CHAPTER I

PRELIMINARIES

In this chapter we shall give some definitions and theorems which will be needed in the sequel.

<u>Fundamental Properties of Sequences</u>

In this section, we restrict ourselves to the sequences of real numbers.

The set of real numbers will be denoted by $\mathbb{R}$.

1.1 <u>Definition</u>. A <u>sequence</u> in a set X is a function from $\mathbb{Z}^+$, the positive integers, into X.

Instead of using the functional notation $f(n)$, $n \in \mathbb{Z}^+$, for a sequence, we will write $\{x_n\}$, where $x_n = f(n)$.

1.2 <u>Definition</u>. A sequence $\{x_n\}$ in $\mathbb{R}$ is said to be

(a) monotonically nondecreasing if $x_n \leq x_{n+1}$ for all n;

(b) monotonically nonincreasing if $x_n \geq x_{n+1}$ for all n.

The class of monotonic sequences consists of the monotonically nondecreasing and the monotonically nonincreasing sequences.

1.3 Definition. A sequence $\{x_n\}$ in $\mathbb{R}$ is said to converge to a point x if for any $\varepsilon > 0$ there is a positive integer N such that $|x_n - x| \leq \varepsilon$ whenever $n \geq N$. In this case we write

$$\lim_{n \to \infty} x_n = x \quad \text{or} \quad x_n \longrightarrow x.$$

If $\{x_n\}$ does not converge, it is said to diverge.

The set of numbers x_n $(n = 1,2,3,\ldots)$ is called the range of $\{x_n\}$. The range may be a finite set, or it may be infinite.

The sequence $\{x_n\}$ is said to be bounded if its range is bounded.

The sequence $\{x_n\}$ is said to be bounded above (below) if $x_n \leq M$ $(x_n \geq m)$ for some M (m) and all n.

A basic criterion for deciding when some sequences converge is as follows.

1.4 Theorem. Every bounded monotone sequence converges.

Proof. It suffices to consider the case where $\{x_n\}$ is nonincreasing and bounded. Let x be the infimum of the range of the sequence. Then for any $\varepsilon > 0$, there is an N such that $x \leq x_N \leq x + \varepsilon$. The monotonicity of $\{x_n\}$ now yields $x \leq x_n \leq x_N \leq x + \varepsilon$ for $n \geq N$, and the convergence is proved.

Q.E.D.

The following theorem is needed later.

1.5 Theorem. Let $\{x_n\}$ be a nonincreasing sequence of real numbers and bounded below by 0. If $\{x_n\}$ does not converge to 0, then there exists an $\varepsilon > 0$ such that $|x_n| \geq \varepsilon$ for all n.

Proof. Since $\{x_n\}$ does not converge to 0, hence there is an $\varepsilon > 0$ such that

(*) $|x_n| \geq \varepsilon$ for infinitely many n.

Now if there is an n_0 such that $|x_{n_0}| < \varepsilon$, then, by the nonincreasing property of $\{x_n\}$,

$$|x_n| < \varepsilon \quad \text{for all } n \geq n_0.$$

This contradicts to (*),hence the result.

Q.E.D.

We will also need the notion of a subsequence of a sequence.

1.6 Definition. Let n(k) be an increasing function from $\mathbb{Z}^+$ into $\mathbb{Z}^+$ (i.e., $n(k+1) > n(k)$ for all k), and let f(n) define the sequence $\{x_n\}$. The function $f[n(k)]$ then defines a sequence which is a subsequence of $\{x_n\}$; it is designated $\{x_{n_k}\}$.

1.7 Definition. A point x is called a cluster point (or accumulation point) of the sequence $\{x_n\}$ if there is a subsequence of $\{x_n\}$ that converges to x.

If a sequence has a limit x, then x is also a cluster point, the converse is not usually true.

1.8 *Example*. The sequence in $\mathbb{R}$ defined by

$$x_n = 0 + 1/n \; , \; n = 1,3,5,7,\ldots$$
$$\quad = 1 - 1/n \; , \; n = 2,4,6,8,\ldots$$

has 0 and 1 as cluster points but has no limit.

1.9 *Definition*. Let E be a subset of $\mathbb{R}$. A *neighborhood* of a point $x_0 \in E$ is a set

$$N(x_0, \varepsilon) = \{ x \in E \mid |x - x_0| < \varepsilon \}.$$

We note that every neighborhood is an open set.

1.10 *Definition*. Let E be a subset of $\mathbb{R}$. A point x_0 is a *limit point* of E if every neighborhood of x_0 contains a point $x \neq x_0$ such that $x \in E$.

1.11 *Theorem*. If x_0 is a limit point of a set E, then every neighborhood of x_0 contains infinitely many points of E.

Proof. Suppose there is a neighborhood N of x_0 which contains only a finite number of points of E. Let $x_1, x_2, \ldots, x_n$ be those points of $N \cap E$, which are distinct from x_0, and put

$$r = \min \{ |x_1 - x_0| , \ldots, |x_n - x_0| \} .$$

Clearly $r > 0$. The neighborhood $N(x_0, r)$ contains no point x of E such that $x \neq x_0$, so that x_0 is not a limit point of E. This contradiction establishes the theorem.

Q.E.D.

We state without prove the following well-known theorem.

1.12 Theorem. (Bolzano-Weierstrass.) Every bounded infinite subset E of $\mathbb{R}$ has a limit point.

For the proof of this theorem see e.g. [10] .

1.13 Theorem. Every bounded sequence $\{x_n\}$ contains a convergent subsequence.

Proof. Let E be the range of $\{x_n\}$. Then E is bounded.

If E is finite, then there is at least one point in E, say x_0, and a sequence $\{n_k\}$, where $n_1 < n_2 < \dots$, such that $x_{n_1} = x_{n_2} = \dots = x_0$. The subsequence $\{x_{n_k}\}$ obtained in this manner clearly converges.

If E is infinite, then E has a limit point (Theorem 1.12). Let x_0 be the limit point of E. Consider the neighborhoods $N(x_0, (\frac{1}{2})^k)$ $k = 1, 2, 3, \dots$, of x_0. Each such neighborhood contains infinitely many points of E (Theorem 1.11). Thus $N(x_0, \frac{1}{2})$ contains some x_{n_1}, $N(x_0, (\frac{1}{2})^2)$ contains some x_{n_2}, $n_2 > n_1$, $N(x_0, (\frac{1}{2})^3)$ contains some x_{n_3}, $n_3 > n_2$, etc. The resulting subsequence $\{x_{n_k}\}$ clearly has limit x_0.

Q.E.D.

It follows from Theorem 1.13 that if $\{x_n\}$ is a bounded sequence then it has at least one cluster point.

1.14 <u>Theorem</u>. Let E be a closed and bounded subset of $\mathbb{R}$. Then any sequence $\{x_n\} \subset E$ which has only one cluster point x_0 converges to x_0.

<u>Proof</u>. Suppose x_0 is not the limit of $\{x_n\}$; then there would exist a number $\varepsilon > 0$ such that there would be a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ whose points belong to $E \cap (N(x_0, \varepsilon))^c$. ($N^c$ denote the complement of N.) By Theorem 1.13, the subsequence has a cluster point y_0, and as $E \cap (N(x_0, \varepsilon))^c$ is closed, $y_0 \in E \cap (N(x_0, \varepsilon))^c$. The sequence $\{x_n\}$ would thus have two distinct cluster points, contrary to our assumption. This proves the theorem.

Q.E.D.

1.15 <u>Definition</u>. A subset E of $\mathbb{R}$ is <u>compact</u> if every infinite subset of E has a limit point in E.

We note that every closed and bounded subset of $\mathbb{R}$ is compact.

1.16 <u>Definition</u>. A sequence $\{x_n\}$ is a <u>Cauchy</u> <u>sequence</u> if, for every $\varepsilon > 0$, there exists an N such that $|x_m - x_n| \leq \varepsilon$ whenever $m,n \geq N$.

We close this section with a few more remarks on sequences.

Let $\{x_n\}$ be a sequence of real numbers. Consider successively the sequences

$$\left.\begin{array}{r} x_1, x_2, x_3, \ldots, x_n, x_{n+1}, \ldots, \\ x_2, x_3, \ldots, x_n, x_{n+1}, \ldots, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n, x_{n+1}, \ldots, \\ \ldots\ldots\ldots\ldots \end{array}\right\} \quad (*)$$

If $\{x_n\}$ is bounded below, these sequences have infima

$$k_1, k_2, \ldots, k_n, \ldots .$$

For all n, $k_n \leq k_{n+1}$, since $\{x_n, x_{n+1}, \ldots\} \supset \{x_{n+1}, x_{n+2}, \ldots\}$.
Hence k_n tends either to a finite limit or to $+\infty$.

If $\{x_n\}$ is not bounded below, none of the sequences (*) is bounded below, i.e. $\inf\limits_{r \geq n} x_r = -\infty$ for all n. We are therefore led to the following definition.

1.17 <u>Definition</u>. Let $\{x_n\}$ be a sequence of real numbers. If $\{x_n\}$ is bounded below,

$$b = \lim_{n \to \infty} (\inf_{r \geq n} x_r)$$

is called the <u>lower limit</u> (or <u>limit inferior</u>) of $\{x_n\}$ and we write

$$\underline{\lim}_{n \to \infty} x_n = b \quad \text{or} \quad \lim_{n \to \infty} \inf x_n = b.$$

If $\{x_n\}$ is not bounded below, we write

$$\underline{\lim}_{n \to \infty} x_n = -\infty \text{ or } \lim_{n \to \infty} \inf x_n = -\infty.$$

We note that lower limit of the sequence $\{x_n\}$ always exists. This follows from the monotonicity of the sequence $\{k_n\}$, where $k_n = \inf\limits_{r \geq n} x_r$.

1.18 Remark. If a real number b is the lower limit of the sequence $\{x_n\}$, then given $\varepsilon > 0$, $x_n > b - \varepsilon$ for all sufficiently large n.

Proof. Suppose that $b = \lim_{n\to\infty} \inf x_n = \lim_{n\to\infty} k_n$, where $k_n = \inf_{r \geq n} x_r$. Then, given $\varepsilon > 0$, there is N such that

$$k_N > b - \varepsilon$$

and so

$$x_n > b - \varepsilon \qquad \text{for } n \geq N.$$

Q.E.D.

• N-dimensional Euclidean Spaces

1.19 Definition. Let N be a positive integer. An ordered set $(x_1, \ldots, x_N)$ of N real numbers is called a vector with N components, or a point of N-dimensional Euclidean space $\mathbb{R}^N$.

For any two points $x = (x_1, \ldots, x_N)$, $y = (y_1, \ldots, y_N)$ of $\mathbb{R}^N$, we define

$$x + y = (x_1 + y_1, \ldots, x_N + y_N)$$

and

$$cx = (cx_1, \ldots, cx_N)$$

for any real c.

The zero element of $\mathbb{R}^N$ is

$$0 = (0, \ldots, 0).$$

We define the <u>norm</u> (or <u>length</u>) of x by $|x| = (x_1^2 + \dots + x_N^2)^{\frac{1}{2}}$.

Metric Spaces

1.20 <u>Definition</u>. A <u>metric space</u> is a set M with a real-valued function $d(x,y)$ on $M \times M$ called a distance function. This function is required to have the properties:

(a) $d(x,y) \geq 0$; $d(x,y) = 0$ if and only if $x = y$.

(b) $d(x,y) = d(y,x)$.

(c) $d(x,y) \leq d(x,z) + d(z,y)$, for any $z \in M$.

1.21 <u>Definition</u>. A sequence $\{x_n\}$ in a metric space (M,d) is said to <u>converge</u> to a point x if for any $\varepsilon > 0$ there is a positive integer N such that $d(x_n,x) < \varepsilon$ whenever $n \geq N$. A sequence $\{x_n\}$ is said to be a <u>Cauchy sequence</u> if for any $\varepsilon > 0$ there exists an N such that $d(x_m,x_n) < \varepsilon$ whenever $m,n \geq N$. A metric space (M,d) is said to be <u>complete</u> if every Cauchy sequence in (M,d) converges to a limit in (M,d).

Normed Vector Spaces

1.22 <u>Definition</u>. A set X of elements is called a <u>vector space</u> over the reals if we have a function + on $X \times X$ to X and a function $\cdot$ on $\mathbb{R} \times X$ to X which satisfy the following conditions:

1. $x + y = y + x$; $x, y \in X$.
2. $(x+y)+z = x+(y+z)$; $x, y, z \in X$.
3. There is a vector 0 in X such that, for all x in X, $x+0 = x$.
4. $a(x+y) = ax + ay$; $a \in \mathbb{R}$, $x, y \in X$.

5. $(a+b)x = ax + bx$; $a, b \in \mathbb{R}$, $x \in X$.

6. $a(bx) = (ab)x$; $a, b \in \mathbb{R}$, $x \in X$.

7. $0 \cdot x = 0$.

8. $1 \cdot x = x$.

1.23 Example. $\mathbb{R}^N$ is a vector space under the operations

$$x + y = (x_1, \ldots, x_N) + (y_1, \ldots, y_N) = (x_1+y_1, \ldots, x_N+y_N), \text{ and}$$

$$c(x_1, \ldots, x_N) = (cx_1, \ldots, cx_N),$$

where $c \in \mathbb{R}$, $x, y \in \mathbb{R}^N$.

1.24 Definition. A normed vector space is a vector space X with a real-valued function $x \rightarrow \| x \|$ satisfying the conditions:

(a) $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$;

(b) $\|ax\| = |a| \, \|x\|$ for each real number a;

(c) $\|x+y\| \leq \|x\| + \|y\|$.

1.25 Theorem. Let X be a normed vector space. Then $d(x,y) = \|x-y\|$ is a metric on X.

Proof. Let x, y, z be any elements in X. Then

$$d(x,y) = \|x-y\| \geq 0,$$

and

$$d(x,y) = \|x-y\| = 0 \text{ if and only if } x-y = 0 \text{ i.e. } x = y.$$

By (c) in definition 1.24, we have

$$\|x-z\| \leq \|x-y\| + \|y-z\|$$

so that $d(x,z) \leq d(x,y) + d(y,z)$.

Finally by taking $a = -1$ in (b) in Definition 1.24, we have

$$\|x-y\| = \|y-x\|$$

so that $d(x,y) = d(y,x)$.

Q.E.D.

1.26 Definition. A Banach space is a normed vector space which is complete in the metric defined by its norm.

1.27 Example. $\mathbb{R}^N$ is a Banach space.

Convex Sets in Normed Vector Spaces

1.28 Definition. A subset E of $\mathbb{R}^N$ is convex if

$$(*) \qquad tx + (1-t)y \in E$$

whenever $x \in E$, $y \in E$, and $0 \leq t \leq 1$.

The set of points (*) is called the segment between x and y.

By the closed N-disk with center at x_0 and radius r we mean the set

$$E = \{ x \in \mathbb{R}^N \mid |x-x_0| \leq r \}$$

where $|x| = \sqrt{x_1^2 + \dots + x_N^2}$.

If we define the norm of $x \in \mathbb{R}^N$ by

$$\|x\|_1 = \max \{ |x_1|, \dots, |x_N| \} ,$$

then the set

$$E = \{ x \in \mathbb{R}^N \mid \|x-x_0\|_1 \leq r \}$$

is called the closed N-cell.

1.29 Theorem. Closed N-disks are convex.

Proof. Let E be a closed N-disk center at x_0 with radius r. If $|x-x_0| \leq r$, $|y-x_0| \leq r$, and $0 \leq t \leq 1$, then

$$\begin{aligned} |tx+(1-t)y-x_0| &= |t(x-x_0)+(1-t)(y-x_0)| \\ &\leq t\,|x-x_0|+(1-t)|y-x_0| \\ &\leq tr + (1-t)r = r. \end{aligned}$$

Q.E.D.

The proof is the same if E is a closed N-cell.

Equivalence of Norms

For any $x = (x_1,\ldots,x_N) \in \mathbb{R}^N$, we define the Euclidean norm of x by

$$|x| = \sqrt{x_1^2 + \ldots + x_N^2} .$$

The following two functions are also norms on $\mathbb{R}^N$:

$$\|x\|_1 = \max\{|x_1|, \ldots, |x_N|\}$$

$$\|x\|_2 = |x_1| + \ldots + |x_N| .$$

The purpose of this section is to prove that in $\mathbb{R}^N$ (or in any finite dimensional vector space) the limit definitions are independent of the choice of norms. For example, if x_0 is a limit point of a set E with respect to one norm, then it is a limit point of E with respect to every norm, and the same goes for the other limit definitions.

We first give the following definition.

1.30 <u>Definition</u>. For any two norms $\| \|_1$ and $\| \|_2$ on $\mathbb{R}^N$, we define $\| \|_1$ to be <u>equivalent</u> to $\| \|_2$ if there exist positive real numbers k and K such that, for any x in $\mathbb{R}^N$,

$$k\|x\|_1 \leq \|x\|_2 \leq K\|x\|_1 \qquad ....(*)$$

It can be shown that this is a true equivalence relation, that is, it satisfies the three requirements:

Reflexivity: Every norm is equivalent to itself.

Symmetry: $\| \|_1$ is equivalent to $\| \|_2$ if and only if $\| \|_2$ is equivalent to $\| \|_1$.

Transitivity: If $\| \|_1$ is equivalent to $\| \|_2$ and if $\| \|_2$ is equivalent to $\| \|_3$, then $\| \|_1$ is equivalent to $\| \|_3$.

We now verify the preceding contention, let $\| \|_1$, and $\| \|_2$ be equivalent norms, and suppose x_0 is a limit point of E with respect to $\| \|_1$. Then, for any $\varepsilon_1 > 0$, there exists a point x in E such that $0 < \|x-x_0\|_1 < \varepsilon_1$. Thus, if $\varepsilon_2 > 0$ is given arbitrarily, we may set $\varepsilon_1 = \varepsilon_2/K$ and obtain, by inequality (*),

$$0 < \|x-x_0\|_2 \leq K\|x-x_0\|_1 < K\varepsilon_1 = \varepsilon_2.$$

Hence, x_0 is a limit point of E with respect to $\| \|_2$.

We now turn to the principal theorem.

1.31 Theorem. Any two norms on $\mathbb{R}^N$ are equivalent.

Proof. Let $\| \ \|$ be an arbitrary norm on $\mathbb{R}^N$. Choose a basis $\{e_1,\ldots,e_N\}$ for $\mathbb{R}^N$, and define a Euclidean norm on $\mathbb{R}^N$ by setting

$$|x| = \sqrt{x_1^2 + \ldots + x_N^2} \ ,$$

for any $x = x_1e_1 + \ldots + x_Ne_N$. We shall show that $\| \ \|$ is equivalent to $| \ |$, that is, there exist positive real numbers k and K such that $k|x| \leq \|x\| \leq K|x|$, for all x in $\mathbb{R}^N$. By the transitivity property of the equivalence relation between norms, it then follows that any two norms on $\mathbb{R}^N$ are equivalent.

For any $x = x_1e_1 + \ldots + x_Ne_N$, we have

$$\|x\| \leq \sum_{i=1}^{N} |x_i| \|e_i\| \leq \left(\sum_{i=1}^{N} \|e_i\| \right) \max_i \{|x_i|\}$$
$$\leq \left(\sum_{i=1}^{N} \|e_i\| \right) \left(\sum_{i=1}^{N} x_i^2 \right)^{\frac{1}{2}} = K|x|$$

where $K = \sum_{i=1}^{N} \|e_i\| > 0$. We now prove that k exists. We contend that as a function of x the real-valued function $\| \ \|$ is continuous with respect to the Euclidean norm $| \ |$. For if $\varepsilon > 0$ is given, we pick $\delta = \varepsilon / K$. Then, if $|x - x_0| < \delta$,

$$\big| \|x\| - \|x_0\| \big| \leq \|x - x_0\| \leq K|x - x_0| < \varepsilon.$$

Let k be the minimum value of the function $\| \ \|$ restricted to the Euclidean unit sphere $|x| = 1$. Then, for any $x \neq 0$, it follows that $\|x/|x|\| \geq k$, and hence that

$$\|x\| \geq k|x|, \text{ for any } x \text{ in } \mathbb{R}^N.$$

Q.E.D.

Inner Product Spaces

1.32 Definition. An inner product on a real vector space X is a real-valued function on $X \times X$, whose values are denoted by $\langle x,y \rangle$, which has the following properties:

(a) $\langle x,x \rangle \geq 0$; $\langle x,x \rangle = 0$ if and only if $x = 0$;

(b) $\langle x,y \rangle = \langle y,x \rangle$;

(c) $\langle x+y,z \rangle = \langle x,z \rangle + \langle y,z \rangle$;

(d) $\langle ax,y \rangle = a\langle x,y \rangle$, for any real number a.

A vector space with inner product defined on it is called an inner product space.

The following properties are immediate consequences of the definition of an inner product.

$$\langle x,y+z \rangle = \langle x,y \rangle + \langle x,z \rangle$$

$$\langle x,ay \rangle = a\langle x,y \rangle$$

$$\langle x,0 \rangle = 0 = \langle 0,x \rangle .$$

1.33 Theorem. If X is an inner product space, then $\| x \| = \sqrt{\langle x,x \rangle}$ defines a norm on X, and for all x and y in X,

$$|\langle x,y \rangle| \leq \| x\| \| y \| \quad \text{(Schwarz inequality)}.$$

Proof. The properties $\| ax \| = |a| \, \| x\|$ and $\|x\| = 0$ if and only if $x = 0$ are immediate. The inequality $|\langle x,y \rangle| \leq \| x\| \| y\|$ is true if either $x = 0$ or $y = 0$; otherwise it can be written

$|\langle x\|x\|^{-1}, y\|y\|^{-1}\rangle| \leq 1$, and thus it suffices to show that $|\langle x,y\rangle| \leq 1$ whenever $\|x\| = \|y\| = 1$.

Now, for any real number a, we have

$$0 \leq \langle x+ay, x+ay\rangle = \langle x,x\rangle + a\langle y,x\rangle + a\langle x,y\rangle + a^2\langle y,y\rangle$$

or

$$0 \leq 1 + 2a\langle x,y\rangle + |a|^2.$$

Choosing $a = -\langle x,y\rangle$, we obtain

$$0 \leq 1 - |\langle x,y\rangle|^2$$

or $|\langle x,y\rangle| \leq 1$, as desired.

The triangle inequality is now easily established; we have

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

so that

$$\|x+y\| \leq \|x\| + \|y\| .$$

Q.E.D.

The norm associated with an inner product satisfies the parallelogram law:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

This is easily verified:

$$\begin{aligned} \|x+y\|^2 + \|x-y\|^2 &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &\quad + \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle \\ &= 2(\|x\|^2 + \|y\|^2). \end{aligned}$$

Geometrically the parallelogram law expresses the fact that in a parallelogram the sum of the squares of the lengths of the diagonals is equal to the sum of the squares of the lengths of the four sides.

1.34 Definition. An inner product space which is complete is called a Hilbert space.

1.35 Rxample. $\mathbb{R}^N$ is a Hilbert space.