9

EPILOGUE

9.1 SOME COMPACTNESS RESULTS

There are two major loose ends in our development of real analysis. We have never formally defined the real numbers, and when we developed the basic topology of Euclidean $n$-space, we assumed the Cantor nested set property (Section 2.2.2) without apology. We now attempt to face these problems; a convenient way to do this is via some general compactness results in metric spaces.

9.1.1 Definitions and Comments

Let $A$ be a subset of the metric space $\Omega$. We say that $A$ is totally bounded if for every $\epsilon > 0$, $A$ can be covered by finitely many open balls of radius $\epsilon$.

It follows from the definition that any totally bounded set is bounded, but the converse is not true. For example, let $\Omega$ be any set, and let $d(x,y) = 1$, $x \neq y$; $d(x,x) = 0$, $x, y \in \Omega$. Then all subsets of $\Omega$ are bounded, but if $\Omega$ has infinitely many points, then $\Omega$ is not totally bounded. (If $0 < r < 1$, then $B_r(x) = \{x\}$.) In $\mathbb{R}^p$, however, a bounded set must be totally bounded. (Intuitively, a bounded set
can be placed inside a large rectangular box, which can be broken up into a finite number of small boxes of maximum dimension \( \epsilon \). Each of the small boxes can be covered by a finite number of open balls of radius \( \epsilon \).

If \( A \) is totally bounded, then \( A \) has a countable dense subset \( E \). To see this, note that for any positive integer \( n \), \( A \) can be covered by finitely many balls \( B_{1/n}(x_{nj}), j = 1, 2, \ldots, m_n \), with centers \( x_{nj} \in A \). The set \( E \) consisting of all \( x_{nj} \) is countable and dense in \( A \) (see the proof of Lemma 7.4.1).

If every sequence in \( A \) has a convergent subsequence (the limit need not be in \( A \)), then \( A \) is totally bounded. For if \( A \) cannot be covered by finitely many open balls of radius \( \epsilon \), select, inductively, a sequence of points \( x_n \in A \) with \( x_2 \notin B_\epsilon(x_1), x_3 \notin B_\epsilon(x_1) \cup B_\epsilon(x_2), \ldots, x_n \notin \bigcup_{i=1}^{n-1} B_\epsilon(x_i), \ldots \). If \( m > n \), then \( x_m \notin B_\epsilon(x_n) \); hence \( d(x_m, x_n) \geq \epsilon \).

Thus, \( \{x_n\} \) has no convergent subsequence.

We may now give two general compactness criteria in metric spaces.

9.1.2 THEOREM. Let \( K \) be a subset of the metric space \( \Omega \). Then \( K \) is compact if and only if every sequence in \( K \) has a subsequence converging to a point of \( K \).

Proof. The "only if" part is Theorem 2.3.1, so assume that every sequence in \( K \) has a subsequence converging to a point of \( K \). As shown in Section 9.1.1, \( K \) is totally bounded and therefore has a countable dense subset. By Lemma 7.4.1, every open covering of \( K \) has a countable subcovering. Thus, we may start with the assumption that \( K \subseteq \bigcup_{i=1}^{\infty} G_i \), where the \( G_i \) are open sets, and attempt to find a finite subcovering.

If for every \( n \), \( K \) is not a subset of \( \bigcup_{i=1}^{n} G_i \), select \( x_n \in K \) with \( x_n \notin \bigcup_{i=1}^{n} G_i \). By hypothesis there is a subsequence converging to a limit \( x \in K \). Now \( x \) belongs to some \( G_j \), and therefore the subsequence is in \( G_j \) eventually. Thus, \( x_n \in G_j \) for some \( n > j \), contradicting \( x_n \notin \bigcup_{i=1}^{n} G_i \).

9.1.3 THEOREM. Let \( K \) be a subset of the metric space \( \Omega \). Then \( K \) is compact if and only if \( K \) is totally bounded and complete. (Note that
completeness of K means that every Cauchy sequence in K converges to a point in K.)

Proof. If K is compact, then K is totally bounded by the definition of compactness. (Note that for each ε > 0, K ⊆ ∪_{x∈K} B_ε(x), and by compactness there is a finite subcovering. Alternatively, Theorem 9.1.2 and Section 9.1.1 can be used, but this is a rather indirect approach.) Let \{x_n\} be a Cauchy sequence in K; by Theorem 9.1.2 there is a subsequence \{x_{n_k}\} converging to a point x ∈ K. We proceed as in the proof of Theorem 2.4.5. Given ε > 0, choose N so that 
d(x_{n_k}, x_m) < ε/2 for all n, m ≥ N, and choose k_0 so that n_k ≥ N and 
d(x_{n_k}, x) < ε/2 for all k ≥ k_0. Then for any fixed k ≥ k_0,

d(x_n, x) ≤ d(x_n, x_{n_k}) + d(x_{n_k}, x) < ε for all n ≥ N.

Thus, x_n → x ∈ K.

Conversely, assume K totally bounded and complete, and let x_1, x_2, ... be a sequence in K. By Theorem 9.1.2, it suffices to produce a subsequence converging to a point of K. Now for any r, K is covered by finitely many open balls of radius r; hence, for some ball B = B_r(y) in the covering we must have x_k ∈ B for infinitely many k. Using this idea successively for r = 1/n, n = 1, 2, ... , we obtain the following array, with row n a subsequence of row n - 1, n = 2, 3, ... , and row 1 a subsequence of \{x_n\}:

\[
\begin{array}{cccc}
x_{11} & x_{12} & x_{13} & \cdots & \in B_1(y_1) \\
x_{21} & x_{22} & x_{23} & \cdots & \in B_{1/2}(y_2) \\
\vdots & & & & \\
x_{n1} & x_{n2} & x_{n3} & \cdots & \in B_{1/n}(y_n) \\
\vdots & & & & 
\end{array}
\]

Let z_n = x_{nn}, n = 1, 2, ... . If m ≥ n, then z_m ∈ B_{1/n}(y_n); hence,

\[
d(z_m, z_n) ≤ d(z_m, y_n) + d(y_n, z_n) ≤ \frac{1}{n} + d(y_n, x_{nn}) ≤ \frac{1}{n} + \frac{1}{n} → 0 \quad \text{as} \quad n → \infty.
\]
Thus, \( \{z_n\} \) is a Cauchy sequence, which by completeness converges to a point of \( K \).

9.2 REPLACING CANTOR'S NESTED SET PROPERTY

We can now examine the problem of replacing our assumption 2.2.2 by something more natural. Suppose we assume instead that \( \mathbb{R}^p \) is complete. The Heine–Borel Theorem can then be proved. For if \( K \) is a closed and bounded subset of \( \mathbb{R}^p \), then \( K \) is totally bounded by Section 9.1.1, and therefore compact by Theorem 9.1.3. (Note that \( K \), a closed subset of the complete space \( \mathbb{R}^p \), must be complete.) From the Heine–Borel Theorem we can prove Cantor's nested set theorem, as follows. Let \( B_1 \supseteq B_2 \supseteq \ldots \) be a nested sequence of closed, bounded, nonempty subsets of \( \mathbb{R}^p \), and pick \( x_n \in B_n, n = 1, 2, \ldots \). Since \( B_1 \) is compact, there is by Theorem 9.1.2 a subsequence \( \{x_{n_j}\} \) converging to a limit \( x \in B_1 \). For any positive integer \( k \), we have \( n_j \geq k \) for all sufficiently large \( j \); hence \( x_{n_j} \in B_{n_j} \subseteq B_k \). It follows that \( x \in B_k \) for every \( k \), as desired.

Now it follows directly from the Euclidean distance formula that a sequence of points \( x_n = (x_{n1}, \ldots, x_{np}) \in \mathbb{R}^p \) is Cauchy in \( \mathbb{R}^p \) if and only if each of the component sequences \( \{x_{nk}, n = 1, 2, \ldots\} \) is Cauchy in \( \mathbb{R} \). Similarly, convergence of \( \{x_n\} \) is equivalent to convergence of each \( \{x_{ni}\}, i = 1, \ldots, p \). Thus, it is sufficient to assume \( \mathbb{R} \) complete. But in fact completeness of \( \mathbb{R} \) follows from the assumption that every nonempty subset of \( \mathbb{R} \) that has an upper bound has a supremum (equivalently, every nonempty subset of \( \mathbb{R} \) that has a lower bound has an infimum). For under this assumption, let \( \{x_n\} \) be a Cauchy sequence in \( \mathbb{R} \), and let \( L = \inf_n \sup_{k \geq n} x_k \) (the inf and the sup must exist because \( \{x_n\} \) is bounded; see Section 2.4.4).

We show that \( x_n \to L \). If \( \epsilon > 0 \), then \( \inf_n \sup_{k \geq n} x_k > L - \epsilon \); hence, for every \( n \), \( \sup_{k \geq n} x_k > L - \epsilon \). Thus,

\[
(\forall n)(\exists k \geq n)(x_k > L - \epsilon).
\]

(1)

Also, \( \inf_n \sup_{k \geq n} x_k < L + \epsilon \), so for some \( n \), \( \sup_{k \geq n} x_k < L + \epsilon \). Therefore,

\[
(\exists n)(\forall k \geq n)(x_k < L + \epsilon).
\]

(2)
Now by the Cauchy property we have, for some \( N \), \( |x_m - x_n| < \epsilon \) for all \( m, n \geq N \). By (2) there is an \( n_0 \) such that \( x_k < L + \epsilon \) for all \( k \geq n_0 \), and by (1) we can find \( k_0 \geq \max(n_0, N) \) such that \( x_{k_0} > L - \epsilon \). Thus, \( L - \epsilon < x_{k_0} < L + \epsilon \). Finally,

\[
|x_n - L| \leq |x_n - x_{k_0}| + |x_{k_0} - L| < 2\epsilon \quad \text{for all } \ n \geq N.
\]

Therefore \( x_n \to L \).

9.3 THE REAL NUMBERS REVISITED

When the real numbers are introduced by means of a set of axioms, the standard assumptions are as follows.

1. \( R \) is a field.

Intuitively, this means that addition, subtraction, multiplication, and division can be carried out without leaving the set \( R \).

2. \( R \) is an ordered field.

The idea is that there is an ordering on \( R \), denoted by "<," such that if \( y < z \), then \( x + y < x + z \); and if \( x > 0 \) and \( y > 0 \), then \( xy > 0 \). Also for all \( x \) and \( y \), exactly one of the three conditions \( x < y \), \( x = y \), \( x > y \) holds.

3. \( R \) has the least upper bound property.

In other words, every nonempty subset of \( R \) that has an upper bound has a least upper bound.

Of course, we have glossed over a host of formal details, but this material belongs more to the areas of logic and algebra than to analysis. The student may consult a standard textbook to fill in the gaps.