valent to the set of m relations $a^{(1)}=\lim_n a_n^{(1)}$, $a^{(2)}=\lim_n a_n^{(2)},\ldots$, $a^{(m)}=\lim_n a_n^{(m)}$, *in the spaces* $\boldsymbol{A}^{(1)},\boldsymbol{A}^{(2)},\ldots,\boldsymbol{A}^{(m)}$, *respectively;*

(b) *if* $A^{(1)},A^{(2)},\ldots,A^{(m)}$ *are closed sets in the corresponding spaces* $\boldsymbol{A}^{(1)},\boldsymbol{A}^{(2)},\ldots,\boldsymbol{A}^{(m)}$, *then the set* $A=A^{(1)}\times A^{(2)}\times\ldots\times A^{(m)}$ *is a closed set in the space* $\boldsymbol{A}$;

(c) *if* $B^{(1)},B^{(2)},\ldots,B^{(m)}$ *are compact sets in the corresponding spaces* $\boldsymbol{A}^{(1)},\boldsymbol{A}^{(2)},\ldots,\boldsymbol{A}^{(m)}$, *then the set* $B=B^{(1)}\times B^{(2)}\times\ldots\times B^{(m)}$ *is a compact set in the space* $\boldsymbol{A}$.

Proof. Part (a) of the theorem is obvious, and part (b) follows immediately from part (a). To prove part (c) let us assume for simplicity that $m=2$ and let $\{b_n=(b_n^{(1)}, b_n^{(2)})\}$ be an arbitrary sequence of points of the set B. From the sequence $\{b_n^{(1)}\}_{n=1,2,\ldots}$, which consists of points of the set $B^{(1)}$, we can extract a subsequence $\{b_{n_k}^{(1)}\}_{k=1,2,\ldots}$ convergent in the space $\boldsymbol{A}^{(1)}$. Next, from the sequence $\{b_{n_k}^{(2)}\}_{k=1,2,\ldots}$, whose points belong to the set $B^{(2)}$, we can extract a subsequence $\{b_{n_{k_j}}^{(2)}\}_{j=1,2,\ldots}$ convergent in the space $\boldsymbol{A}^{(2)}$. The sequence $\{(b_{n_{k_j}}^{(1)}, b_{n_{k_j}}^{(2)})\}_{j=1,2,\ldots}$, extracted from the given sequence $\{(b_n^{(1)}, b_n^{(2)})\}_{n=1,2,\ldots}$, is consequently (by (a)) convergent in the space $\boldsymbol{A}$. The set $B\subset\boldsymbol{A}$ is therefore compact.

We can regard the Cartesian m-th power $\boldsymbol{E}^m$ of the plane $\boldsymbol{E}$ as an example of a Cartesian product of spaces. The system of neighbourhoods for the space $\boldsymbol{E}^m$ is formed by Cartesian products of the type

$$\mathrm{K}(z^{(1)};r_1)\times\mathrm{K}(z^{(2)};r_2)\times\ldots\times\mathrm{K}(z^{(m)};r_m),$$

where $r_1,r_2,\ldots,r_m$ are arbitrary positive real numbers and $z^{(1)}, z^{(2)},\ldots,z^{(m)}$ are points of the plane. In the case $m=2$ we shall also call the neighbourhood $\mathrm{K}(z_1;r_1)\times\mathrm{K}(z_2;r_2)$ a *bicircular neighbourhood with centre* (z_1,z_2) or a *bicircular neighbourhood of the point* (z_1,z_2).

By the *distance* $\varrho(z_1,z_2)$ of two points $z_1=(z_1^{(1)},z_1^{(2)},\ldots,z_1^{(m)})$ and $z_2=(z_2^{(1)},z_2^{(2)},\ldots,z_2^{(m)})$ in the space $\boldsymbol{E}^m$ we shall mean the largest of the numbers $\varrho(z_1^{(k)},z_2^{(k)})$ for $k=1,2,\ldots,m$. From theorem 13.1 (a) it follows immediately that the *relation* $\lim_k z_k=z$ *in the space* $\boldsymbol{E}^m$ *is equivalent to the relation* $\varrho(z_k,z)\to 0$.

CHAPTER I

FUNCTIONS OF A COMPLEX VARIABLE

§ 1. Continuous functions. In this section we shall establish fundamental definitions and notations concerning functions of one and of several variables. The independent variables as well as the functions will assume complex values; the value ∞ will also be admissible provided, of course, that the context does not necessitate its exclusion.

A function $F(z_1,z_2,\ldots,z_n)$ of n complex variables, ranging respectively over n sets $Z_1,Z_2,\ldots,Z_n$ in the plane $\boldsymbol{E}$, can be considered as a function of the point $z=(z_1,z_2,\ldots,z_n)$, defined on the Cartesian product (cf. Introduction, § 13) $Z=Z_1\times Z_2\times\ldots\times Z_n$. Instead of $F(z_1,z_2,\ldots,z_n)$, where $z_1\epsilon Z_1$, $z_2\epsilon Z_2,\ldots$, $z_n\epsilon Z_n$, we can also write $F(z)$, where $z\epsilon Z_1\times Z_2\times\ldots\times Z_n$. The function F is said to be:

1° *bounded* on the set Z, if there exists a finite number M such that $|F(z)|\leqslant M$ for each point $z\epsilon Z$;

2° *uniformly continuous* on Z, if for each number $\varepsilon>0$ there exists a number $\eta>0$ such that the inequality $\varrho(z_1,z_2)<\eta$ implies $|F(z_2)-F(z_1)|<\varepsilon$ for every pair of points z_1,z_2 of the set Z (this definition presupposes that F is finite-valued).

We denote by $\varrho(z_1,z_2)$ the distance between the points z_1 and z_2 in agreement with the definitions in the Introduction, §§ 8, 13.

(1.1) *If the sets* $Z_1,Z_2,\ldots,Z_n$ *are closed, then every finite and continuous function* F *on the set* $Z=Z_1\times Z_2\times\ldots\times Z_n$ *is bounded and uniformly continuous on this set. Moreover, if the function* F *is real, then at a certain point of the set* Z *it attains the upper bound of its values on this set.*

Proof. Let us assume that the function F is not bounded on Z. Then there exists a sequence of points $\{z^{(k)}\}_{k=1,2,\ldots}$ in the set Z such that $F(z^{(k)})\to\infty$. Let $\{z^{(k_j)}\}$ be a convergent subsequence extracted from the sequence $\{z^{(k)}\}$. Such a subsequence exists in virtue of the compactness of the space $\boldsymbol{E}^n$ (Introduction, theorems

8.2 and 13.1(c)). Denoting its limit by z_0 we have $z_0 \epsilon Z$ in virtue of the fact that the sets $Z_1, Z_2, \ldots, Z_n$ are closed. Consequently, $F(z_0) = \lim_j F(z^{(k_j)}) = \infty$, which is contrary to the hypothesis that the function $F(z)$ is finite on Z.

Next, let us assume that the function F is not uniformly continuous on Z. Then we can determine a positive number ε and two sequences of points $\{p^{(k)}\}$ and $\{q^{(k)}\}$ in the set Z such that

$$\lim_k \varrho(p^{(k)}, q^{(k)}) = 0 \tag{1.2}$$

and

$$|F(p^{(k)}) - F(q^{(k)})| \geqslant \varepsilon \quad \text{for} \quad k = 1, 2, \ldots \tag{1.3}$$

Let $\{p^{(k_j)}\}$ be a convergent subsequence extracted from $\{p^{(k)}\}$ and let $p_0 = \lim_j p^{(k_j)}$. By (1.2) we easily verify that also $p_0 = \lim_j q^{(k_j)}$. Hence, substituting $k = k_j$ in (1.3) and passing to the limit as $j \to \infty$, we obtain, in virtue of the continuity and finiteness of the function $F(z)$ at the point p_0, the inequality $0 \geqslant \varepsilon$, obviously false.

Finally, if we assume that the function F is real and that M is the upper bound of its values on Z, then there exists a sequence of points $\{z_k\}$ of the set Z such that $F(z_k) \to M$ when $k \to \infty$. Denoting by $\{z_{k_i}\}$ an arbitrary convergent subsequence of this sequence, and by z^0 the limit of this subsequence, we shall have $z^0 \epsilon Z$ and $F(z^0) = \lim_i F(z_{k_i}) = M$.

In particular, from theorem 1.1 it follows that *every finite and continuous function on the entire plane is uniformly continuous on it* (this theorem is obviously not true for the open plane $\boldsymbol{E}_0$).

We shall now prove two theorems "on the continuity" and "on the differentiability of a definite integral with respect to a parameter". In these theorems the processes of integration and differentiation are understood to be with respect to a real variable, so that in fact we do not go beyond the real domain, although the function itself may assume complex values. For, if $F(t)$ is a complex function of the real variable t, then representing it in the form $F(t) = U(t) + iV(t)$, where $U(t)$, $V(t)$ are real functions, we can define the derivative and the integral of $F(t)$ by the formulae

$$F'(t) = U'(t) + iV'(t), \qquad \int_a^b F(t)\,dt = \int_a^b U(t)\,dt + i\int_a^b V(t)\,dt;$$

these definitions are obviously equivalent to the direct definitions: of the integral as a limit of approximating sums, and of the derivative $F'(t)$ as the finite limit of the quotient $[F(t+h) - F(t)]/h$ as h tends to 0 through real values. New facts appear only when we differentiate with respect to a complex variable (cf. § 6 further on, and Chapter II, § 1; theorem 1.5, which we give below, will be generalized in Chapter II, § 3, to differentiation with respect to a complex parameter).

(1.4) *If $T = [a, b]$ is a finite interval, $Z_1, Z_2, \ldots, Z_n$ are closed sets, and $F(t, z_1, z_2, \ldots, z_n)$ a finite and continuous function on the set $T \times Z_1 \times Z_2 \times \ldots \times Z_n$, then the integral*

$$\Phi(z_1, z_2, \ldots, z_n) = \int_a^b F(t, z_1, z_2, \ldots, z_n)\,dt$$

is a continuous function on the set $Z = Z_1 \times Z_2 \times \ldots \times Z_n$.

Proof. By theorem 1.1 to each number $\varepsilon > 0$ there corresponds a number $\eta > 0$ such that with $t \epsilon T$, $z_1' \epsilon Z_1$, $z_1'' \epsilon Z_1$, $z_2' \epsilon Z_2$, $z_2'' \epsilon Z_2, \ldots$, $z_n' \epsilon Z_n$, $z_n'' \epsilon Z_n$, the set of inequalities $\varrho(z_1', z_1'') < \eta$, $\varrho(z_2', z_2'') < \eta, \ldots$, $\varrho(z_n', z_n'') < \eta$ implies the inequality

$$|F(t, z_1'', z_2'', \ldots, z_n'') - F(t, z_1', z_2', \ldots, z_n')| < \frac{\varepsilon}{b-a},$$

and therefore

$$|\Phi(z_1'', z_2'', \ldots, z_n'') - \Phi(z_1', z_2', \ldots, z_n')| \leqslant \frac{(b-a)\varepsilon}{b-a} = \varepsilon,$$

which gives the continuity of the function Φ on Z.

(1.5) *Let $T = [a, b]$ and $U = [c, d]$ be linear intervals; $Z_1, Z_2, \ldots, Z_n$ closed sets in the plane; and $F(t, u, z_1, z_2, \ldots, z_n)$ a finite and continuous function on the set $T \times U \times Z_1 \times Z_2 \times \ldots \times Z_n$. Then, if the function F has a derivative $F_u'(t, u, z_1, z_2, \ldots, z_n)$ with respect to u, continuous on the set $T \times U \times Z_1 \times Z_2 \times \ldots \times Z_n$, the function*

$$\Phi(u, z_1, z_2, \ldots, z_n) = \int_a^b F(t, u, z_1, z_2, \ldots, z_n)\,dt$$

also has a derivative with respect to u, continuous on the set

$$U \times Z_1 \times Z_2 \times \ldots \times Z_n$$

and given by the formula

$$(1.6)\qquad \Phi'_u(u,z_1,z_2,\ldots,z_n)=\int_a^b F_u(t,u,z_1,z_2,\ldots,z_n)\,dt.$$

Proof. By considering the real and the imaginary part of the function F separately, we may assume that the function is real. Let $(u,z_1,z_2,\ldots,z_n)$ be an arbitrary point of the set $U\times Z_1\times Z_2\times\ldots\times Z_n$, and ε an arbitrary positive number. In virtue of the uniform continuity of the derivative F'_u on the set $T\times U\times Z_1\times Z_2\times\ldots\times Z_n$ (cf. theorem 1.1), there exists a number $\eta>0$ such that for each $t\epsilon T$ and $v\epsilon U$ the inequality $|v-u|<\eta$ implies the inequality

$$|F'_u(t,v,z_1,z_2,\ldots,z_n)-F'_u(t,u,z_1,z_2,\ldots,z_n)|<\frac{\varepsilon}{b-a}.$$

Let Δu be an arbitrary increment of the variable u, such that $u+\Delta u\,\epsilon U$ and $|\Delta u|<\eta$. By the mean-value theorem we have

$$F(t,u+\Delta u,z_1,z_2,\ldots,z_n)-F(t,u,z_1,z_2,\ldots,z_n)$$
$$=\Delta u\cdot F'_u(t,u+\theta\Delta u,z_1,z_2,\ldots,z_n),$$

where $0<\theta<1$. Whence, because $|\theta\Delta u|\leqslant|\Delta u|<\eta$, we obtain

$$\left|\frac{F(t,u+\Delta u,z_1,z_2,\ldots,z_n)-F(t,u,z_1,z_2,\ldots,z_n)}{\Delta u}-F'_u(t,u,z_1,z_2,\ldots,z_n)\right|$$
$$<\frac{\varepsilon}{b-a},$$

and therefore,

$$\left|\frac{\Phi(u+\Delta u,z_1,z_2,\ldots,z_n)-\Phi(u,z_1,z_2,\ldots,z_n)}{\Delta u}-\int_a^b F'_u(t,u,z_1,z_2,\ldots,z_n)\,dt\right|$$
$$=\left|\int_a^b\left[\frac{F(t,u+\Delta u,z_1,z_2,\ldots,z_n)-F(t,u,z_1,z_2,\ldots,z_n)}{\Delta u}\right.\right.$$
$$\left.\left.-F'_u(t,u,z_1,z_2,\ldots,z_n)\right]dt\right|<\varepsilon,$$

from which follows simultaneously the existence of the derivative Φ'_u and formula (1.6). And from this formula, in virtue of theorem 1.4, follows immediately the continuity of the derivative Φ'_u on the set $U\times Z_1\times Z_2\times\ldots\times Z_n$.

§ 2. Uniformly and almost uniformly convergent sequences.

In the remainder of this chapter we shall be concerned primarily with complex functions of one complex variable; however, the majority of the definitions and theorems generalize immediately to functions of several variables.

Every function F of a complex variable can obviously be considered as a function of two real variables; denoting the real and imaginary parts of the complex number z by x and y, respectively, we shall frequently write $F(x,y)$ instead of $F(z)$.

We say that a sequence $\{F_n(z)\}$ of finite functions on a certain set Z is *uniformly convergent* on this set to a certain function $F(z)$, if to each number $\varepsilon>0$ there corresponds a number N such that $|F_n(z)-F(z)|\leqslant\varepsilon$ for every $n\geqslant N$ and for every point $z\epsilon Z$. Clearly, the limit function $F(z)$ will then also be finite. Similarly, a sequence $\{F_n(z)\}$ of functions (finite or not) on the set Z will be said to be *uniformly divergent to* ∞ *on* Z, if to each number M there corresponds a number N such that $|F_n(z)|\geqslant M$ for every $n\geqslant N$ and for every point $z\epsilon Z$.

In many future considerations the following generalization of the above definitions will prove to be convenient. We shall call a sequence $\{F_n(z)\}$ of functions on an open set G *almost uniformly convergent (almost uniformly divergent to* ∞*)* on this set, if it is uniformly convergent (uniformly divergent to ∞) on every closed set contained in G. If $F(z)$ is the limit of an almost uniformly convergent sequence $\{F_n(z)\}$, then we write $F_n(z)\rightrightarrows F(z)$. Similarly, $F_n(z)\rightrightarrows\infty$ will denote the fact that the sequence $\{F_n(z)\}$ is almost uniformly divergent to ∞.

These definitions extend immediately to series of functions. A series of functions is therefore *uniformly convergent, uniformly divergent to* ∞, etc., if the sequence of its partial sums is uniformly convergent, is uniformly divergent to ∞, etc. If the series $\sum_{n=0}^{\infty}|F_n(z)|$ is convergent (uniformly convergent, almost uniformly convergent), then we say that the series $\sum_{n=0}^{\infty}F_n(z)$ is *absolutely convergent (absolutely uniformly convergent, absolutely almost uniformly convergent).* Obviously, absolute convergence of a series implies ordinary convergence; the same applies to absolute uniform convergence and absolute almost uniform convergence.

(2.1) *If a sequence* $\{F_n(z)\}$ *of finite and continuous functions on the set* Z *is uniformly convergent on this set to the function* $F(z)$, *then the function* $F(z)$ *is also finite and continuous on* Z.

In the case when Z is an open set, instead of assuming that the sequence $\{F_n(z)\}$ is uniformly convergent, it is sufficient to assume that it is almost uniformly convergent.

Proof. Let $z_0 \epsilon Z$ and $\varepsilon > 0$. Let N be a positive integer such that $|F_N(z) - F(z)| < \varepsilon/3$ for each point $z \epsilon Z$. Finally, let η be a positive number such that for each point $z \epsilon Z$ the condition

$$\varrho(z_0, z) < \eta \tag{2.2}$$

implies that $|F_N(z) - F_N(z_0)| < \varepsilon/3$. Condition (2.2) implies then that for each point $z \epsilon Z$,

$$|F(z) - F(z_0)| \leqslant |F(z) - F_N(z)| + |F_N(z) - F_N(z_0)| + |F_N(z_0) - F(z_0)| < \varepsilon,$$

which gives the continuity of the function F on the set Z at the point z_0. The second part of the theorem is an obvious consequence of the first.

We shall introduce the following two additional definitions, similar to the definitions given at the beginning of this section. A family of functions will be said to be *bounded* on a set Z (and the functions belonging to this family — *uniformly bounded* on Z), if there exists a finite number M such that $|F(z)| \leqslant M$ for every function F of the family considered and every point $z \epsilon Z$. If a family of functions is bounded on every closed set contained in an open set G, then this family is said to be *almost bounded on G*, and the functions belonging to it — *almost uniformly bounded on G*. By a series *bounded* on the set Z (*almost bounded* on the open set G) we shall mean a series whose partial sums form a sequence uniformly bounded on Z (almost uniformly bounded on G).

(2.3) *If a family $\mathfrak{F}$ of functions is bounded in the neighbourhood of every point of a certain open set G (i. e. if every point of this set has a neighbourhood in which the family $\mathfrak{F}$ is bounded), then this family is almost bounded on the entire set G.*

In fact, let P be an arbitrary closed set contained in G. With each point $z \epsilon P$ we can associate an open circle which contains it and on which the family $\mathfrak{F}$ is bounded. By the theorem of Borel-Lebesgue (Introduction, § 6) the set P can be covered by a finite number of such circles. The family $\mathfrak{F}$ is therefore bounded on the set P.

Similarly we prove that *every sequence of functions uniformly convergent (divergent to ∞) in the neighbourhood of each point of an open set G is almost uniformly convergent (divergent to ∞) on this entire set.*

EXERCISES. 1. Prove that for $|z| < 1$ we have

$$\sum_{n=1}^{\infty} \frac{z^n}{1-z^n} = \sum_{k=1}^{\infty} \tau(k) z^k, \tag{*}$$

where $\tau(k)$ denotes the number of divisors of the number k, and both series in the equation (*) are almost uniformly convergent in the circle $K(0;1)$.

2. Prove that the series $\sum_{n=0}^{\infty} \frac{z^{2^n}}{1-z^{2^{n+1}}}$ is almost uniformly convergent on the set obtained by removing the circumference $C(0;1)$ from the open plane. It converges to $z/(1-z)$, when $|z| < 1$, and to $1/(1-z)$, when $|z| > 1$.

[*Hint.* Note that every positive integer can be represented in one and only one way in the form $2^{h-1}(2k-1) = k \cdot 2^h - 2^{h-1}$, where k and h are positive integers.]

3. Let $\boldsymbol{S}$ be the family of all continuous functions on an open set G in the (closed) plane. We can consider the family $\boldsymbol{S}$ as a metric space (cf. Introduction, § 3, exercise 3), and define a metric for it in the following way:

We represent G as a sum of an increasing sequence of closed sets $\{F_n\}$ by defining the set F_n as the set of all those points of the plane whose distance from the complement of the set G is $\geqslant 1/n$ (cf. Introduction, § 11). Denoting, for every pair of continuous functions $U(z)$, $V(z)$ on G, by $M_n(U,V)$ the upper bound of $|U(z) - V(z)|$ for $z \epsilon F_n$, we define the "distance" between these functions by the formula

$$\varrho(U, V) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot \frac{M_n(U,V)}{1 + M_n(U,V)}.$$

Show that 1° the distance so defined satisfies the conditions of exercise 3, Introduction, § 3; 2° the space $\boldsymbol{S}$ is complete (see Introduction, § 4, exercise 7) and the relation $\lim_{n,m\to\infty} \varrho(W_n, W_m) = 0$ is equivalent to the almost uniform convergence of the sequence $\{W_n(z)\}$ in the set G; 3° the space $\boldsymbol{S}$ is not compact (cf. Introduction, § 6) in any neighbourhood of the point 0, *i. e.* for each number $\varepsilon > 0$ there exists a sequence of functions $\{W_n(z)\}$ such that:

(a) $\varrho(W_n, 0) \leqslant \varepsilon$ for $n = 1, 2, \ldots$,

(b) no subsequence of the sequence $\{W_n\}$ is convergent in the space $\boldsymbol{S}$.

§ 3. Normal families of functions. Elementary considerations of analysis frequently give one an opportunity of applying a method based on the theorem of Bolzano-Weierstrass and depending on the extraction of convergent subsequences from arbitrary sequences of points. Although the theorem of Bolzano-Weierstrass does not generalize directly to arbitrary sequences of functions, nevertheless, the method of "extracting convergent subsequences" can be applied with success in many proofs concerning special sequences of functions.

We find its applications in various branches of analysis. As an example, we point to the proof of the existence of a solution of the differential equation $y'=u(x,y)$ when the only assumption we make concerning the function $u(x,y)$ is that it is continuous; one could also give many other examples from the calculus of variations, the theory of integral equations, etc. (see, *e. g.* E. Kamke, *Differentialgleichungen reeller Funktionen,* Leipzig 1930, pp. 60-66, 126-130; C. Carathéodory, *Variationsrechnung und partielle Differentialgleichungen erster Ordnung,* München 1935, pp. 1-8).

These numerous applications are associated with the names of Ascoli, Arzelà, Hilbert, Carathéodory and others. However, it was Montel who, distinguishing the class of the so-called normal families of functions, first worked out systematically the method we are considering now, pointing out its particular significance in the theory of functions of a complex variable (see, *e. g.* P. Montel, *Leçons sur les familles normales de fonctions analytiques et leurs applications,* Paris 1927).

A family of functions (real or complex) defined on an open set G is said to be *normal* in G if every sequence of functions belonging to this family contains either a subsequence almost uniformly convergent in G, or a subsequence almost uniformly divergent in G to ∞. The theory of normal families rests on the following simple lemma:

(3.1) *Let $\{E_k\}$ be a sequence of sets, and $\mathfrak{F}$ a family of functions defined on the sum of these sets. If for each set E_k there can be extracted from each sequence of functions belonging to $\mathfrak{F}$ either a subsequence uniformly convergent on the set E_k, or a subsequence uniformly divergent on this set to ∞, then from each sequence of functions of the family $\mathfrak{F}$ there can be extracted a subsequence which on each of the sets E_k is either uniformly convergent or uniformly divergent to ∞.*

Proof. Let $\mathfrak{C}=\{F_n(z)\}_{n=1,2,...}$ be an arbitrary sequence of functions of the family $\mathfrak{F}$. We shall define by induction for each $k=0,1,2,\ldots$, a certain subsequence $\mathfrak{C}^{(k)}=\{F_n^{(k)}(z)\}_{n=1,2,...}$ of the sequence $\mathfrak{C}$, taking $\mathfrak{C}^{(0)}=\mathfrak{C}$, *i. e.* $F_n^{(0)}(z)=F_n(z)$ for $n=1,2,\ldots$ Assuming that the sequences $\mathfrak{C}^{(0)},\mathfrak{C}^{(1)},\ldots,\mathfrak{C}^{(k-1)}$ are defined, we shall take as the sequence $\mathfrak{C}^{(k)}=\{F_n^{(k)}(z)\}_{n=1,2,...}$ an arbitrary subsequence extracted from $\mathfrak{C}^{(k-1)}$, which is either uniformly convergent on the set E_k, or uniformly divergent to ∞; we can assume, in addition, that the first $k-1$ terms of the sequence $\mathfrak{C}^{(k)}$ are identical with the corresponding terms of the sequence $\mathfrak{C}^{(k-1)}$, *i. e.* that

$$F_1^{(k)}(z)=F_1^{(k-1)}(z),\quad \ldots,\quad F_{k-1}^{(k)}(z)=F_{k-1}^{(k-1)}(z).$$

Because of this assumption the sequence $\{F_j^{(j)}(z)\}_{j=1,2,...}$ is a subsequence of all the sequences $\mathfrak{C}^{(1)},\mathfrak{C}^{(2)},\ldots,\mathfrak{C}^{(j)},\ldots$, and hence it is uniformly convergent, or uniformly divergent to ∞, on each of the sets E_j.

Let us assume that the sets E_k considered in lemma 3.1 reduce to single points. By the theorem of Bolzano-Weierstrass every family of functions satisfies the hypotheses of the lemma with respect to so specialized a sequence of sets $\{E_k\}$. As a particular case of lemma 3.1 we therefore obtain the following theorem:

(3.2) *If E is a finite or denumerable set, then every sequence $\{F_n(z)\}$ of functions defined on E contains a subsequence which at each point of the set E is either convergent or divergent to ∞.*

Appealing once more to lemma 3.1, we shall prove the following general theorem concerning normal families:

(3.3) *If a family $\mathfrak{F}$ of functions is normal in the neighbourhood of each point of a certain region G, i. e. if each point of this region has a neighbourhood in which the family considered is normal, then the family $\mathfrak{F}$ is normal in the entire region G.*

Proof. To each point z_0 of the region G we can assign a neighbourhood K in which the family $\mathfrak{F}$ is normal. Let us denote by H an arbitrary circle having a rational centre and radius, such that $z_0\epsilon H\subset\overline{H}\subset K$. From each sequence of the family $\mathfrak{F}$ we can therefore extract a subsequence which is either uniformly convergent, or uniformly divergent to ∞ in H. Since the set of all circles with rational centres and radii is denumerable, by associating in this manner a circle H with each point $z_0\epsilon G$ we obtain a sequence of circles $\{H_j\}$ jointly covering G, where for each j every sequence of functions of the family $\mathfrak{F}$ contains a subsequence which in the circle H_j is either uniformly convergent, or uniformly divergent to ∞.

Let $\{F_n(z)\}$ be an arbitrary sequence of functions of the family under consideration. In view of lemma 3.1 this sequence contains a subsequence $\{F_{n_k}(z)\}$ which in each circle H_j is uniformly convergent or uniformly divergent to ∞. We shall show that the sequence $\{F_{n_k}(z)\}$ is either divergent, or convergent, on all the circles simultaneously. In fact, denoting in the contrary case by G_1 the sum of those circles on which the sequence considered is convergent, and by G_2 the sum of those circles on which it is divergent, we should obtain a decomposition of the region G into two non-empty, open and disjoint sets, which is obviously impossible.

Let us assume that the sequence $\{F_{n_k}(z)\}$ is everywhere convergent; we shall show that it is almost uniformly convergent in G. In fact, let $P\subset G$ be an arbitrary closed set. By the theorem of Borel (Introduction, theorem 6.2) there exists then a finite number of circles $H_1, H_2, \ldots, H_m$ covering P jointly. Since the sequence $\{F_{n_k}(z)\}$ is uniformly convergent in each of them, it is uniformly convergent on P.

Similarly, assuming that the sequence $\{F_{n_k}(z)\}$ is everywhere divergent, we prove that it is almost uniformly divergent to ∞ in G.

(3.4) *Every sequence of functions which is normal and convergent everywhere to a finite limit in an open set G, is almost uniformly convergent in G; every sequence of functions which is normal in the open set G and divergent to ∞ at one point of this set at least, is almost uniformly divergent to ∞ in G.*

Proof. Let us assume that the sequence $\{F_n(z)\}$, which is normal and convergent in G to a certain finite function $F(z)$, is not almost uniformly convergent in G, and therefore not uniformly convergent on a certain closed set $P\subset G$. Then there exists a certain number $\varepsilon>0$, an increasing sequence of indices $\{n_k\}$, and a sequence of points $\{z_k\}$ of the set P, such that $|F_{n_k}(z_k)-F(z_k)|\geqslant\varepsilon$ for $k=1,2,\ldots$ In view of the convergence and normality of the sequence $\{F_n(z)\}$ we can extract from the sequence $\{F_{n_k}(z)\}$ a subsequence $\{F_{n_{k_i}}(z)\}$ almost uniformly convergent in G — and hence uniformly convergent on the set P — to $F(z)$. On the other hand, however, we have $|F_{n_{k_i}}(z_{k_i})-F(z_{k_i})|\geqslant\varepsilon$ for $i=1,2,\ldots$, which constitutes an obvious contradiction.

The proof of the second part of the theorem proceeds in a similar manner.

(3.5) *If a sequence of functions $\{F_n(z)\}$ finite in the open set G is normal in the set G and bounded at a certain point of the set G, then this sequence is almost bounded in the entire set G.*

Proof. Let us assume that the sequence $\{F_n(z)\}$ is not almost bounded in G, and hence not bounded on a certain closed set $P\subset G$. Then there exists a sequence of points $\{z_k\}$ of the set P and an increasing sequence of indices n_k such that

(3.6) $$|F_{n_k}(z_k)|>k \quad \text{for} \quad k=1,2,\ldots$$

Next, the sequence $\{F_{n_k}(z)\}$ contains a subsequence $\{F_{n_{k_j}}(z)\}$, which is either almost uniformly convergent in the open set G, or almost uniformly divergent to ∞. In view of (3.6), however, the first alternative drops out and the sequence $\{F_{n_{k_j}}(z)\}$ is almost uniformly divergent to ∞. This however contradicts the assumption that the sequence $\{F_n(z)\}$ is bounded at a certain point of the set G.

EXERCISES. 1. Examples of normal families of functions: (a) A sequence of functions $\{F_n(x,y)\}$ defined in the open plane by means of the equations: $F_n(x,y)=0$ for $x\leqslant 0$ and $F_n(x,y)=n$ for $x>0$, is normal in the region $x>0$ as well as in the region $x<0$; however, it is not normal in the sum of these regions (this example shows that in theorem 3.3 the condition that the set G is a region is essential and cannot be replaced by the condition that the set G is only open). (b) The family of all functions $F(z)=az$, where a is an arbitrary coefficient, is normal in the open plane with the point 0 removed, but it is not normal in any neighbourhood of the point 0. (c) The family of functions $F(z)=az+b$, where a and b are arbitrary coefficients, is not normal in the neighbourhood of any point.

2. If $\{F_n(z)\}$ is a normal and almost bounded sequence of functions on a region G, and if there exists a function $F(z)$ such that every almost uniformly convergent subsequence of this sequence is convergent to $F(z)$, then $F_n(z)\rightrightarrows F(z)$ in G.

§ 4. Equi-continuous functions.

We say that the functions belonging to a family $\mathfrak{F}$ of finite functions are *equi-continuous* on a set A, if to every positive number ε there corresponds a number $\eta>0$ such that for every function $F(z)$ of the family $\mathfrak{F}$ and for every pair of points z_1, z_2 of the set A the inequality $\varrho(z_1,z_2)<\eta$ implies $|F(z_2)-F(z_1)|<\varepsilon$. From this definition it follows immediately that the functions belonging to a family of equi-continuous functions on some set are all uniformly continuous on this set (cf. § 1, p. 43).

(4.1) ASCOLI'S THEOREM. *Every bounded sequence $\{F_n(z)\}$ of functions equi-continuous in an open and bounded set G contains a subsequence uniformly convergent in G.*

Proof. Let $\{w_j\}$ be a sequence of all rational points (cf. Introduction, § 8) of the set G. By theorem 3.2, a subsequence $\{F_{n_p}(z)\}$, convergent at all points of the sequence $\{w_j\}$, can be extracted from the sequence $\{F_n(z)\}$. Writing for brevity $H_p(z)=F_{n_p}(z)$ we shall prove that the sequence $\{H_p(z)\}$ is uniformly convergent on the entire set G.

With this in view, let ε be an arbitrary positive number and let η be a corresponding positive number such that

(4.2) $z_1 \epsilon G, \quad z_2 \epsilon G, \quad |z_2 - z_1| < \eta \quad$ implies $\quad |H_p(z_2) - H_p(z_1)| < \varepsilon/3$

for $p = 1, 2, \ldots$

The denumerable sequence of circles $K_j = K(w_j; \eta)$ covers the closed set $\overline{G}$, and hence by the theorem of Borel (Introduction, theorem 6.2) there exists a finite system $K_1, K_2, \ldots, K_m$ of these circles jointly covering the set $\overline{G}$, and hence the given open set G. Let p_0 be a natural number such that

(4.3) $|H_p(w_j) - H_q(w_j)| < \varepsilon/3 \quad$ for $\quad p > p_0, \quad q > p_0, \quad j = 1, 2, \ldots, m.$

Let us consider an arbitrary point $z \epsilon G$. This point belongs to at least one of the circles $K_1, K_2, \ldots, K_m$. For a certain value of the index $j_0 \leqslant m$ we therefore have $|z - w_{j_0}| < \eta$ and, by (4.2) and (4.3), for $p > p_0$ and $q > p_0$,

$$|H_p(z) - H_p(w_{j_0})| < \varepsilon/3, \qquad |H_p(w_{j_0}) - H_q(w_{j_0})| < \varepsilon/3,$$
$$|H_q(w_{j_0}) - H_q(z)| < \varepsilon/3;$$

by adding these three inequalities we obtain $|H_p(z) - H_q(z)| < \varepsilon$ for every point $z \epsilon G$ when $p > p_0$ and $q > p_0$. The sequence $\{H_p(z)\}$ is therefore uniformly convergent in G, q. e. d.

From theorem 4.1 we deduce the following criterion for the normality of a family of functions:

(4.4) *If $\mathfrak{F}$ is a family of functions equi-continuous in the neighbourhood of every point of a certain region G (i. e. if every point of the region has a neighbourhood in which the functions of the family $\mathfrak{F}$ are equi-continuous), then $\mathfrak{F}$ is a normal family in G.*

Proof. In virtue of theorem 3.3 it is sufficient to prove that the family $\mathfrak{F}$ is normal in the neighbourhood of every point of the region. Hence, let us consider an arbitrary point $z_0 \epsilon G$, and let $K_0 \subset G$ be a neighbourhood of the point z_0 in which the functions of the family $\mathfrak{F}$ are equi-continuous. We can assume that this neighbourhood is so small that $|F(z) - F(z_0)| < 1$ for every function F of the family $\mathfrak{F}$ and for every point $z \epsilon K_0$.

Let us consider an arbitrary sequence $\{F_n(z)\}$ of functions of the family $\mathfrak{F}$ and let $\Phi_n(z) = F_n(z) - F_n(z_0)$. The functions $\Phi_n(z)$ are equi-continuous and at the same time uniformly bounded in the circle K_0. Therefore, by Ascoli's theorem (4.1), the sequence $\{\Phi_n(z)\}$ contains a subsequence $\{\Phi_{n_k}(z)\}$ uniformly convergent in K_0. On the other hand, the sequence of numbers $\{F_{n_k}(z_0)\}$ contains a subsequence $\{F_{n_{k_j}}(z_0)\}$ which is convergent to a finite limit or to ∞. In the first case, the sequence $\{F_{n_{k_j}}(z) = \Phi_{n_{k_j}}(z) + F_{n_{k_j}}(z_0)\}$ is uniformly convergent in K_0; in the second case, it is uniformly divergent to ∞. The family $\mathfrak{F}$ is therefore normal in the neighbourhood of every point z_0 of the region G, q. e. d.

In Chapter II, § 7, we shall give an application of theorem 4.4 in the domain of holomorphic functions.

EXERCISES. 1. Show that theorem 4.1 remains true for every bounded set G (not necessarily open); the assumption of boundedness of the set is essential.

2. If $\mathfrak{F}$ is a normal and almost bounded family of continuous functions in an open set G, then these functions are equi-continuous on every closed set $P \subset G$ (converse of theorem 4.4).

§ 5. The total differential. We say that the expression $M\Delta x + N\Delta y$, linear in the real variables Δx and Δy, with coefficients M and N (in general, complex), is the *total differential* of the function (in general, complex) $W(z) = W(x, y)$ at the point $z_0 = (x_0, y_0)$ if

$$(5.1) \qquad \frac{W(x_0 + \Delta x, y_0 + \Delta y) - W(x_0, y_0) - M\Delta x - N\Delta y}{(\Delta x^2 + \Delta y^2)^{1/2}} \to 0,$$

when $\Delta x \to 0$, $\Delta y \to 0$, *i. e.* if

$$W(x_0 + \Delta x, y_0 + \Delta y) = W(x_0, y_0) + M\Delta x + N\Delta y + \varepsilon(\Delta x, \Delta y)(\Delta x^2 + \Delta y^2)^{1/2},$$
$$(5.2) \qquad \text{where} \quad \varepsilon(\Delta x, \Delta y) \to 0 \quad \text{as} \quad \Delta x \to 0, \ \Delta y \to 0.$$

Substituting in (5.1) $\Delta y = 0$ or $\Delta x = 0$, we obtain $M = W'_x(x_0, y_0)$ and $N = W'_y(x_0, y_0)$. Consequently:

(5.3) *A function $W(x, y)$, having a total differential at the point (x_0, y_0), has partial derivatives at this point with respect to each of the variables x, y, and these derivatives are respectively equal to the coefficients of the differential.*

Substituting $W(x, y) = U(x, y) + iV(x, y)$, $M = A + Bi$, $N = C + Di$, where $U(x, y)$, $V(x, y)$ are real functions, and A, B, C, D real numbers, we verify immediately that relation (5.1) is equivalent to the two relations

$$\frac{U(x_0+\Delta x, y_0+\Delta y)-U(x_0,y_0)-A\Delta x-C\Delta y}{(\Delta x^2+\Delta y^2)^{1/2}}\to 0,$$

$$\frac{V(x_0+\Delta x, y_0+\Delta y)-V(x_0,y_0)-B\Delta x-D\Delta y}{(\Delta x^2+\Delta y^2)^{1/2}}\to 0,$$

when $\Delta x\to 0$, $\Delta y\to 0$. Whence:

(5.4) *In order that* $W(z)=W(x,y)=U(x,y)+iV(x,y)$ *have a total differential at the point* $z_0=(x_0,y_0)$, *it is necessary and sufficient that both its real and imaginary parts,* $U(x,y)$ *and* $V(x,y)$, *have total differentials.*

The converse of theorem 5.3 would obviously be false: a partially differentiable function need not have a total differential. However:

(5.5) *If a function* $W(x,y)$ *has partial derivatives* $W'_x(x,y)$ *and* $W'_y(x,y)$ *everywhere in a neighbourhood of the point* (x_0,y_0) *and if these derivatives are continuous at the point* (x_0,y_0), *then the function has a total differential at this point.*

Proof. In virtue of theorem 5.4 we can assume that the function $W(x,y)$ is real. Applying the theorem of the mean, we have first of all

(5.6) $$W(x_0+\Delta x, y_0+\Delta y)-W(x_0,y_0)$$
$$=[W(x_0+\Delta x, y_0+\Delta y)-W(x_0,y_0+\Delta y)]+[W(x_0,y_0+\Delta y)-W(x_0,y_0)]$$
$$=W'_x(x_0+\vartheta_1\Delta x, y_0+\Delta y)\cdot\Delta x+W'_y(x_0,y_0+\vartheta_2\Delta y)\cdot\Delta y,$$

where $0\leqslant\vartheta_1\leqslant 1$, $0\leqslant\vartheta_2\leqslant 1$.

Hence, taking $M=W'_x(x_0,y_0)$, $N=W'_y(x_0,y_0)$, we can write the last member of relation (5.6) in the form $(M+\varepsilon_1)\Delta x+(N+\varepsilon_2)\Delta y$, where $\varepsilon_1,\varepsilon_2$ tend to zero together with $\Delta x,\Delta y$. Consequently,

$$\frac{W(x_0+\Delta x, y_0+\Delta y)-W(x_0,y_0)-M\Delta x-N\Delta y}{(\Delta x^2+\Delta y^2)^{1/2}}=\frac{\varepsilon_1\Delta x+\varepsilon_2\Delta y}{(\Delta x^2+\Delta y^2)^{1/2}}\to 0,$$

when $\Delta x\to 0$, $\Delta y\to 0$, and the expression $M\Delta x+N\Delta y$ is the total differential of the function $W(x,y)$ at the point (x_0,y_0).

The total differentiability of a function at a point (x,y) is therefore something intermediate between the mere existence of the partial derivatives at this point, and the condition that they be also continuous at this point. In many proofs, however, in which the continuity of the partial derivatives is usually assumed, total differentiability is sufficient. It is important to note that while the assumption of the continuity of the partial derivatives at the point (x_0,y_0) already implies the existence of these derivatives in a certain entire neighbourhood of this point, the assumption of total differentiability does not require this condition at all. On this, in fact, depends the significance of the total differential, which — introduced by Stolz — plays now an important role in the theory of real functions of several variables, as well as in certain more recent investigations of the foundations of the theory of the functions of a complex variable (see D. Menchoff, *Les conditions de monogénéité*, Paris 1936).

§ 6. The derivative in the complex domain. Cauchy-Riemann equations. A function of a complex variable $W(z)$, defined and finite in the neighbourhood of a point $z_0\neq\infty$, is said to be *differentiable* at this point, if the limit of the expression $[W(z_0+\Delta z)-W(z_0)]/\Delta z$, as Δz tends to zero through arbitrary complex values exists and is finite; this limit is then called the *derivative* of the function $W(z)$ at the point z_0, and we denote it by $W'(z_0)$ or $(dW/dz)_{z=z_0}$; in the case when the function W depends on other variables also, we use the usual symbols for the partial derivative: W'_z or $\partial W/\partial z$, to denote its derivative with respect to z.

This definition is formally identical with the definition adopted in real analysis. Therefore it is immediately evident that the formal rules of differentiation (of a sum, product, quotient of two functions, etc.) extend immediately from the real domain to the complex domain. However, the fact that the increment Δz of the independent variable in the above definition ranges over complex values, and not only over real ones, implies more essential consequences than would be expected at first glance. As we shall have the opportunity to emphasize many times, the condition of differentiability in the complex domain is something much stronger than in the real domain. It can be said that the entire theory of analytic functions is the investigation of the consequences of this condition.

Among the formal rules of differentiation that extend immediately from the real domain to the complex domain, we mention the following rule for differentiation of a composite function:

(6.1) *If a function* $W(z)$ *is differentiable at the point* $z_0\neq\infty$, *and assumes there the value* $w_0\neq\infty$ *and if the function* $F(w)$ *is defined in the neighbourhood of the point* w_0 *and is differentiable at this point, then the composite function* $\Phi(z)=F[W(z)]$ *is also differentiable at* z_0, *and* $\Phi'(z_0)=F'(w_0)\cdot W'(z_0)$.

Proof. In fact, the composite function $\Phi(z_0+h)$ is defined for sufficiently small values of h. Supposing at first that h assumes values for which $W(z_0+h)\neq W(z_0)$, we have for $h\to 0$,

$$\frac{\Phi(z_0+h)-\Phi(z_0)}{h}$$

$$=\frac{F[W(z_0+h)-F[W(z_0)]}{W(z_0+h)-W(z_0)}\cdot\frac{W(z_0+h)-W(z_0)}{h}\to F'[W(z_0)]W'(z_0).$$

If there exist values of $h\neq 0$ arbitrarily close to zero, such that $W(z_0+h)=W(z_0)$, then obviously $W'(z_0)=0$; however, for those values of h for which $W(z_0+h)=W(z_0)$, we also have $\Phi(z_0+h)-\Phi(z_0)=0$. Therefore, when h tends to zero in any manner whatsoever, we always have

$$\frac{\Phi(z_0+h)-\Phi(z_0)}{h}\to F'(w_0)W'(z_0).$$

As follows directly from the definition of the derivative, in order that the number Q be the derivative of the function $W(z)$ at the point $z_0=(x_0,y_0)$, it is necessary and sufficient that

$$\frac{W(z_0+\Delta z)-W(z_0)-Q\Delta z}{|\Delta z|}\to 0$$

when $\Delta z\to 0$; in other words, taking $\Delta z=\Delta x+i\Delta y$ and remembering that $|\Delta z|=\sqrt{\Delta x^2+\Delta y^2}$, we can say: in order that $Q=W'(z_0)$ it is necessary and sufficient that the expression $Q\Delta z=Q\Delta x+iQ\Delta y$ be the total differential of the function $W(z)$ at the point z_0. If this condition is satisfied, then by theorem 5.3 we have $W'_x(x_0,y_0)=Q$ and $W'_y(x_0,y_0)=iQ$, and hence

(6.2) $$W'_x(x_0,y_0)=-iW'_y(x_0,y_0).$$

Conversely, if the function $W(z)$ has a total differential at $z_0=x_0+iy_0$ and satisfies equation (6.2), then, denoting the common value of the sides of this equation by Q, we obtain the expression $Q\Delta x+iQ\Delta y=Q\Delta z$ as the total differential of the function $W(z)$ at the point z_0, and therefore $Q=W'(z_0)$. Taking $W(z)=U(x,y)+iV(x,y)$, where $U(x,y)$ and $V(x,y)$ are respectively the real and imaginary parts of the function $W(z)$, we can write the relation (6.2) in the form of two real equations:

(6.3) $$U'_x(x_0,y_0)=V'_y(x_0,y_0),\qquad U'_y(x_0,y_0)=-V'_x(x_0,y_0),$$

and by theorem 5.4 we obtain the following theorem, which reduces differentiation in the complex domain to the differentiation of real functions:

(6.4) *In order that the function $W(z)=U(x,y)+iV(x,y)$ have a derivative at the point $z_0=(x_0,y_0)$, it is necessary and sufficient that both its real and imaginary parts, $U(x,y)$ and $V(x,y)$, have a total differential at (x_0,y_0) and satisfy equations (6.3), or — equivalently — that the function $W(z)$ have a total differential and satisfy equation (6.2). If this condition is satisfied, then*

$$W'(z_0)=W'_x(x_0,y_0)=-iW'_y(x_0,y_0),$$

and $W'(z_0)(\Delta x+i\Delta y)$ is the total differential of the function $W(z)$ at the point z_0.

Relations (6.3), which — when x_0 and y_0 are considered as variables — are partial differential equations, are usually called the *Cauchy-Riemann equations*. Relation (6.2) is obviously only another form of these equations, and we shall call it the *Cauchy-Riemann condition in complex form.*

In the next few sections we shall give examples of several fundamental complex functions which are differentiable in the entire open plane.

EXERCISES. 1. Determine the points at which the following functions have derivatives: (a) $W(z)=z$, (b) $W(z)=|z|$, (c) $W(z)=|z|^2$, (d) $W(z)=\Re z$.

2. If the function $W(z)$ has the total differential $(A+Bi)\Delta x+(C+Di)\Delta y$ at the point z_0, then in order that the function $\overline{W(z)}$, the conjugate of $W(z)$, have a derivative at the point z_0, it is necessary and sufficient that $A=-D$ and $B=C$, *i. e.* that $W'_x(z_0)=iW'_y(z_0)$.

Deduce from this that if the function $W(z)$ has a derivative at the point z_0, then the conjugate $\overline{W(z)}$ of this function has a derivative at this point if and only if, when $W'(z_0)=0$.

3. If for the function $W(z)$, having a total differential at the point z_0, there exists a finite limit of the real part of the expression $[W(z)-W(z_0)]/(z-z_0)$ when $z\to z_0$, then the function has a derivative at the point z_0.

4. If for the function $W(z)$, having a total differential at the point z_0, there exists a finite limit, when $z\to z_0$, of the absolute value of the expression $[W(z)-W(z_0)]/(z-z_0)$, then either the function $W(z)$ or the conjugate function $\overline{W(z)}$ has a derivative at the point z_0.

5. Verify that the function $F(z)=\sqrt{|xy|}$ $(z=x+iy)$ has partial derivatives with respect to x and y equal to zero at the point $z=0$; hence, it satisfies the Cauchy-Riemann equations; however, it has no ordinary derivative at this point (hence it has no total differential either).

[*Hint.* Substituting $z=x(1+ia)$, where a is an arbitrary real number, notice that the quotient $[F(z)-F(0)]/z=F(z)/z$ has a finite limit for each fixed value of a when $x\to 0$, but this limit varies with a.]

6. The Cauchy-Riemann conditions are satisfied at the point 0 for the continuous function $G(z)$, defined by the formulae $G(0)=0$ and $G(z)=x^2y/(x^4+y^2)$ for $z=x+iy\neq 0$, and the limit of the ratio $[G(z)-G(0)]/z=G(z)/z$ exists when z tends to zero along an arbitrary straight line passing through the point 0; besides, this limit has the same value, 0, for every such straight line. Nevertheless, the function $G(z)$ does not have a derivative at the point 0.

[*Hint.* The limit of the ratio $G(z)/z$ when z tends to zero along the parabola $y=x^2$ is $\frac{1}{2}$.]

7. If $W(z)$ is a polynomial of the n-th degree and $z_1, z_2, \ldots, z_n$ denote the roots of this polynomial (if multiple roots exist, the number of the occurrences of each root in the sequence $z_1, z_2, \ldots, z_n$ is equal to its multiplicity), then

$$(*)\qquad \frac{W'(z)}{W(z)}=\sum_k \frac{1}{z-z_k}$$

at each point z for which $W(z)\neq 0$.

8. If $W(z)$ is a polynomial of degree >1, then all the roots of the equation $W'(z)=0$ lie in the convex set determined (cf. Introduction, § 8, exercise 2) by the roots of the equation $W(z)=0$ (G a u s s).

[*Hint.* Make use of the formula (*) of the preceding exercise; note that $1/(z-z_k)=(\overline{z-z_k})/|z-z_k|^2$; use exercise 2, Introduction, § 8.]

9. The equation $1+z+az^n=0$, where n is an integer $\geqslant 2$ and a an arbitrary complex number, always has a root whose absolute value is $\leqslant 2$ (L a n d a u).

[*Hint.* Substitute $z=1/\mathfrak{z}$ and use the theorem of Gauss (exercise 8).]

§ 7. The exponential function. In real analysis we usually define power and the number e directly; then we prove, as a theorem, the well-known expansion of the exponential function e^x in a power series of the variable x for real values of this variable. In the complex domain this method necessarily fails (since the usual definition of power cannot be directly extended to complex values of the exponent). Hence in the present exposition we shall define the exponential function e^z — which we shall also denote by $\exp z$ — directly by the equation

$$(7.1)\qquad \exp z=e^z=\sum_{n=0}^{\infty}\frac{z^n}{n!},$$

understanding $0!$ to be (as usual) the number 1.

It is seen at once (applying *e. g.* the ratio test) that the series appearing in this equation is convergent in the entire open plane and, moreover, convergent absolutely and uniformly

in every circle $K(0;r)$, since in such a circle we obviously have $|z^n/n!|<r^n/n!$. Therefore:

(7.2) *The exponential function is defined and continuous in the entire open plane.*

From definition (7.1) we shall derive fundamental formulae concerning the exponential function (without referring to results known from real analysis). In particular:

$$(7.3)\qquad \exp 0=e^0=1,$$

$$(7.4)\qquad \exp(a+b)=\exp a\cdot\exp b\quad \text{for every pair of values } a, b.$$

Formula (7.3) is obtained directly by substituting $z=0$ in (7.1). In order to obtain formula (7.4) let us note that, by Cauchy's theorem on the multiplication of series, we have

$$\exp a\cdot\exp b=\sum_{n=0}^{\infty}\frac{a^n}{n!}\cdot\sum_{n=0}^{\infty}\frac{b^n}{n!}=\sum_{n=0}^{\infty}c_n,$$

where

$$c_n=\sum_{k=0}^{n}\frac{a^k b^{n-k}}{k!\,(n-k)!}=\frac{1}{n!}\sum_{k=0}^{n}\frac{n!}{k!\,(n-k)!}a^k b^{n-k}=\frac{(a+b)^n}{n!},$$

and therefore

$$\exp a\cdot\exp b=\sum_{n=0}^{\infty}\frac{(a+b)^n}{n!}=\exp(a+b).$$

(7.5) *The exponential function does not vanish at any point of the plane.*

P r o o f. Indeed, were $\exp a=0$, then by (7.3) and (7.4) we should have $1=\exp 0=\exp a\cdot\exp(-a)=0$.

From definition (7.1) it follows immediately that for real values of $x\geqslant 0$ the function $\exp x$ is a constantly increasing function, varying from 1 to $+\infty$ when x varies from 0 to $+\infty$. Making use of the equation $\exp(-x)=1/\exp x$, which is a consequence of (7.3) and (7.4), we establish more generally that

(7.6) *The function* $\exp x$ *in the real domain is a positive and increasing function, and*

$$\exp x\to+\infty \quad \textit{for} \quad x\to+\infty,$$
$$\exp x\to 0 \quad \textit{for} \quad x\to-\infty.$$

Finally, we have from (7.4) and (7.1) for $h\neq 0$

$$(7.7)\qquad \frac{\exp(z+h)-\exp z}{h}=\exp z\,\frac{\exp h-1}{h}=\exp z\sum_{n=1}^{\infty}\frac{h^{n-1}}{n!},$$

whence, passing to the limit as $h\to 0$, we obtain (making use of the fact that the series appearing in the equation (7.7) represents a continuous function of the variable h) the equation

$$\lim_{h\to 0}\frac{\exp(z+h)-\exp z}{h}=\exp z.$$

Consequently:

(7.8) *The exponential function is differentiable and equal to its derivative in the entire open plane.*

EXERCISE. 1. For every real z we have $e^z\geqslant 1+z$; if $0\leqslant z\leqslant 1$, then $1-z\leqslant e^{-z}\leqslant 1-z/2$.

§ 8. Trigonometric functions. As in the case of the exponential function, we define the trigonometric functions $\cos z$ and $\sin z$ in the complex domain as the sums of series. Namely:

$$(8.1)\qquad \cos z=\sum_{n=0}^{\infty}\frac{(-1)^n z^{2n}}{(2n)!},\qquad \sin z=\sum_{n=0}^{\infty}\frac{(-1)^n z^{2n+1}}{(2n+1)!}.$$

The series appearing in these formulae are convergent in the entire open plane, and therefore the functions $\cos z$ and $\sin z$ are defined in the entire open plane, and it is apparent at once from formulae (8.1) that the first of these functions is even and the second odd.

For real values of z the equations (8.1) coincide with the well-known expansions of the functions $\cos z$ and $\sin z$ in power series. In the real domain the present definition, therefore, coincides with the traditional one based on geometric methods. However, we shall not make use of the results of the geometric theory here; the fundamental properties of the trigonometric functions will be deduced directly from formulae (8.1). In this way we shall obtain not only an extension of these properties to the complex domain, but at the same time an arithmetization of the theory of trigonometric functions in the real domain. The results so obtained indicate simultaneously the direct and natural way to the arithmetization of the concept of angle.

We shall begin with the establishment of relations between the trigonometric functions. In virtue of (7.1) we have

$$\exp iz=\sum_{n=0}^{\infty}\frac{i^n z^n}{n!}=\sum_{k=0}^{\infty}\frac{i^{2k}z^{2k}}{(2k)!}+i\sum_{k=0}^{\infty}\frac{i^{2k}\,z^{2k+1}}{(2k+1)!}$$
$$=\sum_{k=0}^{\infty}\frac{(-1)^k z^{2k}}{(2k)!}+i\sum_{k=0}^{\infty}\frac{(-1)^k z^{2k+1}}{(2k+1)!},$$

and hence, by (8.1),

$$(8.2)\qquad \exp iz=\cos z+i\sin z.$$

Substituting $-z$ for z we obtain the equation

$$\exp(-iz)=\cos z-i\sin z;$$

adding it to (8.2), and subtracting it from (8.2), we find:

$$(8.3)\qquad \cos z=\frac{1}{2}[\exp iz+\exp(-iz)],\qquad \sin z=-\frac{1}{2}i[\exp iz-\exp(-iz)].$$

These are the so-called *Euler formulae*. By formal manipulations, in which we make use only of the properties (7.3) and (7.4) of the exponential function, we obtain the following fundamental relations betwen the trigonometric functions:

$$(8.4)\qquad \begin{gathered}\cos^2 z+\sin^2 z=1,\qquad \cos(a\pm b)=\cos a\cos b\mp\sin a\sin b,\\ \sin(a\pm b)=\sin a\cos b\pm\cos a\sin b,\qquad \text{etc.}\end{gathered}$$

Since the exponential function is by theorems 7.2 and 7.8 continuous and differentiable in the entire open plane, it also follows from formulae (8.3) that

(8.5) *The trigonometric functions* $\cos z$ *and* $\sin z$ *are continuous and differentiable in the entire open plane.*

Differentiating formulae (8.3) we obtain, in virtue of theorem 7.8,

$$(8.6)\qquad (\cos z)'=-\sin z,\qquad (\sin z)'=\cos z.$$

From the equality $\cos 0=1$ and from the easy estimate

$$\cos 2=1-\frac{2^2}{1\cdot 2}+\sum_{n=2}^{\infty}\frac{(-1)^n 2^{2n}}{(2n)!}$$
$$<-1+\sum_{n=2}^{\infty}\frac{2^{2n}}{(2n)!}<-1+\frac{2^4}{4!}\sum_{k=0}^{\infty}\left(\frac{2}{5}\right)^{2k}=-1+\frac{50}{63}<0,$$

it follows immediately that the function $\cos z$ has at least one positive real root. The smallest of these roots (*i. e.* their lower bound) will be denoted by $\pi/2$. In the interval $[0,\pi/2]$ the function $\cos z$ is therefore always positive, with the exception of the right end-point, at which it vanishes. In virtue of the second of the formulae (8.6) the function $\sin z$ therefore increases in this entire interval, and hence is always positive in this interval, except at the left end-point, at which it vanishes. Therefore, in view again of the first of the formulae (8.6), the function $\cos z$ is decreasing in the interval $[0,\pi/2]$. Since in virtue of (8.4) we have $1=\cos^2\pi/2+\sin^2\pi/2=\sin^2\pi/2$ and, as we have seen, the function $\sin z$ is non-negative in the interval $[0,\pi/2]$, it follows that $\sin\pi/2=1$. Making use, then, of the second and third formula in (8.4), we obtain in succession:

$$\cos\pi/2=0,\quad \sin\pi/2=1;\quad \cos\pi=-1,\quad \sin\pi=0;$$
$$\cos(3\pi/2)=0,\quad \sin(3\pi/2)=-1;\quad \cos 2\pi=1,\quad \sin 2\pi=0;$$

and more generally, for an arbitrary complex number z,

$$\text{(8.7)}\qquad \cos(z+\pi/2)=-\sin z,\quad \sin(z+\pi/2)=\cos z;\qquad \cos(z+2\pi)=\cos z,\quad \sin(z+2\pi)=\sin z.$$

From the second pair of formulae (8.7) it follows that

(8.8) *The functions* $\cos z$ *and* $\sin z$ *are periodic, with period* 2π, *in the entire open plane.*

On the other hand, the first pair of formulae (8.7) enables one to deduce the behaviour of the functions considered in the intervals ("quadrants") $[\pi/2,\pi]$, $[\pi,3\pi/2]$, etc., in view of the investigation, already made, of their behaviour in the interval $[0,\pi/2]$. Furthermore, since the function $\cos z$ is decreasing, and the function $\sin z$ increasing, in the interval $[0,\pi/2]$, while both functions are continuous and non-negative in this interval, we obtain by (8.7) and by the first of the formulae (8.4) the following theorem:

(8.9) *If a and b are real numbers satisfying the condition $a^2+b^2=1$, then there exists exactly one value of θ such that:*

$$\cos\theta=a,\qquad \sin\theta=b,\qquad -\pi<\theta\leqslant\pi.$$

The trigonometric functions $\tan z$ and $\cot z$ are defined by the formulae:

$$\text{(8.10)}\qquad \tan z=\frac{\sin z}{\cos z},\qquad \cot z=\frac{\cos z}{\sin z}.$$

From this definition and from the properties of the functions $\cos z$ and $\sin z$ already deduced, it follows directly that:

(8.11) *The functions* $\tan z$ *and* $\cot z$ *are continuous functions in the entire open plane, finite and differentiable everywhere, with the exception of the points at which the function* $\cos z$, *or the function* $\sin z$, *vanishes and at which the functions* $\tan z$ *and* $\cot z$, *respectively, assume the value* ∞. *The derivatives of these functions are given by the formulae:*

$$\text{(8.12)}\qquad (\tan z)'=\frac{1}{\cos^2 z},\qquad (\cot z)'=-\frac{1}{\sin^2 z},$$

which we obtain by differentiating the equations (8.10) *with the aid of formulae* (8.6).

Finally, the Euler formulae (8.3) give for the functions $\tan z$ and $\cot z$:

$$\text{(8.13)}\qquad \tan z=-i\,\frac{\exp iz-\exp(-iz)}{\exp iz+\exp(-iz)},\qquad \cot z=i\,\frac{\exp iz+\exp(-iz)}{\exp iz-\exp(-iz)}.$$

From the behaviour of the functions $\cos z$ and $\sin z$ in the interval $[0,\pi/2]$ it follows that the function $\tan z$ is steadily increasing in the interval $[0,\pi/4]$ and assumes the values from 0 to 1. Therefore, making the substitution $t=a\tan u$ in the integral $\int_0^a \frac{dt}{a^2+t^2}$ (where a is an arbitrary positive number), we obtain by simple manipulations, in virtue of the first of the relations (8.4), the equation

$$\text{(8.14)}\qquad \int_0^a \frac{dt}{a^2+t^2}=\frac{1}{a}\int_0^{\pi/4}\frac{1}{1+\tan^2 u}\cdot\frac{d\tan u}{du}\,du=\frac{1}{a}\int_0^{\pi/4}du=\frac{\pi}{4a},$$

which expresses the number π in terms of a definite integral of a rational function.

Similarly,

$$\text{(8.15)}\qquad \int_0^{+\infty}\frac{dt}{a^2+t^2}=\frac{\pi}{2a}.$$

By an analogous substitution $t=a\sin u$ in the integral $\int_0^a \frac{dt}{\sqrt{a^2-t^2}}$ we obtain the formula

$$\text{(8.16)}\qquad \int_0^a \frac{dt}{\sqrt{a^2-t^2}}=\frac{\pi}{2}.$$

EXERCISES. 1. *Hyperbolic functions.* The hyperbolic functions, the hyperbolic cosine and the hyperbolic sine, are defined by formulae analogous to the Euler formulae:

$$\cosh z=\frac{1}{2}[\exp z+\exp(-z)],\qquad \sinh z=\frac{1}{2}[\exp z-\exp(-z)].$$

Verify the following relations between the hyperbolic and trigonometric functions: $\cosh z=\cos iz$, $\sinh z=-i\sin iz$. From this derive formulae analogous to the formulae (8.4):

$$\cosh^2 z-\sinh^2 z=1, \qquad \cosh(a\pm b)=\cosh a\cdot\cosh b\pm\sinh a\cdot\sinh b,$$

etc., as well as formulae analogous to formulae (8.6) on derivatives:

$$(\cosh z)'=\sinh z, \qquad (\sinh z)'=\cosh z.$$

Verify the following expansions of the hyperbolic functions in power series:

$$\cosh z=\sum_{n=0}^{\infty}\frac{z^{2n}}{(2n)!}, \qquad \sinh z=\sum_{n=0}^{\infty}\frac{z^{2n+1}}{(2n+1)!},$$

from which it follows, in particular, that $|\cos z|\leqslant\cosh|z|$ and $|\sin z|\leqslant\sinh|z|$.

2. Prove the following identities, in which m denotes an arbitrary positive integer:

$$\cos mz=\cos^m z-\binom{m}{2}\cos^{m-2}z\sin^2 z+\binom{m}{4}\cos^{m-4}z\sin^4 z-\dots$$

(the last term of the right side is $(-1)^{m/2}\sin^m z$ or $(-1)^{(m-1)/2}m\cos z\sin^{m-1}z$, depending on whether m is an even or odd number);

$$\sin mz=m\cos^{m-1}z\sin z-\binom{m}{3}\cos^{m-3}z\sin^3 z+\dots$$

(the last term of the right side is $(-1)^{(m+2)/2}m\cos z\sin^{m-1}z$ or $(-1)^{(m-1)/2}\sin^m z$, depending on whether m is an even or odd number);

$$2^{m-1}\cos^m z=\cos mz+\binom{m}{1}\cos(m-2)z+\binom{m}{2}\cos(m-4)z+\dots$$

(the last term of the right side is $\frac{1}{2}\binom{m}{m/2}$ or $\binom{m}{(m-1)/2}\cos z$, depending on whether m is an even or odd number);

the analogous formula for $\sin^m z$: for even m,

$$(2i)^m\sin^m z=2\cos mz-2\binom{m}{1}\cos(m-2)z+2\binom{m}{2}\cos(m-4)z-\dots+(-1)^{m/2}\binom{m}{m/2},$$

and for odd m,

$$(2i)^{m-1}\sin^m z=\sin mz-\binom{m}{1}\sin(m-2)z+\binom{m}{2}\sin(m-4)z+\dots+(-1)^{(m-1)/2}\binom{m}{(m-1)/2}\sin z.$$

[*Hint*. Take, first, z real in the identities

$$\cos mz+i\sin mz=(\cos z+i\sin z)^m, \qquad 2^m\cos^m z=(e^{iz}+e^{-iz})^m,$$

which follow directly from formulae (8.2) and (8.3), expand the right sides by Newton's binomial theorem and equate the real and imaginary parts of both sides. The generalization to complex values of z follows from the theorem stating that a power series which vanishes for real values of the variable, vanishes identically.]

3. Show that the trigonometric expression

$$2^n\sin^n\theta+\binom{n}{1}2^{n-1}\sin^{n-1}\theta\cos\left(\theta+\frac{\pi}{2}\right)+\binom{n}{2}2^{n-2}\sin^{n-2}\theta\cos 2\left(\theta+\frac{\pi}{2}\right)+\dots+\cos n\left(\theta+\frac{\pi}{2}\right)$$

is equal to $(-1)^{n/2}\cos n\theta$ or $(-1)^{(n-1)/2}\sin n\theta$, depending on whether n is an even or odd number.

[*Hint*. Consider the expression $\{2\sin\theta+\exp i(\theta+\pi/2)\}^n$.]

4. Substituting $w=\sin z$ show that:

(a) $\cos mz=F_1(w)$, $\sin mz/\cos z=G_1(w)$, for every even positive integral value of m,

(b) $\cos mz/\cos z=F_2(w)$, $\sin mz=G_2(w)$, for every odd positive integral value of m,

where $F_1(w)$ and $G_2(w)$ are polynomials in w of degree m, and $G_1(w)$ and $F_2(w)$ are of degree $m-1$.

[*Hint*. Cf. exercise 2.]

5. Show that the roots of the polynomials $F_1(w)$, $G_1(w)$, $F_2(w)$, and $G_2(w)$, exercise 4, are respectively:

$$\pm\sin\frac{\pi}{2m},\quad \pm\sin\frac{3\pi}{2m},\quad \dots,\quad \pm\sin\frac{(m-1)\pi}{2m};$$

$$0,\quad \pm\sin\frac{\pi}{m},\quad \pm\sin\frac{2\pi}{m},\quad \dots,\quad \pm\sin\frac{m-2}{2}\frac{\pi}{m};$$

$$\pm\sin\frac{\pi}{2m},\quad \pm\sin\frac{3\pi}{2m},\quad \dots,\quad \pm\sin\frac{(m-2)\pi}{2m};$$

$$0,\quad \pm\sin\frac{\pi}{m},\quad \pm\sin\frac{2\pi}{m},\quad \dots,\quad \pm\sin\frac{m-1}{2}\frac{\pi}{m}.$$

Derive from this the following formulae:

(α) for even m:

$$\cos mz=\left(1-\frac{\sin^2 z}{\sin^2\frac{\pi}{2m}}\right)\left(1-\frac{\sin^2 z}{\sin^2\frac{3\pi}{2m}}\right)\dots\left(1-\frac{\sin^2 z}{\sin^2\frac{(m-1)\pi}{2m}}\right),$$

$$\frac{\sin mz}{\cos z}=m\sin z\cdot\left(1-\frac{\sin^2 z}{\sin^2\frac{\pi}{m}}\right)\left(1-\frac{\sin^2 z}{\sin^2\frac{2\pi}{m}}\right)\dots\left(1-\frac{\sin^2 z}{\sin^2\frac{m-2}{2}\frac{\pi}{m}}\right);$$

(β) for odd m:

$$\frac{\cos mz}{\cos z}=\left(1-\frac{\sin^2 z}{\sin^2\frac{\pi}{2m}}\right)\left(1-\frac{\sin^2 z}{\sin^2\frac{3\pi}{2m}}\right)\dots\left(1-\frac{\sin^2 z}{\sin^2\frac{(m-2)\pi}{2m}}\right),$$

$$\sin mz = m \sin z \cdot \left(1-\frac{\sin^2 z}{\sin^2 \frac{\pi}{m}}\right)\left(1-\frac{\sin^2 z}{\sin^2 \frac{2\pi}{m}}\right)\cdots\left(1-\frac{\sin^2 z}{\sin^2 \frac{m-1}{2}\frac{\pi}{m}}\right).$$

6. Substituting z/m for z in formulae (a) of exercise 5 and passing to the limit as $m\to+\infty$, derive the following expansions of the cosine and the sine in infinite products:

$$\cos z = \left(1-\frac{z^2}{(\pi/2)^2}\right)\left(1-\frac{z^2}{(3\pi/2)^2}\right)\left(1-\frac{z^2}{(5\pi/2)^2}\right)\cdots,$$

$$\sin z = z\left(1-\frac{z^2}{\pi^2}\right)\left(1-\frac{z^2}{4\pi^2}\right)\left(1-\frac{z^2}{9\pi^2}\right)\cdots$$

The preceding products are uniformly convergent in every circle of finite radius, *i. e.* almost uniformly in the entire plane (by the convergence of the infinite product $A_1\cdot A_2\cdot\ldots\cdot A_n\cdot\ldots$ we here mean the convergence of the sequence of partial products $P_1=A_1$, $P_2=A_1\cdot A_2$, ..., $P_n=A_1\cdot A_2\cdot\ldots\cdot A_n,\ldots$; a more detailed discussion of the notion of limit of an infinite product will be given in Chapter VII).

7. Let $f(t)$ be a real function, bounded and continuous, or having at most a finite number of discontinuities in the entire interval $(-\infty,+\infty)$. Then at each point τ at which the function is continuous we have

$$\int_{-\infty}^{+\infty}\frac{y f(t)\,dt}{y^2+(t-\tau)^2}\to\pi f(\tau),$$

as y tends to 0 through positive real values.

§ 9. Argument. Let $z=x+iy\neq 0$, where $x=\mathcal{R}z$, $y=\mathcal{J}z$. By theorem 8.9, there exists a real number θ satisfying the equations:

(9.1) $$\cos\theta=\frac{x}{\sqrt{x^2+y^2}}=\frac{\mathcal{R}z}{|z|},\qquad \sin\theta=\frac{y}{\sqrt{x^2+y^2}}=\frac{\mathcal{J}z}{|z|}.$$

Therefore, taking $r=|z|$ and making use of formula (8.2), we have:

(9.2) $z=r(\cos\theta+i\sin\theta)=re^{i\theta}$, *where* $r>0$ *and* θ *is a real number.*

Every real number θ satisfying conditions (9.1) is called an *argument* or on *amplitude* of the complex number z; in virtue of (8.8) and (8.9), it is defined uniquely to within an additive constant of the form $2k\pi$, where k is an arbitrary integer. Therefore every complex number $z\neq 0$ has an infinite number of arguments differing from one another by a multiple of 2π, or — using the terminology of arithmetic and the theory of numbers — congruent mod 2π. (We say that two numbers a and b are *congruent modulo* c, and we write $a\equiv b \pmod{c}$, if $b-a$ is an integral multiple of c.)

It will cause no misunderstanding if occasionally we use the expression "the argument of z", when we mean an arbitrary argument of z. A similar convention will apply later to the logarithm, power and angle.

We denote the argument of z (*i. e.* any of the arguments of z) by $\arg z$. Among the arguments θ of the same number z, exactly one satisfies the inequality $-\pi<\theta\leqslant\pi$; we shall call it the *principal argument* of this number and denote it by $\operatorname{Arg} z$. If z is a real number, then $\operatorname{Arg} z$ is equal to 0 or π, depending on whether z is a positive or negative number.

The function $\operatorname{Arg} z$, defined in this manner in the entire open plane with the exception of the point 0, is obviously bounded in this entire region. Moreover:

(9.3) *The function* $\operatorname{Arg} z$ *is continuous at each point* $z=z_0$ *of the open plane which does not lie on the negative real half-axis.*

Proof. Let $\theta_0=\operatorname{Arg} z_0$. Since the point z_0 does not lie on the negative real half-axis, therefore certainly $\theta_0\neq\pi$, and hence

(9.4) $$-\pi<\theta_0<\pi.$$

Let us assume that the function $\operatorname{Arg} z$ is not continuous at the point z_0. Then we can determine a sequence of points $\{z_n\}$ tending to z_0 in such a way that the sequence of their principal arguments tends to a certain number $\tilde\theta_0$ different from θ_0. Since for every n we have $-\pi<\operatorname{Arg} z_n\leqslant\pi$,

(9.5) $$-\pi\leqslant\tilde\theta_0\leqslant\pi.$$

Writing for brevity $\theta_n=\operatorname{Arg} z_n$, for $n=1,2,\ldots$, we shall have $\cos\theta_n=\mathcal{R}z_n/|z_n|$ and $\sin\theta_n=\mathcal{J}z_n/|z_n|$ for $n=1,2,\ldots$, whence it follows that $\cos\tilde\theta_0=\mathcal{R}z_0/|z_0|=\cos\theta_0$ and $\sin\tilde\theta_0=\mathcal{J}z_0/|z_0|=\sin\theta_0$, as $z_n\to z_0$ and $\theta_n\to\tilde\theta_0$. From this, however, in virtue of (9.4) and (9.5) (cf. theorem 8.9), we obtain $\tilde\theta_0=\theta_0$ and we come to a contradiction.

Returning to the representation (9.2) of complex numbers, we notice that the number $z=0$ also has a representation of this form: it is sufficient to take $r=0$, and for θ an arbitrary real number. For symmetry, we shall understand the *argument of the number* 0 to mean an arbitrary real number. We can therefore say that:

(9.6) *Every complex number* z *has a representation of the form* (9.2). *In such a representation,* r *and* θ *are, respectively, the absolute value and the argument of the number* z.

The second part of this theorem still requires a brief justification. From (9.2), in virtue of (8.4), it follows first of all that $|z|=|r|\cdot|\cos\theta+i\sin\theta|=r$, and then that $\cos\theta=\mathcal{R}z/|z|$ and $\sin\theta=\mathcal{I}z/|z|$; this means that r and θ are the absolute value and the argument of the number z, respectively.

If $z_1=r_1\exp i\theta_1$, $z_2=r_2\exp i\theta_2$, then $z_1z_2=r_1r_2\exp i(\theta_1+\theta_2)$, $z_1/z_2=(r_1/r_2)\exp i(\theta_1-\theta_2)$. From theorem 9.6 it therefore follows that

(9.7) *When complex numbers (different from zero) are multiplied and divided, their arguments are correspondingly added and subtracted; in particular, when we multiply by a real positive number, the argument (more precisely, the set of arguments) of a complex number does not undergo a change.*

For every complex number z we have

$$\exp z=\exp\mathcal{R}z\cdot\exp i\mathcal{I}z=\exp\mathcal{R}z\cdot(\cos\mathcal{I}z+i\sin\mathcal{I}z),$$

and consequently, since (cf. (7.6)) $\exp\mathcal{R}z>0$, we have by theorem 9.6:

$$(9.8)\qquad |\exp z|=\exp\mathcal{R}z,\qquad \arg\exp z=\mathcal{I}z+2k\pi,$$

where k is an arbitrary integer.

In the second of the above formulae the equality sign joins two symbols which do not denote one number, but a certain set of numbers (in this case a set of numbers congruent mod 2π). The equality sign in such instances will indicate that each of the values of one of the sides is a certain value of the other side and conversely. However, one must be on guard against ascribing to the sign "=", used in this sense, properties which it has when used in the ordinary sense. For example, in the equality considered, one cannot, obviously, transpose the term $2k\pi$ from the right side to the left.

From (9.8) it follows immediately, as a supplement to formula 7.3, that

(9.9) $\exp z=1$ *if and only if z is of the form $2k\pi i$, where k is an arbitrary integer.*

Since the relation $\exp z_1=\exp z_2$ is equivalent to the relation $\exp(z_1-z_2)=1$, therefore:

(9.10) $\exp z_1=\exp z_2$ *if and only if z_1,z_2 differ by a multiple of $2\pi i$; consequently, the exponential function is periodic with period $2\pi i$.*

In the preceding section the roots of the real functions $\cos z$ and $\sin z$ were investigated. We shall prove that in the complex domain the trigonometric functions have no roots other than these, and hence that

(9.11) *The only roots of the functions $\cos z$ and $\sin z$ are real points of the form $\pi/2+k\pi$ for $\cos z$, and of the form $k\pi$ for $\sin z$, where k is an arbitrary integer.*

Proof. By the Euler formulae (8.3), the relation $\sin z=0$ is equivalent to the relation $\exp iz-\exp(-iz)=0$, and hence to the relation $\exp 2iz=0$, which in view of theorem 9.9 is satisfied if and only if $2iz$ has the form $2k\pi i$, and hence when z is of the form $k\pi$. From the equation $\cos z=\sin(z+\pi/2)$ (formula (8.7)) a corresponding property follows for $\cos z$.

Theorem 9.11 can be supplemented by the following theorem which will find application in many further considerations:

(9.12) *Let ε be an arbitrary positive number and let E_ε denote the complement of the sum of the circles $\mathrm{K}(n\pi;\varepsilon)$, where $n=0,\pm1,\pm2,\ldots$, with respect to the open plane. Then there exists a positive constant C_ε depending only on ε and such that*

$$(9.13)\qquad |\sin z|\geqslant C_\varepsilon \quad\textit{and}\quad |\tan z|\geqslant C_\varepsilon \quad\textit{for every}\ z\,\epsilon\, E_\varepsilon.$$

Proof. Let $K=\mathrm{K}(0;\varepsilon)$ and let H denote the strip defined by the inequality $-\pi/2\leqslant\mathcal{R}z\leqslant\pi/2$. In view of the relations $\sin(z+\pi)=-\sin z$, $\tan(z+\pi)=\tan z$, it is sufficient to show that there exists a constant C_ε such that the inequalities (9.13) are satisfied for all $z\,\epsilon\, H-K$.

Let us denote by H_1 the set of points $z\,\epsilon\, H-K$ for which $|\mathcal{I}z|\leqslant1$. Since this set is bounded, closed, and neither the function $\sin z$ nor $\tan z$ vanishes anywhere on it, there exists a number $C>0$ such that $|\sin z|\geqslant C$ and $|\tan z|\geqslant C$ for $z\,\epsilon\, H_1$.

On the other hand, by the Euler formulae (8.3) and (8.13) we have, taking $z=x+iy$,

$$(9.14)\qquad |\sin z|=\frac{1}{2}|e^{iz}-e^{-iz}|\geqslant\frac{1}{2}\big||e^{iz}|-|e^{-iz}|\big|=\frac{1}{2}|e^{-y}-e^{y}|,$$

and similarly

$$(9.15)\qquad |\tan z|\geqslant\frac{|e^{y}-e^{-y}|}{e^{y}+e^{-y}}.$$

We easily see that for $|y|\geqslant1$ the expressions on the right in (9.14) and (9.15) are respectively not less than $(e-e^{-1})/2$ and $(e-e^{-1})/(e+e^{-1})$. Hence, denoting by C_ε the least among the last two numbers and C, we verify immediately that the relations (9.13) are satisfied for every $z\,\epsilon\, H-K$, and hence also for every $z\,\epsilon\, E_\varepsilon$.

§ 10. Logarithm. Every number z satisfying the equation $\exp z=\zeta$ is called a *logarithm* of ζ and we denote it by $\log\zeta$. From theorem 9.10 it follows that if a certain number z_0 is $\log\zeta$, then every number $z_0+2k\pi i$, where k is an arbitrary integer, is also $\log\zeta$, and conversely: every number which is $\log\zeta$ has the form $z_0+2k\pi i$. Among these logarithms, therefore, exactly one has an imaginary part y such that $-\pi<y\leqslant\pi$; this logarithm is called the *principal logarithm* of the number ζ and we denote it by $\operatorname{Log}\zeta$.

From theorem 7.6 it follows that every real positive number has exactly one real logarithm; this logarithm is at the same time its principal logarithm. In view of theorem 7.5 the number 0 does not have a logarithm. Now, if $\zeta\neq 0$, then the relation $\exp z=\zeta$ is, in virtue of (9.8), equivalent to a system of two real relations

$$|\zeta|=\exp\mathcal{R}z, \qquad \arg\zeta=\mathcal{I}z+2k\pi,$$

where k is an arbitrary integer.

The first of these relations defines $\mathcal{R}z=\operatorname{Log}|\zeta|$ uniquely, and from the second it follows that $\mathcal{I}z$ is one of the values of $\arg\zeta$. Summarizing:

(10.1) *Every complex number* $\zeta\neq 0$ *has an infinite number of logarithms. The general form of* $\log\zeta$ *is given by the formula* $\log\zeta=\operatorname{Log}|\zeta|+i\arg\zeta$. *In particular:*

$$\operatorname{Log}\zeta=\operatorname{Log}|\zeta|+i\operatorname{Arg}\zeta.$$

By means of the logarithm we define the *power* of an arbitrary base $a\neq 0$ with an arbitrary exponent γ. Namely, we take as the definition of the power a^γ of the base $a\neq 0$ the equation

$$a^\gamma=e^{\gamma\log a}. \tag{10.2}$$

Since $\log a$ has infinitely many values, a^γ also has, in general, an infinite number of values (that one which corresponds to the principal value of $\log a$ will be called the *principal value* of the power). In fact, denoting by θ anyone of the values of $\arg a$, we can, in virtue of theorem 10.1, write formula (10.2) in the form

$$\begin{aligned} a^\gamma &= \exp[\gamma(\operatorname{Log}|a|+\theta i+2k\pi i)] \\ &= \exp[\gamma(\operatorname{Log}|a|+\theta i)]\cdot\exp 2k\gamma\pi i, \end{aligned} \tag{10.3}$$

where k is an arbitrary integer. However, only the second factor on the right side of the equation (10.3) can be multi-valued, and we notice at once that even this factor is single-valued and equal to 1 if γ is a real integer; in this case, the definition of a^γ by formula

(10.2) coincides with the usual arithmetical definition of a power. In general, in view of theorem 9.10, if γ is a real number, to the two values k_1, k_2 of the variable integer k there corresponds the same value of a^γ in equation (10.3) if and only if the number $\gamma(k_1-k_2)$ is an integer. Hence, if γ is an irrational number, then to different values of k there always correspond different values of the power a^γ, which therefore has infinitely many values. On the other hand, if γ is a rational number, then, taking $\gamma=p/q$, where p and q are relatively prime integers, we notice at once that $\gamma(k_1-k_2)$ is an integer if and only if k_1-k_2 is a multiple of the number q; therefore a^γ has exactly q different values, which we obtain *e. g.* for $k=0,1,\ldots,q-1$.

From equation (10.3) we have $|a^\gamma|=\exp(\gamma\operatorname{Log}|a|)=|a|^\gamma$ for every real value of γ, understanding $|a|^\gamma$ in the sense adopted in real analysis; the expression a^γ (for real values of γ) therefore has, in spite of its multi-valuedness, a uniquely defined absolute value.

EXERCISES. 1. The equation $\cos z=a$ has an infinite number of roots for each finite value of a. For what values of a do every two roots of this equation differ by a multiple of 2π? by a multiple of π? Write (in terms of the logarithm) the formula for the solution of the equation $\cos z\neq a$. An analogous problem for the equation $\sin z=a$.

The values of z satisfying respectively the equations $\cos z=a$ and $\sin z=a$ are generally denoted by $\operatorname{arc\,cos} a$ and $\operatorname{arc\,sin} a$.

2. The equations $\tan z=a$ and $\cot z=a$ have infinitely many roots for each value of $a\neq\pm i$, finite or infinite. For $a=\pm i$ these equations have no solutions.

Write (in terms of the logarithm) the formula for the solution of these equations (take into consideration, in particular, the case $a=\infty$).

The values of z satisfying respectively the equations $\tan z=a$ and $\cot z=a$ are generally denoted by $\operatorname{arc\,tan} a$ and $\operatorname{arc\,cot} a$.

3. Show that the roots of the equation

$$\binom{n}{1}x+\binom{n}{3}x^3+\ldots=0$$

(in which the last term is nx^{n-1} or x^n, according as the number n is even or odd) are given by the formula: $x=\pm i\tan(k\pi/n)$, where $k=0,1,\ldots,(n-2)/2$ for even n, and $k=0,1,\ldots,(n-1)/2$ for odd n.

[*Hint.* Write the given equation in the form $(1+x)^n=(1-x)^n$.]

4. Determine all the roots of the equations:

(a) $$\binom{n}{1}x-\binom{n}{3}x^3+\binom{n}{5}x^5-\ldots=0,$$

(b) $$1-\binom{n}{2}x^2+\binom{n}{4}x^4-\ldots=0,$$

(c) $$1+\binom{n}{1}x-\binom{n}{2}x^2-\binom{n}{3}x^3+\ldots(-1)^{n(n-1)/2}x^n=0.$$

For equations (a) and (b) distinguish the cases of n even and odd.

5. If $z_0, z_1, \ldots, z_{n-1}$ is the set of n-th roots of the number 1, then for every integer k the sum $z_0^k+z_1^k+\ldots+z_{n-1}^k$ is equal to n when k is a multiple of n, and equal to 0 in the contrary case.

6. Find all the values of the powers 1^i, i^i, $i^{\sqrt{2}}$ (notice that the first two have only real values, and the third — only imaginary values).

7. Show that in order that the power a^{u+vi} (where u and v are real numbers and $a\neq 0$) have only real values, it is necessary and sufficient that $2u$ be an integer and the number $v\operatorname{Log}|a|+u\operatorname{Arg}a$ be a multiple of π.

8. Investigate the distribution of the points $z\neq 0$ at which the power z assumes real values only (these points lie on a denumerable set of lines parallel to the y-axis and on each of these lines there are infinitely many points having this property).

9. Show that in order that all values of the power a^c have the same absolute value it is necessary and sufficient that $\mathcal{I}c=0$. If $\mathcal{I}c\neq 0$, then the power a^c has infinitely many distinct absolute values whose lower bound is 0 and upper bound $+\infty$.

In order that all the values of the power a^c lie on one half-line with origin at the point 0, it is necessary and sufficient that $\mathcal{R}c$ be an integer; and in order that they lie on a finite number of such half-lines, it is necessary and sufficient that $\mathcal{R}c$ be a rational number.

10. Prove that the points z at which all the terms of the series $\sum_n(1-z^{2^n})$ vanish, beginning from a certain place, form a denumerable set, everywhere dense on the circumference $C(0;1)$. For points other than these the series is divergent.

11. If a and b are two incommensurable numbers, then for each number $\varepsilon>0$ there exists a linear combination $ma+nb$ with integral coefficients m,n, not vanishing simultaneously, such that $|ma+nb|<\varepsilon$. [Proof. To each integer k we can assign a number $x(k)$ such that $0\leqslant x(k)\leqslant|b|$ and $x(k)\equiv ka \pmod b$. There certainly exist two distinct integers k_1, k_2, such that $|x(k_2)-x(k_1)|\leqslant\varepsilon$. We have $x(k_1)=k_1a+p_1b$ and $x(k_2)=k_2a+p_2b$, where p_1 and p_2 are integers. Then $|(k_2-k_1)a+(p_2-p_1)b|\leqslant\varepsilon$.]

In view of the preceding result show that: (a) If c is an arbitrary real irrational number, then the values of the power a^c (where $a\neq 0$) form an everywhere dense set on the circumference with centre 0 and radius equal to the absolute value $|a|^c$; (b) If a continuous function $F(x)$ of a real variable x has two incommensurable periods, then it reduces to a constant.

Many interesting examples in connection with the topics of §§ 7-10 can be found in the books: G. H. Hardy, *Pure Mathematics* and E. W. Hobson, *Plane Trigonometry*.

§ 11. Branches of the logarithm, argument, and power. If $F(z)$ is a function defined on a set E, then by a *single-valued branch of the logarithm* of this function on E we shall mean any function

$L(z)$, finite and continuous on the set E, which satisfies the equation $\exp L(z)=F(z)$ on this set; the last condition states that the values of the function $L(z)$ at the points $z\in E$ are values of $\log F(z)$ at these points.

We define the branches of the argument, of the power, and of other functions defined by multi-valued expressions, in a manner analogous to that used in defining a branch of the logarithm. Therefore, if $F(z)$ is defined and vanishes nowhere on the set E, then by a *single-valued branch of the argument* of the function $F(z)$ on E we shall mean any continuous function on E whose values at the points $z\in E$ are values of $\arg F(z)$ at these points. Similarly, by a *single-valued branch of the power* $[F(z)]^\gamma$ we shall mean any continuous function whose values at the points z of the set E are values of $[F(z)]^\gamma$.

In the subsequent chapters (Chapter VI) we shall also consider multi-valued branches of functions; in this chapter and in the next four, however, we shall limit ourselves to single-valued branches, so that without any fear of misunderstanding we shall be able to say simply "branch" instead of "single-valued branch".

In virtue of theorem 10.1 it is apparent that if $L(z)$ is a branch of $\log F(z)$ on the set E, then the function $A(z)=\mathcal{I}L(z)$ is a branch of $\arg F(z)$ on E; conversely, if $A(z)$ is a branch of the argument of the function $F(z)$ continuous on the set E and vanishing nowhere on this set, then the function $L(z)=\operatorname{Log}|F(z)|+iA(z)$ is a branch of the logarithm of the function $F(z)$.

The existence of a branch of the argument of a function continuous and non-vanishing on a set, is therefore equivalent to the existence of a branch of the logarithm of this function. Consequently, it follows from theorem 9.3 that the functions $\operatorname{Arg}z$ and $\operatorname{Log}z$ are respectively branches of $\arg z$ and $\log z$ in the region obtained by removing the negative real half-axis from the open plane. More generally:

(11.1) *If G_α denotes the region which is obtained by removing from the open plane the points having the argument α, then the functions*

$$\operatorname{Arg}[z\exp(\pi-\alpha)i]-(\pi-\alpha),\qquad \operatorname{Log}[z\exp(\pi-\alpha)i]-(\pi-\alpha)i$$

are respectively branches of $\arg z$ *and* $\log z$ *in the region* G_α.

Proof. Let us note that when z ranges over the region G_α, then the point $\zeta=z\exp(\pi-\alpha)i$ ranges over the open plane with the exception of the points on the negative real half-axis. By theorem

9.3 the function $\operatorname{Arg}[z\exp(\pi-a)i]-(\pi-a)$ is therefore continuous on G_a. On the other hand, from theorem 9.7 it follows immediately that $\operatorname{Arg}[z\exp(\pi-a)i]-(\pi-a)$ is equal to one of the values of $\arg z$ for every z. Therefore the function considered is a branch of $\arg z$ on the set G_a. The proof for the logarithm proceeds similarly.

In Chapter IV we shall prove the existence of branches of $\log z$ and $\arg z$ in every simply connected region not containing the points 0 and ∞. For the present we shall limit ourselves to several more elementary results. Let us observe, first of all, that if $L_1(z), L_2(z)$ are branches of the logarithm of the same function $F(z)$ on the set E, then the function $[L_1(z)-L_2(z)]/2\pi i$ assumes only integral values on E and therefore (cf. Introduction, theorem 11.1) it simply reduces to a constant when E is a connected set. Consequently:

(11.2) *Two branches of the logarithm of the same function on a connected set can differ at most by a constant, namely, by an integral multiple of $2\pi i$. Similarly, two branches of the argument of the same function on a connected set can differ at most by an integral multiple of 2π.*

We prove further that each branch of $\log z$ on an open set has a derivative everywhere in this set. Precisely:

(11.3) *If the function $L(z)$ is a branch of $\log z$ on an open set G, then at each point $z\epsilon G$ we have $L'(z)=1/z$.*

Proof. Let $z_0\epsilon G$, $z\epsilon G$, $w_0=L(z_0)$, $w=L(z)$. Hence we have $z_0=\exp L(z_0)=\exp w_0$, and $z=\exp L(z)=\exp w$. Therefore,

$$\frac{L(z)-L(z_0)}{z-z_0}=\frac{w-w_0}{\exp w-\exp w_0}. \tag{11.4}$$

In virtue of the formula for the derivative of the exponential function (theorem 7.8), we see that when $z\to z_0$ (and hence $w\to w_0$) the right side of the equation (11.4) tends to $1/\exp w_0=1/z_0$. Consequently, $L'(z_0)=1/z_0$ at each point $z_0\epsilon G$.

EXERCISES. 1. If K is a circle not containing the point ∞, then for each point $a\neq\infty$ not belonging to the circle K there exists in K a branch of $\arg(z-a)$ and a branch of $\log(z-a)$.

2. If $W(z)$ is a continuous function on a set E, and if all the values which this function assumes on this set belong to a circle K not containing the points 0 and ∞, then there exists on E a branch of $\arg W(z)$ and a branch of $\log W(z)$.

3. If C is a simple arc (see Introduction, § 12), then for every point a not lying on C there exists on C a branch of $\arg(z-a)$ and a branch of $\log(z-a)$.

In general, for every function $W(z)$, continuous and vanishing nowhere on C, there exists a branch of $\arg W(z)$ and a branch of $\log W(z)$ on C.

4. If for a function $F(z)$, vanishing nowhere on a certain connected set E, there exist on E two branches of the power $[F(z)]^a$ (where a is an arbitrary complex number), then the ratio of these two branches is constant on E.

5. On no circumference with centre $a\neq\infty$ (and hence, of course, in no annular neighbourhood of this point) does there exist a branch of $\log(z-a)$, of $\arg(z-a)$, or of $(z-a)^a$, where a is a real number not an integer.

6. A branch of $\sqrt{(z-a)(z-b)}$, where $a\neq b$, exists on the circumference of every circle which contains both points a and b either inside or outside, but does not exist on the circumference of any circle which contains one of the points a,b inside and the other outside.

7. If two branches of the function $\operatorname{arc\,tan} z$ exist on a connected set E not containing the points $\pm i$ (see § 10, exercise 2), then these branches differ at most by a constant multiple of π.

8. If two branches of the function $\operatorname{arc\,cos} z$ exist on a connected set E not containing the points ± 1 (see § 10, exercise 1), then either the sum or the difference of these two branches is a constant multiple of 2π on E.

9. No branch of $\operatorname{arc\,tan} z$ exists in any annular neighbourhood of the point i or $-i$; similarly, no branch of $\operatorname{arc\,cos} z$ exists in any annular neighbourhood of the point 1 or -1.

10. If $\varphi_\nu(z)$, where $|\nu|>R$, denotes a branch of the power $\left(1+\frac{z}{\nu}\right)^\nu$ in the circle $K(0;R)$, assuming the value 1 for $z=0$, then $\varphi_\nu(z)$ tends uniformly in $K(0;R)$ to $\exp z$, when $\nu\to\infty$.

§ 12. Angle between half-lines. We shall give an application of the above considerations to the definition of an angle between two half-lines.

The set of all points of the form

$$z=z_0+pt, \tag{12.1}$$

where t assumes non-negative values, is called a *half-line* with origin at z_0 and direction number $p\neq 0$.

Taking $z_0=x_0+iy_0$, $p=m+in$, where x_0,y_0,m,n are real numbers, we can write equation (12.1) in the form of a system of two real equations $x=x_0+mt$, $y=y_0+nt$, calling the pair of numbers (m,n) *direction numbers* of the half-line. However, in the sequel we shall rather take advantage of the complex form of the equation of the half-line, characterizing its direction by one complex number which takes the place of a pair of real numbers.

If in equation (12.1) we replace the direction number p by λp, where λ is an arbitrary real positive number, then we obviously obtain the same half-line, since if t assumes all real non-negative values, λt also assumes all non-negative values. Conversely:

(12.2) *If $z=z_1+p_1t$ and $z=z_2+p_2u$, where t and u assume non-negative values, are equations of the same half-line L, then $z_2=z_1$ and the ratio p_2/p_1 is a positive number; in other words, the origin of the half-line is defined uniquely, and the direction number to within an arbitrary real positive factor.*

Proof. Substituting $z_1=x_1+iy_1$, $z_2=x_2+iy_2$, $p_1=m_1+in_1$, $p_2=m_2+in_2$, we can write the two given equations of the half-line L in the form of two systems of real equations, $x=x_1+m_1t$, $y=y_1+n_1t$ and $x=x_2+m_2u$, $y=y_2+n_2u$. It is immediately apparent that x_1 is a bound (lower or upper depending on the sign of m_1) of the set of abscissae of the points of the half-line L, and the only finite bound of this set. On the other hand, x_2 is obviously also a bound of this set; therefore $x_2=x_1$. Similarly we prove that $y_2=y_1$, whence finally $z_2=z_1$.

Returning to the equations of the half-line L in the complex forms $z=z_1+p_1t$ and $z=z_2+p_2u$, let us denote by z_0 any point of this half-line different from the point $z_1=z_2$. Let t_0 and u_0 be the values of the parameters t and u, corresponding to the point z_0 in these equations. We therefore have

$$p_1=\frac{z_0-z_1}{t_0} \qquad \text{and} \qquad p_2=\frac{z_0-z_2}{u_0}=\frac{z_0-z_1}{u_0},$$

whence, since the numbers t_0 and u_0 are real and positive, we obtain $p_2/p_1=t_0/u_0>0$, q. e. d.

If p is an arbitrary direction number of a half-line, then $p/|p|$ is also a direction number of this half-line and its absolute value is 1. As is evident at once, among the direction numbers of a half-line there exists only one whose absolute value is 1. We shall call it the *normalized direction number* or briefly the *direction* of the given half-line.

The argument of the direction of a half-line, being in virtue of theorems 12.2 and 9.7 simultaneously the argument of all its direction numbers, is called briefly the *argument* of the given half-line. Consequently, by theorem 9.6,

(12.4) *The argument θ of a half-line is determined to within an additive constant of the form $2k\pi$; if $p=m+ni$ is a direction number of a half-line, then its argument θ is defined by the formulae*

$$\cos\theta=\frac{m}{\sqrt{m^2+n^2}}, \qquad \sin\theta=\frac{n}{\sqrt{m^2+n^2}}.$$

By the *angle* $\sphericalangle(L_1,L_2)$ between two half-lines L_1,L_2, we mean the difference of their arguments, *i. e.* denoting by p_1 and p_2 respectively direction numbers of the half-lines L_1 and L_2, we take $\sphericalangle(L_1,L_2)=\arg p_2-\arg p_1$. The angle between two half-lines is therefore defined uniquely to within an arbitrary additive constant of the form $2k\pi$.

The positive real half-axis obviously has the argument zero (in general $2k\pi$). On the other hand, if an arbitrary finite point $z\neq 0$ is given, the half-line with origin at the point 0 and passing through the point z has its argument equal to $\arg z$. Hence *the argument of an arbitrary point $z\neq 0$ in the plane is equal to the angle which a half-line with origin at the point 0 and passing through the point z makes with the positive real half-axis.* Furthermore, *the absolute value $|z|$ of an arbitrary point is obviously equal to its distance from the point 0.* In this way we obtain the geometric interpretation of the argument and of the absolute value of an arbitrary finite complex number.

From theorem 12.4 and formulae (8.4) it follows that

(12.5) *The angle $a=\sphericalangle(L_1,L_2)$ between two half-lines L_1,L_2, whose direction numbers are $p_1=m_1+n_1i$, $p_2=m_2+n_2i$, is defined by the equations*

$$\cos a=\frac{m_1m_2+n_1n_2}{(m_1^2+n_1^2)^{1/2}(m_2^2+n_2^2)^{1/2}},$$

$$\sin a=\frac{m_1n_2-m_2n_1}{(m_1^2+n_1^2)^{1/2}(m_2^2+n_2^2)^{1/2}},$$

whence:

$$\tan a=\frac{m_1n_2-m_2n_1}{m_1m_2+n_1n_2}.$$

§ 13. Tangent to a curve.

Let

(13.1) $$z=z(t), \qquad \text{where} \quad a\leqslant t\leqslant b,$$

be a curve in the plane (cf. Introduction, § 12). Let $t_0<b$ be a point of the interval $[a,b]$, in no right-hand neighbourhood of which is the function $z(t)$ constant, and let Δt be a positive number such that $z(t_0+\Delta t)\neq z(t_0)$. The half-line with origin at the point $z(t_0)$, and passing through the point $z(t_0+\Delta t)$, has a direction number $z(t_0+\Delta t)-z(t_0)$, and therefore its direction (*i. e.* the normalized direction number, cf. § 12, p. 78) is equal to

$$\frac{z(t_0+\Delta t)-z(t_0)}{|z(t_0+\Delta t)-z(t_0)|}.$$

If the limit of this direction,

$$l = \lim_{\Delta t \to 0+} \frac{z(t_0+\Delta t)-z(t_0)}{|z(t_0+\Delta t)-z(t_0)|}, \tag{13.2}$$

exists (where we take into consideration only those values of Δt for which the numerator and denominator of the ratio considered do not vanish), then the number l is called the *right-hand direction* of the curve (13.1) at the point t_0, and the half-line with origin $z(t_0)$ and direction l is called the *right-hand half-tangent* of the curve at this point.

We define a *left-hand direction* and a *left-hand half-tangent* of a curve at a given point similarly.

If a curve has a right-hand as well as a left-hand direction at t_0, and if these directions differ only in sign, then both half-tangents form one straight line, which we call the *tangent* to the curve at the point t_0. Instead of "right-hand direction of the curve" we shall then simply say *direction.*

If the function $z(t)$ has a derivative $z'(t_0)\neq 0$ at the point t_0, then the curve $z=z(t)$ has a tangent at this point, and its direction there is equal to $z'(t_0)/|z'(t_0)|$. In fact, dividing the numerator and denominator of the expression following the limit sign in (13.2) by $\Delta t>0$, we obtain $l=z'(t_0)/|z'(t_0)|$; and if Δt tends to 0 through negative values, then the same expression tends to $-l$.

If $z(t)=x(t)+iy(t)$, then the existence of the derivative $z'(t)$ is equivalent to the existence of both derivatives $x'(t), y'(t)$, and the condition $z'(t)\neq 0$ — to the condition $[x'(t)]^2+[y'(t)]^2>0$. The characterization of the tangent by the normalized direction number

$$\frac{z'(t)}{|z'(t)|} = \frac{x'(t)+iy'(t)}{\{[x'(t)]^2+[y'(t)]^2\}^{1/2}}$$

is equivalent to defining its direction by the pair of numbers

$$\frac{x'(t)}{\{[x'(t)]^2+[y'(t)]^2\}^{1/2}}, \qquad \frac{y'(t)}{\{[x'(t)]^2+[y'(t)]^2\}^{1/2}},$$

as is usually done in differential geometry.

§ 14. Homographic transformations. A one-to-one transformation of the (closed) plane, of the form

$$\zeta = az+b, \qquad \text{where} \quad a\neq 0, \tag{14.1}$$

is called a *linear transformation* or a *similarity.* The numbers $|a|$ and $\arg a$ are called the *coefficient of similarity* and the *angle of rotation* of this transformation, respectively.

It is immediately evident that under a linear transformation a straight line, half-line, and segment, go into a straight line, half-line, and segment, respectively. The distance between two finite points is multiplied by the coefficient of similarity; therefore a circumference and a circle (Introduction, § 8) are transformed into a circumference and circle, respectively.

Under the transformation (14.1) the half-line $z=z_0+pt$ with direction number p goes into the half-line $\zeta=(az_0+b)+apt$ with the direction number ap. Since $\arg ap=\arg p+\arg a$, we see that under a linear transformation the argument of a half-line increases by the angle of rotation of the transformation, whence it follows that *under a linear transformation the angle between two half-lines remains unchanged.*

If the coefficient of similarity of the transformation is equal to 1, the transformation is called a *motion.* A motion whose angle of rotation is equal to zero is called a *translation.* The transformation (14.1) is therefore a translation if and only if $a=1$. The general form of a translation is consequently

$$\zeta = z+b; \tag{14.2}$$

the term b is here called the *translation vector.*

A finite point which is invariant under a linear transformation will be called a *centre* of this transformation. For the identity transformation every finite point is obviously a centre. A translation, if it is not an identity transformation, has no centre at all, while every linear transformation (14.1) which is not a translation has exactly one centre, namely, $z_0=b/(1-a)$. Expressing b in terms of z_0 and a, and substituting $k=|a|$ and $\theta=\arg a$ in equation (14.1), we can write every linear transformation which has a centre and is not an identity in the form

$$\zeta - z_0 = ke^{i\theta}(z-z_0), \tag{14.3}$$

where z_0, k, and θ denote the centre, the coefficient of similarity, and the angle of rotation of the transformation, respectively. Since the identity transformation is also included (for $k=1$, $\theta=0$) in formula (14.3), this formula represents the general form of a linear transformation having a centre.

A motion which has a centre (*i. e.* is either an identity transformation or is not a translation) is termed a *rotation.* A linear transformation which has a centre and whose angle of rotation is equal to zero, is called a *dilation*. Particularizing formula (14.3) we obtain as the general form of a rotation, and of a dilation, with centre at z_0, the respective formulae:

(14.4) $\zeta - z_0 = e^{i\theta}(z - z_0)$, where θ is the angle of rotation,

(14.5) $\zeta - z_0 = k(z - z_0)$, where k is the coefficient of similarity.

If $h(z) = az + b$, where $a = re^{i\theta} \neq 0$, then taking

$$h_1(z) = e^{i\theta}z, \qquad h_2(z) = rz, \qquad h_3(z) = z + b,$$

we have $h = h_3h_2h_1$. Consequently:

(14.6) *Every linear transformation is the product of a rotation with centre* 0, *a dilation with centre* 0, *and a translation.*

Linear transformations — as is evident immediately — form a group (cf. Introduction, § 7). This group constitutes a subgroup of a more general class of transformations called homographic transformations.

Every transformation of the plane of the form

$$\zeta = \frac{az + b}{cz + d}, \tag{14.7}$$

where a, b, c, and d are arbitrary finite complex numbers such that $ad - bc \neq 0$, is called a *homographic transformation*. In this transformation we associate the point $z = \infty$ with the point $\zeta = a/c$.

The expression $\Delta = ad - bc$ is called the *determinant* of the transformation (14.7). The assumption that $\Delta \neq 0$ is entirely natural; for were the determinant Δ to vanish, we should have $a/c = b/d$, and the right side of the equation (14.7) would then reduce to a constant (except for the point $z = -d/c$, at which it would assume the indeterminate form 0/0).

Every homographic transformation is uniquely invertible and the inverse transformation of a homographic transformation is also homographic. This is verified immediately by expressing z in terms of ζ from equation (14.7); we obtain $z = (-d\zeta + b)/(c\zeta - a)$.

In a similar manner we easily verify that the product of two homographic transformations is a homographic transformation. In fact, if $h_1(z) = (a_1z + b_1)/(c_1z + d_1)$, and $h_2(z) = (a_2z + b_2)/(c_2z + d_2)$, where $a_1d_1 - b_1c_1 \neq 0$ and $a_2d_2 - b_2c_2 \neq 0$, then we find, for $h = h_2h_1$,

$$h(z) = \frac{a_2(a_1z + b_1) + b_2(c_1z + d_1)}{c_2(a_1z + b_1) + d_2(c_1z + d_1)} = \frac{(a_2a_1 + b_2c_1)z + (a_2b_1 + b_2d_1)}{(c_2a_1 + d_2c_1)z + (c_2b_1 + d_2d_1)}.$$

The determinant of this transformation is equal to

$$(a_1d_1 - b_1c_1)(a_2d_2 - b_2c_2) \neq 0.$$

The homographic transformations therefore form a group of transformations.

As in the case of a linear transformation (cf. theorem 14.6), every homographic transformation can be represented as the product of several more special transformations.

The homographic transformation given by the formula

$$\zeta - a = \frac{r^2}{z - a}$$

will be called an *inversion with respect to the circumference* $C(a;r)$ (where $a \neq \infty$ and $r \neq \infty$).

Under this transformation the circumference $C(a;r)$ is transformed into itself (the point $z = a + r\exp i\theta$ is transformed into the point $\zeta = a + r\exp(-i\theta)$); the region interior to the circumference, *i. e.* the circle $K(a;r)$, and the exterior region, *i. e.* $C\overline{K}(a;r)$, are transformed into each other; in particular, the points a and ∞ are transformed into each other.

We shall also call an inversion with respect to the circumference $C(a;r)$ an *inversion with centre a and radius r*, and, in particular, an inversion with respect to $C(0;1)$ — simply an *inversion*. The inversion, expressed by the formula $\zeta = 1/z$, is used frequently when we are concerned with a transformation of the plane into itself, such that the point ∞ is transformed into the point 0 and conversely. By means of this transformation every point z of the circumference $C(0;1)$ is transformed into its conjugate point $\bar{z}$, *i. e.* into one symmetric to it with respect to the real axis.

(14.8) *Every homographic transformation is the product of a finite number of rotations, dilations with centre* 0, *translations, and inversions.*

Proof. Let $h(z) = (az + b)/(cz + d)$ be a homographic function. We may assume that $c \neq 0$, since for $c = 0$ the function $h(z)$ is linear and the theorem reduces to theorem 14.6. Hence we can write

$$h(z) = \frac{a}{c} + \frac{bc - ad}{c^2} \cdot \frac{1}{z + d/c};$$

therefore, taking

$$h_1(z) = z + \frac{d}{c}, \qquad h_2(z) = \frac{1}{z}, \qquad h_3(z) = \frac{bc - ad}{c^2}z + \frac{a}{c},$$

we have $h=h_3h_2h_1$, where h_2 is an inversion, and h_1 and h_3 linear transformations. Since, by theorem 14.6, every linear transformation is the product of a rotation, a dilation with centre 0, and a translation, theorem 14.8 is proved.

As we have seen, under a linear transformation a circumference is always transformed into a circumference, and a straight line into a straight line. However, under homographic transformations a circumference may be transformed into a straight line and conversely. In order to simplify formulations we shall call a straight line with the point ∞ added to it, a *closed straight line* or an *improper circumference*. In this way we can express the fundamental geometrical property of homographic transformations as follows:

(14.9) *A homographic transformation always transforms a circumference into a circumference.*

Proof. The equation of a circumference can be written in Cartesian coordinates in the form

(14.10) $$A(x^2+y^2)+Bx+Cy+D=0,$$

where A, B, C, and D are real numbers, $|A|+|B|+|C|>0$, and $|B|+|C|+|D|>0$. In view of theorem 14.8, it is sufficient to show that the circumference (14.10) is transformed into a circumference under the inversion $\zeta=1/z$. Taking $z=x+iy$, $\zeta=\xi+i\eta$, we have $x=\xi/(\xi^2+\eta^2)$, $y=-\eta/(\xi^2+\eta^2)$, and from equation (14.10), after easy simplifications, we obtain

$$A+B\xi-C\eta+D(\xi^2+\eta^2)=0,$$

i. e. again an equation of a circumference, proper or improper, depending on whether $D\neq 0$, or $D=0$.

Since under every inversion the centre of inversion is transformed into the point ∞, and on the other hand a circumference is improper if and only if it contains the point ∞, it follows in particular from theorem 14.9 that

(14.11) *Under an inversion with an arbitrary centre, a circumference is transformed into a proper circumference or into an improper circumference (a closed line) according as to whether the centre of inversion does not lie on the circumference, or lies on it.*

EXERCISES. 1. Every transformation by a motion in the plane can be written in the form of two real equations, $\xi=a+x\cos\alpha-y\sin\alpha$, $\eta=b+x\sin\alpha+y\cos\alpha$, where α is the angle of rotation of the transformation.

2. Let R be a set of points one of whose coordinates is rational and the other irrational. Show that by means of a suitable rotation this set can be transformed into a set all of whose points have both coordinates irrational (Sierpiński).

3. A transformation of the plane is given by means of two real equations:

(*) $$\xi=a+u_1x+v_1y,\qquad \eta=b+u_2x+v_2y.$$

Show that in order that the transformation (*) preserve the distance between points it is necessary and sufficient that $u_1^2+u_2^2=1$, $v_1^2+v_2^2=1$, $u_1v_1+u_2v_2=0$. Show that if this condition is satisfied, then the transformation is either a motion, or the product of a symmetry with respect to the x-axis and of a motion, *i. e.* it is a transformation of the form $\zeta=g(\bar{z})$, where $\zeta=g(z)$ is a motion.

4. Show that every transformation of the plane, which preserves the distance between points, is either a motion, or the product of a symmetry with respect to the x-axis and of a motion.

[*Hint.* Denoting by (ξ,η), (a,b), $(a+m_1,b+n_1)$, and $(a+m_2,b+n_2)$, the points into which the points (x,y), $(0,0)$, $(1,0)$ and $(0,1)$ are respectively transformed, express the fact that the distances of the point (ξ,η) from the points (a,b), $(a+m_1,b+n_1)$, and $(a+m_2,b+n_2)$, are respectively equal to the distances of the point (x,y) from the points $(0,0)$, $(1,0)$, and $(0,1)$; then make use of exercise 3.]

5. Show that every homographic transformation, not an identity, has always one, and at most two invariant points (an *invariant point* of a transformation is a point which goes into itself under the transformation).

Homographic transformations which have two distinct invariant points are called *loxodromic*, and transformations which have only one invariant point, *parabolic*. The identity transformation is included in each of these classes.

Distinguish the loxodromic and parabolic transformations among the linear transformations. Investigate to which of these two classes of transformations the inversions belong.

6. Show that the general forms of the loxodromic and parabolic transformations, not identities, with finite invariant points are given by the respective formulae:

(a) $$\frac{\zeta-z_1}{\zeta-z_2}=k\frac{z-z_1}{z-z_2}\quad (k\neq 0,\ k\neq 1),$$ (b) $$\frac{1}{\zeta-z_0}=\frac{1}{z-z_0}+l\quad (l\neq 0),$$

where z_1,z_2 (for the loxodromic transformation) and z_0 (for the parabolic transformation) denote the invariant points of the transformation.

Indicate the change in the formulae (a) and (b) in the case of an invariant point at infinity.

Verify that the loxodromic transformations with common invariant points z_1,z_2, form an Abelian group; the parabolic transformations with a common invariant point also form an Abelian group. (A group of transformations is said to be *Abelian* — named after the Norwegian methematician Abel — if the multiplication of transformations in this group is commutative, *i. e.* if $h_1h_2=h_2h_1$ for every pair of transformations h_1 and h_2 belonging to this group).

7. A homographic transformation is termed *real* if it transforms every real point into a real point. The coefficients of a real transformation are real, or become real after a suitable reduction.

If $h(z)=(az+b)/(cz+d)$ is a real transformation and the coefficients a, b, c, and d are real, then by the *signum* of this transformation ($\operatorname{sign} h$) we mean $+1$ or -1, depending on whether the determinant $ad-bc$ is positive or negative. The signum does not change when the coefficients of the transformation are multiplied by the same real factor, positive or negative.

Verify that: 1° if h is a real homographic transformation, then $\operatorname{sign} h^{-1}=\operatorname{sign} h$; 2° if h_1 and h_2 are real homographic transformations, then $\operatorname{sign} h_2h_1=\operatorname{sign} h_2\cdot\operatorname{sign} h_1$; 3° the real homographic transformations of positive signum form a group.

8. Write the loxodromic transformation (a) of exercise 6 in the form $\zeta=(\alpha z+\beta)/(z+\delta)$ and calculate the coefficients α,β, and δ in the following two cases: 1° when z_1,z_2, and k are real numbers ($z_1\neq z_2$, $k\neq 0$); 2° when $z_1=u+iv$ and $z_2=u-iv$ are conjugate imaginary points, and $k=e^{i\theta}$.

Show that in both cases the transformation is real, and in case 2° the signum of the transformation is always positive, while in case 1° it is positive or negative depending on the sign of k.

9. The loxodromic transformation (a), exercise 6, (with an obvious modification of the formula when one of the points z_1,z_2 is at infinity) is called *hyperbolic* if k is a real number, and *elliptic* if k is of the form $e^{i\theta}$. The identity transformation is included in the hyperbolic as well as in the elliptic transformations.

Show that every real homographic transformation (not an identity) has either 1° two distinct real invariant points, or 2° two conjugate imaginary invariant points, or finally 3° one real invariant point — and that depending on these three cases it is 1° hyperbolic, 2° elliptic, or 3° parabolic.

10. The hyperbolic transformations, with a given pair of invariant points, form an Abelian group. Similarly for elliptic transformations (cf. exercise 6).

11. Let z_1 and z_2 be two distinct points in the plane, and let $\mathfrak{H}$ and $\mathfrak{E}$ denote respectively the group of all homographic hyperbolic transformations and the group of all elliptic transformations, with the same invariant points z_1 and z_2.

Show that if w is an arbitrary point of the plane, different from z_1 and z_2, then the set of all points $h(w)$, where h denotes an arbitrary transformation belonging to the group $\mathfrak{H}$, is a circumference passing through the points z_1 and z_2. When the point w varies over the plane we obtain all the circumferences passing through the points z_1 and z_2. These circumferences are called the *trajectories* of the transformations belonging to the group $\mathfrak{H}$.

Formulate analogous definitions and theorems for the group $\mathfrak{E}$ of elliptic transformations.

12. Let $h(z)$ be an arbitrary homographic transformation of the hyperbolic, elliptic, or parabolic type. Every point w_0, different from the invariant points of the transformation h, determines a sequence of points infinite in both directions

$$\ldots,w_{-n},\ldots,w_{-1},w_0,w_1,\ldots,w_n,\ldots,$$

such that $w_{n+1}=h(w_n)$ for $n=\ldots,-1,0,+1,\ldots$ Show that all the points of this sequence lie on a certain trajectory of the transformation h, and: 1° in the hyperbolic case the sequence $\{w_n\}$ tends to one or the other of the invariant points of the transformation h depending on whether $n\to+\infty$, or $n\to-\infty$; 2° in the elliptic case there either exists an integer p such that $w_{n+p}=w_n$ for every n, or the points of the sequence $\{w_n\}$ form an everywhere dense set on the trajectory (cf. § 10, exercise 11); 3° in the parabolic case the sequence $\{w_n\}$ tends to the invariant point of the transformation when $n\to+\infty$ as well as when $n\to-\infty$.

§ 15. Similarity transformations. Referring to the considerations of § 5 and § 6, we shall now give the fundamental geometric interpretation of the derivative in the complex domain. We shall first prove the following theorem:

(15.1) *If a function*

(15.2) $$\zeta=W(z)$$

is continuous and finite in the neighbourhood of a point $z_0\neq\infty$ and has at this point the total differential $(A+Bi)\Delta x+(C+Di)\Delta y$ with determinant $AD-BC\neq 0$ (where A, B, C, D are real numbers), then to each direction l^0 there corresponds a certain direction λ^0 such that when the curve C

(C) $$z=z(t),\quad \text{where}\quad a\leqslant t\leqslant b,\quad z(a)=z_0,$$

with initial point z_0, has the direction l^0 for $t=a$, then the curve Γ

(Γ) $$\zeta=\zeta(t)=W[z(t)],\quad \text{where}\quad a\leqslant t\leqslant b,\quad \zeta(a)=\zeta_0,$$

which has initial point $\zeta_0=W(z_0)$ and is the image of the curve C under the transformation $\zeta=W(z)$, has the direction λ^0 for $t=a$.

This correspondence between the directions l^0 and λ^0 is given by the formula

(15.3) $$\lambda=\mu+\nu i=h[(Am+Cn)+i(Bm+Dn)],$$

where $l=m+ni$ and $\lambda=\mu+\nu i$ are any direction numbers of the directions l^0 and λ^0 respectively, while h is an arbitrary positive factor.

Proof. Let $l^0=m^0+n^0i$ be the direction (*i. e.* the normalized direction number, cf. § 12) of the curve C for $t=a$. Taking $z(t)=x(t)+iy(t)$ and:

$$\Delta z=z(a+\Delta t)-z(a),\qquad \Delta x=x(a+\Delta t)-x(a),\qquad \Delta y=y(a+\Delta t)-y(a),$$
$$\Delta\zeta=\zeta(a+\Delta t)-\zeta(a),$$

we have

(15.4) $$l^0=m^0+n^0i=\lim_{\Delta t\to 0+}\frac{\Delta z}{|\Delta z|}=\lim_{\Delta t\to 0+}\left(\frac{\Delta x}{|\Delta z|}+i\frac{\Delta y}{|\Delta z|}\right).$$

On the other hand,

$$\Delta\zeta=W[z(a+\Delta t)]-W[z(a)]=W(z_0+\Delta z)-W(z_0)$$
$$=(A+Bi)\Delta x+(C+Di)\Delta y+\varepsilon(\Delta z)|\Delta z|,$$

where $\varepsilon(\Delta z)\to 0$ when $\Delta z\to 0$, and hence when $\Delta t\to 0$. Consequently, from (15.4),

(15.5) $$\lim_{\Delta t\to 0+}\frac{\Delta\zeta}{|\Delta z|}=(A+Bi)m^0+(C+Di)n^0.$$

The expression on the right side of (15.5) certainly does not vanish; in fact, if it were zero, then we should have $Am^0+Cn^0=0$ and $Bm^0+Dn^0=0$, and hence, in virtue of the fact that $AD-BC\neq 0$, both numbers m^0, n^0, would vanish, which is contrary to the equality $|l^0|=|m^0+n^0i|=1$. From the existence of the limit (15.5), therefore, follows the existence of the limit of the expression

$$\frac{\Delta\zeta}{|\Delta\zeta|}=\frac{\Delta\zeta/|\Delta z|}{|\Delta\zeta/\Delta z|},$$

which is the direction λ^0 of the curve Γ for $t=a$; in addition, we obtain from (15.5)

(15.6) $$\lambda^0=\frac{(Am^0+Cn^0)+(Bm^0+Dn^0)i}{|(Am^0+Cn^0)+(Bm^0+Dn^0)i|}.$$

Therefore, if $l=m+ni$ is an arbitrary direction number of the direction l^0, and $\lambda=\mu+\nu i$ an arbitrary direction number of the direction λ^0 (*i. e.* if l and λ differ from l^0 and λ^0, respectively, at most by real positive factors), then in virtue of equation (15.6) they satisfy the relation (15.3), where h is a positive factor.

A transformation $\zeta=W(z)$, continuous in the neighbourhood of a point $z_0\neq\infty$, is said to be a *similarity transformation* at this point if

1) the function $W(z)$ is finite at the point z_0 and has at this point a total differential $(A+Bi)\Delta x+(C+Di)\Delta y$ with determinant $AD-BC\neq 0$,

2) in the correspondence of directions which this transformations determines in virtue of theorem 15.1, the angle between directions is preserved; in other words: denoting by λ_1,λ_2 the pair of directions corresponding to a pair of directions l_1,l_2, we have (in agreement with the definition of angle, § 12)

(15.7) $$\arg\lambda_2-\arg\lambda_1=\arg l_2-\arg l_1.$$

Condition 2) can be replaced by the following:

2^{bis}) in the correspondence of directions determined by the transformation $\zeta=W(z)$ at the point z_0, the angle between corresponding directions has a constant value (obviously to within an arbitrary multiple of 2π).

In fact, for the numbers $l_1, l_2, \lambda_1, \lambda_2$ considered above, condition 2^{bis}) means that

$$\arg\lambda_2-\arg l_2=\arg\lambda_1-\arg l_1,$$

and this relation is equivalent to relation (15.7). The constant angle which the corresponding directions under a similarity transformation at the point z_0 form with each other, will be called the *angle of rotation* of the transformation at this point.

(15.8) *In order that a transformation $\zeta=W(z)$, continuous in the neighbourhood of a point $z_0\neq\infty$, be a similarity transformation at this point, it is necessary and sufficient that the function $W(z)$ be finite at this point and have a derivative $W'(z_0)\neq 0$. If this condition is satisfied, then $\arg W'(z_0)$ is the angle of rotation of the given transformation at the point z_0.*

Proof. Let us assume at first that the given transformation is a similarity transformation at z_0. Then the function $W(z)$ has at this point a total differential of the form $(A+Bi)\Delta x+(C+Di)\Delta y$ with determinant $AD-BC\neq 0$. Therefore we have:

(15.9) $$W'_x(z_0)=A+Bi,\qquad W'_y(z_0)=C+Di.$$

Let θ be the angle of rotation of the transformation $\zeta=W(z)$ at the point z_0, and let $\gamma=\tan\theta$. To the direction whose direction number is $l=m+ni$ there corresponds, in virtue of theorem 15.1, the direction whose direction number is

$$\lambda=(Am+Cn)+(Bm+Dn)i;$$

therefore, by theorem 12.5, we have for every pair of real numbers m, n, not vanishing simultaneously,

(15.10) $$\frac{m(Bm+Dn)-n(Am+Cn)}{m(Am+Cn)+n(Bm+Dn)}=\tan\theta=\gamma,$$

and hence

(15.11) $$(B-A\gamma)m^2+[(D-A)-(B+C)\gamma]mn-(C+D\gamma)n^2=0$$

(in the case when $\gamma=\tan\theta=\infty$, we set the denominator of the fraction (15.10) equal to zero, which makes the calculation even simpler).

From (15.11) we obtain $B=A\gamma$, $C=-D\gamma$, $D-A=(B+C)\gamma$, and substituting the values of B and C from the first two equations into the third, we have $(D-A)(1+\gamma^2)=0$, and hence $A=D$; next, from the first two equations we get $B=-C$. Consequently, in virtue of (15.9), we have $W'_x(z_0)=A+Bi=-i(C+Di)=-iW'_y(z_0)$, *i. e.* the function $W(z)$ satisfies the condition of Cauchy-Riemann at the point z_0, and therefore, by theorem 6.4, it has a derivative at this point. Moreover,

$$|W'(z_0)|^2=|W'_x(z_0)|^2=A^2+B^2=AD-BC\neq 0.$$

Conversely, if the function $W(z)$ has at the point z_0 a derivative

$$W'(z_0)=A+Bi\neq 0,$$

then, by theorem 6.4, it has the total differential $W'(z_0)(\Delta x+i\Delta y)=(A+Bi)\Delta x+(-B+Ai)\Delta y$ with determinant $A^2+B^2>0$ at this point. Therefore, to the direction having the direction number $l=m+ni$ there corresponds, in virtue of theorem 15.1, a direction having the direction number

$$\lambda=(A+Bi)m+(-B+Ai)n=(A+Bi)(m+ni)=W'(z_0)l.$$

By theorem 9.7, the angle between the directions l and λ is equal to $\arg\lambda-\arg l=\arg W'(z_0)l-\arg l=\arg W'(z_0)$, and hence has a constant value. The transformation $\zeta=W(z)$ is therefore a similarity transformation at the point z_0, and $\arg W'(z)$ is the angle of rotation of this transformation.

EXERCISES. 1. If a function $W(z)$ has a total differential at the point z_0, then the ratio $[W(z)-W(z_0)]/(z-z_0)$ has a finite limit when z tends to z_0 along any half-line with origin at the point z_0. Assuming that $(A+Bi)\Delta x+(C+Di)\Delta y$ is the total differential of the function $W(z)$ at the point z_0, calculate this limit, as well as its absolute value, for the half-line $z=z_0+(m+ni)t$.

2. If a function $W(z)$ has a total differential at the point z_0, and if the limit of the real part of the ratio $[W(z)-W(z_0)]/(z-z_0)$ exists, is finite, and the same when $z\to z_0$ along three half-lines with origin at the point z_0, of which no two lie on one straight line, then the function $W(z)$ has a derivative at the point z_0.

This is a generalization of the theorem contained in exercise 3, § 6. Formulate and prove an analogous generalization of the theorem contained in exercise 4, § 6.

3. If a function $W(z)$ has a total differential at the point z_0, and if the limit of the ratio $[W(z)-W(z_0)]/(z-z_0)$ is the same when $z\to z_0$ along two half-lines with origin at the point z_0 and not lying on one straight line, then the function $W(z)$ has a derivative at the point z_0 (Menshov).

[*Hint.* It can be assumed that one of the given half-lines has the direction of the positive real axis.]

This theorem can be considered as a generalization of theorem 6.4, since the Cauchy-Riemann conditions denote the existence of the limit of the ratio under consideration along two half-lines parallel to the real and the imaginary axis, respectively.

4. If the function $W(z)$, continuous in the neighbourhood of a point z_0, has a total differential at this point, and if in the correspondence of the directions determined at the point z_0 by the transformation $\zeta=W(z)$ there exist three pairs of corresponding directions (l_1,λ_1), (l_2,λ_2), (l_3,λ_3), such that 1° no two of the directions l_1,l_2,l_3 are identical or opposite (*i. e.* are not the directions of half-lines with origin at the point z_0 and lying on one straight line) and 2° $\sphericalangle(l_i,l_k)=\sphericalangle(\lambda_i,\lambda_k)$ for $i,k=1,2,3$, — then $\zeta=W(z)$ is a similarity transformation at the point z_0, *i. e.* the function $W(z)$ has a derivative $\neq 0$ at this point (Menshov).

5. Let z_1 and z_2 be two distinct points, and let $\mathfrak{H}$ and $\mathfrak{E}$ be respectively the groups of all hyperbolic and of all elliptic transformations having z_1 and z_2 as invariant points. Then the trajectories of the transformations $\mathfrak{E}$ form a family of circumferences orthogonal to the trajectories of the transformations $\mathfrak{H}$; more precisely: through every point of the plane, different from the given invariant points, there pass two circumferences, orthogonal to each other, which are respectively trajectories of the transformations $\mathfrak{H}$ and $\mathfrak{E}$.

[*Hint.* Cf. § 14, exercise 11; first prove the theorem for the case when the invariant points are the points 0 and ∞; proceed to the general case by appealing to the fact that homographic transformations are similarity transformations and therefore preserve orthogonality (perpendicularity) of circumferences.]

§ 16. Regular curves. A curve C given by the equation (cf. § 13, and Introduction, § 12)

$$(16.1)\qquad z=z(t),\quad \text{where}\quad a\leqslant t\leqslant b,$$

will be called *regular*, if the interval $[a,b]$ of its parameter can be divided into a finite number of subintervals such that in each one of them the function $z(t)$ has a continuous derivative $z'(t)$ (at the end-points of these subintervals the function may have only one-sided derivatives; these end-points may therefore be "angular" points of this curve). It is evident that the regularity of a curve is a property which does not depend on the manner in which the curve is represented parametrically (cf. Introduction, § 12).

If $t_0,t_1,\ldots,t_n$ denotes an arbitrary sequence of values of the parameter, such that $a=t_0<t_1<\ldots<t_n=b$, then the upper bound of

the sums $\sum_{k=0}^{n-1}|z(t_{k+1})-z(t_k)|$ is called the *length* of the curve $z=z(t)$, on the interval $[a,b]$ of its parameter. If the function $z(t)=x(t)+iy(t)$ has a continuous derivative in this interval, then, as is well known from real analysis, the length of the curve (16.1) is given by the integral

$$(16.2)\qquad l=\int_a^b\{[x'(t)]^2+[y'(t)]^2\}^{1/2}dt=\int_a^b|z'(t)|\,dt.$$

Since the length of a sum of several curves is equal to the sum of the lengths of these curves, the above formula extends immediately to every regular curve in the interval $[a,b]$.

§ 17. Curvilinear integrals. Let $g(t)$ and $h(t)$ be two finite functions on the interval $[a,b]$. Let us consider sums of the form

$$(17.1)\qquad \sum_{k=0}^{n-1}g(t_k')[h(t_{k+1})-h(t_k)],$$

where $a=t_0<t_1<\ldots<t_n=b$ and $t_k\leqslant t_k'\leqslant t_{k+1}$.

If a finite limit of these sums exists as the characteristic number (cf. Introduction, § 9, p. 25) of the sequence $t_0,t_1,\ldots,t_n$ tends to zero, then we say that the function $g(t)$ is *integrable in* $[a,b]$ *with respect to* $h(t)$; this limit is called the *Stieltjes integral* of the function $g(t)$ *with respect to* $h(t)$ in $[a,b]$ and will be denoted by $\int_a^b g(t)\,dh(t)$.

If the functions $g(t)$ and $h(t)$ are continuous on the interval $[a,b]$ and if, in addition, the function $h(t)$ has a continuous derivative on this interval, then the limit considered certainly exists and is expressed by an ordinary Riemann integral. In this case we have

$$(17.2)\qquad \int_a^b g(t)\,dh(t)=\int_a^b g(t)\,h'(t)\,dt.$$

In order to prove this, it may be assumed that the function $h(t)$ is real, for in the contrary case we could carry out the proof separately for its real and imaginary parts. Let M be the upper bound of $|g(t)|$ on $[a,b]$, and let l denote the characteristic number of the sequence $t_0,t_1,\ldots,t_n$. Let $\omega(l)$ be the upper bound of the numbers $|h'(\beta)-h'(\alpha)|$ when $|\beta-\alpha|\leqslant l$, $a\leqslant\alpha\leqslant\beta\leqslant b$. Denoting for brevity by S the sum (17.1) and applying the theorem of the mean, we have

$$S=\sum_{k=0}^{n-1}g(t_k')\,h'(t_k'')(t_{k+1}-t_k),\quad\text{where}\quad t_k\leqslant t_k'\leqslant t_{k+1},\ t_k\leqslant t_k''\leqslant t_{k+1}.$$

Consequently,

$$(17.3)\qquad \Big|S-\sum_{k=0}^{n-1}g(t_k')\,h'(t_k')\,(t_{k+1}-t_k)\Big|\leqslant M(b-a)\,\omega(l).$$

In virtue of the continuity of the derivative $h'(t)$, the number $\omega(l)$ tends to zero together with l. On the other hand, the subtrahend between the absolute value signs in formula (17.3) is an approximating sum of the Riemann integral appearing on the right side of equation (17.2), and tends to this integral when $l\to 0$. Therefore, the limit of the sum S also exists when $l\to 0$, and the equation (17.2) is satisfied.

If the functions $g(t)$ and $h(t)$ are continuous on the interval $[a,b]$, and if the function $g(t)$ is integrable with respect to $h(t)$ on each of the two intervals $[a,c]$ and $[c,b]$, where $a<c<b$, then it is also integrable on the entire interval $[a,b]$ and its Stieltjes integral on this interval is equal to the sum of its integrals on the two subintervals mentioned. This observation enables us to formulate the result just obtained in a somewhat more general form:

(17.4) *If the functions $g(t)$ and $h(t)$ are continuous on the interval $[a,b]$, and if this interval can be divided into a finite number of subintervals such that on each of them the function $h(t)$ has a continuous derivative, then the function $g(t)$ is integrable with respect to $h(t)$ on the interval $[a,b]$, and its Stieltjes integral on this interval reduces to a Riemann integral by means of the equation* (17.2).

If C is an arbitrary continuous curve

$$(17.5)\qquad z=z(t),\quad\text{where}\quad a\leqslant t\leqslant b,$$

and $F(z)$ is a finite function defined on this curve (more precisely: on the geometric image of this curve, cf. Introduction, § 12), then by the *integrability* and the *curvilinear integral* of the function $F(z)$ *along the curve* C we mean, respectively, the integrability and the Stieltjes integral of the function $F[z(t)]$ with respect to the function $z(t)$ on the interval $[a,b]$. We shall denote the curvilinear integral of the function $F(z)$ along the curve C by $\int_C F(z)\,dz$.

From theorem 17.4 it follows directly that

(17.6) *A function $F(z)$, finite and continuous on a regular curve C, always has a curvilinear integral along this curve, where, if the curve C is given by equation* (17.5), *then*

$$\int_C F(z)\,dz = \int_a^b F[z(t)]\,z'(t)\,dt. \tag{17.7}$$

From this equation it follows, among other things, that the value of the integral of a continuous function along a regular curve does not depend on the parametric representation of the curve (cf. Introduction, § 12). In fact, let $t=\varphi(\tau)$ be a non-essential change of the parameter, by means of which the curve C given by equation (17.5) is transformed into the curve Γ whose equation is

$$z=z[\varphi(\tau)], \quad \text{where} \quad \alpha\leqslant\tau\leqslant\beta.$$

The function $\varphi(\tau)$ is increasing and continuous on the interval $[\alpha,\beta]$, and is differentiable, with the exception of at most a finite number of points. Therefore, for every function $F(z)$, continuous on the curve C, or — what amounts to the same thing — on the curve Γ, we have

$$\int_C F(z)\,dz = \int_a^b F[z(t)]\frac{dz}{dt}\,dt = \int_a^\beta F[z(t)]\frac{dz}{dt}\frac{dt}{d\tau}\,d\tau$$

$$= \int_a^\beta F\{z[\varphi(\tau)]\}\frac{dz}{d\tau}\,d\tau = \int_\Gamma F(z)\,dz.$$

From theorem 17.6 there also result immediately the following relations, in which C, C_1 and C_2 denote regular curves, and $F(z)$ a continuous function on C:

$$\int_{-C} F(z)\,dz = -\int_C F(z)\,dz; \tag{17.8}$$

$$\int_C F(z)\,dz = \int_{C_1} F(z)\,dz + \int_{C_2} F(z)\,dz, \quad \textit{if} \quad C\equiv C_1+C_2; \tag{17.9}$$

$$\left|\int_C F(z)\,dz\right| \leqslant M\int_a^b |z'(t)|\,dt = ML, \tag{17.10}$$

where M denotes the upper bound of $|F(z)|$ on C, and L the length of the curve C.

§ 18. Examples. In this section we shall establish terminology for a few regular curves with which we shall deal more often. Let us recall first of all that curves which differ from each other at most non-essentially, as well as their geometric image, are usually included under the same term. In § 12 of the Introduction we have already defined a segment and a polygonal line as curves in the plane; these are obviously regular curves. In particular, the perimeter of a rectangle (cf. Introduction, § 8, p. 19) may be considered as a regular closed curve without multiple points. Making use of formula (16.2), we see that the *length of a segment is equal to the distance between its end-points.*

By a *circumference with centre* $z_0\neq\infty$ and *radius* $r\neq\infty$ we mean the closed curve without multiple points

$$z=z_0+re^{it}, \quad \text{where} \quad 0\leqslant t\leqslant 2\pi. \tag{18.1}$$

Its length is given by the integral

$$\int_0^{2\pi}\left|\frac{dz}{dt}\right|dt = r\int_0^{2\pi}|e^{it}|\,dt = 2\pi r.$$

If $K=\mathrm{K}(z_0;r)$, then (K) will denote the circumference (18.1). We shall also denote this circumference by $\mathrm{C}(z_0;r)$, *i. e.* in the same way as its geometric image (which is also called a circumference; cf. Introduction, § 8, p. 20). Writing for brevity $C_r=\mathrm{C}(z_0;r)$, we have, in virtue of theorem 17.6, for every integer $k\neq-1$:

$$\int_{C_r}(z-z_0)^k\,dz = ir^{k+1}\int_0^{2\pi}e^{(k+1)it}dt = r^{k+1}\,\frac{e^{2(k+1)\pi i}-1}{k+1} = 0,$$

while

$$\int_{C_r}\frac{dz}{z-z_0} = i\int_0^{2\pi}dt = 2\pi i.$$

Consequently,

$$\int_{C(z_0;r)}(z-z_0)^k\,dz = \begin{cases} 0 & \text{for} \quad k\neq-1, \\ 2\pi i & \text{for} \quad k=-1. \end{cases}$$

As an example, and for future application, we shall calculate a few curvilinear integrals along polygonal lines.

If $F(z)$ is a finite and continuous function defined on a segment $[z_1,z_2]$, then writing the equation of this segment in the form $z=z_1+(z_2-z_1)t$, where $0\leqslant t\leqslant 1$, we find

$$\int_{[z_1,z_2]} F(z)\,dz=(z_2-z_1)\int_0^1 F[(z_2-z_1)t+z_1]\,dt;$$

in particular, for a segment $[z_1,z_2]$ parallel to the axis of abscissae,

(18.3) $$\int_{[z_1,z_2]} F(z)\,dz=\int_{x_1}^{x_2} F(x+iy_1)\,dx, \qquad \text{if } z_1=x_1+iy_1, \quad z_2=x_2+iy_1,$$

and similarly for a segment $[z_1,z_2]$ parallel to the axis of ordinates,

(18.4) $$\int_{[z_1,z_2]} F(z)\,dz=i\int_{y_1}^{y_2} F(x_1+iy)\,dy, \qquad \text{if } z_1=x_1+iy_1, \quad z_2=x_1+iy_2.$$

Therefore, for every function $F(z)$ continuous on the perimeter (I) of the rectangle $I=[a_1,a_2;b_1,b_2]$,

$$\int_{(I)} F(z)\,dz=\int_{a_1}^{a_2}[F(x+ib_1)-F(x+ib_2)]\,dx+i\int_{b_1}^{b_2}[F(a_2+iy)-F(a_1+iy)]\,dy.$$

In particular, therefore, for the square $Q=[-a,a;-a,a]$ we have, in virtue of formula (18.4),

$$\int_{(Q)}\frac{dz}{z}=2ai\int_{-a}^{a}\frac{dx}{x^2+a^2}+2ai\int_{-a}^{a}\frac{dy}{y^2+a^2}=2\pi i.$$

Substituting here $z-z_0$ for z, where z_0 is an arbitrary complex number, we obtain the formula

(18.5) $$\int_{(Q)}\frac{dz}{z-z_0}=2\pi i, \qquad \text{where } Q \text{ is a square with centre } z_0.$$

EXERCISES. 1. Let $W(z)$ be a continuous function in the neighbourhood of a point z_0, having a total differential at this point. Prove that, in this case

$$\lim_{r\to 0}\frac{1}{\pi r^2}\int_{C_r} W(z)\,dz=i[W'_x(z_0)+iW'_y(z_0)],$$

where C_r denotes the circumference $C(z_0;r)$.

Hence, in order that $W(z)$ have a derivative at the point z_0, it is necessary and sufficient that $\lim_{r\to 0}\frac{1}{\pi r^2}\int_{C_r} W(z)\,dz=0$.

2. Let $P(z)$ and $Q(z)$ be two polynomials, and let $Q(z)$ be of higher degree than $P(z)$. Denoting by I_m, where m is an arbitrary positive integer, the square with vertices at the points $(m+\frac{1}{2})(\pm 1\pm i)$, prove that the integrals $\int_{(I_m)}\frac{1}{\sin\pi z}\cdot\frac{P(z)}{Q(z)}dz$ and $\int_{(I_m)}\cot\pi z\frac{P(z)}{Q(z)}dz$ tend to 0 when $m\to\infty$.

[*Hint.* Make use of theorem 9,12; note that $z\left[\frac{P(z)}{Q(z)}+\frac{P(-z)}{Q(-z)}\right]\to 0$ when $z\to\infty$; when integrating along the perimeter (I_m), combine the integrals along opposite sides.]

3. Calculate the integrals

$$\int_0^{2\pi} e^{\cos\theta}\sin(n\theta-\sin\theta)\,d\theta \quad\text{and}\quad \int_0^{2\pi} e^{\cos\theta}\cos(n\theta-\sin\theta)\,d\theta,$$

where n is an integer.

[*Hint.* Notice that these integrals are the real and imaginary parts respectively, of the integral of the function $e^z z^{-(n+1)}$ along the circumference C(0;1).]

4. Let D_r denote the semi-circumference $z=re^{i\theta}$, where $0\leqslant\theta\leqslant\pi$. Assuming that $a_1,a_2,\ldots,a_n$ are real positive numbers and that $A_1+A_2+\ldots+A_n=0$, calculate the limit of the value of the integral

$$\int_{D_r}\frac{A_1\exp ia_1z+A_2\exp ia_2z+\ldots+A_n\exp ia_nz}{z^2}\,dz,$$

when $r\to 0$ and when $r\to\infty$.

5. If $-1<\nu<1$, then the integral $I_n=\int_{C_n}\frac{\exp\nu z\pi i}{\sin\pi z}\cdot\frac{dz}{z-a}$, where a is an arbitrary number and $C_n=C(0;n+\frac{1}{2})$, tends to 0 as n tends to $+\infty$ ranging over integral values.