{1-x^2} \quad (x \in [0,1])$.
(b) $\operatorname{arc\,tg} x = \arcsin\frac{x}{\sqrt{1+x^2}} \quad (x \in \mathbb{R})$.

(c) $\operatorname{arc\,sin} x = \operatorname{arc\,tg} \frac{x}{\sqrt{1-x^2}} \quad (x \in (-1,1))$.
(d) $\operatorname{arc\,tg} x + \operatorname{arc\,tg}(1/x) = \pi/2 \quad (x > 0)$.

11.39. Solve the following equation: $\sin(2 \operatorname{arc\,tg} x) = 1/x$.

11.6 Hyperbolic Functions and Their Inverses

The so-called hyperbolic functions defined below share many properties with their trigonometric analogues. We call the functions

$$\operatorname{sh} x = \frac{e^x - e^{-x}}{2}, \qquad \text{and} \qquad \operatorname{ch} x = \frac{e^x + e^{-x}}{2}$$

hyperbolic sine and **hyperbolic cosine** respectively. By their definition, it is clear that $\operatorname{sh} x$ and $\operatorname{ch} x$ are defined everywhere and are continuous, and moreover, that $\operatorname{ch} x$ is even, while $\operatorname{sh} x$ is odd.

Since the function e^x is strictly monotone increasing, and the function e^{-x} is strictly monotone decreasing, *the function* $\operatorname{sh} x$ *is strictly monotone increasing on* $\mathbb{R}$ (Figure 11.11). By the limit relation (11.6), it is clear that

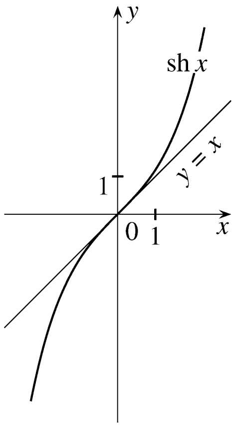

Fig. 11.11

$$\lim_{x \to -\infty} \operatorname{sh} x = -\infty \quad \text{and} \quad \lim_{x \to \infty} \operatorname{sh} x = \infty. \tag{11.53}$$

Since e^x and e^{-x} are strictly convex, it is easy to see that *the function* $\operatorname{ch} x$ *is strictly convex on* $\mathbb{R}$.

The following properties, which follow easily from the definitions, show some of the similarities between the hyperbolic and trigonometric functions. First of all,

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1$$

for all x, that is, the point $(\operatorname{ch} u, \operatorname{sh} u)$ lies on the hyperbola with equation $x^2 - y^2 = 1$ (whence the name "hyperbolic"; how this is analogous to $\cos^2 x + \sin^2 x = 1$ is clear)

Considering that $\operatorname{ch} x$ is positive everywhere,

$$\operatorname{ch} x = \sqrt{1+\operatorname{sh}^2 x} \qquad (11.54)$$

for all x. It immediately follows that *the smallest value of* $\operatorname{ch} x$ *is 1*, and moreover, that $\operatorname{ch} x$ *is strictly monotone increasing on* $[0,\infty)$ *and strictly monotone decreasing on* $(-\infty,0]$ (Figure 11.12).

For all x, y, the addition formulas

$$\begin{aligned}\operatorname{sh}(x+y) &= \operatorname{sh} x \operatorname{ch} y + \operatorname{ch} x \operatorname{sh} y,\\ \operatorname{sh}(x-y) &= \operatorname{sh} x \operatorname{ch} y - \operatorname{ch} x \operatorname{sh} y,\\ \operatorname{ch}(x+y) &= \operatorname{ch} x \operatorname{ch} y + \operatorname{sh} x \operatorname{sh} y,\\ \operatorname{ch}(x-y) &= \operatorname{ch} x \operatorname{ch} y - \operatorname{sh} x \operatorname{sh} y\end{aligned} \qquad (11.55)$$

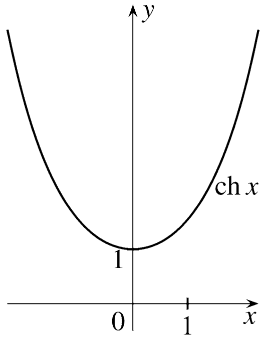

Fig. 11.12

hold, which are simple consequences of the identity $e^{x+y} = e^x \cdot e^y$. The following identities then follow from the addition formulas:

$$\begin{aligned}\operatorname{sh} 2x &= 2\operatorname{sh} x \operatorname{ch} x,\\ \operatorname{ch} 2x &= \operatorname{ch}^2 x + \operatorname{sh}^2 x = 1 + 2\operatorname{sh}^2 x = 2\operatorname{ch}^2 x - 1,\\ \operatorname{ch}^2 x &= \frac{1+\operatorname{ch} 2x}{2},\\ \operatorname{sh}^2 x &= \frac{-1+\operatorname{ch} 2x}{2}.\end{aligned} \qquad (11.56)$$

Remark 11.28. The similarities to trigonometric functions we see above are surprising, and what makes the analogy even more puzzling is how differently the two families of functions were defined. The answer lies in the fact that—contrary to how it seems—exponential functions very much have a connection with trigonometric functions. This connection, however, can be seen only through the complex numbers, and since dealing with complex numbers is not our goal here, we only outline this connection in the second appendix of the chapter (as well as Remarks 13.18 and 11.29).

Much like the functions tg and ctg, we introduce $\operatorname{th} x = \operatorname{sh} x / \operatorname{ch} x$ and $\operatorname{cth} x = \operatorname{ch} x / \operatorname{sh} x$ (Figures 11.13 and 11.14). We leave it to the reader to check that the function $\operatorname{th} x$ (**hyperbolic tangent**) is defined on the real numbers, continuous everywhere, odd, and strictly monotone increasing, and moreover, that

$$\lim_{x\to-\infty} \operatorname{th} x = -1 \quad \text{and} \quad \lim_{x\to\infty} \operatorname{th} x = 1. \qquad (11.57)$$

The function $\operatorname{cth} x$ (**hyperbolic cotangent**) is defined on the set $\mathbb{R}\setminus\{0\}$, and it is continuous there; it is strictly monotone decreasing on the intervals $(-\infty,0)$ and $(0,\infty)$, and moreover,

$$\lim_{x\to\pm\infty} \operatorname{cth} x = \pm 1 \quad \text{and} \quad \lim_{x\to 0\pm 0} \operatorname{cth} x = \pm\infty. \qquad (11.58)$$

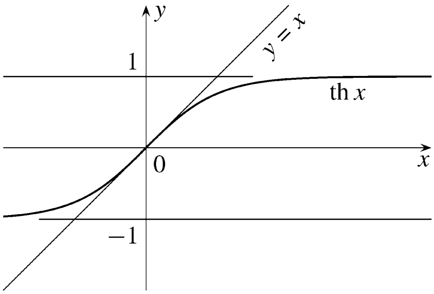

Fig. 11.13

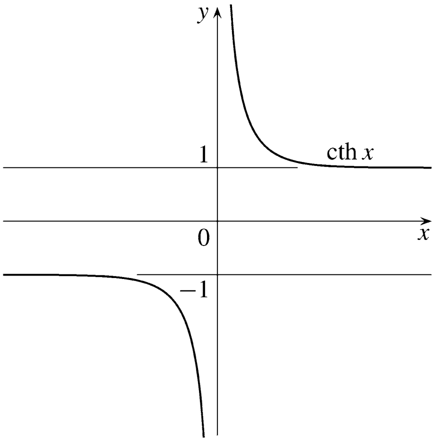

Fig. 11.14

The inverse of the function sh$x$ is called the **area hyperbolic sine** function, denoted by arsh$x$ (Figure 11.15). Since sh$x$ is strictly monotone increasing, continuous, and by (11.53) it takes on every value, arsh$x$ is defined on all real numbers, is continuous everywhere, and is strictly monotone increasing. We can actually express arsh with the help of power and logarithmic functions. Notice that arsh$x = y$ is equivalent to sh$y = x$. If we write the definition of sh$x$ into this and multiply both sides by $2e^y$, we get the equation $e^{2y} - 1 = 2xe^y$. This is a quadratic equation in e^y, from which we get $e^y = x \pm \sqrt{x^2+1}$. Since $e^y > 0$, only the positive sign is considered. Finally, taking the logarithm of both sides, we get that

$$\operatorname{arsh} x = \log\left(x + \sqrt{x^2+1}\right) \tag{11.59}$$

for all x.

The function ch$x$ is strictly monotone increasing and continuous on the interval $[0,\infty)$, and by (11.54), its image there is $[1,\infty)$. The inverse of the function ch$x$ restricted to the interval $[0,\infty)$ is called the **area hyperbolic cosine**, and we denote it by arch$x$ (Figure 11.16). By the above, arch$x$ is defined on the interval $[1,\infty)$, and it is continuous and strictly monotone increasing. It is easy to see that

$$\operatorname{arch} x = \log\left(x + \sqrt{x^2-1}\right) \tag{11.60}$$

for all $x \geq 1$.

The inverse of the function th$x$ is called the **area hyperbolic tangent**, and we denote it by arth$x$ (Figure 11.17). By the properties of th$x$, we see that arth$x$ is defined on the interval $(-1,1)$, is continuous, and is strictly monotone increasing.

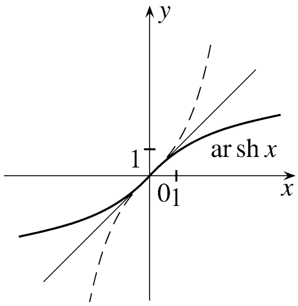

Fig. 11.15

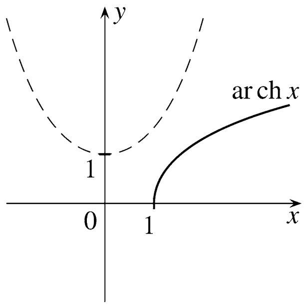

Fig. 11.16

It is easy to see that

$$\operatorname{arth} x = \frac{1}{2} \cdot \log \left(\frac{1+x}{1-x} \right) \tag{11.61}$$

for all $x \in (-1,1)$. We leave the definition of arcth and the verification of its most important properties to the reader.

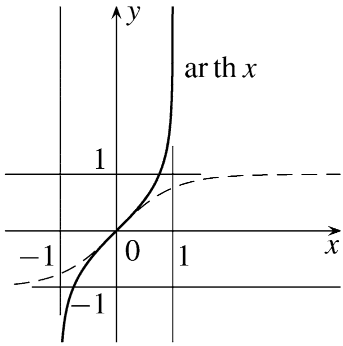

Fig. 11.17

Remark 11.29. The names of trigonometric and hyperbolic functions come from Latin words. The notation $\sin x$ comes from the word *sinus* meaning *bend, fold*.

The notation $\operatorname{tg} x$ comes from the word *tangens*, meaning *tangent*. The name is justified by the fact that if $0 < x < \pi/2$, then $\operatorname{tg} x$ is the length of the segment tangent to the circle at $(1,0)$, starting there and ending where the line going through the origin and the point $(\cos x, \sin x)$ crosses it (see Figure 11.4).

The inverse trigonometric functions all have the *arc* prefix, which implies that $\arccos x$ corresponds to the length of some arc. For inverse hyperbolic functions, instead of *arc*, the word used is *area*. This is justified by the following observation. Let $u \geq 1$ and $v = \sqrt{u^2 - 1}$. The line segments connecting the origin to (u,v) and

$(u,-v)$, and the hyperbola $x^2 - y^2 = 1$ between the points (u,v) and $(u,-v)$ define a section A_u of the plane. One can show that the area of A_u is equal to $\operatorname{arch} u$ (see Exercise 16.9).

Exercises

11.40. Check the addition formulas for $\operatorname{sh} x$ and $\operatorname{ch} x$.

11.41. Find and prove formulas analogous to (11.46) about $\operatorname{th} x$.

11.42. Prove that $\log\left(\sqrt{x^2+1} - x\right) = -\operatorname{arsh} x$ for all x.

11.43. Prove that $\log\left(x - \sqrt{x^2-1}\right) = -\operatorname{arch} x$ for all $x \geq 1$.

11.44. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous, and suppose that $f(x+y) + f(x-y) = 2f(x)f(y)$ for all x, y. Prove that one of the following holds.

(a) f is the function that is identically zero.
(b) There exists a constant c such that $f(x) = \cos cx$ for all x.
(c) There exists a constant c such that $f(x) = \operatorname{ch} cx$ for all x.

11.45. We say that the function f is **algebraic** if there exist polynomials $p_0(x)$, $p_1(x), \ldots, p_n(x)$ such that p_n is not identically zero, and

$$p_0(x) + p_1(x)f(x) + \cdots + p_n(x)f^n(x) = 0$$

for all $x \in D(f)$. We call a function f **transcendental** if it is not algebraic.

(a) Prove that every polynomial and every rational function are algebraic.
(b) Prove that $\sqrt{1+x}$, $\sqrt[3]{\frac{x^2-1}{x+2}}$ and $|x|$ are algebraic functions.
(c) Prove that e^x, $\log x$, $\sin x$, $\cos x$ are transcendental functions.
(d) Can a (nonconstant) periodic function be algebraic?

11.7 First Appendix: Proof of the Addition Formulas

Let O_α denote the positive rotation around the origin by α degrees. Then for arbitrary $\alpha \in \mathbb{R}$ and $x \in \mathbb{R}^2$, $O_\alpha(x)$ is the point that we get by rotating x around the origin by α in the positive direction. We will need the following properties of the rotations O_α.

(i) For arbitrary $\alpha, \beta \in \mathbb{R}$ and $x \in \mathbb{R}^2$, $O_{\alpha+\beta}(x) = O_\alpha\left(O_\beta(x)\right)$.

(ii) The map O_α is linear, that is,

$$O_\alpha(px+qy) = p \cdot O_\alpha(x) + q \cdot O_\alpha(y)$$

for all $x, y \in \mathbb{R}^2$ and $p, q \in \mathbb{R}$.

We will use these properties without proof. (Both properties seem clear; we can convince ourselves of (ii) if we think about what the geometric meaning is of the sum of two vectors and of multiplying a vector by a real number.) We now show that

$$O_\alpha((1,0)) = (\cos\alpha, \sin\alpha). \tag{11.62}$$

Let $O_\alpha((1,0)) = P$, and let h denote the ray starting at the origin and crossing P. Then the angle formed by h and the positive half of the x-axis is α, that is, h intersects the circle C at the point $(\cos\alpha, \sin\alpha)$. Since rotations preserve distances (another geometric fact that we accept), the distance of P from the origin is 1. Thus P is on the circle C, that is, it agrees with the intersection of h and C, which is $(\cos\alpha, \sin\alpha)$.

By the above, $(0,1) = O_{\pi/2}((1,0))$, so by property (i),

$$\begin{aligned} O_\alpha((0,1)) = O_{\alpha+(\pi/2)}((1,0)) = \\ = \left(\cos\left(\alpha+\frac{\pi}{2}\right), \sin\left(\alpha+\frac{\pi}{2}\right)\right) = \\ = (-\sin\alpha, \cos\alpha), \end{aligned} \tag{11.63}$$

where we used (11.28) and (11.32).

Let the coordinates of x be (x_1, x_2). Then by (11.62), (11.63), and (ii), we get that

$$\begin{aligned} O_\alpha(x) = x_1 \cdot O_\alpha((1,0)) + x_2 \cdot O_\alpha((0,1)) = \\ = x_1 \cdot (\cos\alpha, \sin\alpha) + x_2 \cdot (-\sin\alpha, \cos\alpha) = \\ = (x_1 \cdot \cos\alpha - x_2 \cdot \sin\alpha,\ x_1 \cdot \sin\alpha + x_2 \cdot \cos\alpha). \end{aligned}$$

By all these,

$$\begin{aligned} (\cos(\alpha+\beta), \sin(\alpha+\beta)) = O_{\alpha+\beta}((1,0)) = O_\alpha\left(O_\beta((1,0))\right) = \\ = O_\alpha\left((\cos\beta, \sin\beta)\right) = \\ = (\cos\beta \cdot \cos\alpha - \sin\beta \cdot \sin\alpha,\ \cos\beta \cdot \sin\alpha + \sin\beta \cdot \cos\alpha), \end{aligned}$$

so after comparing coordinates, we get the first and third identities of (11.33). If we apply these to $-y$ instead of y, then also using that $\cos x$ is even and $\sin x$ is odd, we get the other two identities.

11.8 Second Appendix: A Few Words on Complex Numbers

The introduction of complex numbers is motivated by the need to include a number whose square is -1 into our investigations. We denote this number by i.[10] We call the formal expressions $a+bi$, where a and b are real numbers, **complex numbers**. We consider the real numbers complex numbers as well by identifying the real number a with the complex number $a+0\cdot i$. Addition and multiplication are defined on the complex numbers in a way to preserve the usual rules, and so that $i^2=-1$ also holds. Thus the sum and product of the complex numbers $z_1=a_1+b_1i$ and $z_2=a_2+b_2i$ are defined by the formulas

$$z_1+z_2=(a_1+a_2)+(b_1+b_2)i$$

and

$$z_1\cdot z_2=(a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)i.$$

One can show that the complex numbers form a field with these operations if we take the zero element to be $0=0+0\cdot i$ and the identity element to be $1=1+0\cdot i$. Thus *the complex numbers form a field that contains the real numbers.*

A remarkable fact is that among the complex numbers, every nonconstant polynomial (with complex coefficients) has a root (so for example, the polynomial x^2+1 has the roots i and $-i$). One can prove that *a degree-n polynomial has n roots counting multiplicity.* This statement is called the **fundamental theorem of algebra.**

To define complex powers, we will use the help of (11.11). Since $(1+z/n)^n$ is defined for all complex numbers z, it is reasonable to define e^z as the limit of the sequence $(1+z/n)^n$ as $n\to\infty$. (We say that the sequence of complex numbers $z_n=a_n+b_ni$ tends to $z=a+bi$ if $a_n\to a$ and $b_n\to b$.) One can show that this is well defined, that is, the limit exists for all complex numbers z, and that the powers defined in this way satisfy $e^{z+w}=e^z\cdot e^w$ for all complex z and w. Moreover—and this is important to us—if $z=ix$, where x is real, then the limit of $(1+z/n)^n$ is $\cos x+i\cdot\sin x$, so

$$e^{ix}=\cos x+i\cdot\sin x \tag{11.64}$$

for all real numbers x.[11] If we apply (11.64) to $-x$ instead of x, then we get that

$$e^{-ix}=\cos x-i\cdot\sin x.$$

We can express both $\cos x$ and $\sin x$ from this expression, and we get the identities

$$\cos x=\frac{e^{ix}+e^{-ix}}{2},\qquad \sin x=\frac{e^{ix}-e^{-ix}}{2i}. \tag{11.65}$$

[10] As the first letter of the word "imaginary."

[11] It is worth checking that $e^{i(x+y)}=e^{ix}\cdot e^{iy}$ holds for all $x,y\in\mathbb{R}$. By (11.64), this is equivalent to the addition formulas for the functions cos and sin.

These are called **Euler's formulas.** These two identities help make the link between trigonometric and hyperbolic functions clear. If we extend the definitions of ch and sh to the complex numbers, we get that

$$\cos x = \operatorname{ch}(ix) \ \text{ and } \ \sin x = \operatorname{sh}(ix)/i \tag{11.66}$$

for all real x.