CHAPTER VI

ANALYTIC FUNCTIONS

§ 1. Introductory remarks. In Chapter I (§§ 9, 10, 11) we have already introduced certain multiple-valued expressions such as *e. g.* $\arg z$, $\log z$, $\sqrt{z}$, and in Chapter IV (§§ 3, 10) we have considered the topological properties of regions in which single-valued continuous (holomorphic) "branches" of these expressions exist. Analogous investigations in the real domain would be trivial. If x assumes only real values, then, taking *e.g.* $f(x)=+\sqrt{x}$ for $x\geqslant 0$ and $f(x)=+i\sqrt{|x|}$ for $x<0$, we obtain a continuous branch of $\sqrt{x}$ on the entire real axis; of course, we could also define still other continuous branches of $\sqrt{x}$ on the real axis, and it would be superfluous to distinguish special linear regions on which such branches can be defined. The essential difference between the complex plane and the real axis is made clear by the fact that on the plane we can, in general, pass from one value of the multiple-valued expression under consideration to another one in a "continuous" manner, starting from some point with one value, and returning to it along a certain closed curve with another value. For example, if we move along the circumference $z=e^{it}$, where $0\leqslant t\leqslant 2\pi$, starting from the point $z=1$ for $t=0$ with the value 1 for $\sqrt{z}$, then, varying the value of this square root in a continuous manner, we return to the point $z=1$ for $t=2\pi$ with the value of $\sqrt{z}$ equal to -1. This example shows that on the circumference under consideration it is not possible to define a single-valued and continuous branch of $\sqrt{z}$, and at the same time it leads in a natural manner to the concept of a multiple-valued function. The definition of such a function obviously must be more then a formal generalization of the notion of a single-valued function which merely assumes that to each value of the "independent variable" there corresponds in general not one but several values of the "dependent variable". In the general definition of a multiple-valued analytic function it is a matter, first of all, of associating with each other the various values of the function in a natural manner, and of establishing methods of passing "continuosly" from one value to another. We owe such a definition to Weierstrass who has based it on the notions of analytic element and analytic continuation. The first sections of this chapter will be devoted to these concepts.

§ 2. Analytic element. A pair $\{F,a\}$, consisting of the point a and the function $F(z)$ meromorphic at this point, will be called an *analytic element;* the point a will be called the *centre* of the element $\{F,a\}$. We shall also say that the fuction $F(z)$ *determines* the analytic element $\{F,a\}$ at the point a.

Let us consider circles K with centre a, to which the function $F(z)$, meromorphic at the point a, can be extended with a preservation of meromorphism, *i. e.* circles in which there exists a meromorphic function identical with $F(z)$ in the neighbourhood of the point a. Among these circles there exists a largest one; we shall call it the *circle of the element* $\{F,a\}$, and its radius the *radius* of this element. The meromorphic function in the circle of the element $\{F,a\}$, identical with F in the neighbourhood of the point a, will be denoted by F_a.

Deviating somewhat from the definition of the circle given in the Introduction, § 8, we shall here regard the closed plane as a circle.

Distinguishing between the function F and the function F_a is sometimes indispensable. For example, when a function $F(z)$ is given in a region G, then at every point $a\epsilon G$ it determines an element $\{F,a\}$ with a certain circle K_a. The function $F(z)$ then coincides with $F_a(z)$ in that component of the open set $K_a\cdot G$ which contains the point a. However, the set $K_a\cdot G$ need not be a region and may have other components in which the functions $F_a(z)$ and $F(z)$ can differ.

Two elements $\{F,a\}$ and $\{G,b\}$ are considered to be *identical*, and we write $\{F,a\}=\{G,b\}$, if $a=b$ and if the functions F and G are identical in a neighbourhood of the point $a=b$; these elements then have a common circle and $F_a=G_a$. In particular, $\{F,a\}=\{F_a,a\}$ always.

The circle of an element $\{F,a\}$ can be the entire plane; as follows from theorem 7.3, Chapter III, this occurs if (and only if) F is a rational function.

Every analytic element $\{F_a, b\}$, where b is an arbitrary point of the circle K of the element $\{F, a\}$, will be called *a direct continuation* of the element $\{F, a\}$. The function $F_a(z)$ is defined in the entire circle K, and hence the element $\{F_a, b\}$ is defined at each point $b \in K$. If $a \neq \infty$, then the radius of the element $\{F_a, b\}$ is at least equal to the positive number $r - |b-a|$, where r denotes the radius of the circle K, since the circle with centre b and radius $r - |b-a|$ is contained in the circle K.

A point z, lying in the interior or on the boundary of the circle of an element $\{F, a\}$, will be called a *point of continuability* of this element if it is contained in the interior of a circle of at least one element which is a direct continuation of the element $\{F, a\}$; in the contrary case, the point z will be called a *point of non-continuability* of the element under consideration. Obviously all the interior points of the circle of an element are points of continuability.

(2.1) *The set of points of non-continuability of an analytic element is closed and, except for the case when the circle of an element is the entire plane, non-empty.*

Proof. Let $\{F, a\}$, $a \neq \infty$, be an analytic element with the circle $K = \mathrm{K}(a; R)$. We may obviously assume that the function $F(z)$ is meromorphic in the entire circle K. It is evident that if some point z of the circumference of the circle K is a point of continuability of the element, then all the points of this circumference situated sufficiently close to the point z are also points of continuability. The set Z of points of non-continuability is therefore a closed set.

In order to show that the set Z is non-empty, let us denote for every point $b \in K$ by K_b the circle of the element $\{F, b\} = \{F_b, b\}$, and by G the sum of all circles K_b for $b \in K$. The functions $F_b(z)$ jointly determine one function $H(z)$ in the entire open set G, *i. e.* if some point z_0 belongs to two circles K_{b_1} and K_{b_2}, where $b_1 \in K$ and $b_2 \in K$, then $F_{b_1}(z_0) = F_{b_2}(z_0)$. Indeed, if $K_{b_1} \cdot K_{b_2} \neq 0$, then, as we see at once (see Fig. 19), also $K \cdot K_{b_1} \cdot K_{b_2} \neq 0$. In the regions $K \cdot K_{b_1}$ and $K \cdot K_{b_2}$ we have, however, $F(z) = F_{b_1}(z)$ and $F(z) = F_{b_2}(z)$, respectively, and hence $F_{b_1}(z) = F_{b_2}(z)$ in the set $K \cdot K_{b_1} \cdot K_{b_2}$. From this, however, it follows

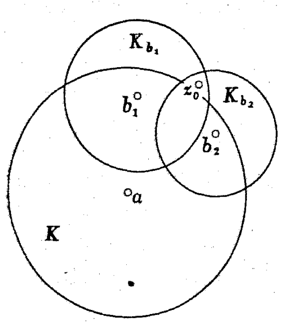

Fig. 19.

(Chapter III, theorem 8.6) that the functions $F_{b_1}(z)$ and $F_{b_2}(z)$ are identical in the entire region $K_{b_1} \cdot K_{b_2}$ and that, in particular, $F_{b_1}(z_0) = F_{b_2}(z_0)$.

The function $H(z)$ is obviously meromorphic in G and identical with $F(z)$ in K. Now, if the circumference of the circle K did not contain points of non-continuability, then we should have $\bar{K} \subset G$. Hence there would exist a circle K^* such that $\bar{K} \subset K^* \subset G$, and in it a meromorphic function $H(z)$, identical with $F(z)$ in the circle K. Therefore the circle K would not be the circle of the element $\{F, a\}$, except in the case when $K = \bar{K}$, *i. e.* when K is the entire plane, and hence, when the function $F(z)$ is a rational function.

The case $a = \infty$ reduces to the case considered above by an inversion.

On the other hand, the set of points of continuability of an element on the circumference of its circle can be empty even when the radius of the element is finite. The simplest example is the element $\{F, 0\}$, where

$$F(z) = \sum_{n=0}^{\infty} z^{2^n}.$$

We have, namely, for every pair of positive integers k, p, and for $r < 1$

$$\left|F\left(r \exp \frac{k\pi i}{2^p}\right)\right| = \left|\sum_{n=0}^{\infty} r^{2^n} \exp 2^{n-p} k\pi i\right| \geqslant -(p+1) + \sum_{n=p+1}^{\infty} r^{2^n},$$

whence

$$\lim_{r \to 1-} \left|F\left(r \exp \frac{k\pi i}{2^p}\right)\right| = \infty. \tag{2.2}$$

The points of the form $z = \exp(k\pi i/2^p)$, where k and p are positive integers, form an everywhere dense set on the circumference $\mathrm{C}(0;1)$; and since a meromorphic function can assume the value ∞ in an isolated set at most, it follows from (2.2) that $\mathrm{K}(0;1)$ is the circle of the element under consideration, and at the same time that none of the points of the circumference of this circle is a point of continuability of this element.

The preceding definitions depart somewhat from the traditional ones. Weierstrass's original definition of an analytic element (and of related notions) differs from the definition given in this section, namely, instead of the condition of meromorphism of the functions under consideration there appears in it the condition of holomorphism, which is more restrictive. Moreover, since every function holomorphic at a given point is expansible in the neighbourhood of this point in a power series, and the circle of convergence of this series is the largest circle with centre at the given point to which the function can be extended with the preservation of holomorphism it follows that an *analytic element, in the sense of Weierstrass, with centre a* may be regardet simply as a power series with centre a and positive radius of convergence. The analogues of the centre, circle and radius of an analytic element are then the centre, circle of convergence and radius of convergence of a power series. Changing similarly the definitions of a direct continuation of an analytic element as well as those of the points of continuability and non-continuability, we obtain the definitions of a *direct continuation of a power series* (which is again a certain power series) as well as those of the *points of continuability* and *non-continuability of a power series.* Further considerations of this chapter (in particular theorem 2.1) carry over formally into the theory of the analytic continuation of a power series. The example considered above (p. 241) of an analytic element without points of continuability on the circumference of its circle is at the same time an example of a power series which does not have points of continuability on the circumference of its circle of convergence.

The points of non-continuability of a series are frequently called *singular* points of the series.

The following examples illustrate the preceding relations. A rational function determines at each point of the plane an analytic element whose circle is the entire plane, and at every point z_0, not a pole of the function — a power series with radius equal to the distance of the point z_0 from the nearest finite pole. A function holomorphic in the entire open plane determines at each of its points an analytic element as well as a power series with an infinite radius. The function $\exp 1/(z-a)$ determines at each point $z_0 \neq a$ an analytic element as well as a power series with radius equal to $\varrho(a, z_0)$ (also when $z_0=\infty$). The only point of non-continuability for this element (as also for the power series) is the point a.

EXERCISES. 1. The functions considered in exercises 5-8, § 2, Chapter IV, determine analytic elements with the circle $\mathrm{K}(0;1)$, non-continuable at any point of the circumference of this circle.

2. Let C be a regular arc, and $f(z)$ a continuous function on C. The function

$$F(z)=\frac{1}{2\pi i}\int_C \frac{f(\mathfrak{z})}{\mathfrak{z}-z}\,d\mathfrak{z}$$

is then holomorphic in the entire completement of C. Show that if the function $f(z)$ is defined and holomorphic in the neighbourhood of a point a of the curve C, which is not an end-point of this curve, then for every point b, not lying on C and sufficiently close to the point a, the circle of the element $\{F, b\}$ contains the point a.

3. In order that the power series $\sum_n a_n z^n$ with circle of convergence $\mathrm{K}(0;1)$ have exactly one point of non-continuability on the circumference of this circle, namely, a simple pole at the point 1, it is necessary and sufficient that $\limsup_n |a_{n+1}-a_n|^{1/n}<1$ (Pringsheim).

(The point a on the circumference of convergence of a power series is said to be a *k-tuple pole* of this series if the function equal to the sum of the given series in its circle of convergence is extensible as a meromorphic function with a k-tuple pole at the point a to a region containing the point a.)

4. If

$$(*)\qquad F(z)=\sum_n a_n z^n$$

is a convergent power series in the circle $\mathrm{K}(0;R)$ with finite radius, then, for every point $z_0=re^{i\alpha}\in\mathrm{K}(0;R)$, the series

$$({*\atop *})\qquad \sum_n \frac{F^{(n)}(z_0)}{n!}(z-z_0)^n$$

is a direct continuation of the series (*) and has a radius of convergence $\geqslant R-r$. If $z_0\neq 0$, then, in order that the point $Re^{i\alpha}$ be a point of non-continuability of the series (*), it is necessary and sufficient that the radius of convergence of the series $({*\atop *})$ be equal to $R-r$.

5. If a series $\sum_n a_n z^n$ with real non-negative coefficients has radius of convergence 1, then the point 1 is a point of non-continuability of this series (Pringsheim, Vivanti, Dienes).

6. Let $\sum_n \alpha_n z^n$ and $\sum_n \beta_n z^n$ be series with real coefficients and with radii of convergence $\geqslant 1$. If the point 1 is a point of continuability of the series $\sum_n (\alpha_n+i\beta_n)z^n$, then it is also a point of continuability for both given series.

Deduce from this that in exercise 5 the condition that the coefficients a_n are real and non-negative can be replaced by a somewhat more general condition, namely, that $|\mathcal{I}a_n|\leqslant C\cdot\mathcal{R}a_n$ for $n=1,2,\ldots$, where C is a positive finite constant, and also by the condition that $\mathcal{R}a_n\geqslant 0$ for $n=1,2,\ldots$ and that the series $\sum_n(\mathcal{R}a_n)z^n$ has radius of convergence 1.

7. Let $\{n_k\}_{k=0,1,\ldots}$ be an increasing sequence of positive integers such that

$$n_{k+1}-n_k\geqslant \alpha n_k \quad\text{for}\quad k=0,1,2,\ldots,\quad\text{where}\quad \alpha>0.$$

Then every power series of the form

$$(*)\qquad \sum_k a_k z^{n_k},$$

provided it has a finite radius of convergence, is non-continuable at each point of the circumference of its circle of convergence (Hadamard: *theorem on "gap" power series*).

[*Hint.* Assuming that the series (*) has radius of convergence 1 and that it is continuable, say, at the point 1, let us denote by $F(z)$ a function holomorphic in a region containing the circle $\mathrm{K}(0;1)$ as well as a neighbourhood of the point 1, and defined in $K(0;1)$ by the series (*). The function $G(\mathfrak{z})=F[\mathfrak{z}^p(\mathfrak{z}+1)/2]$, where $p>1/a$ is a positive integer, is then holomorphic in a certain circle $|\mathfrak{z}|<r$, where $r>1$. We obtain the power series of the function $G(\mathfrak{z})$ in the circle $|\mathfrak{z}|<1$, replacing the terms $a_k z^{n_k}$ in the series (*) by the expansions of the terms $a_k[\mathfrak{z}^p(\mathfrak{z}+1)/2]^{n_k}$ in powers of the variable $\mathfrak{z}$ (see Chapter IV, § 9, exercise 5); and since the series of the function $G(\mathfrak{z})$ is convergent in the entire circle $|\mathfrak{z}|<r$, it follows that returning again to the variable z, we should obtain the result that the series (*) is convergent at all the points of a certain neighbourhood of the point 1, and therefore at points outside its circle of convergence (Faber, Mordell).]

8. Examples of series satisfying the condition of Hadamard's theorem in exercise 7: the series $\sum_k z^{2^k}/2^k$, $\sum_k z^{k!}/2^k$ with radius of convergence 1 (uniformly convergent on the entire closed circle $\overline{K}(0;1)$); more generally: every series of the form $\sum_k a_k z^{m_k}$ and of finite radius of convergence, where $\{m_k\}$ denotes an increasing sequence of positive integers in which every successive number is an integral multiple of the preceding one. Prove directly, without appealing to Hadamard's theorem, that every such series is non-continuable at all points of the circumference of the circle of convergence.

9. Generalize Hadamard's theorem of exercise 7 as follows. Let $\{n_k\}_{k=0,1,\dots}$ be an increasing sequence of positive integers and let

$$(**) \qquad \sum_n a_n z^n$$

be a power series with a finite radius of convergence, such that $a_n=0$ for $n_k<n<(1+a)n_k$, $k=0,1,\dots$, where $a>0$.

If $\{s_n(z)\}_{n=0,1,\dots}$ denotes the sequence of partial sums of the series (**), then the subsequence $\{s_{n_k}(z)\}_{k=0,1,\dots}$ of this sequence is uniformly convergent in the neighbourhood of every point of continuability of the series (**) (Ostrowski).

[*Hint*. The method is analogous to that indicated in exercise 7 for the proof of Hadamard's theorem (Estermann).]

10. A power series is said to be *over-convergent* in a region G containing the circle of convergence of the series, if it is possible to select from the sequence of its partial sums a subsequence which converges almost uniformly in G. If this region extends beyond the circle of convergence, then all the points of the circumference of this circle which are contained in G are points of continuability of the series under consideration.

An example of a power series which is over-convergent in a region extending beyond its circle of convergence. The given series is

$$(*) \qquad F(z)=\sum_{n=1}^{\infty}\frac{[z(1-z)]^{4^n}}{2^{4^n}}.$$

Prove that: 1° this series is almost uniformly convergent in the region $K(0;1)+K(1;1)$ and the function $F(z)$ defined by it is, therefore, holomorphic in this region; 2° the expansion of the function $F(z)$ in a power series with centre 0 is obtained formally by dropping the parentheses in the series (*) and arranging the terms according to powers of the variable z (see Chapter IV, § 9, exercise 5); 3° the circle of convergence of this expansion is the circle $K(0;1)$.

Moreover, since the sequence of partial sums of the series (*) is a subsequence of the sequence of partial sums of the expansion of the function $F(z)$, the power series of the function $F(z)$ is over-convergent in the region $K(0;1)+K(1;1)$ extending beyond the circle of convergence $K(0;1)$ of this series.

[*Hint*. We have $|z(1-z)|<2$ for $z\in K(0;1)+K(1;1)$, and $|z(1-z)|=2$ for $z=-1$.]

11. Let G be a simply connected bounded region containing the circle $K(0;1)$ in its interior and having at least one point of the circumference of this circle on its boundary. Construct a power series with centre 0 and radius of convergence 1, over-convergent in the region G (see exercise 10).

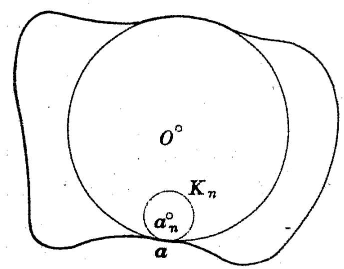

Fig. 20.

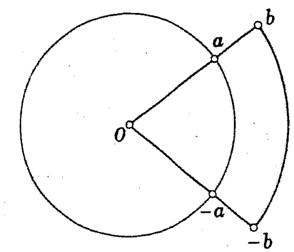

Fig. 21.

[*Hint*. Let a be a point common to the circumference $C(0;1)$ and the boundary of the region G (Fig. 20). Let $a_n=(1-1/3^n)a$, $K_n=K(a_n;1-|a_n|)$. We represent G as the sum of an increasing sequence $\{F_n\}$ of closed sets, taking as F_n the set of all points $z\in G$ such that $\varrho(z,CG)\geqslant 2/3^n$ (then $a_k\in F_n$ for $k=1,2,\dots,n-1$; however, $F_n\cdot K_n=0$ for $n=1,2,\dots$). Making use of Runge's theorem (Chapter IV, theorem 2.2) we define by induction a sequence of polynomials $\{P_n(z)\}_{n=1,2,\dots}$ as well as an increasing sequence of integers $\{m_n\}_{n=1,2,\dots}$ such that: (a) the number m_n is larger than the degree of the polynomial $z^{m_{n-1}}P_{n-1}(z)$, (b) $|z^{m_n}P_n(z)|\leqslant 1/2^n$ for $z\in F_n$, (c) $|z^{m_n}P_n(z)|\geqslant n$ for $z=a_n$.

The series $\sum_n z^{m_n}P_n(z)$ becomes a power series (see Chapter IV, § 9, exercise 5) with the desired property.]

12. A power series can be divergent even at those points of the circumference of its circle of convergence at which it is continuable (example: the geometric series $1+z+z^2+\dots$ is continuable at all the points of its circumference of convergence with the exception of the point 1, but at none of them is it convergent). However:

If a power series $\sum_n a_n z^n$ has radius of convergence 1, and if $a_n \to 0$ as $n \to \infty$, then this series converges at every point of the circumference $C(0;1)$ at which it is continuable (Fatou, M. Riesz).

[*Hint.* Let 1 be a point of continuability of the series under consideration; we can determine a closed circular sector W (Fig. 21), extending beyond the circle $K(0;1)$ and containing the point 1 inside, such that there exists a function $F(z)$ holomorphic on W and equal to the sum of the given series in the part of the sector contained in the circle $K(0;1)$. Taking $s_n(z) = a_0 + a_1 z + \ldots + a_n z^n$, we consider the function

$$R_n(z) = \frac{F(z) - s_n(z)}{z^{n+1}}(z+a)(z-a)$$

(cf. Fig. 21). It is sufficient to prove that $R_n(z)$ tends uniformly to zero on the perimeter of the sector W, and hence (Chapter III, theorem 12.7) on this entire sector; to that end, we estimate $|R_n(z)|$ on the perimeter of the sector, considering successively the segment $[0, -a]$ minus the point $-a$, the segment $[-a, -b]$, the arc $[-b, b]$ and the remaining analogous parts of the perimeter.]

§ 3. Analytic continuation along a curve. We say that a family of analytic elements $\{P(t)\}_{a \leqslant t \leqslant b}$, depending on a real parameter t ranging over the interval $[a,b]$, is a *chain of elements along the curve* C, given by the equation $z = z(t)$, where $a \leqslant t \leqslant b$, if:

1° for every $t \in [a,b]$ the point $z(t)$ is the centre of the element $P(t)$,

2° to every $t \in [a,b]$ there corresponds a number $\varepsilon > 0$ such that if $|h| < \varepsilon$ and $t+h \in [a,b]$, then the element $P(t+h)$ is a direct continuation of the element $P(t)$.

An analytic element P_2 is called a *continuation of the element* P_1 *along the curve* $z = z(t)$, where $a \leqslant t \leqslant b$, if there exists a chain of elements $\{P(t)\}_{a \leqslant t \leqslant b}$ along this curve, such that $P_1 = P(a)$ and $P_2 = P(b)$; then, we also say that the chain $\{P(t)\}$ *joins* the element P_2 to P_1. It is evident that if the element P_2 is a continuation of the element P_1 along the curve C, then the element P_1 is a continuation of the element P_2 along the curve $-C$.

From the above definitions we have immediately the following theorem:

(3.1) *If $F(z)$ is a meromorphic function in an open set G, then for every curve $z = z(t)$, where $a \leqslant t \leqslant b$, lying in G, the family of analytic elements $P(t) = \{F, z(t)\}$ is a chain of elements along this curve.*

An analytic element P_2 is said to be a *continuation of the element* P_1 if it is a continuation of P_1 along some curve. In particular, if P_2 is a continuation of the element P_1 along a curve lying in some open set G, then P_2 is said to be a *continuation of the element P_1 in the set G.*

(3.2) *Every element has at most one continuation along a curve emanating from its centre.*

Proof. Let us assume that $\{P(t)\}$ and $\{R(t)\}$ are two chains of elements along the curve $z = z(t)$, where $a \leqslant t \leqslant b$, and that $P(a) = R(a)$. Let T_1 denote the set of those points t of the interval $[a,b]$ for which $P(t) = R(t)$, and let T_2 be the set of the remaining points of this interval. If t_1 is a point of accumulation of the set T_1, then for values t of the set T_1, sufficiently close to the point t_1, the elements $P(t) = R(t)$ are direct continuations of the element $P(t_1)$ as well as of $R(t_1)$; and since the elements $P(t_1)$ and $R(t_1)$ have a common centre, $P(t_1) = R(t_1)$ and $t_1 \in T_1$. Similarly, if t_2 is a point of accumulation of the set T_2, then $P(t_2) \neq R(t_2)$, and hence $t_2 \in T_2$: indeed, were $P(t_2) = R(t_2)$, then for every value t sufficiently close to the point t_2 the elements $P(t)$ and $R(t)$ would be direct continuations, with a common centre, of the same element; we should therefore have $P(t) = R(t)$, and hence $t \in T_1$, which is contrary to the fact that $t_2 \in \overline{T}_2$. Both sets T_1 and T_2 are therefore closed; hence, one of them is certainly empty. Moreover, $P(a) = R(a)$, *i.e.* $a \in T_1$; consequently, $T_2 = 0$ and therefore $P(t) = R(t)$ for every t of the interval $[a,b]$. In particular, $P(b) = R(b)$.

§ 4. Analytic functions. Every (non-empty) family $\mathfrak{P}$ of analytic elements with centres in a region G, such that:

1° given any two elements of the family $\mathfrak{P}$, one is a continuation of the other in the region G,

2° every analytic element which is a continuation in the region G of an analytic element belonging to $\mathfrak{P}$ belongs to $\mathfrak{P}$,

will be called an *analytic function in the region G.*

If G is the entire plane, an analytic function in G will be called simply *analytic function.*

If P is an arbitrary analytic element with centre in a region G, then the set of all continuations of this element in the region G is an analytic function in this region, containing P. It is the only analytic function in G which contains P. Consequently, every analytic element with centre in a region G, *determines* an analytic function in this region.

If $\mathfrak{P}$ is an analytic function in a region G, and H is a region contained in G, then every element of the function $\mathfrak{P}$ with centre in the region H determines an analytic function in H. Every such function is called a *branch* of the analytic function $\mathfrak{P}$ *in the subregion* H of the region G; obviously all its elements also belong to the function $\mathfrak{P}$.

Conversely, when $\mathfrak{R}$ is an analytic function in a region H, and G an arbitrary region containing H, then the function $\mathfrak{R}$ can be considered as a branch of a certain analytic function $\mathfrak{P}$ in G. This function — which we call the *continuation of the function* $\mathfrak{R}$ in the region G — is obtained as the set of all the continuations in the region G of an arbitrarily chosen element of the function $\mathfrak{R}$.

(4.1) *If $\mathfrak{P}$ is an analytic function in a region G, and $P=\{F,a\}$ is an element of this function, then all the direct continuations of this element, with centres belonging to a certain neighbourhood K of the point a, are also elements of the function $\mathfrak{P}$; if G is the entire plane, then the circle of the element P is such a neighbourhood.*

Proof. Let K be a circle with centre a which is contained simultaneously in G and in the circle of the element P. For every point $b\epsilon K$ the element $\{F_a,b\}$ is therefore (by theorem 3.1) the continuation of the element P along the segment $[a,b]$, contained in G, and hence belongs to $\mathfrak{P}$.

If $\{F,a\}$ is an element of an analytic function $\mathfrak{P}$, then $F(a)$ is called the *value* of the function $\mathfrak{P}$ at the point a. The values of the function $\mathfrak{P}$ at the point a are generally denoted by $\mathfrak{P}(a)$. We also frequently write — traditionally, even though not entirely with justification — $\mathfrak{P}(z)$ instead of $\mathfrak{P}$, denoting by z the variable point of the region in which the function $\mathfrak{P}$ is defined.

The values of the function $\mathfrak{P}$ are defined only at those points which are centres of the elements of this function. In view of theorem 4.1, the set of these points for the function $\mathfrak{P}$, analytic in the region G, is an open subset G_1 of the region G. Moreover, since any two elements of the function $\mathfrak{P}$ are continuations of each other along a curve joining their centres, it is immediately evident that G_1 is a connected open set, *i. e.* a subregion of the region G. We shall call it the *natural subregion* of the analytic function $\mathfrak{P}$ in the region G, and in the case when G is the entire plane — simply the *natural region* of this function.

If $\{F,a\}$ is an element of the analytic function $\mathfrak{P}$, then by theorem 4.1 the values of the function $F(z)$ at the points of a certain neighbourhood of the point a are at the same time the values of the function $\mathfrak{P}$ at these points; in other words,

(4.2) *In a sufficiently small neighbourhood of every point of the natural subregion of an analytic function $\mathfrak{P}$, there exists a meromorphic function whose values at the points of this neighbourhood are at the same time values of the function $\mathfrak{P}$ at these points.*

Let $\mathfrak{P}$ be an analytic function in the region G. If there exists a finite number p such that every point of this region is the centre of at most p elements of the function $\mathfrak{P}$, then this function is said to be *finitely valued* in G, and in the contrary case — *infinitely valued.* If every point of the region G is the centre of at most p elements of the function $\mathfrak{P}$, and in addition there exist points $z\epsilon G$ which are the centres of exactly p elements, then the function $\mathfrak{P}$ is said to be *p-valued* in G. Finally, if every point of the natural subregion of the function $\mathfrak{P}$ is the centre of the same number p of elements of the function, then we say that the function $\mathfrak{P}$ is *strictly p-valued,* or *p-valued in the strict sense,* in the region G.

It is easy to see that if $\mathfrak{P}$ is a p-valued $(p\neq\infty)$ analytic function in the region G, then the set of those points $z\epsilon G$ which are the centres of p distinct elements of the function is an open set.

At a point which is the centre of p elements of an analytic function, the function may assume fewer than p distinct values. However, the set of such points is at most denumerable. More precisely: *if, for $p\neq\infty$, every point of a set Z is the centre of at least p distinct elements of the function $\mathfrak{P}$, then the set of those points $z\epsilon Z$ at which the function assumes fewer than p distinct values is isolated and closed in Z.* In fact, if the point a is the centre of p distinct elements $\{W^{(1)},a\}$, $\{W^{(2)},a\},\ldots,$ $\{W^{(p)},a\}$ of the function $\mathfrak{P}$, then (Chapter III, theorem 8.6) no two of the functions $W^{(j)}$, for $j=1,2,\ldots,p$, assume the same value at any point $z\neq a$ of a sufficiently small neighbourhood of the point a. In particular, therefore, *every strictly p-valued analytic function in a region G assumes exactly p distinct values at every point of its natural subregion, with the exception at most of points of an isolated set closed in this subregion.*

If $W(z)$ is a function meromorphic in a region G, then the set of analytic elements which this function determines at the points of the region G is, in view of theorems 3.1 and 3.2, a single-valued analytic function in G. Its values are obviously identical with the corresponding values of the given meromorphic function W at every point of the region G.

Conversely, if $\mathfrak{W}$ is an analytic function in a region G, having exactly one value at each point of its natural subregion G_1, then, taking $W(z)=\mathfrak{W}(z)$ for $z \in G_1$, we obtain by theorem 4.2 a meromorphic function W in G_1, and we see at once that the function $\mathfrak{W}$ is the set of analytic elements which the function W determines at the points of the region G_1.

The notions of a meromorphic function and an analytic function single-valued in its natural region (or subregion) are therefore equivalent; accordingly, in the future we shall be able to identify without fear of misunderstanding a single-valued analytic function with its corresponding meromorphic function. On the other hand, from the above considerations it follows, at the same time, that *in order that an analytic function be single-valued it is necessary and sufficient that it assume exactly one value at each point of its natural region (or subregion).*

A function analytic in a region G is said to be *univalent*[1]) if it does not assume the same value at any two distinct points of this region. The term "univalent function" and the term "single-valued function" are completely unrelated, since a univalent function need not be single-valued.

As an example of a univalent function (which is at the same time multi-valued), we shall define the logarithm. For every point a different from 0 and ∞, let us denote by $L_a^{(0)}(z)$ the single-valued branch[2]) of $\log z$ in the circle $\mathrm{K}(a;|a|)$, taking the value $\operatorname{Log} a$ at the point a, and let $L_a^{(k)}(z)=L_a(z)+2k\pi i$. In this way we obtain, when k runs through all the integers, all the single-valued branches of $\log z$ in the circle $\mathrm{K}(a;|a|)$ (cf. Chapter I, § 11). Let us denote by $\mathfrak{L}$ the set of all the analytic elements $\{L_a^{(k)},a\}$, where $a\neq 0$, $a\neq\infty$, $k=0,\pm 1,\pm 2,\ldots$ We shall prove that this set is an analytic function.

[1]) The corresponding terms in French and German are: *fonction univalente* and *schlichte Funktion*, respectively. We also say, more generally, that an analytic function is *p-valent* (French: *fonction p-valente* (Montel)), if it assumes each of its values at at most p points, and if there exist values which it assumes at exactly p distinct points.

[2]) The word branch is here understood in the sense established in Chapter I, § 11. As will appear later on, the branch of a logarithm in this sense can be considered as a branch of $\log z$ also in the sense now given for analytic functions in general.

To that end, let us consider one of these elements, *e. g.* $\{L_1^{(0)},1\}$, and let $\mathfrak{L}_0$ be the analytic function determined by it. Since $\exp L_1^{(0)}(z)=z$, it follows that continuing the element $\{L_1^{(0)},1\}$ to any point b along any curve not passing through 0 or ∞ we always obtain an element $\{F,b\}$ such that $\exp F(z)=z$, and therefore an element of the form $\{L_b^{(k)},b\}$. Consequently $\mathfrak{L}_0\subset\mathfrak{L}$. On the other hand, if b is an arbitrary point different from 0 and ∞, then denoting by C_b the circumference $z=b\exp it$, where $0\leqslant t\leqslant 2\pi$, passing through b, we verify immediately that by continuing an arbitrary element $\{L_b^{(k)},b\}$ along C_b, from the point b to this same point, we obtain the element $\{L_b^{(k)}+2\pi i,b\}$, and, more generally, continuing this element along the curves nC_b we obtain all the elements $\{L_b^{(k)}+2\pi ni,b\}$, where $n=0,\pm 1,\pm 2,\ldots$, *i. e.* all the elements of the family $\mathfrak{L}$ with centre b. It follows from this that $\mathfrak{L}\subset\mathfrak{L}_0$, and therefore finally $\mathfrak{L}=\mathfrak{L}_0$.

The family $\mathfrak{L}$ is consequently an analytic function and we define it as the *analytic function* $\log z$. Analogously, we define the *analytic function* $\sqrt{z}$, which is an example of a double-valued function, the *analytic function* $\sqrt[n]{z}$ (where n is an arbitrary integer different from zero), which is an example of an n-valued function (in the strict sense), etc. The natural region for each of these functions is the plane minus the points 0 and ∞.

EXERCISES. 1. A single-valued branch of $\operatorname{arc\,tan} z$ exists in the neighbourhood of every point $a\neq\pm i$ and determines an analytic element at this point. Find the radius of this element. Verify that all the elements determined in this way by the single-valued branches of $\operatorname{arc\,tan} z$ form one infinitely valued analytic function. This function is called the *analytic function* $\operatorname{arc\,tan} z$. Determine its natural region.

A similar exercise for $\operatorname{arc\,cos} z$ and $\operatorname{arc\,sin} z$.

2. The natural region of the function $\exp\{1/(1+\sqrt{z})\}$ is the entire plane minus the points 0 and ∞. This function is double-valued, however, not strictly double-valued: every $z\neq 1$ of its natural region is the centre of two elements, but the point $z=1$ is the centre of only one element of the function.

3. If every point of a set Z is the centre of infinitely many elements of an analytic function $\mathfrak{P}$, and at the same time the function $\mathfrak{P}$ assumes

only a finite number of different values at each of the points of this set, then the set Z is at most denumerable.

4. If $\mathfrak{P}$ is a univalent analytic function in a region G, then the number of values which this function assumes at each point $a \in G$ is equal to the number of elements of the function $\mathfrak{P}$ with centre at this point. Therefore, in order that a univalent analytic function in the region G be strictly p-valued, it is necessary and sufficient that it assume exactly p distinct values at every point of the region.

5. Let $\sum_k \lambda_k$ be an absolutely convergent series with terms different from zero and let $A=\{a_1, a_2, \ldots\}$ be a sequence of (distinct) points on the circumference $C(0;1)$. The series

$$(*) \qquad \sum_k \frac{\lambda_k}{z-a_k}$$

is then almost uniformly convergent in the circle $K(0;1)$ as well as in the interior of this circle, *i. e.* in the circle $K(\infty;1)$. Let $F_1(z)$ and $F_2(z)$ denote the holomorphic functions given by the series (*) in the circles $K(0;1)$ and $K(\infty;1)$, respectively.

Prove that if the set A is everywhere dense on the circumference $C(0;1)$, then these circles are, respectively, the natural regions of the functions $F_1(z)$ and $F_2(z)$; however, if the set A is not everywhere dense on $C(0;1)$, then the set CA' (where A' denotes the set of the points of accumulation of the set A) is a region containing $K(0;1)$ and $K(\infty;1)$, and the series (*) represents a meromorphic function in the region CA', for which this region is its natural region.

6. Let an arbitrary region G be given. Construct a function holomorphic in G, for which the region G is its natural region.

(a) Method I. Let R be the boundary of the region G. It may be assumed that R does not contain the point ∞. Let $A=\{a_n\}$ be the sequence of isolated points of the boundary (provided such points exist), and $B=\{b_n\}$ a sequence everywhere dense in $R-A$. With each point b_n we associate a sequence $\{b_n^{(k)}\}_{k=1,2,\ldots}$ of points of the region G converging to b_n.

Let r_n denote the lower bound of the numbers $|b_m^{(k)}-b_n|$, and ϱ_n the lower bound of the number $\exp[-1/|b_m^{(k)}-a_n|]$, where $m=1,2,\ldots,n-1$ and $k=1,2,\ldots$ Then the function

$$F(z)=\sum_n \frac{\varrho_n \exp[1/(z-a_n)]}{2^n}+\sum_n \frac{r_n}{2^n(z-b_n)}$$

has the desired property.

(b) Method II, based on Runge's theorem (Chapter IV, theorem 2.1). For every n we can cover the boundary R of the region G by a finite system of circles $K_1^{(n)}, K_2^{(n)}, \ldots, K_{m_n}^{(n)}$, with radii $\leqslant 1/n$ and centres on R. In each of the circles $K_j^{(n)}$, where $j=1,2,\ldots,m_n$, we choose two distinct points $p_j^{(n)} \in G$ and $q_j^{(n)} \in G$ (as on Fig. 22 for $j=1$ and $j=2$). Denoting by $H^{(n)}$ the sum

of these circles, we can determine the system $K_1^{(n)}, K_2^{(n)}, \ldots, K_{m_n}^{(n)}$, for $n=1,2,\ldots$, in such a way that the sequence $\{H_n\}$ is decreasing and

$$p_j^{(n)} \in G-H^{(n+1)}, \qquad q_j^{(n)} \in G-H^{(n+1)}$$

for $j=1,2,\ldots,m_n$.

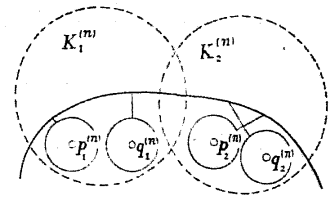

Fig. 22.

Making use of Runge's theorem, we define by induction a sequence of holomorphic functions $W_n(z)$ in the region G in such a way that

1° $|W_n(z)| \leqslant 1/2^n$ for $z \in G-H^{(n)}$,

2° $|W_1(z)+W_2(z)+\ldots+W_n(z)|<1/2^n$ for $z=p_j^{(n)}$, and $>2^n$ for $z=q_j^{(n)}$, where $j=1,2,\ldots,m_n$.

The series $\sum_n W_n(z)$ is then almost uniformly convergent in G and its sum is a holomorphic function with the desired property (this function does not have a limit, finite or infinite, at any point of the boundary R of the region G upon approaching it from the interior of this region).

7. Let $\boldsymbol{H}$ denote a space whose elements are functions holomorphic in an arbitrarily fixed region G (Chapter II, § 7, exercise 3).

Let $K_1, K_2, \ldots, K_n$ be an arbitrary system of circles with centres on the boundary of the region G, and ε an arbitrary positive number. Let $\mathfrak{P}$ denote the family of all functions $P(z)$ holomorphic in G, such that for every $j=1,2,\ldots,n$ the set $K_j \cdot G$ contains points at which $|P(z)|<\varepsilon$, and points at which $|P(z)|>1/\varepsilon$. Prove that the holomorphic functions in G which do not belong to $\mathfrak{P}$ form a nowhere dense closed set in the space $\boldsymbol{H}$.

Deduce from this (not appealing to the result of exercise 6) that holomorphic functions in the region G exist for which G is their natural region and, moreover, all functions holomorphic in G have this property, with the exception, at most, of functions forming a set of the first category in $\boldsymbol{H}$ (Mazurkiewicz).

(Cf. analogous exercises 6-8, Chapter IV, § 2.)

8. Example of a double-valued analytic function such that every point of a certain region is the centre of only one element of the function (cf. exercise 2).

Let H denote the sum of the half-plane $\mathcal{I}z<0$ and the circle $|z|<1$ (Fig. 23) and let $\Phi(z)$ be a holomorphic function for which H is the natural region (cf. exercise 6). On the other hand, let G denote the doubly connected region obtained by removing the point 0 and the interval $[1,+\infty]$ of the real axis from the plane (Fig. 24). The region G is then the natural region of the function $\Phi(\sqrt{z})$. Moreover, every point of the region G belonging to the circle $K(0;1)$ is the centre of two distinct elements of this function,

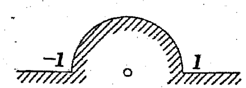

Fig. 23.

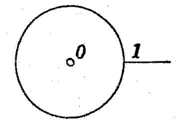

Fig. 24.

while each of the remaining points of the region G is the centre of only one element.

§ 5. Inverse of an analytic function. An analytic element $R=\{F,z_0\}$ is said to be *invertible*, if the function $F(z)$ assumes the value $w_0=F(z_0)$ at the point z_0 once (Chapter III, § 8). The function $F(z)$ then has in a neighbourhood of the point w_0 an inverse function $F^{-1}(w)$ (Chapter III, theorem 12.4); we shall call the element $\{F^{-1},w_0\}$ the *inverse element,* or the *inverse,* of the element $R=\{F,z_0\}$ and denote it by R^{-1}.

It is easy to see that *every analytic function* $\mathfrak{R}$, *which is not a constant, has invertible elements.* In fact, if $\{F,z_0\}$ is an arbitrary element of the function $\mathfrak{R}$, then (cf. Chapter III, theorem 12.1) at every point $z_1\neq z_0$ sufficiently close to the point z_0 the function $F(z)$ assumes the value $F(z_1)$ once, and hence the corresponding element $\{F,z_1\}$, which — as a direct continuation of the element $\{F,z_0\}$ — also belongs to $\mathfrak{R}$, is an invertible element.

(5.1) *If* $R_1=\{F,z_1\}$ *and* $R_2=\{\Phi,z_2\}$ *are invertible elements of a function* $\mathfrak{R}$, *then they can be joined by means of a chain consisting exclusively of invertible elements.*

Proof. Let $R(t)=\{F^{(t)},z(t)\}$, where $a\leqslant t\leqslant b$, be a chain of elements joining the element R_2 to R_1. Let us denote by T the set of those values t of the interval $[a,b]$ for which the element $R(t)$ can be joined to $R_1=R(a)$ by a chain of invertible elements. Let t_0 be the upper bound of the set T and let K_0 denote a sufficiently small neighbourhood of the point $z(t_0)$, so that the function $F^{(t_0)}(z)$ assumes everywhere in K_0, with the exception at most of the point $z(t_0)$, each of its values once. Furthermore, let t_1 be a point of the set T for which $z(t_1)\epsilon K_0$ and for which $R(t_1)$ is a direct continuation of the element $R(t_0)$; finally, let C_1 denote the curve along which the chain of invertible elements joins the element $R(t_1)$ to the element $R(a)$.

Let us note now that $t_0=b$. In fact, supposing that $t_0<b$, we should be able to determine a point t_2 such that $t_0<t_2<b$, $z(t_2)\epsilon K_0$, $z(t_2)\neq z(t_0)$, and such that the element $R(t_2)$ is a direct continuation of the element $R(t_0)$. Let C_2 denote an arbitrary curve lying in K_0 and joining the points $z(t_1)$, $z(t_2)$. Moreover, if $z(t_1)\neq z(t_0)$, we could assume that this curve does not contain the point $z(t_0)$. Then, continuing the element $R(a)$ along the curve C_1+C_2 we should obtain a chain of invertible elements joining the element $R(t_2)$ to $R(a)$. We should therefore have $t_2\epsilon T$, which is impossible in view of the fact that $t_0<t_2$.

Consequently $t_0=b$. The function $F^{(b)}(z)=\Phi(z)$ therefore assumes its values at all points of the segment $[z(t_1),z(b)]\subset K_0$ only once, and the direct continuations of the element $R_2=R(b)$ with centres on this segment are all invertible. Hence, continuing the element $R(a)$ along the curve $C_1+[z(t_1),z(b)]$, we obtain a chain of invertible elements joining the elements $R_2=R(b)$ and $R_1=R(a)$.

(5.2) *If* P_1 *and* P_2 *are two invertible elements of an analytic function* $\mathfrak{P}$, *then the elements* P_1^{-1} *and* P_2^{-1} *also belong to one analytic function.*

Proof. By theorem 5.1 the elements P_1 and P_2 can be joined by a chain of invertible elements $P(t)=\{F^{(t)},z(t)\}$, where $a\leqslant t\leqslant b$. Let $\zeta(t)=F^{(t)}[z(t)]$ and let $H^{(t)}(\zeta)$ denote the inverse of the function $F^{(t)}(z)$ in the neighbourhood of the point $z(t)$. The elements $P^{-1}(t)=\{H^{(t)},\zeta(t)\}$, where $a\leqslant t\leqslant b$, form then a chain along the curve $\zeta=\zeta(t)$ joining the elements $P^{-1}(a)=P_1^{-1}$ and $P^{-1}(b)=P_2^{-1}$. These elements therefore belong to one analytic function.

From theorem 5.2 it follows that all the inverses of the invertible elements of an analytic function $\mathfrak{P}$ determine the same analytic function. We shall call this function the *inverse of the function* $\mathfrak{P}$ and we shall denote it by $\mathfrak{P}^{-1}$. We see at once that the function $\mathfrak{P}$ is in turn the inverse of the function $\mathfrak{P}^{-1}$. In addition to the inverses of the elements of the function $\mathfrak{P}$, the function $\mathfrak{P}^{-1}$ may contain other elements as well; these elements, however, as is easily seen, cannot be invertible. For example, the inverse of the function $\sqrt{z}$ (all of whose elements are invertible) is the function ζ^2, whose elements with centres 0 and ∞ are not the inverses of any element of the function $\sqrt{z}$ and are, moreover, not invertible at all.

EXERCISES. 1. Distinguish the non-invertible elements, if any, of the functions $\exp z$, $\cos z$, $\sin z$, $\tan z$. Verify that the functions $\log z$ (§ 4, p. 251), $\operatorname{arc\,cos} z$, $\operatorname{arc\,sin} z$, $\operatorname{arc\,tan} z$ (§ 4, exercise 1), are, respectively, the inverses of the functions $\exp z$, $\cos z$, $\sin z$, $\tan z$.

2. In order that the inverse of an analytic function $\mathfrak{P}$ be a single-valued function (*i. e.* meromorphic, see § 4, p. 250), it is necessary and sufficient that the function $\mathfrak{P}$ be univalent.

§ 6. Analytic functions arbitrarily continuable in a region. If each element R of an analytic function in a region G has a continuation along every curve emanating from the centre of R and lying in G, then we say that the function is *arbitrarily continuable* in G; the natural subregion of such a function obviously coincides with the entire region G, and every point of this region is the centre of the same number (finite or infinite) of elements of the function.

The simplest example of analytic functions arbitrarily continuable in G are meromorphic functions in G (cf. § 4). Examples of analytic multi-valued functions arbitrarily continuable in a region are the functions $\log z$, $\sqrt{z}$, etc. (§ 4, p. 251), in the region obtained by removing the points 0 and ∞ from the plane. Here, the double connectivity of this region plays an essential role, since, as follows from the *monodromy theorem* (theorem 6.3, below), every analytic function arbitrarily continuable in a simply connected region is single-valued in this region.

(6.1) *If an analytic function $\mathfrak{P}$ in a region G is arbitrarily continuable in this region, then, except for the case when G is the plane, the circle of every element of the function $\mathfrak{P}$ has, inside or on its boundary, points of the complement of the region G.*

If G is the plane, i. e. $CG=0$, then the entire plane is the circle of every element of the function $\mathfrak{P}$.

Proof. Let $P_0=\{F,a\}$ be an arbitrary element of the function $\mathfrak{P}$ and let K be the circle of this element. Let us suppose that K is not the entire plane and that $\bar{K}\subset G$. Let b be an arbitrary point of the circumference of the circle K and let $z=g(t)$, where $0\leqslant t\leqslant 1$, $g(0)=a$, $g(1)=b$, be an arbitrary curve joining the points a and b and lying in the circle K for $0\leqslant t<1$. Continuing the element P_0 along this curve we obtain a chain of elements $P(t)$, such that (cf. theorem 3.1 and 3.2) $P(t)=\{F_a,g(t)\}$ for $0\leqslant t<1$; in particular, $P(0)=P_0$. The elements $P(t)$ are therefore direct continuations of the element P_0 for $t\neq 1$. On the other hand, for values of t sufficiently close to 1, the element $P(t)$ is a direct continuation of the element $P(1)$ with centre b, and the circle of the element $P(t)$ contains the point b. Consequently, every point b on the circumference of the circle of the element P_0 would be a point of continuability of this element, which contradicts theorem 2.1.

From this follows immediately *the monodromy theorem for the circle:*

(6.2) *If an analytic function $\mathfrak{P}$ in a circle K is arbitrarily continuable in this circle, then it is a single-valued function, i. e. a meromorphic function. Hence, if K is the entire plane, then $\mathfrak{P}$ is a rational function.*

Proof. Let a be the centre of the circle K and let $P_0=\{F,a\}$ be an element of the function $\mathfrak{P}$. From theorem 6.1 it follows that the circle K is contained in the circle of the element P_0, and hence that the function $F_a(z)$ is defined in the entire circle K and meromorphic in K. Therefore (cf. theorem 3.1 and 3.2) the function $\mathfrak{P}$ has for each point $b\in K$ exactly one element with centre at this point, namely, $\{F_a,b\}$, and is therefore a single-valued function, q. e. d.

Let $W(z)$ be a uniquely invertible meromorphic function in the region G and let $H=W(G)$. With each element $P=\{F,w_0\}$, where $w_0\in H$, we can then associate uniquely the element $\{FW,z_0\}$, where $z_0=W^{-1}(w_0)\in G$; let us denote this element by PW. We obviously have $P=PW\cdot W^{-1}$ and the established correspondence between elements with centres in the regions H and G is one-to-one.

If $\{P(t)\}$ is a chain of elements along a curve $w=g(t)$ $(a\leqslant t\leqslant b)$ lying in the region H, then the elements $P(t)W$ form a chain of elements in G along the curve $z=W^{-1}[g(t)]$, where $a\leqslant t\leqslant b$. Therefore, if the element P_2 is a continuation in H of the element P_1, then the element P_2W is a continuation in G of the element P_1W; the same occurs when we interchange the regions H and G, replacing simultaneously W by W^{-1}. Hence, if $\mathfrak{P}$ is an analytic function in the region H, then the set of all elements of the form PW, where $P\in\mathfrak{P}$, is an analytic function in G. Let us denote this function by $\mathfrak{P}W$. If the function $\mathfrak{P}$ is arbitrarily continuable in H, then the function $\mathfrak{P}W$ is arbitrarily continuable in G; if the function $\mathfrak{P}$ is single-valued in H, then the function $\mathfrak{P}W$ is single-valued in G. The converse theorems are obviously true, since $\mathfrak{P}=\mathfrak{P}W\cdot W^{-1}$.

Now, let H be a simply connected region, and $\mathfrak{P}$ an analytic function arbitrarily continuable in H. As the region G we take an open circle, and as $W(z)$ a meromorphic function transforming the circle G into the region H conformally (see Chapter V, theorem 6.14). (If the region H were the entire plane or a plane minus one point, then it would also be necessary to take as G the entire plane or the open plane.) The function $\mathfrak{P}W$ is then arbitrarily continuable in the circle G and by theorem 6.2 is single-valued in G. Therefore the function $\mathfrak{P}$ is single-valued in H. Thus, we obtain the general *monodromy theorem:*

(6.3) *An analytic function in a simply connected region H, arbitrarily continuable in this region, is single-valued in this region.*

Theorem 6.3 may be considered as a generalization of theorems 3.1 and 3.2, Chapter IV, on the existence of single-valued branches of the logarithm in simply connected regions. At the same time, the monodromy theorem contains an analytic definition of the simple connectivity of a region. Namely: *In order that the region G be simply connected, it is necessary and sufficient that every analytic function in the region G, arbitrarily continuable in G, be single-valued in this region* (*Radò's criterion*). The necessity of the condition is already contained in theorem 6.3. For the purpose of proving the sufficiency of the condition let us assume that the region G, satisfying the above condition, is not simply connected. Applying, if necessary, an inversion and a translation, we may assume that the points 0 and ∞ belong to different components of the complement of the region G. Every element of the function $\log z$ is then arbitrarily continuable in the region G, and nevertheless determines a multiple-valued function in this region (cf. Chapter IV, theorem 10.3).

EXERCISE. Let G be a plane minus the points 0 and ∞, and $Z\subset G$ an arbitrary denumerable set. Construct an analytic function $\mathfrak{P}$ arbitrarily continuable in G and such that: 1° every point of the set Z is the centre of infinitely many elements of the function $\mathfrak{P}$, 2° at every point of the set Z the function $\mathfrak{P}$ assumes only a finite number of distinct values (cf. § 4, exercise 3).

§ 7. Theorem of Poincaré-Volterra. This theorem states that:

(7.1) *If $\mathfrak{P}$ is an analytic function, then every point of the plane is the centre of at most a denumerable number of elements of the function $\mathfrak{P}$. An analytic function can therefore assume at each point at most a denumerable number of distinct values.*

The proof is based on a few simple auxiliary arguments. We say that a finite sequence of elements $P_1, P_2, \dots, P_n$ joins the analytic elements A and B, if every successive element in this sequence is a direct continuation of its predecessor, and if the extreme elements P_1 and P_n are, respectively, direct continuations of the elements A and B.

(7.2) *If one of the two elements B and C is a direct continuation of the other, and if the element B can be joined to a given element A by a finite sequence of elements with rational centres, then the element C can also be joined to A by a finite sequence of elements with rational centres.*

Proof. Let $P_1, P_2, \dots, P_n$ be a sequence of elements with rational centres joining the elements A and B. Let $p_1, p_2, \dots, p_n$ be the centres of the elements $P_1, P_2, \dots, P_n$, and $K_1, K_2, \dots, K_n, K_B, K_C$, the circles of the elements $P_1, P_2, \dots, P_n, B, C$, respectively. Finally, let $B=\{F,b\}$. We may assume that $F(z)$ is a meromorphic function in the entire circle K_B.

Let us assume that C is a direct continuation of the element B. Consequently, $C=\{F,c\}$, where $c\in K_B$. In addition, $P_n=\{F,p_n\}$, where $p_n\in K_B$. Hence we can define in the circle K_B a finite sequence of circles $K_{n+1}, K_{n+2}, \dots, K_{n+m}$ with rational centres $p_{n+1}, p_{n+2}, \dots, p_{n+m}$ in such a way that $p_{n+j}\in K_{n+j-1}$ for $j=1,2,\dots,m$, and $p_{n+m}\in K_C$. Taking $P_{n+j}=\{F,p_{n+j}\}$ for $j=1,2,\dots,m$, we see immediately that the sequence $P_1, P_2, \dots, P_m, \dots, P_{n+m}$ is the desired sequence of elements with rational centres, joining the elements A and C.

If the element B is a direct continuation of the element C, we obtain the same result by means of a similar argument.

(7.3) *Each two elements of an analytic function can be joined by a finite sequence of elements having rational centres.*

Proof. Denote by A and B two elements of the analytic function $\mathfrak{P}$. Then, there exists (cf. §§ 4, 3) a chain of elements $P(t)$, where $a\leqslant t\leqslant b$, such that $P(a)=A$ and $P(b)=B$. Let T denote the set of those values of t of the interval $[a,b]$ for which the element $P(t)$ can be joined to $A=P(a)$ by a finite sequence of elements with rational centres. Let t_0 be the upper bound of the set T. Since for all values of t of the interval $[a,b]$, sufficiently close to the point t_0, all the elements $P(t)$ are direct continuations of the element $P(t_0)$, it follows from theorem 7.2, first of all, that $t_0\in T$, and next, that all the points t of the interval $[a,b]$, sufficiently close to the point t_0, also belong to T. Since t_0 is the upper bound of the set T, we have $t_0=b$, and therefore $b\in T$, q. e. d.

Proceeding now to the proof of theorem 7.1, let us consider an arbitrary analytic function $\mathfrak{P}$ and let us fix one of its elements P_0. With each element P of the function $\mathfrak{P}$ we can, by theorem 7.3, associate a finite sequence of rational points, such that there exists a sequence of elements with centres at these points, joining the elements P_0 and P. It is evident that to the different elements P of the function $\mathfrak{P}$ with a common centre there will thus correspond different sequences of rational points. Moreover, since the set of all finite sequences of rational points is denumerable, the set of all elements of the function $\mathfrak{P}$ with a common centre is also at most denumerable.

***§ 8. An analytic function as an abstract space.** Let E denote the set of all analytic elements. This set can be thought of as a certain abstract space, where by *neighbourhood* (see Introduction,

§ 3) we mean the set of elements determined at the points of an arbitrary circle $K(a;r)$ by an arbitrary meromorphic function in this circle. In other words, a neighbourhood in the space $\boldsymbol{E}$ is an analytic function arbitrarily continuable in a circle. This circle is called the *circle of the neighbourhood* under consideration. A neighbourhood having a rational circle (Introduction, § 8, p. 21) will be called, for brevity, a *rational neighbourhood.*

It is easy to see that the family of neighbourhoods defined in this manner for the space $\boldsymbol{E}$ satisfies the conditions of the postulates I and II of the Introduction, § 3.

Analytic functions (understood as sets of analytic elements) coincide with the components of the space $\boldsymbol{E}$, and analytic functions in a region G (§ 4) — with the components of the set in the space $\boldsymbol{E}$ formed from the elements with centres belonging to G. The proof does not present any difficulties and may be left to the reader.

Every analytic function $\mathfrak{P}$ itself can also be considered as an abstract space. The family of neighbourhoods for $\mathfrak{P}$ consists of all those neighbourhoods of the space $\boldsymbol{E}$ which contain only elements of the function $\mathfrak{P}$. By the Poincaré-Volterra theorem (theorem 7.1), the rational neighbourhoods in this family constitute a denumerable set which is at the same time, as we easily verify, a denumerable base of the family of neighbourhoods (Introduction, § 3). Consequently:

(8.1) *An analytic function, considered as an abstract space, is a separable space. The elements of this function with rational centres form in it a denumerable everywhere dense set.*

If $F(z)$ is a meromorphic function in the circle $K(a;r)$, then, assigning to each element $\{F,z\}$, where $z\in K(a;r)$, the point $(z-a)/r$, when $a\neq\infty$, and the point $1/rz$, when $a=\infty$, we obtain a homeomorphic mapping of the set of these elements onto the circle $K(0;1)$. In other words:

(8.2) *Every single-valued analytic function $\mathfrak{P}$ in the circle $K(a;r)$, considered as an abstract space, is homeomorphic with the circle $K(0;1)$; we obtain, namely, a homeomorphic correspondence associating with the element of the function $\mathfrak{P}$ with centre $z\in K(a;r)$ the point $(z-a)/r$, or the point $1/rz$, of the circle $K(0;1)$, depending on whether $a\neq\infty$, or $a=\infty$.*

This correspondence between the elements of the function $\mathfrak{P}$ and the points of the circle $K(0;1)$ will be called a *canonical correspondence.*

Theorem 8.2 can also be formulated as follows: *every neighbourhood in the space of analytic elements is homeomorphic with the circle.*

EXERCISE. Every analytic function can have at most a denumerable number of non-invertible elements (§ 5, p. 254) and, if the analytic function is considered as an abstract space, then the set of these elements is an isolated closed set.

§ 9. Analytic functions in an annular neighbourhood of a point. Let $W(z)$ be a function meromorphic in a region G and not reducing to a constant, and let $H=W(G)$. If z_0 is a point of the region G, and $P=\{F,w_0\}$ an arbitrary analytic element with centre $w_0=W(z_0)\in H$, then the element $\{FW,z_0\}$ will be called an *element PW*. This notation was already introduced in § 6 for a uniquely invertible function $W(z)$. If the function $W(z)$ is not uniquely invertible, then there exist in general many points $z_0\in G$, satisfying the condition $w_0=W(z)$ for a fixed point $w_0\in H$, and hence for a given element $P=\{F,w_0\}$ there exist in general many elements PW.

If $\mathfrak{P}$ is an analytic function in H, then every element PW, where $P\in\mathfrak{P}$, determines an analytic function in G. The functions thus obtained will be called *functions* $\mathfrak{P}W$.

If the function $W(z)$ is uniquely invertible, then there exists for a given function $\mathfrak{P}$ only one (cf. § 6, p. 256) function $\mathfrak{P}W$, and this function is the set of all elements of the form PW, where $P\in\mathfrak{P}$. However, if the function $W(z)$ is not uniquely invertible, then there may be many functions $\mathfrak{P}W$ and these functions can contain elements not necessarily of the form PW. For example, taking H and G to be the entire plane, $W(z)=z^2$, and $\mathfrak{P}(w)=\sqrt{w}$, we obtain as $\mathfrak{P}W$ the linear functions z and $-z$.

(9.1) *If $W(z)$ is a meromorpic function, not reducing to a constant, in a region G, and if $\mathfrak{P}$ is an analytic function arbitrarily continuable in the region $H=W(G)$, then each of the functions $\mathfrak{P}W$ is arbitrarily continuable in the region G and is formed exclusively of the elements PW, where $P\in\mathfrak{P}$.*

Proof. It is sufficient to prove that every element of the form P_0W, where $P_0\in\mathfrak{P}$, is continuable along every curve emanating from the centre of P_0 and lying in G, and that the continuation obtained along this curve is also an element of the form P_0W. To that end, let $P_0=\{F,w_0\}$, where $w_0=W(z_0)$ and $z_0\epsilon G$, be an element of the function $\mathfrak{P}$ and let $R_0=\{FW,z_0\}$. Let $z=g(t)$ $(a\leqslant t\leqslant b)$ be a curve in G with initial point z_0. The curve $w=W[g(t)]$ $(a\leqslant t\leqslant b)$ lies in H, having w_0 as its initial point. Therefore there exists a chain of elements $P(t)=\{F^{(t)},W[g(t)]\}$ of the function $\mathfrak{P}$ along the curve $w=W[g(t)]$, such that $P(a)=P_0$. The elements $R(t)=\{F^{(t)}W,g(t)\}$ form then a chain along the curve $z=g(t)$, with $R(a)=\{F^{(a)}W,g(a)\}=\{FW,z_0\}=R_0$; on the other hand, $R(b)=\{F^{(b)}W,g(b)\}$ is the element $P(b)W$, where $P(b)\epsilon\mathfrak{P}$, q. e. d.

If $W(z)$ is a function meromorphic in a region H, then with each analytic element $P=\{F,a\}$ such that $F(a)\epsilon H$, we can associate the element $\{WF,a\}$, which we shall denote by WP. If $\mathfrak{P}$ is an analytic function in a region G, all of whose values belong to the region H, then to each element $P\epsilon\mathfrak{P}$ there corresponds an element WP. It is evident that if P_1 and P_2 are elements of the function $\mathfrak{P}$, and if P_2 is the continuation of the element P_1 along a given curve lying in G, then WP_2 is the continuation of the element WP_1 along the same curve. Therefore all the elements WP, where $P\epsilon\mathfrak{P}$, determine the same analytic function in G; we shall denote it by $W\mathfrak{P}$. It is evident that *if a function $\mathfrak{P}$ is arbitrarily continuable in the region G, then the function $W\mathfrak{P}$ is also arbitrarily continuable in G and is the set of all elements of the form WP, where $P\epsilon\mathfrak{P}$.*

Finally, if $\mathfrak{P}$ and $\mathfrak{R}$ are two analytic functions, and $P=\{F,a\}$, $R=\{\Phi,b\}$, respectively, are elements of these two functions, such that $\Phi(b)=a$, then the function $F\Phi(z)$ is meromorphic in a neighbourhood of the point b and the element $\{F\Phi,b\}$ determines a certain analytic function. An analytic function obtained from the functions $\mathfrak{P}$ and $\mathfrak{R}$ in this manner will be called a *function* $\mathfrak{P}\mathfrak{R}$. Of course, it may happen that the functions $\mathfrak{P}$ and $\mathfrak{R}$ do not have elements P and R satisfying the condition given above; in that case no function $\mathfrak{P}\mathfrak{R}$ exists.

(9.2) *If a function $\mathfrak{R}$, analytic in the annular neighbourhood $P=\mathrm{P}(a;0,r)$ of a point a, is arbitrarily continuable in this neighbourhood, and if some element R of the function $\mathfrak{R}$ is identical with its continuation along the curve $p\cdot C$, where p is an integer, and C denotes the circumference with centre a with its origin at the centre of the element R, then the function $\mathfrak{R}$ is at most p-valued and is of the form $\Phi(\sqrt[p]{z-a})$ (or of the form $\Phi(1/\sqrt[p]{z})$, if $a=\infty$), where $\Phi(z)$ is a function meromorphic in the annulus $\mathrm{P}(0;0,r^{1/p})$.*

Proof. We may obviously assume that $a=0$ (applying a translation or, if $a=\infty$, an inversion). We may also assume (applying, if necessary, a rotation) that $R=\{F,r_0\}$, where $0<r_0<r$, and hence that $C=\mathrm{C}(0;r_0)$.

1° We shall consider first the case $p=1$. Let C_1 and C_2 denote the upper and lower semi-circumferences of the circumference C. Since the element R is identical with its continuation along C, it has the same continuation along each of the semi-circumferences C_1 and C_2; let $R^*=\{F^*,-r_0\}$ be this continuation.

Let us now denote by G_1 and G_2 the simply connected regions which we obtain from the annular neighbourhood P by removing, respectively, the segments $[-ri,0]$ and $[0,ri]$ of the imaginary axis; we shall denote by H and H^* the parts of the annulus P situated to the right and the left of this axis.

By the monodromy theorem (theorem 6.3), the element R determines meromorphic functions in the regions G_1 and G_2; let us denote them by $F_1(z)$ and $F_2(z)$, respectively. We therefore have $F_1(z)=F_2(z)=F(z)$ in the neighbourhood of the point r_0, and $F_1(z)=F_2(z)=F^*(z)$ in the neighbourhood of the point $-r_0$.

Since $r_0\epsilon H\subset G_1\cdot G_2$ and $-r_0\epsilon H^*\subset G_1\cdot G_2$, we have $F_1(z)=F_2(z)$ in the region H as well as in the region H^*, and therefore in the entire open set $G_1\cdot G_2$. The functions $F_1(z)$ and $F_2(z)$ together determine, then, in the entire region $P=G_1+G_2$, one meromorphic function. This function, since it is identical with $F(z)$ in the neighbourhood of the point r_0, coincides with the function $\mathfrak{R}$ as determined by the element $R=\{F,r_0\}$ in the region P.

2° Now, let p be an arbitrary positive integer and let $W(z)=z^p$. The function $W(z)$ transforms the annulus $P_1=\mathrm{P}(0;0,r^{1/p})$ into the annulus $P=\mathrm{P}(0;0,r)$, and, in particular, the circumference $C_1=\mathrm{C}(0;r_0^{1/p})$ into pC. The element $\{FW,r_0^{1/p}\}$ therefore is transformed into itself by a continuation along the circumference C_1, and, on the other hand (cf. theorem 9.1), it determines an arbitrarily continuable analytic function in the annulus P_1. By 1° (for the case $p=1$) this function is a function $\Phi(z)$ meromorphic in the annulus P_1. Consequently, in the neighbourhood of the point $r_0^{1/p}$ we have $F(z^p)=FW(z)=\Phi(z)$ identically, and therefore $F(z)=\Phi[G(z)]$ in the neighbourhood of the point r_0, where $G(z)$

denotes that single-valued branch of $\sqrt[p]{z}$ in the neighbourhood of the point r_0 which assumes the value $r_0^{1/p}$ at this point. Therefore the analytic function $\mathfrak{R}$, determined by the element $R=\{F,r_0\}$, is identical with the function $\Phi(\sqrt[p]{z})$, determined by the element $\{\Phi G,r_0\}$.

(9.3) *Every analytic function $\mathfrak{R}$, arbitrarily continuable and n-valued (where n is a finite number) in the annulus $P=\mathrm{P}(a;0,r)$, has the form $\Phi(\sqrt[n]{z-a})$ if $a\neq\infty$, or the form $\Phi(1/\sqrt[n]{z})$ if $a=\infty$, where $\Phi(z)$ is a function meromorphic in the annulus $\mathrm{P}(0;0,r^{1/n})$.*

Proof. As in the proof of the previous theorem, we may assume that $a=0$. Let $R=\{F,r_0\}$ be an arbitrary element of the function $\mathfrak{R}$ with centre at a point $r_0>0$ and let $C=\mathrm{C}(0;r_0)$. Since the function $\mathfrak{R}$ has only n distinct elements with centre r_0, there exists a number $p\leqslant n$ such that the element R is transformed into itself by a continuation along the curve $p\cdot C$. The function $\mathfrak{R}$ is therefore of the form $\Phi(\sqrt[p]{z})$, where $\Phi(\mathfrak{z})$ is a function meromorphic in the annulus $\mathrm{P}(0;0,r^{1/p})$. We find, at the same time, that $p=n$: indeed, we have on the one hand $p\leqslant n$, and on the other $p\geqslant n$, since the function $\Phi(\sqrt[p]{z})$ is at most p-valued.

EXERCISE. Every arbitrarily continuable analytic function in the annular neighbourhood $\mathrm{P}(0;0,r)$ of the point 0 is of the form $\Phi(\log z)$, where $\Phi(w)$ is a function meromorphic in the half-plane $\mathcal{R}w<\mathrm{Log}\,r$.

***§ 10. An analytic function in an annular neighbourhood as an abstract space.** In § 8 we established a certain (canonical) homeomorphic correspondence between the points of the unit circle and the elements of an analytic function arbitrarily continuable (and hence single-valued) in a given circle. An analogous result can be obtained for an analytic function, finitely valued and arbitrarily continuable in an annular neighbourhood $\mathrm{P}(a;0,r)$.

Let us suppose for simplicity that $a=0$ and let us take into consideration the function $\sqrt[n]{z}$ in the annulus $\mathrm{P}(0;0,r)$. Associating the element $\{G,z\}$ of this function with the point $G(z)/r^{1/n}$, we obtain, as is immediately evident, a homeomorphic correspondence between the elements of the analytic function $\sqrt[n]{z}$ in $\mathrm{P}(0;0,r)$ and the points of the annulus $\mathrm{P}(0;0,1)$. In this correspondence, an element of the function $\sqrt[n]{z}$ with centre $z=r\mathfrak{z}^n$ is associated with the point $\mathfrak{z}\in\mathrm{P}(0;0,1)$.

Let us now consider any analytic function $\mathfrak{R}$, arbitrarily continuable and n-valued in the annulus $\mathrm{P}(0;0,r)$. By theorem 9.3, the function $\mathfrak{R}$ is the set of elements of the form $\{\Phi G,z\}$, where Φ is a certain function meromorphic in the annulus $\mathrm{P}(0;0,r^{1/n})$, and $\{G,z\}$ an arbitrary element of the function $\sqrt[n]{z}$. We shall associate every element $\{G,z\}$ of the function $\sqrt[n]{z}$ with the element $\{\Phi G,z\}$ and show that this correspondence is one-to-one, *i. e.* that if the elements $\{G_1,z_1\}$ and $\{G_2,z_2\}$ of the function $\sqrt[n]{z}$ in the circle $\mathrm{K}(0;r)$ are distinct, then the elements $\{\Phi G_1,z_1\}$, $\{\Phi G_2,z_2\}$ are also distinct. This is obvious when $z_1\neq z_2$. However, if $z_1=z_2$, then, as we see immediately, in the neighbourhood of the point z_1 we have $G_2(z)=G_1(z)\exp(2\pi ki/n)$ identically, where $0<k<n$ in view of $G_1(z)\neq G_2(z)$. Therefore the element $\{G_1,z_1\}$ is transformed into the element $\{G_2,z_1\}$, and the element $\{\Phi G_1,z_1\}$ into the element $\{\Phi G_2,z_1\}$, by a continuation along the curve $k\cdot C$, where $C=\mathrm{C}(0;|z_1|)$. Therefore, if the elements $\{\Phi G_2,z_1\}$ and $\{\Phi G_1,z_1\}$ were identical, then by theorem 9.2 the function $\mathfrak{R}$ would be at most k-valued, which contradicts the fact that $k<n$.

Hence the correspondence established between the elements of the functions $\mathfrak{R}(z)$ and $\sqrt[n]{z}$ is one-to-one and, moreover, as we easily verify, invertibly continuous (Introduction, § 7). On the other hand, we have already established above a homeomorphic correspondence between the elements of the analytic function $\sqrt[n]{z}$ in the annulus $\mathrm{P}(0;0,r)$ and the points of the annulus $\mathrm{P}(0;0,1)$. In this way we obtain a homeomorphism between the function $\mathfrak{R}$ and the annulus $\mathrm{P}(0;0,1)$. As a result:

(10.1) *Between the points $\mathfrak{z}$ of the annulus $\mathrm{P}(0;0,1)$ and the elements of an analytic function $\mathfrak{R}$, arbitrarily continuable and n-valued in an annulus $\mathrm{P}(a;0,r)$, where $r\neq\infty$, a homeomorphic correspondence can be established in such a way that to the point $\mathfrak{z}$ there corresponds an element of the function with centre $z=a+r\mathfrak{z}^n$, if $a\neq\infty$, or with centre $1/r\mathfrak{z}^n$, if $a=\infty$.*

§ 11. Critical points. If $\mathfrak{R}$ is an analytic function in a region G, and a an arbitrary point of this region, then, in order that a branch (see § 4, p. 248) of the function $\mathfrak{R}$ exist in every neighbourhood of this point, it is necessary and sufficient that this point lie in the interior or on the boundary of the natural subregion. A point $a\in G$, lying in the interior or on the boundary of this subregion will be called an *ordinary point* of the function $\mathfrak{R}$ if it has a neighbourhood in which every branch of the function $\mathfrak{R}$ is arbitrarily continuable and therefore, by the monodromy theorem (theorem 6.2), is a meromorphic function; in the contrary case the ponit a will be called a *critical point* of the function $\mathfrak{R}$. As follows immediately from this definition, every ordinary point of an analytic function lies in its natural subregion, and hence by the same token, every point on the boundary of this subregion, belonging to the region G of the function, is its critical point. An ordinary point of the analytic function $\mathfrak{R}$ is called its *pole* if at least one branch of this function, arbitrarily continuable (and hence meromorphic) in a neighbourhood of this point, has a pole at this point.

The set of critical points of an analytic function in a region G is, obviously, closed in this region. The isolated points of this set are called *isolated critical points*.

(11.1) *If $\mathfrak{R}$ is an analytic function in a region G, then every element R of this function is continuable along every curve C emanating from its centre, lying in G and not containing critical points of the function.*

Therefore, in order that the function $\mathfrak{R}$ be arbitrarily continuable in its region G it is necessary and sufficient that G contain no critical points.

Proof. Let the curve C be given by the equation $z=z(t)$, where $a\leqslant t\leqslant b$. Let us denote by T the set of all points τ of the interval $[a,b]$, such that the element R is continuable along the arc $[a,\tau]$ of the curve C (*i. e.* along the curve $z=z(t)$, where $a\leqslant t\leqslant \tau$). Let t_0 be the upper bound of the set T. The point $z(t_0)$ therefore certainly lies in the interior or on the boundary of the natural subregion of the function; and because the curve C does not contain critical points, the point $z(t_0)$ has a neighbourhood K in which every element of the function $\mathfrak{R}$ is arbitrarily continuable. Let t_1 be an arbitrary point of the set T, such that the arc $[t_1,t_0]$ of the curve C is contained entirely in K, and let R_1 be the continuation of the element R along the arc $[a,t_1]$ of this curve. Continuing the element R_1 along the arc $[t_1,t_0]$, we therefore obtain a continuation of the element R along the entire arc $[a,t_0]$, from which it follows that $t_0\epsilon T$. In addition, we have $t_0=b$; for, in the contrary case, we could obtain in this way a continuation of the element R along an arc $[a,t_2]$ of the curve C, choosing the point t_2 such that $t_0<t_2<b$ and such that the arc $[t_0,t_2]$ of the curve C is contained in the circle K; hence, we should have $t_0<t_2\epsilon T$, which contradicts the definition of the point t_0. Consequently, $b=t_0\epsilon T$, and hence the element R is continuable along the entire curve C.

From theorem 11.1 it follows immediately that:

(11.2) *In order that a critical point of an analytic function be an isolated critical point, it is necessary and sufficient that it have an annular neighbourhood in which every branch of the function is arbitrarily continuable.*

(11.3) *In order that the set of critical points of a function $\mathfrak{R}$, analytic in a region G, be isolated, it is necessary and sufficient that every element of this function be continuable along every curve emanating from the centre of the element, contained in G and not containing points of a set E isolated and closed in the region G.*

If this condition is satisfied, then the function $\mathfrak{R}$ does not contain critical points outside the set E.

EXERCISES. 1. A function, analytic in a region G, can have at most a denumerable number of poles in this region.

2. Let G denote the plane minus the points 0 and ∞, and let $B\subset G$ be an arbitrary denumerable set. Construct an arbitrarily continuable analytic function in G which has a pole at every point of the set B.

(Such a function can be constructed for an arbitrary region G having a degree of connectivity $\geqslant 2$.)

§ 12. Algebraic critical points. Let a be a point lying in the interior or on the boundary of the natural region of an analytic function $\mathfrak{R}$. A number A (finite or infinite) is called the *limit* of the function $\mathfrak{R}$ at the point a, if to each number $\varepsilon>0$ there corresponds a number $\eta>0$ such that all the values of the function $\mathfrak{R}$ in the circle $\mathrm{K}(a;\eta)$ belong to the circle $\mathrm{K}(A;\varepsilon)$.

An isolated critical point a of an analytic function $\mathfrak{R}$ is called an *algebraic critical point*, if in some annular neighbourhood of the point a every branch of the function $\mathfrak{R}$ is finitely valued and has a limit at this point. Isolated critical points which are not algebraic are called *transcendental*.

(12.1) *In order that the critical point a of an analytic function $\mathfrak{R}$ be algebraic, it is necessary and sufficient that in some annular neighbourhood $\mathrm{P}(a;0,r)$ of this point every branch of the function $\mathfrak{R}$ have the form $\Phi(\sqrt[n]{z-a})$ if $a\neq\infty$, or the form $\Phi(1/\sqrt[n]{z})$ if $a=\infty$, where Φ is a function meromorphic in the circle $\mathrm{K}(0;r^{1/n})$.*

Proof. The sufficiency of the condition is evident. In order to prove the necessity of the condition, let us take for simplicity $a=0$, and let us note that, when the point 0 is an algebraic critical point of the function $\mathfrak{R}$, then by theorem 9.3 an annular neighbourhood $P(0;0,r)$ of it exists, in which every branch of the function $\mathfrak{R}$ has the form $\Phi(\sqrt[n]{z})$, where $\Phi(z)$ is a function meromorphic in $P(0;0,r^{1/n})$. Since the branch $\Phi(\sqrt[n]{z})$ has a limit at the point 0, the function $\Phi(z)$ also has a limit (finite or infinite) as $z\to 0$, and, after a suitable definition of its value at the point 0, it is also meromorphic at this point (cf. Chapter III, § 6, p. 145).

A function, analytic in the entire plane, finitely valued, and not having critical points other than algebraic points, is said to be *algebraic*. Since an isolated and closed set in the entire plane is finite, an algebraic function can have only a finite number of critical points. An analytic function (in the entire plane), which is not algebraic, is called *transcendental*.

The function $\sqrt[n]{z}$ has only two critical points: 0 and ∞, both algebraic. The function $\log z$ also has only two critical points: 0 and ∞, but both transcendental. The function $\sqrt{\log z}$ has three critical points: algebraic at the point 1 and transcendental at the points 0 and ∞.

EXERCISES. 1. Distinguish the poles and critical points (algebraic, transcendental) of the analytic functions $\arccos z$, $\arctan z$, $\log(\arctan z)$, $\sqrt{\log(\arctan z)}$.

2. Determine the critical points of the analytic function obtained by continuing the holomorphic function

$$F(z)=\frac{1}{2\pi i}\int_0^1\frac{f(\mathfrak{z})}{\mathfrak{z}-z}\,d\mathfrak{z},$$

where $f(z)$ is a function holomorphic in the entire open plane and not vanishing identically (cf. § 2, exercise 2).

§ 13. Auxiliary theorems of algebra. In the sequel we shall make use of several fundamental theorems of algebra concerning the resolution of polynomials into factors. We shall consider polynomials of the form

(13.1) $$T(z,w)=A_0(z)w^n+A_1(z)w^{n-1}+\ldots+A_n(z),$$

with coefficients $A_j(z)$ meromorphic in a fixed region G. If $A_0(z)$ does not vanish identically, then the polynomial (13.1) is called a *polynomial of the n-th degree*. If $T(z,w)=Q(z,w)\cdot P(z,w)$, where

$Q(z,w)$ and $P(z,w)$ are also polynomials of the form considered, then each of the polynomials $Q(z,w)$ and $P(z,w)$ will be called a *divisor* of the polynomial $T(z,w)$. If a polynomial of degree n has a divisor of positive degree less than n, then this polynomial is said to be *reducible*. Polynomials having a common divisor of positive degree are called *co-divisible*.

(13.2) *If for $z\in G$ we have identically*

(13.3) $$w^n+A_1(z)w^{n-1}+\ldots+A_n(z)$$
$$=[w^m+B_1(z)w^{m-1}+\ldots+B_m(z)][w^p+C_1(z)w^{p-1}+\ldots+C_p(z)],$$

where $A_j(z)$, $B_j(z)$ and $C_j(z)$ are functions meromorphic in the region G, then the coefficients $B_j(z)$ and $C_j(z)$ do not have poles other than the poles of the coefficients $A_j(z)$.

Proof. Let us assume that the point $a\neq\infty$ is a pole of one of the functions $B_j(z)$ or $C_j(z)$, but not a pole of any one of the coefficients $A_j(z)$. Let k and h denote the highest multiplicities[1]) of the point a, as a pole of the functions $B_j(z)$ and the functions $C_j(z)$, respectively. Let

$$B_j=\lim_{z\to a}(z-a)^kB_j(z),\qquad C_j=\lim_{z\to a}(z-a)^hC_j(z),$$

where, for symmetry, we have taken $B_0(z)=C_0(z)=1$. Multiplying both sides of the equation (13.3) by $(z-a)^{h+k}$, we obtain identically

$$(B_0w^m+B_1w^{m-1}+\ldots+B_m)(C_0w^p+C_1w^{p-1}+\ldots+C_p)=0,$$

which is impossible, however, since neither all the coefficients B_j, nor all the coefficients C_j vanish simultaneously. We reason similarly if $a=\infty$.

Let us now take into consideration two polynomials:

$$T_1(z,w)=A_0(z)w^n+A_1(z)w^{n-1}+\ldots+A_n(z),$$
$$T_2(z,w)=B_0(z)w^m+B_1(z)w^{m-1}+\ldots+B_m(z),$$

with coefficients meromorphic in a region G. Let us assume that $n\geqslant m>0$. Applying Euclid's algorithm (successive division) to these polynomials, we determine two finite sequences of polynomials, $Q_1,Q_2,\ldots,Q_{p-2}$ and $T_3,T_4,\ldots,T_p$, satisfying the following conditions:

[1]) If a function is holomorphic at a point, then by the *multiplicity* of the point, as a pole of this function, we mean the number 0.

(13.4) $T_1 = Q_1T_2 + T_3, \quad T_2 = Q_2T_3 + T_4, \quad \ldots, \quad T_{p-2} = Q_{p-2}T_{p-1} + T_p,$

(13.5) *the degrees of the polynomials $T_2, T_3, \ldots, T_p$ form a decreasing sequence, and T_p is of degree 0 in w, and hence is a function of the variable z only.*

It is evident that the polynomials Q_j and T_j are defined uniquely by these conditions; their coefficients are expressed rationally in terms of the coefficients of the polynomials $T_1(z,w)$ and $T_2(z,w)$, and hence are also meromorphic with respect to z in the region G; in particular, meromorphic in G is the "last remainder" $T_p = T_p(z)$. Let us denote by E the set (isolated and closed in G) of points at which at least one of the coefficients of the polynomials $T_j(z,w)$ and $Q_j(z,w)$ has a pole.

We verify immediately that if the last remainder $T_p(z)$ vanishes for some value $z = z_0 \epsilon G - E$, then denoting by w_0 a root of the equation $T_{p-1}(z_0, w) = 0$, we have directly from (13.4)

$$T_1(z_0, w_0) = T_2(z_0, w_0) = \ldots = T_{p-1}(z_0, w_0) = 0,$$

which means that the equations $T_1(z_0, w) = 0$ and $T_2(z_0, w) = 0$ have a common root $w = w_0$. Conversely, if w_0 is a common root of these two equations (for $z_0 \epsilon G - E$), then from (13.4) we find in a similar manner that $T_p(z_0) = 0$. On the other hand, either 1° the set of roots of the function $T_p(z)$ is isolated and closed in G, or 2° this function vanishes identically.

In case 2°, as we see directly from (13.4), the next to the last remainder $T_{p-1}(z,w)$, which is a polynomial of degree >0 in w, is a common divisor of the polynomials $T_1(z,w)$ and $T_2(z,w)$. Therefore, if the polynomial $T_{p-1}(z,w)$ is, in addition, of degree $<n$ (with respect to w), then the polynomial $T_1(z,w)$ is certainly reducible; however, if $T_{p-1}(z,w)$ is of degree n, then we verify that $p-1=2$, $n=m$, and the polynomials $T_1(z,w)$ as well as

$$T_2(z,w) = T_{p-1}(z,w)$$

differ, in view of this, at most by a factor depending only on z.

Summarizing the above results, we obtain the following theorem:

(13.6) *If $T_1(z,w)$ and $T_2(z,w)$ are polynomials of degree n and m in w, respectively, with coefficients meromorphic in z in a region G, and $n \geqslant m > 0$, then either 1° for no value of z, except for a set isolated and closed in G, do the polynomials $T_1(z,w)$ and $T_2(z,w)$ have common roots, or 2° these polynomials are co-divisible and hence either differ at most by a factor independent of w, or the polynomial $T_1(z,w)$ is reducible.*

Taking $T_1(z,w) = T(z,w)$ and $T_2(z,w) = T'_w(z,w)$ in the above theorem, and considering separately the trivial case when $T(z,w)$ is of degree 1 in w, we obtain the following corollary:

(13.7) *If $T(z,w)$ is a polynomial of positive degree in w, with coefficients meromorphic in z in a region G, then either for every value of z, except for an isolated set closed in G, the polynomial $T(z,w)$ has no multiple roots, or this polynomial is reducible.*

§ 14. Functions with algebraic critical points. In Chapter III, § 14, we considered the existence of functions $w = W(z)$ determined by an equation of the form $F(z,w) = 0$, where $F(z,w)$ is a holomorphic function of the variables z and w. However, these considerations had an exclusively local character and were limited to the proof of the existence, in the neighbourhood of a point z_0, of a holomorphic function $W(z)$ assuming at this point the value w_0 given initially. The concept of an analytic function permits one, in certain cases, to unite these local solutions into one whole.

Let us consider, for an arbitrary region G, an equation of the form

(14.1) $$T(z,w) = w^n + A_1(z)w^{n-1} + \ldots + A_n(z) = 0,$$

in which the coefficients $A_j(z)$ are functions meromorphic in G.

We shall say that an analytic element $R = \{W, z_0\}$, where $z_0 \epsilon G$, *satisfies* this equation, if in a neighbourhood of the point z_0 we have $T[z, W(z)] = 0$ identically. Concerning an analytic function $\mathfrak{R}$ in G, we shall say that it *satisfies* equation (14.1) if each of its elements satisfies this equation; it is evident that in order that the function $\mathfrak{R}$ satisfy equation (14.1) it is sufficient that it be satisfied by any element of this function.

(14.2) *Every equation of the form* (14.1) *is satisfied by at least one analytic function in the region G.*

Every function $\mathfrak{R}(z)$ satisfying equation (14.1) *is at most n-valued and has neither poles nor critical points other than the points z which are poles of the coefficients of the equation, or for which this equation has multiple roots; moreover, all the critical points of the function $\mathfrak{R}$ are algebraic.*

Proof. Let J denote the set of those points $z \in G$ which are either poles of the coefficients of equation (14.1), or for which this equation has multiple roots. In view of theorem 13.2, we may assume that the polynomial on the left side of equation (14.1) is irreducible, and hence, by theorem 13.7, that the set J is isolated. Let $z_0 \in G-J$.

By theorem 14.6, Chapter III, there exists, in view of the fundamental theorem of algebra (theorem 5.12, Chapter II), a holomorphic function $W(z)$ in the neighbourhood of the point z_0, such that $T[z, W(z)]=0$ identically. The element $\{W, z_0\}$, and therefore also the analytic function determined by it in the region G, then satisfy the equation (14.1).

On the other hand, let $\mathfrak{R}$ be an analytic function in G satisfying this equation. Let us consider an arbitrary point $a \in G-J$. There exist exactly n distinct values of w satisfying the equation $T(a,w)=0$. Let $w_1, w_2, \ldots, w_n$ be these values, and let

$$W_1(z), W_2(z), \ldots, W_n(z)$$

be holomorphic functions in a neighbourhood $K \subset G-J$ of the point a, such that $W_j(a)=w_j$ for $j=1,2,\ldots,n$, and $T[z, W_j(z)]=0$ identically in this neighbourhood of the point a (cf. Chapter III, theorem 14.6). Every analytic element $R \in \mathfrak{R}$ with centre $z \in K$ is of the form $\{W_j, z\}$, where $j=1,2,\ldots,n$, and therefore determines a single-valued analytic function in the circle K, identical with the corresponding holomorphic function $W_j(z)$. Consequently, all the critical points and poles of the function $\mathfrak{R}$ in G are contained in the isolated set J. Moreover, since every element of the function $\mathfrak{R}$ with centre $a \in G-J$ coincides with one of the elements $\{W_j, a\}$, this function is at most n-valued.

Let us next consider an arbitrary point b of the region G. Let $P \subset G$ be an annular neighbourhood of this point, containing neither critical points nor poles of the function $\mathfrak{R}$. Therefore every branch of this function in the neighbourhood P is arbitrarily continuable in this neighbourhood and, by theorem 9.3, is of the form $F(\sqrt[n]{z-b})$, where $F(\mathfrak{z})$ is a function holomorphic in the annular neighbourhood of the point 0 (if $b=\infty$, then $z-b$ should of course be replaced by $1/z$). We shall show that the function $F(\mathfrak{z})$ has at most a pole at the point 0. In fact, multiplying both sides of equation (14.1) by a suitable power of the binomial $z-b$, we can write this equation in the form

$$T_1(z,w)=(z-b)^q T(z,w)=B_0(z)w^n+B_1(z)w^{n-1}+\ldots+B_n(z)=0,$$

with coefficients $B_j(z)=(z-b)^q A_j(z)$, holomorphic in the neighbourhood of the point b and not vanishing simultaneously at this point (if b is not a pole of any one of the coefficients $A_j(z)$, then we take $q=0$). In a sufficiently small annular neighbourhood of the point 0 we have $T[b+\mathfrak{z}^n, F(\mathfrak{z})]=0$ identically, and hence also $T_1[b+\mathfrak{z}^n, F(\mathfrak{z})]=0$. Therefore, if w denotes an arbitrary finite limiting value of the function $F(\mathfrak{z})$ as $\mathfrak{z}\to 0$, then $T_1(b,w)=0$, *i. e.* w satisfies an equation of degree at most n, with coefficients not vanishing identically. The function $F(\mathfrak{z})$ therefore has only a finite number of distinct limiting values at the point 0 and by the Casorati-Weierstrass theorem (Chapter III, theorem 6.1) has a removable singularity or a pole at this point. Therefore every point $b \in G$ is an ordinary point or a critical algebraic point for the function $\mathfrak{R}$.

The converse of theorem (14.2) is the following theorem:

(14.3) *If $\mathfrak{R}$ is an n-valued analytic function in a region G, having only algebraic critical points, then there exists one and only one equation of the n-th degree of the form* (14.1) *which this function satisfies. The left side of this equation is an irreducible polynomial whose coefficients do not have poles other than the poles and critical points of the function $\mathfrak{R}$.*

Proof. Let L be the set of critical points of the function $\mathfrak{R}$. For every point $a \in G-L$, let us denote by $\{W_a^{(1)}, a\}, \{W_a^{(2)}, a\}, \ldots, \{W_a^{(n)}, a\}$ the elements of the function $\mathfrak{R}$ with centre a; let K_a be an arbitrarily fixed circle with centre a, contained in the region $G-L$ and simultaneously in the circles of all the elements $\{W_a^{(j)}, a\}$ for $j=1,2,\ldots,n$.

Furthermore, let $S(x_1, x_2, \ldots, x_n)$ be an arbitrary symmetric polynomial with respect to the variables $x_1, x_2, \ldots, x_n$, and let

$$\tilde{S}_a(z)=S[W_a^{(1)}(z), W_a^{(2)}(z), \ldots, W_a^{(n)}(z)],$$

for $z \in K_a$. We shall show that the meromorphic functions $\tilde{S}_a(z)$, defined in this manner in the circles K_a, jointly determine a meromorphic function in the entire region $G-L$.

To that end, it is sufficient to show that if $c \in K_a \cdot K_b$, where a and b are points of the region $G-L$, then $\tilde{S}_a(c)=\tilde{S}_b(c)$. In fact, since $c \in K_a$, the analytic elements $\{W_a^{(j)}, c\}$, where $j=1,2,\ldots,n$, constitute a set of n distinct elements of the function $\mathfrak{R}$; similarly, the elements $\{W_b^{(j)}, c\}$ also form a system of n distinct elements of the function $\mathfrak{R}$, having the same centre c. These two systems can therefore differ at most in the arrangement of their elements, and in view of the symmetry of the function $S(x_1,x_2,\dots,x_n)$ we have $\tilde{S}_a(c)=S[W_a^{(1)}(c),\dots,W_a^{(n)}(c)]=S[W_b^{(1)}(c),\dots,W_b^{(n)}(c)]=\tilde{S}_b(c)$.

Hence, taking $\tilde{S}(z)=\tilde{S}_z(z)$ for $z\in G-L$, we obtain a function $\tilde{S}(z)$, meromorphic in the region $G-L$, which in this region can have poles only at those points which are poles of the function $\mathfrak{R}$. Moreover, since the function $\mathfrak{R}$ does not have transcendental critical points, every branch of this function in an annular neighbourhood of the arbitrary point $z_0\in G$ has a definite limit at this point; hence, as is easily seen, there exists also a limit of the function $\tilde{S}(z)$ as $z\to z_0$. Consequently, none of the points $z_0\in G$ is an essential singularity of the function $\tilde{S}(z)$, and this function is meromorphic in the entire region G.

We shall apply the preceding result in particular to the fundamental symmetric functions $S^{(j)}(x_1,x_2,\dots,x_n)$, *i. e.* to the coefficients of the polynomial in w

$$(w-x_1)(w-x_2)\dots(w-x_n)=w^n+S^{(1)}(x_1,\dots,x_n)w^{n-1}+\dots+S^{(n)}(x_1,\dots,x_n).$$

These functions determine, in the manner established above, n functions meromorphic in the region G; let us denote them by $\tilde{S}^{(1)},\tilde{S}^{(2)},\dots,\tilde{S}^{(n)}$. We verify immediately that the function $\mathfrak{R}(z)$ satisfies the equation in w

$$w^n+\tilde{S}^{(1)}(z)w^{n-1}+\dots+\tilde{S}^{(n)}(z)=0 \tag{14.4}$$

with coefficients meromorphic in the region G and having no singular points in this region other than the poles and critical points, at most, of the function $\mathfrak{R}$.

It remains to show that equation (14.4) is the only equation of the form (14.1) which is satisfied by the function $\mathfrak{R}$, and that the polynomial on the left side of this equation is irreducible. To that end, let us assume that the function $\mathfrak{R}$ satisfies an equation of degree n in w, whose left side is the product of two polynomials $T_1(z,w)$ and $T_2(z,w)$ of degree $<n$ in w. The function $\mathfrak{R}$ would then satisfy at least one of the equations $T_1(z,w)=0$ or $T_2(z,w)=0$, which is obviously impossible, since this function is by hypothesis n-valued. Therefore, if the function $\mathfrak{R}$ satisfies equation (14.1), then the left side of this equation is an irreducible polynomial in w. At the same time, in view of theorem 13.6, it follows from this that at most one equation of the form (14.1) can exist, which the given function $\mathfrak{R}$ satisfies.

Theorem 14.2 can be completed as follows:

(14.5) *If the left side of equation* (14.1) *is an irreducible polynomial, then there exists only one analytic function in the region G which satisfies this equation.*

Proof. Let us suppose that there exist two analytic functions, a p-valued function $\mathfrak{R}_1$ and a q-valued function $\mathfrak{R}_2$, satisfying equation (14.1). By theorem 14.2 and 14.3, the function $\mathfrak{R}_1$ satisfies a certain equation of degree p in w

$$w^p+B_1(z)w^{p-1}+\dots+B_n(z)=0 \tag{14.6}$$

with coefficients meromorphic in the region G. The left sides of equations (14.1) and (14.6) are therefore, by theorem 13.6, co-divisible polynomials, and because $p<p+q\leqslant n$, the left side of equation (14.1) would be a reducible polynomial.

§ 15. Algebraic functions. In view of theorem 14.3, every algebraic function (cf. § 12) satisfies an equation of the form (14.1) with coefficients $A_j(z)$ which are meromorphic in the entire plane, and hence (cf. Chapter III, theorem 7.3) are rational functions. This equation can then be written in the form

$$B_0(z)w^n+B_1(z)w^{n-1}+\dots+B_n(z)=0, \tag{15.1}$$

where $B_j(z)$ are polynomials. Conversely, if an analytic function $\mathfrak{R}$ satisfies an equation of this form, then by theorem 14.2 this function is finitely valued and has in the entire plane no critical points other than algebraic; moreover, by the same theorem, the function $\mathfrak{R}$ has neither critical points nor poles other than those values of z for which either (a) equation (15.1) has multiple roots, or (b) the coefficient $B_0(z)$ vanishes, or (c) the value $z=\infty$ (this lost only if the degree of the coefficient $B_0(z)$ with respect to z is less than the degree of at least one of the remaining coefficients $B_1(z),\dots,B_n(z)$).

Summarizing, we obtain the following theorem, which contains a new definition of an algebraic function:

(15.2) *In order that an analytic function $\mathfrak{R}$ be algebraic, it is necessary and sufficient that it satisfy an equation of the form* (15.1) *with coefficients which are polynomials in the variable z.*

Equation (15.1) can be arranged according to powers of the variable z, and then the coefficients will be polynomials in w. From theorem 15.2 it therefore follows that

(15.3) *The inverse of an algebraic function is also an algebraic function.*

We also see that

(15.4) *Every algebraic function assumes each of its values only a finite number of times.*

EXERCISES. 1. Let $\mathfrak{F}$ be the analytic function which is obtained by continuing in the entire plane the holomorphic function $F(z)$ given in the half-plane $\mathcal{I}z>0$ by the formula (Chapter V, § 8, p.233)

$$F(z)=\int_0^z \frac{d\mathfrak{z}}{(\mathfrak{z}^2-a^2)^{2/3}} \qquad (a>0).$$

Verify that the only critical points of the function $\mathfrak{F}$ are $-a$, a, and ∞; these points are algebraic critical points.

Show that the inverse function $\mathfrak{F}^{-1}$ (see § 5) is a function meromorphic in the entire open plane (with a transcendental critical point at infinity); the open plane can be covered by a net of non-overlapping equilateral triangles such that: 1° $\mathfrak{F}^{-1}$ transforms each of the triangles of the net in a one-to-one manner either on the half-plane $\mathcal{I}z\geqslant 0$, or on the half-plane $\mathcal{I}z\leqslant 0$, 2° at the vertices of each of the triangles of the net the function $\mathfrak{F}^{-1}$ assumes, respectively, the values $-a$, a and ∞.

Deduce from this that the function $\mathfrak{F}^{-1}$, and hence the function $\mathfrak{F}$, is not algebraic (although the function $\mathfrak{F}$ has only a finite number of critical points, all of which are algebraic).

Similarly, investigate the analytic function $\mathfrak{W}$ which is obtained by continuing the holomorphic function $W(z)$ given in the circle $K(0;1)$ by the formula (Chapter V, § 8, exercise 3)

$$W(z)=\int_0^z \frac{d\mathfrak{z}}{(1-\mathfrak{z}^6)^{1/3}}.$$

[*Hint.* For the continuation of the functions F and W apply the reflection principle of Schwarz — with respect to the straight line and the circumference, respectively.]

2. If an analytic function $\mathfrak{P}$ has only two critical points and these points are algebraic, then the function $\mathfrak{P}$ is algebraic (the theorem ceases to be true when the number of critical points exceeds 2; cf. exercise 1).

3. The natural region of the algebraic function $w=W(z)$ satisfying the equation $w^3+w+z=0$ is the entire open plane, however, this function is not single-valued. (This example points out that in the monodromy theorem 6.3 the assumption that the natural region of the function is a simply connected region would not be sufficient) (Kierst).

[*Hint.* Notice that the left side of the equation is an irreducible polynomial and that for no z does the equation have a triple root.]

***§ 16. Riemann surfaces.** Every analytic function $\mathfrak{R}$ in the annulus $P(a;0,r)$ which:

1° is arbitrarily continuable in $P(a;0,r)$,

2° has an ordinary point or at most an algebraic critical point at the point a,

3° cannot be extended to any larger annulus $P(a;0,\varrho)$ and still preserve condition 1° (*i. e.* it is not a branch of any function, arbitrarily continuable in the annulus $P(a;0,\varrho)$ for $\varrho>r$)

will be called a *Riemannian element with centre a and radius r* (or with *annulus* $P(a;0,r)$).

Analytic functions arbitrarily continuable in the region which is obtained by removing the point a from the plane are also included among the *Riemannian elements with centre a.*

A Riemannian element will be called *smooth,* if it is a single-valued function, and *ramified* in the contrary case. If a Riemannian element is an n-valued function, then the number $n-1$ will be called the *order of ramification of this element.* From conditions 1° and 2° it follows that every Riemannian element is a finitely valued function.

Let $\mathfrak{R}$ be an analytic function in a region G. Let $\mathfrak{S}$ be any branch of this function, arbitrarily continuable in na annular neighbourhood $P(a;0,r)$ and having at most an algebraic critical point at the point a; this branch determines a Riemannian element with centre a and radius $r_0\geqslant r$. Every Riemannian element thus obtained will be said to be *determined* (at the point a) *by the function* $\mathfrak{R}$. Of course, the set of elements determined by a function $\mathfrak{R}$ at a point a may be infinite or empty; it is certainly non-empty if the point a belongs to the natural subregion of the function (it then contains, among others, smooth elements), or if the point a is an algebraic critical point of the function (for then every branch of the function in a certain annular neighbourhood of the point a determines a Riemannian element at this point); at ordinary points (§ 11) an analytic function determines only smooth elements.

If $\mathfrak{E}$ is a Riemannian element with annulus $P(a;0,r)$, then for every positive and finite number $\varrho \leqslant r$ the set of elements determined by $\mathfrak{E}$ at the points of the circle $K(a;\varrho)$ will be called a *neighbourhood of the element* $\mathfrak{E}$. With this definition of a neighbourhood, the set of all Riemannian elements can be considered as a an abstract space; we verify immediately that our definition of a neighbourhood satisfies the fundamental conditions I and II of § 3 of the Introduction.

Since the annulus of a Riemannian element does not contain critical points, every neighbourhood of the Riemannian element $\mathfrak{E}$ contains smooth elements exclusively, with the exception at most of the element $\mathfrak{E}$ itself, which may be ramified. Therefore, *in the space of all Riemannian elements, the ramified elements form an isolated and closed set.*

If $\mathfrak{E}$ is a smooth element with annulus $P(a;0,r)$, then the function $\mathfrak{E}$ is a single-valued function, and hence meromorphic in this annulus; moreover, since it has at most an algebraic critical point at the point a, by a suitable definition of the value of the function $\mathfrak{E}$ at this point we obtain a function $E(z)$ meromorphic in the entire circle $K(a;r)$. This circle is at the same time the circle of the analytic element $\{E,a\}$, which we associate with the smooth Riemannian element $\mathfrak{E}$. Conversely, to each analytic element $\{E,a\}$ with circle $K(a;r)$ there corresponds in this manner exactly one smooth Riemannian element with annulus $P(a;0,r)$, namely, the function $E(z)$, regarded as a single-valued analytic function in this annulus. The one-to-one correspondence between the smooth Riemannian elements and the analytic elements defined in this way is at the same time, as is easily verified, invertibly continuous and establishes a homeomorphism between the space of analytic elements and the set of smooth elements in the space of Riemannian elements. This remark enables us to identify, for simplicity of expression, the smooth Riemannian elements with the analytic elements corresponding to them. From this remark it also follows, in view of theorem 10.1, that every neighbourhood of a Riemannian element is homeomorphic with the circle; therefore theorem 10.1 may now be formulated as follows:

(16.1) *If $\mathfrak{U}$ is a neighbourhood of a Riemannian element $\mathfrak{E}_0$ with order of ramification $n-1$ and centre a, then a homeomorphic correspondence between the circle $K(0;1)$ and the neighbourhood $\mathfrak{U}$ can be established in such a way that to the point $\mathfrak{z} \in K(0;1)$ corresponds the element $\mathfrak{E} \in \mathfrak{U}$ with centre $z=a+r\mathfrak{z}^n$ (with centre $z=1/r\mathfrak{z}^n$, if $a=\infty$), where r is a positive constant depending on the neighbourhood $\mathfrak{U}$.*

The correspondence of the preceding theorem will be called *canonical.* The existence of such a correspondence enables one to establish an angular metric in the space of Riemannian elements. We shall prove at first the following corollary of theorem 16.1:

(16.2) *If $\mathfrak{U}_1$ and $\mathfrak{U}_2$ are, respectively, neighbourhoods of the elements $\mathfrak{E}_1$ and $\mathfrak{E}_2$, such that $\mathfrak{U}=\mathfrak{U}_1\cdot\mathfrak{U}_2\neq 0$, and Φ_1 and Φ_2 are the respective canonical mappings of these neighbourhoods onto the circle $K(0;1)$, then associating with each other the points $\mathfrak{z}_1 \in G_1=\Phi_1(\mathfrak{U})$ and $\mathfrak{z}_2 \in G_2=\Phi_2(\mathfrak{U})$, which in the mappings Φ_1 and Φ_2 correspond to the same element $\mathfrak{E} \in \mathfrak{U}$, we obtain a conformal correspondence between the open sets G_1 and G_2 in the circle $K(0;1)$.*

Proof. If $\mathfrak{E} \in \mathfrak{U}$, $\mathfrak{z}_1=\Phi_1(\mathfrak{E})$ and $\mathfrak{z}_2=\Phi_2(\mathfrak{E})$, then in agreement with 16.1 the centre of the element $\mathfrak{E}$ is the point

$$a_2+r_2\mathfrak{z}_2^m=a_1+r_1\mathfrak{z}_1^n, \tag{16.3}$$

where a_1 and a_2, respectively, are the centres of the elements $\mathfrak{E}_1$ and $\mathfrak{E}_2$, $n-1$ and $m-1$ the respective orders of ramification of these elements, finally r_1 and r_2 constants different from zero. Hence, denoting by $\Phi(\mathfrak{z}_1)$ the point $\mathfrak{z}_2 \in G_2$ associated with the point $\mathfrak{z}_1 \in G_1$ in the correspondence considered, we verify first directly that the function $\Phi(\mathfrak{z}_1)$ is uniquely invertible and continuous in G_1, and next, that the relation $\mathfrak{z}_2=\Phi(\mathfrak{z}_1)$ implies equation (16.3). With the view of proving that this correspondence is conformal, and hence that the function Φ is meromorphic in G_1, we shall distinguish three cases:

1° $m=1$. Then, from (16.3), we have directly

$$\mathfrak{z}_2=\Phi(\mathfrak{z}_1)=\frac{a_1-a_2+r_1\mathfrak{z}_1^n}{r_2}.$$

2° $\mathfrak{E}_1=\mathfrak{E}_2$. Then also $a_1=a_2$, $m=n$, and by (16.3) the function $\Phi(\mathfrak{z}_1)$, being continuous, has the form $\mathfrak{z}_2=\Phi(\mathfrak{z}_1)=k\mathfrak{z}_1$, where k is a constant coefficient ($|k|=(r_1/r_2)^{1/n}$).

3° $m>1$ and $\mathfrak{E}_1 \neq \mathfrak{E}_2$. Then the element $\mathfrak{E}_2$ is ramified; therefore it certainly does not belong to $\mathfrak{U}_1$, and hence to $\mathfrak{U}$. Consequently, the set G_2 does not contain the point $\mathfrak{z}_2=0$ and in the equation

$$[\Phi(\mathfrak{z}_1)]^m = \frac{a_1 - a_2 + r_1\mathfrak{z}_1^n}{r_2}$$

the left — and hence also the right — side does not vanish for any point $\mathfrak{z}_1 \epsilon G_1$. The function $\Phi(\mathfrak{z}_1)$, being continuous in G_1, is therefore a holomorphic branch of $\sqrt[m]{a_1 - a_2 + r_1\mathfrak{z}_1^n}/r_2^{1/m}$ in the open set G_1.

In the preceding proof we have assumed that the points a_1 and a_2 are different from ∞. The case when these points are at infinity is treated similarly.

Let us now consider in the space of Riemannian elements an arbitrary element $\mathfrak{E}_0$ and a curve $\mathfrak{L}$ emanating from $\mathfrak{E}_0$ (*i. e.* with initial point at $\mathfrak{E}_0$; see Introduction, § 12). Let $\mathfrak{U}$ be a neighbourhood containing the element $\mathfrak{E}_0$ (not necessarily as a centre, however), and let $\mathfrak{z}_0$ and L denote, respectively, the point and the curve into which $\mathfrak{E}_0$ and $\mathfrak{L}$ are transformed under a canonical mapping of the neighbourhood $\mathfrak{U}$ onto the circle $K(0;1)$; if the curve $\mathfrak{L}$ goes beyond the neighbourhood $\mathfrak{U}$, then we replace $\mathfrak{L}$ by a sufficiently small part of it emanating at $\mathfrak{E}_0$. We shall say that the curve $\mathfrak{L}$ emanating from the element $\mathfrak{E}_0$ has a *definite direction* if the curve L has a definite direction from the point $\mathfrak{z}_0$ (cf. Chapter I, § 13). In view of theorem 16.2 this property does not depend on the choice of the neighbourhood $\mathfrak{U}$. If $\mathfrak{L}_1$ and $\mathfrak{L}_2$ are curves emanating from the same element $\mathfrak{E}_0$ and having definite directions, and $\mathfrak{z}_0$, L_1 and L_2 are, respectively, the images of the element $\mathfrak{E}_0$ and the curves $\mathfrak{L}_1$ and $\mathfrak{L}_2$, under a canonical mapping of an arbitrary neighbourhood $\mathfrak{U}$ containing $\mathfrak{E}_0$, onto the circle $K(0;1)$, then by the *angle* between the curves $\mathfrak{L}_1$ and $\mathfrak{L}_2$ at $\mathfrak{E}_0$ we shall mean the angle between the curves L_1 and L_2 at the point $\mathfrak{z}_0$. By theorem 16.2, we see again that this angle depends only on the curves $\mathfrak{L}_1$ and $\mathfrak{L}_2$, and not on the choice of the neighbourhood $\mathfrak{U}$. Establishing in this manner a measure of an angle in the space of Riemannian elements, we obtain from theorem 16.2 the fundamental metric property of canonical mappings of neighbourhoods:

(16.4) *A canonical mapping of a neighbourhood of a Riemannian element onto the circle* $K(0;1)$ *maps this neighbourhood in a homeomorphic manner, preserving the angle between curves.*

The set of Riemannian elements determined by an analytic function $\mathfrak{R}$ (in the entire plane) is called the *Riemann surface* of this function. This set, as we see immediately, is simultaneously a continuum and a region (*i. e.* it is open, closed and connected). We also easily prove that the Riemann surfaces of analytic functions coincide with the components of the space of all Riemannian elements.

If $\mathfrak{E}_0$ is an element of the Riemann surface of an analytic function $\mathfrak{R}$, then each of its neighbourhoods is contained entirely in this surface; in fact, if a Riemannian element $\mathfrak{E}$ belongs to some neighbourhood of the element $\mathfrak{E}_0$, then — except at most for the case when $\mathfrak{E}=\mathfrak{E}_0$ — the element $\mathfrak{E}$ is a smooth element and the analytic element corresponding to it belongs to $\mathfrak{R}$. A Riemann surface can therefore be considered as an abstract connected space with a fixed angular measure and system of neighbourhoods, which by theorem 16.4 can be mapped in a homeomorphic manner and with a preservation of angles onto the circle $K(0;1)$. One may propose the problem whether every abstract space having these properties can be considered as a Riemann surface, or as a region on a Riemann surface. The precise and afirmative solution of this problem is the essential part of the famous theorem "on uniformization". The scope of this book does not permit the consideration of these matters which belong to the deeper and more beautiful results of the Theory of Functions[1]).

As in the space of all Riemannian elements, the ramified elements on the Riemann surface of an arbitrary analytic function $\mathfrak{R}$ form a closed and isolated set, while the smooth elements form a set homeomorphic with the function $\mathfrak{R}$ itself. Therefore by theorem 8.1 every Riemann surface contains a denumerable and everywhere dense set of smooth elements. Using this fact, we shall show that *the set of all ramified elements of a Riemann surface is at most denumerable*[2]).

First of all we note that the set of all ramified elements with a common centre is at most denumerable. To see this it is enough to select from a neighbourhood of each of these elements a smooth

[1]) The reader will find a detailed discussion in the work of H. Weyl, *Die Idee der Riemannschen Fläche*, 2nd ed., Leipzig-Berlin 1923; see also: T. Radó, *Über den Begriff der Riemannschen Fläche*, Acta Litt. Scient. Szeged **2** (1925), pp. 101-121.

[2]) This theorem is used in the triangulation of a Riemann surface (Weyl, l. c.).

element belonging to the denumerable and everywhere dense set mentioned above and observe that this selected elements are all different. In particular, therefore, the set of all ramified elements with centre at the point ∞ is at most denumerable. It remains to show that the set of all ramified elements with centres different from ∞ is also at most denumerable.

To show this, let us associate with every such element $\mathfrak{E}$ of the Riemann surface that neighbourhood of $\mathfrak{E}$ whose radius is half the radius of the annulus of $\mathfrak{E}$. If we show that all this neighbourhoods are disjoint, our assertion will be established by considering, as above, a denumerable and everywhere dense set of smooth elements. Let us therefore suppose that the neighbourhoods of two ramified elements $\mathfrak{E}_1$ and $\mathfrak{E}_2$, $\mathfrak{E}_1 \neq \mathfrak{E}_2$, have an element in common. Let $\mathrm{K}(a_1, r_1)$ and $\mathrm{K}(a_2, r_2)$ be the circles of the neighbourhoods under consideration and let *e. g.* $r_1 \geqslant r_2$. The annuli of $\mathfrak{E}_1$ and $\mathfrak{E}_2$ are therefore $\mathrm{P}(a_1; 0, 2r_1)$ and $\mathrm{P}(a_2; 0, 2r_2)$, and it is immediately seen that $\mathrm{K}(a_1, 2r_1)$ contains the point a_2. Since neighbourhoods of $\mathfrak{E}_1$ and $\mathfrak{E}_2$ have elements in common, it follows that $\mathfrak{E}_2$ belongs to a neighbourhood of $\mathfrak{E}_1$. But a neighbourhood cannot contain more than one ramified element. Hence $\mathfrak{E}_1 = \mathfrak{E}_2$, contrary to assumption, and this proves our assertion.

Therefore, identifying, as above, the smooth Riemannian elements with the analytic elements corresponding to them, we may say that the Riemann surface of an analytic function $\mathfrak{R}$ results from the closure of the set of analytic elements of the function $\mathfrak{R}$ by adding to it at most a denumerable number of ramified elements. For example, the Riemann surface of the function $\sqrt[k]{z}$, where k is a positive integer, results from the addition to this function of two ramified elements of order $k-1$ with centres 0 and ∞; the surface obtained in this way is homeomorphic with the plane (closed), while the function $\sqrt[k]{z}$ itself is homeomorphic with the plane minus two points; the function $\log z$ is identical with its Riemann surface and homeomorphic with the open plane (the critical points 0 and ∞ of the function $\log z$ are transcendental and the function does not determine Riemannian elements at these points); the Riemann surface of the function $\sqrt{\log z}$ results from the addition to this function of one ramified element of the first order with centre 1, and it is also homeomorphic with the open plane.

It is easy to see that *in order that an analytic function be algebraic, it is necessary and sufficient that its Riemann surface be a compact set.* In this way we obtain a definition of an algebraic function based on an elementary topological property of its Riemann surface.

Very frequently by a **Riemann surface** of an analytic function we mean a certain intuitive geometrical model of the behaviour of the function, homeomorphic with the Riemann surface understood in the sense of the formal abstract definition given above. We give here constructions of such models in a few simple cases.

1° **The function** $\log z$. We cut the plane along the positive real axis $[0, +\infty]$ (Fig. 25) and consider the sequence $\{a_n\}_{n=\ldots,-1,0,1,\ldots}$ — infinite in both directions — of identical replicas of the plane thus cut. We next join all these replicas in this manner: we remove the points 0 and ∞, and we paste the negative boundary of the cut of the replica a_n (where $n = \ldots, -1, 0, 1, 2, \ldots$) with the positive boundary of the cut of the replica a_{n+1}. The model obtained (Fig. 26) can be considered as a pictorial representation of the Riemann surface for the function $\log z$, illustrating the behaviour of this function. Indeed, exactly as we always pass from one element of $\log z$ to another when making a "circuit" about the point 0, so upon making a circuit about 0 on the surface obtained, we pass from one point of the surface to another, situated "above" or "below", according to the sense of the circuit.

2° **The function** $\sqrt{z}$. When making a circuit about the point 0, *e. g.* along a circumference with centre 0, we either return to the same element, or we pass to another, according as the number of circuits is even or odd. To represent this locus we construct a model of the Riemann surface $\sqrt{z}$, joining two replicas a_1, a_2 of the plane, cut as in example 1°, in such a way that the negative boundary of the cut of the replica a_1 is "pasted" to the positive boundary of the cut of the replica a_2, and the negative boundary of the cut a_2 with the positive boundary of the cut a_1; the points 0 and ∞, which are the (algebraic) critical points of the function $\sqrt{z}$, are not thereby removed (Fig. 27). With the exception of the boundary points of the cuts, none of the points of the planes a_1 and a_2 considered are "pasted" together. Of course this "pasting" cannot be realized in a three-dimensional space; nevertheless the geometrical (topological) character of the model obtained and its relation of the function $\sqrt{z}$ is clear. In particular, we see that if the replicas a_1, a_2 are "pasted" along corresponding (and not opposite) boundaries, the model obtained will be homeomorphic with the closed plane, *i. e.* with the surface of a sphere.

3° **The function** $\sqrt{z(z-1)}$. By a continuation of an element of this function along a circumference enclosing only one of the critical points 0 or 1, we pass — after one circuit — to another element. On the other hand, by a continuation along a circumference enclosing both points 0 and 1 we always return to the initial element (Chapter I, § 11, exercise 6). To depict this behaviour let us consider two identical replicas b_1 and b_2 of the plane cut along the segment $[0,1]$. We "paste" (retaining the points 0 and 1) the negative boundary of the cut b_1 with the positive boundary of the cut b_2, and the negative boundary of the cut b_2 with the positive boundary of the cut b_1. The model obtained (Fig. 28) is the model of the Riemann surface of the function $\sqrt{z(z-1)}$ and is homeomorphic — as we easily verify — with the surface which would be obtained by "pasting" the replicas considered along the corresponding (and not the opposite) boundaries of the cuts. This surface, in turn, is homeomorphic with the surface which would be obtained by "pasting" two identical spherical surfaces along identical circular openings, *i. e.* with a surface which is obviously homeomorphic directly with a spherical surface. Thus we see that the Riemann surface $\sqrt{z(z-1)}$ is, as in the case of the Riemann surface $\sqrt{z}$, homeomorphic with a spherical surface.

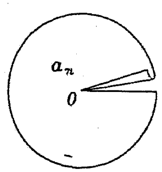

Fig. 25.

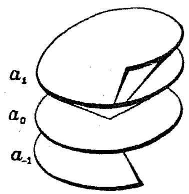

Fig. 26.

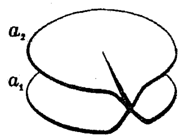

Fig. 27.

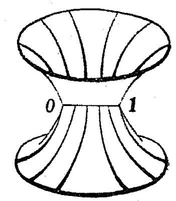

Fig. 28.

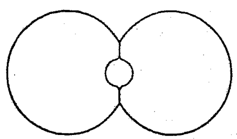

Fig. 29.

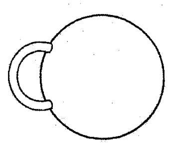

Fig. 30.

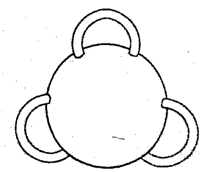

Fig. 31.

4° **The function** $\sqrt{z(z-1)(z-2)}$. We consider two identical replicas of the plane, c_1 and c_2, cut along the segments $[0,1]$ and $[2, +\infty]$. These replicas, as in the preceding constructions, are "pasted" along the opposite boundaries of the corresponding cuts. We obtain a model of the behaviour of the analytic function $\sqrt{z(z-1)(z-2)}$, which is a model of its Riemann surface. By means of transformations analogous to those used in example 3° we verify that the Riemann surface obtained is identical with the surface which we should obtain by "pasting" two spherical surfaces (which we must here imagine to be made of plastic material) along two circular openings (Fig. 29); in other words, the surface considered is homeomorphic with a spherical surface with one ear (handle) (Fig. 30), or — equivalently — with the torus, *i. e.* the surface generated by revolving a circumference about an axis lying in the plane of this circumference but not cutting it.

In this way one can construct generally a model of the Riemann surface for the function $\sqrt{(z-z_1)(z-z_2)\ldots(z-z_n)}$, where $z_1, z_2, \ldots, z_n$ is an arbitrary

system of n distinct points. Such a surface for $n=2k-1$ and $n=2k$ (where k is a positive integer) is homeomorphic with a spherical surface having k ears (handles) (Fig. 30 and 31), or a torus with k "openings". Instead of a square root one could consider here, of course, a root with an arbitrary integral index. It can also be proved, generally, (and this does not present great difficulties) that the Riemann surfaces of all algebraic functions reduce to these topological types.