5. NORMAL FAMILIES

In Chap. 3, Sec. 1 we have already familiarized the reader with the idea of regarding a function as a point in a space. In principle there is thus no difference between a set of points and a set of functions. In order to make a clear distinction we shall nevertheless prefer to speak of *families* of functions, and usually we assume that all functions in a family are defined on the same set.

We are primarily interested in families of analytic functions, defined in a fixed region. Important examples are the families of bounded analytic functions, of functions which do not take the same value twice, etc. The aim is to study convergence properties within such families.

5.1. Equicontinuity. Although analytic functions are our main concern, it is expedient to choose a more general starting point. It turns out that our basic theorems are valid, and equally easy to prove, for families of functions with values in any metric space.

As a basic assumption we shall let $\mathfrak{F}$ denote a family of functions f, defined in a fixed region Ω of the complex plane, and with values in a *metric space* S. As in Chap. 3, Sec. 1, the distance function in S will be denoted by d.

We are interested in the convergence of sequences $\{f_n\}$ formed by functions in $\mathfrak{F}$. There is no particular reason to expect a sequence $\{f_n\}$ to be convergent; on the contrary, it is perhaps more likely that we run into the opposite extreme of a sequence that does not possess a single convergent subsequence. In many situations the latter possibility is a serious disadvantage, and the purpose of the considerations that follow is to find conditions which rule out this kind of behavior.

Let us review the definition of continuity of a function f with values in a metric space. By definition, f is continuous at z_0 if to every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(f(z),f(z_0)) < \varepsilon$ as soon as $|z - z_0| < \delta$. We recall that f is said to be uniformly continuous if we can choose δ independent of z_0. But in the case of a family of functions there is another relevant kind of uniformity, namely, whether we can choose δ independent of f. We choose to require both, and are thus led to the following definition:

Definition 1. *The functions in a family $\mathfrak{F}$ are said to be equicontinuous on a set $E \subset \Omega$ if and only if, for each $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(f(z),f(z_0)) < \varepsilon$ whenever $|z - z_0| < \delta$ and $z_0,z \in E$, simultaneously for all functions $f \in \mathfrak{F}$.*

Observe that, with this definition, each f in an equicontinuous family is itself uniformly continuous on E.

We return now to the question of convergent subsequences. Our second definition serves to characterize families with a regular behavior:

Definition 2. *A family $\mathfrak{F}$ is said to be normal in Ω if every sequence $\{f_n\}$ of functions $f_n \in \mathfrak{F}$ contains a subsequence which converges uniformly on every compact subset of Ω.*

This definition does *not* require the limit functions of the convergent subsequences to be members of $\mathfrak{F}$.

5.2. Normality and Compactness. The reader cannot fail to have noticed the close similarity between normality and the Bolzano-Weierstrass property (Chap. 3, Theorem 7). To make it more than a similarity we need to define a distance on the space of functions on Ω with values in S, and convergence with respect to this distance function should mean precisely the same as uniform convergence on compact sets.

For this purpose we need, first of all, an *exhaustion* of Ω by an increasing sequence of compact sets $E_k \subset \Omega$. By this we mean that every compact subset E of Ω shall be contained in an E_k. The construction is possible in many ways: To be specific, let E_k consist of all points in Ω at distance $\leqq k$ from the origin, and at distance $\geqq 1/k$ from the boundary $\partial\Omega$. It is clear that each E_k is bounded and closed, hence compact. Any compact set $E \subset \Omega$ is bounded and at positive distance from $\partial\Omega$; therefore it is contained in an E_k.

Let f and g be any two functions on Ω with values in S. We shall define a distance $\rho(f,g)$ between these *functions*, not to be confused with the distances $d(f(z),g(z))$ between their values. To do so we first replace d by the distance function

$$\delta(a,b) = \frac{d(a,b)}{1 + d(a,b)}$$

which also satisfies the triangle inequality and has the advantage of being bounded (Chap. 3, Sec. 1.2, Ex. 1). Next, we set

$$\delta_k(f,g) = \sup_{z \in E_k} \delta(f(z),g(z))$$

which may be described as the distance between f and g on E_k. Finally, we agree on the definition

$$\rho(f,g) = \sum_{k=1}^{\infty} \delta_k(f,g)2^{-k}. \tag{65}$$

It is trivial to verify that $\rho(f,g)$ is finite and satisfies all the conditions for a distance function (Chap. 3, Sec. 1.2).

The distance $\rho(f,g)$ has the property we were looking for. Suppose first that $f_n \to f$ in the sense of the ρ-distance. For large n we have then $\rho(f_n,f) < \varepsilon$ and consequently, by (65), $\delta_k(f_n,f) < 2^k\varepsilon$. But this implies that $f_n \to f$ uniformly on E_k, first with respect to the δ-metric, but hence also with respect to the d-metric. Since every compact E is contained in an E_k it follows that the convergence is uniform on E.

Conversely, suppose that f_n converges uniformly to f on every compact set. Then $\delta_k(f_n,f) \to 0$ for every k, and because the series $\Sigma\delta_k(f_n,f)2^{-k}$ has a convergent majorant with terms independent of n it follows readily (as in Weierstrass's M test) that $\rho(f_n,f) \to 0$.

We have shown that convergence with respect to the distance ρ is equivalent to convergence on compact sets. So far we did not assume S to be complete, but if it is, it follows easily that the space of all functions with values in S is complete as a metric space with the distance ρ.

It can be said with some justification that the metric we have introduced is arbitrary and artificial. However, from what we have proved it follows that the open sets are independent of the choices involved in the construction. In other words, the *topology* has an intrinsic meaning, tailored to the needs of the theory of analytic functions.

We now recall the Bolzano-Weierstrass theorem, according to which a metric space is compact if and only if every infinite sequence has a convergent subsequence (Chap. 3, Theorem 7). The theorem is applied to the set $\mathfrak{F}$, equipped with the distance ρ, and we conclude that $\mathfrak{F}$ is compact if and only if $\mathfrak{F}$ is normal, *and* if the limit functions are themselves in $\mathfrak{F}$. On the other hand, if $\mathfrak{F}$ is normal, so is its closure $\mathfrak{F}^-$. Therefore we obtain the following characterization of normal families:

Theorem 12. *A family $\mathfrak{F}$ is normal if and only if its closure $\mathfrak{F}^-$ with respect to the distance function* (65) *is compact.*

It is also customary to say that $\mathfrak{F}$ is *relatively compact* if $\mathfrak{F}^-$ is compact. Thus, normal and relatively compact families are the same.

We shall now relate the notion of normal families to total boundedness. If $\mathfrak{F}$ is normal, then $\mathfrak{F}^-$ is compact, and according to Chap. 3, Theorem 6, $\mathfrak{F}^-$ is totally bounded, and so is consequently $\mathfrak{F}$ (see the footnote on p. 61). By definition, this means that to every $\epsilon > 0$ there exist a finite number of functions $f_1, \ldots, f_n \in \mathfrak{F}$ such that every $f \in \mathfrak{F}$ satisfies $\rho(f,f_j) > \varepsilon$ for some f_j. Conversely, if $\mathfrak{F}$ is totally bounded, so is $\mathfrak{F}^-$. If S is known to be complete, then $\mathfrak{F}^-$ is also complete, and hence compact. In other words, if S is complete, then $\mathfrak{F}$ is normal if and only if it is totally bounded.

The following theorem serves to state the condition of total boundedness in terms of the original metric on S rather than in terms of the auxiliary metric ρ.

Theorem 13. *The family $\mathfrak{F}$ is totally bounded if and only if to every compact set $E \subset \Omega$ and every $\varepsilon > 0$ it is possible to find $f_1, \ldots, f_n \in \mathfrak{F}$ such that every $f \in \mathfrak{F}$ satisfies $d(f,f_j) < \varepsilon$ on E for some f_j.*

If $\mathfrak{F}$ is totally bounded there exist $f_1, \ldots, f_n$ such that, for any $f \in \mathfrak{F}$, $\rho(f,f_j) < \varepsilon$ for some f_j. By (65) this implies $\delta_k(f,f_j) < 2^k\varepsilon$, or $\delta(f,f_j) < 2^k\varepsilon$ on E_k. If we fix k beforehand, we can thus make $\delta(f,f_j)$ arbitrarily small on E_k, and the same is then true of $d(f,f_j)$. This proves that the condition is necessary.

To prove the sufficiency we choose k_0 so that $2^{-k_0} < \varepsilon/2$. By assumption we can find $f_1, \ldots, f_n$ such that any $f \in \mathfrak{F}$ satisfies one of the inequalities $\delta(f,f_j) \leqq d(f,f_j) < \varepsilon/2k_0$ on E_{k_0}. It follows that $\delta_k(f,f_j) < \varepsilon/2k_0$ for $k \leqq k_0$, while trivially $\delta_k(f,f_j) < 1$ for $k > k_0$. From (65) we obtain

$$\rho(f,f_j) < k_0(\varepsilon/2k_0) + 2^{-k_0-1} + 2^{-k_0-2} + \cdots = \varepsilon/2 + 2^{-k_0} < \varepsilon,$$

which is precisely what we wanted to prove.

5.3. Arzela's Theorem. We shall now study the relationship between Definition 1 and Definition 2. The connection is established by a famous and extremely useful theorem known as *Arzela's theorem* (or the Arzela-Ascoli theorem).

Theorem 14. *A family $\mathfrak{F}$ of continuous functions with values in a metric space S is normal in the region Ω of the complex plane if and only if*

(i) *$\mathfrak{F}$ is equicontinuous on every compact set $E \subset \Omega$;*

(ii) *for any $z \in \Omega$ the values $f(z)$, $f \in \mathfrak{F}$, lie in a compact subset of S.*

We give two proofs of the necessity of (i). Assume that $\mathfrak{F}$ is normal and determine $f_1, \ldots, f_n$ as in Theorem 13. Because each of these functions is uniformly continuous on E we can find a $\delta > 0$ such that $d(f_j(z),f_j(z_0)) < \varepsilon$ for $z,z_0 \in E$, $|z - z_0| < \delta$, $j = 1, \ldots, n$. For any given $f \in \mathfrak{F}$ and corresponding f_j we obtain

$$d(f(z),f(z_0)) \leqq d(f(z),f_j(z)) + d(f_j(z),f_j(z_0)) + d(f_j(z_0),f(z_0)) < 3\varepsilon$$

and (i) is proved.

Less elegantly, but without use of Theorem 13, a proof can be given as follows: If $\mathfrak{F}$ fails to be equicontinuous on E there exists an $\varepsilon > 0$, sequences of points $z_n,z'_n \in E$, and functions $f_n \in \mathfrak{F}$ such that $|z_n - z'_n| \to 0$ while $d(f_n(z_n),f_n(z'_n)) \geqq \varepsilon$ for all n. Because E is compact we can choose subsequences of $\{z_n\}$ and $\{z'_n\}$ which converge to a common limit $z'' \in E$, and because $\mathfrak{F}$ is normal there exists a subsequence of $\{f_n\}$ which converges uniformly on E. It is clear that we may choose all three subsequences to have the same subscripts n_k. The limit function f of $\{f_{n_k}\}$ is uniformly continuous on E. Hence we can find k such that the distances from $f_{n_k}(z_{n_k})$ to $f(z_{n_k})$, from $f(z_{n_k})$ to $f(z'_{n_k})$, and from $f(z'_{n_k})$ to $f_{n_k}(z'_{n_k})$ are all $< \varepsilon/3$. It follows that $d(f_{n_k}(z_{n_k}),f_{n_k}(z'_{n_k})) < \varepsilon$, contrary to the assumption that $d(f_n(z_n),f_n(z'_n)) \geqq \varepsilon$ for all n.

To prove the necessity of (ii) we show that the closure of the set formed by the values $f(z)$, $f \in \mathfrak{F}$, is compact. Let $\{w_n\}$ be a sequence in this closure. To each w_n we determine $f_n \in \mathfrak{F}$ so that $d(f_n(z),w_n) < 1/n$. By normality there exists a convergent subsequence $\{f_{n_k}(z)\}$, and the sequence $\{w_{n_k}\}$ converges to the same value.

The sufficiency of (i) together with (ii) is proved by Cantor's famous diagonal process. We observe first that there exists an everywhere dense sequence of points ζ_k in Ω, for instance the points with rational coordinates. From the sequence $\{f_n\}$ we are going to extract a subsequence which converges at all points ζ_k. To find a subsequence which converges at one given point is always possible because of condition (ii). We can therefore find an array of subscripts

$$\begin{array}{c} n_{11} < n_{12} < \cdots < n_{1j} < \cdots \\ n_{21} < n_{22} < \cdots < n_{2j} < \cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ n_{k1} < n_{k2} < \cdots < n_{kj} < \cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \end{array} \tag{66}$$

such that each row is contained in the preceding one, and such that $\lim_{j\to\infty} f_{n_{kj}}(\zeta_k)$ exists for all k. The diagonal sequence $\{n_{jj}\}$ is strictly increasing, and it is ultimately a subsequence of each row in (66). Hence $\{f_{n_{jj}}\}$ is a subsequence of $\{f_n\}$ which converges at all points ζ_k. For convenience we replace the notation n_{jj} by n_j.

Consider now a compact set $E \subset \Omega$ and assume that $\mathfrak{F}$ is equicontinuous on E. We shall show that $\{f_{n_j}\}$ converges uniformly on E. Given $\varepsilon > 0$ we choose $\delta > 0$ such that, for $z,z' \in E$ and $f \in \mathfrak{F}$, $|z - z'| < \delta$ implies $d(f(z),f(z')) < \varepsilon/3$. Because E is compact, it can be covered by a finite number of $\delta/2$-neighborhoods. We select a point ζ_k from each of these neighborhoods. There exists an i_0 such that $i,j > i_0$ implies $d(f_{n_i}(\zeta_k),f_{n_j}(\zeta_k)) < \varepsilon/3$ for all these ζ_k. For each $z \in E$ one of the ζ_k is within distance δ from z; hence $d(f_{n_i}(z),f_{n_i}(\zeta_k)) < \varepsilon/3$, $d(f_{n_j}(z), f_{n_j}(\zeta_k)) < \varepsilon/3$. The three inequalities yield $d(f_{n_i}(z),f_{n_j}(z)) < \varepsilon$. Because all values $f(z)$ belong to a compact and consequently complete subset of S it follows that $\{f_{n_j}\}$ is uniformly convergent on E.

5.4. Families of Analytic Functions. Analytic functions have their values in $\mathbf{C}$, the finite complex plane. In order to apply the preceding considerations to families of analytic functions it is therefore natural to choose $S = \mathbf{C}$ with the usual euclidean distance.

The compact subsets of $\mathbf{C}$ are the bounded and closed sets. For this reason condition (ii) in Arzela's theorem is fulfilled if and only if the values $f(z)$ are bounded for each $z \in \Omega$, with a bound that may depend on z. Suppose now that condition (i) is also satisfied. For a given $z_0 \in \Omega$ determine ρ so that the closed disk $|z - z_0| \leqq \rho$ is contained in Ω. Then $\mathfrak{F}$, the given family of functions, is equicontinuous on the closed disk. If in the definition of equicontinuity $\delta(< \rho)$ corresponds to ε, and if $|f(z_0)| \leqq M$ for all $f \in \mathfrak{F}$, then $|f(z)| \leqq M + \varepsilon$ in $|z - z_0| < \delta$. Because any compact set can be covered by a finite number of neighborhoods with this property, it follows that the functions are uniformly bounded on every compact set, the bound depending on the set. According to Arzela's theorem this is true for all normal families of complex-valued functions.

For analytic functions this condition is also sufficient.

Theorem 15. *A family $\mathfrak{F}$ of analytic functions is normal with respect to* $\mathbf{C}$ *if and only if the functions in $\mathfrak{F}$ are uniformly bounded on every compact set.*

To prove the sufficiency we prove equicontinuity. Let C be the boundary of a closed disk in Ω, of radius r. If z, z_0 are inside C we obtain by Cauchy's integral theorem

$$f(z) - f(z_0) = \frac{1}{2\pi i} \int_C \left(\frac{1}{\zeta - z} - \frac{1}{\zeta - z_0} \right) f(\zeta)\, d\zeta$$

$$= \frac{z - z_0}{2\pi i} \int_C \frac{f(\zeta)\, d\zeta}{(\zeta - z)(\zeta - z_0)}.$$

If $|f| \leqq M$ on C, and if we restrict z and z_0 to the smaller concentric disk of radius $r/2$, it follows that

$$|f(z) - f(z_0)| \leqq \frac{4M|z - z_0|}{r}. \tag{67}$$

This proves equicontinuity on the smaller disk.

Let E be a compact set in Ω. Each point of E is the center of a disk with radius r, as above. The open disks of radius $r/4$ form an open covering of E. We select a finite subcovering and denote the corresponding centers, radii, and bounds by ζ_k, r_k, M_k; let r be the smallest of the r_k and M the largest of the M_k. For a given $\varepsilon > 0$ let δ be the smaller of $r/4$ and $\varepsilon r/4M$. If $|z - z_0| < \delta$ and $|z_0 - \zeta_k| < r_k/4$ it follows that $|z - \zeta_k| < \delta + r_k/4 \leqq r_k/2$. Hence (67) is applicable and we find $|f(z) - f(z_0)| \leqq 4M_k\delta/r_k \leqq 4M\delta/r \leqq \varepsilon$ as desired.

In view of Theorem 15 we may abandon the term "normal with respect to C" which has no historic justification. If a family has the property of the theorem, we say instead that it is *locally bounded.* Indeed, if the family is bounded in a neighborhood of each point, then it is obviously bounded on every compact set. The theorem tells us that every sequence has a subsequence which converges uniformly on compact sets if and only if it is locally bounded.

An interesting feature is that local boundedness is inherited by the derivatives.

Theorem 16. *A locally bounded family of analytic functions has locally bounded derivatives.*

This follows at once by the Cauchy representation of the derivative. If C is the boundary of a closed disk in Ω, of radius r, then

$$f'(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)\, d\zeta}{(\zeta - z)^2}.$$

Hence $|f'(z)| \leqq 4M/r$ in the concentric disk of radius $r/2$ (M is the bound of $|f|$ on C). We see that the f' are indeed locally bounded.

What is true of the first derivatives is of course also true of higher derivatives.

5.5. The Classical Definition. If a sequence tends to ∞ there is no great scattering of values, and it may well be argued that for the purposes of normal families such a sequence should be regarded as convergent. This is the classical point of view, and we shall restyle our definition to conform with traditional usage.

Definition 3. *A family of analytic functions in a region Ω is said to be normal if every sequence contains either a subsequence that converges uniformly on every compact set $E \subset \Omega$, or a subsequence that tends uniformly to ∞ on every compact set.*

We shall show that this definition agrees with Definition 2 if we take S to be the Riemann sphere. If that is what we do, then we can also allow ∞ as a possible value, which means that we may consider families of meromorphic functions. There is no need to rephrase the definition so that it covers normal families of meromorphic functions, for Definition 2 applies without change.

It is necessary, however, to prove a lemma which extends Weierstrass's and Hurwitz's theorems to meromorphic functions (Theorems 1 and 2).

Lemma. *If a sequence of meromorphic functions converges in the sense of spherical distance, uniformly on every compact set, then the limit function is meromorphic or identically equal to* ∞.

If a sequence of analytic functions converges in the same sense, then the limit function is either analytic or identically equal to ∞.

Suppose $f(z) = \lim_{n\to\infty} f_n(z)$ in the sense of the lemma. We know that $f(z)$ is continuous in the spherical metric. If $f(z_0) \neq \infty$, then $f(z)$ is bounded in a neighborhood of z_0, and for large n the functions f_n are $\neq \infty$ in the same neighborhood. It follows by the ordinary form of Weierstrass's theorem that $f(z)$ is analytic in a neighborhood of z_0. If $f(z_0) = \infty$ we consider the reciprocal $1/f(z)$ which is the limit of $1/f_n(z)$ in the spherical sense. We conclude that $1/f(z)$ is analytic near z_0, and hence $f(z)$ is meromorphic. If the f_n are analytic and the second case occurs, then $1/f$ must be identically zero by virtue of Hurwitz's theorem, and f is identically ∞.

The lemma makes it clear that Definition 3 is nothing other than Definition 2 applied to the spherical metric.

It is *not* true that the derivatives of a normal family form a normal family. For instance, consider the family formed by the functions $f_n = n(z^2 - n)$ in the whole plane. This family is normal, for it is clear that $f_n \to \infty$ uniformly on every compact set. Nevertheless, the derivatives $f_n' = 2nz$ do not form a normal family, for $f_n'(z)$ tends to ∞ for $z \neq 0$ and to 0 for $z = 0$.

By Arzela's theorem a family of meromorphic functions is normal if and only if it is equicontinuous on compact sets, for condition (ii) is now trivially fulfilled. The equicontinuity can be replaced by a boundedness condition. We have indeed:

Theorem 17. *A family of analytic or meromorphic functions f is normal in the classical sense if and only if the expressions*

$$\rho(f) = \frac{2|f'(z)|}{1 + |f(z)|^2} \tag{58}$$

are locally bounded.†

The geometric meaning of the quantity $\rho(f)$ is rather evident. Indeed, by use of the formula in Chap. 1, Sec. 2.4

$$d(f(z_1), f(z_2)) = \frac{2|f(z_1) - f(z_2)|}{[(1 + |f(z_1)|^2)(1 + |f(z_2)|^2)]^{\frac{1}{2}}}$$

† This theorem is due to F. Marty.

it is readily seen that f followed by stereographic projection maps an arc γ on an image with length

$$\int_\gamma \rho(f(z))|dz|.$$

If $\rho(f) \leqq M$ on the line segment between z_1 and z_2 we conclude that $d(f(z_1),f(z_2)) \leqq M|z_1 - z_2|$, and this immediately proves the equicontinuity when $\rho(f)$ is locally bounded.

To prove the necessity we remark first that $\rho(f) = \rho(1/f)$ as a simple calculation shows. Assume that the family $\mathfrak{F}$ of meromorphic functions is normal, but that the $\rho(f)$ fail to be bounded on a compact set E. Consider a sequence of $f_n \in \mathfrak{F}$ such that the maximum of $\rho(f_n)$ on E tends to ∞. Let f denote the limit function of a convergent subsequence $\{f_{n_k}\}$. Around each point of E we can find a small closed disk, contained in Ω, on which either f or $1/f$ is analytic. If f is analytic it is bounded on the closed disk, and it follows by the spherical convergence that $\{f_{n_k}\}$ has no poles in the disk as soon as k is sufficiently large. We can then use Weierstrass's theorem (Theorem 1) to conclude that $\rho(f_{n_k}) \to \rho(f)$, uniformly on a slightly smaller disk. Since $\rho(f)$ is continuous it follows that $\rho(f_{n_k})$ is bounded on the smaller disk. If $1/f$ is analytic the same proof applies to $\rho(1/f_{n_k})$, which is the same as $\rho(f_{n_k})$. In conclusion, since E is compact it can be covered by a finite number of the smaller disks, and we find that the $\rho(f_{n_k})$ are bounded on E, contrary to assumption. The contradiction completes the proof of the theorem.

EXERCISES

1. Prove that in any region Ω the family of analytic functions with positive real part is normal. Under what added condition is it locally bounded? *Hint:* Consider the functions e^{-f}.

2. Show that the functions z^n, n a nonnegative integer, form a normal family in $|z| < 1$, also in $|z| > 1$, but not in any region that contains a point on the unit circle.

3. If $f(z)$ is analytic in the whole plane, show that the family formed by all functions $f(kz)$ with constant k is normal in the annulus $r_1 < |z| < r_2$ if and only if f is a polynomial.

4. If the family $\mathfrak{F}$ of analytic (or meromorphic) functions is not normal in Ω, show that there exists a point z_0 such that $\mathfrak{F}$ is not normal in any neighborhood of z_0. *Hint:* A compactness argument.