N70-21801

NASA SP-203

Introduction to

ABSTRACT ANALYSIS

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NASA SP-203

Introduction to ABSTRACT ANALYSIS

by
MARVIN E. GOLDSTEIN
and
BURT M. ROSENBAUM

Lewis Research Center
Cleveland, Ohio

Scientific and Technical Information Division
OFFICE OF TECHNOLOGY UTILIZATION 1969
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C.

For Sale by the Superintendent of Documents,
U.S. Government Printing Office, Washington, D.C. 20402

Price $2.25

Library of Congress Catalog Card Number 70–602068

CONTENTS

PREFACE

This book, which grew out of lectures given at the NASA Lewis Research Center, introduces the scientist and engineer with the usual background in applied mathematics to the concepts of abstract analysis. The emphasis is not on preparing the reader to do research in the field but on giving him some of the background necessary for reading the literature of pure mathematics.

Although the material here is by no means original, the presentation differs in some respects from texts on material of this nature. The proofs are more detailed herein and quite easy to follow. We have attempted to indicate how the material relates to and serves as a foundation for more advanced subjects. We have also attempted at several places to show how the material covered here relates to the more familiar "real mathematics." Enough examples are included to illustrate the concepts. No attempt is made to indicate the original sources of the material or even to point out the originators of all the concepts. Contrary to the usual practice, the relation between convergence and continuity on the one hand and algebraic operations on the other is discussed in the abstract setting of linear spaces. This is done principally to familiarize the reader with these very important concepts in a reasonably simple way.

CHAPTER 1

Elementary Set Concepts

Aside from being one of the principal tools of mathematics, set theory serves also as a unifying principle and foundation upon which mathematics can be based. A few mathematicians might even claim that mathematics is nothing more than set theory. In any event, attempts to put mathematics on set theoretic foundations have led to important contributions to the understanding of some of the more basic concepts of mathematics. However, our interest here in set theory is its use as a tool in mathematics.

The study of sets began with Cantor, around 1874, and grew out of his studies of the fundamental aspects of trigonometric series. Around the turn of the century great progress had been made in the theory of sets by Cantor, Russell, Frege, and others, and it appeared that there could be nothing which would prevent basing all mathematics on set theory alone. However, in 1903, when Frege was about to publish the second volume of his "Grundgesetze der Arithmetik," which was essentially his life work and relied heavily on the theory of sets, Russell sent Frege his ingenious paradox, which seemed so shattering to the foundations of set theory that Frege closed this volume with the following acknowledgment:

> A scientist can hardly encounter anything more undesirable than to have the foundation collapse just as the work is finished. I was put in this position by a letter from Mr. Bertrand Russell when the work was almost through the press.

Immediately set theoretic paradoxes began appearing in large numbers. In some sense these paradoxes always seem to stem from the fact that sets which are "too large" are encountered. From a practical point of view these paradoxes may be avoided by always assuming that there is some possibly large but fixed set from which, roughly speaking, all objects, which are considered in a given discussion, are taken. We will express this principle a little more precisely in the subsequent discussion.

This procedure assures that no known paradoxes will occur, but we can never be absolutely certain that any system will be completely free from contradictions. This was pointed out by Gödel who proved that no consistent system can be used to prove its own consistency.

It is possible to treat set theory itself as a mathematical discipline by taking the concepts of set and membership as undefined and then setting up exact rules to describe their interrelation. However, we make no attempt to develop "axiomatic set theory" here. On the contrary, our aim is only to develop (in a fairly intuitive way) those concepts of set theory which will be useful for the work in the following chapters. In this way, we shall follow the ideas of Halmos as set forth in his "Naive Set Theory" (ref. 1).

Before proceeding with the discussion of sets, let us briefly introduce some terminology which is encountered frequently in mathematics.

Statements which must be either true or false (even though we may not know which) are called *propositions.* For example, "Sauerkraut is better than potato salad" is a statement which cannot be classified as being either true or false. On the other hand, a statement such as "The sauerkraut sold in this supermarket is more expensive per pound than the potato salad" is a statement which must be either true or false. In this chapter propositions will be designated by single letters.

Suppose that p and q are any two propositions. In mathematics, the sentences "*p implies q*", "*if p, then q*", "*p only if q*", "*p is a sufficient condition for q*", and "*q is a necessary condition for p*" occur frequently. They all mean that whenever the proposition p is true, then the proposition q must also be true or, what is the same thing, whenever q is false, p must also be false (for if q were false, p could not be true since this would imply that q had to be true also). The sentences "*p is necessary and sufficient for q*" and "*p if and only if q*" mean both *p implies q* and *q implies p*. The first of these shows that if p is true, q must be true. The second shows that if p is false, then q is false also. Hence, p and q must either both be true or both be false. Thus, for example, if the proposition p is "Paul is taller than Harry" and the proposition q is "Harry is shorter than Paul," it is clear that p if and only if q. If p is the proposition "Paul is taller than Harry and Harry is taller than Mary" and q is the proposition "Paul is taller than Mary," it is clear that p implies q but it is not true that q implies p.

A set is any collection of objects called *elements* or *members.* The only characteristic of a set is the particular objects which it contains. Sets are generally denoted by capital letters. Lowercase letters are used mostly for the members of sets. The notation $x \in E$ means that x is a member of the set E, and x is said to be *contained in* E or to *belong to* E or, simply, to be *in* E. The negation of the statement $x \in E$, denoted by $x \notin E$, means that x is not a member of E. For example, if E is the set of positive integers, then $2 \in E$ but $-2 \notin E$. In general, a diagonal line running through a symbol usually denotes the logical "not statement:" for example, the symbol $\neq$ means "not equal to."

Sets are, in fact, completely determined by the members which they contain. In line with this idea we make the following definition of equality.

Definition 1.1: *Two sets E and D are said to be the* **same set** *or* **equal** *if they contain the same objects. This is denoted by writing $E = D$.*

Stated in a slightly different way, *the two sets E and D are defined to be equal if there is no element of E which is not an element of D and if there is no element of D which is not an element of E.* This means that E and D are equal if every element of E is an element of D and every element of D is an element of E. From a practical point of view this last form of the definition of equality is the most useful one because it is most directly related to the method most used in practice to decide if two sets are equal. It will be useful to have a special name (subset) for the situation when the first half (but not necessarily the second half) of the requirements of this definition is met by two sets.

Definition 1.2: *If there is* **no** *element of a set E which is* **not** *an element of a set D, E is said to be a* **subset** *of D, or E is said to be* **contained** *or* **included** *in D, or D is said to contain E. This is denoted by writing $E \subset D$ or sometimes $D \supset E$.*

This definition means that if $E \subset D$ then every element of E must be an element of D. Note that the symbol $\subset$ only connects sets. If D is the set of positive integers and E is the set whose elements are 1, 2, and 3, then $E \subset D$. However, it is not correct to write $2 \subset D$.

It is often very helpful to visualize sets as regions in the plane. The pictures obtained in this way corresponding to the various types of operations between sets (which will be discussed subsequently) are called Venn diagrams. Figure 1–1 illustrates the meaning of "D is a subset of E."

To avoid any of the known set theoretic paradoxes, mathematical structures are always set up in a manner which assures that there is some large but fixed set X (sometimes referred to as the *universal set*) such that all sets which arise can be considered as being either subsets of X or sets whose elements are

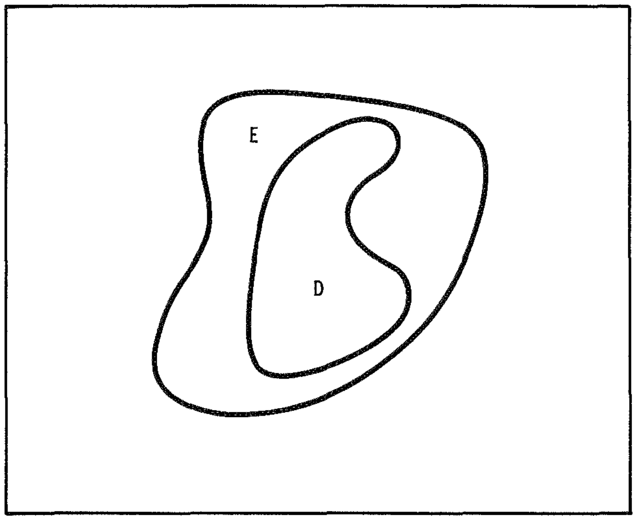


FIGURE 1–1.—Venn diagram for $D \subset E$.

subsets of X, etc. Sometimes this universal set is not mentioned explicitly in a given discussion but it will always be clear from the context that such a set exists. There is no reason why the elements of sets cannot be sets themselves! In fact, this is a situation which frequently arises in mathematics. Sets whose elements are sets are usually called families or collections in order to keep the various levels of set construction clearly in view. Actually, there is no reason why a given set D cannot simultaneously have a set E as one of its elements and an element of E as another. For example, suppose that the universal set is the set of positive integers and that the set E consists of the elements 1 and 2. If D is the set whose elements are 1, 2, 3, and E, then it is not only true that $E \in D$, but it is also true that $E \subset D$. On the other hand, if D is the set whose elements are 1, 3, and E, it is still true that $E \in D$ but it is no longer true that $E \subset D$. Thus, a set E is not a subset of a set D unless all the elements of E are included among the elements of the set D. This example illustrates a difference between elements and sets. It is, however, unfortunate that both the symbols $\in$ and $\subset$ are read as *contained in* even though they refer to *very different* things. Sometimes, then, it is necessary to decide from the context which of these two meanings is to be attributed to the phrase "contained in."

The form of the definition of equality of sets is given in the paragraph immediately following Definition 1.1 shows that, if E and D are any two sets, $E=D$ if and only if $E \subset D$ and $D \subset E$. It also follows directly from Definition 1.2, for arbitrary sets E, D, and A, that if $E \subset D$ and $D \subset A$, then $E \subset A$.

According to the wording of Definition 1.2, $E \subset E$ for every set E. On the other hand, for any reasonable set E, it is never true[1] that $E \in E$. If $E \subset D$ and there is at least one element of D which is not an element of E (i.e., the second half of the requirements of the last form of the definition of equality is not met), E is said to be a **proper subset** of D.

Some authors use the notation $E \subset D$ to mean E is a proper subset of D. If this is done, they write $E \subseteq D$ where we have written $E \subset D$. We shall *not* follow this convention here.

One usually conceives of sets as having at least some elements but as it turns out it is very desirable to consider also the set which has *no* elements. Because a set is completely determined by its elements, there is only one such set and it is denoted by the symbol $\emptyset$ and called the *empty set*. Now $\emptyset$ must be a subset of every set D, since $\emptyset$ contains no elements and therefore there is no element of $\emptyset$ which is *not* an element of D.

Clearly, if E, D, and A are any subsets of a set X, and x is any element of X, statements like "$E \subset D$," "$E \subset D$ and $D \subset A$," "$x \in E$," etc., are propositions. Large parts of mathematical proofs are composed of statements containing propositions of these types.

Before discussing the methods for specifying sets it will be helpful to introduce a certain concept from logic. Propositions usually contain the "names" of (or symbols for) certain objects. For example, the proposition "Paul is taller than Harry" discussed previously contains the names of the objects Paul and Harry. If E is a particular subset of some universal set X and t is a particular element of X, then the statement "$t \in E$" is a proposition which contains the "names" of the objects t and E. In this latter example, we can obtain a different proposition by replacing t by the name of some other element of X and, in general, can obtain an entire collection of propositions by successively replacing t by the names of all the elements of X. This collection of propositions may be described as consisting of all propositions "$x \in E$" as x varies over all the elements of X. Any statement of this type, which contains a symbol x of variable meaning in a place where the "name" of a particular object would normally occur and which becomes a proposition when x is replaced by the "name" of a member of some set D, is called a *propositional scheme* and is denoted by a symbol such as $P(x)$. The set D is called the *domain* of $P(x)$. If s is the "name" of some member of D, the proposition resulting from

[1] In fact, this situation can occur if we do not limit the size of sets as explained previously. Since this is always done in mathematics, for our purposes it is never true that $E \in E$.

replacing x by s in $P(x)$ is denoted by $P(s)$. Thus, in the preceding example, $P(x)$ is the symbol for "$x \in E$" and the domain (i.e., D) of $P(x)$ is X. Note that x serves only to save the place where the name of an object is to be inserted.

Effectively, sets are specified in one of two ways. First, if a set consists of a finite number of elements, since a set is completely determined by its elements, we can specify the set by listing its elements. When this is done, the elements are enclosed by braces and separated by commas. Thus $\{d, 1, 2, 3\}$ is the set whose elements are d, 1, 2, and 3. For sets with an "infinite" number of elements, this procedure cannot be used.

On the other hand, suppose $P(x)$ is some propositional scheme with a domain D. For each particular element $s \in D$, $P(s)$ will either be true or false. There will then be a certain subset of D, say E, which consists of all the elements x of D for which $P(x)$ is true. The set E is denoted by

$$E = \{x \in D \mid P(x)\}$$

which reads "*E is the set of all x contained in D such that P(x) (is true)*," the words in parentheses usually being omitted. Sometimes, when it is understood from the context, the domain D is omitted and we write

$$E = \{x \mid P(x)\}$$

For example, suppose $P(x)$ is the propositional scheme $x^2 = x$ and its domain is the set J of all positive integers. Then the set $\{x \in J \mid x^2 = x\}$ is the set of $x \in J$ (or the set of all positive integers x) such that $x^2 = x$. This is just the one element[2] subset $\{1\}$ of J. On the other hand, the set $\{x \in J \mid x + 1 = x\}$ is the empty set $\emptyset$.

In this manner then, every propositional scheme defines a set and since, for any set E, "$x \in E$" is a propositional scheme, every set determines a propositional scheme. In fact, this method of specifying sets includes the method of listing the elements. For example, if $E = \{1, 2, 3\}$ and J is the set of all positive integers, then

$$E = \{x \in J \mid x \in \{1, 2, 3\}\}$$

since $x \in \{1, 2, 3\}$ is a propositional scheme.

We might point out that two different propositional schemes, say $P(x)$ and

---

[2] Sometimes one element sets are called singleton sets. A distinction is always made between a one element set and the element itself. Thus $\{1\}$ is a different object than 1.

$Q(x)$, with the same domain D may define the same set. For suppose that, for each $d \in D$, $P(d)$ if and only if $Q(d)$. Then $P(x)$ and $Q(x)$ are either both true or both false at every point $x \in D$. Hence, the set of all x for which $P(x)$ is true is the same as the set of all x for which $Q(x)$ is true. That is

$$\{x \in D|P(x)\} = \{x \in D|Q(x)\} \tag{1-1}$$

On the other hand, if equation (1–1) holds, then, for any $d \in D$, either $d \in \{x \in D|P(x)\}$ in which case $d \in \{x \in D|Q(x)\}$ and hence $P(d)$ and $Q(d)$ are both true or $d \notin \{x \in D|P(x)\}$ in which case $d \notin \{x \in D|Q(x)\}$ and hence $P(d)$ and $Q(d)$ are both false. Thus, $P(d)$ if and only $Q(d)$.

In a similar way, it can be seen that the inclusion

$$\{x \in D|P(x)\} \subset \{x \in D|Q(x)\}$$

means that for all $d \in D$, $P(d)$ implies $Q(d)$.

The preceding paragraph illustrates how statements involving propositional schemes can be transformed into relations between sets. In fact, it is generally true that propositional schemes, which arise naturally in any logical reasoning process and which may involve very complicated ideas, can be replaced by sets which are much easier to think about. This is the reason why set theory is such an important tool in mathematics.

We now introduce some elementary ways of combining sets to form new sets. Note that we always assume sets under consideration are subsets of some fixed set X.

Definition 1.3: *The* **union,** $D \cup E$, *of two sets D and E is the set which consists of all elements which are either in D or*[3] *in E. Or,*

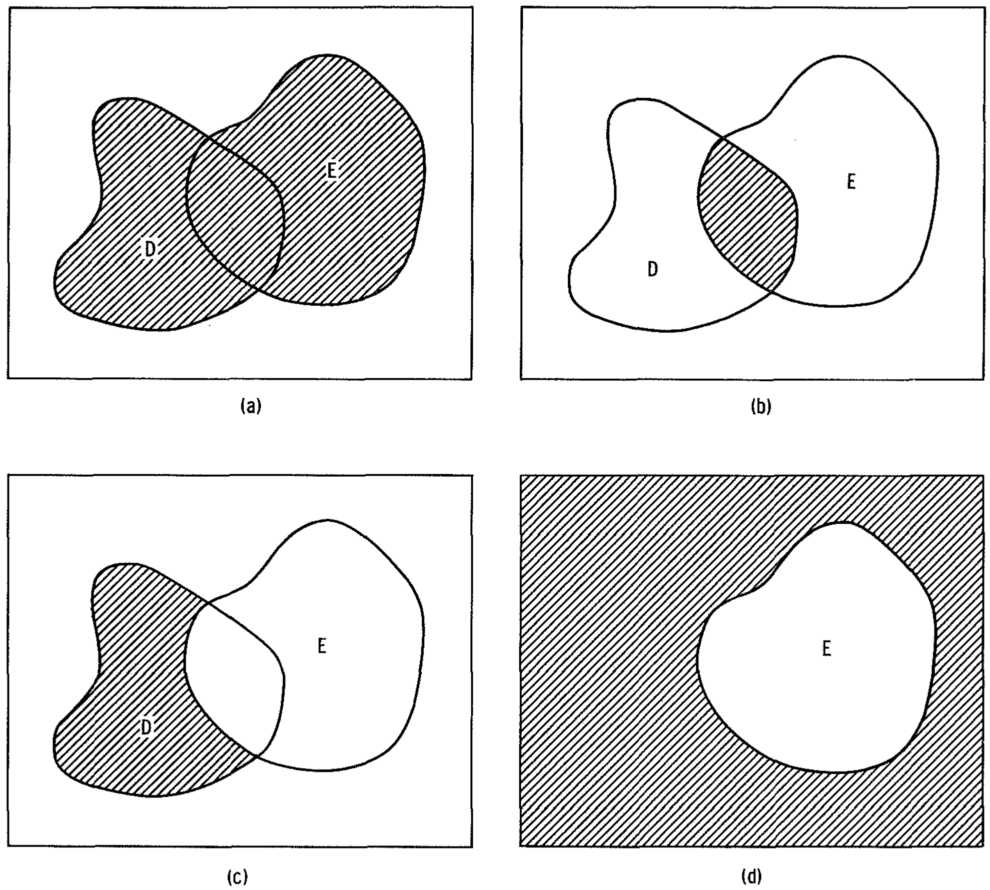
$$D \cup E = \{x|x \in D \text{ or } x \in E\}$$

This definition is illustrated in figure 1–2(a).

Definition 1.4: *The* **intersection,** $D \cap E$, *of two sets D and E is the set which consists of all elements which belong to both D and E. Or,*

$$D \cap E = \{x|x \in D \text{ and } x \in E\}$$

[3] In mathematics the word "or" is always interpreted as meaning "and/or." This is called the "inclusive or." Thus the statement "the colors are yellow or red" means that the colors may be yellow or they may be red or they may be both yellow and red.



(a) $D \cup E$. (b) $D \cap E$.
(c) $D-E$. (d) E^c.

FIGURE 1–2.—Venn diagrams for elementary set operations. (Shaded areas denote the indicated sets.)

If $D \cap E=\emptyset$, D and E are said to be **disjoint** *or* **nonintersecting.** *If $D \cap E \neq \emptyset$, D is said to intersect E.*

This definition is illustrated in figure 1–2(b).

It is easy to prove from these definitions that, for any two sets D and E, $D \subset D \cup E$, $E \subset D \cup E$, $D \cap E \subset D$, and $D \cap E \subset E$.

Definition 1.5: *The* **difference,** *$D-E$, of two sets D and E is the set which consists of all elements of D which are not elements of E. Or,*

$$D-E=\{x|x \in D \text{ and } x \notin E\}$$

If $E \subset D$, $D-E$ is called the **complement** *of E in D, or the* **complement** *of E* **relative** *to D.*

If X is the universal set, $X-E$ is called the **complement** *of E and is denoted by a superscript c; thus, $X-E=E^c$. Clearly,*

$$E^c=\{x|x \notin E\}$$

This definition is illustrated in the Venn diagrams of figures 1–2(c) and (d).

There are a number of relations which connect these operations. Some of the more important ones are listed in table 1–I. The proofs of some of the laws in table 1–I are simple consequences of their definitions. Since the intersection

Table 1–I.—Set Theoretic Identities

	Identity
Idempotent Law	$D \cup D=D \quad D \cap D=D$
Commutativity	$D \cup E=E \cup D \quad D \cap E=E \cap D$
Associativity	$(D \cup E) \cup G=D \cup (E \cup G) \quad (D \cap E) \cap G=D \cap (E \cap G)$
Distributive Law	$D \cup (E \cap G)=(D \cup E) \cap (D \cup G) \quad D \cap (E \cup G)=(D \cap E) \cup (D \cap G)$
Identity	$D \cup \emptyset=D \quad D \cap X=D$ $D \cup X=X \quad D \cap \emptyset=\emptyset$
Complements	$D \cup D^c=X \quad D-E=D \cap E^c \quad D \cap D^c=\emptyset$ $(D^c)^c=D \quad X^c=\emptyset \quad \emptyset^c=X$
DeMorgan's Law	$(D \cup E)^c=D^c \cap E^c \quad (D \cap E)^c=D^c \cup E^c$

and union of sets are just the set theoretic equivalents of the simple logic connectives, "and" and "or," the identities involving only unions and intersections can be proved by expressing them in terms of a corresponding law of logic. There is an essentially equivalent procedure to this which is better suited to our purposes since it demonstrates a frequently used technique. We will demonstrate the procedure by proving the first distributive law. Set $L=D \cup (E \cap G)$ and $R=(D \cup E) \cap (D \cup G)$. If $x \in L$, then x is either in D or in $E \cap G$.

First suppose that $x \in E \cap G$; then $x \in E$ and $x \in G$. It must also be true, therefore, that $x \in D \cup E$ and $x \in D \cup G$; hence, $x \in (D \cup E) \cap (D \cup G)$.

On the other hand, if $x \in D$, then it is certainly true that $x \in D \cup E$ and $x \in D \cup G$; that is, $x \in (D \cup E) \cap (D \cup G)$. In either case, then, $x \in L$ implies that $x \in R$. This shows, since x was an arbitrary element of L, that $L \subset R$.

Conversely, suppose $x \in R$; then, $x \in D \cup E$ and $x \in D \cup G$. Hence, if

$x \notin D$, then $x \in E$ and $x \in G$. That is, $x \in E \cap G$ so it must also be true that $x \in D \cup (E \cap G)$. On the other hand, if $x \in D$, then certainly $x \in D \cup (E \cap G)$. Since either $x \notin D$ or $x \in D$ must be true, we conclude that $x \in L$; hence, $R \subset L$. Combining this with $L \subset R$, we conclude that $L = R$.

This is an extremely detailed proof of a very simple statement, but it was included to illustrate the method.

The proofs of one of DeMorgan's laws and the second distributive law are given for a more general case in chapter 5. The rest of the entries in table 1–I are more or less direct consequences of the preceding definitions. Some of the relations in table 1–I are illustrated in the Venn diagrams of figure 1–3.

The associativity of the unions and intersections of sets shows that we can write such things as $D \cup E \cup G$ or $D \cap E \cap G$ with no danger of misinterpretation. It is clear that

$$(D \cup E) \cup G = \{x | x \in D \text{ or } x \in E \text{ or } x \in G\}$$

The associative law and consequently the omission of parentheses can be extended to the unions and intersections of any finite number of sets.

If D and E are two sets, it is easy to verify that if one of the following three relations is true the other two must be also:

$$D \subset E \tag{1–2a}$$

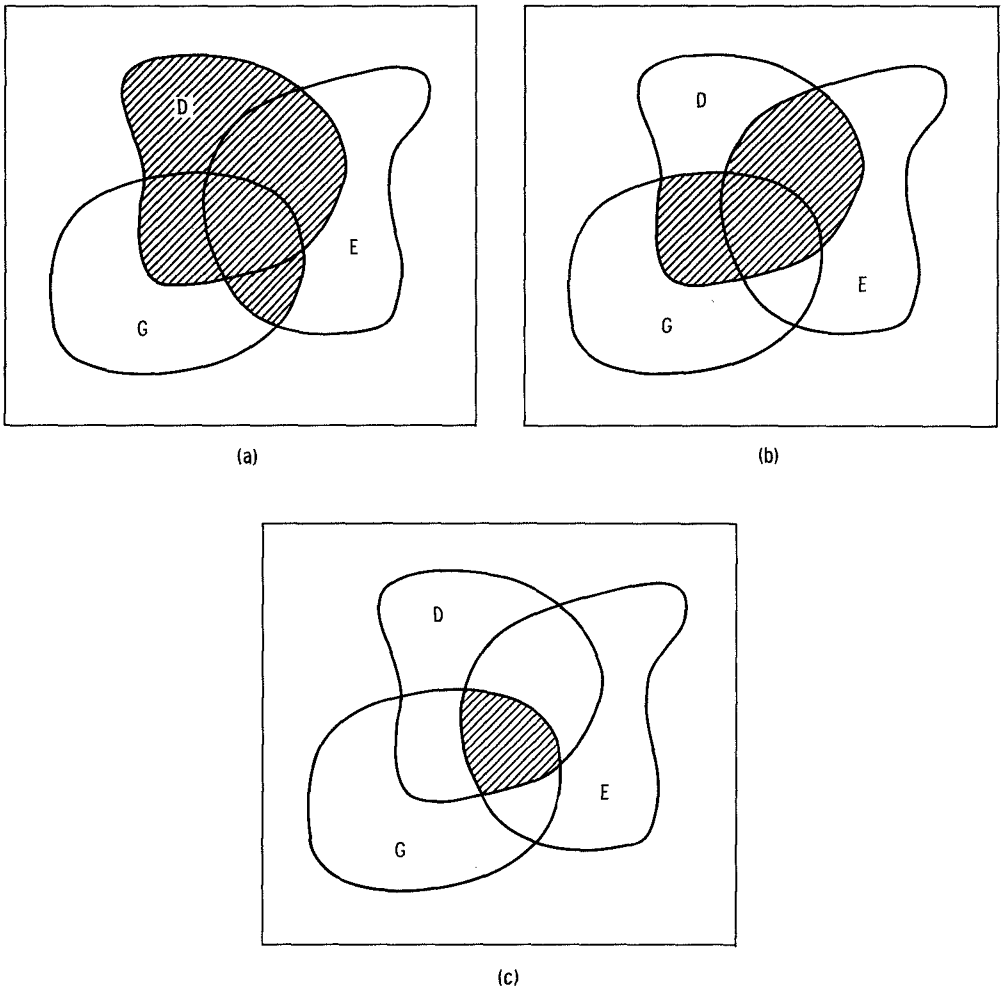
$$D \cap E = D \tag{1–2b}$$

$$D \cup E = E \tag{1–2c}$$

These can be proved very simply by using the same procedure as in the proof of the distributive law. It is quite easy to see, for any sets D, E, and G, that $D \subset G$ and $E \subset G$, if and only if $D \cup E \subset G$, and that $G \subset D$ and $G \subset E$, if and only if $G \subset D \cap E$.

Sets are defined in such a way that they have no internal organization. Thus, the set $\{p, q\}$ is the same as the set $\{q, p\}$. The need, however, arises for "sets" which do have some internal organization; that is, "sets" in which the order of the elements is relevant. A collection of two objects, in which we distinguish between the first object and the second object, is called an ordered pair. The ordered pair whose *first element* is p and whose *second element* is q is denoted by $\langle p, q \rangle$. Thus, according to this definition,

$$\langle p, q \rangle = \langle s, t \rangle \text{ if and only if } p = s \text{ and } q = t \tag{1–3}$$



(a) $D \cup (E \cap G)$. (b) $D \cap (E \cup G)$.
(c) $D \cap E \cap G$.

FIGURE 1–3.—Venn diagrams for set theoretic identities. (Shaded areas denote the indicated sets.)

and therefore

$$\langle p, q \rangle \neq \langle q, p \rangle \text{ if } p \neq q$$

An alternate definition of an ordered pair which does not introduce the concept of order but contains it could have been given. This can be accomplished by setting

$$\langle p, q\rangle = \{p, \{p, q\}\} \tag{1–4}$$

It is not hard to show that equation (1–4) satisfies the condition (1–3) and therefore constitutes an acceptable definition of ordered pair (actually some proof is needed to show this). But this is unnecessary elegance for our purposes and the definition given previously will suffice.

In a similar manner, we can define the *ordered triple* $\langle p, q, s\rangle$ to be the collection of three elements in which the first, second, and third elements are distinguished from one another, and, in general, the *ordered n-tuple* to be the collection $\langle p_1, p_2, \ldots, p_n\rangle$ in which the order of all n elements is distinguished.

It is often useful to distinguish, by special notation, sets whose elements are ordered pairs or, in general, ordered n-tuples. Thus, if D and E are any two sets, the set of all ordered pairs $\langle d, e\rangle$, whose first element is in D and whose second element is in E, is called the *direct product* or *Cartesian product* of D and E and is denoted by $D\times E$. Symbolically then, $D\times E$ is the set

$$D\times E = \{\langle d, e\rangle \mid d \in D \text{ and } e \in E\}$$

and, in general, for any n sets $D_1, D_2, \ldots, D_n$ we can define the direct product $D_1\times \ldots \times D_n$ of $D_1, D_2, \ldots, D_n$ to be the set

$$D_1\times \ldots \times D_n = \{\langle d_1, \ldots, d_n\rangle \mid d_1 \in D_1 \text{ and} \ldots d_n \in D_n\}$$

Sometimes the notation $\bigtimes_{i=1}^{n} D_i$ is used for $D_1\times \ldots \times D_n$.

For example, if $D=\{1, 2\}$ and $E=\{3, 4\}$, then $D\times E=\{\langle 1, 3\rangle, \langle 1, 4\rangle, \langle 2, 3\rangle, \langle 2, 4\rangle\}$.

There is nothing in this definition which requires that D and E be different sets. Since every point in a two-dimensional plane is located by giving exactly two numbers (i.e., the coordinates of that point), we may think of a point in the plane as being an ordered pair of real numbers: the first element of the ordered pair being the first coordinate, and the second element the second coordinate. The entire two-dimensional plane is then the set of all the ordered pairs of real numbers that can be formed. Thus, the two-dimensional plane is the direct product of the set of real numbers with itself. We shall discuss this again in a more general context in chapter 3.

It should be noted that

$$D\times E\times G \neq D\times (E\times G) \neq (D\times E)\times G$$

since

$$D \times E \times G = \{\langle d, e, g\rangle | d \in D \text{ and } e \in E \text{ and } g \in G\}$$
$$D \times (E \times G) = \{\langle d, \langle e, g\rangle\rangle | d \in D \text{ and } \langle e, g\rangle \in E \times G\}$$
$$(D \times E) \times G = \{\langle\langle d, e\rangle, g\rangle | \langle d, e\rangle \in D \times E \text{ and } g \in G\}$$

and $\langle d, e, g\rangle$, $\langle d, \langle e, g\rangle\rangle$ and $\langle\langle d, e\rangle, g\rangle$ are not the same objects. This distinction is, however, unnecessary for our purposes and, *since no contradiction can arise if we do so*, we shall consider $\langle d, e, g\rangle$, $\langle d, \langle e, g\rangle\rangle$, and $\langle\langle d, e\rangle, g\rangle$ to all be the same object—namely, the ordered triple. With this convention, *the direct product is associative.* Of course all these remarks apply to the ordered n-tuple. *The direct product is not commutative* since this contradicts the meaning of the ordered pair.

CHAPTER 2

Real Numbers

The development of a rigorous theory of the abstract concepts of analysis discussed herein requires precise specification of the properties of the real numbers. In addition, many of the abstract concepts of analysis are generalizations of certain properties of the real numbers. A better understanding of these concepts is often obtained when they are compared with the properties of the real numbers from which they came. For these reasons, the properties of the real numbers must be formalized in a manner that will subsequently be useful. Most books concerned with material at the level of this publication construct the real numbers from more fundamental concepts. In fact it is usual to construct the real numbers from the rational numbers (by the use of Dedekind cuts), the rational numbers from the natural numbers, and then to relate the natural numbers to more fundamental set theoretic concepts. For our purposes, however, it is sufficient to consider the real numbers as already given and to state their properties as axioms in a precise way. Much of the material in this chapter will be familiar to the reader although it is quite possible that he has not seen it stated in the form given herein.

Three groups of axioms are given which *completely* characterize the real numbers. The first of these, the field axioms, contains all the algebraic properties of the real numbers. The second group contains all the order properties of the real numbers, that is, those properties which have to do with one number being larger than another. As a consequence, these order properties also contain the concepts of absolute value and distance which are introduced in chapters 3 and 6.

These two groups of axioms and their consequences will be familiar to the reader, and we will use any of their consequences that are needed without making any explicit mention of how they arise. In fact, these axioms are included principally for comparison with the postulates of a normed linear space (introduced in chapter 3), which is a generalization of these groups. On the other hand, the consequences of the third group, which actually consists of a single axiom, will probably be much less familiar. For our purposes, this group is perhaps the most important of the three. This third axiom refers to the completeness properties of the real numbers. It states essentially that there are no "gaps" in the real numbers.

I. Field Axioms: *With any two real numbers, a and b, the two operations + and · each associate unique real numbers, denoted by $a+b$ and $a \cdot b$, respectively, in such a way that, if a, b, c, etc., are real numbers, the following axioms hold:*

Addition axioms:

(*A1*) $a+b=b+a$

(*A2*) $a+(b+c)=(a+b)+c$

(*A3*) *There is a number* 0 *such that for every real number a,* $a+0=a$.

(*A4*) *For every real number a there is a real number denoted by* $-a$ *such that* $a+(-a)=0$.

Multiplication axioms:

(*M1*) $a \cdot (b \cdot c)=(a \cdot b) \cdot c$

(*M2*) *There is a number* 1 *such that* $1 \neq 0$ *and, for every real number a,* $1 \cdot a=a$.

(*M3*) $a \cdot b=b \cdot a$

(*M4*) *For every real number a different from zero there is a number* a^{-1} *such that* $a \cdot a^{-1}=1$.

Distributive axiom:

(*D1*) $a \cdot (b+c)=(a \cdot b)+(a \cdot c)$

II. Order Axioms: *There is a subset $\mathscr{P}$ of the real numbers called the positive numbers such that if a and b are any numbers the following are true:*

(*O1*) *For every real number a at least one of the following must be true:* $a=0$; $a \in \mathscr{P}$; *or* $-a \in \mathscr{P}$.

(*O2*) $a \in \mathscr{P}$ *implies* $-a \notin \mathscr{P}$.

(*O3*) $a, b \in \mathscr{P}$ *implies* $a+b \in \mathscr{P}$.

(*O4*) $a, b \in \mathscr{P}$ *implies* $a \cdot b \in \mathscr{P}$.

These axioms actually define the set $\mathscr{P}$ of positive numbers.[4] However, they are entirely equivalent to defining the concept "larger than" denoted by the symbol $>$, for we need only require, for any two real numbers

$$a + (-b) \in \mathscr{P}$$

and the symbol $>$ will be the same one with which the reader is already familiar. For example, axiom *(O1)* is equivalent to the fact that for any two numbers a and b either $a > b$ or $b > a$ or $a = b$ must hold. To see this, replace $a + (-b)$ in axiom *(O1)*. The symbol $<$ is defined by

$a < b$ if and only if $b > a$

and the symbol $\geqslant$ is defined by

$a \geqslant b$ if and only if $a > b$ or $a = b$

Before giving the completeness axiom we need the following definitions.

Definition 2.1: *A set E of real numbers is said to be* **bounded above** *if there exists a real number p such that $x \leqslant p$ for every $x \in E$, and* **any** *number p with this property is said to be an* **upper bound** *of E. A set E of real numbers is said to be* **bounded below** *if there exists a real number p such that $x \geqslant p$ for every $x \in E$, and any number p with this property is said to be a* **lower bound** *of E.*

Definition 2.2: *If p is an* **upper bound** *of a set E, then p is said to be the* **least upper bound** *of E if no real number which is* **less than** *p is an upper bound of E. If p is a* **lower bound** *of a set E, p is said to be the* **greatest lower bound** *of E if no real number which is* **larger than** *p is a lower bound of E.*

Clearly, there cannot be two least upper bounds of a given set: if p and q were two least upper bounds of E such that $p \neq q$ then either $p < q$, in which case p is not an upper bound of E, or $q < p$, in which case q is not. Hence, we are justified in saying "*the*" least upper bound in Definition 2.2. With these definitions we can now introduce the completeness axiom.

---

[4] The absolute value $|a|$ of a number a is defined by

$$|a| = \begin{cases} -a \text{ if } -a \in \mathscr{P} \\ a \text{ if } -a \notin \mathscr{P} \end{cases}$$

III. Completeness Axiom: *Every nonempty set of real numbers which is bounded above has a least upper bound.*

It is not obvious that this axiom is equivalent to stating that there are no "gaps" in the real numbers. The demonstration of this fact, however, would lead too far into the foundations of number theory.

The terms *supremum* and *infimum* are used interchangeably with the terms least upper bound (lub) and greatest lower bound (glb), respectively. The least upper bound of a set E is sometimes denoted by lub E and sometimes by sup E. Similarly, the greatest lower bound of E is sometimes denoted by glb E and sometimes by inf E. It is also common to write $\operatorname{lub}_{x \in E} x$ or $\sup_{x \in E} x$ for lub E and $\operatorname{glb}_{x \in E} x$ or $\inf_{x \in E} x$ for glb E. If the set E is defined by a propositional scheme $P(x)$, the notations lub $\{x \mid P(x)\}$ or sup $\{x \mid P(x)\}$ or even $\operatorname{lub}_{P(x)} x$ or $\sup_{P(x)} x$ are used for lub E, with a similar convention of course for glb E. If the propositional scheme $P(x)$ or the set E is understood from the context, we sometimes write lub x or sup x or glb x or inf x with the obvious meaning. It is clear from Definition 2.2 and the familiar properties of the real numbers that

$$\operatorname{glb}_{x \in E} x = -\operatorname{lub}_{x \in E} (-x) \tag{2–1a}$$

$$\operatorname{lub}_{x \in E} x = -\operatorname{glb}_{x \in E} (-x) \tag{2–1b}$$

These identities are often useful in transforming statements about the supremum into statements about the infimum.

Clearly, if any set E contains a largest member y, then by definition y is an upper bound of E. But it must also be true that $y = \operatorname{lub} E$ for, if x is any real number such that $x < y$, then the element y of E is larger than x and so x is not an upper bound of E. On the other hand, if any upper bound of E, say p, belongs to E, then p is both the largest member of E and the least upper bound of E. When lub $E \in E$ or glb $E \in E$, we shall sometimes write max E in place of lub E or min E in place of glb E.

Clearly, every nonempty finite set of real numbers, say $\{x_1, x_2, \ldots, x_n\}$, has a largest element x_j and a smallest element x_i. Hence, $x_j = \operatorname{lub} \{x_1, \ldots, x_n\}$ and $x_1 = \operatorname{glb} \{x_1, \ldots, x_n\}$. It is also true that *every nonempty subset of the positive integers has a smallest element.* It is easy to see that, if $E \subset D$ and the least bound y of D exists, lub $E \leqslant$ lub D because every $x \in E$ is also contained in D and hence $x \leqslant y$ for every $x \in E$. Therefore, y is an upper bound of E and so cannot be less than lub E. These very simple properties of the infimum and supremum find many applications in the following chapters.

We shall take for granted the facts that the integers and the rational numbers are subsets of the real numbers and that for any real number p there is an integer n such that $n > p$. Actually, this last statement can easily be proved from the completeness axiom and the fact that, for every integer m, $m+1$ is an integer which is larger than m, but we shall not bother to do so here. While it is also quite easy to prove from these facts that the following statement is true, for our purposes it will be sufficient to merely list it as an axiom.

Axiom of Archimedes: *Between any two real numbers there is a rational number.*

This means that if p and q are any real numbers and, say for definiteness, $p < q$, we can find a rational number r such that $p < r < q$.

One of the properties of the real numbers is that for every real number p there is a real number q such that $q > p$; that is, there is no largest (and no smallest) real number. For many purposes, however, it is convenient to be able to talk about a largest and smallest number. This can be done if the set of real numbers is enlarged in the manner indicated in the following definition.

Definition 2.3: *The* **extended** *real number system is defined to be the set which consists of all the real numbers plus the two symbols* $+\infty$ *and* $-\infty$ *which, for every real number* p, *have the following properties:*

(1) $-\infty < p < +\infty$

(2) $p + (+\infty) = +\infty$, $p + (-\infty) = -\infty$

(3) $p \cdot (+\infty) = \begin{cases} +\infty & p > 0 \\ -\infty & p < 0 \end{cases}$

(4) $\dfrac{p}{+\infty} = \dfrac{p}{-\infty} = 0$

(5) $\begin{array}{ll} (+\infty) + (+\infty) = +\infty, & (-\infty) + (-\infty) = -\infty \\ (+\infty) \cdot (\pm\infty) = \pm\infty, & (-\infty) \cdot (\pm\infty) = \mp\infty \end{array}$

The operation $(+\infty) + (-\infty)$ is not defined. We shall also not define $0 \cdot (+\infty)$ and $0 \cdot (-\infty)$. However, sometimes the convention $0 \cdot (+\infty) = 0$ is adopted, particularly in the theory of integration.

The extended real numbers have many of the algebraic properties of the real numbers especially if the arbitrary convention $0 \cdot (+\infty) = 0$ is adopted. A notable exception is that there is no extended real number p such that $p + (+\infty) = 0$. Also, when we are dealing with the extended real numbers we can *not* conclude from the fact that $a + b = c + b$ that $a = c$. When it is desired to make explicit the distinction between the real numbers and the extended real numbers, the former will be termed *finite.*

The introduction of the symbols $+\infty$ and $-\infty$ with their order properties (i.e., $+\infty$ is larger than any real number and $-\infty$ is smaller) provides two numbers which are upper and lower bounds of every set. The following definition uses this fact to eliminate the restriction "bounded above" in the completeness axiom.

Definition 2.4: *If E is a nonempty set of extended real numbers and if, for every finite number y, there is a p of E such that $p > y$, then the least upper bound of E,* lub E, *is defined to be $+\infty$. A similar convention is adopted for the* glb E. *If $E = \{+\infty\}$, we define* glb E *to be $+\infty$, and if $E = \{-\infty\}$, we define* lub E *to be $-\infty$.*

CHAPTER 3

Vector Spaces

A favorite occupation of the nineteenth century algebraist was the generalization of quadratic and bilinear forms from three to any finite number of variables. The algebra which occurred as a consequence of this was soon interpreted as the geometry of hyperquadratics in n-dimensional space. With this interpretation many of the familiar problems in three-dimensional space suggested very obvious things to do in n-dimensional Euclidean space.

Around the turn of the century the work begun by Cantor and carried on principally by Fréchet, which we shall discuss in somewhat more detail in chapter 6, led to the concept of a *space* as being any set of points of unspecified nature subject to certain postulates. Today, just about all the research in abstract mathematics is concerned with studying one type of space or another. The postulates of the spaces that are studied are usually designed to mirror those properties of the real numbers or of the real functions of a real variable which lead to useful results when taken over into this more general setting. As already mentioned in the preceding chapter, the Field Axioms I contain all the algebraic properties of the real numbers and the Order Axioms II contain the concept of absolute value.

In this chapter, we shall first define a space whose postulates mirror some of the field axioms (the linear properties) and as a result a purely algebraic space will be obtained; that is, there will be no geometric relation between the points of the space. Next, we will add to this space some additional postulates, which mirror some of the order axioms and their consequences, in such a way that these postulates abstract the notion of absolute value. It will then be shown that the abstract spaces obtained in this manner (which are called normed linear spaces) not only contain the real numbers as a special case, but also the n-dimensional Euclidean space. In 1922 Banach introduced the normed linear space with one additional postulate. We shall not deal with this postulate, which mirrors the completeness axiom of the real numbers, until chapter 9.

Our interest here in normed linear spaces is to put them to work to study the relation between algebraic operations on the one hand and continuity and convergence on the other.

Apropos of this, we first define vector or linear space (compare with the field axioms in chapter 2).

Definition 3.1: *A real (complex)* **vector space** *is a set* V, *whose elements are called* **vectors,** *together with the set* $\mathscr{S}$ *of all real (complex) numbers and an algebraic operation, called* **addition,** *which associates with* **any** *two elements* v_1, $v_2 \in V$ *a unique vector, denoted by* v_1+v_2, *in such a way that, if* v_3 *is any vector, the following are true:*

(*A1*) $v_1+v_2=v_2+v_1$ (*Commutativity*)

(*A2*) $v_1+(v_2+v_3)=(v_1+v_2)+v_3$ (*Associativity*)

(*A3*) *There exists a vector denoted by* 0 *and called the zero vector such that, for every* $v \in V$, $v+0=v$ (*Identity*)

(*A4*) *To each* $v \in V$ *there corresponds a vector denoted by* $-v$ *such that* $v+(-v)=0$ (*Inverse*)[5]

Further an operation, called **multiplication,** *which associates with* **any** *element* $v \in V$ *and any real (complex) number* α, *a unique vector denoted by* αv, *is defined in such a way that, for any* $\beta \in \mathscr{S}$, *the following axioms hold:*

(*M1*) $\alpha(\beta v)=(\alpha\beta)v$ (*Associativity*)

(*M2*) $1v=v$ (*Identity*)

Finally, the following two distributive laws hold for any α, β, $\in \mathscr{S}$ *and any* v_1, $v_2 \in V$:

(*D1*) $\alpha(v_1+v_2)=\alpha v_1+\alpha v_2$ (*Distributive*)

(*D2*) $(\alpha+\beta)v_1=\alpha v_1+\beta v_1$ (*Distributive*)

In the context of vector spaces, the elements of $\mathscr{S}$ are often referred to as *scalars*. It should be noted that we have made no distinction between the notation for the zero vector (sometimes called the *origin*) and for the number

[5] The properties (A1) through (A4) are those of an Abelian group. Thus, every vector space is an Abelian group with respect to the operation +.

zero. However, this is common practice since no confusion arises from it. It is a direct consequence of axioms (A1) and (A3) that the vector denoted by 0 is unique: for, suppose $0'$ was another such vector; then, $v+0'=v$ for all $v \in V$. In particular, setting $v=0$ in this equality shows that

$$0+0'=0$$

But setting $v=0'$ in (A3) gives

$$0'+0=0'$$

Now (A1) shows that the left sides of these two equalities are equal; hence,

$$0=0'$$

We are therefore justified in calling 0 *the* zero vector.

It is standard practice, although not really logically consistent, to refer to the set V alone in Definition 3.1 as the vector space. After all, we have not changed V in any way just because we have in mind certain algebraic operations between its members and the members of $\mathscr{S}$. As a matter of fact, this sort of thing is quite common in mathematics and we shall meet it again—for example, in the discussion of metric spaces. Following this procedure tends to avoid an undesirable awkwardness.

Another term for vector space is *linear space*. The two terms will be used interchangeably throughout the text.

It is easily verified that the real numbers satisfy the axioms of a real vector space and the complex numbers satisfy the axioms of both real and complex vector spaces, with the usual method of adding and multiplying complex numbers. The following simple theorem is an immediate consequence of Definition 3.1.

Theorem 3.2: *Let* V *be a real* (*complex*) *vector space and suppose* $v_1, v_2, v_3 \in \mathrm{V}$. *Then*

(a) $v_1+v_2=v_3+v_2$ *implies* $v_1=v_3$ (*Cancellation Law*)

(b) $0v_1=0$

(c) *The additive inverse is unique and* $(-1)v_1=-v_1$

Proof: Part (a). This follows immediately from postulates (A2) to (A4) for, if $-v_2$ is added to both members of $v_1+v_2=v_3+v_2$, we get

$$(v_1+v_2)+(-v_2)=(v_3+v_2)+(-v_2)$$

Using (A2) gives

$$v_1+(v_2+(-v_2))=v_3+(v_2+(-v_2))$$

Using (A4) gives

$$v_1+0=v_3+0$$

and finally we see from (A3) that

$$v_1=v_3$$

Part (b). This follows from part (a) and from postulates (D2) and (A3) for if α is any scalar, then it follows from (D2) that

$$\alpha v_1=(0+\alpha)v_1=0v_1+\alpha v_1$$

and so from (A3) we see that

$$0+\alpha v_1=\alpha v_1=0v_1+\alpha v_1$$

Hence, it follows from part (a) that

$$0=0v_1$$

Part (c). We first show that the vector $-v$ is unique. Suppose w was another additive inverse of v. Then, $v+w=0$. It is clear from (A1) that $v+(-v)=v+w$, and then applying (A4) to this shows that $(-v)+v=w+v$. Thus, the first part of this theorem shows that $-v=w$.

To complete the proof of part (c) we use propositions (M2) and (D2) and part (b) to get

$$(-1)v_1+v_1=(-1)v_1+1v_1=(-1+1)v_1=0v_1=0$$

Since the additive inverse $-v_1$ is unique, we conclude from this that

$$(-1)v_1=-v_1$$

Given a vector space V, it is natural to inquire as to which subsets of V are themselves vector spaces. It is clear that not all subsets of V can be made into vector spaces by imposing the rules of addition and scalar multiplication defined on V. With this in mind we make the following definition.

Definition 3.3: *Let* V *be a vector space and let* $\mathrm{U} \subset \mathrm{V}$. *Suppose that, for every* $v_1, v_2 \in \mathrm{U}$,

$$\alpha v_1 + \beta v_2 \in \mathrm{U}$$

for all scalars α *and* β. *Then,* U *is called a* **linear subspace** *of* V *or more simply a* **subspace** *of* V.

It now follows immediately from Definitions 3.1 and 3.3 that every linear subspace of a vector space is a vector space in its own right with the same definition of addition and multiplication as in the original space.

The concept of vector space is a purely algebraic one and therefore (the reason for this will become clear in the following) it alone cannot be used to discuss continuity and convergence. Roughly speaking we must add some structure to it that allows us to introduce the notions of convergence and continuity. Apropos of this we now define a "normed" vector space.

Definition 3.4: *A real (complex) vector space* V *is said to be* **normed** *if there is associated with every element* v *of* V *a unique real number* $\|v\|$ *which has the following properties:*

(*N1*) $\|v\| \geqslant 0$; $\|v\| = 0$ *if and only if* $v = 0$
(*N2*)[6] $\|\alpha v\| = |\alpha|\ \|v\|$ *for every real (complex) number* α
(*N3*) $\|v_1 + v_2\| \leqslant \|v_1\| + \|v_2\|$ *for all* $v_1, v_2 \in V$

The number $\|v\|$ *is called the* **norm** *of* v.

It is straightforward but tedious to verify that both the real and complex numbers satisfy these axioms with the norm taken to be the absolute value. Proposition (*N3*) is known as the triangle inequality since this is a generalization of that classical concept from Euclidean geometry. This will become clear subsequently when we introduce an important and familiar type of normed vector space.

From an elementary standpoint a vector is taken to be a quantity having both magnitude and direction. When this outlook is adopted, vectors are usually visualized as arrows in a three-dimensional space with the tail of the arrow at the origin 0. The direction of the arrow is then the direction of the vector

[6] $|\alpha|$ denotes the absolute value of the number α.

and the length of the arrow represents the magnitude of the vector. The point at the head of the arrow completely specifies the vector. Therefore, one can say that this point *is* the vector. (This is done in the next definition.) In three dimensions, a vector $\boldsymbol{r}$ is just a point with coordinates x, y, and z. Vectors are added by the parallelogram law which amounts to adding the coordinates of the points corresponding to the vectors. Thus, if the head of the vector $\boldsymbol{r}_1$ is at the point x_1, y_1, z_1 and the head of the vector $\boldsymbol{r}_2$ is at the point x_2, y_2, z_2, then the head of the vector $\boldsymbol{r}_1+\boldsymbol{r}_2$ is at the point $x_1+x_2, y_1+y_2, z_1+z_2$. The product of the vector $\boldsymbol{r}$ with a real number c is defined to be the vector whose head is at the point cx, cy, and cz. The dot product of the two vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ is defined to be the real number $x_1x_2+y_1y_2+z_1z_2$. The magnitude of the vector $\boldsymbol{r}$ is $(x^2+y^2+z^2)^{1/2}$. In the next definition these geometric concepts are formalized by a set of postulates and extended to spaces with an arbitrary number of dimensions called Euclidean spaces.

Definition 3.5: *For each positive integer k, the k-dimensional* **Euclidean space** *is the set R^k of* **all** *ordered k-tuples of the form*

$$\boldsymbol{x}=\langle x_1, x_2, \ldots, x_k\rangle$$

where $x_1, x_2, \ldots, x_k$ are real numbers, with the operations of addition and scalar multiplication between all the elements of R^k defined as follows: Let

$$\boldsymbol{x}=\langle x_1, \ldots, x_k\rangle$$

and

$$\boldsymbol{y}=\langle y_1, \ldots, y_k\rangle$$

be any two members of R^k and let α be any real number. Put

$$\boldsymbol{x}+\boldsymbol{y}=\langle x_1+y_1, \ldots, x_k+y_k\rangle \tag{3-1}$$

and

$$\alpha\boldsymbol{x}=\langle \alpha x_1, \ldots, \alpha x_k\rangle \tag{3-2}$$

It is clear that $\boldsymbol{x}+\boldsymbol{y}\in R^k$ and $\alpha\boldsymbol{x}\in R^k$. We call the first of these operations **addition** *and the second* **scalar multiplication.** *We further define the* **zero vector** *or* **origin** *of R^k (denoted by* $\mathbf{0}$*) to be the k-tuple*

$$\mathbf{0}=\langle 0, \ldots, 0\rangle \tag{3-3}$$

In addition, we associate with any two elements $\boldsymbol{x}$ and $\boldsymbol{y}$ of R^k the real number

$\boldsymbol{x} \cdot \boldsymbol{y}$ defined by

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{k} x_i y_i \tag{3–4}$$

and with any element $\boldsymbol{x}$ of R^k we associate the nonnegative number $|\boldsymbol{x}|$ defined by

$$|\boldsymbol{x}| = \left(\sum_{i=1}^{k} x_i^2\right)^{1/2} = (\boldsymbol{x} \cdot \boldsymbol{x})^{1/2} \tag{3–5}$$

If $\boldsymbol{x} = \langle x_1, x_2, \ldots, x_k\rangle$ is any element of R^k, we call $\boldsymbol{x}$ a **vector** *or a* **point** *of R^k and the numbers $x_1, x_2, \ldots, x_k$ its coordinates. Relation (3–4) is called the* **inner product.**

We shall denote a vector in R^k by using boldface type for the letter, and we shall denote its coordinates by the same letter with subscripts ranging from 1 *to k.*

Now that the Euclidean spaces have been introduced, we still must show that these spaces are in fact normed linear spaces. With the definitions of addition given by equation (3–1) and of scalar multiplication by equation (3–2) and with the zero vector defined by equation (3–3), it is an easy matter (inasmuch as the real numbers satisfy the associative, distributive, and commutative laws) to verify that the postulates (A1) to (A4), (M1) and (M2), and (D1) and (D2) of Definition 3.1 are satisfied by the members of R^k. We therefore conclude that the k-dimensional *Euclidean space R^k* is a *real vector space.* Before showing that in addition R^k is a normed linear space, it is convenient to prove the next theorem.

Theorem 3.6: *Suppose that $\boldsymbol{x}$, $\boldsymbol{y} \in R^k$ and that α* is a real number. *Then,*

(a) $|\boldsymbol{x}| \geqslant 0$
(b) $|\boldsymbol{x}| = 0$ *if and only if* $\boldsymbol{x} = \boldsymbol{0}$
(c) $|\alpha \boldsymbol{x}| = |\alpha|\ |\boldsymbol{x}|$
(d) $|\boldsymbol{x} \cdot \boldsymbol{y}| \leqslant |\boldsymbol{x}|\ |\boldsymbol{y}|$
(e) $|\boldsymbol{x} + \boldsymbol{y}| \leqslant |\boldsymbol{x}| + |\boldsymbol{y}|$

Proof: Parts (a), (b), and (c) are obvious. To prove part (d) we first note that, from the definition of inner product, $(\boldsymbol{x}+\boldsymbol{y}) \cdot \boldsymbol{z} = \boldsymbol{x} \cdot \boldsymbol{z} + \boldsymbol{y} \cdot \boldsymbol{z}$. Assume neither $|\boldsymbol{x}|$ nor $|\boldsymbol{y}|$ nor $\boldsymbol{x} \cdot \boldsymbol{y}$ is zero for otherwise, by part (b), $\boldsymbol{x} \cdot \boldsymbol{y} = 0$ and the inequality is trivial. Now from part (a) we have, for *any* real number λ,

355-525 O-70—3

$$0 \leqslant |\boldsymbol{x}-\lambda\boldsymbol{y}|^2=(\boldsymbol{x}-\lambda\boldsymbol{y})\cdot(\boldsymbol{x}-\lambda\boldsymbol{y})=\boldsymbol{x}\cdot\boldsymbol{x}$$
$$+\lambda^2\boldsymbol{y}\cdot\boldsymbol{y}-2\lambda\boldsymbol{x}\cdot\boldsymbol{y}=|\boldsymbol{x}|^2+\lambda^2|\boldsymbol{y}|^2-2\lambda\boldsymbol{x}\cdot\boldsymbol{y}$$

Since λ is arbitrary, we can set

$$\lambda=\frac{|\boldsymbol{x}|}{|\boldsymbol{y}|}\,\frac{|\boldsymbol{x}\cdot\boldsymbol{y}|}{\boldsymbol{x}\cdot\boldsymbol{y}}$$

and we have

$$\lambda^2=\frac{|\boldsymbol{x}|^2}{|\boldsymbol{y}|^2}$$

Hence,

$$0\leqslant 2|\boldsymbol{x}|^2-2\,\frac{|\boldsymbol{x}|}{|\boldsymbol{y}|}\,|\boldsymbol{x}\cdot\boldsymbol{y}|$$

or

$$|\boldsymbol{x}\cdot\boldsymbol{y}|\leqslant|\boldsymbol{x}|\;|\boldsymbol{y}|$$

and part (d) is proved. It follows from part (d) that

$$|\boldsymbol{x}+\boldsymbol{y}|^2=(\boldsymbol{x}+\boldsymbol{y})\cdot(\boldsymbol{x}+\boldsymbol{y})=\boldsymbol{x}\cdot\boldsymbol{x}$$
$$+2\boldsymbol{x}\cdot\boldsymbol{y}+\boldsymbol{y}\cdot\boldsymbol{y}\leqslant|\boldsymbol{x}|^2+2|\boldsymbol{x}|\;|\boldsymbol{y}|+|\boldsymbol{y}|^2=(|\boldsymbol{x}|+|\boldsymbol{y}|)^2$$

so that part (e) is proved.

Comparing parts (a) to (c) and (e) of Theorem 3.6 with (N1) to (N3) of Definition 3.4 now shows that indeed equation (3–5) defines a norm on R^k and so makes R^k into a normed linear space.

We define R^1 to be the set of all (finite) real numbers. We have already pointed out that R^1 is a real normed linear space if we take the norm of any element in R^1 to be its absolute value.

It is clear from the definition of the direct product given in chapter 1 that

$$R^k=\underbrace{R^1\times R^1\times\ .\ .\ .\ \times R^1}_{\text{taken } k\text{-times}} \tag{3–6}$$

It also follows from the discussion at the end of chapter 1 that, for any $k\geqslant 2$,

$$R^k=R^s\times R^{k-s}\qquad 1\leqslant s\leqslant k-1 \tag{3–7}$$

For $k=2$ and $k=3$ the definitions of equations (3–1), (3–2), and (3–4) just correspond to our usual concepts of vector addition, multiplication of a vector by a scalar, and dot product, respectively. Furthermore, equation

(3–5) is the definition of the magnitude of a vector. What has been done here is that our usual geometrical concepts of vector addition, etc., have been formalized into a set of algebraic rules or postulates, which was used to define a mathematical structure. Looked at in this more abstract setting, the spaces R^k then become natural generalizations of the three-dimensional Euclidean space.

We remind the reader that a point in the plane can now be represented in either of two ways. First, it can be represented as a vector in R^2—that is, as the ordered pair $\boldsymbol{x}=\langle x_1, x_2\rangle$ where x_1 and x_2 are real numbers. Second, it can be represented as the complex number $z=x_1+ix_2$ where $i=\sqrt{-1}$.

Geometrically speaking, when the real part of z is the same number as the first element of the ordered pair $\boldsymbol{x}$ and the imaginary part of z is the same number as the second element of the ordered pair $\boldsymbol{x}$, then $\boldsymbol{x}$ and z refer to the same geometric point. If $\boldsymbol{y}=\langle y_1, y_2\rangle$ and $\eta=y_1+iy_2$, then for any real numbers α and β

$$\boldsymbol{v}=\alpha\boldsymbol{x}+\beta\boldsymbol{y}=\langle \alpha x_1+\beta y_1,\ \alpha x_2+\beta y_2\rangle$$

and

$$\xi=\alpha z+\beta\eta=(\alpha x_1+\alpha y_1)+i(\alpha x_2+\beta y_2)$$

which defines ξ. This shows that the complex number ξ represents the same point as the vector $\boldsymbol{v}$, so that it makes no difference which formalism we use to carry out these algebraic operations. They are entirely equivalent. It is clear from equation (3–5) that $|\boldsymbol{x}|=|z|$ is just the absolute value of the complex number z in the usual sense.

Essentially then, for every operation that can be performed in R^2, there is an equivalent operation with complex numbers that will yield the same result when the complex numbers and the vectors in R^2 are identified with each other in the manner indicated. So all the results obtained below for Euclidean spaces have an immediate counterpart for complex numbers.[7]

We shall denote the real part of a complex number z by $\mathscr{Re}\ z$ and the imaginary part by $\mathscr{Im}\ z$. The set of all complex numbers is usually denoted by $\boldsymbol{C}$.

[7] In fact, the complex numbers are a field, whereas R^2 is only a linear space. This in particular implies that multiplication is defined *between* members of the set of complex numbers.

CHAPTER 4

Functions and Relations

It appears that the word "function" was first used by Descartes in 1637 to refer to a positive integral power of some real variable. After a considerable period of time, the term was taken up by Leibnitz to refer to any numerical quantity connected with a curve. The concept of function was next modified by James Bernoulli who regarded it as being any algebraic expression involving a single variable. Euler generalized this concept of function to include algebraic expressions involving any number of real variables and any expression that could be generated from algebraic expressions by the operations of composition, quadrature, and forming infinite series. Euler did not realize the full implications of this definition, and it was Fourier who demonstrated that much more general functions than Euler had thought possible arose as the sums of infinite series. Attempts to give a definition which was meaningful for this large class of relations led Dirichlet to give a definition of function which no longer required that any explicit formula be involved in the definition but only required that a rule of correspondence between numbers be given. He coined the terms "domain" and "range," and he was the first to impose the restriction that the rule of correspondence assign only a single number to each number in the domain of definition of the function. This is essentially the concept of function that is used in elementary calculus today. Dirichlet, however, still required that a function relate numerical quantities, and he did not seem to apply the term function to the rule of correspondence itself.

With the introduction of set theory into mathematics, the term function has come to refer to any rule which associates with an element of a given set a single element of another set. Finally, attempts to formalize mathematics entirely in terms of set theoretic concepts have led many modern authors to define functions in terms of certain sets of ordered pairs.

Before giving a precise definition of function, let us consider an example in which D denotes the set of all books in the NASA Lewis library. Let E be the set of positive integers. Then we can associate a unique member of E with each book in D, namely, the number of pages which that book contains. This scheme defines a function from the set D into the set E.

Definition 4.1: *If D and E are any two sets and there is some scheme or rule whereby, with* **each** *element $x \in D$, there is associated a* **unique** *element $y \in E$, then this scheme is said to be a* **function** *f from D* **into** *E. The element $y \in E$ that is associated in this way with an element $x \in D$ is denoted by $f(x)$ and is called the* **value** *of f at x. The set D is called the* **domain** *of f. The notation $f: D \rightarrow E$* **means that f is a function from the set D into the set E.** *The function f is said to be* **defined on** *D.*

The sets D and E appearing in this definition may of course both be the same set. We shall sometimes say that f is a function from D *to* E instead of from D *into* E.

In various contexts the terms *mapping*, *transformation*, and *operator* are used for function. We emphasize the fact that a function must associate a **single** value with **each** member of its domain. According to Definition 4.1, the function f is completely specified only when the rule connecting *all* the elements of a set D with the elements of a set E is given. It is important in modern mathematics to think of the **entire** function f as a single object and to always make a distinction between the function f and any one of its values, say $f(x)$.

If the domain of a function f is a subset of the direct product of two sets, say A and E, it is common practice to denote the value of f at the point $\langle a, e\rangle \in A \times E$ by $f(a, e)$ instead of by $f(\langle a, e\rangle)$. The function f is usually referred to as a function of two variables. Hence a function of two real variables is a function whose domain is a subset of $R^2 = R^1 \times R^1$. More generally a function of n variables is one whose domain is a subset of the direct product of n sets, say $A_1, \ldots, A_n$, and its value a point $\langle a_1, \ldots, a_n\rangle \in \bigtimes_{i=1}^{n} A_i$ is denoted by $f(a_1, \ldots, a_n)$.

If $f: D \rightarrow E$ (sometimes read "f maps D into E"), the set $\{\langle x, f(x)\rangle \mid x \in D\}$ is called the *graph* of the function f. In modern writing a function is frequently defined to be its graph. For our purposes it makes no difference which of these concepts is used as the definition of function.

It is *not* consistent with Definition 4.1 to define a function f from the set of all real numbers to the set of all real numbers by the scheme

$$f(x) = \tan^{-1} x \qquad \text{(for every real number } x) \tag{4–1}$$

the reason being that $\tan^{-1}$ associates more than *one* real number with each real number x and therefore does not lead to an *unambiguous* scheme. It is, however, consistent with Definition 4.1 to define a function f from the set of all real numbers to the set of all numbers lying between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$ by equation (4–1). It is also *not* in accord with Definition 4.1 to define a function f from the set of all real numbers to the set of all real or even extended real numbers by the scheme

$$f(x)=\frac{1}{x} \qquad \text{(for every real number } x) \tag{4–2a}$$

for $f(0)$ is not defined by equation (4–2a)[8] and a function must be defined at *every* point of its domain. A function into the extended real numbers is, however, defined in a proper way if we replace equation (4–2a) by

$$f(x)=\begin{cases} \dfrac{1}{x} & \text{for every real number } x\neq 0 \\ +\infty & \text{for } x=0 \end{cases} \tag{4–2b}$$

This is contrary to the older usages of the term function which allowed equation (4–2a) to define a function that was said to have an infinite discontinuity at $x=0$.

For any two sets A and E, the function $f:A\times E\rightarrow A$ defined by

$$f(\langle x, y\rangle)=x \qquad \text{for all } \langle x, y\rangle\in A\times E$$

is sometimes called the first projection in $A\times E$. Similarly, the function $g:A\times E\rightarrow E$ defined by

$$g(\langle x, y\rangle)=y \qquad \text{for all } \langle x, y\rangle\in A\times E$$

is sometimes called the second projection in $A\times E$.

Suppose D and E are any two sets and $f:D\rightarrow E$ and $g:D\rightarrow E$. Then, according to our definition, f and g are the same function if and only if

$$f(x)=g(x) \qquad \text{for every } x\in D$$

In this case we write $f=g$.

In the following definition, the notation $\exists\, x$ is used to mean "there exists an x."

[8] This of course follows from the fact that division by zero is not defined even in the extended number system.

Definition 4.2: *If* $f:D \rightarrow E$ *and* A **is any subset of** D, *the* **image of** A **under** f *is defined to be the set* $f(A)$, *which consists of all* $y \in E$ *such that* $y=f(x)$ *for some* $x \in A$. *More formally, the image of* A *is the set*

$$f(A)=\{y \in E \mid (\exists x \in A) \quad \text{for which } y=f(x)\}$$

The set $f(D)$ *is called the* **range** *of* f. *It is clear that, in general,* $f(D) \in E$. *If it happens that* $f(D)=E$, *then* f *is said to map* D **onto** E *and the function* f *is said to be a* **surjective** *function from* D *to* E.[9]

In other words, a function $f:D \rightarrow E$ is surjective, or onto, if and only if, for every $y \in E$, there is some $x \in D$ such that $y=f(x)$. Notice that the statement "f maps D *onto* E" is more specific than the statement (cf., Definition 4.1) "f maps D *into* E." The concept of image is illustrated in figure 4–1.

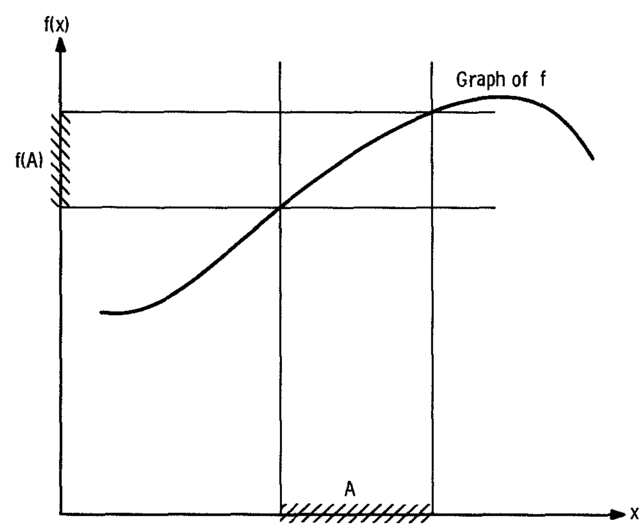


FIGURE 4–1.—Image of a set of real numbers under a real valued function.

If D and E are both sets with a finite number of elements, it is obvious that if E has more elements than D, the mapping $f: D \rightarrow E$ cannot be surjective. For example, the function $g: A \rightarrow G$ where $A=\{a, b, c\}$ and $G=\{1, 2\}$ defined by

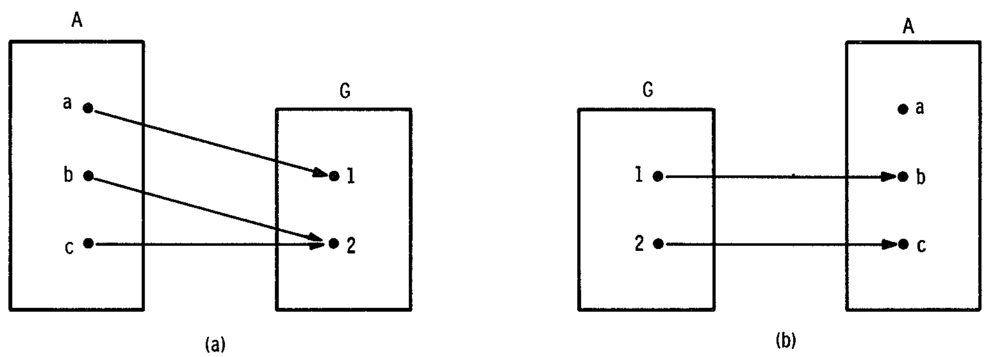
[9] The property of being surjective is not a property of f alone but a property of f and E.

$$g(a)=1$$
$$g(b)=2$$
$$g(c)=2$$

is surjective, whereas the function $h: G \to A$ given by

$$h(1)=b$$
$$h(2)=c$$

is not surjective. These two functions are illustrated in figures 4–2 (a) and (b).



(a) $g: A \to G$.

(b) $h: G \to A$.

FIGURE 4–2.—Schematic representations of g and h.

Clearly any function $f: D \to E$ may be considered as being a surjective function from D to $f(D)$. Also if A_1 and A_2 are any two subsets of the domain D of f, then $A_1 \subset A_2$ implies $f(A_1) \subset f(A_2)$. Finally it is an immediate consequence of Definition 4.2 that $f(\emptyset)=\emptyset$.[10]

The function f from the set of all real numbers to the set of all real numbers defined by

$$f(x)=x^3 \qquad \text{for all real numbers } x$$

is surjective. However, the function g from the set of all real numbers into the set of all real numbers defined by

[10] If $f: D \to E$, $A \subset D$, and E is a subset of the real or extended real numbers, it is common practice to write $\sup\limits_{x \in A} f(x)$ and $\inf\limits_{x \in A} f(x)$ in place of $\sup f(A)$ and $\inf f(A)$, respectively.

$$g(x) = x^2 \qquad \text{for all real numbers } x$$

is not surjective because, given any negative number y, there is no real number x for which $y = x^2$.

Definition 4.3: *For any two sets D and E, the mapping* **induced** *by a function $f: D \to E$ is defined to be the function $\not{f}$ from the* **collection** $\mathscr{P}_D$ **of all subsets** *of D into the* **collection** $\mathscr{P}_E$ **of all subsets** *of E such that, for each subset $A \in \mathscr{P}_D$, the value of the function $\not{f}$ at A is $f(A)$, the image of A under f.*[11]

The symbol f is also used for the induced mapping $\not{f}$ and it is usually necessary to infer from the context which function f refers to. It is clear, from Definition 4.2, that the mapping induced by a function $f: D \to E$ is surjective if and only if its range is the collection of all subsets of E. The following example serves to illustrate some of these concepts.

Let $D = \{a, b\}$ and $E = \{1, 2\}$. The function $f: D \to E$ defined by

$$f(x) = \begin{cases} 1 & \text{for } x = a \\ 1 & \text{for } x = b \end{cases}$$

is not surjective. This function is illustrated in figure 4–3. The subsets of D are

$$\emptyset \quad \{a\} \quad \{b\} \quad \{a, b\}$$

and the subsets of E are

$$\emptyset \quad \{1\} \quad \{2\} \quad \{1, 2\}$$

The values of the induced mapping are

$$f(\emptyset) = \emptyset \qquad f(\{b\}) = \{1\}$$

$$f(\{a\}) = \{1\} \qquad f(\{a, b\}) = \{1\}$$

and this mapping, like the function f, is also not surjective because its range does not include the subsets $\{2\}$ and $\{1, 2\}$ of E.

[11] The collections $\mathscr{P}_D$ and $\mathscr{P}_E$ are called the power sets of D and E, respectively.

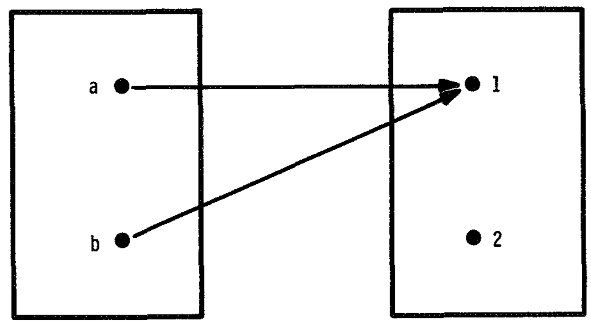


FIGURE 4–3.—Schematic representation of $f:\{a, b\} \rightarrow \{1, 2\}$.

Definition 4:4: *If $f: D \rightarrow E$, and A* **is any subset of** *E, the* **inverse image** *of A under f is defined to be the set $f^{-1}(A)$, which consists of all points $x \in D$ for which $f(x) \in A$. In symbols, the inverse image is the set*

$$f^{-1}(A) = \{x \in D | f(x) \in A\}$$

If the inverse image under f of every one element subset of E contains at most one element of D, then f is said to be **injective** *or f is said to be a* **one-to-one mapping** *from D to E.*

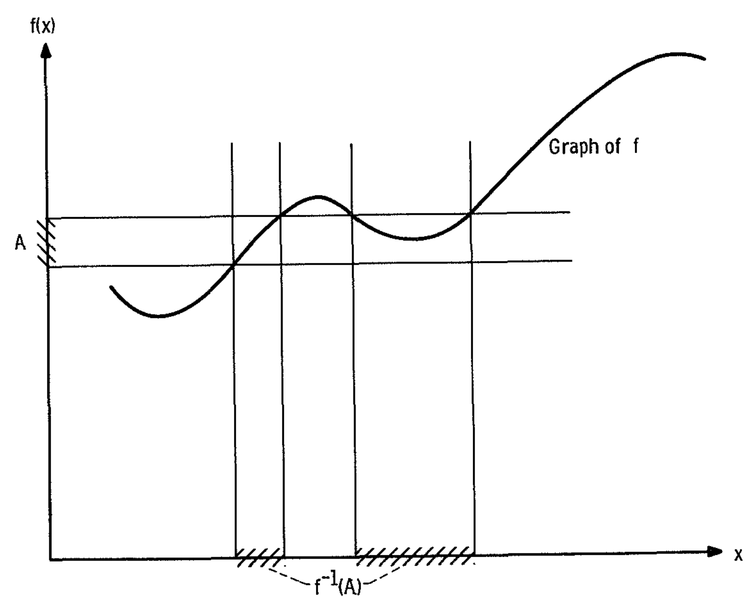


FIGURE 4–4.—Inverse image of a set of real numbers under a real valued function.

The concept of inverse image is illustrated in figure 4–4.

Clearly, if a function $f: D \to E$ is injective, there cannot be two distinct points of D, say x_1 and x_2, such that $f(x_1)=f(x_2)$, for otherwise the inverse image under f of the one element set $\{y\}$ where $y=f(x_1)=f(x_2)$ would contain more than one element, namely, x_1 and x_2. On the other hand, if for any two points of D, say x_1 and x_2, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$, then the inverse image of any one element subset $\{y\}$ of E can have at most one point for, if $f^{-1}(\{y\}) = \{x_1, x_2\}$ and $x_1 \neq x_2$, then we would have $y=f(x_1)$ and $y=f(x_2)$ which is impossible since $f(x_1) \neq f(x_2)$. Hence, we conclude that *a function* $f: D \to E$ *is injective if and only if, for any* $x_1, x_2 \in D$, $x_1 \neq x_2$ *implies* $f(x_1) \neq f(x_2)$.

We can express this last concept in still another way. To this end, suppose that p and q are two propositions (recall that propositions are either true or false but not both). It was shown in chapter 1 (see p. 2) that the two statements "whenever the proposition p is true, then the proposition q must also be true" and "whenever q is false, p must also be false" mean the same thing. They are known as *contrapositives* of one another. Now, we saw that a function f is injective if and only if, for any two points x_1, x_2 in its domain, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$. This statement can be replaced by its contrapositive and we may say that f *is injective if and only if* $f(x_1)=f(x_2)$ *implies* $x_1=x_2$ *for any* x_1, x_2 *in the domain of* f. Situations often occur in the proofs of theorems when it is easier to prove the contrapositive of a statement rather than the statement itself.

We mention in passing that, as in the case of induced mappings, the inverse images of the various subsets of E under a mapping f are the values of a function from the collection of all subsets of E into the collection of subsets of D.

We return to the last example in which $D=\{a, b\}$, $E=\{1, 2\}$, and $f: D \to E$ is defined by

$$f(x)=\begin{cases} 1 & \text{for } x=a \\ 1 & \text{for } x=b \end{cases}$$

The inverse images under f of the subsets of E are

$$f^{-1}(\emptyset)=\emptyset \qquad f^{-1}(\{2\})=\emptyset$$

$$f^{-1}(\{1\})=\{a, b\} \qquad f^{-1}(\{1, 2\})=\{a, b\}$$

The function f is not injective since $f^{-1}(\{1\})$ contains more than one element of D. Also $f(a)=f(b)$, but $a \neq b$.

If D and E are any two sets, $f: D \to E$, and H_1 and H_2 are any two subsets of E, then clearly $H_1 \subset H_2$ implies $f^{-1}(H_1) \subset f^{-1}(H_2)$.

If $f: D \to E$ and $y \in E$, it is common practice to write $f^{-1}(y)$ in place of $f^{-1}(\{y\})$. We shall avoid this convention here since it can easily be confused with the inverse function which is defined subsequently.

Before giving this definition, however, it is helpful to establish the following properties of injective functions.

If $f: D \to E$ is injective, then

(a) For each $y \in f(D)$, there is one and only one $x \in D$ such that $y = f(x)$. This is an immediate consequence of the facts that $f(D)$ is the range of f and that f is injective.

(b) For every $x \in D$, there is a $y \in f(D)$ such that $y = f(x)$. This is a consequence of the fact that $f(x) \in f(D)$ for every $x \in D$.

(c) $x_1 = x_2$ implies $y_1 = y_2$ if $y_1 = f(x_1)$ and $y_2 = f(x_2)$. This is a consequence of the fact that the value of a function at a given point of its domain is unique. It follows from (a) that each element $y \in f(D)$ is uniquely expressible in the form $f(x)$ for exactly one element x of D. We can therefore define in a natural way a mapping from $f(D)$ into D by taking x as the *value* of this mapping at the point $y = f(x) \in f(D)$. It follows from (b) that this mapping is onto D and (c) shows that this mapping is injective. In view of these remarks we make the following definition.

Definition 4.5: *If $f: D \to E$ and f is injective, the function $f^{-1}: f(D) \to D$ which associates, with each $y \in f(D)$, the element $x \in D$ such that $y = f(x)$ is called the* **inverse mapping** *or simply the* **inverse** *of f. If a mapping f is injective, its inverse is said to exist.*

The relation between the mapping f and f^{-1} is suggested by figure 4–5. Notice that the domain of f^{-1} is the range of f and the range of f^{-1} is the domain of f. Also notice from the definition that if the inverse of a given mapping exists, then it is unique. In addition, since the inverse mapping is injective and since $x = f^{-1}(y)$ implies that $y = f(x)$, it is clear that the inverse mapping of f^{-1} (i.e., $(f^{-1})^{-1}$) is just f.

Since the inverse image of a set and the inverse mapping are both denoted by f^{-1}, it is sometimes necessary to exercise some caution so that these two meanings are not confused.

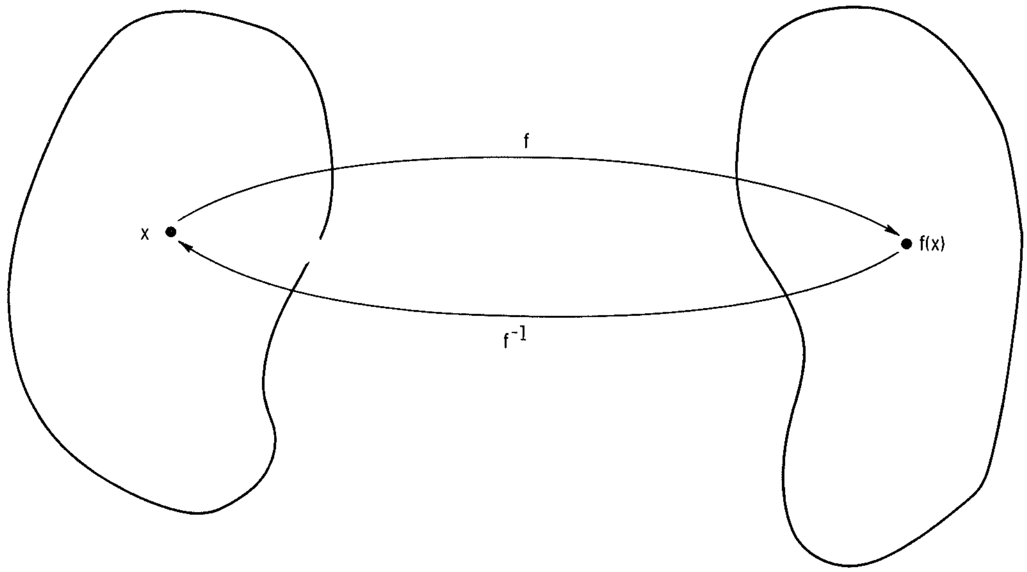


FIGURE 4–5.—Relation between an injective function and its inverse.

Again consider the sets $D=\{a, b\}$ and $E=\{1, 2\}$. This time define the function $g:D\rightarrow E$ by

$$g(a)=2$$
$$g(b)=1$$

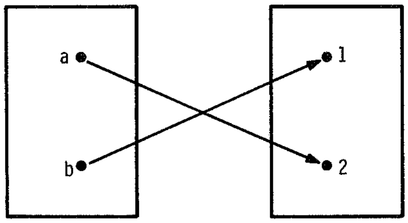
This function is evidently both surjective and injective so the inverse function g^{-1} exists and its domain $g(D)$ is equal to E. Thus, we have

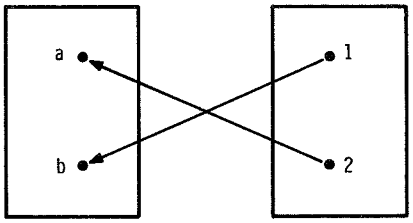
$$g^{-1}(1)=b$$
$$g^{-1}(2)=a$$

This function is shown in figure 4–6.

It is clear that, if D and E are any sets, each containing only a finite number of elements, and D contains fewer elements than E, no surjective function from D to E can be defined whereas, if D contains a larger number of elements than E, no injective function from D to E can be defined.

Definition 4.6: *A function which is both surjective and injective is said to be* **bijective** *or a* **bijection.**





(a) $g : \{a, b\} \rightarrow \{1, 2\}$. (b) $g^{-1} : \{1, 2\} \rightarrow \{a, b\}$.

FIGURE 4–6.—Schematic representations of functions g and g^{-1}.

When a function $f : D \rightarrow E$ is a bijection, the following are true:

(a) The fact that f is a function implies that a *single* element of E is associated with *each* element of D.

(b) The fact that f is surjective implies that every element of E is associated with *at least one* element of D.

(c) The fact that f is injective implies that every element of E is associated with *at most one* element of D.

Thus, every element of E is associated with exactly one element of D and there is no element of D which is not associated with some element of E.

If D is any set, the mapping $i : D \rightarrow D$ which associates each element of D with itself is called the *identity map of* D. This mapping is clearly a bijection.

Clearly if $f : D \rightarrow E$ is a bijection, its inverse mapping f^{-1} is a bijection from E to D. Consider, for example, the function $g : D \rightarrow E$ where D and E are subsets of the real numbers, and g is defined by

$$g(x) = x^2 \qquad \text{for all } x \in D$$

(a) If D and E are both taken to be the set of all nonnegative numbers, then g is a bijection and its inverse g^{-1} is defined on E by

$$g^{-1}(y) = \sqrt{y} \qquad \text{for all } y \in E$$

(b) If D is taken to be the set of all nonnegative numbers and E is taken to be the set of all real numbers, then g is injective but not surjective. The domain of the inverse is still the set of all nonnegative numbers.

(c) If D is taken to be the set of all real numbers and E is taken to be the set of all real numbers, then g is neither surjective nor injective. Its inverse does not exist.

There are a number of relations connecting the images of sets under mappings and the binary set operations. A few of the more common ones are listed in table 4–I.

Table 4–I.—Relations Connecting Images, Inverse Images, and Binary Set Operations

$[f: D \to E; A, A_1, A_2 \subset D; H, H_1, H_2 \subset E]$

....	$f^{-1}(H^c)=(f^{-1}(H))^c$
$f^{-1}(f(A)) \supset A$	$f(f^{-1}(H)) \subset H$
$f(A_1 \cup A_2)=f(A_1) \cup f(A_2)$	$f^{-1}(H_1 \cup H_2)=f^{-1}(H_1) \cup f^{-1}(H_2)$
$f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$	$f^{-1}(H_1 \cap H_2)=f^{-1}(H_1) \cap f^{-1}(H_2)$

In order to demonstrate the procedures involved in obtaining the relations in table 4–I, we shall prove that $f(f^{-1}(H)) \subset H$. To this end, let

$$A=f^{-1}(H)=\{x \in D \mid f(x) \in H\}$$

Then

$$f(f^{-1}(H))=f(A)=\{y \in E \mid (\exists\, x \in A) \text{ for which } y=f(x)\}$$

Now, if y is any point of $f(f^{-1}(H))$, this implies $y=f(x)$ for some $x \in A$. But $x \in A$ means that $f(x) \in H$; hence $y \in H$. Since y was any point of $f(f^{-1}(H))$, we conclude that $f(f^{-1}(H)) \subset H$.

A relation typical of those listed in the third and fourth rows is proved in a more general setting in chapter 5. The other relations listed in the table are obtained in a more or less similar fashion.

To see why the equality sign does not hold for the entry in the fourth row, first column, suppose A_1 and A_2 are disjoint sets. Then $f(A_1 \cap A_2)=\emptyset$. On the other hand, if f is not injective, there are points $x_1 \neq x_2$ of D such that $f(x_1)=f(x_2)$. Suppose that $A_1=\{x_1\}$ and $A_2=\{x_2\}$; then, $f(\{x_1\} \cap \{x_2\})=\emptyset$ but $f(A_1) \cap f(A_2) \neq \emptyset$ since $f(x_1) \in f(A_1) \cap f(A_2)$. The next theorem shows when the inclusion signs ($\subset$ and $\supset$) in the relations listed in the second row of table 4–I can be replaced by equal signs.

Theorem 4.7: *Let $f: D \rightarrow E$.*
(a) *f is surjective if and only if $f(f^{-1}(H))=H$ for every set $H \subset E$.*
(b) *f is injective if and only if $f^{-1}(f(A))=A$ for every set $A \subset D$.*

Proof: Part (a). Suppose f is surjective. Put

$$A=f^{-1}(H)=\{x \in D \mid f(x) \in H\}$$

Let y be any point of H. Since f is surjective, $y=f(x)$ for some $x \in D$ but A is the set of all x for which $f(x) \in H$. Hence $x \in A$. Now $f(A)=\{y \in E \mid (\exists x \in A)$ for which $y=f(x)\}$ and this shows that $y=f(x) \in f(A)$. Since y was any point of H we conclude that $H \subset f(A)=f(f^{-1}(H))$. Combining this with the relation given in the second row, second column of table 4–I shows that $f(f^{-1}(H))=H$.

Conversely, suppose that $f(f^{-1}(H))=H$ for every $H \subset E$. Let y be any point of E. *Then* $f(f^{-1}(\{y\}))=\{y\}$. Since $f(\emptyset)=\emptyset$, we conclude that $f^{-1}(\{y\}) \neq \emptyset$. Hence there exists at least one $x \in D$ such that $y=f(x)$; that is, $y \in f(D)$. Since y was any point of E, we conclude that $E \subset f(D)$. Hence, $E=f(D)$ and so f is surjective.

Part (b). Suppose f is injective. Set

$$H=f(A)=\{y \in E \mid (\exists x \in A) \qquad \text{for which } y=f(x)\}$$

Then

$$f^{-1}(f(A))=f^{-1}(H)=\{x \in D \mid f(x) \in H\}$$

Let x_1 be any member of $f^{-1}(f(A))$. Then $x_1 \in f^{-1}(H)$. Hence, $f(x_1) \in H$ and, because $H=f(A)$, there exists an $x_2 \in A$ such that $f(x_1)=f(x_2)$. It now follows, since f is injective, that $x_1=x_2$. Since x_1 was any point of $f^{-1}(f(A))$, this shows that $f^{-1}(f(A)) \subset A$. Combining this with the entry in the first column, second row of table 4–I shows that $f^{-1}(f(A))=A$.

Conversely, suppose $f^{-1}(f(A))=A$ for every $A \subset D$ and suppose for any two points $x_1, x_2 \in A$ that $f(x_1)=f(x_2)$. Now $f(x_2) \in f(\{x_2\})$ and hence $f(x_1) \in f(\{x_2\})$. Since

$$f^{-1}(f(\{x_2\}))=\{x \in D \mid f(x) \in f(\{x_2\})\}$$

we see that $x_1 \in f^{-1}(f\{x_2\}))$ and by assumption, $f^{-1}(f(\{x_2\}))=\{x_2\}$. Hence, $x_1=x_2$ and this shows that f is injective.

355-525 O-70—4

Theorem 4.8: *Let $f: D \to E$.*

(a) *If f is injective, then the mapping induced by f is an injective function from the collection of all subsets of D to the collection of all subsets of E.*

(b) *If f is surjective, then the mapping induced by f is a surjective function from the collection of all subsets of D to the collection of all subsets of E.*

(c) *If f is bijective, then the mapping induced by f is a bijective function from the collection of all subsets of D to the collection of all subsets of E.*

Proof: Part (a). Suppose A_1 and A_2 are subsets of D such that $f(A_1) = f(A_2)$. We need only show that this implies that $A_1 = A_2$ to prove that the induced mapping is injective. Hence suppose f is injective. If x is any point of A_1, then $f(x) \in f(A_1) = f(A_2) = \{f(y) | y \in A_2\}$. This shows that for some $y \in A_2$, $f(x) = f(y)$ and since f is injective this shows that $A_1 \subset A_2$. The reverse inclusion is obtained in exactly the same way by picking a point in A_2. Hence $A_1 = A_2$.

Part (b). Suppose now that f is surjective. We must show that the mapping induced by f is onto the collection of all subsets of E. That is, we must show that, if H is a subset of E, then $H = f(A)$ for some subset A of D. Since f is surjective, Theorem 4.7(a) shows that this requirement is met if we take A to be the set $f^{-1}(H)$.

Part (c). This follows immediately from (a) and (b).

Definition 4.9: *If $f: D \to E$ and $A \subset D$, a function $g: A \to E$ is said to be the* **restriction of** f **to** A *and f is said to be an* **extension of** g **to** D *if for every $x \in A$, $g(x) = f(x)$.*

It is clear that the restriction of a function to a given set is unique, but there is no natural way of defining a unique extension of a given function. We see that if H is any subset of A and g is the restriction of f to H, then $f(H) = g(H)$. Clearly the restriction of any injective function is also injective. If A is a nonempty subset of a set D, the restriction to A of the identity map of D is denoted by j_A and is called the *natural injection* of A *into* D. It is not hard to see for any $E \subset D$ that

$$j_A^{-1}(E) = E \cap A \tag{4–3}$$

Definition 4.10: *Suppose that there is a scheme whereby some of the elements of a given set A are related in some manner to other elements of A. If an element $x \in A$ is related to an element $y \in A$ by this scheme, we write*

$x \sim y$. In addition, suppose this relation between the elements of A, which is also denoted by $\sim$, has the following properties:

(a) *For every $x \in A$, $x \sim x$.*
(b) *$x \sim y$ implies $y \sim x$.*
(c) *$x \sim y$ and $y \sim z$ implies $x \sim z$.*

Then $\sim$ is called an **equivalence relation in** *A.*

For example, suppose the set A is a family (of people). Then "is the same age as" is a relation between the members of this set which is an equivalence relation. As another example, consider the set of all triangles in a plane. Then each of the following is an equivalence relation in this set:

is similar to
is congruent to
has the same area as
has the same perimeter as

A relation which satisfies (a) is called *reflexive*, one which satisfies (b) is called *symmetric*, and one which satisfies (c) is called *transitive*. Hence, equivalence relations are sometimes called SRT relations.

If A, D, and E are sets and we are given two functions $f : A \to D$ and $g : D \to E$, it is possible to define in a natural way a mapping from A into E in terms of these two functions. For if x is any point of A, $f(x)$ is a uniquely determined point of D. Since g is defined on D, $g(f(x))$ is a uniquely determined point of E. Thus the scheme which associates with each $x \in A$ the unique element $g(f(x))$ of E is a function from A into E. Apropos of these remarks we make the following definition.

Definition 4.11: *If A, D, and E are sets, $f : A \to D$ and $g : D \to E$, the function $h : A \to E$ defined by*

$$h(x) = g(f(x)) \qquad \text{for all } x \in A$$

is called the **composition** *of g and f and is denoted by $g \circ f$.*

It should be clear that the composition of two mappings g and f can only be defined if the range of f belongs to the domain of g.

In order to illustrate the definition, let $A = \{w, x, y, z\}$, $D = \{1, 2, 3\}$, and $E = \{p, q\}$, and let $f : A \to D$ and $g : D \to E$ be defined, respectively, by

$$f(w)=2$$
$$f(x)=1$$
$$f(y)=1$$
$$f(z)=3$$

and

$$g(1)=q$$
$$g(2)=p$$
$$g(3)=q$$

The composition $g\circ f$ is obtained by the following calculation:

$$g\circ f(w)=g(f(w))=g(2)=p$$
$$g\circ f(x)=g(f(x))=g(1)=q$$
$$g\circ f(y)=g(f(y))=g(1)=q$$
$$g\circ f(z)=g(f(z))=g(3)=q$$

We have in fact already encountered an example of the composition of two mappings, for Definition 4.5 shows that if $f:D\to E$ and f^{-1} exists, then

$$f^{-1}(f(x))=x \qquad \text{for every } x\in D$$

The definitions of composition and of equality of two mappings now show that

$$f^{-1}\circ f=i$$

where i is the identity map of D.

If $f:D\to E$, A is a nonempty subset of D and j_A is the natural injection of A into D, then $f\circ j_A$ is the restriction of f to A.

If $f:A\to D$, $g:D\to E$, and $h=g\circ f$, it is clear that $h(A)=g(f(A))$ so that, in particular, if g and f are both surjective, $f(A)=D$ and $g(D)=E$. Hence $h(A)=E$, which shows that h is also surjective.

If g and f are both injective, then, for any $x_1, x_2\in A$, $x_1\neq x_2$ implies $f(x_1)\neq f(x_2)$; since g is also injective, this in turn implies $g(f(x_1))\neq g(f(x_2))$, that is, $h(x_1)\neq h(x_2)$, which shows that h is also injective. Thus, if f and g are bijections, then h is a bijection from A to E.

Definition 4.12: *If A and D are two nonempty sets and there exists a* *bijective mapping f from A to D, then it is said that A and D can be put into* **one-to-one correspondence** *or A and D are said to be* **similar** *or to have the same* **cardinal number.** *This is denoted by $A \sim D$.*

We recall that the identity map of A is a bijection from A to A. Also if f is a bijection from A to D, then f^{-1} is a bijection from D to A. If f is a bijection from A to D and g is a bijection from D to E, then $h = g \circ f$ is a bijection from A to E. These remarks show that one-to-one correspondence is an equivalence relation. Apropos of this discussion we make the following definition.

Definition 4.13: *A nonempty set A is said to be* **finite** *if, for some integer n, $A \sim \{1, 2, \ldots, n\}$; otherwise, A is called* **infinite.** *The empty set is also considered finite.*

It is clear (see remarks preceding Definition 4.6) that any two sets with a finite number of elements can be put in a one-to-one correspondence if and only if they have the same number of elements. We cannot, however, attach any meaning to the statement that two infinite sets have the same number of elements but the concept of one-to-one correspondence has meaning for infinite sets as well as finite sets. This is therefore taken as the appropriate generalization of the concept of number of elements in a set.

Definition 4.14: *Let J be the set of all positive integers. A nonempty set A is said to be* **countable** *if A is either finite or $A \sim J$. If a set is not countable, it is said to be* **uncountable.** *Countable sets are alternatively called* **enumerable** *or* **denumerable.**

For example, the set Z consisting of all the integers is countable. To see this let $f: J \to Z$ be the mapping defined for each $n \in J$ by

$$f(n) = \begin{cases} \dfrac{n}{2} & n \text{ even} \\[2ex] -\dfrac{(n-1)}{2} & n \text{ odd} \end{cases}$$

We shall show that f is a bijection. It is clear that f is one-to-one. To see that it is surjective (or onto), let i be any integer. If i is positive, we can always find a positive even integer n such that $n = 2i$ and, for this n, $f(n) = i$. If i is negative or zero, we can always find a positive odd integer n such that $n=1-2i$ and, for this $n, f(n)=i$.

Clearly no finite set can possibly be put into one-to-one correspondence with one of its proper subsets. We have just shown that this is possible, at least for some infinite sets. As it turns out, every infinite set can be put into one-to-one correspondence with at least one of its proper subsets.

Definition 4.15: *A function f from the set J of positive integers into a set X is called a* **sequence** *in X or, more simply, a* **sequence.** *The range of f is called the range of the sequence. The values of f are called the* **terms** *of the sequence.*

The values of any function defined on the set J have a certain order imparted to them as a consequence of the ordering of the set J by the symbols $<$, $>$, and $=$. A function f defined on J is called a sequence only when the emphasis is to be on the values of f and this ordering. Thus, conceptually, a sequence is usually thought of as a set whose members are listed in some definite order. In line with this idea, it is customary to denote the terms of the sequence by x_n (of course, any letter may be used in place of x) instead of by $f(n)$ for each $n \in J$. Since the emphasis is to be on the values of the function f, it is also customary to denote the sequence f by $\{x_n\}$, which is an abbreviation for the range of f, $\{x_n | n \in J\}$. Sometimes the even more intuitive notation $x_1, x_2, \ldots$ is used to denote the sequence. This latter notation emphasizes the ordering of the terms of the sequence. Sometimes the term sequence is applied to a function whose domain is a finite set of consecutive integers. This is *not* done here! However, occasionally a sequence will be defined as a function on the set of nonnegative integers instead of on J. Thus, we only call a function a sequence if its domain is J or the set of nonnegative integers. Since the function f need not be injective, it is clear that the *range of a sequence may be a finite set,* or it may even consist of a single point.

The proof of the next theorem is based on a principle called *mathematical induction.* This principle can be stated as follows:

Suppose that with each $n \in J$ there is associated some proposition s_n. Suppose further that s_1 is true and that for any positive integer k, s_{k+1} is true whenever s_k is. Then we can conclude that s_n is true for every $n \in J$.

We can justify this principle by using the simple fact (introduced in chapter 2) that every nonempty subset of the positive integers contains a smallest member. We shall not, however, stop to do so here.

As a simple illustration of how a proof by induction can be carried out, let us verify that the formula

$$2+4+...+2n=n(n+1) \tag{4-4}$$

is true for every $n \in J$. (The term on the left denotes the formal sum of n terms.) To this end, for each $n \in J$, let s_n be the proposition that formula (4–4) is correct. Clearly s_1 is true since $2=1(1+1)$. Now suppose s_k is true. Then adding $2(k+1)$ to both sides of (4–4) with $n=k$ we see

$$2+4+...+2k+2(k+1)=k(k+1)+2(k+1)=(k+1)[(k+1)+1]$$

and so we see that for every positive integer k, s_{k+1} is true whenever s_k is. The principle of mathematical induction now tells us that formula (4–4) is true for every $n \in J$.

The principle of mathematical induction can also be used to define functions on J. This procedure is known as a *recursive definition* of the function, and it tells us that we can define a function f by giving $f(1)$ together with a procedure for calculating $f(k+1)$ from $f(k)$ for every $k \in J$. To see how this type of definition is justified by the principle of mathematical induction, we need only let, for each $n \in J$, s_n be the proposition that $f(n)$ is defined by this procedure. Then clearly s_1 is true and s_{k+1} is true whenever s_k is. Hence, the principle of mathematical induction tells us that s_n is true for every $n \in J$ or that f is defined on J.

Theorem 4.16: *If I is an infinite subset of J, then $J \sim I$.*

Proof: In order to prove the theorem we shall construct a bijection f from J to I. To this end we define for each nonnegative integer k the set J_k by

$$J_k=\{j \in J | j \leq k\}$$

Clearly J_k is finite. Therefore $J_k \cap I$ is also finite. Now since I is infinite, for any k, the set $I-J_k$ cannot possibly be empty. It follows from this (since J_k contains all the positive integers which are less than or equal to k and I is a subset of the positive integers) that I contains an integer greater than k for each nonnegative integer k. With this in mind we can define the function f recursively as follows: Let $f(1)$ be the smallest member of I. If $f(n)$ is defined for any $n \in J$, there exists an integer in I greater than $f(n)$ and so $f(n+1)$ can be defined as the smallest integer in I which is greater than $f(n)$. In this way $f(n)$

can be defined recursively for all $n \in J$. Hence, by proceeding in this manner, we define f on J.

By construction $f(n+1)>f(n)$ for every positive integer n. If $f(n+p)>f(n)$ for *any* positive integer p, then $f(n+p+1)>f(n+p)>f(n)$. Hence it follows by induction that $f(n+p)>f(n)$ for *every* positive integer p. Now if $n_1, n_2 \in J$ and $n_1 \neq n_2$, then we can assume that the notation has been chosen so that $n_1 > n_2$. Thus, if we set $m=n_1-n_2$, then $m \in J$ and $n_1=m+n_2$. Therefore

$$f(n_1)=f(m+n_2)>f(n_2)$$

This shows that $f(n_1) \neq f(n_2)$ whenever $n_1 \neq n_2$; that is, f is injective.

It remains only to show that f is surjective. To do this we shall prove that $f(J)=I$. In order to obtain a contradiction suppose that $f(J) \neq I$. Hence the set $I-f(J)$ of positive integers is not empty and so it has a smallest element, say q.

Clearly q cannot be the smallest element of I because if it were, it would imply that $q=f(1) \in f(J)$. Thus I contains at least one integer less than q and therefore $I \cap J_{q-1} \neq \emptyset$. Hence $I \cap J_{q-1}$ (the set of all integers in I which are less than q) is a finite nonempty set of real numbers and so it contains a largest member, say r. Clearly $r \leqslant q-1 < q$ and q is the smallest member of I which is larger than r. (This follows from the fact that r is the largest integer in I which is less than q and so there can be no integers in I lying between r and q.) Since q is the smallest member of I which is not in $f(J)$, it follows that $r \in f(J)$. This shows that there exists a positive integer s such that $r=f(s)$. It now follows from the definition of f that $f(s+1)=q$. Thus $q \in f(J)$ and $q \in I-f(J)$. Since this is impossible we conclude that the assumption $f(J) \neq I$ is incorrect. Hence f is surjective.

Corollary 1: *Every subset E of a countable set A is countable.*

Proof: The proof is trivial if E is finite. Hence assume that E is infinite. Thus A is infinite and countable and so there exists a bijection $f: A \to J$. Let g be the restriction of f to E. Clearly g is injective since f is. Thus g is a bijection from E to $g(E)$. In other words $E \sim g(E)$. Since E is infinite, it is clear that $g(E)$ is also. Theorem 4.16 now shows that $J \sim g(E)$. By using the symmetric and transitive properties of $\sim$, we see that $E \sim J$. Hence E is countable.

Corollary 2: *A nonempty set A is countable if and only if there is an injection from A to J.*

Proof: If A is countable, then, by definition, there is a bijection from A to a subset of J. This mapping is obviously an injection from A to J.

Conversely, if there is an injection f from A to J, then f is a bijection from A to $f(A)$. Hence, $A \sim f(A)$ and $f(A) \subset J$. Since obviously J is countable, Theorem 4.16 shows that $f(A)$ is countable. Hence, $f(A)$ is either finite, in which case $A \sim f(A)$ shows that A is finite, or $f(A) \sim J$, in which case since $\sim$ is an equivalence relation, $A \sim f(A)$ shows that $A \sim J$.

Theorem 4.17: *A nonempty set A is countable if and only if there is a surjection from J to A.*

Proof: Suppose f is a surjection from J to A. For each $x \in A$, the set $f^{-1}(\{x\})$ is not empty. Let $g(x)$ be the smallest integer in $f^{-1}(\{x\})$. Then g is an injection from A to J and corollary 2 of Theorem 4.16 shows that A is countable.

Conversely, assume A is countable. The second corollary to Theorem 4.16 shows that there is an injection f from A to J. Choose any element $a \in A$. Since f^{-1} exists and is a mapping from the subset $f(A)$ of J onto A, define a function $g : J \rightarrow A$ as follows:

$$\begin{aligned} g(n) &= f^{-1}(n) && \text{for all } n \in f(A) \\ g(n) &= a && \text{for all } n \notin f(A) \end{aligned}$$

Then g is a surjection from J to A.

Since every function f from a set A to $f(A)$ is surjective, we conclude from this theorem that the following corollary holds.

Corollary: *A set is countable if and only if it is the range of a sequence.*

Loosely speaking, this corollary states that a set is countable if and only if it can be "arranged in a sequence."

CHAPTER 5

Infinite Collections of Sets

In chapter 4, we have given a certain meaning to the "number of elements in a set" or, to be more precise, the concept of number of elements has been replaced with a more suitable concept which has a precise meaning for sets with more than a finite number of elements. Roughly speaking, the "infinite" sets are those which have at least as many members as there are positive integers. It is with these infinite sets that analysis is principally concerned.

Definitions were given in chapter 1 for the union and intersection of two sets. As a result of these definitions, the union and intersection of any finite collection of sets are defined. In order to deal with infinite collections of sets, it is necessary to extend the concepts of union and intersection to these collections. Then after briefly discussing the relations that hold between various combinations of unions and intersections, we turn to a discussion of the "number of elements" in a set.

Since this chapter is mainly concerned with setting the background for other topics, some of the proofs, although they are quite plausible, are only formal, but they can be made into proper proofs by changing some detail.

First we shall develop a certain formalism, which is frequently used when dealing with infinite collections of sets, to designate these collections. We shall always assume that the sets with which we are dealing are subsets of some universal set X. In order to individuate the various subsets of X we suppose there is some set A, which is called the *index set*, and that, with each element α of A, there is associated a subset of X, say E_α (this defines a function from A to the collection of all subsets of X). The set Ω whose members are these sets E_α is

$$\Omega = \{E_\alpha \mid \alpha \in A\} \tag{5-1}$$

Instead of using the terminology "a set of sets," it is common practice to refer to such a set as a *collection of sets* or a *family of sets*. This helps to make clear the various levels of set construction.

We shall always assume in dealing with infinite collections of sets that a process such as that just described has at least implicitly been carried out. The notation given by equation (5–1) will be used to designate the collection of sets with which we wish to deal. Thus, the terminology "let $\Omega = \{G_\alpha \mid \alpha \in A\}$ be a collection of sets" indicates that the various sets G_α in the collection are individuated in the manner described previously.

Definition 5.1: *Let $\Omega = \{E_\alpha \mid \alpha \in A\}$ be a collection of sets. The set*

$$\{x \mid x \in E_\alpha \quad (\text{for at least one } \alpha \in A)\}$$

is called the **union** *of the collection of sets Ω and is denoted by $\bigcup_{\alpha \in A} E_\alpha$.*

The set

$$\{x | x \in E_\alpha \quad (\text{for every } \alpha \in A)\}$$

is called the **intersection** *of the collection of sets Ω and is denoted by $\bigcap_{\alpha \in A} E_\alpha$.*

Thus, $\bigcup_{\alpha \in A} E_\alpha$ is the set S which has the property that $x \in S$ if and only if $x \in E_\alpha$ for at least one $\alpha \in A$ and $\bigcap_{\alpha \in A} E_\alpha$ is the set I which has the property that $x \in I$ if and only if $x \in E_\alpha$ for every $\alpha \in A$.[12] We see from this that, for every $\alpha \in A$, $E_\alpha \subset S$ and $I \subset E_\alpha$. On the other hand, if V and W are any two sets such that, for every $\alpha \in A$, $E_\alpha \subset V$ and $W \subset E_\alpha$, then $S \subset V$ and $W \subset I$. Thus S is the smallest set which contains every set E_α, and I is the largest set contained in every E_α.

It should be noticed that if the index set A consists of two elements, say α_1 and α_2, Definition 5.1 reduces to

$$\bigcup_{\alpha \in A} E_\alpha = \{x | x \in E_{\alpha_1} \quad \text{or} \quad x \in E_{\alpha_2}\}$$

$$\bigcap_{\alpha \in A} E_\alpha = \{x | x \in E_{\alpha_1} \quad \text{and} \quad x \in E_{\alpha_2}\}$$

and hence reduces to the definitions of union and intersection given in chapter 1.

[12] Notice that we did not require the index set A to be nonempty. We find in fact that $\bigcup_{\alpha \in \emptyset} E_\alpha = \emptyset$ and $\bigcap_{\alpha \in \emptyset} E_\alpha = X$ where X is the universal set. The latter relation often causes some difficulty. What it amounts to is this: an element x of the universal set does not belong to I only if it does not belong to some E_α. Since in this case there are no E_α's there is no element of the universal set which does not belong to I. Hence, $X \subset I$.

If the index set A is countable, the family Ω is called a *countable collection.* The corollary to Theorem 4.17 shows that in this case A is the range of a function g from the positive integers. Hence each $\alpha \in A$ is equal to $g(n)$ for some $n \in J$. It is clear from this that

$$\{E_\alpha | \alpha \in A\} = \{E_{g(n)} | n \in J\}$$

Hence, defining, for each $n \in J$, $F_n = E_{g(n)}$, we see that when the index set is countable it can always be replaced by a subset of J. We will frequently assume that this has been done when dealing with countable collections of sets.

It might be pointed out that the collection of sets $\{F_n | n \in J\}$ obtained in this way is then the range of a sequence $\{F_n\}$ whose terms are sets.

Conversely, any sequence whose terms are sets has a *countable* collection of sets $\{F_n | n \in J\}$ for its range.

If the index set A is the set J of positive integers, it is usual to denote the union by $\bigcup_{n=1}^{\infty} E_n$ instead of by $\bigcup_{n \in J} E_n$ and to denote the intersection by $\bigcap_{n=1}^{\infty} E_n$ instead of by $\bigcap_{n \in J} E_n$. The symbol ∞ serves here only to indicate that the union or intersection of a denumerably infinite collection of sets is taken and has no connections with the symbols $+\infty$ and $-\infty$ introduced in chapter 2. If the index set A consists of the integers $1, 2, \ldots, n$, it is usual to denote the union by $\bigcup_{m=1}^{n} E_m$ or by $E_1 \cup \ldots \cup E_n$ and to denote the intersection by $\bigcap_{m=1}^{n} E_m$ or by $E_1 \cap \ldots \cap E_n$.

To illustrate these ideas, suppose that, for each positive integer n, E_n is the subset of the positive real numbers defined by

$$E_n = \left\{x \mid 0 < x < \frac{1}{n}\right\}$$

Let $\Omega = \{E_n \mid n \in J\}$. As usual J is the set of positive integers. The intersection of Ω satisfies $\bigcap_{n=1}^{\infty} E_n = \emptyset$. To see this, note that for every real number $x > 0$ we can find a positive integer m such that $1/m < x$ and so, for this m, $x \notin E_m$. Hence, $x \notin \bigcap_{n=1}^{\infty} E_n$. Thus, there is no positive real number in $\bigcap_{n=1}^{\infty} E_n$.

Now suppose $A = \{x | 0 < x < 1\}$, $E_x = \{y | 0 \leq y \leq x\}$, and $\Omega = \{E_x | x \in A\}$. Then

$$\bigcup_{x \in A} E_x = \{x \mid 0 \leq x < 1\} \tag{5–2}$$

It is clear that for every $x \in A$

$$E_x \subset \{x \mid 0 \leq x < 1\}$$

Hence

$$\bigcup_{x \in A} E_x \subset \{x \mid 0 \leq x < 1\} \qquad (5\text{–}3)$$

On the other hand, if $t \in \{x \mid 0 \leq x < 1\}$, then $t < 1$ and so $t \in E_t \subset \bigcup_{x \in A} E_x$. Now since t was an arbitrary member of $\{x \mid 0 \leq x < 1\}$, this shows that $\{x \mid 0 \leq x < 1\} \subset \bigcup_{x \in A} E_x$. Combining this equation with (5–3) shows that equation (5–2) holds.

The unions and intersections defined in this chapter satisfy many algebraic identities that are analogous to those given in chapter 1 for the binary set operations. Some of these are listed in table 5–I.

Table 5–I. — Set Theoretic Identities

Associative and commutative laws	
$F \cup \left(\bigcup_{\alpha \in A} E_\alpha\right) = \bigcup_{\alpha \in A} (F \cup E_\alpha)$	$F \cap \left(\bigcap_{\alpha \in A} E_\alpha\right) = \bigcap_{\alpha \in A} (F \cap E_\alpha)$
$\bigcup_{\alpha \in A} \bigcup_{\delta \in D} E_{\alpha,\delta} = \bigcup_{\delta \in D} \bigcup_{\alpha \in A} E_{\alpha,\delta} = \bigcup_{(\alpha,\delta) \in A \times D} E_{\alpha,\delta}$	$\bigcap_{\alpha \in A} \bigcap_{\delta \in D} E_{\alpha,\delta} = \bigcap_{\delta \in D} \bigcap_{\alpha \in A} E_{\alpha,\delta} = \bigcap_{(\alpha,\delta) \in A \times D} E_{\alpha,\delta}$
Distributive laws	
$F \cup \left(\bigcap_{\alpha \in A} E_\alpha\right) = \bigcap_{\alpha \in A} (F \cup E_\alpha)$	$F \cap \left(\bigcup_{\alpha \in A} E_\alpha\right) = \bigcup_{\alpha \in A} (F \cap E_\alpha)$
DeMorgan's laws	
$\left(\bigcup_{\alpha \in A} E_\alpha\right)^c = \bigcap_{\alpha \in A} E_\alpha^c$	$\left(\bigcap_{\alpha \in A} E_\alpha\right)^c = \bigcup_{\alpha \in A} E_\alpha^c$

It is not hard to verify these identities. To illustrate the procedure we will verify the first of DeMorgan's laws and the second distributive law.

In order to show that

$$\left(\bigcup_{\alpha \in A} E_\alpha\right)^c = \bigcap_{\alpha \in A} E_\alpha^c$$

let L and R be the left and right sides, respectively, of this equality. If $x \in L$, then $x \notin \bigcup_{\alpha \in A} E_\alpha$. Hence $x \notin E_\alpha$ for any $\alpha \in A$, and so $x \in E_\alpha^c$ for every $\alpha \in A$. Therefore $x \in \bigcap_{\alpha \in A} E_\alpha^c$. Thus, $L \subset R$.

Conversely, if $x \in R$, then $x \in E_\alpha^c$ for every $\alpha \in A$. Hence $x \notin E_\alpha$ for any $\alpha \in A$, and therefore $x \notin \bigcup_{\alpha \in A} E_\alpha$. Thus $x \in \left(\bigcap_{\alpha \in A} E_\alpha\right)^c$. This shows $R \subset L$, and it follows that $R = L$.

To show that

$$F \cap \left(\bigcup_{\alpha \in A} E_\alpha\right) = \bigcup_{\alpha \in A} \left(F \cap E_\alpha\right)$$

let L and R be the left and right sides, respectively, of this equality. If $x \in L$, then $x \in F$ and $x \in \bigcup_{\alpha \in E} E_\alpha$. Hence $x \in F$, and there exists a $\beta \in A$ such that $x \in E_\beta$. Therefore, $x \in F \cap E_\beta$ for some $\beta \in A$, and so $x \in \bigcup_{\alpha \in E} F \cap E_\alpha$. This shows that $L \subset R$. If $x \in R$, then, for some $\beta \in A$, $x \in F \cap E_\beta$. Hence, $x \in F$ and $x \in E_\beta$ for some $\beta \in A$. Thus $x \in F$ and $x \in \bigcup_{\alpha \in A} E_\alpha$ and so $x \in F \cap \left(\bigcup_{\alpha \in A} E_\alpha\right)$.

There are also many relations connecting images of sets and unions and intersections of arbitrary collections of sets which generalize those given in chapter 4 for binary set operations. We list some of these in table 5–II.

Table 5–II. — Laws Connecting Images with Unions and Intersections

[$f: X \to Y$; $E_\alpha \subset X$ for every $\alpha \in A$; $A_\delta \subset Y$ for every $\delta \in D$.]

$f\left(\bigcup_{\alpha \in A} E_\alpha\right) = \bigcup_{\alpha \in A} f(E_\alpha)$	$f\left(\bigcap_{\alpha \in A} E_\alpha\right) \subset \bigcap_{\alpha \in A} f(E_\alpha)$
$f^{-1}\left(\bigcup_{\delta \in D} A_\delta\right) = \bigcup_{\delta \in D} f^{-1}(A_\delta)$	$f^{-1}\left(\bigcap_{\delta \in D} A_\delta\right) = \bigcap_{\delta \in D} f^{-1}(A_\delta)$

These laws are also easy to verify. For purposes of illustration of the procedure, we will verify one of these.

Let us show that

$$f\left(\bigcup_{\alpha \in A} E_\alpha\right) = \bigcup_{\alpha \in A} f(E_\alpha)$$

Set L equal to the left side of this identity and R equal to the right side. That is,

$$L = \left\{ y \in Y \,\middle|\, \left(\exists x \in \bigcup_{\alpha \in A} E_\alpha\right) \text{ for which } y = f(x) \right\}$$

and $R = \bigcup_{\alpha \in A} f(E_\alpha)$. For each $\alpha \in A$,

$$f(E_\alpha) = \{y \in Y \mid (\exists x \in E_\alpha) \text{ for which } y = f(x)\}$$

Suppose $y \in L$. Then, for some $\beta \in A$, there exists an $x \in E_\beta$ such that $y = f(x)$. Hence, for this β, $y \in f(E_\beta)$, and this shows that $y \in \bigcup_{\alpha \in A} f(E_\alpha) = R$. Since y was any point of L, we have shown that $L \subset R$. Conversely, suppose that $y \in R$. Then, for some $\beta \in A$, $y \in f(E_\beta)$. This means that there exists an $x \in E_\beta$ such that $y = f(x)$. Therefore, there certainly exists an $x \in \bigcup_{\alpha \in A} E_\alpha$ such that $y = f(x)$. This shows that $y \in L$. Since y was any point of R, we see that $R \subset L$. Combining this with the opposite inclusion shows that $L = R$.

In the next theorem we adopt the convention of assuming that matters have already been arranged so that the index set of the countable collection of sets is J.

Theorem 5.2: *If $\Omega = \{E_n \mid n \in J\}$ is a countable collection of countable sets, then the set $S = \bigcup_{n-1}^{\infty} E_n$ is countable.*

Proof: According to the corollary to Theorem 4.17, for each positive integer n, a sequence $\{x_{n,k}\}$ whose range is E_n can be chosen. Having chosen such sequences, we form the infinite array

$$\begin{array}{ccccc}
x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} & \cdot \\
x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} & \\
x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} & \\
x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} & \\
\cdot & \cdot & \cdot & \cdot & \\
 & & & & \\
\cdot & \cdot & \cdot & \cdot &
\end{array}$$

in which the members of E_n make up the nth row. Evidently every member of S is in this array. If the pattern indicated by the arrows is followed, a correspondence can be set up between the members of this array and the positive integers as follows:

$$\begin{array}{cccccccccccc}
x_{1,1}; & x_{2,1}, & x_{1,2}; & x_{3,1}, & x_{2,2}, & x_{1,3}; & x_{4,1}, & x_{3,2}, & x_{2,3}, & x_{1,4}; & \ldots \\
1, & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9, & 10 & \ldots
\end{array}$$

In this manner then, we can find, for any element in the array and hence for any member of S, a positive integer to which it corresponds. This procedure therefore defines a surjective mapping from J to S. Theorem 4.17 now shows that S is countable.

This proof is only formal since we have not actually constructed the surjective function but only indicated how it is to be constructed. However, the mapping $f: J \times J \to J$ defined by

$$f(\langle n, m \rangle) = \frac{m(m+1)+(n-1)(n+2m-2)}{2} \qquad n, m = 1, 2, 3, \ldots$$

can be shown to be a bijection by the theorems of factorization of integers. Hence this function can be used to construct the desired surjection. However, this is all mere detail with which we will not concern ourselves since the basic ideas are indicated in the formal proof.

355-525 O-70—5

Theorem 5.3: *Let A be a countable set; let $D_1 = A$; and, for some integer $n \geqslant 2$,* let $D_n = A \times \ldots \times A$ *(taken n times). Then, for every positive integer n, D_n is countable.*
Proof: D_1 is obviously countable. Suppose that, for any integer $n \geqslant 2$, D_{n-1} is countable. If $x \in D_n$ then $x = \langle d, a \rangle$ with $d \in D_{n-1}$ and $a \in A$. For each fixed $d \in D_{n-1}$, the set $E_d = \{\langle d, a \rangle | a \in A\}$ can be put into one-to-one correspondence with A and is therefore countable. Since $D_n = \bigcup_{d \in D_{n-1}} E_d$, D_n is evidently the union of a countable collection of countable sets and therefore Theorem 5.2 shows that D_n is countable. Hence the conclusion follows by induction.

Corollary: *The set of all rationals is countable.*
Proof: If we apply Theorem 5.3 with $n = 2$ and A the set of all integers, then it follows that the set D_2 of all ordered pairs $\langle a, b \rangle$ where a and b are integers is certainly countable. Now every rational number can be written in the form b/a where a and b are integers. Hence it is clear that the set of all rationals can be put into one-to-one correspondence with a subset E of D_2 which corollary 1 of Theorem 4.16 shows is countable. Hence, $E \sim J$ or else it is finite. Since one-to-one correspondence is an equivalence relation, we conclude that the set of all rational numbers is either similar to J or to a finite set. In either case, the set of all rationals is countable.

The final theorem, called Cantor's diagonalization theorem, shows that there are many uncountable sets of real numbers. Again the proof which we give is only formal but can easily be converted into a proper proof.

Theorem 5.4: *The set A of all real numbers lying between zero and one is uncountable.*
Proof: Every real number lying between zero and one can be written in decimal form as $0.S_1S_2S_3 \ldots$ where the S_j are integers. Let E be any countable subset of A. We can write an arbitrary element of E, S^n, in decimal form as $S^n = 0.S_1^nS_2^nS_3^n \ldots$ where the S_j^n are integers lying between 0 and 9. Now, since E is countable, we can arrange its members in a sequence as follows:

$$
\begin{array}{llll}
S^1 = 0.S_1^1 S_2^1 S_3^1 S_4^1 & \cdot & \cdot & \cdot \\
S^2 = 0.S_1^2 S_2^2 S_3^2 S_4^2 & \cdot & \cdot & \cdot \\
S^3 = 0.S_1^3 S_2^3 S_3^3 S_4^3 & \cdot & \cdot & \cdot \\
S^4 = 0.S_1^4 S_2^4 S_3^4 S_4^4 & \cdot & \cdot & \cdot \\
S^5 = \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & & & \\
\cdot \quad \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & & & \\
\cdot \quad \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & & &
\end{array}
$$

Consider the elements S_j^j along the diagonal of this array. We can choose a number p in A as follows. Let $p = 0.p_1p_2p_3 \ldots$ where $p_1 \neq S_1^1$, $p_2 \neq S_2^2$, . . ., $p_j \neq S_j^j$, ... We can further choose the integers p_j so that they lie between zero and nine and are not all zeros or nines. Then the number p differs from each member of E in at least one decimal place. Hence $p \notin E$ and $p \in A$ so that E is a proper subset of A.

We have thus shown that every countable subset of A is a proper subset of A. It follows that A is uncountable for, otherwise, A would be a *proper* subset of A which is absurd.

CHAPTER 6

Metric Spaces

Modern analysis began when Cantor developed his theory of point sets (which included the important concepts of limit point, derived set, closed set, etc.) from a study of the real valued functions on the real line and the distance properties of the real line. Using Cantor's ideas, Fréchet developed the concept of *metric space* (and, for that matter, abstract spaces in general) when, in 1906, he gave an abstract generalization of continuous functions on point sets. Fréchet's theory was phenomenally successful because nearly all the continuity and convergence arguments that occur in analysis require only the few facts about the concept of distance between points which were embodied in this theory. The actual term "metric space" was first used by Hausdorff in 1914. In fact, he appears to be the first to use the geometrically suggestive word "space" to refer to a set of objects of unspecified nature which are subject to certain postulates.

In the theory of spaces it turns out to be very helpful as well as convenient to use a terminology inspired by classical geometry. Thus the elements of a space are referred to as *points*. A metric space, then, is merely a set of objects, called points, between which a measure of distance is defined in such a way as to single out those properties of the distance between real numbers (contained in the order axioms of chapter 2) which are important for convergence and continuity arguments. Because continuity and convergence are essentially the central concepts in mathematical analysis, this chapter is devoted to the study of the fundamental concepts of metric spaces. In the process of discussing these spaces, we shall develop the geometric language which is currently used to discuss mathematical analysis. The reason for introducing some of the concepts in this chapter will become apparent subsequently.

First, we define metric space. When we think of the distance between two points in a plane or two points on a line, we think of a number associated with these two points—say the number of inches read from a ruler placed between the points. Now if we wish to assign a unique number to each distinct pair of elements of an arbitrary set X, we can accomplish this by defining a function on $X \times X$ with values in the real number system (see remarks following Definition 4.1). With this in mind, we make the following definition.

Definition 6.1: *A metric space $\langle X, d\rangle$ is a set X, whose members are called* **points,** *together with a function $d: X \times X \to R^1$ which, for all p, q, $t \in X$, has the following properties*

(a) $\quad d(p, q) > 0 \quad$ if $p \neq q$; $d(p, p) = 0$

(b) $\quad d(p, q) = d(q, p)$

(c) $\quad d(p, q) \leqslant d(p, t) + d(t, q)$

The function d is called the **metric** *or* **distance function** *or simply the* **distance.** *The value of d at $\langle p, q\rangle$ is called the distance between the points p and q.*

Postulate (a) expresses the fact that the distance between two points is always a positive number and equal to zero if and only if *the two points coincide;* postulate (b) expresses the fact that the distance between two points is the same measured in either direction; and postulate (c) expresses the fact that the distance between two points is not decreased if it is measured via a third point. In fact, postulate (c) is a reflection of the fact that the sum of the lengths of two sides of a triangle is greater than or equal to the length of the third side, and so it is commonly known as the triangle inequality.

As in the case of linear spaces discussed in chapter 3 (see remarks following Definition 3.1), it is common practice to refer to a metric space by the name of the underlying set. However, when it is convenient to have the symbol for the metric given explicitly, we shall employ the more correct procedure of referring to the metric space as the ordered pair $\langle X, d\rangle$ where X is the underlying set and d is the distance function which, in theory, contains the information as to which numbers are to be assigned to each pair of elements in X.

Among the most important examples of metric spaces are the normed linear spaces defined in chapter 3 and, in particular, the Euclidean spaces R^k. If v_1 and v_2 are any two vectors, it is customary to write $v_1 - v_2$ in place of $v_1 + (-v_2)$. To see that the normed linear space V is in fact a metric space, it is only necessary to define the function $d: V \times V \to R^1$ in terms of the norm by

$$d(v_1, v_2) = \|v_1 - v_2\| \qquad \text{for all } v_1, v_2 \in V \tag{6-1}$$

Then, comparing postulates (N1) to (N3) of Definition 3.4 with postulates (a) to (c) of Definition 6.1, we see that d is in fact a distance on the set V. *We shall* *always define the distance between points in a normed linear space in this manner.* In fact, if V is a normed linear space, we shall use the terminology "the metric space V" to refer to the metric space which is obtained by using equation (6–1) to define a metric on the set V.

Specializing equation (6–1) to the Euclidean space R^k we see that the distance is defined by[13]

$$d(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}| \qquad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in R^k \tag{6–2}$$

Therefore, in the Euclidean spaces R^2 and R^3 this is just the magnitude of the vector joining the points $\boldsymbol{x}$ and $\boldsymbol{y}$, as it should be. Also, if the points in the plane are represented by complex numbers, and if the correspondence between the norm of the vectors in the Euclidean space R^2 and the absolute value of the corresponding complex numbers pointed out in the discussion following Theorem 3.6 is used, then, according to this definition, the distance between two complex numbers is just the absolute value of the difference of the two complex numbers. *We shall always define the distance between points in the complex plane in this manner.*

As we pointed out in chapter 4, the set R^1 of real numbers is a normed linear space with the usual arithmetic if we define the norm by the absolute value. We shall sometimes refer to this normed linear space as the Euclidean space R^1. The metric on this normed linear space is then defined by equation (6–1). The distance between any two real numbers is just the absolute value of the difference of the numbers. The metric defined on the real numbers in this manner is called the **usual metric** for R^1.

There are metric spaces more abstract than the Euclidean spaces but equally important. For example, let X be any set and let $\mathscr{B}(X)$ be the set of all real valued functions defined on X such that, for any function $f \in \mathscr{B}(X)$, $\text{lub}\,\{|f(x)| \mid x \in X\} < \infty$. Then $\mathscr{B}(X)$ is a metric space if we define a metric d on it by

$$d(f_1, f_2) = \text{lub}\,\{|f_1(x) - f_2(x)| \mid x \in X\} \qquad \text{for every } f_1, f_2 \in \mathscr{B}(X) \tag{6–3}$$

We shall not stop here to verify that equation (6–3) is indeed a metric since this is a special case of the metric spaces constructed in chapter 11. The set $\mathscr{B}(X)$ is a particular example of a large class of metric spaces known

[13] Notice that according to the convention adopted in mathematics the set R^k is merely a set of ordered k-tuples of real numbers, but the *Euclidean space* R^k is the set R^k *together* with the algebraic operations and the norm defined in Definition 3.5. Thus the *Euclidean space* R^k *is a metric space with the metric defined by equation* (6–2).

as function spaces. Assigning metrics to sets of functions gives these sets a certain geometric nature which is a great help to our intuition about them. In fact much of the success in the theory of functions can be attributed in no small measure to the insight gained through this geometric point of view.

If $\langle X, d\rangle$ is a metric space and x, y, and z are any elements of X, then it follows from postulate (c) of Definition 6.1 that

$$d(x, z) \leqslant d(x, y) + d(y, z)$$

and

$$d(y, z) \leqslant d(y, x) + d(x, z)$$

We can write these relations as

$$d(x, z) - d(y, z) \leqslant d(x, y)$$

and

$$-d(y, x) \leqslant d(x, z) - d(y, z)$$

But postulate (b) shows that $d(x, y) = d(y, x)$. Hence

$$-d(x, y) \leqslant d(x, z) - d(y, z) \leqslant d(x, y)$$

or

$$|d(x, z) - d(y, z)| \leqslant d(x, y) \tag{6-4}$$

Let X be an arbitrary set. We can define a function $d: X \times X \to R^1$ as follows:

$$d(x, y) = \begin{cases} 1 \text{ if } x \neq y \\ 0 \text{ if } x = y \end{cases} \qquad \text{for every } x, y \in X$$

It is clear that conditions (a) and (b) of Definition 6.1 are satisfied. If any two of the three elements x, y, z of X are equal, it is easy to see that condition (c) is satisfied; if this is not the case, then $d(x, z) = 1$ and $d(x, y) + d(y, z) = 2$ and so condition (c) is satisfied in all cases. Hence, $\langle X, d\rangle$ is a metric space. It is said to be a *discrete* metric space. Discrete metric spaces are very useful as counterexamples.

In certain situations it is useful to allow a metric to take on the value $+\infty$ with the arithmetic given in Definition 2.3 (i.e., d is taken to be a function with values in the extended real number system[14]). This is never done, however, unless it is stated explicitly. Although most of the following theorems about metric spaces would go through with little change if we allowed an infinite metric, we shall restrict ourselves to finite metrics, and only in chapter 11 will we consider metric spaces with possibly infinite metrics.

If $\langle X, d\rangle$ is a metric space and Y is *any* subset of X, let d_Y be the restriction of d to $Y \times Y$. Thus

$$d_Y(p, q) = d(p, q) \qquad \text{for all } p, q \in Y$$

It is easy to see that $\langle Y, d_Y\rangle$ is a metric space. For if d satisfies conditions (a) to (c) of Definition 6.1 for all $p, q, t \in X$, then d_Y must certainly satisfy these conditions for all $p, q, t \in Y$. The metric space $\langle Y, d_Y\rangle$ is called a **metric subspace** of $\langle X, d\rangle$ or, when no confusion can occur, simply a **subspace** of $\langle X, d\rangle$. It is common practice (though logically incorrect) not to make any distinction between the metric d and its restriction d_Y. Thus, we say d and d_Y are the "same" metrics, drop the subscript Y, and write d in place of d_Y. No confusion can result from this convention since the set Y is indicated explicitly in the notation $\langle Y, d\rangle$. It is important to note that, in contrast to the situation for linear spaces (as discussed preceding Definition 3.3) every subset Y of a metric space $\langle X, d\rangle$ is a metric space in its own right with the "same" metric as X. For example, every subset of the Euclidean space R^k is a *metric* subspace of R^k but certainly not every subset of R^k is a *linear* subspace of R^k.

As it turns out there are other ways of defining a metric on the set R^k (with

---

[14] Some authors do not call this function a metric. There is, in fact, a somewhat less restrictive concept than that of a metric which is called an *écart* (French for "separation"). This is defined as follows.

Definition: *An* **écart** *on a set X is a function $e: X \times X \to R^1 \cup \{+\infty\}$ such that, for all $p, q, t \in X$,*

(a) $e(p, q) \geqslant 0$; $e(p, p) = 0$

(b) $e(p, q) = e(q, p)$

(c) $e(p, q) \leqslant e(p, t) + e(t, q)$

The only differences between an écart and a metric is that an écart can take on the value $+\infty$ and two distinct points can have écart zero. In order to verify that an écart e is a metric, it is only necessary to establish that $e(p, q)$ is finite for all $p, q \in X$ and that $e(p, q) \neq 0$ if $p \neq q$.

What we will call a possibly infinite metric, then, is an écart e with the restriction that $e(p, q) \neq 0$ if $p \neq q$. An écart e of this type can always be replaced by a finite metric. In fact, it is not hard to verify that the function $d: X \times X \to R^1$ defined by

$$d(p, q) = \frac{e(p, q)}{1 + e(p, q)} \qquad \text{for all } p, q \in X$$

is indeed a metric.

$k > 1$) than that given by equation (6–2). Before discussing this further, let us look at a somewhat more general situation. Let $\langle X, d\rangle$ and $\langle Y, \delta\rangle$ be metric spaces [15] and consider the set $X \times Y$ of ordered pairs. We may ask whether there is any way of defining a distance on $X \times Y$ in terms of the metrics d and δ. Actually this can be done in several ways. Suppose $\langle x_1, y_1\rangle$ and $\langle x_2, y_2\rangle$ are any two points of $X \times Y$. If we define the functions $d_\times$, d_1, and d_2 (from $(X \times Y) \times (X \times Y)$ into R^1) by

$$d_\times(\langle x_1, y_1\rangle, \langle x_2, y_2\rangle) = \max\{d(x_1, x_2), \delta(y_1, y_2)\} \tag{6–5}$$

$$d_1(\langle x_1, y_1\rangle, \langle x_2, y_2\rangle) = d(x_1, x_2) + \delta(y_1, y_2) \tag{6–6}$$

$$d_2(\langle x_1, y_1\rangle, \langle x_2, y_2\rangle) = ([d(x_1, x_2)]^2 + [\delta(y_1, y_2)]^2)^{1/2} \tag{6–7}$$

then it is easily verified that each of the three satisfies conditions (a) to (c) of Definition 6.1. In other words, $\langle X \times Y, d_\times\rangle$, $\langle X \times Y, d_1\rangle$, and $\langle X \times Y, d_2\rangle$ are *all* metric spaces. *The metric space $\langle X \times Y, d_\times\rangle$ is called the* **direct product** *of the metric space $\langle X, d\rangle$ and $\langle Y, \delta\rangle$. All the results which we will prove for the direct product $\langle X \times Y, d_\times\rangle$ of two* metric spaces *X and Y* are also true in the metric spaces $\langle X \times Y, d_1\rangle$ and $\langle X \times Y, d_2\rangle$. Equations (6–5), (6–6), and (6–7), which define the functions $d_\times$, d_1, and d_2, respectively, can be extended in an obvious way to the product of any finite number of metric spaces.

Now let us look at the case of R^k with $k > 1$. Since $R^k = R^s \times R^{k-s}$ with $1 \leq s \leq k-1$, we can construct a metric on the set R^k from the metrics on the Euclidean spaces R^s and R^{k-s} by any of the procedures just described. The metric defined by equation (6–7) is the same as that defined by equation (6–2) for the Euclidean space R^k but the other two are certainly different. Thus the direct product of the Euclidean spaces R^s and R^{k-s} is a different metric space than the Euclidean space R^k. However, in a certain more general sense which we will not go into here (i.e., in the sense of topological spaces), these metric spaces are essentially the same. The preceding discussion points out the fact that the same set can give rise to more than one metric space, so that the convention of referring to the metric space only by the name of the underlying set can sometimes lead to confusion.

Definition 6.2: *Let $\langle X, d\rangle$ be a metric space and suppose $p \in X$. We define for any* **positive number** *r* **the open ball of radius** *r* **about** *p (or*

---

[15] Of course $\langle X, d\rangle$ and $\langle Y, \sigma\rangle$ can be the same metric space.

more simply the **ball of radius** r **about** p) $B(p; r)$ *to be the set*

$$B(p; r) = \{q \in X \mid d(p, q) < r\} \tag{6–8}$$

We emphasize that *the radius of a ball is always a* **finite** *number and is* **never equal to zero.** An arbitrary ball $B(p; r)$ is always a *nonempty* set since it contains the point p. It is often helpful to think of the ball $B(p; r)$ as being the set of all points "close to" p—the degree of closeness being given by r.

The notation $B(p; r)$ used in Definition 6.2 for balls is more or less standard. It should be noted that this notation makes no provision for indicating the metric space to which the ball belongs. Thus, when one considers two (or more) metric spaces simultaneously, the same letter B is used to designate the balls in both spaces. When this is done, it is always made clear to which space the center of the ball (p in eq. (6–8)) belongs, and this is used to determine which space the ball is in. Occasionally, when the location of the center of the ball and its radius are immaterial to the discussion (and when no confusion is likely to result), a ball is denoted simply by the letter B with the argument omitted.

If V is a normed linear space (the metric is defined by eq. (6–1)), then the ball of radius r about the point v is the set

$$B(v; r) = \{w \in V \mid \ \|w - v\| < r\}$$

So that, in particular, balls in the Euclidean space R^k (the metric is given by eq. (6–2)) are the sets

$$B(\boldsymbol{x}; r) = \{\boldsymbol{y} \in R^k \mid \ |\boldsymbol{x} - \boldsymbol{y}| < r\} \qquad \boldsymbol{x} \in R^k$$

According to the discussion following Definition 6.1, *balls in the complex plane are the interiors of circles,*[16] as are the balls in the Euclidean space R^2. Naturally, when the complex numbers are identified with the points of R^2 in the manner described following Theorem 3.6, the balls in the complex plane and those in the Euclidean space R^2 consist of the same points.

In the Euclidean space R^3 balls are the interiors of spheres.

When equations (6–5) to (6–7) are used to define metrics on R^2 in terms of the metric on the Euclidean space R^1, the appearance of the balls are different in each case. The balls arising from the metric $d_\times$ are "interiors" of

[16] In this context balls are often referred to as *disks*.

squares,[17] those arising from d_1 are "interiors" of diamond-shaped regions, and those arising from d_2 are "interiors" of circles.

More generally, *let* $\langle X, d\rangle$ *and* $\langle Y, \delta\rangle$ *be any two metric spaces and let* $B(\langle x, y\rangle; r)$ *be a ball about any point* $\langle x, y\rangle$ *in the direct product* $\langle X\times Y, d_\times\rangle$ *of* $\langle X, d\rangle$ *and* $\langle Y, \delta\rangle$. *Then*

$$B(\langle x, y\rangle; r) = B(x; r)\times B(y; r) \tag{6–9}$$

As far as the material discussed in this book is concerned, it is not important what detailed "shapes" balls have. Roughly speaking, it is only necessary that the balls in any metric space have no dimension which is excessively large compared with its other dimensions, that they consist only of the interiors of sets, and that each point of the space is contained in a ball of arbitrarily "small" size.

If $\langle X, d\rangle$ is a discrete metric space, then for any $p\in X$

$$B(p; r) = \begin{cases} \{p\} & \text{if } r \leqslant 1 \\ X & \text{if } r > 1 \end{cases}$$

Definition 6.3: *If a and b, with $a<b$, are any real (or extended real) numbers, the* **segment** (a, b) *is defined to be the set of real numbers*

$$\{x|a<x<b\}$$

and the **interval** $[a, b]$ *is defined to be the set of real (or extended real) numbers*

$$\{x|a\leqslant x\leqslant b\}$$

We shall also sometimes encounter the **half-open intervals** $[a, b)$ *and* $(a, b]$: *the first is defined to be the set of real (or extended real) numbers*

$$\{x|a\leqslant x<b\}$$

and the second is defined to be the set of real (or extended real) numbers

$$\{x|a<x\leqslant b\}$$

Let $a_i\leqslant b_i$ for $i=1, 2, \ldots, k$ be (finite) real numbers. We shall define the

[17] A precise definition of the interior of a set will be given subsequently.

***k*-cell** *Q to be the set*

$$Q=\{\boldsymbol{x}\in R^k|a_i\leq x_i\leq b_i,\ 1\leq i\leq k\}$$

If we let $I_i=[a_i, b_i]$ for $1\leq i\leq k$, then, as shown at the end of chapter 1, we can also write this as

$$Q=\bigtimes_{i=1}^{k} I_i$$

Thus, for a and b finite, the segment (a, b) is a ball in the metric space R^1 with the usual metric. We remind the reader that points in the spaces R^k must have *finite* real numbers for their coordinates. A 1-cell is an interval, a 2-cell is a rectangle, etc.

The concept in the following definition will not be used here but is given for the sake of completeness.

Definition 6.4: *Let V be a normed linear space. A set $E\subset V$ is* **convex** *if*

$$\lambda v_1+(1-\lambda)v_2\in E$$

whenever $v_1\in E$, $v_2\in E$ and $0<\lambda<1$.

It is easily seen that every subspace of V is a convex set. Every ball in V is also a convex set. For if $\|v_1-v\|<r$ and $\|v_2-v\|<r$ and $0<\lambda<1$, then it follows from (N3) of Definition 3.4 that

$$\begin{aligned}\|\lambda v_1+(1-\lambda)v_2-v\|&=\|\lambda(v_1-v)+(1-\lambda)(v_2-v)\|\\&\leq\lambda\|v_1-v\|+(1-\lambda)\|v_2-v\|<\lambda r+(1-\lambda)r=r\end{aligned}$$

It is also readily shown that in R^k every k-cell is convex.

Definition 6.5: *Let $\langle X, d\rangle$ be a metric space, let E be any subset of X, and let $p\in X$.*

(a) *A point p is called an* **adherence point** *of the set E if every ball about p contains at least one point of E (p is also said to* **adhere to** *E).*

(b) *A point p is called a* **limit point** *of the set E if every ball about p contains at least one point of $E-\{p\}$.*

(c) *A point p is called an* **interior point** *of the set E if there exists a ball B about p such that $B\subset E$.*

Intuitively, we may think of an interior point of a set E as being a point which has *only* points of E in its immediate vicinity. In the complex plane this corresponds to a point being "inside" the set.

Notice that $\{p\}$ is the one point set that consists of the point p alone. Therefore, the set $E-\{p\}$ is the set of all points of E except possibly the point p itself, if p happens to be a point of E. Thus we may think of a limit point of E as being a point which has at least *some* points of E, other than itself, arbitrarily close to it. If p is a limit point of E, then p is an adherence point of $E-\{p\}$.

Suppose the metric space X is the set of real numbers R^1 with the usual metric, $E=\{1/n|n=1, 2, 3, \ldots\}$, and $p=0$. Then p is not a point of E. Since, in R^1, the ball about 0 of radius δ is the segment $(-\delta, \delta)$, it is clear that, for any $\delta>0$, we can choose n so large that $1/n<\delta$. Thus every ball about p contains at least one point of E. Since $p \notin E$, $E-\{p\}=\{x \in R^1|x \in E$ and $x \notin \{p\}\}=E$, and we see that p is a limit point of E.

If we let E and X be the same as above but now set $p=1$, then $p \in E$. Since the segment $(3/4, 1\frac{1}{4})$ is the ball about $p=1$ of radius $1/4$ and since the only point of E contained in this ball is p, we see that $B(p, 1/4)$ contains a point of E but no point of $E-\{p\}=\{1/n|n=2, 3, 4, \ldots\}$. Hence p is not a limit point of E. On the other hand, since every ball about p contains p and p is a point of E, it is clear that every ball about p contains a point of E. Hence p is an adherence point of E.

Definition 6.6: *Let X be a metric space and let $E \subset X$.*

(a) *The set of all adherence points of a set E is called the* **closure** *of E and is denoted by $\bar{E}$.*

(b) *The set of all limit points of a set E is called the* **derived set** *of E and is denoted by E'.*

(c) *The set of all interior points of a set E is called the* **interior** *of E and is denoted by E^0.*

To illustrate the concepts involved in this definition, suppose that the metric space X is R^1 with the usual metric and suppose that E is the half-open interval $(0, 1]$; that is,

$$E=\{x \in R^1|0<x \leqslant 1\}$$

In this case, because the only limit point of E not belonging to E is 0, the set $\bar{E}$ of all adherence points of E is the closed interval $[0, 1]=\{x \in R^1|x=0$ or $x \in (0, 1]\}$. Furthermore, since every point of E is a limit point, the set E' of all limit points of E is also equal to $[0, 1]$. Now the only point of E that is not an interior point is 1 and hence the interior E^0 of E is $(0, 1)$.

To obtain another illustration, let the metric space X be the discrete set of points given by $X=\{\langle x, y\rangle \in R^2 | x, y \in J\}$ with the metric defined in terms of the absolute value as in equation (6–2). As always, J is the set of positive integers. Let $E=\{\langle m, n\rangle \in X | m \leq M \text{ and } n \leq N\}$ where M and N are fixed integers such that $M, N \geq 1$. Since E has no limit points, $\bar{E}=E$ and $E'=\emptyset$. Also, since, for each point $\langle m, n\rangle \in E$, any ball about $\langle m, n\rangle$ of radius less than 1 only contains a single point of E, namely, $\langle m, n\rangle$ itself, every point of E is an interior point. Hence $E^0=E$. Thus, in this case, $\bar{E}=E^0=E$. Now, if the metric space X is changed to R^2 (with metric still defined by eq. (6–2)) while E is left unchanged, it is still true that $E'=\emptyset$ and that $\bar{E}=E$. However, the points of E are no longer interior points (every ball about a point of E contains points of R^2 not belonging to E) and $E^0=\emptyset$.

Definition 6.7: *Let X be a metric space and let E be any subset of X:*
(a) *E is said to be* **dense** *in X if $\bar{E}=X$.*
(b) *E is said to be* **closed** *if $E' \subset E$.*
(c) *E is said to be* **open** *if $E^0=E$.*

Intuitively, a set is closed if none of its points are arbitrarily close to points outside the set. In the complex plane a set is open if all of its points are "inside" the set.

Clearly, the entire space X is both a closed and an open set. It follows from Definition 6.5 that *every point of a set E is an adherence point of E* and that every limit point of E is an adherence point of E. Conversely, if p adheres to a set E and is not a limit point of E, then there is a ball B about p which contains no points of the set $E-\{p\}$ but contains at least one point of E. We conclude (since $p \notin E$ implies $E=E-\{p\}$) that p is a point of E. Points for which this occurs are called *isolated points* of E. Thus a point is an adherence point of a set E if and only if it is either a point of E or a limit point of E. This can be written in symbols as

$$\bar{E}=E \cup E'$$

From this and the equivalence of relations (1–2a) and (1–2c) of chapter 1, it follows immediately that a set E is closed if and only if $\bar{E}=E$. We could then have used this condition as the definition of a closed set. Compare this with Definition 6.7(c).

It is also clear that, for any set E,

$$E^0 \subset E \subset \bar{E}$$

In the preceding discussion we have pointed out some almost immediate consequences of Definitions 6.5 to 6.7. We now prove as theorems some less direct consequences of these definitions.

Theorem 6.8: *$E \subset F$ implies $\bar{E} \subset \bar{F}$ and, for any two sets E_1 and E_2,*

$$\overline{E_1 \cup E_2} = \bar{E}_1 \cup \bar{E}_2 \tag{6–10}$$

Proof: Suppose $E \subset F$. If p is a point of E, it is also a point of F and therefore a point of $\bar{F}$. If p is a limit point of E, then it is also a limit point of F and therefore a point of $\bar{F}$. Thus $p \in \bar{E}$ implies $p \in \bar{F}$. Hence we conclude that $E \subset F$ implies $\bar{E} \subset \bar{F}$. For any two sets E_1 and E_2, this shows that $\bar{E}_1 \subset \overline{E_1 \cup E_2}$ and $\bar{E}_2 \subset \overline{E_1 \cup E_2}$. Therefore, $\bar{E}_1 \cup \bar{E}_2 \subset \overline{E_1 \cup E_2}$. Now suppose $p \in \overline{E_1 \cup E_2}$ but $p \notin \bar{E}_1 \cup \bar{E}_2$; that is $p \notin \bar{E}_1$ and $p \notin \bar{E}_2$. Then there exist balls $B(p; r_1)$ and $B(p; r_2)$ about p such that $B(p; r_1) \cap E_1 = \emptyset$ and $B(p; r_2) \cap E_2 = \emptyset$. Let r be the smaller of the positive numbers r_1 and r_2. Then $B(p; r) \subset B(p; r_1)$ and $B(p; r) \subset B(p; r_2)$. Hence $B(p; r) \cap E_1 = \emptyset$ and $B(p; r) \cap E_2 = \emptyset$. Thus $B(p; r) \cap (E_1 \cup E_2) = \emptyset$. This shows that p is not an adherence point of $E_1 \cup E_2$; that is, $p \notin \overline{E_1 \cup E_2}$, which is a contradiction. We see, then, that $p \in \overline{E_1 \cup E_2}$ implies $p \in \bar{E}_1 \cup \bar{E}_2$ (i.e., $\overline{E_1 \cup E_2} \subset \bar{E}_1 \cup \bar{E}_2$) and so we conclude that $\bar{E}_1 \cup \bar{E}_2 = \overline{E_1 \cup E_2}$.

It is easily seen that the argument used in the proof can be extended to any finite union of sets. We must point out, however, that it *cannot* be extended to any infinite union of sets since the proof depends very strongly on the fact that the smallest member of a finite set of positive numbers is a positive number. In the case of an infinite set, of course, the proper extension of the concept of smallest member is the concept of greatest lower bound, and it is not true, in general, that the greatest lower bound of a set of positive numbers is positive. We shall frequently encounter the principle involved here.

Corollary: *If F is closed and $E \subset F$, then $\bar{E} \subset F$.*

Proof: This follows from the theorem and the fact that, if F is closed, $F = F$.

Theorem 6.9: (a) *Every ball about a limit point p of a set E contains infinitely many points of E.* (b) *If A is a dense subset of the metric space $\langle X, d\rangle$ and p is a limit point of X, then p is a limit point of A.*

Proof: Part (a). The proof is by contradiction. Hence assume that there exists a ball B about p which contains only finitely many points of E and, therefore, only finitely many points of $E-\{p\}$ (B must contain at least one point of $E-\{p\}$ since p is a limit point of E). Denote those points of $E-\{p\}$ which belong to B by $q_1, q_2, \ldots, q_n$. Now there exists an integer m such that $1 \leqslant m \leqslant n$ and

$$\min_{1 \leqslant j \leqslant n} d(p, q_j) = d(p, q_m)$$

Since by construction $q_j \neq p$ for any j, it follows that $d(p, q_j) > 0$ for $1 \leqslant j \leqslant n$ and in particular that $d(p, q_m) > 0$. It is clear that, for $1 \leqslant j \leqslant n$,

$$q_j \notin B(p; d(p, q_m))$$

So the ball $B(p; d(p, q_m))$ about p contains no point of $E-\{p\}$, and p cannot be a limit point of E. This is contrary to hypothesis and therefore the assumption must be false. This proves part (a).

Part (b). Since p is a limit point of X, any ball $B(p; \epsilon)$ about p contains a point of $X-\{p\}$. Let q be such a point. Clearly $0 < d(p, q) < \epsilon$. Hence upon putting $\rho = \min \{d(p, q), \epsilon - d(p, q)\}$, we find that $\rho > 0$. If s is any point of $B(q; \rho)$, then

$$d(p, s) \leqslant d(p, q) + d(q, s) < d(p, q) + \rho \leqslant d(p, q) + \epsilon - d(p, q) = \epsilon$$

Therefore $s \in B(p, \epsilon)$. Since s was an arbitrary point of $B(q, \rho)$, we conclude from this that $B(q; \rho) \subset B(p; \epsilon)$. Now $p \notin B(q; \rho)$ since $d(p, q) \geqslant \rho$. But it follows from the fact that A is a dense subset of X that $B(q; \rho)$ contains a point of A, say t. Thus $t \in B(p, \epsilon)$ and $t \neq p$. This shows that $B(p; \epsilon)$ contains a point of $A-\{p\}$. Since $B(p; \epsilon)$ was any ball about p, we conclude that p is a limit point of A.

The following corollary is an immediate consequence of this theorem.

Corollary: *No finite set can have a limit point.*

If the set E has no limit points, then $E' = \emptyset$ and, since $\emptyset$ is a subset of every set, we see $E' \subset E$. We therefore conclude that every finite set is closed. If every ball B about p contains infinitely many points of E, then certainly it contains at least one point of $E-\{p\}$. It therefore follows from Theorem 6.9(a) that a point p is a limit point of a set E if any only if every ball about p contains infinitely many points of E. There is no reason why this could not have

355-525 O-70—6

been taken as the definition of a limit point instead of the one given in Definition 6.5(b).

Theorem 6.10: *A subset E of a metric space is open if and only if its complement is closed.*

Proof: Let E^c be closed and let p be any point of E. By definition $p \notin E^c$ and therefore p cannot be a limit point of E^c. This means that there is some ball B about p such that B contains no points of $E^c - \{p\}$. Since $p \notin E^c$, this shows that B contains no points of E^c. Thus, $q \in B$ implies $q \notin E^c$, which in turn implies $q \in E$. It follows that $B \subset E$. Thus p is an interior point of E. Since p was any point of E, we conclude that E is open.

Conversely, let E be an open set and let p be any limit point of E^c. Then every ball about p must contain at least one point of E^c; that is, no ball about p contains only points of E. This shows that p cannot be an interior point of E. But the fact that E is open then implies that p cannot be a point of E which means that it must be a point of E^c. Since p was any limit point of E^c, we conclude that E^c contains all its limit points.

Now let F be any set. Since F is the complement of F^c, the theorem shows that if F is closed then F^c is an open set and that if F is an open set then F^c is closed. Thus, the following corollary is an immediate consequence of this theorem.

Corollary: *A subset F of a metric space is closed if and only if its complement is open.*

We see from this that, having defined open sets, we could have defined the *closed sets to be just those sets which are the complements of the open sets.*

If $\langle X, d\rangle$ is any metric space, the empty set $\emptyset$ being equal to X^c must be both open and closed. It is easy to show that, in R^k, $\emptyset$ and R^k are the only subsets which are both open and closed.

The proof of the next theorem is illustrated in figure 6–1.

Theorem 6.11: *Balls are open sets.*

Proof: Suppose q is an arbitrary point of the ball $B(p; r)$. Since $d(p, q) < r$, we can find a positive number u such that

$$d(p, q) = r - u$$

Now if $t \in B(q;\ u)$, then $d(t,\ q) < u$. Therefore $d(p,\ t) \leqslant d(p,\ q) + d(q,\ t) < r - u + u = r$. This shows that $t \in B(p;\ r)$ which implies

$$B(q;\ u) \subset B(p;\ r)$$

Thus q is an interior point of $B(p;\ r)$. Since q was an arbitrary point of $B(p;\ r)$, we conclude that $B(p;\ r)$ is open.

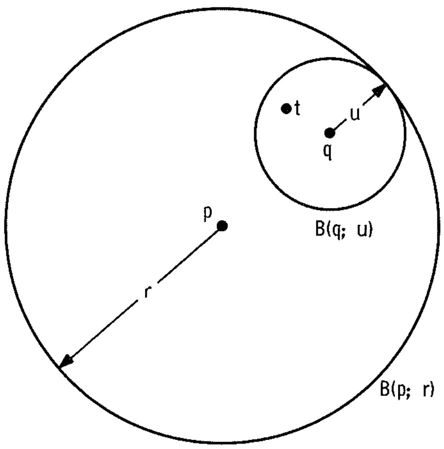


FIGURE 6–1.—Venn diagram illustrating Theorem 6.11.

The technique used in the proof of part (b) of the next theorem can easily be adapted to show that any intersection or union of a finite collection of balls about a single point p is also a ball about p. This theorem gives the principal "internal" properties of open sets.

Theorem 6.12: (a) *The union of an* **arbitrary** *collection of open sets is open.* (b) *The intersection of a* **finite** *collection of open sets is open.* (c) *The empty set and the entire space are open.*

Proof: To begin with, part (c) has already been established, and we include it here only for later reference.

Part (a). Let $\{G_\alpha | \alpha \in A\}$ be an arbitrary collection of open sets. If

$$G = \bigcup_{\alpha \in A} G_\alpha$$

$p \in G$ means that, for some $\alpha \in A$, $p \in G_\alpha$. Since G_α is an open set, there must be a ball B about p such that $B \subset G_\alpha$ and, therefore, it must also be true that $B \subset G$. This shows that p is an interior point of G and since p is an arbitrary point of G, it follows that G is open.

Part (b). Now let $\{G_1, G_2, \ldots, G_n\}$ be a finite collection of open sets. If

$$E = \bigcap_{i=1}^{n} G_i$$

$p \in E$ means that $p \in G_i$ for every $i = 1, 2, \ldots, n$. Since each G_i is an open set, there exist balls $B(p; r_i)$ about p such that $B(p; r_i) \subset G_i$ for $i = 1, 2, \ldots, n$.

If we set

$$r = \min_{1 \leq i \leq n} r_i$$

then $r > 0$ and for every $i = 1, 2, \ldots, n$,

$$B(p;\, r) \subset B(p;\, r_i) \subset G_i$$

But this shows that

$$B(p;\, r) \subset E$$

and therefore that p is an interior point of E. Hence (since p was any point of E) E is open.

Corollary 1: *Every subset of a discrete metric space is open.*

Proof: If $\langle X, d \rangle$ is a discrete metric space, then $\{p\} = B(p; 1/2)$ for every $p \in X$. If E is any subset of X, it is clear that $E = \bigcup_{p \in E} \{p\}$. Hence $E = \bigcup_{p \in E} B(p; 1/2)$. Thus E is the union of open sets. Theorem 6.12(a) now shows that E is open.

Corollary 2: *The intersection of an* **arbitrary** *collection of closed sets is closed. The union of a* **finite** *collection of closed sets is closed.*

Proof: Let $\{F_\alpha | \alpha \in A\}$ be an arbitrary collection of closed sets. The corollary to Theorem 6.10 shows that, for each $\alpha \in A$, F_α^c is open. Thus, Theorem 6.12 shows that $\bigcup_{\alpha \in A} F_\alpha^c$ is open, and so Theorem 6.10 shows that $\left(\bigcup_{\alpha \in A} F_\alpha^c\right)^c$ is closed. But taking the complement of the second DeMorgan law of table 5–I yields the identity

$$\bigcap_{\alpha\in A} F_\alpha = \left(\bigcup_{\alpha\in A} F_\alpha^c\right)^c$$

The second part of the corollary follows in the same way from part (b) of Theorem 6.12 and from the following identity obtained from table 5–I:

$$\bigcup_{i=1}^{n} F_i = \left(\bigcap_{i=1}^{n} F_i^c\right)^c$$

We might mention that, since in the Euclidean space R^1 the segments $(-1/n, 1/n)$ are open sets for $n=1, 2, 3, \ldots$ and since $\bigcap_{n=1}^{\infty} (-1/n, 1/n) = \{0\}$, part (b) of Theorem 6.12 cannot be extended to infinite collections of open sets.

Theorems 6.11 and 6.12 show that an arbitrary union of balls is an open set. On the other hand, if G is any open set and p is any point of G, there is a ball $B(p; r_p)$ about p such that $B(p; r_p) \subset G$ and therefore $\bigcup_{p\in G} B(p; r_p) \subset G$. But every point of G is in one of these balls so we conclude

$$\bigcup_{p\in G} B(p; r_p) = G \qquad (6\text{–}11)$$

Therefore every open set is a union of balls. We thus arrive at the following conclusion: *A set is open if and only if it is a union of balls.* This gives us then another way of defining open sets. Upon combining this with the remarks following Theorem 6.10, it becomes clear that once it is known which subsets of a given metric space are balls all the open and closed sets can be found.

Thus, in the case of R^1 with the usual metric, all the segments are open sets and any set which is a union of segments is an open set. In the complex plane, the interiors of disks are open sets; that is, if z is the complex variable $x+iy$ (x and y real), all sets of the form $\{z \mid |z| < r\}$ (with $r > 0$) are open. Also, any union of sets of this type is open. In three-dimensional Euclidean space, the interiors of spheres are open sets, etc. On the other hand, even though every segment is an open set when considered as a subset of the Euclidean space R^1, segments are not open sets if they are considered as subsets of the Euclidean space R^2. To see this, notice that every circle around a point on the line contains points of R^2 in its interior which are not on the line and these points do not belong to any segment.

Suppose that E_1 is an open subset of the metric space $\langle X, d\rangle$ and E_2 is an open subset of the metric space $\langle Y, \delta\rangle$. If $\langle x, y\rangle$ is any point of $E_1 \times E_2$, then there exist balls $B(x; r_1) \subset E_1$ and $B(y; r_2) \in E_2$. If we set $r = \min\{r_1, r_2\}$, then $B(x; r) \subset B(x; r_1) \subset E_1$, and $B(y; r) \subset B(y; r_2) \subset E_2$. Equation (6–9) shows that if $B(\langle x, y\rangle; r)$ is the ball about $\langle x, y\rangle$ of radius r in the direct product $\langle X \times Y, d_\times\rangle$ of $\langle X, d\rangle$ and $\langle Y, \delta\rangle$, then

$$B(\langle x, y\rangle; r) = B(x; r) \times B(y; r)$$

Hence

$$B(\langle x, y\rangle; r) \subset E_1 \times E_2$$

Since $\langle x, y\rangle$ was any point of $E_1 \times E_2$, we have just proved that *if E_1 is an open subset of a metric space $\langle X, d\rangle$ and E_2 is an open subset of the metric space $\langle Y, \delta\rangle$, then $E_1 \times E_2$ is an open subset of the direct product of the two spaces.*

We have already pointed out that, although we have proved this result only for the metric space $\langle X \times Y, d_\times\rangle$, it also holds true for the metric spaces $\langle X \times Y, d_1\rangle$ and $\langle X \times Y, d_2\rangle$ where d_1 and d_2 are the metrics defined by equations (6–6) and (6–7), respectively.

Theorem 6.12 gives us a means of forming a mathematical structure which is more general than the metric space. The statement of this theorem contains three properties that open sets must have. If these are taken as postulates, they can be used in a certain sense to define the open sets. In this manner the notion of open sets can be taken as basic instead of the notion of distance as in the case of metric space. To be more specific, suppose we are given a set X and a certain collection $\mathcal{O}$ of subsets of X. Suppose the members of $\mathcal{O}$ are called open sets and they satisfy conditions (a) to (c) of Theorem 6.12. Then, $\langle X, \mathcal{O}\rangle$ is called a *topological* space and $\mathcal{O}$ is called a *topology* for X. The concept of topological space grew out of Hausdorff's work in 1914.

Thus, given any metric space $\langle X, d\rangle$, there is a topological space $\langle X, \mathcal{O}\rangle$ (having the same basic set X) *associated* with it in such a way that the members of $\mathcal{O}$ are just the open sets in $\langle X, d\rangle$. Given any set X there are usually many ways of defining a distance on that set. Thus, for the same set X, we may have two different metric spaces $\langle X, d\rangle$ and $\langle X, d'\rangle$ corresponding to the different metrics d and d'. We have already encountered this situation when we defined a metric on the direct product of two metric spaces in terms of the metrics defined on each of the component metric spaces. It may turn out, however, that these different metric spaces $\langle X, d\rangle$ and $\langle X, d'\rangle$ are associated with the same topological space.[18] If this is the case, the metrics d and d' are said to *give rise to the same topology* for X or to be *topologically equivalent.* We shall give a more precise definition of this concept in chapter 9.

On the other hand, not every topological space can be associated with a metric space in this manner. Those which can be are said to be *metrizable.* Since a topological space is really such a general structure, it turns out that metrizable spaces possess many desirable properties which general topological spaces do not. This condition is ameliorated in practice by restricting the topological spaces further by the imposition of one or more postulates in addition to those already discussed. In this way several different types of topological spaces arise which are still more general than metrizable spaces but have many of their desirable features.

In topological spaces, the concept of ball is replaced by the concept of *neighborhood.* A *neighborhood* of a point is defined to be any open set which contains that point. It turns out, as we shall see, that much of what will be said about general metric spaces can be expressed by using the concept of ball instead of referring to the metric explicitly. In turn, since every ball about a point is also a neighborhood and every neighborhood of a point contains a ball about that point, much of this still goes through when balls are replaced by neighborhoods. In this way a large part of theory of metric spaces developed here can be applied to topological spaces. The preceding discussion shows that many of the properties of a given metric space can be completely specified in terms of the open sets (neighborhoods) of that metric space. Such properties are then also intrinsic properties of the associated topological space and are therefore called the *topological properties* of the metric space (as opposed to the purely metric properties). For example, if p is an adherence point of a set E and V is any neighborhood of p, there is a ball B about p such that $B \subset V$. Since B contains a point of E, so does V. Thus every neighborhood of p contains a point of E. Conversely, since balls are neighborhoods, if every neighborhood of p contains a point of E, then certainly every ball about p must also. Hence a point p is an adherence point of a set E if and only if every neighborhood of p

[18] Actually the three metrics given in equations (6–5), (6–6), and (6–7) determine the same topological space. In fact, it can be shown that there exist constants c_1 and c_2 such that

$$c_1 d_i(p, q) \leqslant d_j(p, q) \leqslant c_2 d_i(p, q)$$

where $i, j = \times, 1, 2$. We shall see in chapter 9 that any two metrics that satisfy a relation of this type determine the same topological space.

contains a point of E. We see then that whether or not a given point is an adherence point of a given set is a topological property since it, in effect, depends only on which sets are open. For a further discussion of topological spaces, the reader is referred to references 2 to 4.

Theorem 6.13: *For any set E, $\bar{E}$ is closed.*

Proof: According to the corollary to Theorem 6.10, it is sufficient to prove $\bar{E}^c$ is open. Let p be any point of $\bar{E}^c$. If we can show that p is an interior point of $\bar{E}^c$, we are done. Evidently $p \notin \bar{E}$. Therefore, there exists a ball $B(p; r)$ about p which contains no points of E; that is, $B(p; r) \subset E^c$. Theorem 6.11 shows that $B(p; r)$ is open. Hence, if q is any point of $B(p; r)$, there exists a ball $B(q; \rho)$ about q such that $B(q; \rho) \subset B(p; r) \subset E^c$. This shows that $B(q; \rho)$ contains no points of E and therefore that q is not an adherence point of E. Since q was arbitrary, we conclude that no point of $B(p; r)$ is an adherence point of E. Hence no point of $B(p; r)$ is a point of $\bar{E}$; that is, $B(p; r) \subset \bar{E}^c$. Thus p is an interior point of $\bar{E}^c$.

If p is a point of a metric space $\langle X, d\rangle$ and r is a nonnegative number then the set

$$C(p; r) = \{x \in X \mid d(p, x) \leqslant r\}$$

is often called the **closed** *ball of radius r about p*. Let q be any point of $C^c(p; r)$, the complement of $C(p; r)$. Then $d(p, q) > r$. Hence upon setting $\rho = d(p, q) - r$, we find that $\rho > 0$. Now if $y \in B(q; \rho)$ then

$$d(p, y) \geqslant d(p, q) - d(q, y) > d(p, q) - \rho = r + \rho - \rho = r$$

Hence, $y \notin C(p; r)$; that is, $y \in C^c(p; r)$. Since y was any point of $B(q; \rho)$, we conclude that $B(q; \rho) \subset C^c(p; r)$. And, since q was any point of $C^c(p; r)$, we conclude that $C^c(p; r)$ is open. The corollary to Theorem 6.10 now shows that $C(p; r)$ is indeed closed. However, it is *not* true in general that $C(p; r) = \overline{B(p; r)}$. For if $\langle X, d\rangle$ is a discrete metric space then for any point $p \in X$

$$\overline{B(p; 1)} = \overline{\{p\}} = \{p\}$$

But $C(p; 1) = X$.

Let us emphasize that *when the term "ball" is used it always refers to an* **open** *ball*. The closed balls in R^1 are intervals.

We see from Theorem 6.13 and the corollary of Theorem 6.8 that $\bar{E}$ is the smallest closed set which contains E. Let $\Omega = \{F_\alpha | \alpha \in A\}$ be the collection of all closed sets which contain E. The collection Ω is certainly not empty since the whole space itself is in Ω. The second corollary to Theorem 6.12 shows that $\bigcap_{\alpha \in A} F_\alpha$ is closed and, since $E \subset F_\alpha$ for every $\alpha \in A$, we see that $E \subset \bigcap_{\alpha \in A} F_\alpha$. On the other hand, $\bigcap_{\alpha \in A} F_\alpha$ is a subset of every closed set which contains E; that is, $\bigcap_{\alpha \in A} F_\alpha$ is also the smallest closed set which contains E. Evidently then

$$\bar{E} = \bigcap_{\alpha \in A} F_\alpha$$

Thus *we could have defined* the closure of a set E to be the smallest closed set which contains E or, equivalently, the intersection of all closed sets which contain E.

It is easy to see that in R^1 (with the usual metric) every interval is a closed set and in Euclidean space R^k **every** k-cell **is a closed set.**

Theorem 6.14: (a) *Every closed set of real numbers which is bounded above contains its least upper bound.* (b) *Every closed set of real numbers which is bounded below contains its greatest lower bound.*

Proof: Part (a). Let E be any closed set of real numbers which is bounded above and set $b = \sup E$ (which exists by axiom III of chapter 2). For every positive number δ, $b - \delta$ is not an upper bound of E. Hence there exists a point $p \in E$ such that $b - \delta < p$. Now if we assume $b \notin E$, we can conclude that $b - \delta < p < b$. Since in R^1 balls are segments (i.e., $B(b; \delta) = (b - \delta, b + \delta)$), we see that every ball about b contains a point of $E - \{b\}$. This shows that b is a limit point of E. But since E is closed, this implies $b \in E$.

Part (b). The proof is similar to that of part (a).

We have already observed that, if Y is any subset of the metric space $\langle X, d\rangle$, Y itself is a metric space with the same metric; that is, $\langle Y, d\rangle$ is a metric space. We have seen by example, however, that, if $E \subset Y \subset X$ and E is an open subset of the metric space $\langle Y, d\rangle$, E need not be an open subset of the metric space $\langle X, d\rangle$. Of course, the same must be true for closed sets since they are merely the complements of the open sets. Actually it turns out that there is a simple relation between the open sets in $\langle X, d\rangle$ and those in $\langle Y, d\rangle$. Since the property of being open is really defined in terms of balls, we shall first discuss the relation between the balls in the metric space $\langle X, d\rangle$ and those in the metric space $\langle Y, d\rangle$. Let us temporarily denote balls in the metric space $\langle X, d\rangle$ by attaching the superscript X to the usual notation and those in the metric space $\langle Y, d\rangle$ by attaching a superscript Y. Thus, for example, $B^X(p; r)$ denotes a ball in $\langle X, d\rangle$ and $B^Y(q; \rho)$ denotes a ball in $\langle Y, d\rangle$. We now look at the definition of balls. Equation (6–8) of Definition 6.2 tells us that the balls in $\langle X, d\rangle$ are sets of the form

$$B^X(p; r) = \{q \in X \mid d(p, q) < r\} \qquad r > 0$$

and the balls in $\langle Y, d\rangle$ are sets of the form

$$B^Y(p; r) = \{q \in Y \mid d(p, q) < r\} \qquad r > 0$$

where of course it is understood here that $p \in Y$. It follows from these relations that if $B^Y(p; r)$ is any ball in $\langle Y, d\}$ about any point $p \in Y$, then

$$\begin{aligned} B^Y(p; r) &= \{q \in Y \mid d(p; q) < r\} = \{q \in X \mid d(p, q) < r\} \cap Y \\ &= B^X(p; r) \cap Y \end{aligned}$$

Of course, this also shows that if p is any point of Y and $B^X(p; r)$ is any ball about p in the metric space $\langle X, d\rangle$, then $B^X(p; r) \cap Y$ is a ball about p (of radius r) in the subspace $\langle Y, d\rangle$. Thus the following theorem has been proved.

Theorem 6.15: *Let $\langle X, d\rangle$ be a metric space and suppose $Y \subset X$. Then, for any point $p \in Y$, the set A is a ball about p of radius r in the subspace $\langle Y, d\rangle$ if and only if there is a ball $B^X(p; r)$ about p of radius r in the metric space $\langle X, d\rangle$ such that*

$$A = Y \cap B^X(p; r)$$

This theorem is illustrated in figure 6–2.

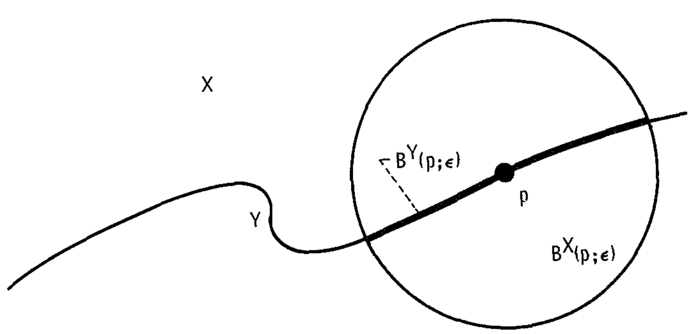


FIGURE 6–2.—Balls in subspaces.

We now turn to the general case of open sets, and prove the following theorem.

Theorem 6.16: *Let $\langle X, d\rangle$ be a metric space and suppose $Y \subset X$. Then a subset G of Y is open in the subspace $\langle Y, d\rangle$ if and only if there exists an open subset H of the metric space $\langle X, d\rangle$ such that*

$$G = Y \cap H \tag{6–12}$$

Proof: First suppose that G is an open set in the metric space $\langle Y, d\rangle$. We have already shown in the remarks following the corollaries to Theorem 6.12 (eq. (6–11)) that, for every $p \in G$, there exists a positive number r_p such that

$$G = \bigcup_{p \in G} B^Y(p; r_p)$$

Theorem 6.15 now shows that, for each $p \in Y$,

$$B^Y(p; r_p) = Y \cap B^X(p; r_p)$$

Hence

$$G = \bigcup_{p \in G} [Y \cap B^X(p; r_p)] = Y \cap \left[\bigcup_{p \in G} B^X(p; r_p)\right]$$

and Theorems 6.11 and 6.12 show that the set $H = \bigcup_{p \in G} B^X(p; r)$ is an open set in the metric space $\langle X, d\rangle$.

Conversely, suppose G is given by equation (6–12) and H is an open set in the metric space $\langle X, d\rangle$. If p is any point of G, then $p \in H$ and, since H is open in the metric space $\langle X, d\rangle$, we can find a ball $B^X(p; \rho)$ about p such that $B^X(p; \rho) \subset H$. Hence

$$Y \cap B^X(p; \rho) \subset Y \cap H = G$$

but Theorem 6.15 shows that $Y \cap B^X(p; \rho)$ is a ball in the metric space $\langle Y, d\rangle$ about p. Thus in the metric space $\langle Y, d\rangle$ all the points of G are interior points.

Corollary 1: *If $\langle X, d\rangle$ is a metric space and $E \subset X$, a necessary and sufficient condition that every subset D of E which is open in the metric space $\langle E, d\rangle$ be open in the metric space $\langle X, d\rangle$ is that E is an open subset of $\langle X, d\rangle$.*

Proof: To see that this condition is necessary, we need only consider the case when $D = E$. The sufficiency of the condition follows from the theorem and the fact that (Theorem 6.12(b)) the intersection of two open sets is open.

Corollary 2: *If $\langle X, d\rangle$ is a metric space and E is any subset of X, then the set $D \subset E$ is closed in the metric space $\langle E, d\rangle$ if and only if there exists a closed subset F of $\langle X, d\rangle$ such that $D = F \cap E$.*

Proof: By the corollary to Theorem 6.10, D is a closed subset of $\langle E, d\rangle$ if and only if the complement of D in E, $E-D$, is an open subset of $\langle E, d\rangle$. Therefore it follows from the theorem that D is a closed subset of $\langle E, d\rangle$ if and only if there exists an open set G of $\langle X, d\rangle$ such that

$$E-D=G \cap E \tag{6–13}$$

Since $E-D=E \cap D^c$, equation (6–13) is equivalent to

$$E^c \cup (E \cap D^c)=E^c \cup (G \cap E)$$

and hence by the distributive law (table 1–I) to

$$E^c \cup D^c=E^c \cup G$$

DeMorgan's law now shows that equation (6–13) is equivalent to

$$E \cap D=E \cap G^c$$

But since $D \subset E$, equation (6–13) is also equivalent to

$$D=E \cap G^c$$

The corollary to Theorem 6.10 now shows that the assertion is true if we take $F=G^c$.

Definition 6.17: *An* **open cover** *of a subset E of a metric space is any family $\{G_\alpha | \alpha \in A\}$ of open subsets of the metric space such that $E \subset \bigcup_{\alpha \in A} G_\alpha$.*

Definition 6.18: *A subset K of a metric space is called* **compact** *if every open cover of K contains a* **finite** *subcover of K. A metric space is called a* **compact space** *if it is a compact subset of itself.*

This means that if $\{G_\alpha | \alpha \in A\}$ is any open cover of K then there is a finite number of the α's, say $\alpha_1, \alpha_2, \ldots, \alpha_n$, such that

$$K \subset G_{\alpha_1} \cup G_{\alpha_2} \cup \cdots \cup G_{\alpha_n}$$

The definition of compactness given in this form shows clearly that it is a topological property. Although compact sets are extremely important, it is not easy to give a physical picture of the property of compactness. It is clear that every finite set is compact and that the union of a finite number of compact sets is compact. In fact, compact sets have many of the properties of finite sets even though they are frequently uncountable sets.[19] That there is a large class of uncountable compact sets in the very important spaces R^k will be shown subsequently. It is not surprising then that compact sets have some very desirable features (especially in connection with continuity which is discussed in chapter 8). Among these is the fact that, in contrast to the properties of being open or closed, the property of compactness is independent of the metric space in which the set is embedded. The next theorem shows this.

Theorem 6.19: *Let $\langle X, d\rangle$ be a metric space, and suppose $K \subset Y \subset X$. Then K is a compact subset in the metric space $\langle X, d\rangle$ if and only if it is a compact subset of the subspace $\langle Y, d\rangle$.*

Proof: Let K be a compact subset of the metric space $\langle X, d\rangle$ and let $\{G_\alpha | \alpha \in A\}$ be any family of open subsets of the metric space $\langle Y, d\rangle$ such that $K \subset \bigcup_{\alpha \in A} G_\alpha$.

It follows from Theorem 6.16 that there is a family $\{H_\alpha | \alpha \in A\}$ of open subsets of the metric space $\langle X, d\rangle$ such that, for each $\alpha \in A$, $G_\alpha = Y \cap H_\alpha$. Since $G_\alpha \subset H_\alpha$, it is obvious that $K \subset \bigcup_{\alpha \in A} H_\alpha$ and, since K is a compact subset of the metric space $\langle X, d\rangle$, there must be a finite collection of indices, say $\alpha_1, \alpha_2, \ldots, \alpha_n \in A$, such that

$$K \subset H_{\alpha_1} \cup \ldots \cup H_{\alpha_n}$$

Because it is also true that $K \subset Y$, we see

$$\begin{aligned} K \subset Y \cap (H_{\alpha_1} \cup \ldots \cup H_{\alpha_n}) &= (Y \cap H_{\alpha_1}) \cup \ldots \cup (Y \cap H_{\alpha_n}) \\ &= G_{\alpha_1} \cup \ldots \cup G_{\alpha_n} \end{aligned}$$

This proves that K is a compact subset of the metric space $\langle Y, d\rangle$.

Now suppose that K is a compact subset of the metric space $\langle Y, d\rangle$ and let $\{H_\alpha | \alpha \in A\}$ be a family of open subsets of the metric space $\langle X, d\rangle$ such that $K \subset \bigcup_{\alpha \in A} H_\alpha$. For each $\alpha \in A$, we define G_α by

[19] For a fuller discussion of this point, see ref. 5.

$$G_\alpha = Y \cap H_\alpha$$

and Theorem 6.16 shows that every G_α is an open subset of the metric space $\langle Y, d \rangle$.

Now, since $K \subset \bigcup_{\alpha \in A} H_\alpha$ and $K \subset Y$, it follows that

$$K \subset Y \cap \left(\bigcup_{\alpha \in A} H_\alpha \right) = \bigcup_{\alpha \in A} (Y \cap H_\alpha) = \bigcup_{\alpha \in A} G_\alpha$$

where one of the distributive laws of table 5–I has been used. Hence $\{G_\alpha | \alpha \in A\}$ is a collection of open subsets of $\langle Y, d \rangle$ which covers K and, since K is a compact subset of $\langle Y, d \rangle$, there is a finite subcollection of indices, say $\alpha_1, \ldots, \alpha_n \in A$, such that

$$\begin{aligned} K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n} &= (Y \cap H_{\alpha_1}) \cup \cdots \cup (Y \cap H_{\alpha_n}) \\ &= Y \cap (H_{\alpha_1} \cup \cdots \cup H_{\alpha_n}) \subset H_{\alpha_1} \cup \cdots \cup H_{\alpha_n} \end{aligned}$$

Since $\{H_\alpha | \alpha \in A\}$ was any open cover of K in the metric space $\langle X, d \rangle$, this shows that K is a compact subset of $\langle X, d \rangle$.

In Theorem 6.19 we can, in particular, let the sets Y and K be the same set. Thus, if we prove under certain conditions that the metric space $\langle K, d \rangle$ is compact, or if some subset E of $\langle K, d \rangle$ is compact, we can conclude that, if $\langle K, d \rangle$ is a subspace of X, then K is a compact subset of X or that E is a compact subset of X, etc. Thus, as a consequence of this theorem, it is sufficient to prove many theorems about compact sets only for compact spaces. We might point out that although we defined compact metric spaces it would be absurd to define a closed or open metric space. After all, every metric space is both a closed and an open subset of itself.

We can use the concept of distance between points in a metric space to associate certain "distances" with sets.

Definition 6.20: *Let $\langle X, d \rangle$ be a metric space and suppose that E is a nonempty subset of X. For any $s \in X$, the number* $\operatorname{lub}_{p \in E} d(s, p)$ *is called the* **distance between the set** E **and the point** s. *It is denoted by* $d(s, E)$. *The number* $\sup_{p, q \in E} d(p, q)$ *is called the* **diameter** *of E and is denoted by $d(E)$. If H is any other subset of X we call the number* $\inf_{\substack{p \in E \\ q \in H}} d(p, q)$ *the* **distance between the two sets** E *and H. It is denoted by $d(E, H)$.*

Clearly the diameter of any set is a nonnegative number or $+\infty$. If a set A contains more than one point, let p and q be any two distinct points of A. Then $d(A) \geq d(p, q) > 0$. Hence, we conclude that $d(A)=0$ if and only if A contains precisely one point.

According to the definition, the diameter of a ball in the Euclidean space R^k is equal to twice its radius and the diameter of a k-cell is equal to the length of the "diagonal" of the k-cell.

In the Euclidean plane R^2 the distance from a point p to a line L is the perpendicular distance from p to L.

If A is any nonempty subset of the metric space $\langle X, d\rangle$ and p and q are any two points of X, then

$$|d(p, A) - d(q, A)| \leq d(p, q) \tag{6–14}$$

In order to see this, notice that for every $t \in A$

$$d(p, t) \leq d(p, q) + d(q, t)$$

Hence

$$d(p, A) = \inf_{t \in A} d(p, t) \leq \inf_{t \in A} (d(p, q) + d(q, t)) = d(p, q) + \inf_{t \in A} d(q, t) = d(p, q) + d(q, A)$$

In precisely the same way we can prove that

$$d(q, A) \leq d(p, q) + d(p, A)$$

Therefore

$$-d(p, q) \leq d(p, A) - d(q, A) \leq d(p, q)$$

and relation (6–14) follows.

Definition 6.21: *A subset E of the metric space $\langle X, d\rangle$ is* **bounded** *if there is a* **finite** *real number M such that $d(E) \leq M$.*

It is not hard to show that the bounded subsets of a metric space are just those sets which are subsets of the balls (recall that a ball has a *finite* radius).

There is a relation between compactness and the properties of being closed and of being bounded.

Theorem 6.22: *Compact subsets of metric spaces are closed and bounded.*

Proof: Suppose K is any compact subset of any metric space $\langle X, d\rangle$. We first show that K is closed. To accomplish this, we need only prove that K^c is an open subset of X and apply the corollary to Theorem 6.10.

Hence fix any point p of K^c and, for every $q \in K$, let $B(p; r_q)$ and $B(q; \rho_q)$ be balls about p and q, respectively, such that both r_q and ρ_q are less than $d(p, q)/2$. Now

$$K \subset \bigcup_{q \in K} B(q; \rho_q)$$

Since K is compact, there must be a finite number of points of K, say $q_1, \ldots, q_n$, such that

$$K \subset B(q_1; \rho_{q_1}) \cup \ldots \cup B(q_n; \rho_{q_n}) = G$$

Let r^* be the smallest of the numbers $r_{q_1}, r_{q_2}, \ldots, r_{q_n}$. Then

$$B(p; r^*) = B(p; r_{q_1}) \cap \ldots \cap B(p; r_{q_n})$$

Now if s is any point of G then, for some $j(1 \leq j \leq n)$, $s \in B(q_j; \rho_{q_j})$. Therefore $s \notin B(p; r_{q_j})$. Hence, $s \notin B(p; r^*)$. This shows $B(p; r^*) \cap G = \emptyset$. Therefore, $B(p; r^*) \cap K = \emptyset$; that is, $B(p; r^*) \subset K^c$. Thus p is an interior point of K^c. Since p was any point of K^c, this proves that K^c is open.

We will now show that K is bounded. To this end consider the collection $\{B(p; 1) | p \in K\}$. Since balls are open and $K \subset \bigcup_{p \in K} B(p; 1)$, it is clear that this collection is an open cover of K and, since K is compact, must contain a finite subcover; that is, there must be a finite number of points of K, say $p_1, \ldots, p_n$, such that

$$K \subset B(p_1; 1) \cup \ldots \cup B(p_n; 1)$$

We now define

$$\delta = \max_{1 \leq i \leq j \leq n} d(p_i, p_j)$$

Of course, $M = \delta + 2$ is a finite number. If q and s are any two points of K, there must be elements of the set $\{p_1, \ldots, p_n\}$, say p_i and p_j, such that $q \in B(p_i; 1)$ and $s \in B(p_j; 1)$ (note that we may have $p_i = p_j$). Then,

$$\begin{aligned} d(q, s) &\leq d(q, p_i) + d(p_i, p_j) + d(p_j, s) \\ &< 1 + \delta + 1 = M \end{aligned}$$

so M is finite and is an upper bound of the set $\{d(q, s) \mid s, q \in K\}$. Therefore the least upper bound $d(K)$ of this set must be less than or equal to M which proves the theorem.

The converse of this theorem is by no means always true! We shall see however that it is true for Euclidean spaces.

We might point out that the proof of boundedness *cannot* be restated purely in terms of neighborhoods and so the fact that compact sets are bounded is a metric and not a topological property. On the other hand, the proof of the fact that compactness implies that a set is closed depends only on the fact that we can always find for any two distinct points two nonoverlapping open sets each of which contains one of the points. Topological spaces with this property are called Hausdorff spaces.

Theorem 6.23: *If F is a closed subset of a compact subset K of a metric space X, then F is compact.*

Proof: Let $\{G_\alpha \mid \alpha \in A\}$ be any open cover of F. We shall show that it has a finite subcover. Since F is closed, F^c is open and hence the collection $\Omega = \{G_\alpha \mid \alpha \in A\} \cup \{F^c\}$ is an open cover of K. Since K is compact, there is a finite subcollection $\Gamma \subset \Omega$ which covers K and hence also $F \subset K$. If $F^c \in \Gamma$, then $\Gamma - \{F^c\}$ is still an open cover of F. But $\Gamma - \{F^c\}$ is a finite subcollection of $\{G_\alpha \mid \alpha \in A\}$, and this proves the theorem.

The following is an almost immediate corollary of this theorem.

Corollary: *If K is compact and F is closed, then $F \cap K$ is compact.*

Proof: Theorem 6.22 shows K is closed and the second corollary to Theorem 6.12 shows that $K \cap F$ is closed. Hence $K \cap F$ is a closed subset of K and Theorem 6.23 now shows that $K \cap F$ is compact.

Definition 6.24: *If every infinite subset of a set E has a limit point in E, then E is said to be* **countably compact.**

Theorem 6.25: *Compact sets are countably compact.*

Proof: The proof is by contradiction. Hence, let K be compact and let E be an infinite subset of K which has no limit point in K. Then there is a ball $B(q; r_q)$ about every point $q \in K$ which contains no points of $E - \{q\}$. That is, $B(q; r_q)$ contains at most one point of E, viz., q if $q \in E$. It is clear that $K \subset \bigcup_{q \in K} B(q; r_q)$,

355-525 O-70—7

and Theorem 6.11 shows that each $B(q; r_q)$ is open. Hence, $\Omega=\{B(q; r_q)\,|\,q\in K\}$ is an open cover of K but no finite subcollection of Ω can cover E. Evidently the same must be true of K for if any finite subcollection of Ω covered K, it would also cover $E \subset K$. This contradicts the hypothesis that K is compact and so proves the theorem.

We shall now turn to a consideration of some of the properties of intersections of compact sets.

Definition 6.26: *A sequence of sets* $\{F_i\}$ *is a* **nested sequence of sets** *if*

$$F_1 \supset F_2 \supset F_3 \supset \ . \ . \ .$$

Theorem 6.27: *Let* $\{K_\alpha|\alpha \in A\}$ *be a collection of compact subsets of a metric space and suppose that the intersection of every* **finite** *subcollection of* $\{K_\alpha|\alpha \in A\}$ *is nonempty. Then* $\bigcap_{\alpha\in A} K_\alpha$ *is not empty.*

Proof: Set $G_\alpha=K_\alpha^c$ and choose any member, say K_1, of $\{K_\alpha|\alpha \in A\}$. The proof is by contradiction. Hence assume $\bigcap_{\alpha\in A} K_\alpha=\emptyset$. This implies that there is no point of K_1 that belongs to every K_α for otherwise this point would be in the intersection. Hence, $K_1 \subset \bigcup_{\alpha\in A} G_\alpha$ and the G_α are open by Theorem 6.22 and the corollary to Theorem 6.10. Thus, $\{G_\alpha|\alpha \in A\}$ is an open cover of K_1 and, since K_1 is compact, there are finitely many α's, say $\alpha_1, \ . \ . \ ., \alpha_n$, such that

$$K_1 \subset G_{\alpha_1} \cup \ . \ . \ . \ \cup G_{\alpha_n}=(K_{\alpha_1} \cap \ . \ . \ . \ \cap K_{\alpha_n})^c$$

by DeMorgan's law (table 5–I). But this means that

$$K_1 \cap (K_{\alpha_1} \cap \ . \ . \ . \ \cap K_{\alpha_n})=\emptyset$$

which contradicts to the hypothesis.

Corollary: *If* $\{K_i\}$ *is a nested sequence of nonempty compact sets in the metric space* $\langle X, d\rangle$, *then* $K=\bigcap_{i=1}^{\infty} K_i$ *is also not empty.*

Proof: This follows immediately from Theorem 6.27 and from the fact that the intersection of any finite subcollection of $\{K_i\}$ is just equal to the smallest K_j in the collection and, hence, is not empty.

Theorem 6.28: If $\{I_i\}$ is a nested sequence of intervals (in R^1), then $\bigcap_{i=1}^{\infty} I_i$ is not empty.

Proof: Let a_i and b_i be the left and right ends of I_i, respectively, and let $A=\{a_i \mid i=1, 2, 3, \ldots\}$. Certainly A is not empty, and it is clear that it must be bounded above by b_1. By axiom III of chapter 2 (see p. 18), lub A exists. Hence let $a=\text{lub } A$. If i and j are any two positive integers, it is clear that

$$a_i \leqslant a_{i+j} \leqslant b_{i+j} \leqslant b_j$$

Hence, for every i, $a_i \leqslant b_j$; that is, b_j is an upper bound for A for each integer j. It follows from Definition 2.2 that lub $A=a \leqslant b_j$ and $a_j \leqslant a$. Evidently $a \in I_j$ for every j which shows that $a \in \bigcap_{j=1}^{\infty} I_j$.

We can easily generalize this theorem to k-cells in R^k.

Theorem 6.29: *If $\{I_i\}$ is a nested sequence of k-cells (in R^k), then $\bigcap_{i=1}^{\infty} I_i$ is not empty.*

Proof: Definition 6.3 shows that for each i there are k intervals, $I_{i,1}, I_{i,2}, \ldots, I_{i,k}$, such that

$$I_i = \bigtimes_{j=1}^{k} I_{i,j}$$

Now it is clear that the condition

$$I_{i+1} \subset I_i$$

implies

$$I_{i+1,j} \subset I_{i,j} \qquad 1 \leqslant j \leqslant k$$

Hence, for each fixed j, $\{I_{i,j}\}$ is a nested sequence of intervals and the preceding theorem shows that there is a real number $a_j \in I_{i,j}$ for every $i=1, 2, 3, \ldots$ So if we set $\boldsymbol{a}^* = \langle a_1, a_2, \ldots, a_k \rangle$, it is clear that, for every $i=1, 2, 3, \ldots, \boldsymbol{a}^* \in I_i$ which shows that $\boldsymbol{a}^* \in \bigcap_{i=1}^{\infty} I_i$.

We shall now use this theorem to prove a very important fact about R^k.

Theorem 6.30: *All k-cells are compact.*

Proof: Suppose Q is an arbitrary k-cell. Definition 6.3 shows that there are real numbers a_i and b_i for $1 \leqslant i \leqslant k$ such that $Q = \bigtimes_{i=1}^{k} [a_i, b_i]$. Now we see that (from the definition of the norm in the Euclidean space R^k), for any $\boldsymbol{x}, \boldsymbol{y} \in Q$ with $\boldsymbol{x} = \langle x_1, \ldots, x_k \rangle$ and $\boldsymbol{y} = \langle y_1, \ldots, y_k \rangle$

$$|\boldsymbol{x} - \boldsymbol{y}| = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} \leqslant \sqrt{\sum_{i=1}^{k} (b_i - a_i)^2} = d(Q)$$

($d(Q)$ is the length of the "diagonal" of Q). Now suppose we are given a k-cell

$$Q^0 = \bigtimes_{i=1}^{k} [a_i^0, b_i^0] \tag{6–15}$$

and suppose in order to arrive at a contradiction that we can find some collection of open sets $\{G_\alpha | \alpha \in A\}$ which covers Q^0 and which contains no finite subcover. The numbers $c_i^0 = (a_i^0 + b_i^0)/2$ (for $1 \leqslant i \leqslant k$) lie between a_i^0 and b_i^0. By replacing some of the intervals in equation (6–15), say $[a_{i_1}^0, b_{i_1}^0], \ldots, [a_{i_n}^0, b_{i_n}^0]$, by the intervals $[a_{i_1}^0, c_{i_1}^0], \ldots, [a_{i_n}^0, c_{i_n}^0]$ and the remainder, say $[a_{i_{n+1}}^0, b_{i_{n+1}}^0], \ldots, [a_{i_k}^0, b_{i_k}^0]$, by the intervals $[c_{i_{n+1}}^0, b_{i_{n+1}}^0], \ldots, [c_{i_k}^0, b_{i_k}^0]$, we can form 2^k different k-cells whose union is Q^0, and each of them has the length of its diagonal equal to half the length of the diagonal of Q^0. We can find at least one of these k-cells which we denote by $Q^1 = \bigtimes_{i=1}^{k} [a_i^1, b_i^1]$ that cannot be covered by any finite subcollection of $\{G_\alpha | \alpha \in A\}$ for if they all could be so covered then so could Q^0. Hence we have $Q^1 \subset Q^0$, $d(Q^1) = \frac{1}{2} d(Q^0)$, and Q^1 is not covered by any finite subcollection of $\{G_\alpha | \alpha \in A\}$. Having obtained the k-cells

$$Q^0 \supset Q^1 \supset \cdots \supset Q^n$$

$$Q^j = \bigtimes_{i=1}^{k} [a_i^j, b_i^j] \qquad 0 \leqslant j \leqslant n$$

none of which can be covered by any finite subcollection of $\{G_\alpha | \alpha \in A\}$ and for which $d(Q^j) = (1/2^j) d(Q^0)$ for $1 \leqslant j \leqslant n$, we can define $c_i^n = (a_i^n + b_i^n)/2$ $(1 \leqslant i \leqslant k)$ and form the intervals $[a_i^n, c_i^n]$ and $[c_i^n, b_i^n]$ $(1 \leqslant i \leqslant k)$. From these intervals we can, in the same way as just described, form 2^k different k-cells whose union is Q^n, and each of these cells has the length of its diagonal equal to $(1/2)d(Q^n)$. Again one of these k-cells, say Q^{n+1}, cannot be covered by any finite subcollection of $\{G_\alpha | \alpha \in A\}$ and $Q^{n+1} \subset Q^n$, so that $d(Q^{n+1})$

$= (1/2)d(Q^n) = (1/2^{n+1})d(Q^0)$. In this way we construct a nested sequence of k-cells $\{Q^n\}$ such that none of these can be covered by any finite subcollection of $\{G_\alpha | \alpha \in A\}$ and such that, for any positive integer n, the condition $\boldsymbol{x}, \boldsymbol{y} \in Q^n$ implies $|\boldsymbol{x}-\boldsymbol{y}| \leqslant d(Q^n)=(1/2^n)d(Q^0)$. This method of construction of a nested sequence of k-cells is illustrated in figure 6–3 for $k=2$.

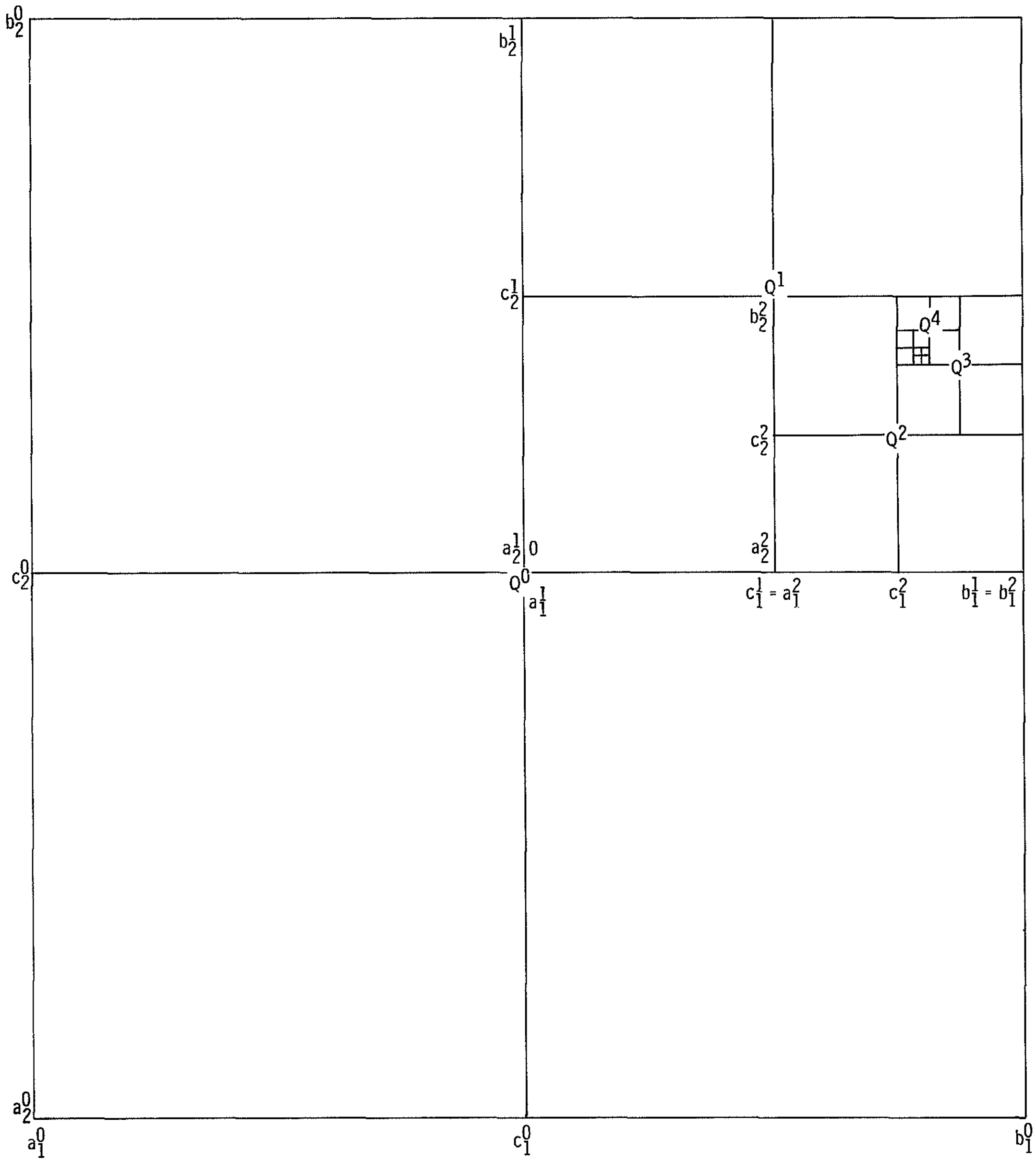


FIGURE 6–3.—Nested sequence of 2-cells.

Theorem 6.29 now shows that there is a least one point z such that, for *every* n, $z \in Q^n$. Since $\{G_\alpha | \alpha \in A\}$ covers Q^0, there must be at least one α, say α_1, such that $z \in G_{\alpha_1}$. Since G_{α_1} is certainly an open set, there exists a ball $B(z; r)$ about z such that $B(z; r) \subset G_{\alpha_1}$. Now choose m so large that $2^{-m}d(Q^0) < r$. (If there were no such m, then for every positive integer p we would have $2^p \leqslant d(Q^0)/r$ which is absurd.) Now $|z - y| \leqslant (1/2^m)d(Q^0) < r$ for every $y \in Q^m$. This shows that $Q^m \subset B(z; r) \subset G_{a_1}$ which contradicts the conclusion that Q^m cannot be covered by any finite subcollection of $\{G_\alpha | \alpha \in A\}$, and this proves the theorem.

We have seen from Theorem 6.22 that in metric spaces compact sets must be closed and bounded, but it was pointed out that the converse is not always true. It is a very important fact that in Euclidean spaces, however, the compactness of a set is equivalent to its being closed and bounded. This is the statement of the well-known theorem called the Heine-Borel theorem.

Theorem 6.31: *A subset of the Euclidean space R^k is compact if and only if it is closed and bounded.*

Proof: Theorem 6.22 of course shows that every compact set is closed and bounded. Therefore we need only prove that every closed and bounded subset of R^k is compact. But every bounded subset of R^k must be contained in some k-cell. Since Theorem 6.30 shows that this k-cell must be compact and since Theorem 6.23 shows that closed sets contained in compact sets are themselves compact, the assertion is proved.

There is another well-known theorem, the Bolzano-Weierstrass theorem, which also follows easily from the fact that k-cells are compact subsets of R^k.

Theorem 6.32: *Every bounded infinite subset of the Euclidean space R^k has a limit point in R^k.*

Proof: Let E be a bounded infinite subset in R^k. Since E is bounded we can find a k-cell Q such that $E \subset Q$. Theorem 6.30 shows that Q is compact and Theorem 6.25 then shows that it is therefore countably compact. This means that E has a limit point in Q and therefore also in R^k.

Definition 6.33: *A metric space is called* **separable** *if it contains a countable dense subset.*

Theorem 6.34: *The Euclidean space R^k is separable.*

Proof: Let S be the set of all points of the Euclidean space R^k whose coordinates are rational numbers. It follows from Theorem 5.3 and its corollary that S is countable. We shall show that S is dense in R^k. To this end, let $\boldsymbol{p}=\langle p_1, p_2, \ldots, p_k\rangle$ be any point of R^k and let $B(\boldsymbol{p}; r)$ be any ball about $\boldsymbol{p}$. The axiom of Archimedes (chapter 2) shows that for each i such that $1 \leqslant i \leqslant k$ there exists a rational number y_i such that

$$p_i-\frac{r}{k^{1/2}}<y_i<p_i+\frac{r}{k^{1/2}}$$

Hence

$$|y_i-p_i|<\frac{r}{k^{1/2}} \tag{6–16}$$

Set $\boldsymbol{y}=\langle y_1, \ldots, y_k\rangle$. Clearly $\boldsymbol{y}\in S$. It follows from inequality (6–16) that

$$|\boldsymbol{y}-\boldsymbol{p}|=\sqrt{\sum_{i=1}^{k}(y_i-p_i)^2}<r$$

Hence $\boldsymbol{y}\in B(\boldsymbol{p}; r)$. Since $B(\boldsymbol{p}; r)$ was any ball about $\boldsymbol{p}$, we have shown that *every* ball about $\boldsymbol{p}$ contains a point of S and therefore that $\boldsymbol{p}$ is an adherence point of S. Since $\boldsymbol{p}$ was any point of R^k, we conclude that S is a dense subset of R^k.

Definition 6.35: *Let E be a subset of the metric space $\langle X, d\rangle$ and let ϵ be any positive number. A set $D_\epsilon \subset X$ is called an* **ϵ-net for the set** *E if for any point $p\in E$, there exists a point $x\in D_\epsilon$ such that $d(x,p)<\epsilon$. If $E=X$, then D_ϵ is said to be an* **ϵ-dense** *subset of X.*

It is clear that D_ϵ is an ϵ-net for the set E if and only if

$$E\subset\bigcup_{x\in D_\epsilon}B(x;\epsilon)$$

For every positive integer n, the set

$$\left\{\left\langle\frac{i_1}{2^n}, \ldots, \frac{i_k}{2^n}\right\rangle \middle| i_1, \ldots, i_k=0, \pm 1, \pm 2, \ldots\right\}$$

is a $\beta(\sqrt{k}/2^{n+1})$-dense subset of R^k for any $\beta>1$. A $\beta(\sqrt{2}/2^{n+1})$-dense subset of R^2 is illustrated in figure 6–4.

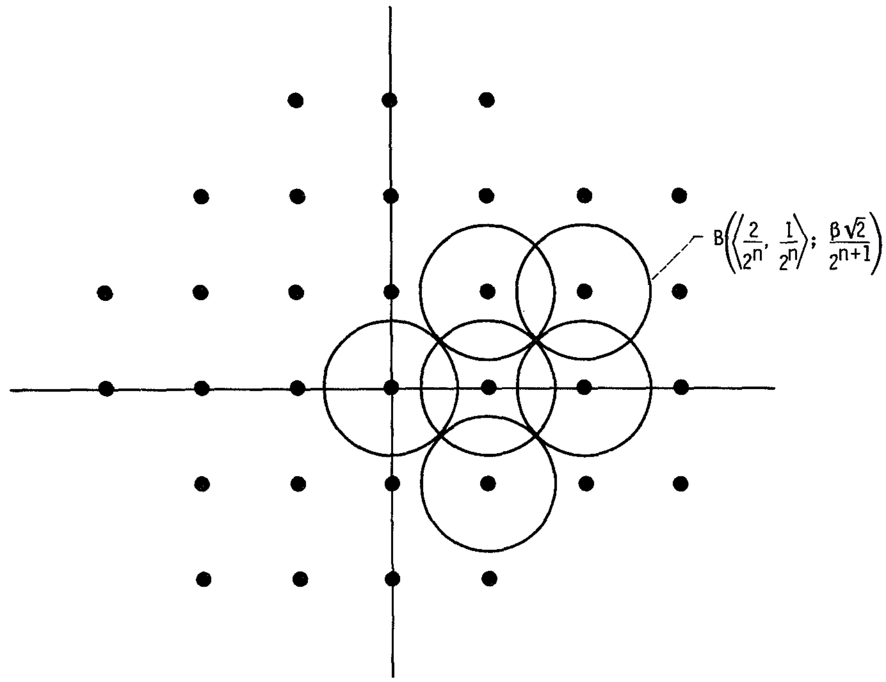


FIGURE 6–4.—$\beta(\sqrt{2}/2^{n+1})$—Dense subset of R^2 with balls of radius $\beta(\sqrt{2}/2^{n+1})$ about points of $\beta(\sqrt{2}/2^{n+1})$—net for β slightly greater than 1.

Definition 6.36: *A subset E of a metric space is said to be* **totally bounded** *if for each positive number* ϵ *there is a* **finite** ϵ*-net for E.*

It is clear that, for every k-cell in R^k, there is a finite subset of the $\beta(\sqrt{k}/2^{n+1})$-dense set just described which is a $\beta(\sqrt{k}/2^{n+1})$-net for this k-cell. This in fact shows that k-cells are totally bounded.

Since every bounded set in R^k is contained in some k-cell we see that in R^k bounded sets are totally bounded. It is not hard to see that in any metric space totally bounded sets are bounded. Thus in R^k boundedness and total boundedness are equivalent. This is not however true in general. It turns out, although we shall not be able to go into this here, that in the cases of most interest for the function spaces which we shall discuss in chapter 11, boundedness and total boundedness are not equivalent.

The material in the remainder of this chapter is not necessary for understanding the rest of the text.

Theorem 6.37: *A metric space is separable if and only if it contains a countable ϵ-dense subset for every positive number ϵ.*

Proof: If D is a countable dense subset of the metric space $\langle X, d\rangle$ then for any point $p \in X$ and for any $\epsilon > 0$ there is a point $x \in D$ such that $x \in B(p; \epsilon)$. Hence, $d(p, x) < \epsilon$ which shows that D is a countable ϵ-dense subset.

Conversely, suppose there is a countable ϵ-dense subset in X for every ϵ, and for each positive integer n let D_n be a countable $1/n$-dense subset of X. According to Theorem 5.2, $D = \bigcup_{n=1}^{\infty} D_n$ is a countable set. Now, if $p \in X$ and $\epsilon > 0$ are given, choose an integer n such that $n > 1/\epsilon$. Then there is a point $x \in D_n$ (and hence also in D) such that $d(x, p) < 1/n < \epsilon$. Hence, $x \in B(p; \epsilon)$ which shows that D is a countable dense subset of X.

Corollary: *If a metric space is totally bounded, it is also separable.*

Definition 6.38: *Let $\{G_\alpha | \alpha \in A\}$ be a family of open sets in a metric space X. If for every open set $V \subset X$ there is a subset C of A such that $V = \bigcup_{\alpha \in C} G_\alpha$, then $\{G_\alpha | \alpha \in A\}$ is said to be* **a base for the open sets of** X.

Theorem 6.39: *There is a countable base for the open sets in the metric space X if and only if X is separable.*

Proof: Let X be separable, let M be a countable dense subset of X, and let

$$\Omega = \{B(x; 1/n) | \langle x, n\rangle \in M \times J\}$$

where, as usual, J is the set of positive integers. Theorem 5.3 shows that $M \times J$ is countable. Hence Ω is a countable collection of open sets. Fix an open set $V \subset X$. Put $C = \{\langle x, n\rangle | B(x; 1/n) \subset V\}$ and $W = \bigcup_{\langle x, n\rangle \in C} B(x; 1/n)$. Evidently, $W \subset V$. Now let p be any point of V. Because V is open, there is a ball $B(p; \epsilon) \subset V$. Choose a positive integer $n > 2/\epsilon$. The fact that M is dense shows that there is a $y \in M$ such that $y \in B(p; 1/n)$. If $t \in B(y; 1/n)$, then

$$d(t, p) \leqslant d(t, y) + d(y, p) < (1/n) + (1/n) < \epsilon$$

Hence $t \in B(p; \epsilon)$, and so

$$B(y; 1/n) \subset B(p; \epsilon) \subset V$$

Since $d(p, y) < 1/n$, we see $p \in B(y; 1/n)$. The fact that $B(y; 1/n) \subset V$ and that $\langle y, n\rangle \in M \times J$ shows that $B(y; 1/n) \subset W$.

Therefore $p \in W$ and, since p was any point of V, this shows that $V \subset W$. So we conclude that $W = V$ and, since V was any open set, this shows that Ω is a countable base for the open sets of X.

Now let $\Omega = \{G_\alpha | \alpha \in A\}$ be a countable base for the open sets of X. Let M be the set which consists of exactly one point from each nonempty $G_\alpha \in \Omega$. Then M is countable. Fix any point $p \in X$ and let ϵ be any positive number. The fact that $B(p; \epsilon)$ is an open set shows that there is a subset C of A such that $B(p; \epsilon) = \bigcup_{\alpha \in C} G_\alpha$. Then, for some $\beta \in C$, $p \in G_\beta$. Hence G_β is not empty and therefore, by construction, there exists a point $y \in M$ such that $y \in G_\beta \subset B(p; \epsilon)$. Since ϵ was arbitrary, this shows that p is an adherence point of M. Because p was any point of X, this implies that M is a countable dense subset of X.

We have actually proved more than the statement of the theorem. We have shown that the countable base can always be chosen so that its members are balls. Since Theorem 6.34 shows that the Euclidean space R^1 is separable and since the balls in R^1 are segments, we conclude that *every open set in* R^1 is the countable union of segments.

The next theorem is known as the Lindelöf covering theorem.

Theorem 6.40: *If* $\{G_\alpha | \alpha \in A\}$ *is any family of open sets in the* **separable** *metric space* X, *there is a countable subset* $D \subset A$ *for which*

$$\bigcup_{\alpha \in A} G_\alpha = \bigcup_{\alpha \in D} G_\alpha$$

Proof: Choose a countable base $\Omega = \{H_n | n \in J\}$ for the open sets of X. Set

$$N = \{n \in J | (\exists \alpha \in A) \text{ for which } H_n \subset G_a\}$$

We can define a function $f\colon N \to A$ as follows: for each $n \in N$, there exists at least one $\beta \in A$ such that

$$H_n \subset G_\beta$$

Pick one such β and set

$$f(n) = \beta$$

Then $f(N)$ is evidently countable and

$$\bigcup_{\alpha \in f(N)} G_\alpha \subset \bigcup_{\alpha \in A} G_\alpha$$

If $p \in \bigcup_{\alpha \in A} G_\alpha$, there is some index $\gamma \in A$ such that $p \in G_\gamma$. Since G_γ is open, it is the union of some subfamily of Ω. Hence we can find a set $H_m \in \Omega$ such that $p \in H_m \subset G_\gamma$. This shows that $m \in N$. Thus, $p \in H_m \subset G_{f(m)} \subset \bigcup_{\alpha \in f(N)} G_\alpha$. Therefore, since p was arbitrary,

$$\bigcup_{\alpha \in A} G_\alpha \subset \bigcup_{\alpha \in f(N)} G_\alpha$$

As a result of this theorem, we are now able to show the relation between compact spaces and separable spaces.

Corollary: *Every open cover of a separable metric space has a countable subcover.*

CHAPTER 7

Limits of Sequences

The concept of limit dates back to the early Greek mathematicians. Our modern understanding of this concept, however, is principally due to Cauchy (1821) whose work, like Cantor's, grew out of a study of trigonometric series. Fréchet later incorporated Cauchy's notion of the limit of a sequence of points into his theory of metric spaces. Indeed, much of the material in this chapter is an outgrowth of this development.

Intuitively, the point p is a limit of a sequence [20] $\{p_n\}$ if p_n is "close" to p whenever n is sufficiently large. This idea is made precise in the next two definitions.

Definition 7.1(a): *Let $\{p_n\}$ be a sequence in the metric space X and let p be a point of X. The sequence $\{p_n\}$ is said to converge to p if, for every $\epsilon > 0$, there is a positive integer N such that $n \geqslant N$ implies that*

$$p_n \in B(p; \epsilon)$$

Then p is called the **limit of** *$\{p_n\}$ (we shall see that the limit is unique), and we write*

$$p_n \to p$$

or

$$\lim_{n\to\infty} p_n = p$$

If there exists no point $p \in X$ to which the sequence converges, it is said to **diverge.**

We may restate the definition of convergence in the following way.

Definition 7.1(b): *If $\{p_n\}$ is a sequence in the metric space $\langle X, d\rangle$, a necessary and sufficient condition that $p = \lim_{n\to\infty} p_n$ is that, for every $\epsilon > 0$, there*

[20] See Definition 4.15.

exists an integer N such that $n \geqslant N$ implies

$$d(p, p_n) < \epsilon$$

We shall have occasion to use both forms of the definition of convergence. Notice that Definition 7.1(b) shows $p_n \to p$ if and only if the sequence of real numbers $\{d(p_n, p)\}$ converges to zero.

It can be seen from the discussion following Definition 6.1 that, when applied to a sequence of complex numbers $\{z_n\}$, Definition 7.1 means that $\{z_n\}$ converges to a complex number ζ if and only if, for every $\epsilon > 0$, there is a positive integer N such that for all $n \geqslant N$

$$|z_n - \zeta| < \epsilon$$

or, equivalently,

$$z_n \in \{z \mid |z - \zeta| < \epsilon\}$$

Thus, for every ϵ, there is a finite number N such that all the terms of $\{z_n\}$ except the first $N-1$ must lie in the interior of a circle of radius ϵ. The concept of convergence of a sequence of complex numbers is illustrated in figure 7–1.

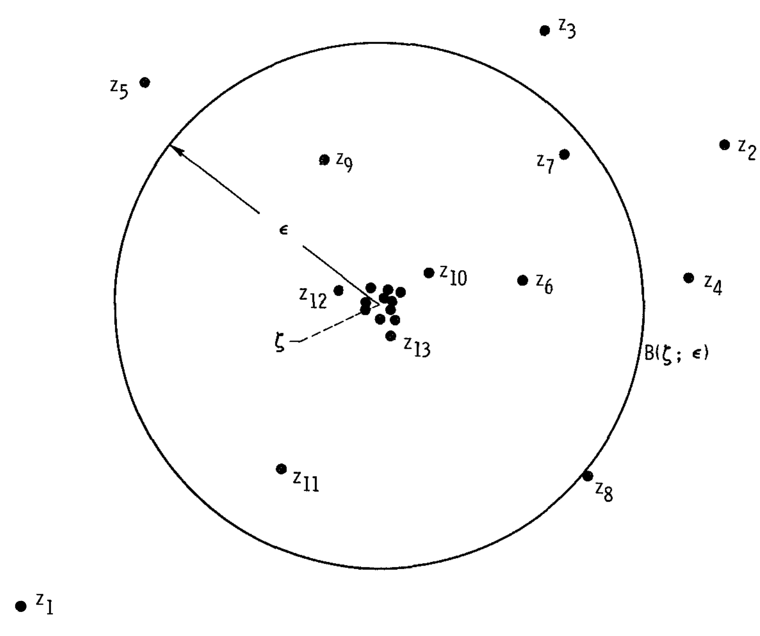


FIGURE 7–1.—Convergent sequence of complex numbers.

The definition of convergence depends strongly on the metric space to which the sequence is assumed to belong. Suppose $\langle X, d\rangle$ is a metric space, $E \subset X$ and $\{p_n\}$ is a sequence in E. Then it may happen that the sequence converges in the metric space $\langle X, d\rangle$ to a point $p \notin E$ and therefore does not converge in the metric space $\langle E, d\rangle$ (which has the same metric as $\langle X, d\rangle$). For example, if d is the usual metric for R^1, the sequence $\{1/n\}$ converges to 0 in the metric space $\langle R^1, d\rangle$. But if $E=\{x \in R^1 | x>0\}$, this sequence does not converge in the metric space $\langle E, d\rangle$. Hence, when a sequence is said to converge, the metric space in which it converges should be specified. This will only be done, however, in cases of possible ambiguity. On the other hand, if $\{p_n\}$ is a sequence of points in a subset E of a metric space $\langle X, d\rangle$ and if this sequence converges to a point $p \in E$ as a sequence in the metric space $\langle X, d\rangle$, it also converges to p as a sequence in the metric space $\langle E, d\rangle$.

It is sometimes said that the convergence of a sequence depends on the "infinite tail" of the sequence. This means that no amount of alteration of a finite number of terms of a divergent sequence can make it converge and, if a convergent sequence is changed by omitting or adding a finite number of terms, the resulting sequence still converges to the same limit as the original sequence.

Recalling Definition 4.15 we see that a sequence $\{p_n\}$ in a set X is just a function $f: J \rightarrow X$ (where J is the set of positive integers) such that, for each $n \in J, f(n)=p_n$. The range of the sequence $\{p_n\}$ is just the range of the function f—that is, the set of all the points p_n. This set may be finite or infinite. In particular, it may consist of one point p (i.e., for every n, $p_n=p$). We shall sometimes use the notation $p, p, p, \ldots$ for sequences of this type. It is clear that such sequences always converge. They are called constant sequences. In fact, any sequence that becomes a constant sequence upon the omission of a finite number of terms converges.

Definition 7.2: *If the range of a sequence is a bounded set, the sequence is said to be* **bounded.**

Let us look at a few sequences in R^1 (with the usual metric). The sequence $\{n^2\}$ is unbounded and diverges, and its range is an infinite set. On the other hand, the sequence $\{1+(-1)^n/n\}$ converges to 1 and is bounded, and its range is also an infinite set. Finally the sequence $\{(-1)^n\}$ *diverges; yet it is bounded*, and its range is a finite set.

It follows from Definition 7.1(a) that, if $\{p_n\}$ is a sequence in a metric space

X, then $p_n \to p$ if and only if, for each neighborhood V of p, there exists a positive integer N such that $p_n \in V$ as soon as $n \geq N$. This can be seen by reasoning as follows. Since all balls are neighborhoods, we see that if, for a given sequence $\{p_n\}$, we can find for each neighborhood V of a point p an integer N such that $p_n \in V$ for all $n \geq N$, then certainly Definition 7.1(a) is satisfied. On the other hand, if V is any neighborhood of p and $p_n \to p$, then, since V is open, we can always find an $\epsilon > 0$ such that $B(p; \epsilon) \subset V$. And, for this ϵ, we can find an N such that $n \geq N$ implies $p_n \in B(p; \epsilon) \subset V$. This shows that convergence could have been defined by using only the concept of open set (neighborhood), and so it is a topological property.

We shall now prove that the limit of a convergent sequence is unique.

Theorem 7.3: *If $\{p_n\}$ is a sequence in the metric space $\langle X, d\rangle$, there is at most one point p such that $\lim_{n\to\infty} p_n = p$.*

Proof: Suppose $p, q \in X$ and $p_n \to p$ and $p_n \to q$. Then, for every $\epsilon > 0$ there are positive integers N_1 and N_2 such that, whenever $n \geq N_1$, $d(p_n, p) < \epsilon/2$, and whenever $n \geq N_2$, $d(p_n, q) < \epsilon/2$. Set $N = \max\{N_1, N_2\}$. Then, for every $n \geq N$,

$$d(p, q) \leq d(p, p_n) + d(p_n, q) < \epsilon/2 + \epsilon/2 = \epsilon$$

Since ϵ is any positive number, we conclude that $d(p, q) = 0$, and this shows that $p = q$.

The simple principle used in this proof, namely that the only nonnegative number which is less than every positive number is zero, is a very useful one, and we shall encounter it frequently.

Theorem 7.4: *If E is a subset of a metric space $\langle X, d\rangle$, then the point $p \in X$ is an adherence point of E if and only if there is a sequence $\{p_n\}$ in E such that $p = \lim_{n\to\infty} p_n$.*

Proof: If $\lim_{n\to\infty} p_n = p$, then Definition 7.1(a) shows that every ball about p contains a term of the sequence $\{p_n\}$. Since every term of $\{p_n\}$ is a point of E, this proves that p is an adherence point of E. Now suppose p is an adherence point of E. Then every ball about p contains a point of E. So for each positive integer n choose a point p_n of E such that $p_n \in B(p; 1/n)$. The sequence $\{p_n\}$ so obtained evidently converges to p for, if $\epsilon > 0$ is given, we can choose a positive integer N such that $N\epsilon > 1$ and, for every $n \geq N$,

$$p_n \in B(p; 1/n) \subset B(p; \epsilon)$$

The statement of Theorem 7.4 is sometimes used as the definition of adherence point. The remarks immediately following Definition 6.7 show that a closed set contains all its adherence points. Thus we have the next corollary.

Corollary: *Every convergent sequence in a metric space X with terms in a closed subset E of X converges to a point of E.*

In the following theorem, as in most theorems about boundedness, we make explicit use of the fact that the distance between any two points of the metric space is finite. It therefore does not hold in metric spaces with possibly infinite metrics.

Theorem 7.5: *Every convergent sequence* $\{p_n\}$ *in a metric space* $\langle X, d\rangle$ *is bounded.*

Proof: Since there is a point p such that $p_n \to p$, we can find a positive integer N such that $n \geqslant N$ implies

$$p_n \in B(p; 1)$$

Now set

$$\rho = \max\{1, d(p_1, p), \ldots, d(p_N, p)\}$$

Since ρ is the least upper bound of a finite set of finite numbers, it is also finite. Then, for every positive integer n

$$p_n \in B(p; \rho)$$

That is, if E is the range of $\{p_n\}$, then

$$E \subset B(p; \rho)$$

Now for any two points $x, y \in E$

$$d(x, y) \leqslant d(x, p) + d(p, y) < \rho + \rho = 2\rho$$

Hence, 2ρ is an upper bound of the set $\{d(x, y) | x, y \in E\}$. Therefore the least upper bound of this set $d(E)$ must be less than or equal to the finite number 2ρ. This shows that the sequence $\{p_n\}$ is bounded.

Theorem 7.6: *If* $\{p_n\}$ *is a sequence in the metric space X, then* $\{p_n\}$ *converges to a point* $p \in X$ *if and only if every ball about p contains all but a finite number of terms of* $\{p_n\}$.

355-525 O-70—8

Proof: Suppose $\{p_n\}$ converges to p. Let $B(p;\epsilon)$ be any ball about p. Then there is a positive integer N such that the only terms of $\{p_n\}$ possibly not in $B(p;\epsilon)$ are $p_1, p_2, \ldots, p_{N-1}$.

Conversely, suppose that every ball about p contains all but finitely many terms of $\{p_n\}$ and $\epsilon>0$ is given. Set $N'=\max\{n \in J \mid p_n \notin B(p;\epsilon)\}$, and put $N=N'+1$. Then $p_n \in B(p;\epsilon)$ for all $n \geqslant N$.

In order to study the relation between convergence and algebraic operations, we turn to the normed linear spaces introduced in chapter 3. As we mentioned in chapter 6, the metric in these spaces will always be given in terms of the norm by equation (6–1).

The next theorem shows the relation between addition and scalar multiplication on the one hand and convergence on the other.

Theorem 7.7: *Suppose $\{u_n\}$ and $\{v_n\}$ are sequences in the normed linear space V and $\{\alpha_n\}$ is a sequence of scalars (i.e., real or complex numbers depending on whether V is a real or complex vector space).* If

$$\lim_{n\to\infty} u_n = u$$

$$\lim_{n\to\infty} v_n = v$$

$$\lim_{n\to\infty} \alpha_n = \alpha$$

then,

(a) $\lim_{n\to\infty} (u_n+v_n)=u+v$

(b) $\lim_{n\to\infty} (\beta u_n)=\beta u$ *for every scalar* β

(c) $\lim_{n\to\infty} (w+u_n)=w+u$ *for every* $w \in V$

(d) $\lim_{n\to\infty} (\alpha_n u_n)=\alpha u$

Proof: Part (a). Given $\epsilon>0$, there exist positive integers N_1 and N_2 such that $n \leqslant N_1$ implies $\|u_n-u\|<\epsilon/2$ and $n \geqslant N_2$ implies $\|v_n-v\|<\epsilon/2$. Set $N=\max\{N_1, N_2\}$. Then $n \geqslant N$ implies $\|(u_n+v_n)-(u+v)\| \leqslant \|u_n-u\|+\|v_n-v\|<\epsilon$.

Part (b). First suppose $\beta \neq 0$. Then given $\epsilon>0$, there exists positive integer N such that $n \geqslant N$ implies $\|u_n-u\|<\epsilon/|\beta|$. Hence

$$\|\beta u_n-\beta u\|=\|\beta(u_n-u)\|=|\beta|\,\|u_n-u\|<\epsilon$$

If $\beta=0$, the result is trivially true.

Part (c). Trivial.

Part (d). Suppose $\alpha \neq 0$ and $u \neq 0$. Then, by (N1) of Definition 3.4, $\|u\| \neq 0$. We shall use the identity

$$\alpha_n u_n - \alpha u = (\alpha_n - \alpha)(u_n - u) + \alpha(u_n - u) + (\alpha_n - \alpha)u$$

Given $\epsilon > 0$, there are positive integers N_1 and N_2 such that $n \geq N_1$ implies $|\alpha_n - \alpha| < \min\{\sqrt{\epsilon}/\sqrt{3}, \epsilon/(3\|u\|)\}$ and $n \geq N_2$ implies $\|u_n - u\| < \min\{\sqrt{\epsilon}/\sqrt{3}, \epsilon/(3|\alpha|)\}$. Set $N = \max\{N_1, N_2\}$. Then $n \geq N$ implies

$$\begin{aligned}\|\alpha_n u_n - \alpha u\| &\leq \|(\alpha_n - \alpha)(u_n - u)\| + \|\alpha(u_n - u)\| + \|(\alpha_n - \alpha)u\| \\ &= |\alpha_n - \alpha|\,\|u_n - u\| + |\alpha|\,\|u_n - u\| + |\alpha_n - \alpha|\,\|u\| \\ &< \frac{\sqrt{\epsilon}}{\sqrt{3}}\frac{\sqrt{\epsilon}}{\sqrt{3}} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon\end{aligned}$$

The proofs for the cases when $\alpha = 0$ or $u = 0$ follow easily from the fact that a convergent sequence must be bounded and from part (b).

Since the complex numbers themselves form a complex normed linear space, it is an immediate consequence of part (d) of this theorem that, if $\{\alpha_n\}$ and $\{\beta_n\}$ are sequences of complex numbers such that $\alpha_n \to \alpha$ and $\beta_n \to \beta$, then

$$\alpha_n \beta_n \to \alpha\beta$$

In the case of complex numbers, we can moreover deduce a relation between division and convergence.

Theorem 7.8: *Suppose $\{s_n\}$ is a sequence of complex numbers such that $\lim_{n\to\infty} s_n = s$, $s_n \neq 0$ for $n = 1, 2, 3, \ldots$ and $s \neq 0$. Then $\lim_{n\to\infty} 1/s_n = 1/s$.*

Proof: Since $|s| \neq 0$, let us choose m so large that $|s_n - s| < |s|/2$ whenever $n \geq m$. Since

$$\tfrac{1}{2}|s| > |s - s_n| \geq |s| - |s_n|$$

it follows that

$$|s_n| > \tfrac{1}{2}|s| \qquad n \geq m$$

Given $\epsilon > 0$, there is an integer $N > m$ such that $n \geq N$ implies

$$|s_n - s| < \tfrac{1}{2}|s|^2\epsilon$$

Hence, for $n \geqslant N$,

$$\left|\frac{1}{s_n}-\frac{1}{s}\right|=\left|\frac{s_n-s}{s_n s}\right|=\frac{|s_n-s|}{|s_n||s|}<\frac{2}{|s|^2}|s_n-s|<\epsilon$$

In the Euclidean spaces R^k, there is an intimate relation between the convergence of sequences of vectors and the convergence of their components.

Theorem 7.9: (a) *Let $\langle X, d\rangle$ and $\langle Y, \delta\rangle$ be metric spaces. Let $\{p_n\}$ be a sequence in X and $\{q_n\}$ be a sequence in Y. Then the sequence $\{\langle p_n, q_n\rangle\}$ converges in $\langle X\times Y, d_\times\rangle$ the direct product of $\langle X, d\rangle$ and $\langle Y, \delta\rangle$, if and only if $\{p_n\}$ converges in X and $\{q_n\}$ converges in Y. In addition*

$$\lim_{n\to\infty} \langle p_n, q_n\rangle = \langle \lim_{n\to\infty} p_n, \lim_{n\to\infty} q_n\rangle$$

whenever the limit exists.

(b) *Let $\{\boldsymbol{x}_n\}$, with $\boldsymbol{x}_n=\langle x_{1,n}, \ldots, x_{k,n}\rangle$ for every positive integer n, be a sequence in the Euclidean space R^k. Then $\{\boldsymbol{x}_n\}$ converges to a point $\boldsymbol{x}=\langle x_1, \ldots, x_k\rangle$ of R^k if and only if*

$$\lim_{n\to\infty} x_{j,n}=x_j \qquad 1\leqslant j\leqslant k$$

Proof: Part (a). First suppose that $p_n\to p$ in X and $q_n\to q$ in Y. Let $\epsilon>0$ be given. Then there exist integers N_1 and N_2 such that for all $n\geqslant N_1$ and all $m\geqslant N_2$, $d(p_n, p)<\epsilon$ and $d(q_m, q)<\epsilon$. Set $N=\max\{N_1, N_2\}$. Then for each $n\geqslant N$ we have

$$d_\times(\langle p_n, q_n\rangle, \langle p, q\rangle)=\max\{d(p_n, p), \delta(q_n, q)\}<\epsilon$$

Hence,

$$\lim_{n\to\infty} \langle p_n, q_n\rangle=\langle p, q\rangle$$

Conversely, suppose that $\lim_{n\to\infty} \langle p_n, q_n\rangle=\langle p, q\rangle$ in $\langle X\times Y, d_\times\rangle$. Then there exists an integer N such that for all $n\geqslant N$

$$\max\{d(p_n, p), \delta(q_n, q)\}<\epsilon$$

Hence, for all $n\geqslant N$, $d(p_n, p)<\epsilon$ and $d(q_n, q)<\epsilon$; that is, $p_n\to p$ and $q_n\to q$.

Part (b). Suppose first that $\lim_{n\to\infty} \boldsymbol{x}_n=\boldsymbol{x}$. It is easily seen from the definition of the norm in R^k that

$$|x_{j,n}-x_j| \leqslant |\boldsymbol{x}_n-\boldsymbol{x}| \qquad \text{for } 1 \leqslant j \leqslant k$$

Now for every $\epsilon > 0$ there is an integer N such that $n \geqslant N$ implies

$$|x_{j,n}-x_j| \leqslant |\boldsymbol{x}_n-\boldsymbol{x}| < \epsilon \qquad 1 \leqslant j \leqslant k$$

which shows

$$\lim_{n\to\infty} x_{j,n} = x_j$$

On the other hand, if

$$\lim_{n\to\infty} x_{j,n} = x_j \qquad \text{for } 1 \leqslant j \leqslant k$$

then for every $\epsilon > 0$ there is a positive integer N such that $n \geqslant N$ implies

$$|x_{j,n}-x_j| < \frac{\epsilon}{\sqrt{k}} \qquad 1 \leqslant j \leqslant k$$

Therefore, whenever $n \geqslant N$,

$$|\boldsymbol{x}_n-\boldsymbol{x}| = \left(\sum_{j=1}^{k} |x_{j,n}-x_j|^2\right)^{1/2} < \epsilon$$

and this shows that

$$\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{x}$$

Since the absolute value of a complex number is the same as the norm of the corresponding vector in the Euclidean space R^2, the following is an immediate corollary of this theorem.

Corollary 1: *If $\{z_n\}$ is a sequence of complex numbers, then $\{z_n\}$ converges to the complex number ζ if and only if*

$$\lim_{n\to\infty} \mathcal{R}e\ z_n = \mathcal{R}e\ \zeta$$

$$\lim_{n\to\infty} \mathcal{I}m\ z_n = \mathcal{I}m\ \zeta$$

The next corollary is an immediate consequence of Theorems 7.7 and 7.9.

Corollary 2: *Suppose $\{\boldsymbol{x}_n\}$ and $\{\boldsymbol{y}_n\}$ are sequences in R^k and $\boldsymbol{x}_n \to \boldsymbol{x}$* and *$\boldsymbol{y}_n \to \boldsymbol{y}$. Then*

$$\boldsymbol{x}_n \cdot \boldsymbol{y}_n \to \boldsymbol{x} \cdot \boldsymbol{y}$$

It follows from Theorem 7.5 and the example following Definition 7.2 that convergent sequences are indeed bounded but, on the other hand, bounded sequences need not converge. Since it is very desirable to be able to assert that the limit of a sequence exists without knowing what it is and since boundedness of a given sequence is often easy to establish, it is fortunate that there is at least one important case in which boundedness is equivalent to convergence. This occurs in the case of *monotone* sequences which we now define.

Definition 7.10: *A sequence of real numbers $\{s_n\}$ is said to be* **increasing** *or* **monotonically increasing** *if, for every pair of integers m and n, $m > n$ implies $s_m \geqslant s_n$. It is said to be* **decreasing** *or* **monotonically decreasing** *if, for every pair of integers m and n, $m > n$ implies $s_m \leqslant s_n$. We call a sequence* **monotone** *or* **monotonic** *if it is either increasing or decreasing and, when only the inequality signs hold in the definition (e.g., $s_m > s_n$), the adverb* **strictly** *is used to indicate this.*

Theorem 7.11: *A monotone sequence $\{s_n\}$ converges if and only if it is bounded.*

Proof: We shall give the proof only for the case of increasing sequences since the proof is completely analogous in the other case. Hence suppose $s_m \geqslant s_n$ whenever $m > n$, and let E be the range of $\{s_n\}$. Since E is bounded, its least upper bound s exists. Thus, for every integer n,

$$s_n \leqslant s$$

Let $\epsilon > 0$ be given. Since $s - \epsilon$ is not an upper bound of E, there exists a positive integer N such that

$$s - \epsilon < s_N \leqslant s$$

It follows from the definition of an increasing sequence that $n \geqslant N$ implies $s_n \geqslant s_N$, and since s is an upper bound of E, we conclude that

$$s - \epsilon < s_n \leqslant s$$

for all $n \geqslant N$ which shows that $\{s_n\}$ converges to s. The converse is already given by Theorem 7.5.

We have in fact proved slightly more than the statement of the theorem; that is, we have proved that the series converges *to* the least upper bound of its range.

There is a natural way of extending the definition of convergence of sequences in the Euclidean space R^1. The set R^1 of finite real numbers with the usual metric is a metric space (with a finite metric) embedded in the set of *extended* real numbers. The set of extended real numbers however cannot be made into a metric space with a finite metric by simply extending the usual metric for R^1. We might however still attempt to extend the concept of ball to the set of extended real numbers. Very crudely speaking, we may think of the real numbers as lying along a line with the points $+\infty$ and $-\infty$ at its ends. Then all the finite numbers would be to the left of the point $+\infty$. Now since balls in R^1 (with the usual metric) are segments and since each ball about a point p also contains p, it is natural to take the "balls" about the points $+\infty$ and $-\infty$ to be the half-open intervals $(a, +\infty]$ and $[-\infty, b)$, respectively, where a and b are finite real numbers. Now if we wish to allow the points $+\infty$ and $-\infty$ to be possible limits of sequences of *finite* real numbers, we can give a definition of convergence of these sequences similar to Definition 7.1(a) but which uses the extended concept of ball. In this way we allow sequences of points of the Euclidean space R^1 to "converge" to points which lie outside this space. We can thus extend the concept of convergence in a natural way to certain types of divergent sequences. All sequences which are convergent (in the sense of Definition 7.1(a) or equivalently Definition 7.1(b)) would still be "convergent" in this extended sense. It must be emphasized however that this does not change the definitions of convergence and divergence already given. *Only those sequences of points of the Euclidean space R^1 which converge in the sense of Definition* 7.1(a) *or Definition* 7.1(b) *are said to converge.* The sequences which would "converge" in this extended sense but not in the sense of Definition 7.1(b) will still be called divergent sequences. In view of this discussion we make the following definition.

Definition 7.12: *If $\{s_n\}$ is a sequence of finite real numbers such that, for every finite number M, there is a positive integer N such that $n \geqslant N$ implies $s_n > M$, we say $\{s_n\}$* **diverges** *to $+\infty$ and write*

$$s_n \to +\infty$$

If, on the other hand, for every finite number M, there is a positive integer N such that $n \geqslant N$ implies $s_n < M$, we say $\{s_n\}$ **diverges** *to $-\infty$ and write*

$$s_n \to -\infty$$

In either case, we say that $\{s_n\}$ **converges improperly** *or is* **improperly convergent.**

Notice that we now use the symbol $\to$ *for certain types of divergent sequences, as well as for convergent sequences, but the notation* $\lim_{n\to\infty} p_n = p$ *is used only for convergent sequences.*

Theorem 7.13: *If* E *is a nonempty set of finite real numbers, then there is a sequence* $\{s_n\}$ *in* E *such that* $s_n \to \operatorname{lub} E$. *Of course a similar result holds for* glb E.

Proof: Set $t = \operatorname{lub} E$. Then t is either finite or $+\infty$. Assume first that t is finite. Let $\epsilon > 0$ be given. Then $t-\epsilon$ is not an upper bound of E. Hence we can find an $s \in E$ such that

$$t-\epsilon < s \leqslant t < t+\epsilon$$

So we conclude that every ball about t contains a point of E which means that t is an adherence point of E. The statement of the theorem now follows from Theorem 7.4.

If $t = +\infty$, then E is not bounded above. Hence, for every positive integer n, choose a point $s_n \in E$ such that $s_n > n$. In this manner, we obtain a sequence $\{s_n\}$ in E such that, if M is any finite number and N is the smallest positive integer larger than or equal to M, then $s_n > M$ for all $n \geqslant N$; that is, $s_n \to +\infty = t$.

We might point out that, if $\{s_n\}$ is an increasing sequence which is not bounded above, then, for every finite number M we can find a member, say s_N, of the sequence such that $s_N > M$. So we conclude from Definition 7.10 that $s_n > M$ for every $n \geqslant N$; that is, $s_n \to +\infty$. Since a similar conclusion holds for decreasing sequences, we arrive at the following result. *A monotone sequence either converges or it converges improperly.*

We have in a sense associated limits with certain types of divergent sequences of real numbers. We shall now see that in the more general setting of metric spaces there is under certain conditions another method of associating limits with sequences even though they diverge. We first introduce the notion of subsequence.

Definition 7.14: *If* $\{p_{n_k}\}$ *is a sequence of points in a set* X *and* $n_1, n_2, n_3, \ldots$ *is a* **strictly increasing** *sequence of positive integers, then the sequence* $\{p_{n_k}\}$ *is called a* **subsequence** *of* $\{p_n\}$.

Thus, for example, let the sequence $\{p_n\}$ be defined by

$$p_n = (\alpha + 1/n)^{1/2} \qquad \text{for } n = 1, 2, 3, \ldots$$

and let the strictly increasing sequence of positive integers $\{n_k\}$ be defined by

$$n_k = 2k \qquad \text{for } k = 1, 2, 3, \ldots$$

Then the sequence $\{p_{n_k}\}$ defined by

$$p_{n_k} = (\alpha + 1/2k)^{1/2} \qquad \text{for } k = 1, 2, 3, \ldots$$

is a subsequence of $\{p_n\}$.

If every ball about a point p of a metric space contains all but finitely many terms of the sequence $\{p_n\}$, the same must be true for any subsequence $\{p_{n_k}\}$ of $\{p_n\}$. Thus Theorem 7.6 and the fact that $\{p_n\}$ must be a subsequence of itself show that a sequence $\{p_n\}$ in a metric space X converges to a point $p \in X$ *if and only if every subsequence of* $\{p_n\}$ converges to p.

It is clear that, if $\{p_{n_k}\}$ is a subsequence of $\{p_n\}$ and $\{p_{n_{k_j}}\}$ is a subsequence of $\{p_{n_k}\}$, then $\{p_{n_{k_j}}\}$ is also a subsequence of $\{p_n\}$.

Evidently if E is the range of a sequence $\{p_n\}$ and $\{p_{n_k}\}$ is a subsequence of $\{p_n\}$, then $\{p_{n_k}\}$ is a sequence in E. However, it is *not* true in general that every sequence in E is a subsequence of $\{p_n\}$. For example, the constant sequence $1, 1, 1, \ldots$ is a sequence of points in the range of the sequence $\{1/n\}$ which is not a subsequence of $\{1/n\}$.

The following definition now shows how we can associate limits with divergent sequences.

Definition 7.15: *Let* $\{p_{n_k}\}$ *be a subsequence of the sequence* $\{p_n\}$ *in the metric space* X. *If* $\{p_{n_k}\}$ *converges, its limit is called a* **subsequential limit** *of* $\{p_n\}$.

The remarks following Definition 7.14 combined with Theorem 7.4 show that if p is a subsequential limit of a sequence $\{p_n\}$ then p is an adherence point of the range E of $\{p_n\}$. Thus if F is the set of all subsequential limits of $\{p_n\}$ we have

$$F \subset \bar{E}$$

However, the reverse inclusion does not in general hold. The next theorem shows that we *can* assert

$$E' \subset F \subset \bar{E}$$

Theorem 7.16: *Let E and F be, respectively, the range of the sequence $\{p_n\}$ and the set of all subsequential limit points of the sequence $\{p_n\}$ in the metric space $\langle X, d\rangle$. Then (a) every limit point of E belongs to F and (b) F is closed.*

Proof: Part (a). If p is a limit point of E, Theorem 6.9(a) shows that every ball about p contains infinitely many terms of $\{p_n\}$. We construct a subsequence of $\{p_n\}$ inductively as follows: Choose n_1 to be any of the infinitely many indices n for which $d(p_n, p) < 1$. Having chosen $n_1 < n_2 < \ldots < n_{k-1}$ with $d(p, p_i) < 1/i$ for $1 \leqslant i \leqslant k-1$, choose n_k from the infinitely many n for which $d(p, p_n) < 1/k$ to be the smallest such n which is larger than n_{k-1}. Thus $\{p_{n_k}\}$ is a subsequence of $\{p_n\}$ such that, for any $\epsilon > 0$, $d(p, p_{n_k}) < \epsilon$ for every $k \geqslant 1/\epsilon$. Hence $p_{n_k} \to p$. This shows that $p \in F$.

Part (b). If p is a limit point of F, then, for every $\epsilon > 0$ there is a point $x \in F$ such that

$$0 < d(x, p) < \epsilon/2$$

Since x is a subsequential limit of $\{p_n\}$, there is a point $p_k \in E$ such that

$$d(x, p_k) < d(x, p)$$

From these two inequalities we see that $p_k \neq p$ and

$$d(p_k, p) \leqslant d(p_k, x) + d(x, p) < 2d(x, p) < \epsilon$$

Since ϵ was arbitrary, we conclude that every ball about p contains a point of $E - \{p\}$. Hence p is a limit point of E and part (a) shows that $p \in F$. Since p was any limit point of F, this shows that F is closed.

Definition 7.17: *If every sequence in a set E has a subsequence which converges to a point of E, then E is said to be* **sequentially compact.**

Theorem 7.18: *Every countably compact set K is sequentially compact.*

Proof: Let E be the range of a sequence $\{p_n\}$ with values in K. If E is a finite set, then the proof is easy for there must be at least one point of $E \subset K$, say p, and a strictly increasing sequence of integers $\{n_i\}$ such that

$$p_{n_1} = p_{n_2} = \ . \ . \ . = p$$

This subsequence evidently converges to $p \in K$.

Hence suppose E is infinite. Then E has a limit point $p \in K$. Hence Theorem 7.16 shows that p is a subsequential limit of $\{p_n\}$. That is, there exists a subsequence $\{p_{n_i}\}$ of $\{p_n\}$ which converges to $p \in K$.

Theorem 7.19: *If every sequence in a subset E of a metric space $\langle X, d\rangle$ has a subsequence which is convergent in X, then for every positive number ϵ there is a finite ϵ-net, D_ϵ, for E such that $D_\epsilon \subset E$.*

Proof: Suppose that every sequence in E has a convergent subsequence and that, for some positive number ϵ, there is no finite ϵ-net for E which is also a subset of E. Then, if $p_1 \in E$, there must be a point $p_2 \in E$ such that $d(p_1, p_2) \geqslant \epsilon$ because, otherwise, $\{p_1\}$, which is a subset of E, would be a finite ϵ-net for E contrary to the assumption. Having chosen $p_1, p_2, \ . \ . \ ., p_{n-1} \in E$ such that $d(p_i, p_j) \geqslant \epsilon$ for $1 \leqslant i < j \leqslant n-1$, we can choose a point $p_n \in E$ such that $d(p_i, p_n) \geqslant \epsilon$ for $1 \leqslant i \leqslant n-1$ because, otherwise, $\{p_i, \ . \ . \ ., p_{n-1}\}$, which is a subset of E, would be a finite ϵ-net for E contrary to the assumption. In this manner, we construct a sequence $\{p_n\}$ in E such that $d(p_n, p_m) \geqslant \epsilon$ for $m \neq n$.

Now suppose $\{p_{n_k}\}$ is any subsequence of $\{p_n\}$. Since $\{n_k\}$ is a strictly increasing sequence, it follows that $n_k \neq n_j$ if $k \neq j$. Hence

$$d(p_{n_k}, p_{n_j}) \geqslant \epsilon \qquad k \neq j$$

If p is any point of X and if k and j are any two different positive integers, we conclude that

$$\epsilon \leqslant d(p_{n_k}, p_{n_j}) \leqslant d(p_{n_k}, p) + d(p, p_{n_j})$$

Hence, there is no positive integer I such that $d(p_{n_i}, p) < \epsilon/3$ for every $i \geqslant I$ for, otherwise, we could conclude that $\epsilon \leqslant 2\epsilon/3$ which is clearly impossible since ϵ is a positive finite number. Thus $\{p_{n_k}\}$ cannot converge. Since $\{p_{n_k}\}$ was any subsequence of $\{p_n\}$, this shows that there must be at least one sequence in E which contains no convergent subsequence. Since this is a contradiction, the theorem is proved.

Corollary: *If a subset E of a metric space is sequentially compact, then it is totally bounded.*

Convergence of sequences is actually an exceptional occurrence. In view of this fact, the following theorem may seem rather surprising.

Theorem 7.20: *If K is a subset of the metric space $\langle X, d\rangle$, the following statements are equivalent:*

(a) *K is compact.*

(b) *K is countably compact.*

(c) *K is sequentially compact.*

Proof: Theorem 6.25 shows that (a) implies (b) and Theorem 7.18 shows that (b) implies (c). We will now show that (c) implies (a). If $K=\emptyset$ there is nothing to prove. Hence assume $K\neq\emptyset$ and let $\{G_\alpha|\alpha\in A\}$ be any open cover of K. Now, for any $x\in K$, there exists an $\alpha\in A$ such that $x\in G_\alpha$. Since G_α is open, there is a positive number δ such that $B(x;\delta)\subset G_\alpha$.

Let

$$\rho(x)=\text{lub }\{r|(\exists\alpha\in A)\text{ and }B(x;r)\subset G_\alpha\}$$

Thus, roughly speaking, $\rho(x)$ is the radius of the largest ball about x that can fit into any of the open sets G_α. Since $\rho(x)$ is the least upper bound of a nonempty set of strictly positive numbers, it is clear that

$$0<\rho(x)$$

This shows that, for every $x\in K$, $\rho(x)$ is a strictly positive number or $+\infty$. Now set

$$\rho_0=\underset{x\in K}{\text{glb}}\ \rho(x)$$

It is clear that, being the greatest lower bound of a set of extended real numbers which are all larger than zero, $\rho_0\geqslant 0$. We shall now proceed to show that ρ_0 is strictly greater than zero. Once this is done, the remainder of the proof follows easily.

Suppose $\rho_0\neq+\infty$. Then the set $\{\rho(x)|x\in K\}-\{+\infty\}$ is a nonempty set of finite real numbers and ρ_0 is the greatest lower bound of this set also. Hence Theorem 7.13 shows that there is a sequence of points from K, say $\{p_n\}$, such that $\lim_{n\to\infty}\rho(p_n)=\rho_0$ (since ρ_0 is finite). By hypothesis, there is a subsequence $\{p_{n_k}\}$ of $\{p_n\}$ and a point $p\in K$ such that

$$\lim_{k\to\infty}p_{n_k}=p$$

Since $\{G_\alpha|\alpha\in A\}$ is a cover of K, there exists an $\gamma\in A$ such that $p\in G_\gamma$ and since G_γ is open there exists a $\delta > 0$ such that

$$B(p;\ \delta) \subset G_\gamma$$

Since

$$\lim_{k\to\infty} p_{n_k} = p$$

there is a positive integer N such that for all $k \geqslant N$

$$d(p_{n_k}, p) < \delta/2$$

Now if for any $k \geqslant N$

$$y \in B(p_{n_k};\ \delta/2)$$

then

$$d(p, y) \leqslant d(p, p_{n_k}) + d(p_{n_k}, y) < \delta/2 + \delta/2 = \delta$$

Therefore $y \in B(p;\ \delta)$. Thus, for every $k \geqslant N$,

$$B(p_{n_k};\ \delta/2) \subset B(p;\ \delta) \subset G_\gamma$$

This shows that, for every $k \geqslant N$,

$$\rho(p_{n_k}) \geqslant \delta/2 > 0$$

Now, since $\lim_{n\to\infty} \rho(p_n) = \rho_0$, the remarks following Definition 7.14 show that

$$\lim_{k\to\infty} \rho(p_{n_k}) = \rho_0$$

Thus, for every $\epsilon_1 > 0$, there is an integer N_1 such that for $k \geqslant N_1$

$$\rho_0 - \epsilon_1 < \rho(p_{n_k}) < \rho_0 + \epsilon_1$$

and so, for $k \geqslant \max\ \{N, N_1\}$

$$\delta/2 < \rho_0 + \epsilon_1$$

Since this must be true for every $\epsilon_1 > 0$, we conclude that

$$0 < \delta/2 \leqslant \rho_0$$

Thus we have shown that if $\rho_0 \neq +\infty$ then ρ_0 is a strictly positive number. In any case, there is a positive number ϵ such that

$$0 < \epsilon < \rho_0$$

Because K is sequentially compact, Theorem 7.19 shows that for this ϵ

there is a finite ϵ-net, D_ϵ, such that $D_\epsilon \subset K$. That is, there is a finite number of points of K, say $y_1, \ldots, y_s$, such that

$$K \subset B(y_1; \epsilon) \cup \ldots \cup B(y_s; \epsilon)$$

Now, since $\rho_0 > \epsilon$, it is also true that for $1 \leqslant i \leqslant s$

$$\rho(y_i) \geqslant \rho_0 > \epsilon$$

and this shows that for each $i = 1, 2, 3, \ldots, s$ we can find an $\alpha_i \in A$ such that

$$B(y_i; \epsilon) \subset G_{\alpha_i}$$

Hence

$$K \subset B(y_1; \epsilon) \cup \ldots \cup B(y_s; \epsilon) \subset G_{\alpha_1} \cup \ldots \cup G_{\alpha_s}$$

Thus $\{G_{\alpha_i} | 1 \leqslant i \leqslant s\}$ is a finite subcover of K, and, since $\{G_\alpha | \alpha \in A\}$ was any open cover of K, this proves that K is compact.

Corollary: *If a sequence in the Euclidean space R^k is bounded, it must contain a convergent subsequence.*

Proof: Since the range A of the sequence is a bounded set, there is a k-cell Q which contains A. But Q is compact. The theorem now shows that the sequence contains a subsequence which converges to a point of Q and hence to a point of R^k.

There is yet another way of associating limits with sequences of real numbers even if they diverge, which, as will be seen, is really a combination of the preceding two ways.

Definition 7.21: *Let $\{s_n\}$ be a sequence of (finite) real numbers. Put*

$$\bar{t}_k = \sup \{s_n | n \geqslant k\}$$

$$s^* = \inf_k \bar{t}_k$$

$$\underline{t}_k = \inf \{s_n | n \geqslant k\}$$

$$s_* = \sup_k \underline{t}_k$$

The numbers s^* and s_* exist in the extended real number system (see *remarks following Definition 2.3) and are called the* **upper** *and* **lower limits** *of* $\{s_n\}$*, respectively. The following notation is used:*

$$s^* = \limsup_{n\to\infty} s_n$$

$$s_* = \liminf_{n\to\infty} s_n$$

The abbreviated notations, such as $\limsup_n$ *or* $\limsup$*, are also commonly used for* $\limsup_{n\to\infty}$ *and similarly for* $\liminf_{n\to\infty}$*. The classical notations* $\overline{\lim}$ *and* $\underline{\lim}$ *for* $\limsup$ *and* $\liminf$*, respectively, are also still in use. The numbers* s^* *and* s_* *are also referred to as the* **superior** *and* **inferior limits** *of* $\{s_n\}$.

From equations (2–1) it is clear that

$$-\bar{t}_k = \inf\{-s_n | n \geqslant k\}$$

and

$$-s^* = \sup_k (-\bar{t}_k)$$

It follows from this that

$$-s^* = \liminf_{n\to\infty} (-s_n) \tag{7–1}$$

and similarly

$$-s_* = \limsup_{n\to\infty} (-s_n) \tag{7–2}$$

Because of these relations we shall often be able to prove theorems about the inferior limit of a sequence immediately from the corresponding theorems about the superior limit of a sequence.

Theorem 7.22: *For any sequence of real numbers* $\{s_n\}$,

$$\liminf_{n\to\infty} s_n \leqslant \limsup_{n\to\infty} s_n$$

Proof: Set $T_k = \{s_n | n \geqslant k\}$, $\bar{t}_k = \sup T_k$, and $\underline{t}_k = \inf T_k$. Then $\{\bar{t}_k\}$ is a monotonically decreasing sequence since $T_{k+1} \subset T_k$ implies $\sup T_{k+1} \leqslant \sup T_k$. In the same way we see that $\{\underline{t}_k\}$ is a monotonically increasing sequence. It is also clear that for any k

$$\underline{t}_k \leqslant \bar{t}_k$$

So for any two positive integers i and j

$$\underline{t}_i \leqslant \underline{t}_{i+j} \leqslant \bar{t}_{i+j} \leqslant \bar{t}_j$$

Thus, for each i,

$$\underline{t}_i \leqslant \inf_j \bar{t}_j = \limsup_{n\to\infty} s_n$$

and so we see

$$\liminf_{n\to\infty} s_n = \sup_i \underline{t}_i \leqslant \limsup_{n\to\infty} s_n$$

We shall now prove two theorems which are very useful for finding properties of the superior and inferior limits.

Theorem 7.23: *If $\{s_n\}$ is a sequence of real numbers and a is an (extended) real number, then*

(a) $a < \limsup_{n\to\infty} s_n$ *implies* $a < s_n$ *for infinitely many* n.

(b) $a > \limsup_{n\to\infty} s_n$ *implies* $a > s_n$ *for all but finitely many* n.

(c) $a > \liminf_{n\to\infty} s_n$ *implies* $a > s_n$ *for infinitely many* n.

(d) $a < \liminf_{n\to\infty} s_n$ *implies* $a < s_n$ *for all but finitely many* n.

Proof: Let $\bar{t}_k$ and s^* be as in Definition 7.21.

Part (a). Suppose $a < s^*$. Then, since $s^* \leqslant \bar{t}_k$ for every k, it follows that $a < \bar{t}_k$ for every k; that is, a is not an upper bound of $\{s_n | n \geqslant k\}$. Hence there is an integer $n \geqslant k$ such that $a < s_n$. Thus, for every k there exists an $n \geqslant k$ such that $a < s_n$. If there were only finitely many n, say $n_1, n_2, \ldots, n_j$ for which $a < s_n$, we could choose $k > \max\{n_1, n_2, \ldots, n_j\}$. But since we can find an $n \geqslant k$ for which $a < s_n$, this is impossible. Therefore there must be infinitely many such n's.

Part (b). Suppose $a > s^*$. Then a is not a lower bound of $\{\bar{t}_k\}$. We can therefore find a j such that $a > \bar{t}_j$ and hence $a > s_n$ for all $n \geqslant j$. In other words, $a > s_n$ for all n except those in the set $\{1, 2, \ldots, j-1\}$, that is, for all but a finite set.

Now parts (c) and (d) follow from parts (a) and (b) and equation (7–2).

Theorem 7.24: *If $\{s_n\}$ is a sequence of real numbers and b is an (extended) real number, then*

(a) $b \leqslant s_n$ *for infinitely many* n *implies* $b \leqslant \limsup s_n$.

(b) $b \geqslant s_n$ *for all but finitely many* n *implies* $b \geqslant \limsup s_n$.

(c) $b \geqslant s_n$ *for infinitely many* n *implies* $b \geqslant \liminf s_n$.

(d) $b \leqslant s_n$ *for all but finitely many* n *implies* $b \leqslant \liminf s_n$.

Proof: Let $\bar{t}_k$ and s^* be as in Definition 7.21.

Part (a). Suppose $b \leqslant s_n$ for infinitely many n. Then certainly $b \leqslant \bar{t}_k$ for every k since each set $\{s_n | n \geqslant k\}$ contains all but finitely many s_n. Therefore, there must be infinitely many members of $\{s_n | n \geqslant k\}$ greater than or equal to b. And as a consequence of this, the least upper bound of this set $\bar{t}_k$ must also be greater than or equal to b. We conclude that $b \leqslant \inf_k \bar{t}_k = \limsup s_n$.

Part (b). Suppose $b \geqslant s_n$ for all but finitely many n, say $n_1, \ldots, n_j$. Choose $k > \max \{n_1, \ldots, n_j\}$. Then, $b \geqslant s_n$ for $n \geqslant k$. So b is an upper bound of the set $\{s_n | n \geqslant k\}$ and, since $\bar{t}_k$ is the least upper bound, we conclude that $b \geqslant \bar{t}_k$. Therefore $\inf_k \bar{t}_k \leqslant \bar{t}_k \leqslant b$. Since $\limsup s_n = \inf_k \bar{t}_k$, this proves part (b).

Finally parts (c) and (d) follow from parts (a) and (b) and equation (7–2).

The next theorem gives us yet another way of asserting the existence of a limit of a sequence without knowing what it is.

Theorem 7.25: *Let $\{s_n\}$ be a sequence of real numbers and let s be an (extended) real number. Then*

$$s_n \to s \tag{7–3}$$

if and only if

$$\limsup s_n = \liminf s_n = s \tag{7–4}$$

Proof: First suppose that equation (7–4) holds and s is finite. Let $\epsilon > 0$ be given. Then $s - \epsilon < \liminf s_n$ and therefore Theorem 7.23(d) shows that $s - \epsilon < s_n$ for all but finitely many n, say $n_1, n_2, \ldots, n_r$. Choose $N_1 > \max \{n_1, n_2, \ldots, n_r\}$. Then, for all $n \geqslant N_1$, $s - \epsilon < s_n$. Now, since $s + \epsilon > \limsup s_n$, we find in the same manner from Theorem 7.23(b) that there exists an N_2 such that, for all $n \geqslant N_2$, $s_n < s + \epsilon$. So setting $N = \max \{N_1, N_2\}$, we see that, for every $n \geqslant N$, $|s - s_n| < \epsilon$ which proves $s = \lim_{n\to\infty} s_n$.

Next suppose equation (7–4) holds and $s = +\infty$. Then, for any finite real number M, we have $M < \liminf s_n$ and so Theorem 7.23(d) shows that $s_n > M$ for all but finitely many n, say $n_1, n_2, \ldots, n_r$. If we choose $N > \max \{n_1, n_2, \ldots, n_r\}$, then $s_n > M$ for all $n \geqslant N$. This shows that equation (7–3) holds. The case where $s = -\infty$ follows from Theorem 7.23(b) in exactly the same way.

Conversely, suppose that equation (7–3) holds and s is finite. Let $\epsilon > 0$ be given, and choose N so that for all $n \geqslant N$

$$|s - s_n| < \epsilon$$

355-525 O-70—9

Then $s+\epsilon > s_n$ for all but finitely many n and, therefore, Theorem 7.24(b) shows $s+\epsilon \geqslant \limsup s_n$. Also, $s-\epsilon < s_n$ for all but finitely many n and so Theorem 7.24(d) shows that

$$s-\epsilon \leqslant \liminf s_n$$

Combining these results with Theorem 7.22 shows that

$$s-\epsilon \leqslant \liminf s_n \leqslant \limsup s_n \leqslant s+\epsilon$$

and, since ϵ was arbitrary, we conclude

$$s = \liminf s_n = \limsup s_n$$

Next suppose equation (7–3) holds and $s=+\infty$. Then, for every real number M, we can find an integer N such that $n \geqslant N$ implies $s_n > M$. Hence $s_n > M$ for all but finitely many n. Therefore, Theorem 7.24(d) implies $M \leqslant \liminf s_n$ for every real number M and so we conclude that $\liminf s_n = +\infty$. Theorem 7.22 now shows that equation (7–4) holds.

Finally, the case for $s=-\infty$ follows from Theorem 7.24(b) in exactly the same way.

The next two theorems show how inferior and superior limits are related to subsequential limits.

Theorem 7.26: *Let $\{s_{n_k}\}$ be a subsequence of the sequence of real numbers $\{s_n\}$. Then*

$$\liminf s_n \leqslant \liminf s_{n_k} \leqslant \limsup s_{n_k} \leqslant \limsup s_n \tag{7–5}$$

Proof: If $\limsup s_n = +\infty$, it is clear that $\limsup s_{n_k} \leqslant \limsup s_n$. So suppose $\limsup s_n \neq +\infty$. Then, for any number $a > \limsup s_n$, Theorem 7.23(b) shows that $s_n < a$ for all but finitely many n. We conclude that $s_{n_k} < a$ for all but finitely many k. Theorem 7.24(b) now shows that $\limsup s_{n_k} \leqslant a$. Now a is any number greater than $\limsup s_n$. If $\limsup s_n = -\infty$, we can conclude that $\limsup s_{n_k}$ is less than or equal to every finite number and therefore that $\limsup s_{n_k} = -\infty$. Hence $\limsup s_{n_k} \leqslant \limsup s_n$. Finally suppose $\limsup s_n$ is finite and let $\epsilon > 0$ be given. Then setting $a = \limsup s_n + \epsilon$, we find that $\limsup s_{n_k} \leqslant \limsup s_n + \epsilon$. But since ϵ was arbitrary, we conclude that $\limsup s_{n_k} \leqslant \limsup s_n$.

Now since $\{-s_{n_k}\}$ is a subsequence of $\{-s_n\}$, this also shows that

$\limsup(-s_{n_k}) \leqslant \limsup(-s_n)$ and so equation (7–2) shows that $\liminf s_n \leqslant \liminf s_{n_k}$. Finally combining these results with Theorem 7.22 gives equation (7–5).

Theorem 7.27: *If $\{s_n\}$ is any sequence of real numbers, then there exists a subsequence $\{s_{n_k}\}$ of $\{s_n\}$ such that*

$$s_{n_k} \to \limsup s_n$$

(Application of equation (7–2) shows that the same result must hold with lim inf in place of lim sup.)

Proof: Let $s = \limsup s_n$. First suppose s is finite. Since $s-1 < \limsup s_n$, Theorem 7.23(a) shows that there are infinitely many n for which $s-1 < s_n$. Let n_1 be any one of these. Suppose we have chosen $n_1, n_2, \ldots, n_k$ so that $n_1 < n_2 < \ldots < n_k$ and

$$s - \frac{1}{j} < s_{n_j} \qquad \text{for } 1 \leqslant j \leqslant k$$

Since $s - \dfrac{1}{k+1} < \limsup s_n$, Theorem 7.23(a) shows that there are infinitely many n such that $s - \dfrac{1}{1+k} < s_n$. Let n_{k+1} be the smallest such n which is larger than n_k. Continuing in this manner, we construct a subsequence $\{s_{n_k}\}$ of $\{s_n\}$ such that for every positive integer N

$$s_{n_k} > s - \frac{1}{N}$$

for all $n_k \geqslant N$. Theorem 7.24(d) now shows that

$$\liminf s_{n_k} \geqslant s - \frac{1}{N}$$

for any positive integer N. We can thus conclude that

$$\liminf s_{n_k} \geqslant s$$

But Theorem 7.26 shows that

$$\limsup s_{n_k} \leqslant s$$

and so, from Theorem 7.22 or 7.26, we see that

$$\limsup s_{n_k} = \liminf s_{n_k} = s$$

Hence Theorem 7.25 shows

$$s_{n_k} \to s$$

Next suppose $s = +\infty$. Then Theorem 7.23(a) shows that for any finite real number M' there are infinitely many n such that $s_n > M'$. Choose n_1 so that $s_{n_1} > 1$. Having chosen $n_1, n_2, \ldots, n_k$ in such a way that $n_1 < n_2 < \ldots < n_k$ and $s_{n_k} > k$, choose n_{k+1} from the infinitely many n for which $s_n > k+1$ to be the smallest such n which is larger than n_k. Continuing in this way we construct a subsequence $\{s_{n_k}\}$ of $\{s_n\}$ such that, for any positive integer K, $s_{n_k} > K$ whenever $k \geqslant K$. Thus, if M is any finite number, we can choose $K \geqslant M$ Then $s_{n_k} > M$ whenever $k \geqslant K$. That is, $s_{n_k} \to +\infty$.

Finally if $s = -\infty$ it follows from Theorems 7.22 and 7.25 that $s_n \to -\infty$ and, since $\{s_n\}$is a subsequence of itself, we are done.

Theorems 7.25 and 7.26 show that all the subsequential limits of a real sequence $\{s_n\}$ must be less than or equal to $\limsup s_n$ and greater than or equal to $\liminf s_n$. In fact, if we let S be the set of all extended real numbers s such that $s_{n_k} \to s$ for some subsequence $\{s_{n_k}\}$ of $\{s_n\}$, it is clear that $\limsup s_n$ is an upper bound of S and $\liminf s_n$ is a lower bound of S. Also, Theorem 7.27 shows that they are the least upper bound and greatest lower bound, respectively, of S and that they are both members of S.

The next theorem introduces two of the many inequalities relating the superior and inferior limits. Most of the others can be obtained by the method used in proving the theorem.

Theorem 7.28: *Let $\{s_n\}$ and $\{t_n\}$ be sequences of finite real numbers.*
(a) *Then* $\limsup s_n + \limsup t_n \leqslant \limsup (s_n + t_n)$.
(b) *If $s_n \leqslant t_n$ for every n, then* $\limsup s_n \leqslant \limsup t_n$.

Proof: Part (a). Set $s^* = \limsup s_n$ and $t^* = \limsup t_n$. Let $\epsilon > 0$ be given. Theorem 7.23(a) shows that, for infinitely many n, $s_n > s^* - \epsilon/2$ and $t_n > t^* - \epsilon/2$. Thus, for infinitely many n,

$$s_n + t_n > s^* + t^* - \epsilon$$

and so Theorem 7.24(a) shows that

$$s^* + t^* - \epsilon \leqslant \limsup (s_n + t_n)$$

Since ϵ was arbitrary, part (a) follows.

Part (b). Let s^* be as in part (a). Notice that, given $\epsilon > 0$, $t_n \geqslant s_n > s^* - \epsilon$ for infinitely many n. So Theorem 7.24(a) shows that

$$s^* - \epsilon \leqslant \limsup t_n$$

Since ϵ was arbitrary, part (b) follows.

Let $\{s_n\}$ be the sequence whose nth term is

$$s_n = (-1)^n + \frac{1}{n^2}$$

Then $\liminf s_n = -1$ and $\limsup s_n = 1$.

We have shown in chapter 5 that the set of all rational numbers can be "arranged" in a sequence. It is easy to show by use of the axiom of Archimedes (chapter 2) that every real number is a subsequential limit of this sequence.

We conclude this chapter by computing the limits of two real valued sequences which occur frequently in practice and will be referred to later in the text. In order to accomplish this we shall employ the following device: If $\{s_n\}$ and $\{t_n\}$ are numerical sequences and if there is some integer N such that, for all $n \geqslant N$, $0 \leqslant s_n \leqslant t_n$, then $\{s_n\}$ converges to zero if $\{t_n\}$ does.

Let a be a positive number. It is clear that, for any $\epsilon > 0$, $|1/n^a| < \epsilon$ for every integer n larger than $(1/\epsilon)^{1/a}$. Hence, $\lim_{n\to\infty} 1/n^a = 0$.

On the other hand, for $n \geqslant 2m$ and $d > 1$, we see from the binomial theorem that

$$d^n = [1 + (d-1)]^n > \frac{n(n-1)\ \ldots\ (n-m+1)}{m!}(d-1)^m > \frac{n^m(d-1)^m}{2^m m!}$$

Hence, for any number r

$$0 < \frac{n^r}{d^n} < \left(\frac{2^m m!}{(d-1)^m}\right)\frac{1}{n^{m-r}}$$

for $n > 2m$. If we take m larger than r in this inequality we see that

$$\lim_{n\to\infty} \frac{n^r}{d^n} = 0 \qquad \text{for } d > 1$$

CHAPTER 8

Continuity and Function Algebras

According to Cauchy's definition of continuity, a function is continuous at a point p if it has a limit at p. The modern point of view is somewhat different from this, but, as we shall see, the current definition of continuity is equivalent to Cauchy's.

Although the topological space is actually the most natural setting for discussing continuous functions, the somewhat less general metric space will better serve our purposes. By discussing the concept of continuity in the abstract setting of a metric space instead of talking about the continuity of real or complex valued functions on the line or plane, we not only obtain greater generality (which is by no means without application) but many of the proofs of the theorems are actually simplified.

We shall first introduce the concept of limit which is closely related to the concept of continuity.

Definition 8.1: *If $\langle X, d\rangle$ and $\langle Y, d'\rangle$ are metric spaces, $E \subset X, f: E \to Y$, and p is a limit point of E, we shall write*

$$f(x) \to q \qquad \text{as } x \to p$$

or

$$\lim_{x \to p} f(x) = q$$

if there exists a point $q \in Y$, with the property that, for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$f(B(p; \delta) \cap E) \subset B(q; \epsilon)$$

In this event f is said to **have a limit** *at p, and the point q is called the* **limit of f at** *p.*

Definition 8.1 does *not* require that the limit point p belong to E (the domain of definition f). However, if it happens that p is in E, then $B(p; \delta) \cap E$ is nothing more than the ball about p of radius δ in the subspace $\langle E, d\rangle$ (see Theorem 6.15). Hence, if we had imposed the restriction $p \in E$, the definition could have been made entirely within the metric space $\langle E, d\rangle$ without mention of the embedding space $\langle X, d\rangle$. Naturally the domain E of f can be equal to X. We shall now show that this definition could have been stated in terms of sequences.

By definition, a sequence $\{p_n\}$ in a set X is a function $f: J \to X$ where J is the set of positive integers. If Y is another set and $g: X \to Y$, then the composition $h = g \circ f$ is a function from J to Y and, hence, is a sequence in Y. Moreover, for each integer $n \in J$, $h(n) = g(p_n)$.

Theorem 8.2: *If X and Y are metric spaces, $E \subset X, f: E \to Y$, and p is a limit point of E, then*

$$\lim_{x \to p} f(x) = q$$

if and only if,

$$\lim_{n \to \infty} f(p_n) = q$$

for every sequence $\{p_n\}$ in E such that $p_n \to p$.

Proof: Assume that equation (8–1) is true and let $\{p_n\}$ be any sequence in E which converges to p. Given $\epsilon > 0$, we can find a $\delta > 0$ such that

$$f(B(p; \delta) \cap E) \subset B(q; \epsilon)$$

and, for this δ, we can find a positive integer N such that, for every $n \geq N$, $p_n \in B(p; \delta)$. Since $p_n \in E$, it is clear that $p_n \in B(p; \delta) \cap E$. Hence, $f(p_n) \in B(q; \epsilon)$ for all $n \geq N$ and equation (8–2) holds.

On the other hand, if equation (8–1) is not true, then, there is some $\epsilon > 0$ such that, for each $\delta > 0$, we can find at least one $x \in B(p; \delta) \cap E$ for which $f(x) \notin B(q; \epsilon)$. Hence, for each positive integer n, choose $p_n \in B(p; 1/n) \cap E$ such that $f(p_n) \notin B(q; \epsilon)$. Then $\{p_n\}$ is a sequence of points of E which converges to p for which equation (8–2) does not hold.

The following corollary is an immediate consequence of Theorems 7.3 and 8.2.

Corollary 1: *If a function f from a subset E of a metric space X to a metric space Y has a limit at some limit point of E, then this limit is unique.*

Corollary 2: *If a function f from a subset E of metric space X to a metric space Y has a limit at some limit point p of E, then $p \in E$ implies $\lim_{x \to p} f(x) = f(p)$.*

Proof: Let $\{p_n\}$ be the constant sequence $p, p, p, \ldots$. The conclusion follows from the fact that $\{p_n\}$ is a sequence in E such that $\lim_{n \to \infty} f(p_n) = f(p)$.

Corollary 3: *Let $\langle X, d\rangle$ and $\langle Y, d'\rangle$ be two metric spaces, let $E \subset A \subset X$, and let $f: A \to Y$. Suppose that p is a limit point of both A and E and that $\lim_{x \to p} f(x) = q$. If $g: E \to Y$ is the restriction of f to E, then $\lim_{x \to p} g(x) = q$.*

Proof: If $\{p_n\}$ is any sequence of points of E such that $p_n \to p$, then $\{p_n\}$ is also a sequence of points of A with this property. Hence the theorem shows that $\lim_{n \to \infty} f(p_n) = q$. Since each p_n belongs to E, it follows that $f(p_n) = g(p_n)$ for every n. Hence $\lim_{n \to \infty} g(p_n) = q$. Since $\{p_n\}$ was *any* sequence in E which converges to p, the theorem shows that $\lim_{x \to p} g(x) = q$.

Corollary 4: *Let X and Y be metric spaces and let A be a nonempty subset of X. If $f: A \to Y$, p is a limit point of A and $\lim_{x \to p} f(x) = q$, then $q \in \overline{f(A)}$.*

Proof: If p is a limit point of A it is certainly an adherence point of A. Hence it follows from Theorem 7.4 that there is a sequence $\{p_n\}$ of points of A such that $p_n \to p$. Therefore Theorem 8.2 shows that $f(p_n) \to q$. Since every p_n belongs to A it is clear that $f(p_n) \in f(A)$ for every n. Hence $\{f(p_n)\}$ is a sequence of points of $f(A)$ which converges to q. Thus Theorem 7.4 now shows that q is an adherence point of $f(A)$; that is, $q \in \overline{f(A)}$.

Definition 8.3: *If $\langle X, d\rangle$ and $\langle Y, d'\rangle$ are metric spaces and $f: X \to Y$, the function f is said to be* **continuous at the point** *$p \in X$ if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that*

$$f(B(p; \delta)) \subset B(f(p); \epsilon) \tag{8–3a}$$

If a function is not continuous at a point $p \in X$ it is said to be **discontinuous** *at p or to have a discontinuity at p. If f is continuous at every point of X, then f is simply said to be* **continuous.**

Note that $B(p; \delta)$ is a ball in $\langle X, d\rangle$ while $B(f(p); \epsilon)$ is a ball in $\langle Y, d'\rangle$. If we are willing to agree that the mathematical concept of "ball about a point" corresponds to our intuitive idea of "proximity" to that point, then the preceding definition says that *f is continuous at p if $f(q)$ is arbitrarily close to $f(p)$ as soon as q is sufficiently close to p.* The concept of continuity is illustrated in figure 8–1.

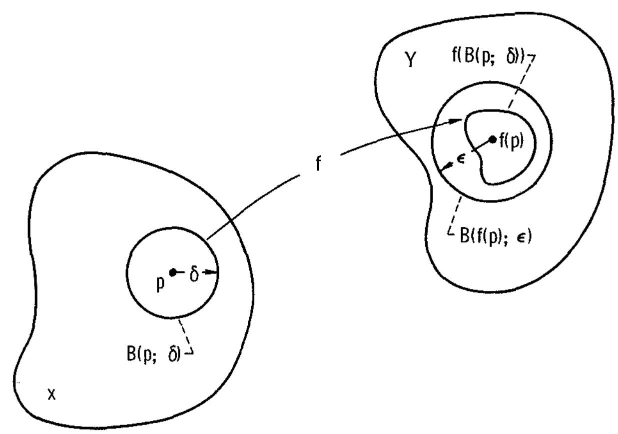


FIGURE 8–1.—Illustration of continuity concept.

If $f: X \rightarrow Y$, $q \in Y$ and $f(x) = q$ for every $x \in X$, then f is called a *constant* mapping. Clearly every constant mapping is continuous.

Suppose $\langle Z, d\rangle$ and $\langle Y, d'\rangle$ are metric spaces, $E \subset Z$, and $f: E \rightarrow Y$. Now $\langle E, d\rangle$ is a metric space (with the same metric as $\langle Z, d\rangle$) and Definition 8.3 applies to this metric space and not to the metric space $\langle Z, d\rangle$. In this situation, then, we must realize that the ball $B(p;\ \delta)$ in relation (8–3a) is to be interpreted as a ball in the metric space $\langle E, d\rangle$ and not as a ball in the metric space $\langle Z, d\rangle$. To point out this difference, let us temporarily return to the notation of Theorem 6.15 and use the superscript E to denote balls in the metric space $\langle E, d\rangle$ and the superscript Z to denote balls in the metric space $\langle Z, d\rangle$. With this notation, the condition (8–3a) is

$$f(B^E(p;\ \delta)) \subset B(f(p);\ \epsilon) \tag{8–3b}$$

But Theorem 6.15 shows us that there is a simple relation between balls in the metric space $\langle E, d\rangle$ and those in the metric space $\langle Z, d\rangle$ which for the present case is

$$B^E(p;\ \delta) = B^Z(p;\ \delta) \cap E$$

If this relation is used, the inclusion (8–3b) can also be written as

$$f(B^Z(p;\ \delta) \cap E) \subset B(f(p);\ \epsilon) \tag{8-3c}$$

It is not hard to verify that Definition 8.3 is entirely equivalent to the following more familiar definition of continuity:

If $\langle X, d\rangle$ and $\langle Y, d'\rangle$ are metric spaces and $f: X \to Y$, then the function f is continuous at the point $p \in X$ if and only if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$d'(f(x), f(p)) < \epsilon$$

whenever

$$d(x, p) < \delta$$

Instead of explicitly mentioning the radii ϵ and δ of the balls in Definition 8.3, we could have said that, if f is to be continuous at p, then, for every ball B about $f(p)$, there must be a ball about p whose image under f is a subset of B. In fact, since every ball about a point is also a neighborhood and since, for every neighborhood V of a point, there is a ball about that point which is in V, we could have stated Definition 8.3 completely in terms of neighborhoods. In this way we see that continuity at a point, like convergence, is a topological property.

Definition 8.3 requires, in contrast to Definition 8.1, that f must be defined at a point p in order to be continuous at p. If the point p in Definition 8.3 is not a limit point (i.e., if it is an isolated point), then we can find a ball $B(p;\ \delta)$ about p that contains only the point p and, for this δ,

$$f(B(p;\ \delta)) = f(\{p\}) \subset B(f(p);\ \epsilon)$$

for every $\epsilon > 0$. Thus *every function is continuous at the isolated points of its domain.* On the other hand, if p is a limit point of X, then there is a close relation between Definitions 8.1 and 8.3 which is given by the following theorem.

Theorem 8.4: *If X and Y are metric spaces, $f: X \to Y$ and $p \in X$ is a limit point of X, then, f is continuous at p if and only if $\lim_{x \to p} f(x) = f(p)$.*

Proof: In view of the fact that the domain of definition of f is all of X, it follows from Definition 8.1 that $\lim_{x \to p} f(x) = f(p)$ if and only if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$f(B(p;\ \delta)) = f(B(p;\ \delta) \cap X) \subset B(f(p);\ \epsilon)$$

that is if and only if f is continuous at p.

Let f map the metric space X into the metric space Y. If p is an isolated point of X and $\{p_n\}$ is any sequence in X which converges to p, it is clear that $\{p_n\}$ differs from the constant sequence $p, p, p, \ldots$ by at most a finite number of terms. It follows from this that

$$\lim_{n\to\infty} f(p_n) = f(p)$$

Hence, in view of the remarks preceding Theorem 8.4, the following theorem is an immediate consequence of Theorems 8.2 and 8.4:

Theorem 8.5: *A function f from a metric space X to a metric space Y is continuous at a point p of X if and only if, for every sequence $\{p_n\}$ in X which converges to p, $\lim_{n\to\infty} f(p_n) = f(p)$.*

This theorem shows that the continuous functions are precisely those which map convergent sequences into convergent sequences or in other words which "preserve convergence."

Corollary: *Let $\langle X, d\rangle$, $\langle Y, \delta\rangle$, and $\langle S, \rho\rangle$ be metric spaces. Let $f: S \to X$ and $g: S \to Y$. Then the function $h: S \to X \times Y$ defined by*

$$h(s) = \langle f(s), g(s)\rangle \qquad \textit{for all } s \in S$$

is a function into the direct product $\{X \times Y, d_X\rangle$ of $\langle X, d\rangle$ and $\langle Y, \delta\rangle$ which is continuous at the point $p \in S$ *if and only if both f and g are continuous at p.*
Proof: This is an immediate consequence of the theorem and Theorem 7.9(a).

Theorem 8.6: *Let $\langle X, d\rangle$, $\langle Y, d'\rangle$, and $\langle Z, d''\rangle$ be metric spaces. Suppose that $f: X \to Y$ and $g: Y \to Z$. If f is continuous at a point $p \in X$ and g is continuous at $f(p)$, then the composition $h = g \circ f$ is continuous at p.*
Proof: (See fig. 8–2.) Let $\epsilon > 0$ be given. Since g is continuous at $f(p)$, there is an $\eta > 0$ such that

$$g(B(f(p); \eta)) \subset B(g(f(p)); \epsilon) = B(h(p); \epsilon)$$

Now the continuity of f at p shows that we can find a $\delta > 0$ such that

$$f(B(p; \delta)) \subset B(f(p); \eta)$$

Hence,

$$g(f(B(p; \delta))) \subset g(B(f(p); \eta)) \subset B(h(p); \epsilon)$$

But,

$$g(f(B(p;\ \delta)))=h(B(p;\ \delta))$$

which proves the theorem.

Roughly speaking, this theorem states that a continuous function of a continuous function is continuous.

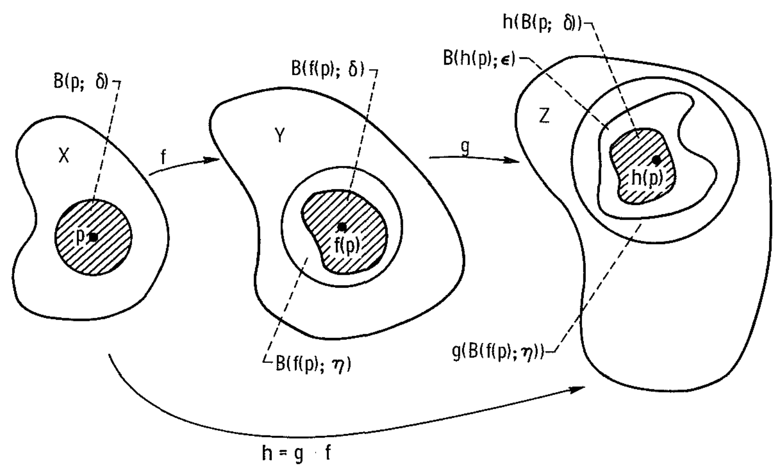


FIGURE 8–2.—Illustration of Theorem 8.8 proof.

The following useful characterization of continuity points out the topological nature of this property.

Theorem 8.7: *If X and Y are metric spaces and $f: X \rightarrow Y$, then f is continuous if and only if, for every open set $V \subset Y$, $f^{-1}(V)$ is an open subset of X.*

Proof: First let f be continuous and choose an open set $V \subset Y$. If p is any point of $f^{-1}(V)$, then $f(p) \in V$. Since V is an open set, we can find a ball $B(f(p);\ \epsilon)$ about $f(p)$ of radius ϵ such that $B(f(p);\ \epsilon) \subset V$. Now the continuity of f at p implies that there is a $\delta > 0$ such that $f(B(p;\ \delta)) \subset B(f(p);\ \epsilon) \subset V$. Then from table 4–I we see that $B(p;\ \delta) \subset f^{-1}(f(B(p;\ \delta))) \subset f^{-1}(V)$ and this shows that $f^{-1}(V)$ is open.

On the other hand, suppose $f^{-1}(V)$ is an open set in X whenever V is an open set in Y. Let p be any point of X and fix $\epsilon > 0$. Since $B(f(p);\ \epsilon)$ is an open set, the set $f^{-1}(B(f(p);\ \epsilon))$ is also. Therefore the fact that $p \in f^{-1}(B(f(p);\ \epsilon))$ shows that there is a $\delta > 0$ for which

$$B(p;\,\delta) \subset f^{-1}(B(f(p);\,\epsilon)$$

It follows from table 4–I that, for this δ,

$$f(B(p;\,\delta)) \subset f(f^{-1}(B(f(p);\,\epsilon))) \subset B(f(p);\,\epsilon)$$

Since ϵ was arbitrary, this shows that f is continuous at p and, since p was any point of X, we conclude that f is continuous.

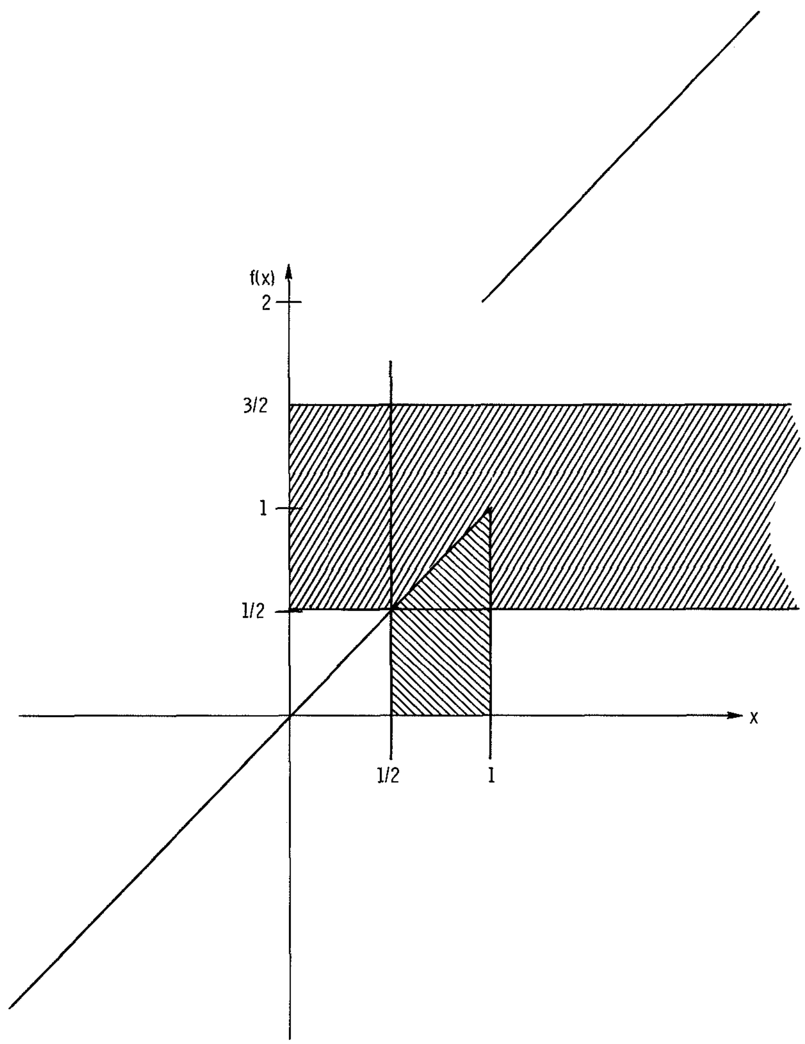


FIGURE 8–3.—Inverse image of an open set under a discontinuous function.

Consider the function $f: R^1 \to R^1$ defined by

$$f(x) = \begin{cases} x & \text{for } x \leqslant 1 \\ 1+x & \text{for } x > 1 \end{cases}$$

This function is clearly discontinuous at 1 and $f^{-1}((1/2, 3/2)) = (1/2, 1]$. Evidently $(1/2, 3/2)$ is an open set and $(1/2, 1]$ is not since 1 is a point of this interval which is not an interior point. This example is illustrated in figure 8–3.

Corollary 1: *If E is any nonempty subset of a metric space $\langle X, d\rangle$, then j_E, the natural injection of E into X, is a continuous mapping of the metric space $\langle E, d\rangle$ into $\langle X, d\rangle$.*

Proof: Let G be any *open* subset of X. Equation (4–3) shows that $j_E^{-1}(G) = G \cap E$. Hence Theorem 6.16 implies that $j_E^{-1}(G)$ is an open subset of $\langle E, d\rangle$. The conclusion now follows from the theorem.

Notice that even though the inverse image of every open set under a continuous mapping is an open set, the image of an open set need not be open. For, if $f: R^1 \to R^1$ is defined by

$$f(x) = x^2 \qquad \text{for every } x \in R^1$$

then f is continuous and the image under f of the open set $(-1, 1)$ is the half-open interval $[0, 1)$ which is not an open subset of R^1.

Corollary 2: *If a mapping f of a metric space $\langle X, d\rangle$ into a metric space $\langle Y, d'\rangle$ is continuous at a point p and if E is a subset of X which contains p, then the restriction of f to E is a mapping of the metric space $\langle E, d\rangle$ into Y which is continuous at p.*

Proof: According to the remark following Definition 4.11, the restriction of f to E is the mapping $f \circ j_E$ where j_E is the natural injection of E into X. Corollary 1 to Theorem 8.7 shows that j_E is a mapping of the metric space $\langle E, d\rangle$ into the metric space $\langle X, d\rangle$ which is certainly continuous at p. Hence Theorem 8.6 now shows that $f \circ j_E$, the restriction of f to E, is a mapping from $\langle E, d\rangle$ into $\langle Y, d'\rangle$ which is continuous at p.

Notice however that the restriction of a mapping f of a metric space $\langle X, d\rangle$ to a subspace $\langle E, d\rangle$ may be a continuous mapping of $\langle E, d\rangle$ even though the mapping f defined on $\langle X, d\rangle$ is not continuous at any point of E. An example of this is provided by Dirichlet's function, which is the mapping $f: R^1 \to R^1$

defined by

$$f(x) = \begin{cases} 0 & \text{for } x \text{ a rational number} \\ 1 & \text{for } x \text{ an irrational number} \end{cases}$$

Since the restriction of f to the set Q of all rational numbers is a constant mapping of Q into R^1, it is clearly continuous even though f is certainly not continuous at any point of R^1.

Theorem 8.8: *Let f and g be two continuous mappings of a metric space $\langle X, d\rangle$ into a metric space $\langle Y, d'\rangle$. Then the set*

$$A = \{x \in X \mid f(x) = g(x)\}$$

is closed.

Proof: The corollary to Theorem 6.10 shows it is sufficient to prove that A^c is open. To this end let p be any point of A^c. Evidently $f(p) \neq g(p)$. Hence if we set $\epsilon = d'(f(p), g(p))$, then $\epsilon > 0$. Since f and g are both continuous at p, there exists a $\delta > 0$ such that $d'(f(p), f(x)) < \epsilon/2$ and $d'(g(p), g(x)) < \epsilon/2$ for all $x \in B(p; \delta)$. Then, whenever $x \in B(p; \delta)$, $\epsilon = d'(f(p), g(p)) \leq d'(f(p), f(x)) + d'(f(x), g(x)) + d'(g(x), g(p)) < \epsilon/2 + d'(f(x), g(x)) + \epsilon/2$. Hence

$$0 < d'(f(x), g(x))$$

This shows that $f(x) \neq g(x)$ for any $x \in B(p; \delta)$. Therefore $B(p; \delta) \subset A^c$. Since p was any point of A^c, it follows that A^c is open.

Corollary: *Let f and g be two continuous functions from a metric space X to a metric space Y and suppose there is a dense subset $E \subset X$ such that, for all $x \in E$, $f(x) = g(x)$. Then $f = g$.*

Proof: Since $\bar{E} = X$ and since $\bar{E}$ is the smallest closed set containing E, the theorem shows that $f(x) = g(x)$ for every $x \in X$; that is, $f = g$.

Theorem 8.9: *Let A be a dense subset of the metric space $\langle X, d\rangle$ and let f map A into the metric space $\langle Y, d'\rangle$. Then f has a continuous extension $\hat{f}$ to X if and only if f has a limit at every limit point p of X. The extension $\hat{f}$ is then unique.*

Proof: First suppose that a continuous extension $\hat{f}$ of f exists. Then f is the restriction of $\hat{f}$ to A. Theorem 8.4 shows that at every limit point p of X

$$\lim_{x \to p} \hat{f}(x) = \hat{f}(p)$$

Since p is a limit point of X and A is a dense subset of X, Theorem 6.9(b) shows that p is a limit point of A. Hence the third corollary to Theorem 8.2 shows that $\lim_{x \to p} f(x)$ exists, which proves the assertion.

Conversely, suppose that $\lim_{x \to p} f(x)$ exists at every limit point p of X. If q is not a limit point of X there exists a ball $B(q; \epsilon)$ about q such that q is the only point of X belonging to $B(q; \epsilon)$. Since A is dense in X, $B(q; \epsilon)$ must contain a point of A. Hence $q \in A$. Thus every point of X is either a limit point of X or a point of A. In view of this, let us define $\hat{f}: X \to Y$ by

$$\hat{f}(p) = \begin{cases} \lim_{x \to p} f(x) & \text{if } p \text{ is a limit point of } X \\ f(p) & \text{if } p \text{ is not a limit point of } X \end{cases}$$

Let p be any point of A. If p is not a limit point of X, $\hat{f}(p) = f(p)$. If p is a limit point of X, then Corollary 2 of Theorem 8.2 shows that $\hat{f}(p) = \lim_{x \to p} f(x) = f(p)$. Hence $\hat{f}$ is an extension of f.

Let us now show that $\hat{f}$ is continuous at every point $p \in X$. If p is not a limit point of X, then $\hat{f}$ is automatically continuous at p. Hence suppose that p is a limit point of X. Fix $\epsilon > 0$. By construction, there exists a $\delta > 0$ such that

$$f(B(p; \delta) \cap A) \subset B(\hat{f}(p); \epsilon/2)$$

Since $B(p; \delta) \cap A$ is a subset of A and $\hat{f}$ is an extension of f,

$$f(B(p; \delta) \cap A) = \hat{f}(B(p; \delta) \cap A)$$

Hence

$$\hat{f}(B(p; \delta) \cap A) \subset B(\hat{f}(p); \epsilon/2) \tag{8–4}$$

Let y be any point of $B(p; \delta)$. We shall show that $\hat{f}(y) \in B(\hat{f}(p); \epsilon)$. If y belongs to A, then

$$\hat{f}(y) \in B(\hat{f}(p); \epsilon/2) \subset B(\hat{f}(p); \epsilon)$$

Hence suppose that $y \notin A$. We have already established that y must be a limit point of X. Theorem 6.9(b) now shows that y is a limit point of A. Hence y

355-525 O-70—10

must also be a limit point of $B(p;\,\delta) \cap A$.[21] Corollary 3 of Theorem 8.2 now shows that if g is the restriction of f to $B(p;\,\delta) \cap A$, then

$$\hat{f}(y) = \lim_{x \to y} g(x)$$

Hence Corollary 4 of Theorem 8.2 shows that $\hat{f}(y) \in \overline{g(B(p;\,\delta) \cap A)}$. Since g is the restriction of f to $B(p;\,\delta) \cap A$ and f is the restriction of $\hat{f}$ to $A \supset B(p;\,\delta) \cap A$, it follows that

$$g(B(p;\,\delta) \cap A) = f(B(p;\,\delta) \cap A) = \hat{f}(B(p;\,\delta) \cap A)$$

We conclude from the inclusion 8.4 and Theorem 6.8 that

$$\hat{f}\,(y) \in \overline{\hat{f}\,(B(p;\,\delta) \cap A)} \subset \overline{B(\hat{f}\,(p);\,\epsilon/2)}$$

Since

$$B(\hat{f}\,(p);\,\epsilon/2) \subset \overline{B(\hat{f}\,(p);\,\epsilon/2)} \subset B(\hat{f}\,(p);\,\epsilon)$$

we have established that

$$\hat{f}\,(y) \in B(\hat{f}\,(p);\,\epsilon) \qquad \text{for every } y \in B(p;\,\delta)$$

Hence $\hat{f}$ is continuous at p for every $p \in X$.

It remains only to prove that there is only one continuous extension of f. But this is an immediate consequence of the corollary to Theorem 8.8 and the fact that A is a dense subset of X.

In discussing sequences we examined the relation between algebraic operations on the one hand and convergence on the other. Now we shall consider the relation between algebraic operations and continuity. To this end we introduce a new algebraic structure called an algebra which is linear space whose elements can be "multiplied together."

---

[21] For if $y \in B(p;\,\delta)$ were not a limit point of $B(p;\,\delta) \cap A$, there would be an $\epsilon_1 > 0$ such that $B(y;\,\epsilon_1)$ contained no points of $B(p;\,\delta) \cap A - \{y\}$. Since $y \notin A$, this means that $B(y;\,\epsilon_1)$ would contain no points of $B(p;\,\delta) \cap A$. But since y belongs to the open set $B(p;\,\delta)$, there is an $\epsilon_2 > 0$ such that $B(y;\,\epsilon_2) \subset B(p;\,\delta)$. Set $\epsilon_3 = \min\{\epsilon_1, \epsilon_2\}$. Since y is a limit point of A, $B(y;\,\epsilon_3)$ contains a point of A. But every point of $B(y;\,\epsilon_3)$ must be a point of $B(p;\,\delta)$. Hence $B(y;\,\epsilon_3)$ contains a point of $B(p;\,\delta) \cap A$. Since $B(y;\,\epsilon_3) \subset B(y;\,\epsilon_1)$, $B(y;\,\epsilon_1)$ must also contain a point of $B(p;\,\delta) \cap A$ which is a contradiction.

Definition 8.10: *A real (complex)* **algebra** *is a real (complex) linear space A and an operation called* **multiplication** *which associates with any two elements a_1, $a_2 \in A$ a unique element of A, denoted by $a_1 a_2$, in such a way that if a_3 is any element of A and α is any real (complex) number, the following are true:*

(M1)	$(a_1a_2)a_3 = a_1(a_2a_3)$	*(Associativity)*
(M2)	$\alpha(a_1a_2) = (\alpha a_1)a_2 = a_1(\alpha a_2)$	*(Associativity)*
(D1)	$a_1(a_2 + a_3) = a_1a_2 + a_1a_3$	*(Distributive law)*
(D2)	$(a_2 + a_3)a_1 = a_2a_1 + a_3a_1$	*(Distributive law)*

If

$$a_1a_2 = a_2a_1 \qquad \text{for all } a_1, a_2 \in A$$

then A is said to be a **commutative algebra.**

If there exists an element $e \in A$ such that

$$ea = ae = a \qquad \text{for all } a \in A$$

then e is said to be **unity** *of A.*

If a is a **nonzero** *element*[22] *of an algebra A with a unity e and if there exists an $x \in A$ such that*

$$xa = ax = e$$

then x is called the **multiplicative inverse**[23] *of a and is denoted by a^{-1}.*

If a is a nonzero element of a commutative algebra A (with a unity) which has a multiplicative inverse and b is any element of A, we write b/a in place of $a^{-1}b$ or ba^{-1} since $a^{-1}b = ba^{-1}$.

It is not hard to prove that if an algebra A has unity e, then e is unique, and if a nonzero element $a \in A$ has a multiplicative inverse, then this inverse is also unique.

It is easily verified that the set of real numbers (with the usual method of adding and multiplying numbers) satisfies the axioms of a real commutative algebra with a unity in which every nonzero element has a multiplicative inverse. The zero of this algebra is of course the number zero and the unity

[22] The zero element of an algebra is of course the zero of the underlying linear space.

[23] Notice that since A is a linear space there is an element $-a \in A$ for every element $a \in A$. We have also called this element $-a$ the inverse of a. It is often called the additive inverse of a to distinguish it from the multiplicative inverse.

is the number 1. Also the set C of all complex numbers (with the usual definitions of addition and multiplication) satisfies the axioms of complex commutative algebra with a unity in which every nonzero element has a multiplicative inverse.

If a is an element of an algebra we write $a^1 = a$ and for any positive integer n we define a^n by

$$a^n = a^{n-1}a$$

Thus a^n is defined inductively for every positive integer n.

It can be proved by induction that for any positive integers m and n the following laws of exponents hold:

$$a^n a^m = a^{(n+m)}$$
$$(a^n)^m = a^{nm}$$

Let E be a subset of a real (complex) algebra A. For any $a_1, \ldots, a_n \in E$, any real (complex) number λ and any positive integers $i_1, i_2, \ldots, i_n$, the element [24]

$$m = \lambda a_1^{i_1} a_2^{i_2} \ldots a_n^{i_n}$$

of A is called a *monomial in E*. If $m_1, m_2, \ldots, m_k$ are any k monomials in E, the element

$$p = m_1 + m_2 + \ldots + m_k$$

is called a *polynomial in E*.

We now are ready to consider those algebraic operations which combine functions. We shall require that functions be combined in a pointwise manner. The reader is reminded that if f is a function from a set X to a linear space M, then, for each $x \in X$, $f(x)$ is a point of M. That is, for each x, $f(x)$ is a vector. With this in mind, we make the following definition.

Definition 8.11: *Let X be any set and M a real (complex) vector space. Let $\mathscr{F}(X, M)$ be the family of all functions from X into M. For any two functions $f, g \in \mathscr{F}(X, M)$ the function $f+g$ is defined pointwise by*

[24] If A has a unity e, we can define in a consistent way, for any nonzero $a \in A$, $a^0 = e$. Then if $e \in E$ we can allow the integers $i_1, \ldots, i_n$ in this definition to be any *nonnegative* integers. Thus, for any real (complex) number λ, λe is a monomial called the "constant" monomial. Any polynomial which has the constant monomial as one of its terms is said to have a "constant" term.

$$(f+g)(x)=f(x)+g(x) \qquad \text{for all } x \in X \tag{8-5}$$

For any real (complex) number α and any function $f \in \mathscr{F}(X, M)$ the function αf is defined by

$$(\alpha f)(x)=\alpha f(x) \qquad \text{for all } x \in X \tag{8-6}$$

The function $(-1)f$ is denoted by $-f$.

For any point $a \in M$ the function $f \in \mathscr{F}(X, M)$ defined by

$$f(x)=a \qquad \text{for all } x \in X \tag{8-7}$$

is called a **constant function** *and we write $f=a$.*

If in addition M is an algebra, then, for any two functions $f, g \in \mathscr{F}(X, M)$, we define the function fg by

$$(fg)(x)=f(x)g(x) \qquad \text{for all } x \in X \tag{8-8}$$

If M is a commutative algebra in which every nonzero element has a multiplicative inverse, then, for any two functions f, $g \in \mathscr{F}(X, M)$ such that 0, *the zero vector of the vector space M, does not belong to the range of f, we define the function g/f by*

$$(g/f)(x)=g(x)/f(x) \qquad \text{for all } x \in X \tag{8-9}$$

If $\boldsymbol{f}$ and $\boldsymbol{g}$ map X into the Euclidean space R^k, we define $\boldsymbol{f} \cdot \boldsymbol{g}$ by

$$(\boldsymbol{f} \cdot \boldsymbol{g})(x)=\boldsymbol{f}(x) \cdot \boldsymbol{g}(x) \qquad \text{for all } x \in X$$

These definitions can be extended in an obvious way to functions with values in the *extended* real number system.

Theorem 8.12: *Let X be any set and M a real (complex) vector space. Then $\mathscr{F}(X, M)$, the family of all functions from X into M, is a vector space with the operations of addition and scalar multiplication defined by equations* (8–5) *and* (8–6), *respectively.*

If M is also an algebra, then $\mathscr{F}(X, M)$ is an algebra[25] *when multiplication on $\mathscr{F}(X, M)$ is defined by equation* (8–8).

[25] We shall call any such algebra whose elements are functions a *function algebra.*

Proof: Clearly the constant function 0, which associates the zero vector with every $x \in X$, is the zero of $\mathscr{F}(X, M)$. Since

$$(f+(-f))(x)=f(x)+(-1)f(x)=0f(x)=0 \qquad \text{for every } x \in X$$

it is clear that $-f$ is the additive inverse of f for any $f \in \mathscr{F}(X, M)$.

It is an immediate consequence of the fact that the remaining axioms of Definition 3.1 hold in the vector space M that they also hold in $\mathscr{F}(X, M)$. Let us show, for example, that the commutative law of addition ((A1) of Definition 3.1) holds in $\mathscr{F}(X, M)$. We have, by definition, for any two functions $f, g \in \mathscr{F}(X, M)$

$$(f+g)(x)=f(x)+g(x) \qquad \text{for all } x \in X$$

and

$$(g+f)(x)=g(x)+f(x) \qquad \text{for all } x \in X$$

Since M is a vector space we must have

$$f(x)+g(x)=g(x)+f(x) \qquad \text{for every } x \in X$$

Hence,

$$(f+g)(x)=(g+f)(x) \qquad \text{for every } x \in X$$

That is,

$$f+g=g+f$$

which is the desired result.

If the points of M satisfy the axioms of Definition 8.10 in addition to those of Definition 3.1, it follows that the elements of $\mathscr{F}(X, M)$ do also when multiplication is defined by equation (8–8).

In particular this theorem shows that *the set of all real* (*complex*) *valued functions which are defined on a given set X is a real* (*complex*) *algebra.*

It follows from the preceding definitions that if $f \in \mathscr{F}(X, M)$ and M is a real (complex) algebra, then for any positive integer n the function f^n is defined by

$$f^n(x)=[f(x)]^n \qquad \text{for all } x \in X$$

or, more generally, if $\alpha_1, \alpha_2, \ldots, \alpha_n$ are n real (complex) numbers, the function $\alpha_1 f^1+\ldots+\alpha_n f^n$ is defined by

$$(\alpha_1 f^1+\ldots+\alpha_n f^n)(x)=\alpha_1 f^1(x)+\ldots+\alpha_n f^n(x) \qquad \text{for all } x \in X$$

Let the functions $p_i : R^k \to R^1 (1 \leq i \leq k)$ be defined by

$$p_i(\boldsymbol{x}) = x_i \qquad \text{for every } \boldsymbol{x} = \langle x_1, \ldots, x_k \rangle \in R^k \tag{8-10}$$

and let 1 be the function which assigns the number 1 to each point $\boldsymbol{x}$ of R^k. Then the set

$$\mathscr{E}_k = \{p_1, \ldots, p_k, 1\} \tag{8-11}$$

is a subset of the real function algebra $\mathscr{F}(R^k, R^1)$.

If λ is any real number and $i_1, \ldots, i_k$ are any nonnegative integers, the function $m = \lambda p_1^{i_1} p_2^{i_2} \ldots p_k^{i_k}$ is a monomial in $\mathscr{E}_k$. Evidently

$$m(x) = \lambda x_1^{i_1} x_2^{i_2} \ldots x_k^{i_k} \qquad \text{for every } x = \langle x_1, \ldots x_k \rangle \in R^k$$

If $m_1, m_2, \ldots, m_j$ are any j monomials in $\mathscr{E}_k$, the function $p = m_1 + m_2 + \ldots + m_j$ is a polynomial in $\mathscr{E}_k$. Thus if $k = 3$ the function $p_3 : R^3 \to R^1$ defined by

$$p_3(\boldsymbol{x}) = 5 + 2x_1 + x_2^5 + 3x_1x_2 + 7x_1x_2^2x_3^5 \qquad \text{for every } x = \{x_1, x_2, x_3\} \in R^3$$

is a polynomial in $\mathscr{E}_3$.

Definition 8.13: *Let A be a real (complex) algebra and let $E \subset A$. Suppose that for every $a_1, a_2 \in E$ and for any two real (complex) numbers α and β*

$$\alpha a_1 + \beta a_2 \in E \tag{8-12}$$

and

$$a_1 a_2 \in E \tag{8-13}$$

Then E is called a **subalgebra** *of A.*

It now follows from Definitions 3.1 and 8.10 that every subalgebra is an algebra in its own right with the same definitions of addition, multiplication, and scalar multiplication as in the original algebra. It is easy to see that if E and F are subalgebras of an algebra A, then $E \cap F$ is also a subalgebra of A.

If A is an algebra and E is any subset of A it can easily be shown that the set P of all polynomials in E is a subalgebra of A which contains E. The set P is said to be the subalgebra of A which is *generated* by E. It is the smallest subalgebra of A which contains E. That is, every subalgebra of A which contains E also contains P. Consider, for example, the subset $\mathscr{E}_k$ of the function algebra $\mathscr{F}(R^k, R^1)$ which is defined by equation (8–11). The set $\mathscr{P}$ of all polynomials in $\mathscr{E}_k$ is a subalgebra of $\mathscr{F}(R^k, R^1)$.

Definition 8.14: *An algebra A is said to be* normed *if the linear space A has a norm and if this norm also has the property that*

$$\|ab\| \leqslant \|a\| \, \|b\| \qquad \textit{for all } a, b \in A$$

Since the absolute value of the product of two real (complex) numbers is equal to the product of their absolute values, it follows that the set of all real (complex) numbers is a normed algebra when, as is usually the case, the norm is defined as the absolute value.

It is easy to adapt the proof of part (d) of Theorem 7.7 to show that if $\{a_n\}$ and $\{b_n\}$ are sequences in a normed algebra A such that $a_n \to a$ and $b_n \to b$, then

$$a_n b_n \to ab$$

If, in addition, A is a commutative algebra with a unity, and if, for every n, a_n^{-1} and a^{-1} exist, then the proof of Theorem 7.8 can be adapted to show that

$$a_n^{-1} \to a^{-1}$$

Theorem 8.15: *Let M be a subalgebra of a real (complex) normed algebra A. Then $\bar{M}$, the closure of M, is also a subalgebra.*

Proof: According to Definition 8.13, it is sufficient to prove that if a and b are any two points of $\bar{M}$, then for any two real (complex) numbers α, β

$$\alpha a + \beta b \in \bar{M} \tag{8–14}$$

and

$$ab \in \bar{M} \tag{8–15}$$

Theorem 7.4 shows that there are sequences $\{a_n\}$ and $\{b_n\}$ in M such that $a_n \to a$ and $b_n \to b$. Since M is an algebra, it follows that for every n

$$\alpha a_n + \beta b_n \in M$$

and

$$a_n b_n \in M$$

Parts (a) and (b) of Theorem 7.7 and the preceding remarks show that

$$\alpha a_n + \beta b_n \to \alpha a + \beta b$$

and

$$a_n b_n \to ab$$

Hence, by applying Theorem 7.4 once again, we conclude that relations (8–14) and (8–15) hold.

Theorem 8.16: *Let X be a metric space and M be a real (complex) normed linear space. Suppose that $f: X \to M$ and $g: X \to M$ and that f and g are continuous at a point $p \in X$.*

(a) *If α_1 and α_2 are scalars, then $\alpha_1 f + \alpha_2 g$ is continuous at p.*

(b) *If M is also a normed algebra, then fg is continuous at p. If, in addition, M is a commutative normed algebra with a unity in which every nonzero element has a multiplicative inverse and if the zero of M does not belong to the range of g, then f/g is continuous at p.*

Proof: Part (a). The proof follows immediately from Theorems 7.7 and 8.5.

Part (b). The proof is an immediate consequence of Theorems 7.7, 7.8, and 8.5 and the remarks following Definition 8.14.

The following corollary follows directly from the theorem and Definitions 3.3 and 8.13.

Corollary: *Let X be any set and M be a real (complex) normed linear space. Then $\mathscr{C}(X, M)$, the family of all* **continuous** *functions from X into M, is a linear subspace of $\mathscr{F}(X, M)$, the family of all functions from X into M. If, in addition, M is a real (complex) normed algebra, then $\mathscr{C}(X, M)$ is a subalgebra of $\mathscr{F}(X, M)$.*

It is clear that, if $\boldsymbol{f}$ is a function from a set X into R^k, there are k functions, say $f_1, \ldots, f_k$, each of which maps X into R^1 such that, at each point $x \in X$, $\boldsymbol{f}(x) = \langle f_1(x), \ldots, f_k(x) \rangle$. The functions $f_1, \ldots, f_k$ are said to be the components of $\boldsymbol{f}$ and the notation $\boldsymbol{f} = \langle f_1, \ldots, f_k \rangle$ is used to indicate this relationship.

Theorem 8.17: (a) *A function $\boldsymbol{f}: X \to R^k$ is continuous at a point $p \in X$ if and only if its coordinates $f_1, \ldots, f_k$ are continuous at p.*

(b) *If $\boldsymbol{f}: X \to R^k$ and $\boldsymbol{g}: X \to R^k$ are continuous at a point $p \in X$, then the function $\boldsymbol{f} \cdot \boldsymbol{g}: X \to R^1$ is continuous at p.*

Proof: Part (a). The inequalities

$$|f_j(q)-f_j(p)| \leqslant |\boldsymbol{f}(q)-\boldsymbol{f}(p)| = \left(\sum_{i=1}^{k} \left|f_i(q)-f_i(p)\right|^2\right)^{1/2} \qquad 1 \leqslant j \leqslant k$$

show that, if $|\boldsymbol{f}(q)-\boldsymbol{f}(p)| < \epsilon$, then, for $1 \leqslant j \leqslant k$, $|f_j(q)-f_j(p)| < \epsilon$ and that, if, for $1 \leqslant j \leqslant k$, $|f_j(q)-f_j(p)| < \epsilon/\sqrt{k}$, then $|\boldsymbol{f}(q)-\boldsymbol{f}(p)| < \epsilon$.

Part (b). The second part of the theorem now follows from the first part and Theorem 8.16.

It is not difficult to find continuous functions on R^k. For example, the functions $p_i : R^k \to R^1 (1 \leqslant i \leqslant k)$ defined by equation (8–10) are continuous on R^k. To see this we merely note that the inequalities

$$|p_i(\boldsymbol{y})-p_i(\boldsymbol{x})| = |y_i - x_i| \leqslant |\boldsymbol{y}-\boldsymbol{x}| \qquad 1 \leqslant i \leqslant k$$

imply that

$$|p_i(\boldsymbol{y})-p_i(\boldsymbol{x})| < \epsilon \qquad \text{whenever } |\boldsymbol{y}-\boldsymbol{x}| < \delta = \epsilon$$

Since the constant functions are continuous, this shows that the set $\mathscr{E}_k$ defined by equation (8–11) is a subset of the function algebra $\mathscr{C}(R^k, R^1)$. Since the set $\mathscr{P}$ of all polynomials in $\mathscr{E}_k$ is the smallest subalgebra of $\mathscr{F}(R^k, R^1)$ which contains $\mathscr{E}_k$, we conclude that $\mathscr{P} \subset \mathscr{C}(R^k, R^1)$. This shows that all the polynomials in $\mathscr{E}_k$ are continuous functions.

Suppose V is a normed linear space and $u, v \in V$. Then the triangle inequality $\|u\| \leqslant \|u-v\| + \|v\|$ shows (after interchanging u and v)

$$\big|\, \|u\| - \|v\| \,\big| \leqslant \|u-v\|$$

Hence the function $f: V \to R^1$ defined by

$$f(u) = \|u\| \qquad \text{for all } u \in V$$

is a continuous mapping of V into R^1 (with the usual metric).

If $f: X \to V$ is a continuous mapping of the metric space X into V and g is defined on X by setting $g(x) = \|f(x)\|$ for all $x \in X$, it follows from Theorem 8.6 that the mapping $g: X \to R^1$ is continuous on X.

Definition 8.18: *A function from a set X to a metric space $\langle Y, d\rangle$ is said to be* **bounded** *if its range is bounded. The set of all bounded functions from X into Y is denoted by $\mathscr{B}(X, Y)$. If Y is the real or complex numbers with the usual metric, the simpler notation $\mathscr{B}(X)$ is sometimes used.*

Thus, if the function $f: X \to Y$ is bounded, there is a finite number M such that $d(f(X)) \leqslant M$. Hence, for every $x, y \in X$, $d(f(x), f(y)) \leqslant M$. If Y is a normed linear space, then this means that

$$\operatorname*{lub}_{x,y \in X} \|f(x)-f(y)\| \leqslant M$$

Hence, for any x and any fixed $y_0 \in X$,

$$\|f(x)\| \leqslant \|f(x)-f(y_0)\|+\|f(y_0)\| \leqslant M+\|f(y_0)\|$$

Since $\|f(y_0)\|$ is a finite number, so is $P=M+\|f(y_0)\|$. Thus, for every $x \in X$, there is a finite number P such that

$$\|f(x)\| \leqslant P$$

Therefore

$$\operatorname*{lub}_{x \in X} \|f(x)\| \leqslant P \tag{8–16}$$

On the other hand, if equation (8–16) holds, then, for any $x, y \in X$,

$$\|f(x)-f(y)\| \leqslant \|f(x)\|+\|f(y)\| \leqslant 2P$$

Hence,

$$\operatorname*{lub}_{x,\, y \in X} \|f(x)-f(y)\| \leqslant 2P$$

which shows that f is bounded.

Thus, we see that *a function f from a set X to a normed linear space is bounded if and only if there is a finite number P such that*

$$\operatorname*{lub}_{x \in X} \|f(x)\| \leqslant P$$

For real or complex valued functions, where the norm is the absolute value, this definition reduces to the usual concept of boundedness.

For example, suppose $f: (0, 1) \to R^1$ and, for each $x \in (0, 1)$, $f(x)=1/x$. Then $\operatorname*{lub}_{x \in (0,1)} |1/x|=+\infty$ and f is not bounded. But if $g: (0, 1) \to R^1$ is defined by $g(x)=1/(1+x)$ for every $x \in (0, 1)$, then $\operatorname*{lub}_{x \in (0,1)} |1/(1+x)|=1$. Hence g is bounded.

Let M be a real (complex) normed linear space and let X be any set. If $f: X \to M$ and $g: X \to M$ and α and β are any real (complex) numbers, then for every $x \in X$

$$\|(\alpha f+\beta g)(x)\|=\|\alpha f(x)+\beta g(x)\|\leq\|\alpha f(x)\|+\|\beta g(x)\|=|\alpha|\ \|f(x)\|+|\beta|\ \|g(x)\|$$

Hence the function $\alpha f+\beta g$ is bounded whenever f and g are. If in addition M is a normed algebra, then for every $x \in X$

$$\|(fg)(x)\|=\|f(x)g(x)\|\leq\|f(x)\|\ \|g(x)\|$$

Hence

$$\operatorname*{lub}_{x\in X}\|(fg)(x)\|\leq\operatorname*{lub}_{x\in X}\|f(x)\|\,\|g(x)\|\leq\operatorname*{lub}_{x\in X}\|f(x)\|\operatorname*{lub}_{y\in X}\|g(y)\|$$

Thus the function fg is bounded whenever f and g are bounded. We therefore arrive at the following conclusion.

Theorem 8.19: *If X is any set and M is a normed linear space, then $\mathscr{B}(X, M)$ is a linear subspace of the linear space $\mathscr{F}(X, M)$. If in addition M is a normed algebra, then $\mathscr{B}(X, M)$ is a subalgebra of the algebra $\mathscr{F}(X, M)$.*

The set of all continuous bounded functions from a set X into a set Y is denoted by $\mathscr{C}^{\infty}(X, Y)$. Clearly

$$\mathscr{C}^{\infty}(X, Y)=\mathscr{C}(X, Y)\cap\mathscr{B}(X, Y)$$

Since the intersection of two linear subspaces is also a linear subspace and the intersection of two subalgebras is also a subalgebra, it follows that *if M is a normed linear space and X is any set, then $\mathscr{C}^{\infty}(X, M)$ is a linear space. If in addition M is a normed algebra, then $\mathscr{C}^{\infty}(X, M)$ is an algebra.*

There are some interesting relations between continuity and compactness.

Theorem 8.20: *If f is a continuous function from a compact metric space X into a metric space Y, then $f(X)$ is a compact subset of Y.*

Proof: Choose any open cover $\{V_\alpha|\alpha\in A\}$ of $f(X)$. Then

$$f(X)\subset\bigcup_{\alpha\in A}V_\alpha$$

It follows from tables 4–I and 5–II that

$$X\subset f^{-1}[f(X)]\subset f^{-1}\left[\bigcup_{\alpha\in A}V_\alpha\right]=\bigcup_{\alpha\in A}f^{-1}(V_\alpha)$$

Now Theorem 8.7 shows that, for each $\alpha \in A$, the set $f^{-1}(V_\alpha)$ is open. Hence the collection $\{f^{-1}(V_\alpha) | \alpha \in A\}$ is an open cover of X and since X is compact there is a finite subset of A, say $\{\alpha_1, \ldots, \alpha_n\}$, such that

$$X \subset f^{-1}(V_{\alpha_1}) \cup \ldots \cup f^{-1}(V_{\alpha_n})$$

Table 4–I shows that

$$\begin{aligned} f(X) &\subset f[f^{-1}(V_{\alpha_1}) \cup \ldots \cup f^{-1}(V_{\alpha_n})] \\ &= f(f^{-1}(V_{\alpha_1})) \cup \ldots \cup f(f^{-1}(V_{\alpha_n})) \subset V_{\alpha_1} \cup \ldots \cup V_{\alpha_n} \end{aligned}$$

and this shows that $f(X)$ is compact.

The following corollaries give some useful consequences of this theorem.

Corollary 1: *If $f: X \to Y$, f is continuous on X, X is a compact metric space, and Y is a metric space, then f is bounded and its range is a closed subset of Y.*

Proof: Theorem 8.20 shows us that the range of f, $f(X)$, is compact and Theorem 6.22 shows that, as a consequence, it is closed and bounded.

The remarks following Definition 8.18 show that, if in addition Y is a normed linear space, there is a finite number M such that $\sup_{x \in X} ||f(x)|| \leq M$. The next corollary shows that this result is even more significant when the normed linear space is R^1.

Corollary 2: *Let f be a real-valued continuous function defined on the compact metric space X and let $U = \sup_{x \in X} f(x)$ and $L = \inf_{x \in X} f(x)$. Then there are points p and q of X such that $f(p) = U$ and $f(q) = L$.*

Proof: The preceding corollary and the remarks following it imply that $f(X) = \{f(x) | x \in X\}$ is closed and that there is a finite number M such that $\operatorname{lub}_{x \in X} |f(x)| \leq M$. Since, for every $x \in X$,

$$f(x) \leq |f(x)| \leq M \qquad \text{and} \qquad -f(x) \leq |f(x)| \leq M$$

we see that $\sup_{x \in X} f(x) \leq M$ and $-\inf_{x \in X} f(x) = \sup_{x \in X} (-f(x)) \leq M$; that is, $-M \leq \inf_{x \in X} f(x)$. Since M is finite, these inequalities show that $f(X)$ is bounded above and below. Hence it follows from Theorem 6.14 that $f(X)$ contains its least upper bound U and its greatest lower bound L and this proves the assertion.

Thus the corollary says that there are points p and q in X such that $f(q) \leqslant f(x) \leqslant f(p)$ for all $x \in X$.

These two corollaries to Theorem 8.20 would not remain true if the compactness of the metric space X were replaced by some weaker condition. In order to show this, suppose that the metric space X is a noncompact subset of the real numbers with the usual metric and that $Y = R^1$. According to the Heine-Borel theorem (Theorem 6.31), X is either not closed or not bounded.

Suppose first that X is not closed and let x_0 be a limit point of X which is not a point of X. We can define functions $f_1 : X \to Y$ and $f_2 : X \to Y$ by

$$\left.\begin{aligned} f_1(x) &= \frac{1}{x - x_0} \\ f_2(x) &= \frac{1}{1 + (x - x_0)^2} \end{aligned}\right\} \qquad \text{for every } x \in X$$

Since the denominators of f_1 and f_2 never vanish on X (since $x_0 \notin X$), Theorem 8.16 shows that f_1 and f_2 are continuous. Now $|f_1(x)| = 1/|x - x_0|$ and $\operatorname*{lub}_{x \in X} f_2(x) = 1$. Since x_0 is a limit point of X, for any positive number ϵ, we can find an $x \in X$ for which $|x - x_0| < \epsilon$. This shows that $\operatorname*{lub}_{x \in X} |f_1(x)| = +\infty$; that is, f_1 is not bounded. Since $x_0 \notin X$, it is clear that $f_2(x) < 1$ for every $x \in X$; thus $f_2(x) \neq \operatorname*{lub}_{x \in X} f_2(x)$ for any $x \in X$.

Now suppose that X is not bounded. This means that, for any finite number M, there is a point $x \in X$ such that $|x| \geqslant M$. In this case let us define the functions $f_1 : X \to Y$ and $f_2 : X \to Y$ by

$$\left.\begin{aligned} f_1(x) &= x \\ f_2(x) &= \frac{x^2}{1 + x^2} \end{aligned}\right\} \qquad \text{for every } x \in X$$

Since the denominators of f_1 and f_2 also do not vanish, Theorem 8.16 again shows that f_1 and f_2 are continuous. Clearly,

$$\operatorname*{lub}_{x \in X} |f_1(x)| = +\infty$$

and

$$\operatorname*{lub}_{x \in X} f_2(x) = 1$$

Also,

$$f_2(x) < 1 \qquad \text{for every } x \in X$$

Thus, f_1 is not bounded and f_2 never attains its maximum.

In any case, we have shown that there exists on any *noncompact* set of real numbers a continuous function which is not bounded and a continuous function which does not attain its maximum.

Theorem 8.21: *If $f: X \to Y$ is a continuous bijection from the compact metric space X to the metric space Y, then the inverse mapping f^{-1} (which is defined* **on** *Y) is continuous.*

Proof: Since $(f^{-1})^{-1} = f$, we see from Theorem 8.7 that the conclusion of the theorem will follow if we can prove that $f(V)$ is an open set in Y for every open set V in X.

Therefore suppose V is any open set in X. Then its complement V^c is closed and so Theorem 6.23 implies that V^c is compact. If g is the restriction of f to V^c, then corollary 2 of Theorem 8.7 shows that g is continuous. Hence it follows from Theorem 8.20 that $g(V^c)$ is compact and, since $f(V^c) = g(V^c)$, we conclude that $f(V^c)$ is a compact subset of Y. Theorem 6.22 therefore shows that $f(V^c)$ is closed. It is not hard to verify that, since f is a bijection, $f(V^c) = [f(V)]^c$. Hence we conclude from Theorem 6.10 that $f(V)$ is open.

If the space X is not compact, we can show by example that the theorem does not hold. Let C be the unit circle and let the function $f: [0, 2\pi) \to C$ be defined by

$$\mathbf{f}(t) = \langle \cos t, \sin t \rangle \qquad \text{for all } t \in [0, 2\pi)$$

Since the sine and cosine are continuous functions, Theorem 8.17 shows that $\boldsymbol{f}$ is also. Clearly $\boldsymbol{f}$ is a bijection but its inverse $\boldsymbol{f}^{-1}$ is not continuous at the point $\langle 1, 0 \rangle$.

Definition 8.22: *If $\langle X, d \rangle$ and $\langle Y, d' \rangle$ are metric spaces and $f: X \to Y$, the function f is said to be* **uniformly continuous** *on X if, for every $\epsilon > 0$, there exists* **a single** *$\delta > 0$ such that*

$$f(B(p\,;\delta)) \subset B(f(p);\,\epsilon) \qquad \textit{for } \textbf{every } p \in X$$

There is one very important difference between the concepts of continuity and uniform continuity. Continuity is defined at each point of a set and, if a function is continuous at some point of the set, then, for each $\epsilon > 0$, there is a

$\delta > 0$ depending on ϵ *and on the point* p which satisfies the requirements of the definition. On the other hand, if a function is uniformly continuous on a set, then, for each $\epsilon > 0$, one can find a single number δ *depending only on* ϵ which satisfies the requirements of the definition at every point of the set. Of course, uniformly continuous functions are continuous at every point of their domain.

Constant mappings and natural injections are uniformly continuous functions.

For any nonempty subset A of a metric space $\langle X, d\rangle$, the mapping $f: X \to R^1$ defined by

$$f(x) = d(x, A) \qquad \text{for all } x \in X$$

is uniformly continuous. This follows immediately from the inequality (6–14).

Theorem 8.23: *If f is a continuous mapping of the compact metric space $\langle X, d\rangle$ into the metric space $\langle Y, d'\rangle$, then f is uniformly continuous on X.*

Proof: Fix $\epsilon > 0$. Since f is continuous on X, we can find a positive number $\lambda(p)$ for each point $p \in X$ such that

$$f[B(p;\, 2\lambda(p))] \subset B(f(p);\, \epsilon/2)$$

Since X is compact and $\{B(p;\, \lambda(p)) \mid p \in X\}$ is an open cover of X, we can find finitely many points of X, say $p_1, p_2, \ldots, p_n$, such that

$$X \subset \bigcup_{i=1}^{n} B(p_1;\, \lambda(p_i))$$

Set $\delta = \min\limits_{1 \leqslant i \leqslant n} \lambda(p_i)$. Clearly $\delta > 0$. Now let q be any point of X. Then, for some integer j, such that $1 \leqslant j \leqslant n$,

$$q \in B(p_j;\, \lambda(p_j))$$

If $y \in B(q;\, \delta)$, then

$$d(y, p_j) \leqslant d(y, q) + d(q, p_j) < \delta + \lambda(p_j) \leqslant 2\lambda(p_j)$$

Hence,

$$B(q;\, \delta) \subset B(p_j;\, 2\lambda(p_j))$$

and

$$f(B(q;\, \delta)) \subset f(B(p_j;\, 2\lambda(p_j))) \subset B(f(p_j);\, \epsilon/2) \tag{8–17}$$

So that in particular,

$$f(q) \in B(f(p_j);\, \epsilon/2)$$

If $y \in B(f(p_j); \epsilon/2)$, then,

$$d'(x, f(q)) \leqslant d'(x, f(p_j)) + d'(f(p_j), f(q)) < \epsilon/2 + \epsilon/2 = \epsilon$$

which shows that

$$B(f(p_j); \epsilon/2) \subset B(f(q); \epsilon)$$

Upon combining this with equation (8–17) we conclude that

$$f(B(q; \delta)) \subset B(f(q); \epsilon)$$

and, since q was any point of X, this proves the theorem.

For functions defined on segments of the real line, the concepts of right and left hand limits are sometimes useful.

Definition 8.24: *Let f be a function on the segment (a, b) and let g_x be the restriction of f to (x, b) for $a \leqslant x < b$. If $\lim_{t \to x} g_x(t)$ exists, it is called the* **right hand limit of** f **at** x *and is denoted by $f(x+)$. The* **left hand limit** *is defined in a similar way by using the restriction of f to (a, x).*

It is clear that $\lim_{t \to x} f(t)$ exists if and only if $f(x+) = f(x-) = \lim_{t \to x} f(t)$. When these concepts are used it is possible to characterize two types of discontinuities for functions defined on a segment of the real numbers.

Definition 8.25: *If $f\colon (a, b) \to X$, f is discontinuous at $x \in (a, b)$, and both $f(x+)$ and $f(x-)$ exist, then f is said to have a* **discontinuity** *of the* **first kind** *at x. All other discontinuities of f are called discontinuities of the* **second kind.**

Evidently a function can have a discontinuity of the first kind in the following two ways:

(a) $f(x+) \neq f(x-)$

(b) $f(x+) = f(x-) \neq f(x)$

The function defined in the example following Theorem 8.7 has a discontinuity of the first kind (see fig. 8–3). Dirichlet's function (see example following corollary 2 of Theorem 8.7) has a discontinuity of the second kind at every point.

As in the case of sequences we introduce the concept of monotonic functions.

355–525 O–70—11

Definition 8.26: *A function $f: (a, b) \to R^1$ is said to be* **monotonically increasing** *on (a, b) if $a < x < y < b$ implies that $f(x) \leqslant f(y)$. Reversing the latter inequality yields the definition of a* **monotonically decreasing** *function.*

Theorem 8.27: *If f is monotonically increasing on (a, b), then, for every $x \in (a, b)$,*

(a) $f(x-) = \sup_{s \in (a, x)} f(s) \leqslant f(x) \leqslant \inf_{s \in (x, b)} f(s) = f(x+)$

(b) $f(x+) \leqslant f(y-)$, *for $a < x < y < b$*

Proof: Part (a). Let g_x be the restriction of f to (a, x). Since f is monotonically increasing, $E_x = \{f(t) | t \in (a, x)\}$ is bounded above by $f(x)$. The least upper bound A of this set therefore exists, and $A \leqslant f(x)$. Fix $\epsilon > 0$. Evidently $A - \epsilon$ is not an upper bound of E_x. Hence there exists a $\delta > 0$ such that $a < x - \delta < x$ and

$$A - \epsilon < f(x - \delta) \leqslant A \tag{8–18}$$

Since f is monotonic it follows that $f(x-\delta) \leqslant f(t) \leqslant A$ for all $t \in (x-\delta, x)$. Combining this with equation (8–18) shows that $f(t) \in (A-\epsilon, A]$ whenever $t \in (x-\delta, x)$. Hence,

$$f((x-\delta, x)) \subset (A-\epsilon, A] \subset (A-\epsilon, A+\epsilon) \tag{8–19}$$

Since $(x-\delta, x) \subset (a, x)$,

$$g_x((x-\delta, x)) = f((x-\delta, x)) \tag{8–20}$$

Evidently

$$(x-\delta, x) = (x-\delta, x+\delta) \cap (a, x) - \{x\} \tag{8–21}$$

It now follows from equations (8–19), (8–20), and (8–21) that

$$g_x((x-\delta, x+\delta) \cap (a, x) - \{x\}) \subset (A-\epsilon, A+\epsilon)$$

Since ϵ was arbitrary, this implies that

$$f(x-) = \lim_{t \to x} g_x(t) = A = \text{lub } E_x = \sup_{s \in (a, x)} f(s)$$

An almost identical argument proves the other half of part (a).

Part (b). This follows almost immediately from part (a) since

$$f(x+) = \inf_{t \in (x, b)} f(t) = \inf_{t \in (x, y)} f(t)$$

and

$$f(y-) = \sup_{t \in (a, y)} f(t) = \sup_{t \in (x, y)} f(t)$$

Corollary: *A monotonic function cannot have a discontinuity of the second kind.*

Theorem 8.28: *The set E of points at which a monotonic function f defined on (a, b) is discontinuous is countable.*

Proof: Let f be increasing. Theorem 8.27(a) implies that if $x \in E$, then $f(x-) < f(x+)$. Hence, there exists a rational number $r(x)$ such that

$$f(x-) < r(x) < f(x+)$$

Since Theorem 8.27(b) shows that $x < y$ implies $f(x+) \leq f(y-)$, we conclude that $x \neq y$ implies $r(x) \neq r(y)$. In this way we associate a rational number with each $x \in E$ and no two members of E are associated with the same rational number. This defines an injective mapping from E to the set of rationals and hence (since every mapping onto its range is surjective) a bijection from E to a subset of the rationals which according to corollary 1 to Theorem 4.16 must be countable since the rationals are. Hence E is countable.

In discussing the convergence of sequences we have shown how we could extend the notion of convergence to include certain types of divergent sequences (improper convergence) by extending the notion of ball about a point to include balls about the points $+\infty$ and $-\infty$ in the extended real number system. In view of the close relation shown in Theorem 8.2 between the existence of limits of functions and the convergence of sequences, it is not surprising that we do the same thing with the limits of a function. Thus we make the following definition.

For the purposes of this definition let us temporarily call the half-open intervals $[-\infty, a)$ and $(b, +\infty]$ balls for all finite real numbers a and b.

Definition 8.29: *Let E be any set of finite real numbers and let $f: E \to R^1$. If p is any extended real number such that, for every ball B about p, $B \cap E - \{p\} \neq \emptyset$, then we write*

$$f(x) \to q \text{ as } x \to p$$

if there exists an extended real number q with the property that, for every ball B' about q, there exists a ball B about p such that

$$f(B \cap E) \subset B'$$

CHAPTER 9

Cauchy Completeness of a Metric Space

This chapter begins by introducing two types of mappings between metric spaces. The first and most general of these, the homeomorphism, preserves the topological properties of the spaces and the second, the isometry, preserves the metric properties. A study of these mappings helps to give a certain amount of insight into the nature of both the topological and the purely metric properties of metric spaces. We then turn to a discussion of an important type of sequence called the Cauchy sequence. This discussion leads, in a natural way, to the purely metric concept of a complete metric space. At this point the concept of isometry is used to show that every metric space can in a certain sense be embedded in a complete metric space.

If there is a mapping f between two metric spaces which, aside from putting these two spaces into one-to-one correspondence, preserves the distance between the corresponding elements, then as metric spaces they will have the same properties. Thus, if a particular subset of one of these spaces is bounded, the image of this subset under f will also be a bounded subset in the other metric space. If one of these spaces has any property that can be defined in an abstract metric space the other will have it also. Before defining this type of mapping which preserves metric properties, we shall define a more general mapping which preserves topological properties.

Definition 9.1: *Two metric spaces $\langle X, d\rangle$ and $\langle Y, d'\rangle$ are said to be* **homeomorphic** *if there is a bijection f from X to Y such that both f and f^{-1} are continuous. The mapping f is then called a* **homeomorphism.**

It can be seen, for example, that the mapping $f: R^1 \to R^1$ defined by

$$f(x) = x^3 \qquad \text{for every } x \in R^1$$

is a homeomorphism.

Suppose $f: X \to Y$ is a homeomorphism. From Theorem 4.8(c) we see that, since f is bijective, the mapping induced by f is a bijective mapping from the collection of all subsets of X to the collection of all subsets of Y. If E is any open subset of Y, Theorem 8.7 shows that $f^{-1}(E)$ is an open subset of X. It was pointed out in chapter 4 that $f(f^{-1}(E))=E$. Hence we conclude that, for every open subset E of Y, there is an open subset G of X such that $E=f(G)$. On the other hand, since $(f^{-1})^{-1}=f$ and since f^{-1} is continuous, Theorem 8.7 implies that, if G is any open subset of X, there is an open subset E of Y such that $E=f(G)$. It follows from these remarks that the mapping induced by f puts the open subsets of X into one-to-one correspondence with the open subsets of Y.

A little thought now shows that, if the metric spaces $\langle X, d\rangle$ and $\langle X, d'\rangle$ are homeomorphic, they must also have the same topological properties. Since many of the properties we have discussed in connection with metric spaces were in fact topological, this type of mapping is rather significant.

For example, from the discussion of topological spaces in chapter 6, we see that, if x is an adherence point of a set $E \subset X$, then $f(x)$ is an adherence point of the set $f(E) \subset Y$.

Suppose that d_1 and d_2 are two distances defined on a set X. Then, as mentioned in chapter 6, $\langle X, d_1\rangle$ and $\langle X, d_2\rangle$ are distinct metric spaces even though they both have the same underlying set. If i is the identity map on X, then i can be considered a mapping from the metric space $\langle X, d_1\rangle$ onto the metric space $\langle X, d_2\rangle$. Clearly i is a bijection and is its own inverse. Hence the inverse of this mapping is i, now regarded as being a function from the metric space $\langle X, d_2\rangle$ onto the metric space $\langle X, d_1\rangle$. Thus if i is both a continuous mapping from the metric space $\langle X, d_1\rangle$ into the metric space $\langle X, d_2\rangle$ and a continuous mapping from the metric space $\langle X, d_2\rangle$ into the metric space $\langle X, d_1\rangle$, it must be a homeomorphism. In this case the distances d_1 and d_2 are said to be (*topologically*) *equivalent* since they both determine the same topology for X. If i is both a *uniformly* continuous mapping from $\langle X, d_1\rangle$ onto $\langle X, d_2\rangle$ and a *uniformly* continuous mapping from $\langle X, d_2\rangle$ onto $\langle X, d_1\rangle$, then the metrics d_1 and d_2 are said to be *uniformly equivalent*. A sufficient condition that d_1 and d_2 be uniformly equivalent metrics on X is that there exist two (finite) real numbers $a>0$ and $b>0$ such that

$$ad_1(p, q) \leq d_2(p, q) \leq bd_1(p, q) \qquad \text{for all } p, q \in X$$

As pointed out in chapter 6 any two of the three metrics defined by equations (6–5) to (6–7) on the product space of two metric spaces satisfy these inequalities. It was also pointed out in chapter 6 that the Euclidean space R^k and the product space of the Euclidean spaces R^s and R^{k-s} are different metric spaces even though they have the same underlying set. However, the preceding discussion shows that these two metric spaces have uniformly equivalent metrics.

Suppose that f is a homeomorphism from a discrete metric space X onto a metric space Y. The first corollary to Theorem 6.12 shows that every subset of X is open. Since every subset of Y is also the image under f of a subset of X, we see that every subset of Y is also an open set. On the other hand, if $\langle Y, d\rangle$ is any metric space having the property that all of its subsets are open, we can define a discrete metric d' on Y as indicated in the examples following Definition 6.1. Then $\langle Y, d'\rangle$ is a discrete metric space. It is not hard to verify that the identity map of Y is then a homeomorphism of the metric space $\langle Y, d'\rangle$ to the metric space $\langle Y, d\rangle$. Hence we conclude that a metric space is homeomorphic to a discrete space if and only if all its subsets are open sets.

It is standard practice to call any metric space which is homeomorphic to a discrete metric space a discrete space. This convention is actually an incorrect use of the language since a metric space which is homeomorphic to a discrete space may have a distance between points which is different from zero or one.

If a metric space X is both homeomorphic to a discrete space and compact, then it must be finite. In order to see this, notice that each one-element subset of X is open. Hence the family $\Omega=\{\{x\} \mid x \in X\}$ is an open cover of X and Ω can only have a finite subcover if X has finitely many points.

Conversely, if X is a finite metric space, then it must be both compact and homeomorphic to a discrete metric space. In order to see this, notice that each one-element set of X is certainly closed. Since every subset of X is a finite union of such sets, corollary 2 to Theorem 6.12 shows that every subset of X is closed. Hence, Theorem 6.10 shows that every subset of X is open. Thus X is homeomorphic to a discrete space. Finally, since there are only a finite number of open sets, every open cover of X is finite and so X is trivially compact. Thus a metric space is finite if and only if it is both compact and homeomorphic to a discrete metric space.

Definition 9.2: *Two metric spaces $\langle X, d\rangle$ and $\langle Y, d'\rangle$ are said to be* **isometric** *if there is a bijection f from X to Y such that, for any two points $x, y \in X$,*

$$d(x, y)=d'(f(x), f(y))$$

The mapping f is then called an **isometry**. *If E is any subset of X, $f(E)$ is called the* **isometric image** *of E (under f).*

It is easily seen that an isometry f and its inverse f^{-1} both satisfy the conditions of Definition 8.3 with $\delta=\epsilon$. This shows that an isometry is also a homeomorphism (which is as it should be since isometries preserve metric properties and homeomorphisms preserve topological properties, and every topological property is also a metric property).

Let the two metric spaces in Definition 9.2 both be R^1, and let f be the mapping which takes each real number y into the real number $y+c$ where c is some fixed real number. Then f is an isometry.

An example of an isometry that has already been discussed in a much less formal way is the mapping f from the Euclidean space R^2 to the set of all complex numbers (with the distance defined in the usual way) which associates with each point $\boldsymbol{x}=\langle x_1, x_2\rangle$ of R^2 the complex number $z=x_1+ix_2$. It has already been established[26] that if $z_1=f(\boldsymbol{x}_1)$ and $z_2=f(\boldsymbol{x}_2)$, then $|z_1-z_2| = |\boldsymbol{x}_1-\boldsymbol{x}_2|$ and the mapping f is clearly a bijection. We have in fact considered these two spaces to be the same. This can often be done when two metric spaces are isometric to one another.

If X and X' are any two isometric metric spaces, then for any theorem proved in X that involves *only distances* between points of X there is a corresponding theorem that holds in the metric space X'.

If $\langle X, d\rangle$ is a metric space and f is a bijection from an arbitrary set X' to X, then it is not hard to see that we can define a distance d' on X' by

$$d'(x, y)=d(f(x), f(y)) \qquad \text{for every } x, y\in X'$$

The bijection f is then an isometry from the metric space $\langle X', d'\rangle$ to the metric space $\langle X, d\rangle$. The distance d' is said to be *transported* from X to X' by f.

Let $\bar{R}$ be the set of *extended* real numbers. The function $f: \bar{R}\to[-1, 1]$ defined by

$$f(x)=\begin{cases} \dfrac{x}{1+|x|} & -\infty<x<+\infty \\ 1 & x=+\infty \\ -1 & x=-\infty \end{cases}$$

is a bijection from $\bar{R}$ onto $[-1, 1]$.

[26] See discussion following Definition 6.1.

Since $[-1, 1]$ is a metric space when the distance is defined (in the usual way) in terms of the absolute value, we can use the process described in the preceding paragraph to define a distance $\bar{d}$ on $\bar{R}$ by

$$\bar{d}(x, y) = |f(x) - f(y)| \qquad \text{for all } x, y \in \bar{R}$$

The metric space $\langle \bar{R}, \bar{d} \rangle$ is sometimes referred to as the *extended real line.* Notice that if $x \geq 0$, $\bar{d}(x, +\infty) = 1/(1+|x|)$ and if $x \leq 0$ then $\bar{d}(-\infty, x) = 1/(1+|x|)$. Hence $\bar{d}(x, +\infty) < \epsilon$ and $x \geq 0$ implies $x > 1/\epsilon - 1$ and $\bar{d}(-\infty, x) < \epsilon$ and $x \leq 0$ implies $x < 1 - 1/\epsilon$. It is now easy to see that *a sequence of points of R^1 converges improperly if and only if it converges to $+\infty$ or $-\infty$ as a sequence in the metric space $\langle \bar{R}, \bar{d} \rangle$.*

We now introduce the important concept of Cauchy sequence.

Definition 9.3: *A sequence $\{x_n\}$ in a metric space $\langle X, d \rangle$ is called a* **Cauchy sequence** *if, for every $\epsilon > 0$, there is an integer N such that, for every m and n greater than N,*

$$d(x_n, x_m) < \epsilon$$

If d_1 and d_2 are topologically equivalent metrics on a set X, a sequence $\{p_n\}$ in X may be a Cauchy sequence in the metric space $\langle X, d_1 \rangle$ but not in the metric space $\langle X, d_2 \rangle$. However, it is easy to show that if d_1 and d_2 are uniformly equivalent, then both metric spaces have the same Cauchy sequences.

Suppose $\{p_n\}$ is a sequence in the metric space $\langle X, d \rangle$. For every positive integer N, set

$$E_N = \{p_n \mid n \geq N\}$$

It follows from a comparison of Definitions 6.20 and 9.3 that $\{p_n\}$ is a Cauchy sequence if and only if

$$\lim_{N \to \infty} d(E_N) = 0$$

The next theorem will be useful in our study of Cauchy sequences.

Theorem 9.4: *For any subset E of the metric space $\langle X, d \rangle$, $d(\bar{E}) = d(E)$.*

Proof: Since $E \subset \bar{E}$, it is clear that

$$d(E) \leq d(\bar{E})$$

To show that the opposite inequality holds, let $\epsilon > 0$ be given and let p and q be any two points of $\bar{E}$. Since p and q are adherence points of E, we can find points p^* and q^* in E such that $p^* \in B(p; \epsilon)$ and $q^* \in B(q; \epsilon)$. Then,

$$d(p, q) \leqslant d(p, p^*) + d(p^*, q^*) + d(q^*, q) < 2\epsilon + d(p^*, q^*) \leqslant 2\epsilon + d(E)$$

Hence, $2\epsilon + d(E)$ is an upper bound of $\{d(p, q) | p, q \in \bar{E}\}$, and

$$d(\bar{E}) = \operatorname*{lub}_{p, q \in \bar{E}} d(p, q) \leqslant 2\epsilon + d(E)$$

Since this must be true for every $\epsilon > 0$, we conclude that

$$d(\bar{E}) \leqslant d(E)$$

Therefore,

$$d(\bar{E}) = d(E)$$

Theorem 9.5: *Every convergent sequence is a Cauchy sequence.*

Proof: [27] Suppose $\{p_n\}$ converges to p in the metric space $\langle X, d \rangle$. Then for every $\epsilon > 0$, we can find an integer N such that $n \geqslant N$ implies $d(p_n, p) < \epsilon/2$. Hence, for $m, n \geqslant N$,

$$d(p_m, p_n) \leqslant d(p_m, p) + d(p, p_n) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Definition 9.6: *A metric space is said to be* **complete** *if every Cauchy sequence in this space converges. If a normed linear space is also a complete metric space, it is called a* **Banach space.** *If a normed algebra is a complete metric space, it is called a* **Banach algebra.**

In a complete metric space, we can assert whether or not a given sequence converges without specific knowledge of its limit. For example, in the case of infinite series (which we shall see are nothing more than sequences in normed linear spaces), all the tests for convergence depend on completeness.

In complete metric spaces, then, one has only to establish whether or not

[27] Notice that, in proving this theorem, it is not necessary to require the distance between points in the metric space X to be finite. This fact will be used in chapter 11.

a given sequence is a Cauchy sequence to find whether or not it converges. This is known as the Cauchy criterion of convergence.

The remarks following Definition 9.3 show that, like boundedness, the definition of a Cauchy sequence can be expressed in terms of the diameters of certain sets. Since we have already pointed out that boundedness involves the metric too intimately to be a topological property (see comments following Theorem 6.22), it is not too surprising that completeness does also.

The next theorem gives a useful characterization of completeness which will be used to show that two very important classes of metric spaces are complete.

Theorem 9.7: *A metric space $\langle X, d\rangle$ is complete if and only if* $\bigcap_{i=1}^{\infty} F_i \neq \emptyset$ *for every nested sequence of nonempty closed and bounded sets* $\{F_i\}$ *such that* $\lim_{n\to\infty} d(F_n) = 0$.

Proof: Let $\langle X, d\rangle$ be complete and let $\{F_i\}$ be a nested sequence of nonempty closed and bounded sets such that $\lim_{n\to\infty} d(F_n) = 0$. For each integer n choose a point $p_n \in F_n$. Since $F_m \subset F_n$ for $m \geqslant n$, we see that $p_m \in F_n$ whenever $m \geqslant n$. Let $\epsilon > 0$ be given and choose N such that $d(F_N) < \epsilon$. Then p_m, $p_n \in F_N$ for all $m, n \geqslant N$. Hence, $d(p_m, p_n) \leqslant d(F_N) < \epsilon$. This shows that the sequence $\{p_n\}$ is a Cauchy sequence. Now because X is complete, there exists a $p \in X$ such that $\lim_{n\to\infty} p_n = p$. Evidently, for each n, the sequence $p_{n+1}, p_{n+2}, p_{n+3}, \ldots$ is a subsequence of $\{p_n\}$ and therefore must also converge to p. Since (according to the remarks preceding Theorem 7.16) p is an adherence point of the range of $p_{n+1}, p_{n+2}, p_{n+3}, \ldots$ and since every term of this sequence belongs to F_n, Theorem 6.8 shows that p is also an adherence point of F_n. But F_n is closed and therefore $p \in F_n$. Since n was any positive integer, this shows that $p \in F_n$ for every n. Hence,

$$p \in \bigcap_{n=1}^{\infty} F_n$$

which proves that

$$\bigcap_{n=1}^{\infty} F_n \neq \emptyset$$

Conversely, suppose that $\bigcap_{i=1}^{\infty} F_i \neq \emptyset$ for every nested sequence of nonempty closed and bounded sets $\{F_i\}$ such that $d(F_i) \to 0$, and let $\{p_n\}$ be a Cauchy sequence. Set $E_n = \{p_i | i \geq n\}$. According to the remarks following Definition 9.3,

$$\lim_{n \to \infty} d(E_n) = 0$$

Clearly $E_n \supset E_{n+1}$. Hence Theorem 6.8 shows that $\bar{E}_n \supset \bar{E}_{n+1}$. Theorem 9.4 implies that $\lim_{n \to \infty} d(\bar{E}_n) = 0$ and that $d(\bar{E}_1) = d(E_1)$. A proof analogous to that of Theorem 7.5 [28] shows that E_1, the range of the Cauchy sequence, is bounded. Hence, $\bar{E}_1$ is bounded. Thus $\{\bar{E}_n\}$ is a nested sequence of nonempty closed and bounded sets such that $d(\bar{E}_n) \to 0$, and so, by hypothesis, there exists a point $p \in \bigcap_{n=1}^{\infty} \bar{E}_n$. That is, $p \in \bar{E}_n$ for every n. Now let $\epsilon < 0$ be given and choose N such that

$$d(\bar{E}_N) < \epsilon$$

Since $p \in E_N$ and $E_N \subset E_N$,

$$d(p_n, p) \leq d(\bar{E}_N) < \epsilon$$

for every $p_n \in E_N$ and, therefore, for every p_n with $n \geq N$. This shows that $p_n \to p$. Since $\{p_n\}$ is an arbitrary Cauchy sequence, it follows that X is complete.

Corollary 1: *The Euclidean space R^k is complete.*
Proof: The Heine-Borel theorem (Theorem 6.31) shows that every closed and bounded set in R^k is compact. Therefore the corollary to Theorem 6.27 shows

[28] The sequence $\{p_n\}$ is a Cauchy sequence. Hence, there exists an N such that, for $n \geq N$,

$$d(p_n, p_N) < 1$$

Let

$$\rho = \max\{d(p_1, p_N), d(p_2, p_N), \ldots, d(p_{N-1}, p_N), 1\}$$

Then, since, for every n

$$p_n \in B(p_N; \rho)$$

the range E_1 of $\{p_n\}$ is a subset of $B(p_N; \rho)$. Now

$$d(B(p_N; \rho)) = 2\rho$$

Hence,

$$d(E_1) \leq 2\rho$$

that every nonempty nested sequence of closed and bounded sets whose diameters tend to zero has a nonempty intersection.

Corollary 2: *Every compact metric space is complete.*

Proof: According to Theorem 6.23, every closed subset of a compact metric space is compact. Hence, the corollary to Theorem 6.27 shows that every nonempty nested sequence of closed and bounded sets whose diameters tend to zero has a nonempty intersection.

There are in fact certain analogies between compact metric spaces and complete metric spaces.

Theorem 9.8: *If E is a complete subspace of the metric space X, then E is a closed subset of X.*

Proof: Let p be any adherence point of E. By Theorem 7.4, there is a sequence of points of E which converges to p. Theorem 9.5 shows that this sequence must be a Cauchy sequence. By hypothesis, every Cauchy sequence in E converges to a point of E and so $p \in E$. This, of course, shows that E is closed.

Theorem 9.8 should be compared with Theorem 6.22.

Theorem 9.9: *If E is a closed subset of the complete metric space $\langle X, d\rangle$, then the subspace $\langle E, d\rangle$ is complete.*

Proof: If $\{p_n\}$ is a Cauchy sequence in E, then $\{p_n\}$ is also a Cauchy sequence in X. Since $\langle X, d\rangle$ is complete, there exists a point $p \in X$ such that $p_n \to p$. Theorem 7.4 shows that p is an adherence point of E. Since E is closed, this means that $p \in E$. Hence $\{p_n\}$ converges in the metric space $\langle E, d\rangle$.

Theorem 9.10: *The direct product of two complete metric spaces is complete.*

Proof: Let $\langle X, d\rangle$ and $\langle Y, \delta\rangle$ be complete metric spaces. Suppose that $\{\langle p_n, q_n\rangle\}$ is *any* Cauchy sequence in the direct product $\langle X \times Y, d_\times\rangle$ of $\langle X, d\rangle$ and $\langle Y, \delta\rangle$. Since for all m and n,

$$d(p_n, p_m) \leq d_\times(\langle p_n, q_n\rangle, \langle p_m, q_m\rangle)$$

and

$$\delta(q_n, q_m) \leq d_\times(\langle p_n, q_n\rangle, \langle p_m, q_m\rangle)$$

it follows immediately that $\{p_n\}$ and $\{q_n\}$ are also Cauchy sequences. Because $\langle X, d\rangle$ and $\langle Y, \delta\rangle$ are complete, $\{p_n\}$ converges in X and $\{q_n\}$ converges in Y.

Hence Theorem 7.9(a) shows that $\{\langle p_n, q_n\rangle\}$ converges in $\langle X \times Y, d_\times\rangle$. Since $\{\langle p_n, q_n\rangle\}$ is an arbitrary Cauchy sequence, this completes the proof.

If E is any nonclosed subset of a metric space $\langle X, d\rangle$, then there must be a limit point p of E such that $p \notin E$. Since p is also an adherence point of E, Theorem 7.4 shows that there is a sequence $\{p_n\}$ of points of E which converges to p and to *no other point*. Theorem 9.5 shows that $\{p_n\}$ is a Cauchy sequence. Now $\{p_n\}$ is also a Cauchy sequence in the subspace $\langle E, d\rangle$ but does not converge to any point of E; hence, E is not a complete metric space. This shows that there exists a large class of metric spaces which are not complete. For example, let E be the interval $(0, 1)$ and let X be R^1 with the usual metric. Then the sequence $\{1/n\}$ is certainly a Cauchy sequence in R^1 and it converges to $0 \notin (0, 1)$. Therefore $\{1/n\}$ is not a convergent sequence in the *subspace* E. In cases like this, it is always clear how we can extend the metric space to obtain a complete metric space.

On the other hand, there are many cases of incomplete spaces where the situation is not so simple. For example, let $\mathscr{R}$ be the set of all Riemann integrable functions on the interval $[a, b]$. It is not hard to see[29] that the function $d: \mathscr{R} \times \mathscr{R} \to R^1$ defined in terms of the Riemann integral (this is the usual integral we learn about in elementary calculus) by

$$d(f, g) = \int_a^b |f(x) - g(x)|\,dx \qquad \text{for all } f, g \in \mathscr{R}$$

is a metric on $\mathscr{R}$. But it turns out that this space (or any subspace of this space) is not complete. If one introduces the concept of Lebesgue integration and increases the space of functions[30] to include all Lebesgue integrable functions, then complete spaces of the previous type are obtained.[31] This is one of the important reasons for the abandonment by mathematicians of the Riemann integral in favor of the Lebesgue integral.

This is our second encounter (see remarks following Definition 6.1) with this type of function space. Since the Lebesgue integral is not developed here, we do not pursue this topic any further. However, other types of function spaces will be encountered subsequently, and we shall discuss them in some detail.

[29] Provided we take for granted the fact that the integral in question always exists.

[30] Every Riemann integrable function is also Lebesgue integrable.

[31] Actually, the points in these spaces are not functions themselves, but each point is an equivalence class whose members are those functions which differ from one another on sets that are in a certain sense negligibly small.

In any event, it is very important in mathematics to be able to assert the existence of limits of certain sequences and, as a result, completeness is a very desirable property. It is not in the least unimportant then that, in a certain sense which we now define, every metric space can be regarded as a subset of a complete metric space.

Definition 9.11: *Let $\langle X, d\rangle$ be a metric space. If there is a complete metric space $\langle X^*, d^*\rangle$ and if there is a dense subset X_0 of X^* such that $\langle X, d\rangle$ is isometric to the subspace $\langle X_0, d^*\rangle$, then $\langle X^*, d^*\rangle$ is said to be a* **completion** *of $\langle X, d\rangle$.*

It shall be proved that every metric space has a completion in the sense of this definition. In the proof the metric space $\langle X^*, d^*\rangle$ is constructed from the metric space $\langle X, d\rangle$. It must be emphasized, however, that with this construction the metric space $\langle X, d\rangle$ is not a subspace of $\langle X^*, d^*\rangle$ but is only isometric to a subspace of $\langle X^*, d^*\rangle$; that is, $\langle X, d\rangle$ and $\langle X_0, d^*\rangle$ have the same abstract properties when considered as metric spaces. For many purposes this is good enough. However, one can go even further and actually embed $\langle X, d\rangle$ in $\langle X^*, d^*\rangle$ since, from the metric spaces $\langle X, d\rangle$ and $\langle X^*, d^*\rangle$, we can form the set

$$X \cup (X^* - X_0)$$

and define a suitable metric on this set from the metrics d and d^*. Even if this is done, those limits of the Cauchy sequences in X which are points of the set $X^* - X_0$ are still different objects from the points in the original set X and so, for many purposes, it is much more desirable to start out with a complete metric space.

In Definition 4.10 we introduced the concept of equivalence relation. Before proving that every metric space has a completion, we must establish some properties of this relation.

Definition 9.12: *A collection $\mathscr{P}$ of nonempty subsets of a set S is called a* **partition of** *S if, for each element $s \in S$, there is* **exactly one** *set $E \in \mathscr{P}$ such that $s \in E$.*

Thus, each element of S is in at least one member of $\mathscr{P}$ and no element of S is in more than one member of $\mathscr{P}$.

Definition 9.13: *Let $\sim$ be an equivalence relation in S. For each $s \in S$,* *the subset $\psi(s)$ of S defined by*

$$\psi(s)=\{t \in S|s \sim t\}$$

is called an **equivalence class** *of $\sim$. The collection of all equivalence classes is called the* **quotient set** *of S by $\sim$.*

Theorem 9.14: *Let $\sim$ be an equivalence relation in a set S and, for each $s \in S$, let $\psi(s)$ be the equivalence class $\{t \in S|s \sim t\}$ of $\sim$. Then, for any two elements s and t of S, $s \sim t$ if and only if $\psi(s)=\psi(t)$.*

Proof: First let $s \sim t$ and let x be any member of $\psi(t)$. Then $t \sim x$. From the transitivity of $\sim$, it follows that $s \sim x$, which shows that $x \in \psi(s)$. Since x was arbitrary, we conclude that $\psi(t) \subset \psi(s)$. Since $\sim$ is symmetric, we see that $t \sim s$. Hence, we can repeat the argument just given to show that the opposite inclusion holds and, therefore, that $\psi(t)=\psi(s)$.

Now let $\psi(s)=\psi(t)$. Since $\sim$ is reflective, it follows that $t \in \psi(t)=\psi(s)$. That is, $s \sim t$.

Theorem 9.15: *If $\sim$ is an equivalence relation in a set S and $Q=\{\psi(s) \mid s \in S\}$ is the quotient set of S by $\sim$, then Q is a partition of S.*

Proof: Choose any element $s \in S$. Evidently $s \sim s$. Hence $s \in \psi(s)$. Therefore there is at least one element of Q which contains s. Hence suppose that E is any member of Q such that $s \in E$. It follows from the way Q was constructed that we can find a $y \in S$ such that $E=\psi(y)$. So $s \in \psi(y)$. This means that $y \sim s$ and therefore Theorem 9.14 shows that $\psi(s)=\psi(y)=E$. We conclude from this that the only member of Q which contains s is ψ(s). And this proves the theorem.

We are now ready to prove that every metric space has a completion.

Theorem 9.16: *Every metric space $\langle X, d\rangle$ has a completion $\langle X^*, d^*\rangle$ and any other completion of $\langle X, d\rangle$ is isometric to $\langle X^*, d^*\rangle$.*

Proof: Let $\{x_n\}$ and $\{y_n\}$ be any two Cauchy sequences in $\langle X, d\rangle$. We write $\{x_n\} \sim \{y_n\}$ if and only if

$$\lim_{n\to\infty} d(x_n, y_n)=0 \tag{9–1}$$

Since the proof of the theorem is quite long, it is divided into parts.

Part (*a*). *The relation $\sim$ is an equivalence relation in the set S of all Cauchy sequences in $\langle X, d\rangle$.*

First, since, for every sequence $\{x_n\} \in S$,

$$\lim_{n\to\infty} d(x_n, x_n) = 0$$

we conclude that $\{x_n\} \sim \{x_n\}$; that is, $\sim$ is reflexive. Since $d(x_n, y_n) = d(y_n, x_n)$, it follows that $\{x_n\} \sim \{y_n\}$ implies $\{y_n\} \sim \{x_n\}$.

Finally, suppose $\{x_n\} \sim \{y_n\}$ and $\{y_n\} \sim \{z_n\}$. Then given $\epsilon > 0$, we can find integers N_1 and N_2 such that $n \geqslant N_1$ implies

$$d(x_n, y_n) < \epsilon/2$$

and $m \geqslant N_2$ implies

$$d(y_m, z_m) < \epsilon/2$$

Hence, if we set $N = \max\{N_1, N_2\}$, then, for all $n \geqslant N$,

$$d(x_n, z_n) \leqslant d(x_n, y_n) + d(y_n, z_n) < \epsilon$$

Therefore

$$\lim_{n\to\infty} d(x_n, z_n) = 0$$

That is, $\{x_n\} \sim \{z_n\}$. Thus $\sim$ is transitive. This proves the assertion.

Part (b). For any two Cauchy sequences $\{x_n\}$ and $\{y_n\}$ in $\langle X, d\rangle$, $\lim_{n\to\infty} d(x_n, y_n)$ exists.

Let $\{s_n\}$ denote the sequence of real numbers $\{d(x_n, y_n)\}$ and let $\epsilon > 0$ be given. Since $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences, we can find integers N_1 and N_2 such that $d(y_{n'}, y_{m'}) < \epsilon/2$ for all $n', m' \geqslant N_1$ and $d(x_{n'}, x_{m'}) < \epsilon/2$ for all $n', m' \geqslant N_2$. Hence if we choose $N = \max\{N_1, N_2\}$, then, for all $n, m \geqslant N$,

$$\begin{aligned}
|s_n - s_m| = |d(x_n, y_n) - d(x_m, y_m)| \\
&= |d(x_n, y_n) - d(x_n, y_m) + d(x_n, y_m) - d(x_m, y_m)| \\
&\leqslant |d(x_n, y_n) - d(x_n, y_m)| + |d(x_n, y_m) - d(x_m, y_m)| \\
&\leqslant d(y_n, y_m) + d(x_n, x_m) < \epsilon
\end{aligned}$$

where the next-to-last inequality follows from equation (6–4). We have thus shown that $\{s_n\}$ is a Cauchy sequence (of real numbers). Now corollary 1 of

355-525 O-70—12

Theorem 9.7 shows that the Euclidean space R^1 is complete. Hence, $\lim_{n\to\infty} s_n = \lim_{n\to\infty} d(x_n, y_n)$ exists.

Part (c). If $\{x_n\}$ and $\{y_n\}$ are two Cauchy sequences in $\langle X, d\rangle$ and if $\{x'_n\}$ and $\{y'_n\}$ are any two Cauchy sequences such that $\{x'_n\} \sim \{x_n\}$ and $\{y'_n\} \sim \{y_n\}$, then $\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(x'_n, y'_n)$.

Given $\epsilon > 0$, it follows from the definition of $\sim$ that we can find integers N_1 and N_2 such that $n \geqslant N_1$ implies $d(x_n, x'_n) < \epsilon/4$ and $n \geqslant N_2$ implies $d(y_n, y'_n) < \epsilon/4$. Let $r = \lim_{n\to\infty} d(x_n, y_n)$ and $r' = \lim_{n\to\infty} d(x'_n, y'_n)$ (which exist by part (b)). Then we can find integers N_3 and N_4 such that, for $n \geqslant N_3$,

$$|r - d(x_n, y_n)| < \epsilon/4$$

and, for $n \geqslant N_4$,

$$|r' - d(x'_n, y'_n)| < \epsilon/4$$

Choose $N = \max\{N_1, \ldots, N_4\}$. Then, for $n \geqslant N$,

$$\begin{aligned}
|r - r'| &= |r - d(x_n, y_n) - r' + d(x'_n, y'_n) + d(x_n, y_n) - d(x'_n, y'_n)| \\
&\leqslant |r - d(x_n, y_n)| + |r' - d(x'_n, y'_n)| + |d(x_n, y_n) - d(x'_n, y'_n)| \\
&< \epsilon/2 + |d(x_n, y_n) - d(x'_n, y'_n)| \\
&= \epsilon/2 + |d(x_n, y_n) - d(x'_n, y_n) + d(x'_n, y_n) - d(x'_n, y'_n)| \\
&\leqslant \epsilon/2 + |d(x_n, y_n) - d(x'_n, y_n)| + |d(x'_n, y_n) - d(x'_n, y'_n)| \\
&\leqslant \epsilon/2 + d(x_n, x'_n) + d(y_n, y'_n) < \epsilon
\end{aligned}$$

Since ϵ was arbitrary, we conclude that $r = r'$.

Let X^* be the quotient set of the set $\bar{S}$ of all Cauchy sequences in $\langle X, d\rangle$ by $\sim$ (i.e., the collection of all equivalence classes of $\sim$). We denote the elements of X^* by x^*, y^*, etc. Each element x^* of X^* is a collection of Cauchy sequences which are equivalent to each other—that is, they satisfy equation (9–1). If we choose a Cauchy sequence $\{x_n\}$ from x^* and a Cauchy sequence $\{y_n\}$ from y^*, then part (b) shows that $\lim_{n\to\infty} d(x_n, y_n)$ exists and part (c) shows that this limit is independent of which Cauchy sequence from x^* and which Cauchy sequence from y^* we use to calculate it. That is, the number $\lim_{n\to\infty} d(x_n, y_n)$ depends only on the sets x^* and y^* from which the Cauchy sequences were taken and is independent of the particular Cauchy sequences used to calculate it. Hence, for any two points x^* and y^* of X^*, let us define $d^*(x^*, y^*)$ by

$$d^*(x^*, y^*) = \lim_{n\to\infty} d(x_n, y_n) \tag{9–2}$$

where $\{x_n\}$ is any Cauchy sequence in x^* and $\{y_n\}$ is any Cauchy sequence in y^*.

Part (d). Equation (9–2) defines a metric on X^*.

We must show that d^* satisfies axioms (a) to (c) of Definition 6.1.

It is clear that $d^*(x^*, y^*) \geqslant 0$ since a sequence, all of whose terms are nonnegative, can certainly not converge to a negative number. We must also show that equality prevails if and only if $x^* = y^*$. To this end suppose $d^*(x^*, y^*) = 0$. This means that $\lim_{n\to\infty} d(x_n, y_n) = 0$ for Cauchy sequences $\{x_n\} \in x^*$ and $\{y_n\} \in y^*$. Hence $\{x_n\} \sim \{y_n\}$. Since $\sim$ is an equivalence relation by part (a), x^* is the equivalence class of $\{x_n\}$ by $\sim$, and y^* is the equivalence class of $\{y_n\}$ by $\sim$, Theorem 9.14 shows that $x^* = y^*$. It is clear that the steps of this argument can be reversed to show that $x^* = y^*$ implies $d^*(x^*, y^*) = 0$.

The symmetry of d^* is clear because $d(x_n, y_n) = d(y_n, x_n)$ for all n.

Finally we must show that the triangle inequality is satisfied. For this purpose, let $x^*, y^*, z^* \in X^*$ and suppose that $\{x_n\}$, $\{y_n\}$, and $\{z_n\}$ are, respectively, members of these equivalence classes. Since d is a distance function on X, it is certainly true that

$$d(x_n, z_n) \leqslant d(x_n, y_n) + d(y_n, z_n)$$

Since this holds for every n, it must also be true in the limit; that is,

$$\begin{aligned} d^*(x^*, z^*) &= \lim_{n\to\infty} d(x_n, z_n) \leqslant \lim_{n\to\infty} [d(x_n, y_n) + d(y_n, z_n)] \\ &= \lim_{n\to\infty} d(x_n, y_n) + \lim_{n\to\infty} d(y_n, z_n) = d^*(x^*, y^*) + d^*(y^*, z^*) \end{aligned}$$

(see Theorem 7.7(a)). We have now shown that $\langle X^*, d^* \rangle$ is a metric space.

For each $x \in X$ the constant sequence $x, x, x, \ldots$ is certainly a Cauchy sequence. Theorem 9.15 shows that there is one and only one equivalence class in X^* which contains $x, x, x, \ldots$. We shall use the notation x^+, y^+, etc., for the points of X^* which, respectively, contain the Cauchy sequences $x, x, x, \ldots$; $y, y, y, \ldots$; etc. Let X_0 be the subset of X^* which consists of all those equivalence classes which contain Cauchy sequences of this type. Since, for each $x \in X$, there is one and only one equivalence class x^+ which contains the Cauchy sequence $x, x, x, \ldots$, the scheme which associates with x the equivalence class x^+ certainly defines a function from X to X^*. We call this function f. Thus,

$$\left.\begin{aligned} &f: X \to X^* \\ &f(x) = x^+ \qquad \text{for every } x \in X \end{aligned}\right\} \tag{9–3}$$

and by construction

$$X_0 = f(X)$$

That is, f is onto X_0.

Part (e). The function f is an isometry from X to X_0 (i.e., X_0 is the isometric image of X under f).

If the Cauchy sequences $\{x_n\} = x, x, x, \ldots$ and $\{y_n\} = y, y, \ldots$ are both elements of x^+, then $\{x_n\} \sim \{y_n\}$ so that

$$\lim_{n\to\infty} d(x_n, y_n) = d(x, y) = 0$$

Hence $x = y$. This shows that there cannot be two different Cauchy sequences of this type in the same equivalence class. Thus f is one-to-one, and f is certainly onto. We need only verify that f preserves distance to prove that it is an isometry.

But, if we compute the distance between any two points of X_0, say x^+ and y^+, we find

$$d^*(x^+, y^+) = d(x, y)$$

Part (f). X_0 is dense in X^*.

We must show that the closure $\bar{X}_0$ of X_0 is all of X^*. To this end let x^* be any point of X^* and choose a Cauchy sequence $\{x_n\} \in x^*$. Given $\epsilon > 0$, there exists an N such that $m \geq N$ implies $d(x_N, x_m) < \epsilon/2$. Since

$$\{x_n\} \in x^*$$

$$x_N, x_N, \ldots \in x_N^+$$

it follows that

$$d^*(x^*, x_N^+) = \lim_{m\to\infty} d(x_m, x_N) \leq \epsilon/2 < \epsilon$$

Now in view of the fact that $x_N^+ \in X_0$ we have shown that every ball about x^* contains a point of X_0 and, hence, we have shown that $\bar{X}_0 = X^*$.

Part (g). If $\{x_n^+\}$ is any Cauchy sequence of elements of X_0, then $\{x_n^+\}$ converges to a point $x^* \in X^*$.

It follows from the construction that, for each $n = 1, 2, \ldots$, the constant sequence $x_n, x_n, \ldots$ is an element of x_n^+ and that

$$\left.\begin{aligned} x_n^+ &= f(x_n) \\ x_n &= f^{-1}(x_n^+) \end{aligned}\right\} \qquad \text{for } n=1, 2, 3, \ldots$$

Hence the inverse image of the range of the sequence $\{x_n^+\}$ under the isometry f is the range of the sequence $\{x_n\}$ in X. (This is sometimes called the isometric preimage of the range of $\{x_n^+\}$).

Since $\{x_n^+\}$ is a Cauchy sequence in X^*, it follows that, given $\epsilon > 0$, we can find an integer N such that $m, n \geqslant N$ implies

$$d^*(x_m^+, x_n^+) < \epsilon$$

But since

$$x_n^+ = f(x_n)$$

and since f is an isometry, we conclude that

$$d(x_m, x_n) = d^*(x_m^+, x_n^+) < \epsilon$$

for $m, n \geqslant N$. Thus $\{x_n\}$ (with $x_n = f^{-1}(x_n^+)$ for $n=1, 2, 3, \ldots$) is a Cauchy sequence in X and, therefore, belongs to one of the equivalence classes, say $x^* \in X^*$. We shall show that the sequence $\{x_n^+\}$ converges to x^*. Evidently

$$d^*(x_n^+, x^*) = \lim_{m\to\infty} d(x_n, x_m)$$

Hence,

$$\lim_{n\to\infty} d^*(x_n^+, x^*) = \lim_{n\to\infty} \lim_{m\to\infty} d(x_n, x_m)$$

Since $\{x_n\}$ is a Cauchy sequence, it follows that $\lim\limits_{n\to\infty} \lim\limits_{m\to\infty} d(x_n, x_m) = 0$. Thus $x_n^+ \to x^*$, which is the desired result.

Part (h). The metric space X^* is complete.

To show that X^* is complete, we must demonstrate that any Cauchy sequence of points in X^* converges to a point in X^*.

Hence let $\{x_n^*\}$ be an arbitrary Cauchy sequence in X^*. Since X_0 is dense in X^*, for each $n = 1, 2, 3, \ldots$, we can find an $x_n^+ \in X_0$ such that

$$d^*(x_n^*, x_n^+) < 1/n$$

Thus we have constructed a sequence $\{x_n^+\}$ of points in X_0. Now let $\epsilon > 0$ be given. We can find an N such that $m, n \geqslant N$ implies $d^*(x_n^*, x_m^*) < \epsilon/3$. Hence, for $m, n \geqslant \max\{3/\epsilon, N\}$,

$$d^*(x_n^+, x_m^+) \leqslant d^*(x_n^+, x_n^*) + d^*(x_n^*, x_m^*) + d^*(x_m^*, x_m^+) < 1/n + \epsilon/3 + 1/m \leqslant \epsilon$$

so that $\{x_n^+\}$ is a Cauchy sequence. Now, as shown in part (g), $\{x_n^+\}$ must converge to some $y^* \in X^*$. We now claim that $\{x_n^*\}$ must also converge to y^*. To prove this, we note that, given $\epsilon > 0$, we can find an N' such that $n \geqslant N'$ implies

$$d^*(x_n^+, y^*) < \epsilon/2$$

Let N^* be the smallest integer which is larger than both N' and $2/\epsilon$. Then,

$$d^*(y^*, x_n^*) \leqslant d^*(y^*, x_n^+) + d^*(x_n^+, x_n^*) < \epsilon/2 + 1/n \leqslant \epsilon$$

for $n \geqslant N^*$. That is,

$$x_n^* \rightarrow y^*$$

We have now established that $\langle X^*, d^* \rangle$ is a completion of $\langle X, d \rangle$. Suppose $\langle X^{**}, d^{**} \rangle$ is another completion of $\langle X, d \rangle$. There is a dense subset X_0^* of X^* which is isometric to X and a dense subset X_0^{**} of X^{**} which is isometric to X.[32] Since, just as in the case of one-to-one correspondence, isometry is an equivalence relation, we see that X_0^* must be isometric to X_0^{**}. So let h be an isometry from X_0^* to X_0^{**}. Clearly the image of the range of any sequence $\{x_n^*\}$ in X_0^* under the isometry h is the range of a sequence $\{x_n^{**}\}$ in X^{**}.

Part (i). Let $\{x_n^*\}$ be any sequence in X_0^*. If there is a point $x^* \in X^*$ such that $x_n^* \rightarrow x^*$, then the sequence $\{x_n^{**}\}$ in X_0^{**} defined by $x_n^{**} = h(x_n^*)$ $(n = 1, 2, 3, \ldots)$ converges to a point $x^{**} \in X^{**}$.

Theorem 9.5 shows that the sequence $\{x_n^*\}$ must be a Cauchy sequence. Since h is an isometry, we see that, for any integers m and n,

$$d^*(x_n^*, x_m^*) = d^{**}(x_n^{**}, x_m^{**})$$

which shows that $\{x_n^{**}\}$ must also be a Cauchy sequence. Since X^{**} is complete, this sequence must have a limit $x^{**} \in X^{**}$.

Part (j). Let $\{x_n^*\}$ and $\{y_n^*\}$ be sequences in X_0^* and suppose there is a point $x^* \in X^*$ such that $x_n^* \rightarrow x^*$ and $y_n^* \rightarrow x^*$. If the sequences $\{x_n^{**}\}$ and $\{y_n^{**}\}$ in X_0^{**} are defined by

[32] The set X_0 is now denoted by X_0^*. The notation has been changed to distinguish between the sets X_0^* and X_0^{**}.

$$\left.\begin{aligned} x_n^{**} &= h(x_n^*) \\ y_n^{**} &= h(y_n^*) \end{aligned}\right\} n = 1, 2, 3, \ldots$$

and if $x_n^{**} \to x^{**}$ *and* $y_n^{**} \to y^{**}$, *then* $x^{**} = y^{**}$.

Since h is an isometry, it follows that, for any integers m and n,

$$d^{**}(x_n^{**}, y_m^{**}) = d^*(x_n^*, y_m^*)$$

which shows, when combined with the triangle inequality, that

$$\begin{aligned} d^{**}(x^{**}, y^{**}) &\leq d^{**}(x^{**}, x_n^{**}) + d^{**}(x_n^{**}, y_m^{**}) + d^{**}(y_m^{**}, y^{**}) \\ &= d^{**}(x^{**}, x_n^{**}) + d^{**}(y_m^{**}, y^{**}) + d^*(x_n^*, y_m^*) \\ &\leq d^{**}(x^{**}, x_n^{**}) + d^{**}(y_m^{**}, y^{**}) + d^*(x_n^*, x^*) + d^*(x^*, y_m^*) \end{aligned}$$

Since $x_n^{**} \to x^{**}$, $y_m \to y^{**}$, $x_n^* \to x^*$, and $y_m^* \to x^*$, we conclude that $d^{**}(x^{**}, y^{**}) = 0$; that is, $x^{**} = y^{**}$

Since X_0^* is dense in X^*, every point x^* of X^* must be an adherence point of X_0^*. Therefore, Theorem 7.4 shows that there must be at least one sequence of points in X_0^*, say $\{x_n^*\}$, which converges to x^*. Part (i) shows that the sequence $\{x_n^{**}\}$ in X_0^{**} defined by $x_n^{**} = h(x_n^*)$ $(n = 1, 2, 3, \ldots)$ converges to a point $x^{**} \in X^{**}$. Thus, with each point $x^* \in X^*$, we can associate a point $x^{**} \in X^{**}$. Part (j) shows that, if any other sequence in X_0^* that converges to x^* was used in this construction, it would still lead to the same point x^{**}. In this way then, we associate with each point $x^* \in X^*$ a unique point $x^{**} \in X^{**}$. This scheme defines a function $\Phi: X^* \to X^{**}$.

Part (k). The function Φ is an isometry.

The mapping is certainly injective, for suppose $y^{**} = \Phi(y^*)$, $x^{**} = \Phi(x^*)$, and $x^{**} = y^{**}$. Let $\{x_n^*\}$ and $\{y_n^*\}$ be Cauchy sequences in X_0^* such that $x_n^* \to x^*$ and $y_n^* \to y^*$, and set $x_n^{**} = h(x_n^*)$ and $y_n^{**} = h(y_n^*)$ $(n = 1, 2, \ldots)$. Then,

$$\begin{aligned} d^*(x^*, y^*) &\leq d^*(x^*, x_n^*) + d^*(x_n^*, y_n^*) + d^*(y_n^*, y^*) = d^*(x^*, x_n^*) \\ &\quad + d^*(y_n^*, y^*) + d^{**}(x_n^{**}, y_n^{**}) \leq d^*(x^*, x_n^*) \\ &\quad + d^*(y_n^*, y^*) + d^{**}(x_n^{**}, x^{**}) + d^{**}(x^{**}, y_n^{**}) \end{aligned}$$

Since $x_n^* \to x^*$ and $y_n^* \to y^*$ and since these imply $x_n^{**} \to x^{**}$ and $y_n^{**} \to x^{**}$, we conclude $d^*(x^*, y^*) = 0$; that is $x^* = y^*$.

In order to show that Φ is surjective, let x^{**} be any point of X^{**}. Since X_0^{**} is dense in X^{**}, there is a Cauchy sequence of points of X_0^{**}, say $\{x_n^{**}\}$, such that

$x_n^{**} \to x^{**}$. Since h^{-1} must be an isometry (because h is), the sequence $\{x_n^*\}$ defined by $x_n^* = h^{-1}(x_n^{**})$ $(n=1, 2, 3, \ldots)$ is also a Cauchy sequence and so converges to a point $x^* \in X^*$ (because X^* is complete). Since $x_n^{**} = h(x_n^*)$ $(n=1, 2, 3, \ldots)$, we see from this that $x^{**} = \Phi(x^*)$.

In order to show that Φ is an isometry, it only remains to show that distances are preserved under this map. To this end let x^* and y^* be any elements of X^* and let $x^{**} = \Phi(x^*)$ and $y^{**} = \Phi(y^*)$ be their images under Φ. Choose sequences $\{x_n^*\}$ and $\{y_n^*\}$ in X_0^* such that $x_n^* \to x^*$ and $y_n^* \to y^*$ so that, in X^{**}, $y_n^{**} \to y^{**}$ and $x_n^{**} \to x^{**}$ where $\{y_n^{**}\}$ and $\{x_n^{**}\}$ are defined by

$$\left.\begin{aligned} x_n^{**} &= h(x_n^*) \\ y_n^{**} &= h(y_n^*) \end{aligned}\right\} n = 1, 2, 3, \ldots$$

Now let $\epsilon > 0$ be given and choose $N_1, \ldots, N_4$ such that $d^*(x^*, x_n^*) < \epsilon/2$ for $n \geqslant N_1$, $d^*(y^*, y_n^*) < \epsilon/2$ for $n \geqslant N_2$, $d^{**}(x^{**}, x_n^{**}) < \epsilon/2$ for $n \geqslant N_3$ and $d^{**}(y^{**}, y_n^{**}) < \epsilon/2$ for $n \geqslant N_4$. Setting $N = \max\{N_1, \ldots, N_4\}$, we find, for $n \geqslant N$,

$$d^*(x^*, y^*) \leqslant d^*(x^*, x_n^*) + d^*(x_n^*, y_n^*) + d^*(y_n^*, y^*) < d^*(x_n^*, y_n^*) + \epsilon$$
$$d^*(x_n^*, y_n^*) \leqslant d^*(x_n^*, x^*) + d^*(x^*, y^*) + d^*(y^*, y_n^*) < d^*(x^*, y^*) + \epsilon$$

Similarly

$$d^{**}(x^{**}, y^{**}) < d^{**}(x_n^{**}, y_n^{**}) + \epsilon$$
$$d^{**}(x_n^{**}, y_n^{**}) < d^{**}(x^{**}, y^{**}) + \epsilon$$

So we conclude that

$$|d^*(x^*, y^*) - d^*(x_n^*, y_n^*)| < \epsilon$$
$$|d^{**}(x^{**}, y^{**}) - d^{**}(x_n^{**}, y_n^{**})| < \epsilon$$

and hence

$$\begin{aligned} |d^*(x^*, y^*) - d^{**}(x^{**}, y^{**})| &\leqslant |d^*(x^*, y^*) - d^*(x_n^*, y_n^*)| \\ &\quad + |d^*(x_n^*, y_n^*) - d^{**}(x_n^{**}, y_n^{**})| + |d^{**}(x_n^{**}, y_n^{**}) - d^{**}(x^{**}, y^{**})| \\ &< 2\epsilon + |d^*(x_n^*, y_n^*) - d^{**}(x_n^{**}, y_n^{**})| \end{aligned}$$

But since x_n^{**} and y_n^{**} are the *isometric images of* x_n^* and y_n^*, respectively, we see that $d^{**}(x_n^{**}, y_n^{**}) = d^*(x_n^*, y_n^*)$. Therefore

$$|d^*(x^*, y^*) - d^{**}(x^{**}, y^{**})| < 2\epsilon$$

Since ϵ was *arbitrary*, this means that

$$d^*(x^*, y^*) = d^{**}(x^{**}, y^{**})$$

which completes the proof.

Many applied problems in mathematics reduce to finding a solution to an equation of the form

$$f(x) = x \tag{9–4}$$

where f is a function whose domain is some metric space $\langle X, d\rangle$ and $f: X \to X$. If X is R^1 (or some subset of R^1), equation (9–4) may be an algebraic equation. However, the metric space X can be much more general than this. For example, it may be a function space. In this case equation (9–4) could represent an integral or differential equation.

If X is a *complete* metric space, it is frequently possible, with certain restrictions on the function f, to construct a convergent iterative procedure for calculating the solution to equation (9–4) which at the same time yields a proof of the existence and uniqueness of the solution. The theorems which yield general iterative procedures of this type are known as *algebraic fixed point theorems.* Before introducing some of these theorems, the following definitions are needed.

Definition 9.17: *Let E be a subset of a metric space $\langle X, d\rangle$. A function $f: E \to X$ is said to satisfy a* **Lipschitz condition** *with* **modulus** *α if there exists a* **positive** *number α such that for every $p, q \in E$*

$$d(f(p), f(q)) \leqslant \alpha d(p, q)$$

If, in addition, $\alpha < 1$, the function f is said to be a **contraction mapping of E into X with modulus α.**

Definition 9.18: *Let $f: X \to X$. If E is a subset of X such that*

$$f(x) \in E \qquad \text{for all } x \in E$$

then E is said to be an **invariant subset** *of f. If, in addition, X is a metric space, then E is called an* **invariant subspace** *of f.*

Suppose $f: X \to X$ and E is an invariant subset of f. Let h be the restriction of f to E. Then the range of h is a subset of E. Hence *h can be considered as a function from E* **into** *E.*

The following theorem, due to Banach, is the most well known of all the algebraic fixed point theorems. It is sometimes called the Principle of Contraction Mapping. Many of the other algebraic fixed point theorems are attempts to weaken the hypothesis of Banach's theorem.

Theorem 9.19: *Let f be a contraction mapping of the* **complete** *metric space $\langle X, d\rangle$ into itself. Then there is a unique point $p \in X$ such that $f(p)=p$ and, for any point $x_0 \in X$, the sequence $x_0, f(x_0), f(f(x_0)), \ldots$ converges to p.*

Proof: Choose any point $x_0 \in X$ and define a sequence $\{x_n\}$ recursively by

$$x_{n+1}=f(x_n) \qquad n=0, 1, 2, 3, \ldots \tag{9–5}$$

Since f is a contraction mapping, there exists a positive number $\alpha < 1$ such that, for every $p, q \in X$,

$$d(f(p), f(q)) \leqslant \alpha d(p, q) \tag{9–6}$$

Therefore, for every positive integer n,

$$d(x_{n+1}, x_n) \leqslant \alpha d(x_n, x_{n-1}) \tag{9–7}$$

In particular,

$$d(x_2, x_1) \leqslant \alpha d(x_1, x_0)$$

Suppose that for $n \geqslant 1$,

$$d(x_{n+1}, x_n) \leqslant \alpha^n d(x_1, x_0) \tag{9–8}$$

Then equation (9–7) shows that

$$d(x_{n+2}, x_{n+1}) \leqslant \alpha^{n+1} d(x_1, x_0)$$

Hence, by induction, we conclude that equation (9–8) must be true for every n.

By successively applying the triangle inequality, we find that, for $k > n$,

$$d(x_k, x_n) \leqslant \sum_{r=n}^{k-1} d(x_{r+1}, x_r)$$

Upon substituting equation (9–8) into this expression, we see that

$$d(x_k, x_n) \leqslant \sum_{r=n}^{k-1} \alpha^r d(x_1, x_0) = \left[\frac{d(x_1, x_0)}{1-\alpha}\right](\alpha^n - \alpha^k) \tag{9–9}$$

Since $\{\alpha^n\}$ converges (recall that $\alpha < 1$) and since every convergent sequence is a Cauchy sequence, equation (9–9) shows that $\{x_n\}$ is also a Cauchy sequence. But X is a complete metric space. Hence there exists a point $p \in X$ such that

$$\lim_{n\to\infty} x_n = p \tag{9–10}$$

Since this implies that given $\epsilon > 0$ there exists a positive integer N such that

$$d(x_k, p) < \epsilon \qquad \text{for } k \geqslant N$$

it follows from equation (6–4) that

$$|d(x_n, p) - d(x_k, x_n)| < \epsilon$$

whenever $k \geqslant N$. That is,

$$\lim_{k\to\infty} d(x_k, x_n) = d(p, x_n)$$

Hence taking the limit as $k \to \infty$ in equation (9–9) yields (since $\lim_{k\to\infty} \alpha^k = 0$)

$$d(p, x_n) \leqslant \left[\frac{d(x_1, x_0)}{1-\alpha}\right] \alpha^n \tag{9–11}$$

Upon combining this result with equations (9–5) and (9–6) we find that

$$d(f(p), x_{n+1}) = d(f(p), f(x_n)) \leqslant \alpha d(p, x_n) \leqslant \left[\frac{d(x_1, x_0)}{1-\alpha}\right] \alpha^{n+1}$$

And since $\lim_{n\to\infty} \alpha^n = 0$, this implies that $\lim_{n\to\infty} x_n = f(p)$. Thus it follows from equation (9–10) that

$$f(p) = p$$

This proves the existence of the point p and gives a constructive procedure for finding it.

It remains to show that this point is unique. To this end suppose $f(p) = p$ and $f(q) = q$. Then equation (9–6) shows that

$$d(p, q) \leqslant \alpha d(p, q)$$

Since $0 < \alpha < 1$, this shows that $d(p, q) = 0$; that is, $p = q$.

Notice that equation (9–11) gives us a means for determining an upper bound for the error after n steps of the iteration process described in this theorem have been carried out. The point x in equation (9–4) is referred to as a **fixed point** of f.

The hypothesis of the theorem requires that the range of the contraction mapping f be a subset of its domain. It frequently occurs in practice that one is interested in obtaining solutions to equations of the form given in equation (9–4) when this condition cannot be met. The following corollary gives conditions under which these restrictions on the contraction mapping can be weakened.

Corollary 1: *Let $\langle E, d\rangle$ be a complete subspace of the metric space $\langle X, d\rangle$ and let f be a contraction mapping of E into X with modulus α. If there exists an $x^* \in E$ such that the set $S(x^*)$ defined by*[33]

$$S(x^*) = \left\{x \in X \middle| d(x, x_0) \leqslant \frac{\alpha}{1-\alpha} d(x_0, x^*)\right\} \qquad \textit{with } x_0 = f(x^*)$$

is a subset of E, then there is a unique point $p \in E$ such that $f(p) = p$ and the sequence $x_0, f(x_0), f(f(x_0)), \ldots$ converges to p. In addition, p must belong to $S(x^*)$.

Proof: Suppose there is an $x^* \in E$ such that $S(x^*) \subset E$ and let x be any point of $S(x^*)$. Then $x \in E$. Since f is a contraction mapping of E with modulus α, it follows that

$$d(f(x), f(x^*)) \leqslant \alpha d(x, x^*)$$

Hence

$$d(f(x), x_0) = d(f(x), f(x^*)) \leqslant \alpha d(x, x^*) \leqslant \alpha[d(x, x_0) + d(x_0, x^*)]$$
$$\leqslant \alpha\left[\frac{\alpha}{1-\alpha} + 1\right] d(x_0, x^*) = \frac{\alpha}{1-\alpha} d(x_0, x^*)$$

[33] The set $S(x^*)$ is the closed ball of radius $\left(\frac{\alpha}{1-\alpha}\right) d(x_0, x^*)$ about x_0.

This shows that $x \in S(x^*)$ implies $f(x) \in S(x^*)$. Thus $S(x^*)$ is an invariant subspace of f. Let h be the restriction of f to h. According to the remarks following Definition 9.18, h can be considered as a function from the subset $S(x^*)$ into itself. Since f is a contraction mapping of E into X, it follows that h is a contraction mapping of the subspace $\langle S(x^*), d\rangle$ into itself. Clearly $S(x^*)$ is a closed[34] subset of the complete metric space $\langle E, d\rangle$. Hence Theorem 9.9 shows that $\langle S(x^*), d\rangle$ is a complete metric space. Since $x_0 \in S(x^*)$, Banach's theorem now shows that there is a unique point $p \in S(x^*)$ such that $f(p) = h(p) = p$ and that the sequence $x_0, f(x_0), f(f(x_0)), \ldots$ converges to p. On the other hand p is also the only point of E such that $f(p) = p$. For if q was another such point, it follows from the fact that f is a contraction mapping with constant α that

$$d(p, q) \leqslant \alpha d(p, q)$$

which shows, since $0 < \alpha < 1$, that $p = q$.

Let X be any set and let $f: X \to X$. For each positive integer n we define inductively a function $f^{(n)}: X \to X$ as follows: Put $f^{(1)} = f$ and for each positive integer n put $f^{(n+1)} = f \circ f^{(n)}$. It is easy to prove by induction that for any two positive integers m and n

$$f^{(m)} \circ f^{(n)} = f^{(m+n)}$$

Banach's fixed point theorem now tells us that if f is a contraction mapping of a complete metric space $\langle X, d\rangle$ into itself then, for every $x_0 \in X$, $\lim_{n\to\infty} f^{(n)}(x_0) = p$ where p is the unique fixed point of f.

It may turn out that a continuous function f is not a contraction mapping but for some integer m the function $f^{(m)}$ is. The next corollary shows that in this case the conclusions of Banach's theorem still hold.

Corollary 2: *Let $\langle X, d\rangle$ be a complete metric space and let f be a continuous function from X into itself. If, for some positive integer m, the function $f^{(m)}$ is a contraction mapping, then there exists a point $p \in X$ such that $f(p) = p$ and for any $x_0 \in X$ the sequence $\{f^{(n)}(x_0)\}$ converges to p.*
Proof: Set $g = f^{(m)}$. Banach's theorem shows that there is a *unique* point

[34] See example following Theorem 6.13.

$p \in X$ such that $g(p)=p$. Hence $f(g(p))=f(p)$. But

$$f(g(p))=f(f^{(m)}(p))=f\circ f^{(m)}(p)=f^{(m+1)}(p)=f^{(m)}\circ f(p)=f^{(m)}(f(p))$$

Hence $g(f(p))=f(p)$. This shows that $f(p)$ is also a fixed point of g. Since the fixed point is unique, we conclude that $f(p)=p$. It also follows from Banach's theorem that the sequence $\{g^{(n)}(x_0)\}$ converges to p for every $x_0 \in X$. However,

$$g^{(n)}(x_0)=f^{(mn)}(x_0)$$

And, since f is continuous, Theorem 8.5 shows that

$$\lim_{n\to\infty} f(g^{(n)}(x_0))=f(p)=p$$

Hence the sequence $\{f^{(mn+1)}(x_0)\}$ converges to p. Continuing in this manner we can show that for any $j<m$ the sequence $\{f^{(mn+j)}(x_0)\}$ converges to p. But the theorem of factorization of integers shows that every positive integer i can be written in the form

$$i=mn+j \qquad \text{with } 0\leqslant j<m$$

Hence we conclude that the sequence $\{f^{(i)}(x_0)\}$ converges to p.

We shall first give a very simple example to illustrate how Banach's theorem can be applied. Suppose f is a real valued function defined on the closed interval $[0, 1]$ and suppose that f satisfies the Lipschitz condition with modulus K

$$|f(p_2)-f(p_1)|\leqslant K|p_2-p_1| \qquad \text{for all real numbers } p_1, p_2 \in [0, 1] \tag{9–12}$$

If $0<K<1$ and the values of f lie in $[0, 1]$, then f is a contraction mapping of the metric space $[0, 1]$ (with the usual metric) into itself. Since (Theorem 9.9) closed subspaces of complete metric spaces are complete, the requirements of Theorem 9.19 are met. Hence, for any $x_0 \in [0, 1]$, the sequence $\{x_n\}$ defined by

$$x_{n+1}=f(x_n) \tag{9–13}$$

converges to the unique number x^* which is a solution of the equation

$$f(x)=x$$

Figure 9–1 illustrates how the successive approximations given by equation

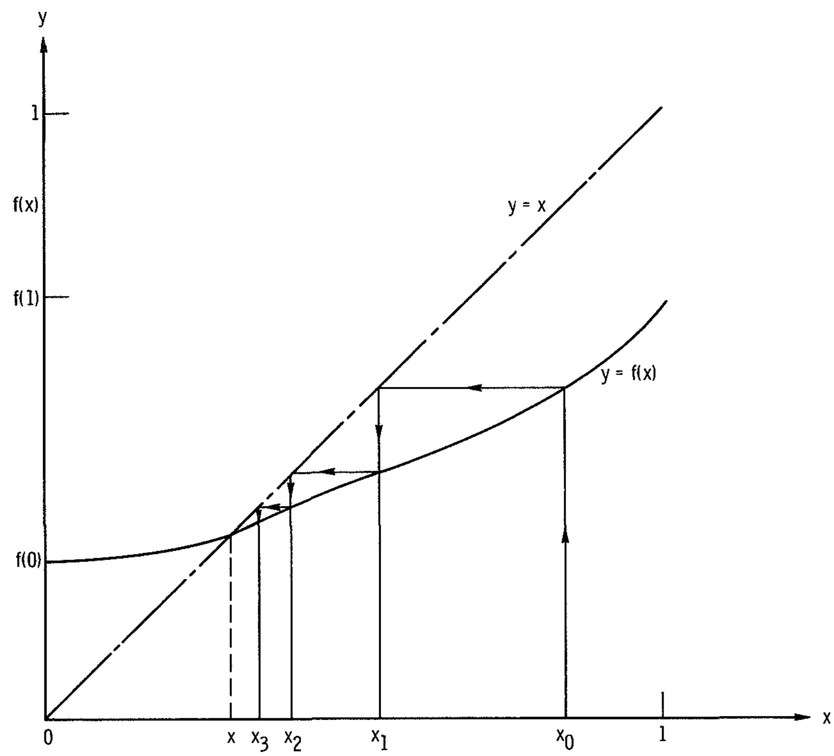


FIGURE 9–1.—Convergence of contraction mapping iteration procedure.

(9–13) converge to the fixed point x^* for the case where f has a positive slope.

Now, for each $a \in [0, 1]$, consider the function $f_a:[0, 1] \to R^1$ defined by

$$f_a(x) = x + \tfrac{1}{2}(a - x^2) \quad \text{for all } x \in [0, 1] \tag{9–14}$$

For any real number x,

$$x + \tfrac{1}{2}(a - x^2) = \tfrac{1}{2}(a + 1 - x^2 + 2x - 1) = \tfrac{1}{2}[(a+1) - (1-x)^2]$$

Hence

$$0 \leq \tfrac{1}{2}a \leq f_a(x) \leq \tfrac{1}{2}(a+1) \leq 1 \quad \text{for all } a, x \in [0, 1]$$

Therefore f_a is a mapping of the complete space $[0, 1]$ into itself. Since

$$|f_a(x) - f_a(y)| = \tfrac{1}{2}|(1-y)^2 - (1-x)^2| = \tfrac{1}{2}|x-y|\,|2-(x+y)|$$

it is easy to see that there is no positive number $\alpha < 1$ such that

$$|f_a(x)-f_a(y)| \leqslant \alpha|x-y| \quad \text{for all } x, y \in [0, 1]$$

However, let us consider the function $f_a^{(2)}$. Clearly, for any $x \in [0, 1]$,

$$\begin{aligned} f_a^{(2)}(x) = f_a(f_a(x)) &= \tfrac{1}{2}\{(a+1)-[1-\tfrac{1}{2}(a+1)+\tfrac{1}{2}(1-x)^2]^2\} \\ &= \tfrac{1}{2}(a+1)-\tfrac{1}{8}[1-a+(1-x)^2]^2 \end{aligned}$$

Hence for any $x, y \in [0, 1]$

$$\begin{aligned} |f_a^{(2)}(x)-f_a^{(2)}(y)| &= \frac{1}{8}|(1-a+(1-y)^2)^2-(1-a+(1-x)^2)^2| \\ &= \frac{1}{8}|(1-y)^2-(1-x)^2|\,|2(1-a)+(1-x)^2+(1-y)^2| \\ &= \frac{1}{8}|x-y|\,|2-x-y|\,|2(1-a)+(1-x)^2+(1-y)^2| \\ &\leqslant \frac{1}{8}2\cdot 2(2-a)\,|x-y| = \frac{2-a}{2}|x-y| \end{aligned}$$

Thus $f_a^{(2)}$ is a contraction mapping of the complete metric space $[0, 1]$ into itself with constant $(2-a)/2$ for $0 < a \leqslant 1$.

Corollary 2 of the theorem now tells us that for each $a \in (0, 1]$ there is a point $p(a) \in [0, 1]$ such that

$$p(a) = f_a(p(a)) = p(a) + \frac{1}{2}(a-[p(a)]^2)$$

and that the sequence $\{p_n(a)\}$ which is defined recursively by

$$\left.\begin{aligned} p_0(a) &= 0 \\ p_{n+1}(a) = f_a(p_n(a)) &= p_n(a) + \frac{1}{2}(a-[p_n(a)]^2) \end{aligned}\right\} \tag{9–15}$$

converges to $p(a)$.

Evidently $[p(a)]^2 = a$. That is, $p(a) = \sqrt{a}$. We have therefore shown that $\{p_n(a)\}$ converges to $\sqrt{a}$ for $0 < a \leqslant 1$. However it is easy to see from equation (9–15) that $p_n(0) = 0$ for every n. Since $\sqrt{0} = 0$, we conclude that *the sequence* $\{p_n(a)\}$ *defined recursively by equation* (9–15) *converges to* $\sqrt{a}$ *for* $0 \leqslant a \leqslant 1$. *In addition* $p_n(0) = 0$ *for every* n.

The following is another relatively simple application of Theorem 9.19. Consider the set of k linear algebraic equations

$$b_i = \sum_{j=1}^{k} a_{ij} x_j \qquad i = 1, 2, \ldots, k \tag{9–16}$$

for the k unknowns $x_1, \ldots, x_k$. This system of equations can be transformed into the system

$$x_i = \sum_{j=1}^{k} C_{ij} x_j + b_i \qquad i = 1, 2, \ldots, k \tag{9–17}$$

by setting $C_{ij} = \delta_{ij} - a_{ij}$ where δ_{ij} is the Kronecker delta defined by

$$\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

Let $\boldsymbol{f} = \langle f_1, \ldots, f_k \rangle$ be the function from R^k into R^k defined by

$$f_i(\boldsymbol{y}) = \sum_{j=1}^{k} C_{ij} y_j + b_j \qquad i = 1, 2, \ldots, k \text{ for all } \boldsymbol{y} = \langle y_1, \ldots, y_k \rangle \in R^k$$

Then finding the solution to the system of equations (9–17) is equivalent to finding a vector $\boldsymbol{x} \in R^k$ such that

$$\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}) \tag{9–18}$$

that is, finding a fixed point of $\boldsymbol{f}$. Since the Euclidean space R^k is complete, Theorem 9.19 gives a convergent iterative procedure for calculating the fixed point of equation (9–18) or, equivalently, the solution of the equation (9–17) whenever $\boldsymbol{f}$ is a contraction mapping. We shall now obtain a sufficient condition for $\boldsymbol{f}$ to be a contraction mapping. To this end notice that for any two points $\boldsymbol{p} = \langle p_1, \ldots, p_k \rangle$ and $\boldsymbol{q} = \langle q_1, \ldots, q_k \rangle$ of R^k

$$\begin{aligned} d(\boldsymbol{f}(\boldsymbol{p}), \boldsymbol{f}(\boldsymbol{q})) &= |\boldsymbol{f}(\boldsymbol{p}) - \boldsymbol{f}(\boldsymbol{q})| \\ &= \sqrt{\sum_{i=1}^{k} [f_i(\boldsymbol{p}) - f_i(\boldsymbol{q})]^2} \\ &= \sqrt{\sum_{i=1}^{k} \left(\sum_{j=1}^{k} C_{ij} p_j - \sum_{j=1}^{k} C_{ij} q_j \right)^2} \\ &= \sqrt{\sum_{i=1}^{k} \left[\sum_{j=1}^{k} C_{ij} (p_j - q_j) \right]^2} \end{aligned}$$

Since any set of k real numbers are the components of a vector in R^k, it follows from Theorem 3.6(d) after using the definitions of norm and scalar product in the Euclidean space R^k that for any real numbers $d_1, \ldots, d_k$ and $e_1, \ldots, e_k$

$$\left|\sum_{j=1}^{k} d_j e_j\right| \leqslant \left(\sum_{j=1}^{k} d_j^2\right)^{1/2}\left(\sum_{n=1}^{k} e_n^2\right)^{1/2}$$

Hence for each $i=1, \ldots, k$

$$\left[\sum_{j=1}^{k} C_{ij}(p_j - q_j)\right]^2 \leqslant \sum_{j=1}^{k} C_{ij}^2 \sum_{n=1}^{k} (p_n - q_n)^2$$

Therefore

$$d(\boldsymbol{f}(\boldsymbol{p}), \boldsymbol{f}(\boldsymbol{q})) \leqslant \sqrt{\sum_{i=1}^{k}\sum_{j=1}^{k} C_{ij}^2 \sum_{n=1}^{k} (p_n - q_n)^2} = \left(\sum_{i=1}^{k}\sum_{j=1}^{k} C_{ij}^2\right)^{1/2} |\boldsymbol{p} - \boldsymbol{q}| = \alpha d(\boldsymbol{p}, \boldsymbol{q})$$

where we have put $\alpha^2 = \sum_{i=1}^{k}\sum_{j=1}^{k} C_{ij}^2$. We may now conclude that $\boldsymbol{f}$ is a contraction mapping whenever

$$\sum_{i=1}^{k}\sum_{j=1}^{k} C_{ij}^2 < 1$$

A less trivial application of Banach's theorem is given in chapter 11.

CHAPTER 10

Infinite Series

We have already indicated (see introductions to chapters 1 and 7) that perhaps the two most important events in the development of modern analysis (discovery of sets and definition of convergence) occurred as a result of the study of trigonometric series. Aside from their role in the development of mathematics, infinite series are also important in their own right.

The classical theory of convergence of infinite series as presented in elementary texts is due to Cauchy and Abel. With the advent of a consistent theory of the real number system, their work passed into the standard expositions of analysis practically unchanged. Upon recognizing the importance of convergence, the mathematicians of the nineteenth century proceeded to develop numerous convergence tests.

In the eighteenth century, mathematicians freely rearranged terms of infinite series until, in 1833, Cauchy gave an example of a conditionally convergent series which, upon rearrangement, converged to a different sum. In 1837, Dirichlet proved that every rearrangement of an absolutely convergent series converges to the same sum.

Since infinite series involve both the concept of addition and the concept of convergence, the natural setting for discussing this topic is the normed linear space, or the Banach space if the completeness property is needed for existence statements. We therefore begin by defining an infinite series in a normed linear space.

If $\{v_n\}$ is any sequence in a normed linear space, we can, of course, for any integers p and q with $p < q$, form the sum $v_p + v_{p+1} + \ldots + v_q$. The familiar notation

$$\sum_{n=p}^{q} v_n$$

is used for this sum.

Definition 10.1: *Let $\{v_n\}$ be any sequence in a normed linear space V. The sequence $\{s_n\}$ whose* nth *term is defined by*

$$s_n = \sum_{k=1}^{n} v_k \qquad n = 1, 2, 3, \ldots$$

is called an **infinite series in** V. *The symbol*

$$\sum_{n=1}^{\infty} v_n \tag{10–1}$$

is used for $\{s_n\}$. *The* nth *term of* $\{s_n\}$ *is called the* nth **partial sum** *of the infinite series. If, in addition, the sequence $\{s_n\}$ converges, its limit s is called the* **sum** *of the infinite series and we write*

$$s = \sum_{n=1}^{\infty} v_n$$

As indicated in the comments following Definition 4.15, the sequence $\{v_n\}$ may be defined as a function from the set of nonnegative integers (instead of the set J of positive integers) into V. In this case the notation

$$\sum_{n=0}^{\infty} v_n \tag{10–2}$$

is used instead of notation (10–1) and, sometimes, when the distinction is immaterial, we simply write

$$\sum v_n$$

Occasionally we use the symbolic notation

$$v_1 + v_2 + v_3 + \ldots$$

in place of notation (10–1).

It is important to point out that, *when an infinite series converges, its sum is not obtained by simple addition but is, in fact, the limit of a sequence of sums.*

We remind the reader that the Euclidean spaces R^k are normed linear spaces. In fact, according to corollary 1 to Theorem 9.7 they are Banach spaces. Thus, in particular, any general results about infinite series in normed linear spaces can be applied to sequences of real or complex numbers (with the norm taken to be the absolute value).

Although it is true that most theorems about sequences can be restated as theorems about infinite series and vice versa, it is still helpful to consider both concepts. For example, we may restate Theorem 9.5 as follows:

Theorem 10.2: *If* $\sum v_n$ *is a convergent infinite series in a normed linear space, then, for every* $\epsilon > 0$, *there is an integer* N *such that*

$$\left\| \sum_{k=n}^{m} v_k \right\| < \epsilon \tag{10–3}$$

whenever $m \geqslant n \geqslant N$.

So, in particular, if we put $m = n$, this implies $\|v_n\| < \epsilon$ whenever $n \geqslant N$. Thus we have the following theorem.

Theorem 10.3: *If* $\sum v_n$ *is a convergent infinite series in a normed linear space, then*

$$\lim_{n \to \infty} v_n = 0$$

We must point out, however, that *the condition* $v_n \to 0$ *is not a sufficient condition for the convergence of* $\sum v_n$. For example, it will be seen that the real series $\sum\limits_{n=1}^{\infty} 1/n$ diverges.

However, in a Banach space condition (10–3) is equivalent to convergence. Thus, we have the next theorem.

Theorem 10.4: *A series* $\sum b_n$ *in a Banach space converges if and only if, for every* $\epsilon > 0$, *there is an integer* N *such that* $m \geqslant n \geqslant N$ *implies*

$$\left\| \sum_{k=n}^{m} b_k \right\| < \epsilon$$

Theorem 7.11 about monotone sequences also has a counterpart for series with real terms.

Theorem 10.5: *An infinite series with (real) nonnegative terms converges if and only if its sequence of partial sums is bounded above.*

When Theorem 7.7 and the first corollary to Theorem 7.9 are translated into the language of infinite series, we arrive at the following theorem.

Theorem 10.6: (a) *If $\sum z_n$ is an infinite series of complex numbers, then $\sum z_n$ converges if and only if both $\sum \mathscr{R}e\, z_n$ and $\sum \mathscr{I}m\, z_n$ converge and, when these series converge, it must be true that*

$$\mathscr{R}e \sum z_n = \sum \mathscr{R}e\, z_n \quad \text{and} \quad \mathscr{I}m \sum z_n = \sum \mathscr{I}m\, z_n$$

(b) *If $\sum u_n$ and $\sum v_n$ are convergent infinite series in a normed linear space with sums u and v, respectively, and if α and β are complex numbers, then $\sum (\alpha u_n + \beta v_n)$ is convergent and has the sum $\alpha u + \beta v$.*

The value of Theorems 10.4 and 10.5 is that they allow us to establish the convergence of an infinite series without a knowledge of its sum. The next theorem (called the comparison test) gives us another way of doing this.

Theorem 10.7: (a) *Let $\{b_n\}$ be a sequence in a Banach space and suppose that for some integer N_0, $\|b_n\| \leqslant c_n$ whenever $n \geqslant N_0$ and that the real infinite series with nonnegative terms $\sum c_n$ converges. Then $\sum b_n$ converges.*

(b) *Let $\{a_n\}$ be a sequence of nonnegative numbers and suppose that there exists a sequence $\{d_n\}$ of nonnegative numbers such that, for all $n \geqslant N_0$, $a_n \geqslant d_n$ and that the infinite series $\sum d_n$ diverges. Then $\sum a_n$ diverges.*

Proof: Part (a). Let $\epsilon > 0$. It follows from Theorem 10.4 that there exists an integer $N \geqslant N_0$ such that $m \geqslant n \geqslant N$ implies

$$\sum_{k=n}^{m} c_k < \epsilon$$

So we see from the triangle inequality that

$$\left\| \sum_{k=n}^{m} b_k \right\| \leqslant \sum_{k=n}^{m} \|b_k\| \leqslant \sum_{k=n}^{m} c_k < \epsilon \tag{10–4}$$

and part (a) follows from Theorem 10.4.

Part (b). This is a direct consequence of part (a) for, if $\sum a_n$ converges, then so must $\sum d_n$.

Theorem 10.7 is frequently used to establish the convergence of a given infinite series. It is clear, however, that this cannot be done unless we know something about the convergence of a certain number of infinite series whose terms are nonnegative numbers.

First let us consider the so-called geometric series $\sum_{n=0}^{\infty} x^n$. The convergence of this series can be established by considering its partial sums. It is easy to see that, if $x \neq 1$, then

$$s_n = \sum_{k=0}^{n} x^k = \frac{1-x^{(n+1)}}{1-x} \tag{10–5}$$

Hence, upon taking the limit, we immediately establish that, for $0 \leqslant x < 1$,

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \tag{10–6}$$

and that, if $x > 1$, $\sum_{n=0}^{\infty} x^n$ diverges. For $x = 1$ we see that

$$\sum_{n=0}^{\infty} x^n = 1 + 1 + \ . \ . \ .$$

which certainly diverges.

There is a result due to Cauchy which allows us to establish the convergence of many real infinite series with monotonically decreasing terms by examining certain rather "small" subsequences. In fact, if $\{a_n\}$ is a monotonically decreasing sequence of nonnegative numbers, the infinite series $\sum_{n=1}^{\infty} a_n$ converges if and only if the infinite series $\sum_{k=0}^{\infty} 2^k a_{(2^k)}$ does.

To obtain this result, it is sufficient to show, in view of Theorem 10.5, that the sequence of partial sums $\sum_{k=0}^{\infty} 2^k a_{(2^k)}$ is bounded above whenever the sequence of partial sums $\sum_{n=1}^{\infty} a_n$ is and vice versa. Hence, let s_n be the nth partial sum of $\sum_{n=1}^{\infty} a_n$ and t_k be the kth partial sum of $\sum_{k=0}^{\infty} 2^k a_{(2^k)}$. Then,

$$s_n = a_1 + a_2 + \ldots + a_n$$

and

$$t_k = a_1 + 2a_2 + 4a_4 + \ldots + 2^k a_{2^k}$$

Clearly, for $n < 2^k$,

$$s_n \leqslant a_1 + (a_2 + a_3) + \ldots + (a_{2^k} + \ldots + a_{2^{k+1}-1}) \leqslant a_1 + 2a_2 + \ldots + 2^k a_{2^k} = t_k$$

Thus, $s_n \leqslant t_k$ whenever $n < 2^k$. Hence, if there is a finite number M such that, for every k, $t_k \leqslant M$, then certainly $s_n \leqslant M$. Thus if $\{t_k\}$ is bounded above, $\{s_n\}$ is also. Conversely, if $n > 2^k$, it is clear that

$$\begin{aligned} s_n &\geqslant a_1 + a_2 + (a_3 + a_4) + \ldots + (a_{2^{k-1}-1} + \ldots + a_{2^k}) \\ &\geqslant \frac{1}{2} a_1 + a_2 + 2a_4 + \ldots + 2^{k-1} a_{2^k} = \frac{1}{2} t_k \end{aligned}$$

Therefore, $2s_n \geqslant t_k$ whenever $n > 2^k$. If there is a finite number M which is an upper bound of $\{s_n\}$, then clearly $2M$ must be an upper bound of $\{t_k\}$ and this establishes the result.

We can apply this to the series $\sum_{n=1}^{\infty} 1/n^p$ where p is any number. Clearly, if $p \leqslant 0$, Theorem 10.3 shows this series cannot possibly converge. On the other hand, if $p > 0$, Cauchy's result is applicable. We see that

$$\sum_{k=0}^{\infty} 2^k \frac{1}{(2^k)^p} = \sum_{k=0}^{\infty} 2^{(1-p)k}$$

but this is just a geometric series and so we know that it converges if and only if $1 - p < 0$. Thus we have established that $\sum_{n=1}^{\infty} 1/n^p$ converges if and only if $p > 1$.

There is perhaps an even more important result contained in the proof of

Theorem 10.7 than that given in the statement. For even if it has been established that a given series of numbers converges, it may converge so slowly that it is useless for computational purposes. Thus it is important to be able to decide how closely the "first n terms of a series" approximates its sum if this series is to be used for computation. Hence, if $\sum_{n=1}^{\infty} a_n$ is a series of numbers and s is its sum, we would like to know an upper bound to the number

$$\left| s - \sum_{m=1}^{n} a_m \right|$$

for various values of n. Such estimates can be obtained by using the inequality (10–4) in the proof of Theorem 10.7. Suppose s is the sum of the series $\sum b_n$ in Theorem 10.7.

Since, for $m > n$,

$$\left\| s - \sum_{k=1}^{n} b_k \right\| \leqslant \left\| s - \sum_{k=1}^{m} b_k \right\| + \left\| \sum_{k=n+1}^{m} b_k \right\| \tag{10–7}$$

and since by the definition of convergence,

$$\lim_{m \to \infty} \left\| s - \sum_{k=1}^{m} b_k \right\| = 0$$

we see, by taking the limit as $m \to \infty$ in equations (10–4) and (10–7), that

$$\left\| s - \sum_{k=1}^{n} b_k \right\| \leqslant \lim_{m \to \infty} \left\| \sum_{k=n+1}^{m} b_k \right\| \leqslant \lim_{m \to \infty} \sum_{k=n+1}^{m} c_k = \sum_{k=n+1}^{\infty} c_k$$

where the notation $\sum_{k=n+1}^{\infty} c_k$ is interpreted to mean $\sum_{k=1}^{\infty} c_k - \sum_{k=1}^{n} c_k$.

In particular, if a_n is a series of real numbers whose sum is s, and if $|a_k| < c_k$ for all k greater than a fixed integer N_0, then

$$\left| s - \sum_{k=1}^{n} a_k \right| \leqslant \sum_{k=n+1}^{\infty} c_k \tag{10–8}$$

when $n + 1 \geqslant N_0$.

For example, suppose we wish to determine what the error will be in evaluating the series $\sum_{k=1}^{\infty} \frac{1}{k!3^k}$ by keeping only the first n terms. Then in equation (10–8) we take $a_k = \frac{1}{(k!3^k)}$ and set $c_k = \frac{1}{(n+1)!3^k}$. Then $a_k \leqslant c_k$ for $k \geqslant n+1$. Therefore, the error will be less than

$$\frac{1}{(n+1)!} \sum_{k=n+1}^{\infty} \frac{1}{3^k}$$

or, using equations (10–5) and (10–6), we see that the error introduced by keeping only n terms will be less than

$$\frac{1}{2(n+1)!3^n}$$

We now turn to yet another method of establishing the convergence or divergence of a given infinite series called the "root test." This result was first discovered by Cauchy in 1821. It was lost for a while in the prodigious jungle of his work and rediscovered by Hadamard in 1892.

Theorem 10.8: *Let $\sum b_n$ be an infinite series in a Banach space. Set*

$$\rho = \limsup_{n\to\infty} (\|b_n\|)^{1/n}$$

Then

(a) *If $\rho < 1$, $\sum b_n$ converges.*
(b) If $\rho > 1$, $\sum b_n$ *diverges.*

Proof: Suppose $\rho < 1$ and choose a so that $\rho < a < 1$. Theorem 7.23(b) shows that we can find an integer N such that $n \geqslant N$ implies

$$\|b_n\|^{1/n} < a$$

Hence,

$$\|b_n\| < a^n \qquad \text{for } n \geqslant N$$

Since the series $\sum a^n$ converges for $a < 1$, Theorem 10.7(a) shows that $\sum b_n$ also converges.

Now suppose that $\rho > 1$. Theorem 7.23(a) shows that $\|b_n\|^{1/n} > 1$ for infinitely many n and so $\|b_n\| > 1$ for infinitely many n. We therefore conclude that for some $\epsilon > 0$ there is *no* integer N such that $n \geqslant N$ implies $\|b_n\| < \epsilon$.

Thus, $\{b_n\}$ does not converge to zero and Theorem 10.3 shows that $\sum b_n$ does not converge.

It is easy to see by an example that, in general, nothing can be said about the case $\rho=1$. Thus, $\rho=1$ for both of the infinite series $\sum 1/n$ and the infinite series $\sum 1/n^2$, but we have already shown that the first of these diverges and the second converges.

Definition 10.9: *Let $\sum v_n$ be an infinite series in a normed linear space. If the infinite series $\sum \|v_n\|$ converges, then the infinite series $\sum v_n$ is said to* **converge absolutely.** *If, on the other hand, $\sum v_n$ converges but $\sum \|v_n\|$ diverges, it is said that $\sum v_n$* **converges nonabsolutely** *or is* **nonabsolutely convergent.**

For example, the special logarithmic series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$ is nonabsolutely convergent. (Its sum is ln 2.)

Theorem 10.10: *If $\sum b_n$ is an absolutely convergent infinite series in a Banach space, then $\sum b_n$ converges.*

Proof: This theorem is an immediate consequence of Theorem 10.7(a).

Clearly there is no difference between convergence and absolute convergence for numerical series with positive terms.

We mention that, in addition to the tests for convergence discussed here, there are also the ratio test and Raabe's test which we have not included. An important feature that these tests have in common is that they are all tests for absolute convergence and do not show whether or not a given series is nonabsolutely convergent. In practice it is often difficult to establish except in certain special cases (such as the alternating series which is not discussed here) whether or not a given series is nonabsolutely convergent. Even if it is established that a series is nonabsolutely convergent, the series is still much less desirable for many purposes than an absolutely convergent series. The reason for this is that absolutely convergent series can be manipulated pretty much as finite sums. Thus it is shown in Theorem 10.14 that these series may be rearranged arbitrarily without affecting the sums. We have already shown in Theorem 10.6(b) that any two convergent series can be added together term by term. It is also easy to show, although we shall not do so here, that two absolutely convergent series of complex numbers can be multiplied term by term to obtain a series which converges to the product of the sums. We shall see, however, at least in the case of series with complex numbers (see remarks following Theorem 10.15), that the situation is not nearly as good in the case of nonabsolutely convergent series.

Definition 10.11: *Let* $f: J \to J$ *be a* **bijection** *from the set of positive integers to itself (note that, according to Definition 4.15,* f *is a sequence with values in* J*) and let* $\{v_n\}$ *be a sequence in a normed linear space. Define the sequence* $\{v_n^*\}$ *by*

$$v_n^* = v_{f(n)} \qquad \text{for } n = 1, 2, 3, \ldots$$

Then the infinite series $\sum_{n=1}^{\infty} v_n^*$ *is said to be a* **rearrangement** *of the infinite series* $\sum_{n=1}^{\infty} v_n$.

It is clear that, in general, the partial sums of the series $\sum v_n^*$ will be different from the partial sums of the infinite series $\sum v_n$. Thus it should not be surprising if a rearrangement of a given series does not converge to the same sum as the original series does. We shall now establish under what conditions we might expect all rearrangements of a given infinite series to converge and whether or not they converge to the same sum as the original series.

Definition 10.12: *An infinite series in a normed linear space* **converges unconditionally** *if every rearrangement of this series converges to the same sum.*

The notation used for infinite series suggests that they behave much as finite sums and we have already indicated that, as far as absolutely convergent series are concerned, this is pretty much the case. To see that it is not true for all series, consider the infinite series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$. Writing this out we have

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots \tag{10–9}$$

It is well known, although we shall not prove it here, that this series converges. If we denote its sum by t, an inspection of its terms shows that

$$t < 1 - \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$$

On the other hand, we can rearrange the terms of this series in such a way that two positive terms are always followed by one negative one and obtain the series

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots \tag{10–10}$$

Each consecutive grouping of three terms is represented by the formula

$$\frac{1}{4n-3} + \frac{1}{4n-1} - \frac{1}{2n} \qquad n = 1, 2, 3, \ldots$$

It is clear that

$$\frac{1}{4n-3} + \frac{1}{4n-1} - \frac{1}{2n} = \frac{1}{4n}\left(\frac{3}{4n-3} + \frac{1}{4n-1}\right) > 0$$

If we denote the nth partial sum of this series by σ_n, we see from the preceding considerations that

$$\sigma_3 < \sigma_6 < \sigma_9 < \cdots$$

and this shows that

$$\limsup_{n \to \infty} \sigma_n > \sigma_3 = \frac{5}{6}$$

Therefore the infinite series (10–9) and (10–10) cannot possibly converge to the same sum. It is clear, from the discussion of the series $\sum 1/n^p$, that the series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$ does not converge absolutely. We will see from the next theorem, which was first proved by Riemann, that the type of behavior illustrated here is in fact quite general.

Theorem 10.13: *If $\sum c_n$ is any nonabsolutely convergent series of real numbers and $\alpha \leqslant \beta$ are any two extended real numbers, there is a rearrangement* *of* $\sum c_n$ *whose sequence of partial sums* $\{A_n\}$ *has the property that*

$$\liminf_{n\to\infty} A_n = \alpha \tag{10–11}$$

and

$$\limsup_{n\to\infty} A_n = \beta \tag{10–12}$$

Proof: For each positive integer n, let us define the numbers c_n^+ and c_n^- by

$$c_n^+ = \max\{c_n, 0\}$$

$$c_n^- = -\min\{c_n, 0\}$$

It is clear that $c_n = c_n^+ - c_n^-$, $|c_n| = c_n^+ + c_n^-$ and that c_n^+, c_n^- are nonnegative. Theorem 10.6(b) shows that, if both $\sum c_n^+$ and $\sum c_n^-$ converge, then $\sum (c_n^+ + c_n^-) = \sum |c_n|$ must also converge. Since this is contrary to hypothesis, at least one of these series must diverge. On the other hand, for every positive integer n,

$$\sum_{k=1}^{n} c_k = \sum_{k=1}^{n} (c_k^+ - c_k^-) = \sum_{k=1}^{n} c_k^+ - \sum_{k=1}^{n} c_k^-$$

and, thus, if only one of the series $\sum c_n^+$ and $\sum c_n^-$ diverges, then $\sum c_n$ would have to diverge also. Since this is also contrary to the hypothesis, we conclude that both $\sum c_n^+$ and $\sum c_n^-$ diverge.

Let $d_1^+, d_2^+, d_3^+, \ldots$ be the nonnegative terms of $\sum c_n$ taken in the order in which they occur and let $-d_1^-, -d_2^-, -d_3^-, \ldots$ be the negative terms of $\sum c_n$ taken in the order in which they occur in the series. It is clear that the only difference between the series $\sum d_n^+$ and $\sum c_n^+$ and between the series $\sum d_n^-$ and $\sum c_n^-$ are terms which are zeros, and since these cannot affect the sum, we conclude that both $\sum d_n^+$ and $\sum d_n^-$ diverge. Let D_n^+ and D_n^- denote the nth partial sums of $\sum d_n^+$ and $\sum d_n^-$, respectively. Clearly, $D_n^+ \to +\infty$ and $D_n^- \to +\infty$ since they are both divergent monotonically increasing sequences. Hence, by taking n sufficiently large, we can find a D_n^+ and a D_n^- larger than any real number.

Let $\{\alpha_n\}$ and $\{\beta_n\}$ be sequences of finite real numbers such that $\alpha_n \to \alpha$, $\beta_n \to \beta$, $\beta_1 > 0$ and, for every n, $\alpha_n < \beta_n$. We can now define two increasing sequences of integers $\{r_n\}$ and $\{s_n\}$ inductively by the following procedure.

Let r_1 and s_1 be the smallest positive integers such that

$$D^+_{r_1} > \beta_1$$

$$D^+_{r_1} - D^-_{s_1} < \alpha_1$$

Having chosen $r_1, \ldots, r_{n-1}$ and $s_1, \ldots, s_{n-1}$ such that, for $1 \leqslant j \leqslant n-1$, r_j and s_j are the smallest integers for which

$$D^+_{rj} - D^-_{sj-1} > \beta_j$$

$$D^+_{rj} - D^-_{sj} < \alpha_j$$

we let r_n and s_n be the smallest integers such that

$$D^+_{rn} - D^-_{sn-1} > \beta_n$$

$$D^+_{rn} - D^-_{sn} < \alpha_n$$

With these sequences of integers, we construct the following rearrangement of $\sum c_n$:

$$d_1^+ + d_2^+ + \ldots + d^+_{r_1} - d_1^- - d_2^- - \ldots - d_{s_1} + d^+_{r_1+1}$$
$$+ d_{r_1+2} + \ldots + d_{r_2} - d_{s_1+1} - d_{s_1+2} - \ldots - d_{s_2} + \ldots$$

Now let A_n denote the nth partial sum of this series. Then for every positive integer n,

$$\left.\begin{aligned} A_{(rn+sn-1)} &= D^+_{rn} - D^-_{sn-1} > \beta_n \\ A_{(rn+sn)} &= D^+_{rn} - D^-_{sn} < \alpha_n \end{aligned}\right\} \tag{10–13}$$

and, since r_n and s_n are the smallest integers for which this is true,

$$\left.\begin{aligned} A_{(rn+sn-1)} - d^+_{rn} &\leqslant \beta_n \\ A_{(rn+sn)} + d^-_{sn} &\geqslant \alpha_n \end{aligned}\right\} \tag{10–14}$$

Clearly $\{A_{(rn+sn-1)}\}$ and $\{A_{(rn+sn)}\}$ are subsequences of $\{A_n\}$. Hence, equations (10–13) show that

$$\limsup_{n\to\infty} A_n \geqslant \limsup_{n\to\infty} A_{(r_n+s_{n-1})} \geqslant \limsup_{n\to\infty} \beta_n = \beta \tag{10–15}$$

$$\liminf_{n\to\infty} A_n \leqslant \liminf_{n\to\infty} A_{(r_n+s_n)} \leqslant \liminf_{n\to\infty} \alpha_n = \alpha \tag{10–16}$$

If $\beta=+\infty$, we see that equation (10–12) is true and, if $\alpha=-\infty$, equation (10–11) is true.

Assume first that β is finite and let $\epsilon>0$ be given. Since $\sum c_n$ converges, Theorem 10.3 shows that there exists an integer N_1 such that $d_{r_n}^+<\epsilon/2$ for all $n\geqslant N_1$ (recall that $\{r_n\}$ is monotonic). The condition $\beta_n\to\beta$ shows that there is an integer N_2 such that for all $n\geqslant N_2$, $\beta_n<\beta+\epsilon/2$. Hence, if $N=\max\{N_1, N_2\}$, it follows from equations (10–14) that, for $n\geqslant N$,

$$A_{(r_n+s_{n-1})}<\beta_n+\frac{\epsilon}{2}<\beta+\epsilon \tag{10–17}$$

For any $k\geqslant r_N+s_{N-1}$, there is a unique integer $n\geqslant N$ such that $r_n+s_{n-1}\leqslant k<r_{n+1}+s_n$. The method of construction shows that

$$A_k\leqslant \max\{A_{(r_n+s_{n-1})}, A_{(r_{n+1}+s_n)}\}$$

and, since equation (10–17) shows that both of these are less than $\beta+\epsilon$, we conclude that $A_k<\beta+\epsilon$ for all $k\geqslant r_N+s_{N-1}$. Hence,

$$\limsup_{k\to\infty} A_k\leqslant\beta+\epsilon$$

and, since ϵ was arbitrary,

$$\limsup_{k\to\infty} A_k\leqslant\beta$$

Combining this with equation (10–15) shows that

$$\limsup_{n\to\infty} A_n=\beta$$

An almost identical argument shows that, when α is finite,

$$\liminf_{n\to\infty} A_n=\alpha$$

The proofs for the two cases $\beta=-\infty$ and $\alpha=+\infty$ follow by similar arguments.

Theorem 10.14: *If $\sum b_n$ is an absolutely convergent infinite series in a Banach space, then it also converges unconditionally.*

Proof: Suppose $\sum b_n^*$ is a rearrangement of $\sum b_n$. Let the nth partial sum of $\sum b_n$ be s_n and the nth partial sum of $\sum b_n^*$ be s_n^*. Let $\epsilon>0$ be given. Then there is an integer N such that $m \geqslant n \geqslant N$ implies

$$\sum_{i=n}^{m} \|b_i\| < \frac{\epsilon}{2} \tag{10–18}$$

Now, there is a bijective function $f: J \to J$ such that $b_n^* = b_{f(n)}$ for every positive integer n. Since f is surjective, it follows that $f(f^{-1}(Y)) = Y$ for every $Y \subset J$. In particular, let $Y = \{1, 2, 3, \ldots, N\}$. Since f is injective, $f^{-1}(Y)$ contains at most N positive integers; that is, $f^{-1}(Y)$ is a finite set of positive integers. Hence we can choose an integer p such that $p > m$ for all $m \in f^{-1}(Y)$. Then,

$$f^{-1}(Y) \subset \{1, 2, 3, \ldots, p\}$$

Therefore

$$\{1, 2, 3, \ldots, N\} = f(f^{-1}(Y)) \subset f(\{1, 2, 3, \ldots, p\}) = \{f(1), f(2), \ldots, f(p)\}$$

Thus, if $i > p$, the vectors $b_1, \ldots, b_N$ will cancel in the difference

$$s_i - s_i^* = \sum_{j=1}^{i} b_j - \sum_{j=1}^{i} b_{f(j)}$$

Hence, we conclude from equation (10–18) that, for $i > p$,

$$\|s_i - s_i^*\| \leqslant \sum_{j=N+1}^{k} \|b_j\| < \frac{\epsilon}{2}$$

where k is some integer greater than N. Theorem 10.10 shows that $\sum b_n$ converges. If we let s be its sum then there is an integer P such that $i \geqslant P$ implies

$$\|s_i - s\| < \frac{\epsilon}{2}$$

Hence, for $j \geqslant \max\{P, p\}$ we find that

$$\|s_j^* - s\| \leqslant \|s_j^* - s_j\| + \|s_j - s\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

That is, the infinite series $\sum b_n^*$ converges to the same sum as $\sum b_n$.

In the case of infinite series of complex numbers we have the following converse to Theorem 10.14.

Theorem 10.15: *Let $\{a_n\}$ be a sequence of complex numbers. If $\sum a_n$ converges unconditionally, then it converges absolutely.*

Proof: Suppose $\sum a_n$ converges nonabsolutely and let x_n and y_n be the real and imaginary parts of a_n, respectively; that is, $a_n = x_n + iy_n$.

Since $|a_n| \leqslant |x_n| + |y_n|$ and since, by assumption, $\sum |a_n|$ diverges, Theorem 10.7(b) shows that $\sum (|x_n| + |y_n|)$ diverges. We conclude that either $\sum |x_n|$ diverges or $\sum |y_n|$ diverges for if both converged, Theorem 10.6(b) would show that $\sum (|x_n| + |y_n|)$ converged also.

Since $\sum a_n$ converges by assumption, we conclude from Theorem 10.6(a) that $\sum x_n$ converges and $\sum y_n$ converges. So either $\sum x_n$ or $\sum y_n$ converges nonabsolutely. We see, therefore, from Theorem 10.13 that there is a rearrangement $\sum a_n^*$ of $\sum a_n$ such that if x_n^* and y_n^* are the real and imaginary parts, respectively, of a_n^*, either $\sum x_n^*$ or $\sum y_n^*$ diverges. Now since Theorem 10.6(a) shows that the convergence of $\sum a_n^*$ implies the convergence of both $\sum x_n^*$ and $\sum y_n^*$, we conclude that $\sum a_n^*$ diverges. Hence $\sum a_n$ does not converge unconditionally.

We have in fact proved more than the statement of the theorem. We have shown that if a sequence of complex numbers $\sum a_n$ converges nonabsolutely, then there is a divergent rearrangement of $\sum a_n$. Now Theorem 10.14 shows that, if $\sum a_n$ converges absolutely, every rearrangement of $\sum a_n$ converges to the same sum. Since, if any given infinite series converges, it either converges absolutely or it converges nonabsolutely, we conclude that the following theorem holds.

Theorem 10.16: *If $\sum a_n$ is an infinite series of complex numbers all of whose rearrangements converge, then they necessarily all converge to the same sum.*

This theorem shows that for infinite series of complex numbers the phrase "to the same sum" can be omitted in Definition 10.12.

CHAPTER 11

Sequences of Functions and Function Spaces

It was shown in chapter 8 that if E is any set and X is a set with a certain "algebraic structure," then this "algebraic structure" on X "induces" an "algebraic structure" on the family $\mathscr{F}(E, X)$ of all functions from E to X. This led in a natural way to the construction of linear spaces whose points are functions.[35] In this chapter we show that if E is an arbitrary set and X is *any metric space* then the "metric structure" on X "induces" a "metric structure" on $\mathscr{F}(E, X)$. As in chapter 8 this leads to the construction of metric spaces whose points are functions. The "powerful machinery" which has been developed in the course of our study of metric spaces is then applied to these function spaces and some very useful results are obtained. This is one of the major justifications for studying the theory of metric spaces.

As in chapter 8 we shall also be concerned herein with the family $\mathscr{B}(E,X)$ of all bounded functions from E to X and, when E is also a metric space, with the families $\mathscr{C}(E,X)$ of all continuous functions from E to X and $\mathscr{C}^{\infty}(E,X)$ of all bounded continuous functions from E to X.[36]

We shall approach the material in this chapter by considering sequences of functions. To this end let E be any set and let $\langle X, d\rangle$ be an arbitrary metric space. Let $\{f_n\}$ be a sequence in $\mathscr{F}(E,X)$. Evidently, *each term of this sequence is a function*.[35] We now ask how we might define convergence for this sequence in a useful way. Perhaps the most obvious way to do this is to say that $\{f_n\}$ converges if there is a function $g:E\rightarrow X$ such that, for every $x\in E$,

$$\lim_{n\to\infty} f_n(x)=g(x)$$

This is the notion of *pointwise* convergence. There is, however, another way

[35] See the last sentence of the paragraph immediately following Definition 4.1.

[36] The notation $\mathscr{F}(E,X)$, $\mathscr{B}(E,X)$, $\mathscr{C}(E,X)$, and $\mathscr{C}^{\infty}(E,X)$ will always be used to denote, respectively, the family of all functions from E to X, the family of all bounded functions from E to X, the family of all continuous functions from E to X, and the family of all bounded continuous functions from E to X except when X is the real or complex numbers with the usual metric in which case the notation $\mathscr{F}(\bar{E})$, $\mathscr{B}(E)$, $\mathscr{C}(E)$, and $\mathscr{C}^{\infty}(E)$ will sometimes be used.

to define convergence which is analogous to uniform continuity of functions. In view of this we make the following definition.

Definition 11.1: *Let $\{f_n\}$ be a sequence of functions from a set E to a metric space $\langle X, d\rangle$. If there exists a $g: E \to X$ such that, for every $x \in E$ and for every $\epsilon > 0$, there exists an integer N* **depending on** *x and ϵ for which*

$$d(f_n(x), g(x)) < \epsilon \tag{11–1}$$

whenever $n \geq N$, we say that $\{f_n\}$ converges to g **pointwise** *on E.*

If, on the other hand, there exists a **single** *integer N for* **each** *$\epsilon > 0$ such that, whenever $n \geq N$, equation (11–1) holds for* **all** *$x \in E$, we say that $\{f_n\}$ converges to g* **uniformly** *on E and that g is the* **uniform limit** *of $\{f_n\}$.*

Of course, uniform convergence implies pointwise convergence. For uniform convergence, we can for each ϵ find a single N which will do for all x; whereas, for pointwise convergence, we have to use different N's for different x's in order for equation (11–1) to hold. Uniform convergence of a sequence is, generally speaking, a more desirable property than just pointwise convergence. This is due, at least partially, to a fact which we shall prove subsequently; namely, uniform convergence is equivalent to the convergence of a sequence of points in a suitable metric space. In general, this is not true for sequences of functions which are only pointwise convergent.

Apropos of this remark we will prove the following theorems.

Theorem 11.2: *Let $\{f_n\}$ be a sequence of functions from a set E to a metric space $\langle X, d\rangle$. Then $\{f_n\}$ converges uniformly to g if and only if, for every $\epsilon > 0$, there exists an integer N such that*

$$\operatorname*{lub}_{x \in E} d(f_n(x), g(x)) < \epsilon \tag{11–2}$$

whenever $n \geq N$.

Proof: It is clear that, if equation (11–2) holds,

$$d(f_n(x), g(x)) < \epsilon$$

for all $x \in E$. Hence equation (11–2) implies $\{f_n\}$ converges uniformly to g.

Suppose $\{f_n\}$ converges uniformly to g. Then, given $\epsilon > 0$, we can find an N such that, for every $x \in E$,

$$d(f_n(x), g(x)) < \frac{\epsilon}{2}$$

whenever $n \geqslant N$. Hence,

$$\operatorname*{lub}_{x \in E} d(f_n(x), g(x)) \leqslant \frac{\epsilon}{2} < \epsilon$$

which shows that the uniform convergence of $\{f_n\}$ implies equation (11–2).

In everything done so far, we have always taken the distance between any two points in a metric space to be a finite number, but, as pointed out in chapter 6, we can allow the distance to take on values in the extended real number system. In what follows we shall allow possibly infinite metrics but when this is done it will always be stated explicitly! We note, in passing, that all the definitions given for metric spaces with finite metrics carry over to the case where the metric can be infinite and that the proofs of most of the theorems do also.

Theorem 11.3: (a) *Let $\mathscr{F}(E, X)$ be the family of* **all** *functions from the set E to the metric space $\langle X, d\rangle$. Then $\mathscr{F}(E, X)$ together with the function Δ defined by*

$$\Delta(f, g) = \operatorname*{lub}_{x \in E} d(f(x), g(x)) \qquad \text{for all } f, g \in \mathscr{F}(E, X) \qquad (11\text{–}3)$$

is a metric space (with possibly infinite metric).

(b) *Let $\mathscr{B}(E, X)$ be the family of all* **bounded** *functions from the set E to the metric space $\langle X, d\rangle$. Then $\mathscr{B}(E, X)$ together with the restriction*[37] *of the function Δ defined by equation (11–3) is a metric space (with finite distance).*

Proof: Part (a). We must show that Δ satisfies conditions (a) to (c) of Definition 6.1.

(a) It is clear that $\Delta(f, g) \geqslant 0$ since, for each $x \in E, d(f(x), g(x))$ is. Since $0 \leqslant d(f(x), g(x)) \leqslant \Delta(f, g)$ for every $x \in E$, it is also clear that $\Delta(f, g) = 0$ implies $f(x) = g(x)$ for every $x \in E$ and, therefore, that $f = g$. In view of the fact that $d(f(x), f(x)) = 0$ for every $x \in E$, it follows that $\Delta(f, f) = 0$.

[37] To $\mathscr{B}(E, X) \times \mathscr{B}(E, X)$.

(b) It is obvious from equation (11–3) that

$$\Delta(f, g) = \Delta(g, f)$$

(c) Since, for any $f, g, h \in \mathscr{F}(E, X)$,

$$d(f(x), g(x)) \leq d(f(x), h(x)) + d(h(x), g(x))$$

for every $x \in E$, we see that

$$\begin{aligned}\operatorname*{lub}_{x \in E} d(f(x), g(x)) &\leq \operatorname*{lub}_{x \in E} [d(f(x), h(x)) + d(h(x), g(x))] \\ &\leq \operatorname*{lub}_{x \in E} d(f(x), h(x)) + \operatorname*{lub}_{x \in E} d(h(x), g(x))\end{aligned}$$

Hence,

$$\Delta(f, g) \leq \Delta(f, h) + \Delta(h, g)$$

for all $f, g, h \in \mathscr{F}(E, X)$.

Part (b). It follows from part (a) that the restriction of Δ to $\mathscr{B}(E, X) \times \mathscr{B}(E, X)$ is a metric for $\mathscr{B}(E, X)$.[38] Hence we need only show that this restriction of Δ is a finite metric. To this end, choose f and g in $\mathscr{B}(E, X)$. Evidently $d(f(E)) < \infty$ and $d(g(E)) < \infty$. Fix $x_0 \in E$. Then

$$\begin{aligned}d(f(x), g(x)) &\leq d(f(x), f(x_0)) + d(f(x_0), g(x_0)) \\ &\quad + d(g(x_0), g(x)) \leq d(f(E)) + d(f(x_0), g(x_0)) + d(g(E))\end{aligned}$$

Hence

$$\Delta(f, g) = \sup_{x \in E} d(f(x), g(x)) \leq d(f(E)) + d(g(E)) + d(f(x_0), g(x_0))$$

Since d is a finite metric for X, the right side of the inequality is clearly a finite number and the theorem is proved.

Figure 11–1 illustrates the concept of a ball in the metric space $\langle \mathscr{F}(E), \Delta \rangle$, where $\mathscr{F}(E)$ is the family of all real functions defined on the interval $[0, 1]$.

In this chapter the symbol Δ will always be used to denote the metric (or any of its restrictions) defined by equation (11–3). It is now clear that, *if $\{f_n\}$ is a sequence of functions from a set E to a metric space $\langle X, d \rangle$, then the uniform convergence of $\{f_n\}$ to a function $g: E \to X$ is equivalent to the convergence of the sequence $\{f_n\}$ of points of the metric space $\langle \mathscr{F}(E, X), \Delta \rangle$ to the point g of $\langle \mathscr{F}(E, X), \Delta \rangle$* in the sense of Definition 7.1. For this reason Δ is sometimes called the *metric of uniform convergence.* It should be noted

[38] $\langle \mathscr{B}(E, X), \Delta \rangle$ is, of course, a subspace of $\langle \mathscr{F}(E, X), \Delta \rangle$.

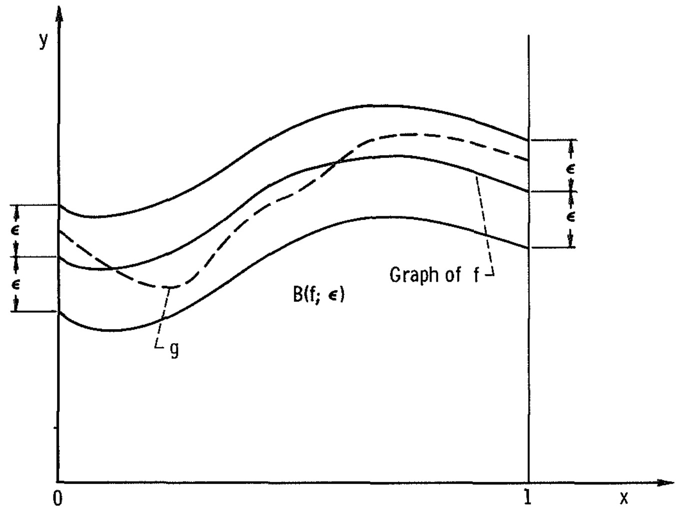


FIGURE 11–1.—Illustration of balls in $\mathscr{F}([0, 1])$ (family of functions defined on interval $[0, 1]$) $B(f; \epsilon)=\{g \in \mathscr{F}([0, 1]) \mid \Delta(f, g) < \epsilon\}$; $\Delta(f, g)= \operatorname{lub}_{x\in[0, 1]} |f(x)-g(x)|$.

that although all of the theorems about metric spaces which we have developed previously apply to the metric space $\langle \mathscr{B}(E, X), \Delta\rangle$, they cannot always be directly applied to $\langle \mathscr{F}(E, X), \Delta\rangle$ since we have allowed an infinite distance between points in $\mathscr{F}(E, X)$. However, many of these theorems do hold for $\langle \mathscr{F}(E, X), \Delta\rangle$. For example, as indicated in chapter 9, *every convergent sequence is a Cauchy sequence even in a metric space with a possibly infinite metric.*

Theorem 11.4: *Suppose $\langle X, d\rangle$ is a* **complete** *metric space and E is any set. Then:*

(a) *$\langle \mathscr{F}(E, X), \Delta\rangle$ is a complete metric space (with a possibly infinite metric).*

(b) *$\langle \mathscr{B}(E, X), \Delta\rangle$ is a complete metric space (with finite metric).*

Proof: Part (a). Let $\{f_n\}$ be any Cauchy sequence in $\langle \mathscr{F}(E, X), \Delta\rangle$. Since, for every $x \in E$, $d(f_m(x), f_n(x)) \leqslant \Delta(f_m, f_n)$, it follows that, for each fixed $x \in E$, the sequence $\{f_n(x)\}$ is a Cauchy sequence of points of X. Since X is complete, this sequence converges. Hence we can define a function $f:E \to X$ by

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad \text{for each } x \in E$$

We shall show that $f_n \to f$ in $\langle \mathscr{F}(E, X), \Delta\rangle$. To this end let $\epsilon' > 0$ be given.

For each $x \in E$, we can find an integer $N'(x)$ such that $d(f(x), f_m(x)) < \epsilon'$ whenever $m \geq N'(x)$. Now, for any integer n and for any integer $m \geq N'(x)$ equation (6–4) implies

$$|d(f_n(x), f_m(x)) - d(f_n(x), f(x)) \leq d(f_m(x), f(x)) < \epsilon'$$

This shows that for each $x \in E$

$$\lim_{m \to \infty} d(f_n(x), f_m(x)) = d(f_n(x), f(x)) \tag{11–4}$$

Now fix $\epsilon > 0$ and choose a positive integer N such that, for all m, $n \geq N$,

$$\Delta(f_m, f_n) < \frac{\epsilon}{2}$$

It follows from this that, for every $x \in E$, and every $m, n \geq N$,

$$d(f_m(x), f_n(x)) < \frac{\epsilon}{2}$$

Combining this with equation (11–4) shows that for each fixed $n \geq N$

$$d(f_n(x), f(x)) = \lim_{m \to \infty} d(f_m(x), f_n(x)) \leq \frac{\epsilon}{2}$$

for every $x \in E$. Hence, $\epsilon/2$ is an upper bound of the set

$$\{d(f_n(x), f(x)) \mid x \in E\}$$

and so

$$\Delta(f_n, f) = \operatorname*{lub}_{x \in E} d(f_n(x), f(x)) \leq \frac{\epsilon}{2} < \epsilon$$

for any $n \geq N$. This proves that $f_n \to f$ in $\langle \mathscr{F}(E, X), \Delta \rangle$ and, since $\{f_n\}$ was any Cauchy sequence in $\langle \mathscr{F}(E, X), \Delta \rangle$, this proves that $\langle \mathscr{F}(E, X), \Delta \rangle$ is complete.

Part (b). Let $\{f_n\}$ be a Cauchy sequence in $\langle \mathscr{B}(E, X), \Delta \rangle$. Clearly $\{f_n\}$ is also a Cauchy sequence in $\langle \mathscr{F}(E, X), \Delta \rangle$. Hence it follows from part (a) that $\{f_n\}$ converges to a point f of $\mathscr{F}(E, X)$. Thus if we can show that $f \in \mathscr{B}(E, X)$, then we can conclude that $\{f_n\}$ converges in $\langle \mathscr{B}(E, X), \Delta \rangle$.

Evidently there exists an integer N such that $d(f_N(x), f(x)) < \epsilon$ for all $x \in E$. Hence

$$d(f(x), f(y)) \leqslant d(f(x), f_N(x)) + d(f_N(x), f_N(y)) + d(f_N(y), f(y)) < \epsilon + d(f_N(E)) + \epsilon$$

for every $x, y \in E$. Therefore,

$$d(f(E)) = \sup_{\substack{f(x) \in f(E) \\ f(y) \in f(E)}} d(f(x), f(y)) \leqslant 2\epsilon + d(f_N(E))$$

Since f_N is bounded, the right side of this inequality is a finite number and this implies

$$f \in \mathscr{B}(E, X)$$

When the set E of Definition 11.1 is also a metric space, we can consider sequences of continuous functions. This leads us to a consideration of the family $\mathscr{C}(E, X)$ of all continuous functions from E to X and of the family $\mathscr{C}^\infty(E, X)$ of all bounded continuous functions from E to X. Corollary 1 of Theorem 8.20 shows that **if E is a compact metric space then $\mathscr{C}(E, X) = \mathscr{C}^\infty(E, X)$.**

There is however another family of continuous functions which is also of some interest. This is introduced in the following definition.

Definition 11.5: *Let f be a continuous function from a metric space $\langle E, d'\rangle$ to a metric space $\langle X, d\rangle$. We say f* **vanishes at infinity** *if, for any $\epsilon > 0$, there exists a compact set $K \subset E$ such that $d(f(K^c)) < \epsilon$. Let $\mathscr{C}_0(E, X)$ be the family of all continuous functions from E to X which vanish at infinity.*

Theorem 11.6: *If $\{f_n\}$ is a sequence of continuous functions from a metric space $\langle E, d'\rangle$ to a metric space $\langle X, d\rangle$ which converges uniformly to a function $f: E \to X$, then f is continuous.*

Proof: Choose a point x of E and fix $\epsilon > 0$. Since $\{f_n\}$ converges uniformly, we can find a positive integer N such that, for *all* points $y \in E$,

$$d(f_n(y), f(y)) < \frac{1}{3}\epsilon \tag{11–5}$$

as soon as $n \geqslant N$. Now, for each n, the continuity of f_n at x shows that there is a $\delta > 0$ such that

$$f_n(B(x;\,\delta)) \in B\left(f_n(x);\frac{\epsilon}{3}\right)$$

This shows that for any $t \in B(x;\,\delta)$,

$$d(f_n(t),f_n(x)) < \frac{\epsilon}{3} \tag{11–6}$$

Now for $n \geq N$ and $t \in B(x;\,\delta)$

$$d(f(x),f(t)) \leq d(f(x),f_n(x)) + d(f_n(x),f_n(t)) + d(f_n(t),f(t))$$

Equation (11–5) shows that the first and last terms on the right of this inequality are both less than $\epsilon/3$ and equation (11–6) shows that the second term is also. Hence $d(f(x),\,f(t)) < \epsilon$ for all $t \in B(x;\,\delta)$; that is, $f(t) \in B(f(x);\,\epsilon)$ which proves that f is continuous, since x was an arbitrary point of E.

To see that the sequence *must* be uniformly convergent in order for the theorem to hold, we need only consider the pointwise convergent sequence of continuous real valued functions $\{f_n\}$ defined on R^1 by

$$f_n(x) = \frac{1}{1+nx^2} \qquad \text{for all } x \in R^1;\ n=1,2,3\ \ldots \tag{11–7}$$

Since

$$\lim_{n\to\infty} f_n(0) = \lim_{n\to\infty} 1 = 1$$

and since, for $x \equiv 0$,

$$\lim_{n\to\infty} f_n(x) = 0$$

we see that this sequence converges to a discontinuous function.

Corollary 1: *Let $\langle E,\,d'\rangle$ and $\langle X,\,d\rangle$ be metric spaces.*

(a) *The family $\mathscr{C}^\infty(E,X)$ is a closed subset of the metric space $\langle \mathscr{B}(E,X),\Delta\rangle$.*

(b) *If $\langle X,\,d\rangle$ is a complete metric space, then $\langle \mathscr{C}^\infty(E,X),\,\Delta\rangle$ is a complete metric space.*

(c) *The family $\mathscr{C}_0(E,X)$ is a closed subset of the metric space $\langle \mathscr{C}^\infty(E,X),\Delta\rangle$.*

(d) *If $\langle X,\,d\rangle$ is a complete metric space, then $\langle \mathscr{C}_0(E,X),\,\Delta\rangle$ is a complete metric space.*

(e) *If $\langle X,\,d\rangle$ is a complete metric space and $\langle E,\,d'\rangle$ is a compact metric space, then $\langle \mathscr{C}(E,X),\,\Delta\rangle$ is a complete metric space.*

Proof: Part (a). Clearly $\mathscr{C}^\infty(E, X)$ is a subset of the metric space $\langle \mathscr{B}(E, X), \Delta \rangle$. Let f be a limit point of $\mathscr{C}^\infty(E, X)$. Then Theorem 7.4 shows that there is a sequence $\{f_n\}$ of points of $\mathscr{C}^\infty(E, X)$ which converges to f (recall that every limit point is an adherence point). Since f_n is continuous for every n, Theorem 11.6 shows that f is continuous. Now f is obviously bounded and so $f \in \mathscr{C}^\infty(E, X)$. Hence, in view of the fact that f was any limit point of $\mathscr{C}^\infty(E, X)$, this shows that $\mathscr{C}^\infty(E, X)$ is closed.

Part (b). Theorem 11.4 shows that $\langle \mathscr{B}(E, X), \Delta \rangle$ is complete. Hence part (a) and Theorem 9.9 show that the subspace $\langle \mathscr{C}^\infty(E, X), \Delta \rangle$ is complete.

Part (c). Let g be any point of $\mathscr{C}_0(E, X)$. Clearly g is continuous and there exists a compact subset K_1 of E such that $d(g(K_1^c)) < 1$. Corollary 1 of Theorem 8.20 shows that there exists a finite real number M such that $d(g(K_1)) < M$. Since

$$g(E) = g(K_1 \cup K_1^c) = g(K_1) \cup g(K_1^c)$$

it is easy to verify that

$$d(g(E)) \leqslant d(g(K_1)) + d(g(K_1^c)) + d(g(K_1), g(K_1^c)) < 1 + M \\ + d(g(K_1), g(K_1^c))$$

if both $g(K_1)$ and $g(K_1^c)$ are nonempty and that

$$d(g(E)) \leqslant d(g(K_1)) + d(g(K_1^c)) < 1 + M$$

if either $g(K_1)$ or $g(K_1^c)$ is empty. Since $d(g(K_1), g(K_1^c))$ is the greatest lower bound of a set of finite positive numbers it is certainly finite and so these last two inequalities show that $d(g(E))$ is finite. Hence g is bounded and continuous. Since g was arbitrary this shows that $\mathscr{C}_0(E, X)$ is a subset of the metric space $\langle \mathscr{C}^\infty(E, X), \Delta \rangle$.

Now let f be a limit point of $\mathscr{C}_0(E, X)$ (in the metric space $\langle \mathscr{C}^\infty(E, X), \Delta \rangle$). Then f is certainly continuous and Theorem 7.4 shows that there exists a sequence $\{f_n\}$ of points of $\mathscr{C}_0(E, X)$ which converges to f. Let $\epsilon > 0$ be given and choose N so large that $\Delta(f, f_N) < \epsilon/3$. Since f_N vanishes at infinity there exists a compact subset K of E such that $d(f_N(K^c)) < \epsilon/3$. Hence for every $x, y \in K^c$

$$d(f(x), f(y)) \leqslant d(f(x), f_N(x)) + d(f_N(x), f_N(y)) + d(f_N(y), f(y)) \\ \leqslant \Delta(f, f_N) + d(f_N(K^c)) + \Delta(f_N, f) < \epsilon$$

Therefore,

$$d(f(K^c)) = \operatorname*{lub}_{\substack{x \in K^c \\ y \in K^c}} d(f(x), f(y)) \leqslant \epsilon$$

Hence f vanishes at infinity and therefore $f \in \mathscr{C}_0(E, X)$. Since f was any limit point of $\mathscr{C}_0(E, X)$ this shows that $\mathscr{C}_0(E, X)$ is closed.

Part (d). We see from part (b) that $\langle \mathscr{C}^\infty(E, X), \Delta \rangle$ is a complete metric space. Hence part (c) and Theorem 9.9 show that the subspace $\langle \mathscr{C}_0(E, X), \Delta \rangle$ is complete.

Part (e). This is an immediate consequence of the remarks preceding Definition 11.5 and of part (b).

The space of continuous functions was first discussed by F. Riesz in 1918.

Let x be a limit point of the metric space $\langle E, d' \rangle$. According to Theorem 8.4, a function $f: E \to X$ is continuous at x if and only if

$$\lim_{t \to x} f(t) = f(x)$$

Now if $\{f_n\}$ is a sequence of continuous functions from E to X, it is clear that $f_n(x) = \lim_{t \to x} f_n(t)$. But if the sequence $\{f_n\}$ converges either pointwise or uniformly, then there is a function $f: E \to X$ such that $\lim_{n \to \infty} f_n(t) = f(t)$ for every $t \in E$. So, in particular, $f(x) = \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \lim_{t \to x} f_n(t)$. Hence, if the limit f of the sequence $\{f_n\}$ is continuous at x, then

$$\lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{n \to \infty} \lim_{t \to x} f_n(t)$$

That is, the order in which the limit processes are carried out is immaterial. In view of these remarks the following is an immediate corollary to Theorem 11.6.

Corollary 2: *Let $\{f_n\}$ be a uniformly convergent sequence of continuous functions from the metric space $\langle E, d' \rangle$ to a metric space $\langle X, d \rangle$ and let $x \in E$ be a limit point of E. Then*

$$\lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{n \to \infty} \lim_{t \to x} f_n(t)$$

If $\{f_n\}$ is the sequence of continuous functions defined for every real number x by equation (11–7), then

$$\lim_{t \to 0} \lim_{n \to \infty} f_n(t) = \lim_{t \to 0} 0 = 0$$

and

$$\lim_{n\to\infty} \lim_{t\to 0} f_n(t) = \lim_{n\to\infty} 1 = 1$$

so that *uniform* convergence of a sequence is in general necessary to be able to interchange the limit processes. However, this is not always the case. In fact there is no converse to Theorem 11.6; that is, a sequence of continuous functions may converge to a continuous function even though the convergence is not uniform. This can be verified by considering the following example.

Let $\{f_n\}$ be the sequence in $\mathscr{C}([x, 1])$, which is defined by

$$f_n(x) = n^2x(1-x^2)^n \qquad 0 \leqslant x \leqslant 1,\ n = 1, 2, 3, \ldots$$

For $0 < x \leqslant 1$, it follows from the last equation of chapter 7 that

$$\lim_{n\to\infty} f_n(x) = 0$$

Since $f_n(0) = 0$, we see that

$$\lim_{n\to\infty} f_n(x) = 0 \qquad 0 \leqslant x \leqslant 1$$

Suppose the convergence of $\{f_n\}$ is uniform. Then Theorem 11.2 shows that, given $\epsilon > 0$, there exists an N such that $n \geqslant N$ implies

$$\operatorname*{lub}_{0\leqslant x\leqslant 1} |f_n(x)| < \epsilon \tag{11–8}$$

Since $[0, 1]$ is compact, we know, from the second corollary to Theorem 8.20, that for each $n, f_n(x)$ is equal to $\operatorname*{lub}_{0\leqslant x\leqslant 1} |f_n(x)|$ at some point $x \in [0, 1]$. This point obviously occurs where $f_n(x)$ is a maximum, that is, at $x = 1/\sqrt{2n+1}$. Thus,

$$\operatorname*{lub}_{0\leqslant x\leqslant 1} |f_n(x)| = \frac{2^n n^{(n+2)}}{(2n+1)^{(n+1/2)}} = \frac{1}{\sqrt{2}} \frac{n^{(n+2)}}{\left(n+\frac{1}{2}\right)^{(n+1/2)}}$$

$$= \frac{1}{\sqrt{2}} \frac{n^{3/2}}{\left(1+\frac{1}{2n}\right)^{1/2}\left(1+\frac{1}{2n}\right)^n} \geqslant \frac{1}{\sqrt{2}} \frac{n^{3/2}}{\left(1+\frac{1}{2}\right)^{1/2}\left(1+\frac{1}{2n}\right)^n}$$

$$> \frac{n^{3/2}}{\sqrt{3}\left(1+\frac{1}{n}\right)^n} > \frac{n^{3/2}}{\sqrt{3}e} \geqslant \frac{n}{\sqrt{3}e} \qquad \text{for } n = 1, 2, 3, \ldots$$

Hence equation (11–8) cannot be satisfied and therefore we conclude the convergence of $\{f_n\}$ cannot be uniform.

There is, however, at least one case where we conclude, from the fact that a sequence of continuous functions converges to a continuous function, that the convergence is uniform. This important result is known as Dini's theorem.

Theorem 11.7: *If $\{f_n\}$ is a sequence of real-valued continuous functions defined on a compact set K which converges pointwise to a real-valued continuous function f on K and if it is also true that, for every $x \in K, f_n(x) \geqslant f_{n+1}(x)$ for $n=1, 2, 3, \ldots$, then $\{f_n\}$ converges to f uniformly on K.*

Proof: In view of Definition 8.11 we can define a sequence of real functions $\{g_n\}$ on K by

$$g_n = f_n - f \qquad \text{for } n=1, 2, 3, \ldots$$

Then Theorem 8.16(a) shows that $\{g_n\}$ is a sequence of continuous functions and Theorem 7.7 shows that, for every $x \in E$, $\lim_{n\to\infty} g_n(x)=0$; that is, $\{g_n\}$ converges to zero pointwise on E. Now since, for each n, $|g_n(x)-0|=|f_n(x) - f(x)|$, we can establish the theorem by proving that $\{g_n\}$ converges to zero uniformly on K.

It is clear that, for each n and every $x \in K$,

$$g_n(x) \geqslant g_{n+1}(x) \geqslant 0 \tag{11–9}$$

So given $\epsilon > 0$, we can find, for each $x \in K$, a positive integer $n(x)$ such that

$$0 \leqslant g_{n(x)}(x) \leqslant \frac{\epsilon}{2}$$

Since each $g_{n(x)}$ is a continuous function which is always nonnegative, we can find a $\delta(x) > 0$ such that

$$g_{n(x)}(B(x;\, \delta(x))) \subset \left(g_{n(x)}(x) - \frac{\epsilon}{2},\, g_{n(x)}(x) + \frac{\epsilon}{2}\right) \cap [0, \infty)$$

$$\subset \left[0,\, g_{n(x)}(x) + \frac{\epsilon}{2}\right) \subset [0, \epsilon)$$

Thus, for every $y \in B(x;\, \delta(x))$,

$$0 \leqslant g_{n(x)}(y) < \epsilon$$

It follows from equation (11–9) that whenever $n \geqslant n(x)$,

$$g_n(y) \leqslant g_{n(x)}(y) \qquad \text{for all } y \in B(x;\, \delta(x))$$

so that

$$0 \leqslant g_n(y) < \epsilon \tag{11–10}$$

whenever $y \in B(x;\, \delta(x))$ and $n \geqslant n(x)$. Now since K is compact and $\{B(x;\, \delta(x)) \mid x \in K\}$ is an open covering of K, we can find finitely many points of K, say $x_1, \ldots, x_m$, such that

$$K \subset B(x_1;\, \delta(x_1)) \cup \ldots \cup B(x_m;\, \delta(x_m)) \tag{11–11}$$

Set $N = \max\limits_{1 \leqslant i \leqslant m} n(x_i)$. Equations (11–10) and (11–11) now show that

$$0 \leqslant g_n(y) < \epsilon$$

for every $y \in K$ and for every $n \geqslant N$. This implies that $\{g_n\}$ converges to zero uniformly on K.

We might point out that the sequence of functions $\{f_n\}$ defined by equation (11–7) forms a monotonically decreasing sequence at each point of R^1. If, for each n, we let g_n be the restriction of f_n to (0, 1), then $\{g_n\}$ converges to zero at each point of (0, 1) and hence to a continuous function. But the convergence is not uniform since $\operatorname*{lub}\limits_{x \in (0,1)} |g_n(x) - 0| = 1$ and so cannot be made less than any positive number ϵ. The reason this sequence does not satisfy the requirements of Theorem 11.7 is that (0, 1) is not a compact set.

Consider the sequence $\{p_n(a)\}$ defined recursively by equation (9–15). We have shown that this sequence converges to $\sqrt{a}$ for each $a \in [0, 1]$ and that $p_n(0) = 0$ for every n. It is clear from equation (9–15) that $p_{n+1}(a)$ is a polynomial in $\{a\}$ whenever $p_n(a)$ is. Clearly $p_1(a) = \frac{1}{2}a$ is a polynomial in $\{a\}$. Hence we conclude by induction that $p_n(a)$ is a polynomial in $\{a\}$ for every positive integer n.

If $0 \leqslant p_n(a) \leqslant a$, then equation (9–15) shows that

$$\begin{aligned} p_{n+1}(a) = p_n(a) + \tfrac{1}{2}(a - [p_n(a)]^2) &= \tfrac{1}{2}[(1+a) - (1 - p_n(a))^2] \\ &\leqslant \tfrac{1}{2}[(1+a) - (1-a)^2] = \tfrac{1}{2}(1+a)a \leqslant a \end{aligned}$$

and

$$p_{n+1}(a) = \tfrac{1}{2}[(1+a) - (1 - p_n(a))^2] \geqslant \tfrac{1}{2}[1 + a - 1] = \tfrac{1}{2}a \geqslant 0$$

Clearly $p_1(a) = (1/2)a$ lies between 0 and a. It therefore follows by induction that $0 \leqslant p_n(a) \leqslant a$ for every n.

Equation (9–15) now shows that

$$p_{n+1}(a)-p_n(a)=\tfrac{1}{2}(a-[p_n(a)]^2)\geqslant\tfrac{1}{2}(a-a^2)=\tfrac{a}{2}(1-a)\geqslant 0$$

Hence $p_{n+1}(a)\geqslant p_n(a)$ for all $a\in[0, 1]$. Thus we have shown that $\{p_n(a)\}$ is a monotonically increasing sequence[39] of polynomials which converges to $\sqrt{a}$ for each $a\in[0, 1]$. The discussion following 8.17 now shows that for each n the function p_n from $[0, 1]$ into R^1, whose value at a is $p_n(a)$, is continuous. It is not hard to show that the function h which assigns to each $a\in[0, 1]$ the value $\sqrt{a}$ is also continuous. Hence we may conclude after applying Dini's theorem that *the sequence* $\{p_n\}$ *defined recursively by equation (9–15) is a monotonically increasing sequence of polynomials*[40] *on* $[0, 1]$, *which converges uniformly to the function h which assigns to each* $x\in[0, 1]$ the value $\sqrt{x}$. In addition, $p_n(0)=0$ for *every n.*

Now let $\theta:[-1, 1]\to[0, 1]$ be defined by

$$\theta(x)=x^2 \qquad \text{for all } x\in[-1, 1]$$

Then

$$h\circ\theta(x)=h(\theta(x))=|x| \qquad \text{for all } x\in[-1, 1]$$

Let $\epsilon>0$ be given. We have shown that there exists a single integer N such that for every $n\geqslant N$

$$|h(t)-p_n(t)|<\epsilon \qquad \text{for all } t\in[0, 1]$$

Since $\theta([-1, 1])\subset[0, 1]$, it follows that for every $n\geqslant N$

$$|h(\theta(x))-p_n(\theta(x))|<\epsilon \qquad \text{for all } x\in[-1, 1]$$

In view of the fact that $p_n(t)$ is a polynomial in $\{t\}$ for every $t\in[0, 1]$, it is clear that

$$P_n(x)=(p_n\circ\theta)(x)=p_n(\theta(x))=p_n(x^2)$$

is a polynomial in $\{x\}$ for every $x\in[-1, 1]$. Hence the function P_n whose

---

[39] The convergence of the sequence of polynomials $\{p_n(a)\}$ can now also be deduced from Theorem 7.11.

[40] If S is any subset of R^1 we use the terminology "polynomial (defined) on S" to mean a polynomial in $\{1, j_s\}$ where j_s is the natural injection of S into R^1 and 1 is the constant function which assigns the number 1 to every $x\in S$. Thus a function p of the form

$$p(x)=a_0+a_1x+\ .\ .\ .\ +a_nx^n \quad \text{for all } x\in S$$

is a polynomial on S.

value at each point $x \in [1, 1]$ is $P_n(x)$ is a polynomial defined on $[-1, 1]$. Since $p_n(0)=0$, we see that $P_n(0)=0$. Evidently for every $n \geqslant N$

$$||x|-P_n(x)|<\epsilon \qquad \text{for all } x \in [-1, 1]$$

We have therefore proved that *there exists a sequence* $\{P_n\}$ *of polynomials defined on* $[-1, 1]$ *which converges uniformly to the function g which associates with each* $x \in [-1, 1]$ *the value* $|x|$, *and that for every* n, $P_n(0)=0$. We shall use this result subsequently.

Theorem 8.12 shows that the family $\mathscr{F}(E, M)$ of functions which map the set E into the linear space (algebra) M is itself a linear space (algebra) with the pointwise definitions of addition, multiplication, and scalar multiplication given in Definition 8.11. Moreover $\mathscr{B}(E, M)$, $\mathscr{C}(E, M)$, and $\mathscr{C}^{\infty}(E, M)$ are linear subspaces (subalgebras) of $\mathscr{F}(E, M)$. If, in addition, M is a normed linear space (normed algebra), we can define a norm on $\mathscr{B}(E, M)$ by

$$\|f\|=\operatorname*{lub}_{x \in E} \|f(x)\| \qquad \text{for every } f \in \mathscr{B}(E, M) \tag{11-12}$$

Then $\mathscr{B}(E, M)$ is a normed linear space (normed algebra). The proof that equation (11–12) defines a norm on $\mathscr{B}(E, M)$—that is, that it satisfies axioms (N1) to (N3) of Definition 3.4 (and Definition 8.14 if M is also an algebra)—can be carried out in almost exactly the same way as the proof of Theorem 11.3 and so we will not do so here. The norm defined on $\mathscr{B}(E, M)$ by equation (11–12) is called the *supremum norm.* With this definition of norm it is clear that, when a metric is defined on $\mathscr{B}(E, M)$ in terms of this norm in the usual way (i.e., by eq. (6–1)), we arrive at the metric space $\langle \mathscr{B}(E, M), \Delta \rangle$ introduced in Theorem 11.3. Thus all theorems proved about the metric space $\langle \mathscr{B}(E, M), \Delta \rangle$ can be applied to this normed linear space (normed algebra). In particular, the following theorem is an immediate consequence of Theorem 11.4(b).

Theorem 11.8: *Let M be a Banach space and let E be any set. Then* $\mathscr{B}(E, M)$ *is itself a Banach space with supremum norm defined by equation* (11–12). *If, in addition, M is a Banach algebra, then* $\mathscr{B}(E, M)$ *is also a Banach algebra.*

Suppose that M is a normed linear space and E is any set.

In view of Definition 10.1, it is clear that, with each sequence $\{f_n\}$ in $\mathscr{B}(E, M)$, we can associate an infinite series $\sum f_n$ and that we can apply the

355-525 O-70—15

theorems of chapter 10 to this series. We see that, if s_n is the nth partial sum of $\sum f_n$, then, for every $x \in E$,

$$s_n(x) = \sum_{i=1}^{n} f_i(x)$$

So this definition of an infinite series of functions reduces to the familiar one in the case when M is the complex numbers with the usual metric or the Euclidean space R^1. We see also that the sequence of partial sums $\{s_n\}$ is a sequence of functions from E to M and so we say that *the infinite series* $\sum f_n$ *converges pointwise or uniformly on E if* $\{s_n\}$ *converges pointwise or uniformly on E.* It follows from Theorem 11.2 that the infinite series $\sum f_n$ converges in the normed linear space $\mathscr{B}(E, M)$ (i.e., in the supremum norm) if and only if it converges uniformly on E. We may, for example, apply Theorem 10.7(a) to the sequence of functions $\{f_n\}$. Thus, by combining Theorems 11.8 and 10.7(a) with the remarks following Theorem 11.3 and using the fact that, if for some finite number a, $\|f(x)\| \leqslant a$ for all $x \in E$, then $\|f\| = \operatorname{lub}_{x \in E} \|f(x)\| \leqslant a$, we arrive at the following theorem.

Theorem 11.9: *Suppose* $\{f_n\}$ *is a sequence of functions from a set E to a Banach space M and suppose that, for every* $x \in E$,

$$\|f_n(x)\| \leqslant a_n \qquad (n = 1, 2, 3, \ldots)$$

Then $\sum f_n$ *converges uniformly on E if* $\sum a_n$ *converges.*

We shall now use the material developed in this section combined with Theorem 9.19 to obtain a result from the theory of differential equations known as Picard's theorem.

Consider the differential equation

$$\frac{dy}{dx} = f(x, y) \tag{11–13}$$

where f is continuous on the 2-cell

$$Q = \{\langle x, y\rangle \mid a_1 \leqslant x \leqslant b_1 \text{ and } a_2 \leqslant y \leqslant b_2\}$$

and for each fixed $x \in [a_1, b_1]$ satisfies the Lipschitz condition

$$|f(x, y_1)-f(x, y_2)| \leqslant M|y_1-y_2| \qquad \text{for all } y_1, y_2 \in [a_2, b_2] \quad (11\text{–}14)$$

If $I \subset [a_1, b_1]$ is an interval containing x_0, then $g: I \to R^1$ is a solution of the differential equation (11–13) satisfying the initial condition

$$g(x_0)=y_0 \qquad (11\text{–}15)$$

for $\langle x_0, y_0\rangle \in Q$ if and only if g satisfies the integral equation

$$g(x)=y_0+\int_{x_0}^{x} f(t, g(t))\,dt \qquad \text{for all } x \in I \qquad (11\text{–}16)$$

We shall now show that the integral equation (11–16) and hence the differential equation (11–13) with the initial condition (11–15) has a unique solution, when the interval I is sufficiently small.

To this end we note that since f is a continuous function on the compact set Q, corollary 1 to Theorem 8.20 shows that f is bounded on Q. Hence there exists a constant K such that

$$|f(x, y)| \leqslant K \qquad \text{for all } \langle x, y\rangle \in Q \qquad (11\text{–}17)$$

Choose a positive number a such that $Ma < 1$ and such that the 2-cell

$$Q' = \{\langle x, y\rangle | x_0 - a \leqslant x \leqslant x_0 + a,\ y_0 - Ka \leqslant y \leqslant y_0 + Ka\}$$

is a subset of Q. Now set $I=[x_0-a, x_0+a]$ and let $\mathscr{E}$ be the set of all continuous real valued functions on I with values in $[y_0-Ka, y_0+Ka]$. That is, $g: I \to R^1$ belongs to $\mathscr{E}$ if and only if

$$\Delta(g, y_0) = \sup_{x \in I} |g(x) - y_0| \leqslant Ka$$

Thus $\mathscr{E}$ is the closed ball with center at the constant function y_0 in the metric space $\langle \mathscr{C}(I, R^1), \Delta\rangle$. Hence $\mathscr{E}$ is closed and corollary 1(e) of Theorem 11.6 shows that $\langle \mathscr{C}(I, R^1), \Delta\rangle$ is a complete metric space. Therefore Theorem 9.9 shows that $\langle \mathscr{E}, \Delta\rangle$ is a complete metric space. Now consider the mapping $T: \mathscr{E} \to \mathscr{E}$ defined by

$$(T(g))(x) = y_0 + \int_{x_0}^{x} f(t, g(t))\,dt \qquad \text{for all } x \in I \text{ and all } g \in \mathscr{E}$$

To see that T maps $\mathscr{E}$ into itself we notice that it follows from equation

(11–17) that

$$\Delta(T(g), y_0) = \sup_{x \in I} \left| \int_{x_0}^{x} f(t, g(t))\,dt \right| \leq \sup_{x \in I} \int_{x_0}^{x} \left| f(t, g(t)) \right| dt$$

$$\leq \sup_{x \in I} K|x - x_0| \leq Ka$$

for every $g \in \mathscr{E}$. Hence $T(g) \in \mathscr{E}$.

Evidently equation (11–16) can be written as

$$g = T(g)$$

Furthermore, it follows from equation (11–14) that for all $g_1, g_2 \in \mathscr{E}$

$$\Delta(T(g_1), T(g_2)) = \sup_{x \in I} |(T(g_1))(x) - (T(g_2))(x)|$$

$$= \sup_{x \in I} \left| \int_{x_0}^{x} [f(t, g_1(t)) - f(t, g_2(t))]\,dt \right|$$

$$\leq \sup_{x \in I} \int_{x_0}^{x} |f(t, g_1(t)) - f(t, g_2(t))|\,dt$$

$$\leq M \sup_{x \in I} \int_{x_0}^{x} |g_1(t) - g_2(t)|\,dt$$

$$\leq M \sup_{x \in I} |x - x_0| \sup_{t \in I} |g_1(t) - g_2(t)|$$

$$\leq Ma \sup_{t \in I} |g_1(t) - g_2(t)| = Ma\Delta(g_1, g_2)$$

Since $Ma < 1$, this shows that T is a contraction mapping on $\mathscr{E}$ and so Theorem 9.19 shows that there is a unique function $g \in \mathscr{E}$ such that

$$g = T(g)$$

Hence the function g is a unique solution of equation (11–13). Notice that in addition Theorem 9.19 gives us a convergent iterative procedure for solving equation (11–16). The ideas developed in this example have very wide application.

We now introduce a concept which is particularly important for the space of continuous functions.

Definition 11.10: *A subset K of the metric space $\langle X, d\rangle$ is said to be relatively compact if its closure $\bar{K}$ is compact.*

It follows from Theorem 6.22 and the fact that every closed set is equal to its closure that all compact sets are relatively compact. The Heine-Borel theorem (Theorem 6.31) shows us that the compact sets in R^k are simply those sets which are closed and bounded. However, Theorem 9.4 shows that every bounded set has bounded closure. Thus, the subsets of R^k with compact closure and the bounded subsets of R^k are the same sets. For this reason it is generally true that, in arbitrary metric spaces (and, for that matter, topological spaces), the relatively compact sets play the same role as the bounded sets do in R^k.

The next theorem gives us another way of characterizing relatively compact sets.

Theorem 11.11: *A subset K of a metric space $\langle X, d\rangle$ is relatively compact if and only if every sequence in K contains a convergent subsequence.*

Proof: If K is relatively compact, then, by definition, $\bar{K}$ is compact. Hence Theorem 7.20 shows that $\bar{K}$ is sequentially compact. Since every sequence in K is also a sequence in $\bar{K}$, it follows that every sequence in K contains a convergent subsequence.

Conversely, suppose that every sequence in K has a convergent subsequence and let $\{y_n\}$ be any sequence in $\bar{K}$. Since every point of $\bar{K}$ is an adherence point of K, we can find, for each integer n, a point x_n of K such that $x_n \in B(y_n;\ 1/n)$; that is, $d(x_n, y_n) < 1/n$. By hypothesis, this sequence $\{x_n\}$ of points of K contains a convergent subsequence, say $\{x_{n_k}\}$. So let $x = \lim\limits_{k\to\infty} x_{n_k}$. Theorem 7.4 shows that x is an adherence point of K. Thus $x \in \bar{K}$.

Now let $\epsilon > 0$ be given. Evidently for each k, $d(x_{n_k}, y_{n_k}) < 1/n_k$. Since $\{n_k\}$ is an increasing sequence of integers and $\{x_{n_k}\}$ converges to x, we can find an integer N such that $k \geqslant N$ implies $1/n_k < \epsilon/2$ and $d(x, x_{n_k}) < \epsilon/2$. Hence, for $k \geqslant N$,

$$d(y_{n_k}, x) \leqslant d(y_{n_k}, x_{n_k}) + d(x_{n_k}, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Therefore, the subsequence $\{y_{n_k}\}$ of $\{y_n\}$ converges to $x \in \bar{K}$. Since $\{y_n\}$ was an arbitrary sequence in $\bar{K}$, this shows that $\bar{K}$ is sequentially compact and, therefore, by Theorem 7.20, also compact. Thus K is relatively compact.

Although relatively compact subsets of the spaces of continuous functions are particularly useful, it is generally not easy to tell in practice whether a given subset is relatively compact. It is therefore important to develop a useful criterion for relative compactness in these spaces. The most frequently used criterion of this type is given by what is known alternatively as Arzela's theorem, Ascoli's theorem, or the generalized Arzela's or Ascoli's theorem. Before stating this theorem it is necessary to introduce the following concept.

Definition 11.12: *A family $\mathscr{E}(X, Y)$ of functions from a metric space $\langle X, d\rangle$ to a metric space $\langle Y, d'\rangle$ is said to be* **equicontinuous** *if, for every $\epsilon > 0$, there exists a* **single** *$\delta > 0$ such that $d'(f(x_1), f(x_2)) < \epsilon$ for every $f \in \mathscr{E}(X, Y)$ and for all $x_1, x_2 \in X$ for which $d(x_1, x_2) < \delta$.*

It follows from this definition that every member of an equicontinuous family is uniformly continuous. The important additional thing here is that a *single* number δ can be found which will make $d'(f(x_1), f(x_2)) < \epsilon$ for *every* $f \in \mathscr{E}(X, Y)$; that is, one number δ will serve for all $f \in \mathscr{E}(X, Y)$.

If $\mathscr{E}_K$ is the set of all functions from X to Y which satisfy a Lipschitz condition on X with modulus K, then $\mathscr{E}_K$ is an equicontinuous family.

Theorem 11.13: *Let $\langle X, d\rangle$ and $\langle Y, d'\rangle$ be compact metric spaces. Then a subset $\mathscr{K}$ of $\langle \mathscr{C}(X, Y), \Delta\rangle$ is relatively compact if and only if it is equicontinuous.*

Proof: Suppose $\mathscr{K}$ is equicontinuous. Theorem 7.20 and the corollaries to Theorems 7.19 and 6.37 show that every compact set is separable. Hence let E be a countable dense subset of X. The corollary to Theorem 4.17 shows that the points of E can be arranged in a sequence $\{x_n\}$. Let $\{f_i\}$ be any sequence in $\mathscr{K}$. Since $\{f_i(x_1)\}$ is a sequence of points in the compact set Y, Theorem 7.20 shows that $\{f_i\}$ must contain a subsequence, which we will denote by $S_1 = \{f_{i,k}\}$ such that the sequence $\{f_{1,k}(x_1)\}$ of points in Y converges (i.e., $\lim_{k\to\infty} f_{1,k}(x_1)$ exists). Now since $\{f_{1,k}(x_2)\}$ is also a sequence of points in the compact set Y, we see in the same way as before that $\{f_{1,k}\}$ must contain a subsequence, which we denote by $S_2 = \{f_{2,k}\}$, such that the sequence of points $\{f_{2,k}(x_2)\}$ of Y converges. By proceeding inductively in this way, we obtain a countable collection of sequences $S_1, S_2, S_3, \ldots$ which can be symbolically represented by the array

$$
\begin{array}{l}
S_1 = f_{1,1}, f_{1,2}, f_{1,3}, f_{1,4}, \cdots \\
S_2 = f_{2,1}, f_{2,2}, f_{2,3}, f_{2,4}, \cdots \\
S_3 = f_{3,1}, f_{3,2}, f_{3,3}, f_{3,4}, \cdots \\
S_4 = f_{4,1}, f_{4,2}, f_{4,3}, f_{4,4}, \cdots \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdots \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdots \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdots
\end{array}
$$

and which have the following properties:

(a) For every positive integer n, S_{n+1} is a subsequence of S_n.

(b) The $\lim_{k\to\infty} f_{n,k}(x_n)$ exists for every positive integer n.

(c) If one function precedes another in S_1, then these two functions are in the same order in every S_n until one of them is deleted. Property (c) shows that, in moving down from one row to the next in the array, a given function can only move to the left and never to the right.

Now consider the sequence S whose terms are the diagonal elements of this array; that is, the sequence

$$f_{1,1}, f_{2,2}, f_{3,3}, f_{4,4}, \cdots$$

It is easy to see that, for each n, this sequence is, except for possibly its first $n-1$ terms, a subsequence of S_n. Since a finite number of terms cannot affect the convergence or divergence of a sequence and since every subsequence of a convergent sequence converges, we see from property (b) that $\lim_{n\to\infty} f_{n,n}(x_i)$ exists for each $x_i \in E$. Let $\epsilon > 0$ be given. Since S is a sequence in $\mathscr{K}$ and since $\mathscr{K}$ is equicontinuous, there exists a $\delta > 0$ such that, whenever $d(x, x') < \delta$

$$d'(f_{n,n}(x), f_{n,n}(x')) < \frac{\epsilon}{3} \tag{11-18}$$

Since the range E of the sequence $\{x_i\}$ is a dense subset of X, we see that [41]

$$X = \bigcup_{i=1}^{\infty} B(x_i; \delta)$$

[41] If $p \in X$, then there exists an $x_i \in E$ such that $d(p, x_i) < \delta$ and so, for this i, $p \in B(x_i; \delta)$.

But X is compact and $\{B(x_i;\ \delta)|i \in J\}$ is an open cover for X. Therefore, we can find an integer r such that

$$X = \bigcup_{i=1}^{r} B(x_i;\ \delta)$$

Since every convergent sequence is a Cauchy sequence, it is clear that, for $1 \leqslant i \leqslant r$, the sequence

$$f_{1,1}(x_i), f_{2,2}(x_i), f_{3,3}(x_i), f_{4,4}(x_i), \ . \ . \ .$$

is a Cauchy sequence. Hence we can find, for each i, an integer N_i such that

$$d'(f_{n,n}(x_i), f_{m,m}(x_i)) < \frac{\epsilon}{3} \tag{11–19}$$

for all n, $m \geqslant N_i$.

Now let $N = \max\limits_{1 \leqslant i \leqslant r} N_i$. Then equation (11–19) must hold for all n, $m \geqslant N$. If x is any point of X, we can choose one of the points $x_1, x_2, \ . \ . \ ., x_r$, say x_p, such that $x \in B(x_p;\ \delta)$. Since this means that $d(x, x_p) < \delta$, it follows from equations (11–18) and (11–19) that, whenever m, $n \geqslant N$,

$$\begin{aligned} d'(f_{n,n}(x), f_{m,m}(x)) \leqslant d'(f_{n,n}(x), f_{n,n}(x_p)) + d'(f_{n,n}(x_p), f_{m,m}(x_p)) \\ + d'(f_{m,m}(x_p), f_{m,m}(x)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \end{aligned}$$

Hence, for every $x \in X$, $d(f_{n,n}(x), f_{m,m}(x)) < \epsilon$ for all n, $m \geqslant N$ and the same N will do for all x. Hence

$$\Delta(f_{n,n}, f_{m,m}) = \operatorname*{lub}_{x \in X} d'(f_{n,n}(x), f_{m,m}(x)) \leqslant \epsilon$$

for all n, $m \geqslant N$ and $f_{1,1}, f_{2,2}, f_{3,3}, \ . \ . \ .$ is a Cauchy sequence in the metric space $\langle \mathscr{C}(X,\ Y),\ \Delta \rangle$. Since, by corollary 2 to Theorem 9.7 and part (e) of corollary 1 to Theorem 11.6, this is a complete metric space, $f_{1,1}, f_{2,2}, f_{3,3}, \ . \ . \ .$ must converge. In view of the fact that $\{f_i\}$ was any sequence in $\mathscr{K}$, we conclude that every sequence in $\mathscr{K}$ contains a convergent subsequence. It now follows from Theorem 11.11 that $\mathscr{K}$ is relatively compact.

Now suppose that $\overline{\mathscr{K}}$ is compact and choose $\epsilon > 0$. Theorem 7.20 and the corollary to Theorem 7.19 show that $\overline{\mathscr{K}}$ is totally bounded. Hence let $\mathscr{N}_\epsilon = \{f_1, f_2, \ . \ . \ ., f_n\}$ be an $(\epsilon/3)$-net for $\overline{\mathscr{K}}$. Since X is compact, Theorem 8.23

shows that, for each k $(1 \leqslant k \leqslant n)$, f_k is uniformly continuous. We can therefore find, for each k, a $\delta_k > 0$ such that

$$d'(f_k(x), f_k(x')) < \frac{\epsilon}{3}$$

whenever $d(x, x') < \delta_k$. Set $\delta = \min_{1 \leqslant k \leqslant n} \delta_k$. Thus $\delta > 0$. Now if f is any function which belongs to $\mathscr{K} \subset \overline{\mathscr{K}}$, we can find an $f_k \in \mathscr{N}_\epsilon$ such that

$$\Delta(f, f_k) < \frac{\epsilon}{3}$$

We see from these two inequalities that, whenever $d(x, x') < \delta$,

$$\begin{aligned} d'(f(x), f(x')) &\leqslant d'(f(x), f_k(x)) + d'(f_k(x), f_k(x')) + d'(f_k(x'), f(x')) \\ &\leqslant \Delta(f, f_k) + \frac{\epsilon}{3} + \Delta(f, f_k) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \end{aligned}$$

Since this must be true for every $f \in \mathscr{K}$ and the same δ will do for every such f, we conclude that $\mathscr{K}$ is equicontinuous.

Corollary: *Let $\langle X, d\rangle$ be a compact metric space. Then a closed subset $\mathscr{K}$ of $\langle \mathscr{C}(X), \Delta\rangle$ is compact if and only if it is equicontinuous and bounded.*
Proof: If $\mathscr{K}$ is bounded, then there is a finite number M such that

$$\operatorname*{lub}_{f,g \in \mathscr{K}} \operatorname*{lub}_{x \in X} |f(x) - g(x)| = \Delta(\mathscr{K}) < M$$

Hence the ranges of the functions belonging to $\mathscr{K}$ all lie in some fixed bounded subset Y of the complex plane (or the real line) and we can choose Y to be closed (since $Y \subset \bar{Y}$ and $d(Y) = d(\bar{Y})$). The Heine-Borel theorem shows that Y is compact. Clearly $\mathscr{K} \subset \mathscr{C}(X, Y) \subset \mathscr{C}(X)$.

Now suppose $\mathscr{K}$ is a compact subset of $\mathscr{C}(X)$. Then Theorem 6.22 shows that $\mathscr{K}$ is bounded and so we can find a compact set Y of complex (or real) numbers such that $\mathscr{K} \subset \mathscr{C}(X, Y)$. It follows from Theorem 6.19 that $\mathscr{K}$ is a compact subset of $\langle \mathscr{C}(X, Y), \Delta\rangle$ and, since compact sets are also relatively compact, Theorem 11.13 shows that $\mathscr{K}$ is equicontinuous.

Conversely, if $\mathscr{K}$ is equicontinuous and bounded, then we can find a compact set of complex (or real) numbers Y such that $\mathscr{K} \subset \mathscr{C}(X, Y)$. Hence Theorem 11.13 shows that $\mathscr{K}$ is a relatively compact set of the metric space $\langle \mathscr{C}(X, Y), \Delta\rangle$. Thus, by Theorem 11.11, every infinite sequence in $\mathscr{K}$ has a limit point in $\mathscr{C}(X, Y)$. But since $\mathscr{C}(X, Y) \subset \mathscr{C}(X)$, we see that every infinite sequence in $\mathscr{K}$ has a limit point in $\mathscr{C}(X)$ and, using Theorem 11.11 again, we see that $\mathscr{K}$ is a relatively compact subset of the metric space $\langle \mathscr{C}(X), \underline{\Delta} \rangle$. Since, by hypothesis, $\mathscr{K}$ is a closed subset of this metric space (i.e., $\mathscr{K} = \overline{\mathscr{K}}$), this shows that $\mathscr{K}$ is compact.

We are now going to establish a result (Theorem 11.19) which shows that every family of real valued functions (on a compact metric space X) which is closed under certain operations and which is sufficiently "rich" can be used to approximate uniformly every continuous real valued function on X. This very famous theorem is known as the Stone-Weierstrass theorem. The original form of the theorem, discovered by Weierstrass, concerned the real valued continuous functions defined on an interval $[a, b]$. Now since any polynomial p (with real coefficients) defined on $[a, b]$,[40] say

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n \qquad x \in [a, b]$$

is a continuous function, we know from Theorem 11.6 that the limit of any uniformly convergent sequence of such polynomials is also a continuous real function. Weierstrass showed that the converse of this result is also true; that is, every continuous real valued function on $[a, b]$ is the limit of a uniformly convergent sequence of polynomials.

Stone generalized this result by first replacing the interval $[a, b]$ by a compact topological space X (we shall limit ourselves here to replacing $[a, b]$ by a compact metric space) and then by finding a subfamily of the continuous real valued functions on X which has all those properties of the polynomials that made the Weierstrass theorem possible.

Now let 1 be the constant function which assigns the number 1 to every x in $[a, b]$ and let j be the natural injection of $[a, b]$ into R^1; that is,

$$j(x) = x \qquad \text{for every } x \in [a, b]$$

The family $\mathscr{P}$ of all polynomials on $[a, b]$ [40] is just the set of all functions which can be built up from these two functions by successively applying the operations of addition, multiplication, and multiplication by real numbers (scalars). Thus if $a_0, \ldots, a_n$ are real numbers, the members of $\mathscr{P}$ are func-

[40] See page 218.

tions of the form

$$a_0 \cdot 1 + a_1 j + a_2 j^2 + \ . \ . \ . \ + a_n j^n$$

As pointed out in chapter 8 $\mathscr{P}$ is a subalgebra of the continuous real valued functions on $[a, b]$. Aside from the fact that $\mathscr{P}$ is an algebra there are two additional properties of $\mathscr{P}$ that make the Weierstrass theorem possible. These properties are given in Definition 11.18. Theorem 7.4 and the remarks following Theorem 11.2 show that the Weierstrass theorem is equivalent to the assertion that $\overline{\mathscr{P}}$, the closure of $\mathscr{P}$ in the metric space $\langle \mathscr{C}([a, b]), \Delta \rangle$, be equal to $\mathscr{C}([a, b])$. Before proving the Stone-Weierstrass theorem, it is necessary to obtain some preliminary results.

Definition 11.14: *Let I be a subset of the positive integers and let $\Omega = \{f_i | i \in I\}$ be a family of real-valued functions defined on a set X. The* **upper envelope** *of the family Ω, which is denoted by $\sup\limits_{i \in I} f_i$, is the function h defined by*

$$h(x) = \sup_{i \in I} f_i(x) \qquad \text{for all } x \in X$$

The **lower envelope** *is defined similarly and is denoted by $\inf\limits_{i \in I} f_i$. If Ω contains only two functions, say f and g, then the upper and lower envelopes of Ω are, respectively, denoted by* $\sup(f, g)$ *and* $\inf(f, g)$.

If f_1, f_2, f_3 are real valued functions defined on a set X and $h = \sup(f_1, f_2)$, then $\sup\limits_{i \in \{1, 2, 3\}} f_i = \sup(h, f_3)$.

Definition 11.15: *A subset $\mathscr{A}$ of $\mathscr{C}(X, R^1)$ is said to be a* **lattice subset** *if, for every $f, g \in \mathscr{A}$, the upper and lower envelopes,* $\sup(f, g)$ *and* $\inf(f, g)$, *also belong to $\mathscr{A}$.*

It is easy to see from the remark following the preceding definition that if $\mathscr{A}$ is a lattice subset and $\{f_i | i \in I\}$ is any *finite* subcollection of the elements of $\mathscr{A}$, then $\sup\limits_{i \in I} f_i$ also belongs to $\mathscr{A}$. Of course, a similar conclusion holds for $\inf\limits_{i \in I} f_i$.

If f and g are real valued functions on a set X, we shall write $f > g$ if $f(x) > g(x)$ for every $x \in X$. And we define the function $|f|$ by

$$|f|(x) = |f(x)| \qquad \text{for every } x \in X$$

Theorem 11.16: *Let K be compact and let $\mathscr{A}$ be a lattice subset of $\langle \mathscr{C}(K, R^1), \Delta \rangle$. If $f \in \mathscr{C}(K, R^1)$ and if, for any $\epsilon > 0$ and any $x, y \in K$, there exists a function $g_{x,y} \in \mathscr{A}$ such that*

$$|f(x) - g_{x,y}(x)| < \frac{\epsilon}{2}$$

and $\qquad\qquad$ (11–20)

$$|f(y) - g_{x,y}(y)| < \frac{\epsilon}{2}$$

then $f \in \overline{\mathscr{A}}$.

Proof: Let $\epsilon > 0$ be given. For each $x, y \in K$ there exists by hypothesis a function $g_{x,y}$ which satisfies the inequalities (11–20). Hence let us set

$$V_{x,y} = \left\{ p \in K \,\middle|\, g_{x,y}(p) - f(p) < \frac{\epsilon}{2} \right\}$$

Evidently,

$$V_{x,y} = (g_{x,y} - f)^{-1}\left(\left(-\infty, \frac{\epsilon}{2}\right)\right)$$

Since the function $g_{x,y} - f$ is continuous and $(-\infty, \epsilon/2)$ is an open subset of R^1, Theorem 8.7 shows that $V_{x,y}$ is open. In addition, the second inequality (11–20) shows that $y \in V_{x,y}$. Hence the family $\Omega_x = \{V_{x,y} | y \in K\}$ is an open cover of K. Since K is compact, there exists finitely many points of K, say $y_1, \ldots y_n$, such that $K \subset \sum_{i=1}^{n} V_{x,y_i}$. Set

$$g_x = \inf_{i \in \{1, \ldots, n\}} g_{x,y_i}$$

Then $g_x(p) < f(p) + \epsilon/2$ for every $p \in K$. Since $\mathscr{A}$ is a lattice subset, $g_x \in \mathscr{A}$. Also, since the first inequality (11–20) shows that $g_{x,y_i} > f(x) - \epsilon/2$, it follows that $g(x) > f(x) - \epsilon/2$.

Now set

$$V_x = \left\{ p \in K \,\middle|\, g_x(p) - f(p) > -\frac{\epsilon}{2} \right\}$$

Since $g_x - f$ is continuous, we see as before from Theorem 8.7 that V_x is open and that $x \in V_x$. Hence $\Omega = \{V_x | x \in K\}$ is an open cover of K. Since K is compact there exists finitely many points of K, say $x_1, \ldots, x_k$, such that $\{V_{x_i} | 1 \leq i \leq k\}$ is an open cover of K. Set

$$g = \sup_{i \in \{1, \ldots, k\}} g_{x_i}$$

Then,

$$g(p) > f(p) - \frac{\epsilon}{2} \qquad \text{for all } p \in K \tag{11–21}$$

Since $\mathscr{A}$ is a lattice subset, $g \in \mathscr{A}$. Also, since, for every $i = 1, 2, \ldots, k$ and every $p \in K$, $g_{x_i}(p) < f(p) + \epsilon/2$, it follows that

$$g(p) < f(p) + \frac{\epsilon}{2} \qquad \text{for every } p \in K$$

When this equation is combined with equation (11–21) we find,

$$|g(p) - f(p)| < \frac{\epsilon}{2} \qquad \text{for all } p \in K$$

Hence

$$\Delta(g, f) = \sup_{p \in X} |g(p) - f(p)| \leq \frac{\epsilon}{2} < \epsilon$$

Thus we have found a point $g \in \mathscr{A}$ such that $g \in B(f; \epsilon)$. Since ϵ was arbitrary, this shows that f is an adherence point of $\mathscr{A}$.

Theorem 11.17: *Let K be compact. Then every closed subalgebra $\mathscr{A}$ of $\mathscr{C}(K, R^1)$ is a lattice subset of $\mathscr{C}(K, R^1)$.*

Proof: Suppose $f \in \mathscr{A}$. Since K is compact, f is bounded (by corollary 1 of Theorem 8.20) and

$$b = \sup_{x \in K} |f(x)|$$

is a finite number. Set

$$g = \begin{cases} f/b; & f \neq 0 \\ 0; & f = 0 \end{cases}$$

Since $\mathscr{A}$ is an algebra, $g \in \mathscr{A}$. Now fix $\epsilon > 0$. We have shown in the example following Dini's theorem (Theorem 11.7) that there exists a sequence of polynomials $\{P_n\}$ on $[-1, 1]$ which converges uniformly to the function which associates with every $x \in [-1, 1]$ its absolute value $|x|$, and which has the property that for every n, $P_n(0) = 0$. We can therefore find an integer n such that

$$\big||x| - P_n(x)\big| < \frac{\epsilon}{2} \qquad \text{for all } x \in [-1, 1]$$

Since P_n is a polynomial and $P_n(0) = 0$, there must be real numbers, say $\alpha_1, \ldots, \alpha_k$, such that

$$P_n(x) = \alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_k x^k \qquad \text{for all } x \in [-1, 1]$$

It is clear that for every $y \in K$, $g(y) \in [-1, 1]$. Hence for every $y \in K$

$$\left| |g(y)| - (\alpha_1 g^1(y)) + \ldots + \alpha_k g^k(y) \right| < \frac{\epsilon}{2}$$

Since $\mathscr{A}$ is an algebra which contains g, it follows that the function

$$P_n(g) = \alpha_1 g^1 + \ldots + \alpha_k g^k$$

belongs to $\mathscr{A}$. Now

$$\Delta(|g|, P_n(g)) = \sup_{y \in K} \left| |g(y)| - P_n(g(y)) \right|$$

$$= \sup_{y \in K} \left| |g(y)| - (\alpha_1 g^1(y) + \ldots + \alpha_k g^k(y)) \right| \leq \frac{\epsilon}{2} < \epsilon$$

Since ϵ was arbitrary we have shown that every ball $B(|g|; \epsilon)$ about $|g|$ contains a point of $\mathscr{A}$. Thus $|g|$ is an adherence point of $\mathscr{A}$. Using the fact that $\mathscr{A}$ is closed, we see that $|g| \in \mathscr{A}$. Now $f = bg$ implies $|f| = b|g|$. Since $\mathscr{A}$ is an algebra we conclude that $|f| \in \mathscr{A}$. Hence we have shown that $|f|$ belongs to $\mathscr{A}$ whenever f does.

Now suppose f_1 and f_2 are any two members of $\mathscr{A}$. Clearly

$$\sup(f_1, f_2) = \tfrac{1}{2}[(f_1 + f_2) + |f_1 - f_2|]$$

and

$$\inf(f_1, f_2) = \tfrac{1}{2}[(f_1 + f_2) - |f_1 - f_2|]$$

Since $\mathscr{A}$ is an algebra, $f_1 - f_2 \in \mathscr{A}$. Hence the preceding result shows $|f_1 - f_2| \in \mathscr{A}$. It therefore follows from the fact that $\mathscr{A}$ is an algebra that sup (f_1, f_2) and inf (f_1, f_2) also belong to $\mathscr{A}$, and this shows $\mathscr{A}$ is a lattice subset.

Definition 11.18: *Let $\mathscr{A}$ be a family of functions from a set X into a set Y. Then $\mathscr{A}$ is said to separate points of X if, for each pair of distinct points $x, y \in X$, there exists an $f \in \mathscr{A}$ such that $f(x) \neq f(y)$.*

If for each $x \in X$ there exists a $g \in \mathscr{A}$ such that $g(x) \neq 0$, then $\mathscr{A}$ is said to vanish at no point of X.

Theorem 11.19: *Let K be a compact metric space. If $\mathscr{A}$ is a subalgebra of $\langle \mathscr{C}(K, R^1), \Delta \rangle$ such that $\mathscr{A}$ separates points of K and $\mathscr{A}$ vanishes at no point of K, then*

$$\overline{\mathscr{A}} = \mathscr{C}(K, R^1)$$

Proof: Theorems 8.15 and 6.13 show that $\overline{\mathscr{A}}$ is a closed subalgebra of $\mathscr{C}(K, R^1)$. Hence Theorem 11.17 shows that $\overline{\mathscr{A}}$ is a lattice subset. Since the closure of $\overline{\mathscr{A}}$ is $\overline{\mathscr{A}}$, it follows from Theorem 11.16 that it is sufficient to prove that for every $f \in \mathscr{C}(K, R^1)$ and for each pair of distinct points $x, y \in K$, there exists a function $g_{x,y} \in \overline{\mathscr{A}}$ such that

$$g_{x,y}(x) = f(x)$$

and $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (11–22)

$$g_{x,y}(y) = f(y)$$

We shall show that there exists a function $g_{x,y} \in \mathscr{A} \subset \overline{\mathscr{A}}$ with this property. By hypothesis there exist functions u and h in $\mathscr{A}$ such that $u(x) \neq u(y)$ and $h(x) \neq 0$. Put

$$v = u + \beta h$$

where β is a real number which we choose as follows:

If $u(x) \neq 0$, set $\beta = 0$. If $u(x) = 0$, then $u(y) \neq 0$ and so there is a $\beta \neq 0$ such that

$$\beta[h(x) - h(y)] \neq u(y)$$

Since $\mathscr{A}$ is an algebra, $v \in \mathscr{A}$ and our choice of β shows that $v(x) \neq v(y)$ and $v(x) \neq 0$. Now set

$$\gamma = v^2(x) - v(x)v(y)$$

Clearly $\gamma \neq 0$ and if we set

$$g_x = \gamma^{-1}[v^2 - v(y)v]$$

it follows that $g_x \in \mathscr{A}$, $g_x(x) = 1$, and $g_x(y) = 0$.

In a similar way we can construct a function $g_y \in \mathscr{A}$ such that

$$g_y(y) = 1 \text{ and } g_y(x) = 0$$

Put

$$g_{x,y} = f(x)g_x + f(y)g_y$$

Then $g_{x,y} \in \mathscr{A}$ and satisfies equation (11–22).

Theorem 11.19 does not hold for algebras of complex valued functions. However the conclusion does hold for a subalgebra $\mathscr{A}$ of $\mathscr{C}(X, \boldsymbol{C})$ if we impose an extra condition on $\mathscr{A}$—namely, that $\mathscr{A}$ be self-adjoint. This means that if $f \in \mathscr{A}$ then its complex conjugate $\bar{f}$ also belongs to $\mathscr{A}$. The complex conjugate of a function f on a set K is defined to be the function $\bar{f}$ such that

$$\bar{f}(x) = \overline{f(x)} \qquad \text{for all } x \in K$$

Corollary: Let K be compact and $\mathscr{A}$ be a complex self-adjoint subalgebra of $\mathscr{C}(K, \boldsymbol{C})$ which separates points of K and vanishes at no point of K. Then

$$\overline{\mathscr{A}} = \mathscr{C}(K, \boldsymbol{C})$$

Proof: Let $\mathscr{A}_R$ be the set of all real valued continuous functions on K which belong to $\mathscr{A}$. Since the sum and product of two real functions are real functions and since the product of a real function and a real number is a real function, it is easy to see that $\mathscr{A}_R$ is a subalgebra of $\mathscr{A}$ over the real numbers. If $f \in \mathscr{A}$ then there exist real functions u and v such that $f = u + iv$ and $2u = f + \bar{f}$. Since $\mathscr{A}$ is a self-adjoint algebra we see that $u \in \mathscr{A}_R$.

If x_1 and x_2 are distinct points of K, there exists an $f \in \mathscr{A}$ such that $f(x_1) = 1$ and $f(x_2) = 0$. Hence $0 = u(x_2) \neq u(x_1) = 1$, which shows that $\mathscr{A}_R$ separates points of K. If $x \in K$ then there exists a $g \in \mathscr{A}$ such that $g(x) \neq 0$. Therefore we can find a complex number λ such that $\lambda g(x) > 0$. Now set

$f=\lambda g$. Then $f\in\mathscr{A}$ and $u=(1/2)(f+\bar{f})$ belongs to $\mathscr{A}_R$. Clearly $u(x)>0$. Hence $\mathscr{A}_R$ vanishes at no point of K.

Thus $\mathscr{A}_R$ satisfies the hypothesis of Theorem 11.19, which implies

$$\overline{\mathscr{A}}_R=\mathscr{C}(K, R^1)$$

Hence every real valued continuous function on K belongs to $\overline{\mathscr{A}}_R$. Clearly $\overline{\mathscr{A}}_R\subset\overline{\mathscr{A}}$. Therefore every real valued continuous function on K belongs to $\overline{\mathscr{A}}$. Now if $f\in\mathscr{A}(K, C)$ then there are functions $u, v\in\mathscr{C}(K, R^1)$ such that

$$f=u+iv$$

Hence $u\in\overline{\mathscr{A}}$ and $v\in\overline{\mathscr{A}}$. Since the closure of an algebra is also an algebra (Theorem 8.15) we see that $f\in\overline{\mathscr{A}}$; that is,

$$\mathscr{C}(K, C)=\overline{\mathscr{A}}$$

Theorem 11.19 and its corollary are among the most important facts in modern analysis. For applications it is convenient to restate this theorem in the following way:

If $\{f_\alpha|\alpha\in A\}=\Omega$ *is a family of elements of* $\mathscr{C}(K, R^1)$ *which separates points of the compact set* K *and vanishes at no point of* K, *then given any real valued continuous functions* f *on* K *there is a sequence* $\{g_n\}$ *of polynomials in* Ω *which converges uniformly to* f.

If we associate the set of all polynomials in Ω with the subalgebra of $\mathscr{C}(K, R^1)$ generated by Ω then we see from Theorem 7.4 and the remarks following Theorem 11.3 that this result follows directly from Theorem 11.19.

Let K be a compact subset of R^k and let $\mathscr{A}$ be the algebra whose points are the restrictions to K of the polynomials in the set $\mathscr{E}_k$ defined by equation (8–11). Since all the coordinates of two distinct points of K cannot be the same, $\mathscr{A}$ separates the point of K. Also since $\mathscr{A}$ contains the constant functions, it vanishes at no point of K. Hence, any real valued continuous function defined on a compact subset $K\subset R^k$ is the uniform limit of a sequence of these polynomials.

Let K be the unit circle in the complex plane parametrized by the angle $\pi\theta$, so that the continuous functions on K can be identified with the continuous functions on R^1 having period 2; thus,

$$K=\{e^{i\pi\theta}|0<\theta\leq 2\}$$

Let $\mathscr{A}$ be the complex algebra generated by the constant functions and the functions whose values are $e^{i\pi\theta}$ and $e^{-i\pi\theta}$. Then the elements of $\mathscr{A}$ are the trigonometric polynomials whose values are of the form

$$\sum_{n=-N}^{N} C_n e^{\pi i n\theta} \qquad 0 < \theta \leqslant 2$$

It is easy to see that $\mathscr{A}$ is a self-adjoint algebra which separates points of K and vanishes at no point of K. Hence any continuous complex-valued function defined on R^1 which is periodic with a period of 2 is the uniform limit of a sequence of trigonometric polynomials.

COMMONLY USED SYMBOLS AND SPECIAL NOTATIONS

B or $B(p;\epsilon)$	ball about the point p with radius ϵ
$\mathscr{B}(X, Y)$	set of all bounded functions from X to Y
$\boldsymbol{C}$	set of all complex numbers
$\mathscr{C}(X, Y)$	set of all continuous functions from X to Y
$\mathscr{C}^{\infty}(X, Y)$	set of all bounded continuous functions from X to Y
$\mathscr{C}_0(X, Y)$	set of all continuous functions from X to Y which vanish at infinity
d or $d(p, q)$	metric (also diameter and distance between sets)
$d_\times$	metric on product space
$\mathscr{F}(X, Y)$	family of all functions from X to Y
glb	greatest lower bound
$\mathscr{Im}$	imaginary part (of complex number)
i	identity map
inf	greatest lower bound (infimum)
J	set of positive integers
j_A	natural injection of set A
lim, $\lim_{n\to\infty}$	limit
$\liminf_{n\to\infty}$	inferior limit
$\limsup_{n\to\infty}$	superior limit
lub	least upper bound
max	least upper bound (of finite set)
min	greatest lower bound (of finite set)
R^k	k-dimensional Euclidean space
R^1	real numbers
$\mathscr{Re}$	real part (of complex number)
sup	least upper bound (supremum)
ϵ	small positive number
$\emptyset$	empty set
$\subset\ \supset$	set inclusions
$\in$	membership in a set

$\bigcap$, $\cap$	intersection
$\bigcup$, $\cup$	union
$\overline{}$	(overbar) closure of a set
c	(superscript) complement of a set
$\langle\ ,\ \rangle, \langle\cdot,\cdot,\ldots,\cdot\rangle$	ordered pair, ordered n-tuple
$\lvert\cdot\rvert$	norm in R^k, absolute value in R^1 and R^2
$\lVert\cdot\rVert$	norm in general linear space
$'$	(prime) derived set
$\{\ \ \}$	set notation, or sequence
$\times$, $\times$	direct product
$\exists$	"exists," "for at least one"
o	(superscript) interior of a set
$\circ$	composition (written as $f \circ g$)
Δ, $\Delta(f, g)$	metric in function space
$+\infty, -\infty$	upper and lower bounds for the real numbers
$\sim$	equivalence
$\rightarrow$	convergence, improper convergence, function from one set to another (written as $f : X \rightarrow Y$)
$(\ ,\ )$	segment
$[\ ,\ ]$	interval (closed)
$(\ ,\ ], [\ ,\ )$	intervals (half open)

REFERENCES

1. HALMOS, PAUL R.: Naive Set Theory. D. Van Nostrand Co., Inc., 1960.
2. PERVIN, WILLIAM J.: Foundations of General Topology. Academic Press, 1964.
3. CHOQUET, GUSTAV: Topology. Academic Press, 1966.
4. SIMMONS, GEORGE F.: Introduction to Topology and Modern Analysis. McGraw-Hill Book Co., Inc., 1963.
5. HEWITT, EDWIN: The Rôle of Compactness in Analysis. Am. Math. Monthly, vol. 67, 1960, pp. 499–516.

BIBLIOGRAPHY

BACHMAN, GEORGE; and NARICI, L.: Functional Analysis. Academic Press, 1966.

DIEUDONNÉ, JEAN: Foundations of Modern Analysis. Academic Press, 1960.

GLEASON, ANDREW M.: Fundamentals of Abstract Analysis. Addison-Wesley Publ. Co., 1966.

KOLMOGOROV, A.; and FOMIN, S.: Metric and Normed Spaces. Vol. 1 of Elements of the Theory of Functions and Functional Analysis. Graylock Press, 1957.

ROYDEN, H. L.: Real Analysis. Macmillan Co., 1963.

RUDIN, WALTER: Principles of Mathematical Analysis. Second ed., McGraw-Hill Book Co., Inc., 1964.

INDEX

U.S. GOVERNMENT PRINTING OFFICE : 1970 OL—355-525

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

— NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Notes, and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546