CHAPTER I

Convergence

1. Introduction

Analysis is essentially the mathematics of approximation. Mathematical entities that are difficult to deal with directly are approximated by others that are more convenient, and an attempt is made to apply to the original entities the results obtained for the approximate ones. Our first task will be to illustrate this statement for sequences, which involve the most elementary but also the most basic concepts of analysis. Many of the theorems about sequences can be generalized in far-reaching ways and can be set in a much broader framework.

It is not our intention here to repeat well-known details. Rather we are interested in bringing to the foreground certain dominating aspects of the subject. Thus from the very beginning the relevant concepts will be presented in a somewhat more general (and therefore more abstract) form than is customary in textbooks. From the pedagogical point of view it would be desirable to make as much use as possible of concrete examples, but in a systematic treatment of the subject we must not lose sight of the advantages offered by abstraction; it enables us to break up the longer proofs into shorter steps and thereby gain a better understanding of their logical structure.

Let us give a brief summary of the organization of the present chapter.

For the description of limit processes in analysis it is customary to introduce sequences. Our purpose is to analyze the resulting concepts and theorems from several points of view with the idea in mind that in mathematics such a procedure always leads to valuable insights into the structure of the subject and to productive generalizations.

The concept of *convergence* does not in itself require the whole structure of the real or complex numbers but can already be formulated in arbitrary *topological spaces* (§2.1). But even the condition that a convergent sequence shall have only one limit already forces us to introduce our first restriction: this condition holds only in so-called *Hausdorff spaces* (§2.2).

We are accustomed to the idea of *calculating* with the limits of numerical sequences, but in a general topological space such calculations may not be possible, since none of the relevant operations are necessarily defined in such a space. But if they are defined, then calculation with limit values has a meaning, and the question arises: under what circumstances can we prove the well-known theorems about interchange of a limit process with addition or multiplication? In other words, when are addition and multiplication consistent with the topology of the space? This question leads to the concept of a *topological group* and to related concepts (§§2.3 to 2.6).

The order structure of the real numbers makes them into a topological space with striking properties; consequently the sequences of real numbers (and also, to some extent, of complex numbers, which can be reduced to them) have properties that require the special structure of this topological space for their very formulation (§3).

The *fundamental Cauchy theorem* for numerical sequences, which is proved in §3, does not make use of the full order structure of the real numbers but only of its weaker metric structure; but even the latter is definitely stronger (that is, more special) than a merely topological structure, in which the *Cauchy condition* cannot even be formulated. In a metric space, however, this condition has meaning and the validity of the fundamental Cauchy theorem is an additional property, called *completeness*. An important criterion for completeness is given at the close of this section (§4).

In the following section (§5) we focus our attention, not on the set of values of a sequence, but on its index set. For a sequence in the classical sense this index set is the set of natural numbers, which proves for many purposes to be far too "thin": for example, in the symbol $\lim_{x \to a} f(x)$ the "index" x runs through all the numbers in a neighborhood of a. Of course this particular example of a "continuous" limit process can be reduced to the convergence of sequences in the classical sense, but that fact does not make it any more natural to insist that index sets must be countable. The natural generalization is provided by the *Moore-Smith sequences*, or by the equivalent, and often more convenient, concept of a *filter*.

From the theoretical point of view the most important application of filters is the following: just as the topological properties of a space can be described without reference to any metric but merely in terms of subsets of the space, we may also introduce a general structure in which the *Cauchy condition* can be formulated without reference to a metric. Such a "*uniform structure*" induces a natural topology and allows us, in particular, to give the most general formulation (§6) of the extremely important concept of uniform convergence.

2. Sequences

2.1. *The Limit Concept*

The definition of a *sequence* involves three entities: the index set, namely (in the classical case) the set N of non-negative whole numbers $0, 1, 2, \ldots$; the set of values M, which may be any non-empty set, and finally a mapping of N into M.[1] Thus to each number n from N there corresponds a well-defined element of M, denoted by a_n. A sequence is not merely a subset; it is an indexed (enumerated) subset. If the elements of M are numbers, we speak of a numerical sequence, if they are functions, we speak of a sequence of functions, and so forth. If no further properties are assigned to the set M, the sequences are of no great interest, but the situation is quite different if M is a *topological space*, a concept that arises in the following way. Certain subsets of M are distinguished, in the manner described below, as *neighborhoods*, whereupon M is said to have a *topological structure*; or equivalently, the set M, along with the topological structure thus defined for it, is said to be a topological space.[2] The elements of such a set are usually called *points*, so that these "points" may, for example, be functions. In the set M we distinguish certain subsets which, as already mentioned, are called neighborhoods. They are required to have the following properties.

H 1) *Every point $a \in M$ has at least one neighborhood $U(a)$ and is contained in each of its neighborhoods: $a \in U(a)$.*

H 2) *For any two neighborhoods $U(a)$ and $V(a)$ there exists a neighborhood $W(a)$ which is contained in the intersection of $U(a)$ and $V(a)$: $W(a) \subset U(a) \cap V(a)$.*

H 3) *For every point b in the neighborhood $U(a)$ there exists at least one neighborhood $U(b)$ which is contained in $U(a)$.*

Thus a set, together with a system of subsets, which satisfies H 1 *through* H 3 *is called a topological space.*

It is now possible to introduce one of the most important concepts of analysis, namely that of a *limit*. It may happen that the elements of the sequence (which, as we have seen, are points of the topological space M) are ultimately contained in an arbitrarily chosen neighborhood of a point a of the space. We then say that *the sequence converges to the point a* or that a is a *limit* of the sequence. The situation can be described more precisely as follows: we say that *the sequence $a_n (a_n \in M, n \in N)$ converges to the point $a \in M$ if for every neighborhood $U(a)$ of a there exists a number n_0 such that $a_n \in U(a)$ for all $n \geqq n_0$.*

Thus convergence is a so-called infinitary property of the sequence. In other words, any valid statement about the sequence remains valid if only finitely many elements are changed or omitted. We shall often regard a given property as valid for all the elements of a sequence, even though there may be finitely many exceptions, since it is usually irksome to keep mentioning the exceptional cases.

[1] For the concept of "mapping" see IA, §8.4.

[2] A set can be given a topological structure in many different ways, so that one and the same set can be interpreted as a topological space in many ways.

2.2 *Uniqueness of the Limit*

In a completely general topological space it is possible for a sequence to have several limits, an undesirable feature of the space which can be eliminated only if we specialize further to a so-called *Hausdorff space*, that is, if we also require the following *Hausdorff separation axiom.*

H 4) *For any two distinct points a and b there exist mutually exclusive neighborhoods* $U(a)$, $U(b)$, *with* $U(a) \cap U(b) = \varnothing$.[3]

A topological space with the properties H 1) *through* H 4) *is called a Hausdorff space.* As mentioned above, the following theorem holds in such a space:

Theorem 1. *In a Hausdorff space any sequence has at most one limit.*

The proof is very easy. If a and a' are limits of a sequence a_n, $n \in N$, and $U(a)$ and $U(a')$ are arbitrary neighborhoods of a and a', there exist numbers n_1, n_2 such that $a_n \in U(a)$ for all $n \geqq n_1$, and $a_n \in U(a')$ for all $n \geqq n_2$. If n_0 is the greater of the two numbers n_1, n_2, then a_n is in the intersection $U(a) \cap U(a')$, for all $n \geqq n_0$. Thus no two neighborhoods have an empty intersection, which by H 4) is possible only if a and a' coincide.

In a Hausdorff space[4] the uniquely determined limit of a sequence is denoted by

$$a = \lim_{n \to \infty} a_n. \tag{1}$$

An alternative notation is

$$a_n \to a \quad \text{for} \quad n \to \infty. \tag{2}$$

As a final remark we note that the elements of a sequence may be considered as approximations to the limit, so that in the sense of the topology of M the limit can be approximated with arbitrary exactness.

2.3. *Topological Groups*

A great deal more can be said about sequences if the space M, in addition to being a Hausdorff space, has an algebraic structure that is "consistent" with its topology. A good example is given by the real numbers, which form not only a Hausdorff space but also a field, so that we can speak of a *topological field.* In the same way we can also speak of *topological groups*, *topological rings*, and so forth.

Before giving an exact definition of a topological group, let us introduce the following useful notation:

If A and B are subsets of the group[5] M, we denote by AB the set of

[3] The symbol $\varnothing$ denotes the empty set.

[4] For simplicity we shall assume below (often tacitly) that the space is Hausdorff.

[5] For the axioms defining a group see IB2, §1.1.

"points" ab with $a \in A$, $b \in B$. Similarly, $1/A = A^{-1}$ denotes the set of "points" $1/a = a^{-1}$ inverse to a, with $a \in A$. Compare the rules for calculation with complexes in IB2, §3.1.

When we say that the group property is "*consistent*" with the topology of M we mean that the group operations are continuous mappings of the space into itself, in the following precise sense:

G 1) *If $c = ab$ and $U(c)$ is an arbitrary neighborhood of c, there exist neighborhoods $U(a)$, $U(b)$ of a and b such that $U(a)U(b) \subset U(c)$.*

G 2) *For an arbitrary neighborhood* $U(1/a)$ of $1/a$ *there exists a neighborhood $U(a)$ of a such that* $1/U(a) \subset U(1/a)$.

In more intuitive language the meaning of G 1) and G 2) is as follows: the product $z = xy$ is arbitrarily close to c if x is sufficiently close to a and y is sufficiently close to b; and the inverse $1/x$ is arbitrarily close to $1/a$ if x is sufficiently close to a.

Definition 1. *A topological space which is also a group and which satisfies the axioms* G 1) *and* G 2) *is called a topological group.*

Note. These remarks are independent of the notation used for the operation of the group. For example, in Abelian groups it is customary to regard the operation as addition and to denote it by $+$, and consequently (cf. IB2, §1) to write $-a$ for a^{-1}.

2.4. *Calculation with Limits*

Theorem 2. *Let M be a topological group, the group operation being written multiplicatively. If the sequences a_n and b_n ($a_n, b_n \in M$, $n \in N$) are convergent, then the sequence $c_n = a_n b_n$ is also convergent and has the limit $c = ab$; or:* $\lim c_n = \lim a_n \lim b_n$.

Proof. Let $U(c)$ be an arbitrary neighborhood of c. In view of the continuity of the product operation, there exist neighborhoods $U(a)$, $U(b)$ of a and b such that $U(a)U(b) \subset U(c)$; that is, $xy \in U(c)$, if $x \in U(a)$, $y \in U(b)$. By hypothesis, there exist numbers n_1, n_2 such that $a_n \in U(a)$ for $n > n_1$ and $b_n \in U(b)$ for $n > n_2$. For all $n > n_0 = \max(n_1, n_2)$ we then have $c_n = a_n b_n \in U(c)$, as was to be proved.

Remark. If the group operation is written additively, the theorem becomes:

Theorem 3. $\lim(a_n + b_n) = \lim a_n + \lim b_n$.

Theorem 4. *Let M be a topological group. If the sequence a_n converges to a ($a_n, a \in M$; $n \in N$), then the sequence $1/a_n$ is also convergent and has the limit $1/a_n$ that is,* $\lim(1/a_n) = 1/\lim a_n$.

Proof. For an arbitrary neighborhood $U(1/a)$ of $1/a$ there exists a neighborhood $U(a)$ of a such that $1/U(a) \subset U(1/a)$; that is, $1/x \in U(1/a)$, if $x \in U(a)$. Moreover, there exists a number n_0 such that $a_n \in U(a)$ for $n > n_0$ and also such that $1/a_n \in U(1/a)$.

Since $\frac{a_n}{b_n}$ can be written in the form $a_n \frac{1}{b_n}$, theorems 2 and 4 imply the following theorem for a topological group M with multiplicatively written group operation:

Theorem 5. $\lim \frac{a_n}{b_n} = \frac{\lim a_n}{\lim b_n}$ *if the right side exists.*

2.5. *Topological Fields*

If we proceed still further with this specialization of the space M and require that it be a *topological field,* we mean thereby that the following conditions are satisfied:

1. *M is a Hausdorff space;*
2. *M is a field* (cf. IB5, §§1.2 *and* 1.10); *in other words, M is an additive group, $M^* = M - 0$ is a multiplicative group, and the two group operations are connected by the distributive laws;*
3. *The consistency conditions* G 1) *and* G 2) *are satisfied, both for addition and for multiplication, where in* G 2) *we must, of course, assume $a \neq 0$.*

In a topological field we have the formulas

$$\lim (a_n + b_n) = \lim a_n + \lim b_n,$$
$$\lim a_n b_n = \lim a_n \lim b_n,$$
$$\lim \frac{a_n}{b_n} = \frac{\lim a_n}{\lim b_n}, \qquad b_n \neq 0, \quad b = \lim b_n \neq 0, \tag{3}$$

if the sequences a_n and b_n converge; and in each case the limit is uniquely determined.

2.6. *Application to the Field C of Complex Numbers*

It is natural to ask why the proofs in the textbooks are so much more complicated than the ones given here. The explanation is that in the proofs of the theorems about limits it is customary to include, in implicit form, a proof of the continuity of the two operations, addition and multiplication. Let us illustrate in more detail for the case of the field C of complex numbers. It is clear that the limit theorems from §2.5 will have been proved if we show that this field is a topological field. Regarded as a metric space it is certainly also a Hausdorff space, and thus we need only prove that the addition and multiplication are continuous.

Continuity of addition. Let $c = a + b$ and let ϵ be a given positive number. The ϵ-neighborhood of c is then the set $U(c)$ of numbers z with $|z - c| < \epsilon$. For $U(a)$ we take the set of numbers x with $|x - a| < \frac{\epsilon}{2}$ and for $U(b)$ the set of numbers y with $|y - b| < \frac{\epsilon}{2}$. From $|(x + y) - (a + b)| = |(x - a) +$

$(y - b)| \leqq |x - a| + |y - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ it follows at once that $z = x + y$ is contained in the ϵ-neighborhood of $c = a + b$; that is,

$$U(a) + U(b) \in U(a + b).$$

Continuity of multiplication. After choice of the arbitrary positive number ϵ we choose two other positive numbers ϵ_1 and ϵ_2 in the manner described below. We note that $|xy - ab| = |b(x - a) + a(y - b) + (x - a)(y - b)| \leqq |b|\,|x - a| + |a|\,|y - b| + |x - a|\,|y - b|$. Let $U(a)$ be an ϵ_1-neighborhood of a and $U(b)$ an ϵ_2-neighborhood of b. For $x \in U(a)$ and $y \in U(b)$ we have $|b|\,|x - a| \leqq |b|\epsilon_1$, $|a|\,|y - b| \leqq |a|\epsilon_2$ (with equality if either a or b is equal to 0) and $|x - a|\,|y - b| < \epsilon_1\epsilon_2$. If we take $\epsilon_1 = \frac{1}{3}\epsilon/(|b| + 1)$, $\epsilon_2 = \frac{1}{3}\epsilon/(|a| + 1)$ with $\epsilon_2 < 1$, then obviously $|xy - ab| < \frac{1}{3}\epsilon + \frac{1}{3}\epsilon + \frac{1}{3}\epsilon$, which completes the proof.

Continuity of the operation of forming the inverse. Assume $a \neq 0$ and let η be a positive number to be determined below. If we first assume $\eta < \frac{1}{2}|a|$, then for every number x in the η-neighborhood $U(a)$ of a we have the inequality $|x - a| < \frac{1}{2}|a|$, and consequently $|x| > \frac{1}{2}|a|$. Thus the numbers in $U(a)$ are different from 0. Furthermore,

$$\left|\frac{1}{x} - \frac{1}{a}\right| = \frac{|x - a|}{|x|\,|a|} < \frac{\eta}{\frac{1}{2}|a|^2}.$$

If we now take $\eta < \frac{1}{2}|a|^2\epsilon$, the proof is complete.

The above proof is a good illustration of the remark in §1 to the effect that a proof can often be divided up into shorter steps.

3. Monotone Sequences and Limits of Indeterminacy

3.1. *Monotone Sequences*

We shall make use of the fact (see IB1, §4.4) that in the field of real numbers every set A that is bounded above has a least upper bound

(4) $$\mu = \sup_{x \in A} x,$$

and every set B that is bounded below has a greatest lower bound

(5) $$\lambda = \inf_{x \in B} x.$$

If the sets are unbounded above or below, then for μ, λ it is convenient to take the numbers $+\infty$ and $-\infty$, which are adjoined to the field as improper elements.

For monotone sequences it is easy to establish convergence a priori, without knowing the value of the limit, where, as usual, we say that a sequence a_n, $n \in N$ is *monotone increasing* if $a_n \leqq a_{n+1}$ for all n, and *monotone decreasing* if $a_n \geqq a_{n+1}$ for all n (where it is obviously convenient to allow a finite number of exceptions to these inequalities).

The fundamental theorem on monotone sequences runs as follows:

Theorem 6. *If a monotone increasing sequence is bounded above, then the sequence is convergent, and similarly for a monotone decreasing sequence bounded below.*

It will be sufficient to prove the first part of the theorem. Since the set of numbers a_n, $n \in N$ is bounded above, it has a least upper bound

$$a = \sup_{n \in N} a_n.$$

Thus for all n we have $a_n < a + \epsilon$, with arbitrary positive ϵ, since it is true that $a_n \leqq a$. Furthermore, there exists a number n_0 with $a_{n_0} > a - \epsilon$ and thus, since the sequence is monotone, $a_n > a - \epsilon$ for all $n \geqq n_0$.

3.2. *Limits of Indeterminacy*

All sequences of real numbers, including those that do not converge, display certain regularities closely connected with monotone sequences.

Let μ_k be the least upper bound of the set of terms a_n, $n \geqq k$ in a given sequence a_n, $n \in N$; that is,

$$\mu_k = \sup_{n \geqq k} a_n, \qquad k = 0, 1, 2, \ldots. \tag{6}$$

If the sequence is not bounded above, we set $\mu_k = +\infty$. Otherwise the numbers μ_k are the terms of a monotone decreasing sequence, the *majorizing sequence* of the given sequence. The number

$$\mu = \inf_{k \in N} \mu_k \tag{7}$$

is either $-\infty$ or the limit of the sequence μ_k. This number is called the *upper limit* (limes superior) of the sequence a_n. We write

$$\mu = \limsup_{n \to \infty} a_n. \tag{8}$$

If one of the $\mu_k = +\infty$, we write $\mu = +\infty$.

Correspondingly we can form

$$\lambda_k = \inf_{n \geqq k} a_n, \qquad k = 0, 1, 2, \ldots \tag{9}$$

and

$$\lambda = \sup_{k \in N} \lambda_k. \tag{10}$$

The sequence λ_k, $k \in N$ is called the *minorizing sequence* of the given sequence if the λ_k are finite. The number (10) is equal to $-\infty$ if a $\lambda_k = -\infty$, is equal to $+\infty$ if the sequence λ_k is unbounded above, and is otherwise the limit of the monotone sequence λ_k. This number is called the *lower limit* (limes inferior) of the sequence a_n. We write

$$\lambda = \liminf_{n \to \infty} a_n. \tag{11}$$

It is easy to show that

(12) $$\liminf_{n\to\infty} a_n = -\limsup_{n\to\infty} (-a_n).$$

Since $$\lambda_0 \leqq \lambda_k \leqq \mu_k \leqq \mu_0, \qquad k \in N,$$

we have $$\lambda_0 \leqq \lambda \leqq \mu \leqq \mu_0,$$

and thus

(13) $$\inf_{n\in N} a_n \leqq \liminf_{n\to\infty} a_n \leqq \limsup_{n\to\infty} a_n \leqq \sup_{n\in N} a_n.$$

The upper and lower limits of a sequence are called its *limits of indeterminacy*. They are not necessarily identical with the least upper and greatest lower bounds.

The following theorem is useful in many applications:

Theorem 7. *If a number β is such that ultimately*

$$a_n < \beta + \epsilon,$$

for every fixed choice of the positive number ϵ, then

$$\limsup_{n\to\infty} a_n \leqq \beta.$$

If a number α is such that ultimately

$$a_n > \alpha - \epsilon,$$

then $$\liminf_{n\to\infty} a_n \geqq \alpha.$$

It is sufficient to prove the first part of the theorem. If $\beta = +\infty$, the assertion is trivial. But if β is finite, then by hypothesis the sequence is bounded above and thus the majorizing sequence exists. We can find a number n_0 with $a_n < \beta + \epsilon$ for all $n \geqq n_0$. Thus $\beta + \epsilon$ is an upper bound for the set of numbers a_n, $n \geqq n_0$, and consequently $\mu_k \leqq \beta + \epsilon$, for $k \geqq n_0$, and a fortiori $\mu \leqq \beta + \epsilon$. But since ϵ is arbitrarily small, we have $\mu \leqq \beta$. The proof requires only slight changes for the case $\beta = -\infty$, since the hypothesis then states that ultimately the numbers are less than any preassigned bound and consequently are unbounded below.

It is easy to show that λ is the smallest and μ is the largest number with the property stated in the hypothesis of the preceding theorem. Equivalently we have

Theorem 8. *Let ϵ be an arbitrary positive number. If λ and μ are the limits of indeterminacy of the sequence a_n, $n \in N$, then ultimately*

$$a_n > \lambda - \epsilon, \qquad a_n < \mu + \epsilon$$

and infinitely often

$$a_n < \lambda + \epsilon, \qquad a_n > \mu - \epsilon.$$

Again it is sufficient to consider the case of the upper limit. Since μ is the greatest lower bound of the numbers μ_k, there exists a number k with $\mu_k < \mu + \epsilon$, so that certainly $a_n < \mu + \epsilon$ for $n \geqq k$. If it were not true that $a_n > \mu - \epsilon$ infinitely often, then ultimately we would have $a_n \leqq \mu - \epsilon$, or in view of the preceding theorem $\mu \leqq \mu - \epsilon$, which is impossible.

The connection between limits of indeterminacy and convergence can be stated as follows.

Theorem 9. *A sequence is convergent if and only if its limits of indeterminacy are finite and are equal to each other. Then the sequence converges to this common value.*

The condition stated in the theorem is necessary, since if a_n converges to a, then ultimately $a_n > a + \epsilon$, so that $\mu \leqq a$ and $a_n > a - \epsilon$, and thus $\lambda \geqq a$. Since $\lambda \leqq \mu$, we thus have $\lambda = \mu = a$. Conversely, if $\lambda = \mu = a$, then ultimately $a_n > \lambda - \epsilon = a - \epsilon$ and ultimately $a_n < \mu + \epsilon = a + \epsilon$, so that a is the limit of the sequence.

3.3. *The Fundamental Theorem of Cauchy*

This theorem states a criterion for the convergence of a sequence of complex[6] numbers.

Let c_n, $n \in N$, be a sequence of complex numbers. We first formulate a necessary condition for convergence. Let c denote the limit of the sequence. Then if ϵ is a preassigned positive number, there exists a number n_0 such that $|c_n - c| < \frac{1}{2}\epsilon$ for $n \geqq n_0$. Thus

$$|c_m - c_n| \leqq |c_m - c| + |c_n - c| < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon$$

for all $m, n \geqq n_0$.

But this necessary condition, to the effect that ultimately the terms must remain arbitrarily close to one another, is also sufficient. The two assertions constitute the *fundamental theorem of Cauchy*:

Theorem 10. *For the convergence of a sequence c_n, $n \in N$, of complex numbers it is necessary and sufficient that for an arbitrarily preassigned positive ϵ there exists a number n_0 such that*

$$|c_m - c_n| < \epsilon \tag{14}$$

for all $m, n \geqq n_0$.

It remains to prove the sufficiency of the condition. If we write $c_n = a_n + ib_n$, with real a_n and b_n, then obviously c_n, $n \in N$, is convergent if both the sequences a_n and b_n, $n \in N$, are convergent. From $|c_m - c_n| \leqq |a_m - a_n| + |b_m - b_n|$ and $|a_m - a_n| \leqq |c_m - c_n|$, $|b_m - b_n| \leqq |c_m - c_n|$ we see that the Cauchy condition for c_n implies the same condition for a_n and b_n, and conversely. Thus it is sufficient to prove the theorem for real sequences.

[6] And thus, in particular, of real numbers.

Let a_n, $n \in N$, be a real sequence satisfying the Cauchy condition. If ϵ is an arbitrary positive number, there exists a number n_0 such that for some $m \geqq n_0$ and every $n \geqq n_0$ we have $|a_m - a_n| < \epsilon$ or $a_m - \epsilon < a_n < a_m + \epsilon$. If m is regarded as fixed, these inequalities hold for all numbers a_n with finitely many exceptions. But this statement means that the sequence is bounded and that the two limits of indeterminacy are finite. By §3.2, theorem 8, we have $a_m - \epsilon \leqq \liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n \leqq a_m + \epsilon$, so that

$$\limsup_{n\to\infty} a_n - \liminf_{n\to\infty} a_n \leqq 2\epsilon.$$

But since the positive number ϵ is arbitrary, the two limits of indeterminacy must be equal, which completes the proof of convergence.

4. Metric Spaces

4.1. *Axiomatic Definition*

In III 2, §1.2 we shall be making use of special topological spaces, namely metric spaces. They are characterized by the fact that for every two points there is defined a real non-negative number $d(a, b)$, called the "*distance*" between a and b, with the following properties:

M 1) *$d(a, b) = 0$ if and only if $a = b$,*
M 2) *$d(a, b) = d(b, a)$,*
M 3) *For every choice of elements a, b, c we have*

$$d(a, b) \leqq d(a, c) + d(c, b) \qquad \textit{(triangle inequality)}.$$

The connection here with the classical arguments of analysis is obvious.

As an example of a metric space let us take the set of rational numbers, or the set of real numbers, with the distance function $d(a, b) = |a - b|$. The axioms M 1) to M 3) can be verified at once. Thus theorem 1 in §2.2 is valid if we are dealing with sequences of rational or real numbers.

4.2. *Complete Metric Spaces*

If a sequence a_n converges to a point a in a metric space, then by the triangle inequality (cf. §4.1, M 3)) and the definition of convergence we have, as in §3.3, the following *necessary condition for convergence*:

$$(15) \quad d(a_n, a_m) \leqq d(a_n, a) + d(a, a_m) \leqq \epsilon \qquad \text{for} \qquad n, m > n_0, \quad \epsilon > 0.$$

If the space is such that (15) is also a *sufficient* condition for convergence of sequences, it is said to be *complete*. Thus a sufficient condition for the completeness of a space is that in it every bounded infinite subset has at least one *limit point*, i.e., a point such that every neighborhood of it contains infinitely many points of the set.

We can now prove the following theorem:

Theorem 11. *If in a metric space every bounded infinite subset has at least one limit point, then the space is complete.*

The proof is as follows. For a sequence a_n let $d(a_m, a_n) < \frac{\epsilon}{2}$ for arbitrary $\epsilon > 0$ and $n, m > n_0(\epsilon)$. Then it follows that the sequence is bounded and thus has a limit point a. Consequently, $d(a, a_m) < \frac{\epsilon}{2}$ for infinitely many m; and by hypothesis $d(a_n, a_m) < \frac{\epsilon}{2}$ for $n, m > n_0$. Consequently $d(a_n, a) \leqq d(a_n, a_m) + d(a_m, a) < \epsilon$ for all $n > n_0$, which means that the sequence a_n converges to a.[7]

This criterion for convergence in an arbitrary metric space is obviously a *generalization of the Cauchy criterion for convergence* in classical analysis.

5. Filters

5.1. *Moore-Smith Sequences*

A sequence is a mapping of the set N of natural numbers into a set of values M. This mapping determines a structure in M, which can be described in terms of the set of segments a_n, $n \geqq n_0$, of M, where n_0 is an arbitrary non-negative number.

But in the above discussion the order properties of the set N have not been used to their full extent. Our statements remain correct if for the index set N we take a set in which there is defined a *relation* ρ with the following properties:

gM 1) *If $p \rho q$ and $q \rho r$, then $p \rho r$; namely, the relation is transitive.*
gM 2) *For any two elements p and q in N there exists in N an element r with $p \rho r$ and $q \rho r$.*
gM 3) We have $p \rho p$ for all p.

Such a set is called a *directed set.* For the non-negative integers the relation "ρ" usually means "$\leqq$," but for the same set of numbers we can easily define other relations with the above properties, e.g., the relation "divisible," in which $p \rho q$ means that q is divisible by p.

A set M, a directed set N and a mapping of N into M define a *Moore-Smith sequence.* If for M we take a topological space R, the sequence is said to be *convergent to a point a if for every neighborhood $U(a)$ of a there exists an element n_0 in the index set such that $a_n \in U(a)$ for all n with $n_0 \rho n$.*

The following examples show that we are dealing here with an essential extension of the concept of a sequence.

[7] G. Aumann, Reelle Funktionen. Springer, Berlin-Göttingen-Heidelberg 1954, p. 126.

Let the index set be a (non-empty) set X of real numbers with a limit point a. The set becomes a directed set if we agree that $p \rho q$ means $|p - a| \geqq |q - a|$. The above conditions are then easily verified. A mapping f of the set X into another set Y of real numbers, or in other words a real function, determines a Moore-Smith sequence. The terms of the sequence are the values $f(x)$ of the function. In the sense of Moore-Smith convergence this sequence has a limit b if for every assigned number ϵ there exists a number $x_0 \in X$ such that $|f(x) - b| < \epsilon$ for all x with $x_0 \rho x$, which means that $|x_0 - a| \geqq |x - a|$, or in other words $|x - a| \leqq \delta$, with $\delta = |x_0 - a|$. Since the definition would otherwise be trivial, we assume that a is not an element of X. Then $\delta > 0$, and we have the usual definition of the limit of a function for $x \to a$. Of course it is also possible that $f(a)$ exists. If the value of the function at this point coincides with the limit, then $f(x)$ is said to be *continuous* for $x = a$.

Another striking example[8] is provided by the *Riemann integral*. Here we let the function f be defined in the interval $a \leqq x \leqq b$ and for a partition $\mathfrak{z}$ of the interval defined by the intermediate points

$$a = x_0 \leqq x_1^* \leqq x_1 \leqq x_2^* \leqq x_2 \leqq \cdots \leqq x_{n-1} \leqq x_n^* \leqq x_n = b$$

we consider the Riemann sum

$$S(f, \mathfrak{z}) = \sum_{i=1}^{n} f(x_i^*)(x_i - x_{i-1}).$$

Then the set $\mathfrak{Z}$ of all partitions $\mathfrak{z}$ forms a directed set if by the norm of the partition $\mathfrak{z}$ we mean

$$d_{\mathfrak{z}} = \max_{1 \leqq i \leqq n} |x_i - x_{i-1}|$$

and define the relation $\mathfrak{z} \rho \mathfrak{z}'$ by $d_{\mathfrak{z}} \geqq d_{\mathfrak{z}'}$. Consequently $S(f, \mathfrak{z})$ is a Moore-Smith sequence with the index set $\mathfrak{Z}$ and the image set R of real numbers. If this sequence converges, the function f is said to be integrable in the sense of Riemann.

5.2. *Filters*

The concept of a Moore-Smith sequence can be translated into purely set-theoretic language, since the ordering relation ρ can be replaced by relations between sets.

Let there be given a sequence a_n, $n \in N$, where N is a directed set. In the image set M we consider a nonempty system $\mathfrak{F}$ of subsets with the following property: a subset U of M is an element of the system $\mathfrak{F}$ if and only if the

[8] Cf. G. Pickert, Folgen und Filter in der Infinitesimalrechnung, Der mathematische und naturwissenschaftliche Unterricht **13**, pp. 150–153 (1960).

elements of the sequence "ultimately" belong to U, e.g., if in the index set there exists an element n_0 such that $a_n \in U$ for all n with $n_0 \rho n$.

The system $\mathfrak{F}$ obviously has the following properties:

F 1) *The empty set is not an element of* $\mathfrak{F}$.
F 2) *Every set that includes an element of* $\mathfrak{F}$ *is an element of* $\mathfrak{F}$.
F 3) *The intersection of two elements of* $\mathfrak{F}$ *is an element of* $\mathfrak{F}$.

Such a system is called a *filter*. Let us give some examples.

In the set of non-negative integers we consider the subsets whose complements have only finitely many elements. This filter is called a *Fréchet filter* and is denoted symbolically by $n \to \infty$. A similar filter arises in any set M in which a sequence a_n, $n \in N$, is defined if as subsets we admit those subsets of n that contain "almost" all terms of the sequence, i.e., all terms with the exception of at most finitely many.

In a topological space we may consider the system of subsets that include a neighborhood of a given point a. This filter is called the neighborhood filter of a and is denoted symbolically by $x \to a$.

We now wish to show conversely that for a given filter $\mathfrak{F}$ in a set M we can always construct a Moore-Smith sequence related to $\mathfrak{F}$ in the above way.

Thus our first task is to construct from $\mathfrak{F}$ a suitable index set N. We construct this set in the following manner, which is natural enough but may seem somewhat artificial at first sight.

Let the elements of N be the pairs (a, A), where A is an element of $\mathfrak{F}$ (and is thus a subset of M) and a is an element of M that is contained in A. In N we define an order as follows:

$$(a, A) \rho (b, B) \quad \text{means} \quad A \supset B,$$

so that N is now a directed set. Since the properties gM 1) and gM 3) are immediately clear, we need only verify gM 2): for $A \in \mathfrak{F}$, $B \in \mathfrak{F}$ we see by F 3) that $C = A \cap B \in \mathfrak{F}$, and by F 1) the set C contains an element c; thus if $n = (a, A)$ and $m = (b, B)$ are two elements of N, then $p = (c, C)$ is also an element of N and we have a $n \rho p$ and $m \rho p$. We now obtain a Moore-Smith sequence by means of the mapping

$$n = (a, A) \to a \text{ of } N \text{ into } M.$$

It remains to show that the filter consisting of those subsets U of N which "ultimately" contain all the elements of the sequence is identical with the given filter $\mathfrak{F}$. For such a U there exists a $n_0 \in N$ such that $a_n \in U$ follows from $n_0 \rho n$. If n_0 is the pair (a_0, A_0) and if $a \in A_0$, then in particular we have $n_0 \rho n$ for $n = (a, A_0)$; consequently $a \in U$ for every $a \in A_0$, so that U contains A_0. Since $A_0 \in \mathfrak{F}$ we see from F 2) that $U \in \mathfrak{F}$. Conversely, every element A of the filter is a subset of M "ultimately" containing all the elements of the sequence, and for $a_0 \in A$ and $n_0 = (a_0, A)$ we see that for $n = (b, B)$ the relation $n_0 \rho n$ means precisely that $B \subset A$, so that in fact $a_n = b \in B \subset A$ for $n_0 \rho n$.

Two Moore-Smith sequences corresponding to the same filter obviously behave in the same way with respect to convergence, so that it must be possible to express convergence or nonconvergence solely in terms of filters.

In fact, a Moore-Smith sequence converges to the limit a if and only if every neighborhood $U(a)$ of a "ultimately" contains all the elements of the sequence; in other words, if and only if every $U(a)$ belongs to the filter $\mathfrak{F}$; or expressed still differently, if and only if $\mathfrak{F}$ includes the neighborhood filter $\mathfrak{U}(a)$ of a. This statement remains meaningful if $\mathfrak{F}$ is an arbitrary filter, not necessarily arising from a sequence. Thus it is natural to regard convergence not as a property of sequences but of filters and to make the following definition.

A filter $\mathfrak{F}$ in a topological space M is said to be convergent to $a \in M$ if it is finer[9] *than the neighborhood filter $\mathfrak{U}(a)$.*[10]

Thus a sequence $f: N \to M$ is convergent if and only if the corresponding filter converges.

We have interpreted the behavior of a sequence in terms of filters in the image space M. But the structure of the index set N can also be defined by means of a filter: in N the subsets that "ultimately" contain all elements of N form a filter $\mathfrak{F}_0$ which is obviously a natural generalization of the Fréchet filter for the set of natural numbers.

The *image filter* $\mathfrak{F}$ in M generated by the mapping $f: N \to M$ of the "direction filter" $\mathfrak{F}_0$ in N can now be described as the set of all subsets of M that contain images of sets in $\mathfrak{F}_0$. These image sets themselves form a system of sets $\mathfrak{B}$ in M with the properties

FB 1) *The empty set does not belong to $\mathfrak{B}$,*
FB 2) *The intersection of any two sets in $\mathfrak{B}$ contains a set in $\mathfrak{B}$.*

Such a system of sets $\mathfrak{B}$ is called a *filter basis*. Thus the mapping $f: N \to M$ maps the sets of a filter $\mathfrak{F}_0$ in N onto a filter basis $\mathfrak{B}$ in M, and the image filter $\mathfrak{F}$ arises from $\mathfrak{B}$ by the adjunction of all the sets that include sets in $\mathfrak{B}$.[11]

[9] A filter $\mathfrak{F}_1$ is said to be *finer* than the filter $\mathfrak{F}_2$ if it contains more sets, i.e., if $\mathfrak{F}_1 \supset \mathfrak{F}_2$.

[10] If the space M is a Hausdorff space (cf. §2.2) the filter $\mathfrak{F}$ has at most one limit; for then $\mathfrak{U}(a) \subset \mathfrak{F}$, $\mathfrak{U}(b) \subset \mathfrak{F}$ with $b \neq a$ is impossible, since for the two disjoint neighborhoods $U(a)$ and $U(b)$ it would follow from F 3 that $\varnothing = U(a) \cap U(b) \in \mathfrak{F}$, in contradiction to F 1.

[11] A filter is itself a filter basis and as such generates itself.

Since a sequence is a mapping $f: N \to M$ in which a filter $\mathfrak{F}_0$ is distinguished in N, the concept of convergence in the sense of Moore-Smith is subsumed in the following definition:

The mapping $f: N \to M$ is said to be convergent on the filter $\mathfrak{F}_0$ in N if the image filter $\mathfrak{F}$ of $\mathfrak{F}_0$ under f is convergent.[12]

Finally, the concept of *limits of indeterminacy of sequences* can be taken over for *filters of real numbers*:

If $\mathfrak{F}$ is a filter in the set R of real numbers and if

$$\mu_A \quad \text{and} \quad \lambda_A$$

denote, respectively, the least upper and greatest lower bounds of a set $A \subset R$, we define

$$\mu = \limsup \mathfrak{F} = \inf_{U \in \mathfrak{F}} \mu_U, \qquad \lambda = \liminf \mathfrak{F} = \sup_{U \in \mathfrak{F}} \lambda_U.$$

It is obvious that a real filter is convergent if and only if $\mu = \lambda$.

6. Uniform Spaces

6.1. Our discussion up to now has served the purpose of generalizing the elementary concept of convergence in such a way that it can be defined, not only for metric spaces, but in purely topological terms.

However, the concept of convergence as defined in terms of a metric accomplishes somewhat more than the purely topological definition, as we may show by the example of sequences of real numbers: the *fundamental Cauchy theorem* (§3.3) allows us to express the convergence of a sequence in such a way that the limit itself is not mentioned.

The statement contained in formula (14) in §3.3, namely that "ultimately" two elements of the sequence in question are arbitrarily close to each other, is immediately expressible in terms of a metric but cannot be formulated in terms of convergence alone: for neighborhoods of *one* point it is meaningful to assert that one neighborhood is greater than another (namely, if one of them includes the other), but we cannot make any such comparison between neighborhoods of *distinct* points. A comparison of this sort, as it occurs in the formulation of the fundamental Cauchy theorem, can nevertheless be made without any reference to a metric, provided that the space M has a so-called uniform structure, which means simply that we have some way of measuring the "nearness to each other" of two points.

In order to distinguish the various concepts as clearly as possible, we at first pay no attention to the topology of M, regarding it merely as a set

[12] Thus the definition assumes that M is a topological space, which is not necessarily true for N.

without structure. The definition we now proceed to develop for the concept of a *uniform structure as the totality of nearness relations* in M will then define a topology in M in a natural way. Only those topologies that arise from a uniform structure, in a sense to be made more precise below, allow us to formulate the fundamental Cauchy theorem.

In order to construct the desired definition of nearness we allow ourselves to be guided by the special case of a metric space M. The statements of interest are those that refer to pairs of points, so that it is convenient to formulate them as statements concerning the product space $M \times M$ consisting of the ordered[13] pairs (a, b) of points a and b in M. If $d(a, b)$ is the distance function that makes M into a metric space, the pairs (a, b) with $d(a, b) < \epsilon$ for a preassigned $\epsilon > 0$ form a set V_ϵ in $M \times M$. The system of these sets V_ϵ has the following properties, which are only restatements of the properties of the distance function:

Since $d(a, a) = 0$ for every point $a \in M$, every set V_ϵ contains the set Δ of all pairs (a, a); this set is called the "diagonal."

The equality $d(a, b) = d(b, a)$ states that every set V_ϵ is symmetric: $V_\epsilon^{-1} = V_\epsilon$, where V^{-1} is the set arising from V by reflection in the diagonal:

$$(a, b) \in V^{-1} \text{ if and only if } (b, a) \in V.$$

Finally, the triangle inequality $d(a, c) \leqq d(a, b) + d(b, c)$ states that $(a, c) \in V_\epsilon$ if $(a, b) \in V_{\epsilon/2}$ and $(b, c) \in V_{\epsilon/2}$. This latter statement suggests that we should introduce the following operation (called *composition*) for the subsets of $M \times M$:

For $A \subset M \times M$ and $B \subset M \times M$ let $A \circ B$ be the subset of $M \times M$ consisting of those pairs (a, c) for which there exists a *$b \in M$ such that $(a, b) \in A$ and $(b, c) \in B$.*

The above assertion then has the following form:

$$V_{\epsilon/2} \circ V_{\epsilon/2} \subset V_\epsilon.$$

The system of sets V_ϵ obviously forms a filter basis, since in view of $V_\epsilon \supset \Delta$ no set V_ϵ is empty and also $V_{\epsilon 1} \cap V_{\epsilon 2} = V_{\min(\epsilon_1, \epsilon_2)}$. Thus if to every V_ϵ we adjoin all the sets that include it, we obtain a filter $\mathfrak{G}$ of subsets U from $M \times M$ with the following properties:

F* 1) *Every $U \in \mathfrak{G}$ contains the diagonal Δ.*
F* 2) *If $U \in \mathfrak{G}$, then also $U^{-1} \in \mathfrak{G}$.*
F* 3) *For every $U \in \mathfrak{G}$ there exists a $V \in \mathfrak{G}$ with $V \circ V \subset U$.*

In this way we have gained a vantage point from which we can make the following definition, without reference to any metric:

Let M be a set and let $\mathfrak{G}$ be a filter in $M \times M$ with the properties F* 1)

[13] I.e., the pairs (a, b) and (b, a) are distinct if $a \neq b$.

to F* 3)*; then a uniform structure is thereby defined in M; a set U in $\mathfrak{G}$ is called an entourage.*

If a uniform structure has been defined on a set M, we can formulate the Cauchy condition (as was our original purpose) for sequences, or more generally for filters:

A filter $\mathfrak{F}$ in M is called a Cauchy filter if it contains arbitrarily "small" sets, i.e., if for every entourage U there exists a set $A \in \mathfrak{F}$ with $A \times A \subset U$.

Finally, in order to state the desired generalization of the fundamental Cauchy theorem we must define what we mean by saying that a Cauchy filter converges; in other words, we must convert the set M into a topological space. As was pointed out above, the given uniform structure will enable us to define the desired topology in a natural way:

For every point $a \in M$ and every entourage $U \in \mathfrak{G}$ let $U(a)$ be the set of those $x \in M$ for which $(a, x) \in U$.[14] These sets $U(a)$ will then form the system of neighborhood filters for the desired topology on M. Thus we make the following definition:

A set $A \subset M$ is said to be open if for every $a \in A$ there exists a $U(a)$ with $U(a) \subset A$; the empty set is also said to be open.

This definition satisfies the axioms under which the system of open sets on M defines a topology, namely:

The entire space M is open, since $M \times M \in \mathfrak{G}$ and $M = (M \times M)(a)$.

The union of arbitrarily many open sets is obviously open; and finally, the intersection of two open sets A and B is also open, since for $a \in A \cap B$ and $U_1(a) \subset A$, $U_2(a) \subset B$ we have

$$A \cap B \supset U_1(a) \cap U_2(a) = (U_1 \cap U_2)(a),$$

and with U_1 and U_2 their intersection $U_1 \cap U_2$ also belongs to $\mathfrak{G}$.

Thus from the uniform structure we have constructed on M the topology known as the *uniform topology*, under which the space M itself becomes a *uniform space*.

The question of deciding when a given topological space is uniform will not be discussed here. Let us simply state that every metric space, and thus in particular the space R of real numbers, is uniform; moreover, we shall later construct important examples of function spaces that are uniform and shall use them to illustrate the far-reaching importance of these concepts.

For the present, let us fix our attention again on the *fundamental Cauchy theorem*, which for certain uniform spaces can now be stated in the following way:

Every Cauchy filter converges.

[14] When the uniform structure arises from a metric $d(x, y)$ on M and $U = V_\epsilon$ is an ϵ-entourage, this definition of $U(a)$ obviously means that $U(a)$ is an ϵ-neighborhood of a.

This theorem establishes a particular property[15] of those uniform spaces in which it is valid, namely their *completeness*. Here again we shall not investigate the criteria for completeness in a uniform space but shall content ourselves with the most important example:

Theorem 12. *The space R of real numbers, regarded as a uniform space, is complete.*[16]

Proof. Let $\mathfrak{F}$ be a real Cauchy filter. For $\epsilon > 0$ let V_ϵ again be the ϵ-entourage in $R \times R$. Since $\mathfrak{F}$ is a Cauchy filter, there exists for every $\epsilon > 0$ a $U_\epsilon \subset \mathfrak{F}$ with $U_\epsilon \times U_\epsilon \subset V_\epsilon$, i.e., $|x - y| < \epsilon$ for $x, y \in U_\epsilon$. If y is kept fixed, we have $y - \epsilon < x < y + \epsilon$ for all $x \in U_\epsilon$. Consequently,

$$y - \epsilon \leqq \lambda_{U_\epsilon} \leqq \lambda \leqq \mu \leqq \mu_{U_\epsilon} \leqq y + \epsilon \quad \text{and thus} \quad 0 \leqq \mu - \lambda \leqq 2\epsilon.$$

Since this result holds for arbitrary $\epsilon > 0$, it follows that $\mu = \lambda$, so that the filter is convergent.

Thus we have laid the basis for all the general applications of the fundamental Cauchy theorem in analysis. As an example let us again consider a real function f defined on a set of real numbers with the limit point a. The images of neighborhoods of a and sets including them in R are the elements of a filter, which is a Cauchy filter if for preassigned $\epsilon > 0$ we can always find a δ-neighborhood $U_\delta(a)$ of a such that $f(U_\delta(a)) \times f(U_\delta(a)) \subset V_\epsilon$, where V_ϵ is the set of pairs of numbers $(f(x), f(y))$ with $|f(x) - f(y)| < \epsilon$. But this condition obviously means that $|f(x) - f(y)| < \epsilon$ for all x, y for which $|x - a| < \delta$, $|y - a| < \delta$. Thus under these conditions $\lim_{x \to a} f(x)$ exists.

Although the Cauchy condition is not sufficient for convergence in all uniform spaces, still it is always necessary:

A convergence filter $\mathfrak{F}$ in a uniform space M is a Cauchy filter.

Proof. Let G be the entourage filter in $M \times M$ and let a be the limit of $\mathfrak{F}$. Then $\mathfrak{F}$ is finer than the neighborhood filter $\mathfrak{U}(a)$, i.e., for every neighborhood $V(a)$ of a there exists a set $A \in \mathfrak{F}$ with $A \subset V(a)$. Let $U \in \mathfrak{G}$ be arbitrary and let $V \in \mathfrak{G}$ be such that $V \circ V \subset U$. Here we may take $V^{-1} = V$, since with V the inverse V^{-1} and consequently the intersection $V \cap V^{-1}$ also belong to $\mathfrak{G}$. Then if $A \in \mathfrak{F}$ and $A \subset V(a)$, it follows that $A \times A \subset U$, since for arbitrary elements $x \in A$, $y \in A$ we have $(a, x) \in V$ and $(a, y) \in V$, and thus $(x, a) \in V^{-1} = V$, so that $(x, y) \in V \circ V \subset U$.

[15] For example, in the space of rational numbers with the usual metric, the Cauchy sequence $\left(1 + \frac{1}{n}\right)^n$—and consequently the Cauchy filter generated by it—is not convergent.

[16] It would be easy to reduce this theorem to theorem 11 by showing that every metrically complete space is also complete when regarded as a uniform space.

6.2. *Uniform Convergence*

Let f_n, $n \in N$, be a sequence of real functions of a real variable x. Let the sequence be convergent for every value of x in a certain domain. Then the limit values are the values $g(x)$ of a well-defined function g. For given x and given positive ϵ we can find a number n_0 such that $|g(x) - f_n(x)| < \epsilon$ for $n \geqq n_0$. But these inequalities are not necessarily satisfied for a different value of x. Thus to every value of x there corresponds a suitable number n_0. If all these numbers have a (finite) upper bound m, then $|g(x) - f_n(x)| < \epsilon$ for all $n \geqq m$, where m no longer depends on x. In this case we say that the sequence f_n converges *uniformly* to g.

This situation can be put in a general setting by means of the concepts introduced in the preceding sections.

Let X be an arbitrary set and Y a uniform space. The mappings f of X in Y will again simply be called *functions*. They form the elements of a new set F. We shall show that F can be converted into a uniform space in a natural way. For this purpose we must make a sharp distinction between functions and their values. A function is an element of F, but the value of a function is an element of Y. To an element x of X and an element f of F there is assigned an element $f(x)$ of Y, namely the value of the function for the argument x.

We again let $\mathfrak{S}$ denote the filter of the uniform structure of the space Y. If U is a given subset of $Y \times Y$, we let U^* denote the subset of $F \times F$ such that (f, g) is an element of U^* if and only if the element $(f(x), g(x))$ belongs to U for every choice of x from X. It is clear that $U \subset V$ implies $U^* \subset V^*$. We now let U run through the filter $\mathfrak{S}$. Then U^* runs through a system of subsets in $F \times F$ which together with the sets that include them form a filter $\mathfrak{S}^*$. We now prove that $\mathfrak{S}^*$ defines a uniform structure in F.

Every set U^* obviously includes the diagonal in $F \times F$, since $(f(x), f(x))$ is always an element of U, independently of the choice of x.

Moreover, $(U^{-1})^* = U^{*-1}$ and $(U \circ V)^* = U^* \circ V^*$, since U^{*-1} consists of the pairs (f, g) with $(g, f) \in U^*$, i.e., $(g(x), f(x)) \in U$ or $(f(x), g(x)) \in U^{-1}$ for all values of x. Then (f, g) belongs to $U^* \circ V^*$ if and only if there exists an h with $(f, h) \in U^*$ and $(h, g) \in V^*$, which means that $(f(x), h(x)) \in U$ and $(h(x), g(x)) \in V$ for every x, i.e., $(f(x), g(x)) \in U \circ V$ or $(f, g) \in (U \circ V)^*$.

The uniform structure defined by $\mathfrak{S}^*$ defines a uniform topology for the function space F. Thus it is now possible to define the concept of *uniform convergence*, since it can now be reduced to ordinary convergence:

A filter Φ in the function space F converges uniformly to a function g if Φ converges in the uniform topology of F.

Let us analyze this definition. A neighborhood $U^*(g)$ of g is the set of functions f with $(f, g) \in U^*$, and thus $(f(x), g(x)) \in U$ for all x, where U denotes an element in the filter of the uniform structure of Y. Convergence means that in the filter Φ there exists an element A such that for all $f \in A$

$$(f(x), g(x)) \in U \qquad \text{for all} \qquad x \in X.$$

Let us now suppose that we are dealing with a sequence, in other words with a mapping of the set N of natural numbers into the function space F. The function corresponding to the number n will be denoted, as usual, by f_n. We obtain a filter Φ if we take those subsets of the sequence which contain all but finitely many of its elements. If X and Y denote the set of real numbers, the above definition of uniform convergence means, in terms of the usual uniform structure, that a sequence f_n, $n \in N$, has the limit g if for every positive number ϵ there exists a number n_0 such that

$$|f_n(x) - g(x)| < \epsilon$$

for $x \in X$ and $n \geqq n_0$.

But now the set A in the above discussion is exactly the set of functions f_n with $n \geqq n_0$, so that this specialization to the case of a sequence of real functions is identical with the classical definition with which we began our investigation of uniform convergence.