Facts about metric spaces.

Math. 212a

August 28, 2001

Contents

Chapter 1

Metric spaces

1.1 Metric spaces

A **metric** for a set X is a function d from X to the real numbers $\mathbf{R}$,

$$d : X \times X \to \mathbf{R}$$

such that for all $x, y, z \in X$

1. $d(x, y) = d(y, x)$
2. $d(x, z) \leq d(x, y) + d(y, z)$
3. $d(x, x) = 0$
4. If $d(x, y) = 0$ then $x = y$.

The inequality in 2) is known as the **triangle inequality** since if X is the plane and d the usual notion of distance, it says that the length of an edge of a triangle is at most the sum of the lengths of the two other edges. (In the plane, the inequality is strict unless the three points lie on a line.)

Condition 4) is in many ways inessential, and it is often convenient to drop it, especially for the purposes of some proofs. For example, we might want to consider the decimal expansions $.49999\ldots$ and $.50000\ldots$ as different, but as having zero distance from one another. Or we might want to "identify" these two decimal expansions as representing the same point.

A function d which satisfies only conditions 1) - 3) is called a **pseudo-metric**.

A **metric space** is a pair (X, d) where X is a set and d is a metric on X. Almost always, when d is understood, we engage in the abuse of language and speak of "the metric space X".

Similarly for the notion of a **pseudo-metric space**.

In like fashion, we call $d(x, y)$ the **distance** between x and y, the function d being understood.

If r is a positive number and $x \in X$, the (open) **ball of radius** r about x is defined to be the set of points at distance less than r from x and is denoted by $B_r(x)$. In symbols,

$$B_r(x) := \{y | \ d(x, y) < r\}.$$

If r and s are positive real numbers and if x and z are points of a pseudo-metric space X, it is possible that $B_r(x) \cap B_s(z) = \emptyset$. This will certainly be the case if $d(x, z) > r + s$ by virtue of the triangle inequality. Suppose that this intersection is not empty and that

$$w \in B_r(x) \cap B_s(z).$$

If $y \in X$ is such that $d(y, w) < \min[r - d(x, w), s - d(z, w)]$ then the triangle inequality implies that $y \in B_r(x) \cap B_s(z)$. Put another way, if we set $t := \min[r - d(x, w), s - d(z, w)]$ then

$$B_t(w) \subset B_r(x) \cap B_s(z).$$

Put still another way, this says that the intersection of two (open) balls is either empty or is a union of open balls. So if we call a set in X **open** if either it is empty, or is a union of open balls, we conclude that the intersection of any finite number of open sets is open, as is the union of any number of open sets. In technical language, we say that the open balls form a base for a topology on X.

A map $f : X \to Y$ from one pseudo-metric space to another is called **continuous** if the inverse image under f of any open set in Y is an open set in X. Since an open set is a union of balls, this amounts to the condition that the inverse image of an open ball in Y is a union of open balls in X, or, to use the familiar ϵ, δ language, that if $f(x) = y$ then for every $\epsilon > 0$ there exists a $\delta = \delta(x, \epsilon) > 0$ such that

$$f(B_\delta(x)) \subset B_\epsilon(y).$$

Notice that in this definition δ is allowed to depend both on x and on ϵ. The map is called **uniformly continuous** if we can choose the δ independently of x.

An even stronger condition on a map from one pseudo-metric space to another is the **Lipschitz condition**. A map $f : X \to Y$ from a pseudo-metric space (X, d_X) to a pseudo-metric space (Y, d_Y) is called a **Lipschitz map** with **Lipschitz constant** C if

$$d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2) \quad \forall x_1, x_2 \in X.$$

Clearly a Lipschitz map is uniformly continuous.

For example, suppose that A is a fixed subset of a pseudo-metric space X. Define the function $d(A, \cdot)$ from X to $\mathbf{R}$ by

$$d(A, x) := \inf\{d(x, w), \ w \in A\}.$$

The triangle inequality says that

$$d(x,w) \leq d(x,y) + d(y,w)$$

for all w, in particular for $w \in A$, and hence taking lower bounds we conclude that

$$d(A,x) \leq d(x,y) + d(A,y).$$

or

$$d(A,x) - d(A,y) \leq d(x,y).$$

Reversing the roles of x and y then gives

$$|d(A,x) - d(A,y)| \leq d(x,y).$$

Using the standard metric on the real numbers where the distance between a and b is $|a-b|$ this last inequality says that $d(A,\cdot)$ is a Lipschitz map from X to $\mathbf{R}$ with $C=1$.

A closed set is defined to be a set whose complement is open. Since the inverse image of the complement of a set (under a map f) is the complement of the inverse image, we conclude that the inverse image of a closed set under a continuous map is again closed.

For example, the set consisting of a single point in $\mathbf{R}$ is closed. Since the map $d(A,\cdot)$ is continuous, we conclude that the set

$$\{x|d(A,x)=0\}$$

consisting of all point at zero distance from A is a closed set. It clearly is a closed set which contains A. Suppose that S is some closed set containing A, and $y \notin S$. Then there is some $r>0$ such that $B_r(y)$ is contained in the complement of S, which implies that $d(y,w) \geq r$ for all $w \in S$. Thus $\{x|d(A,x)=0\} \subset S$. In short $\{x|d(A,x)=0\}$ is a closed set containing A which is contained in all closed sets containing A. This is the definition of the **closure** of a set, which is denoted by $\overline{A}$. We have proved that

$$\overline{A} = \{x|d(A,x)=0\}.$$

In particular, the closure of the one point set $\{x\}$ consists of all points u such that $d(u,x)=0$.

Now the relation $d(x,y)=0$ is an equivalence relation, call it R. (Transitivity being a consequence of the triangle inequality.) This then divides the space X into equivalence classes, where each equivalence class is of the form $\overline{\{x\}}$, the closure of a one point set. If $u \in \overline{\{x\}}$ and $v \in \overline{\{y\}}$ then

$$d(u,v) \leq d(u,x) + d(x,y) + d(y,v) = d(x,y).$$

since $x \in \overline{\{u\}}$ and $y \in \overline{\{v\}}$ we obtain the reverse inequality, and so

$$d(u,v) = d(x,y).$$

In other words, we may define the distance function on the quotient space X/R, i.e. on the space of equivalence classes by

$$d(\overline{\{x\}}, \overline{\{y\}}) := d(u, v), \quad u \in \overline{\{x\}}, v \in \overline{\{y\}}$$

and this does not depend on the choice of u and v. Axioms 1)-3) for a metric space continue to hold, but now

$$d(\overline{\{x\}}, \overline{\{y\}}) = 0 \Rightarrow \overline{\{x\}} = \overline{\{y\}}.$$

In other words, X/R is a *metric* space. Clearly the projection map $x \mapsto \overline{\{x\}}$ is an isometry of X onto X/R. (An **isometry** is a map which preserves distances.) In particular it is continuous. It is also open.

In short, we have provided a canonical way of passing (via an isometry) from a pseudo-metric space to a metric space by identifying points which are at zero distance from one another.

A subset A of a pseudo-metric space X is called *dense* if its closure is the whole space. From the above construction, the image A/R of A in the quotient space X/R is again dense. We will use this fact in the next section in the following form:

If $f : Y \to X$ is an isometry of Y such that $f(Y)$ is a dense set of X, then f descends to a map F of Y onto a dense set in the metric space X/R.

1.2 Completeness and completion.

The usual notion of convergence and Cauchy sequence go over unchanged to metric spaces or pseudo-metric spaces Y. A sequence $\{y_n\}$ is said to **converge** to the point y if for every $\epsilon > 0$ there exists an $N = N(\epsilon)$ such that

$$d(y_n, y) < \epsilon \quad \forall\, n > N.$$

A sequence $\{y_n\}$ is said to be **Cauchy** if for any $\epsilon > 0$ there exists an $N = N(\epsilon)$ such that

$$d(y_n, y_m) < \epsilon \quad \forall\, m, n > N.$$

The triangle inequality implies that every convergent sequence is Cauchy. But not every Cauchy sequence is convergent. For example, we can have a sequence of rational numbers which converge to an irrational number, as in the approximation to the square root of 2. So if we look at the set of rational numbers as a metric space R in its own right, not every Cauchy sequence of rational numbers converges in R. We must "complete" the rational numbers to obtain $\mathbf{R}$, the set of real numbers. We want to discuss this phenomenon in general.

So we say that a (pseudo-)metric space is **complete** if every Cauchy sequence converges. The key result of this section is that we can always "complete" a metric or pseudo-metric space. More precisely, we claim that

Any metric (or pseudo-metric) space can be mapped by a one to one isometry onto a dense subset of a complete metric (or pseudo-metric) space.

By the italicized statement of the preceding section, it is enough to prove this for a pseudo-metric spaces X. Let X_{seq} denote the set of Cauchy sequences in X, and define the distance between the Cauchy sequences $\{x_n\}$ and $\{y_n\}$ to be

$$d(\{x_n\}, \{y_n\}) := \lim_{n \to \infty} d(x_n, y_n).$$

It is easy to check that d defines a pseudo-metric on X_{seq}. Let $f : X \to X_{seq}$ be the map sending x to the sequence all of whose elements are x;

$$f(x) = (x, x, x, x, \cdots).$$

It is clear that f is one to one and is an isometry. The image is dense since by definition

$$\lim d(f(x_n), \{x_n\}) = 0.$$

Now since $f(X)$ is dense in X_{seq}, it suffices to show that any Cauchy sequence of points of the form $f(x_n)$ converges to a limit. But such a sequence converges to the element $\{x_n\}$. QED

1.3 Normed vector spaces and Banach spaces.

Of special interest are vector spaces which have metric which is compatible with the vector space properties and which is complete: Let V be a vector space over the real or complex numbers. A **norm** is a real valued function

$$v \mapsto \|v\|$$

on V which satisfies

1. $\|v\| \geq 0$ and > 0 if $v \neq 0$,

2. $\|cv\| = |c|\|v\|$ for any real (or complex) number c, and

3. $\|v + w\| \leq \|v\| + \|w\| \ \forall \ \ v, w \in V$.

Then $d(v, w) := \|v - w\|$ is a metric on V, which satisfies $d(v+u, w+u) = d(v, w)$ for all $v, w, u \in V$. The ball of radius r about the origin is then the set of all v such that $\|v\| < r$. A vector space equipped with a norm is called a **normed vector space** and if it is complete relative to the metric it is called a **Banach space**.

Our construction shows that any vector space with a norm can be completed so that it becomes a Banach space.

1.4 Compactness.

A topological space X is said to be **compact** if it has one (and hence the other) of the following equivalent properties:

- Every open cover has a finite subcover. In more detail: if $\{U_\alpha\}$ is a collection of open sets with

  $$X \subset \bigcup_\alpha U_\alpha$$

  then there are finitely many $\alpha_1, \ldots, \alpha_n$ such that

  $$X \subset U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}.$$

- If $\mathcal{F}$ is a family of closed sets such that

  $$\bigcap_{F \in \mathcal{F}} = \emptyset$$

  then a finite intersection of the F's are empty:

  $$F_1 \cap \cdots \cap F_n = \emptyset.$$

1.5 Total Boundedness.

A metric space X is said to be **totally bounded** if for every $\epsilon > 0$ there are finitely many open balls of radius ϵ which cover X.

Theorem 1.5.1 *The following assertions are equivalent for a metric space:*

1. *X is compact.*
2. *Every sequence in X has a convergent subsequence.*
3. *X is totally bounded and complete.*

Proof that 1. $\Rightarrow$ 2. Let $\{y_i\}$ be a sequence of points in X. We first show that there is a point x with the property for every $\epsilon > 0$, the open ball of radius ϵ centered at x contains the points y_i for infinitely many i. Suppose not. Then for any $z \in X$ there is an $\epsilon > 0$ such that the ball $B_\epsilon(x)$ contains only finitely many y_i. Since $z \in B_\epsilon(z)$, the set of such balls covers X. By compactness, finitely many of these balls cover X, and hence there are only finitely many i, a contradiction.

Now choose i_1 so that y_{i_1} is in the ball of radius $\frac{1}{2}$ centered at x. Then choose $i_2 > i_1$ so that y_{i_2} is in the ball of radius $\frac{1}{4}$ centered at x and keep going. We have constructed a subsequence so that the points y_{i_k} converge to x. Thus we have proved that 1) implies 2.

Proof that 2. $\Rightarrow$ 3. If $\{x_j\}$ is a Cauchy sequence in X, it has a convergent subsequence by hypothesis, and the limit of this subsequence is (by the triangle inequality) the limit of the original sequence. Hence X is complete. We must show that it is totally bounded. Given $\epsilon > 0$, pick a point $y_1 \in X$ and let $B_\epsilon(y_1)$ be open ball of radius ϵ about y_1. If $B_\epsilon(y_1) = X$ there is nothing further to prove. If not, pick a point $y_2 \in X - B_\epsilon(y_1)$ and let $B_\epsilon(y_2)$ be the ball of radius ϵ about y_2. If $B_\epsilon(y_1) \cup B_\epsilon(y_2) = X$ there is nothing to prove. If not, pick a point $y_3 \in X - (B_\epsilon(y_1) \cup B_\epsilon(y_2))$ etc. This procedure can not continue indefinitely, for then we will have constructed a sequence of points which are all at a mutual distance $\geq \epsilon$ from one another, and this sequence has no Cauchy subsequence.
Proof that 3. $\Rightarrow$ 2. Let $\{x_j\}$ be a sequence of points in X which we relabel as $\{x_{1,j}\}$. Let $B_{1,\frac{1}{2}}, \dots, B_{n_1,\frac{1}{2}}$ be a finite number of balls of radius $\frac{1}{2}$ which cover X. Our hypothesis 3. asserts that such a finite cover exists. Infinitely many of the j must be such that the $x_{i,j}$ all lie in one of these balls. Relabel this subsequence as $\{x_{2,j}\}$. Cover X by finitely many balls of radius $\frac{1}{3}$. There must be infinitely many j such that all the $x_{2,j}$ lie in one of the balls. Relabel this subsequence as $\{x_{3,j}\}$. Continue. At the ith stage we have a subsequence $\{x_{i,j}\}$ of our original sequence (in fact of the preceding subsequence in the construction) all of whose points lie in a ball of radius $1/i$. Now consider the "diagonal" subsequence

$$x_{1,1},\ x_{2,2}, x_{3,3}, \dots .$$

All the points from $x_{i,i}$ on lie in a fixed ball of radius $1/i$ so this is a Cauchy sequence. Since X is assumed to be complete, this subsequence of our original sequence is convergent.

We have shown that 2. and 3. are equivalent. The hard part of the proof consists in showing that these two conditions imply 1. For this it is useful to introduce some terminology:

1.6 Separability.

A metric space X is called **separable** if it has a countable subset $\{x_j\}$ of points which are dense. For example $\mathbf{R}$ is separable because the rationals are countable and dense. Similarly, $\mathbf{R}^n$ is separable because the points all of whose coordinates are rational form a countable dense subset.

Proposition 1.6.1 *Any subset Y of a separable metric space X is separable (in the induced metric).*

Proof. Let $\{x_j\}$ be a countable dense sequence in X. Consider the set of pairs (j,n) such that

$$B_{1/2n}(x_j) \cap Y \neq \emptyset.$$

For each such (j,n) let $y_{j,n}$ be any point in this non-empty intersection. We claim that the countable set of points $y_{j,n}$ are dense in Y. Indeed, let y be any point of Y. Let n be any positive integer. We can find a point x_j such that $d(x_j,y) < 1/2n$ since the x_j are dense in X. But then $d(y,y_{j,n}) < 1/n$ by the triangle inequality. QED

Proposition 1.6.2 *Any totally bounded metric space X is separable.*

Proof. For each n let $\{x_{1,n}, \dots, x_{i_n,n}\}$ be the centers of balls of radius $1/n$ (finite in number) which cover X. Put all of these together into one sequence which is clearly dense. QED

A **base** for the open sets in a topology on a space X is a collection $\mathcal{B}$ of open set such that every open set of X is the union of sets of $\mathcal{B}$

Proposition 1.6.3 *A family $\mathcal{B}$ is a base for the topology on X if and only if for every $x \in X$ and every open set U containing x there is a $V \in \mathcal{B}$ such that $x \in V$ and $V \subset U$.*

Proof. If $\mathcal{B}$ is a base, then U is a union of members of $\mathcal{B}$ one of which must therefore contain x. Conversely, let U be an open subset of X. For each $x \in U$ there is a $V_x \subset U$ belonging to $\mathcal{B}$. The union of these over all $x \in U$ is contained in U and contains all the points of U, hence equals U. So $\mathcal{B}$ is a base. QED

1.7 Second Countability.

A topological space X is said to be **second countable** or to satisfy the **second axiom of countability** if it has a base $\mathcal{B}$ which is (finite or) countable.

Proposition 1.7.1 *A metric space X is second countable if and only if it is separable.*

Proof. Suppose X is separable with a countable dense set $\{x_i\}$. The open balls of radius $1/n$ about the x_i form a countable base: Indeed, if U is an open set and $x \in U$ then take n sufficiently large so that $B_{2/n}(x) \subset U$. Choose j so that $d(x_j, x) < 1/n$. Then $V := B_{1/n}(x_j)$ satisfies $x \in V \subset U$ so by Proposition 1.6.3 the set of balls $B_{1/n}(x_j)$ form a base and they constitute a countable set. Conversely, let $\mathcal{B}$ be a countable base, and choose a point $x_j \in U_j$ for each $U_j \in \mathcal{B}$. If x is any point of X, the ball of radius $\epsilon > 0$ about x includes some U_j and hence contains x_j. So the x_j form a countable dense set. QED

Proposition 1.7.2 Lindelof's theorem. *Suppose that the topological space X is second countable. Then every open cover has a countable subcover.*

Let $\mathcal{U}$ be a cover, not necessarily countable, and let $\mathcal{B}$ be a countable base. Let $\mathcal{C} \subset \mathcal{B}$ consist of those open sets V belonging to $\mathcal{B}$ which are such that $V \subset U$ where $U \in \mathcal{U}$. By Proposition 1.6.3 these form a (countable) cover. For each $V \in \mathcal{C}$ choose a $U_V \in \mathcal{U}$ such that $V \subset U_V$. Then the $\{U_V\}_{V \in \mathcal{C}}$ form a countable subset of $\mathcal{U}$ which is a cover. QED

1.8 Conclusion of the proof of Theorem 1.5.1.

Suppose that condition 2. and 3. of the theorem hold for the metric space X. By Proposition 1.6.2, X is separable, and hence by Proposition 1.7.1, X is second countable. Hence by Proposition 1.7.2, every cover $\mathcal{U}$ has a countable subcover. So we must prove that if $U_1,\ U_2,\ U_3,\ \dots$ is a sequence of open sets which cover X, then $X = U_1 \cup U_2 \cup \cdots \cup U_m$ for some finite integer m. Suppose not. For each m choose $x_m \in X$ with $x_m \notin U_1 \cup \cdots \cup U_m$. By condition 2. of Theorem 1.5.1, we may choose a subsequence of the $\{x_j\}$ which converge to some point x. Since $U_1 \cup \cdots \cup U_m$ is open, its complement is closed, and since $x_j \notin U_1 \cup \cdots \cup U_m$ for $j > m$ we conclude that $x \notin U_1 \cup \cdots \cup U_m$ for any m. This says that the $\{U_j\}$ do *not* cover X, a contradiction. QED

Putting the pieces together, we see that a closed bounded subset of $\mathbf{R}^m$ is compact. This is the famous Heine-Borel theorem. So Theorem 1.5.1 can be considered as a far reaching generalization of the Heine-Borel theorem.

1.9 Dini's lemma.

Let X be a metric space and let L denote the space of real valued continuous functions of compact support. So $f \in L$ means that f is continuous, and the closure of the set of all x for which $|f(x)| > 0$ is compact. Thus L is a real vector space, and $f \in L \rightarrow |f| \in L$. Thus if $f \in L$ and $g \in L$ then $f + g \in L$ and also max $(f, g) = \frac{1}{2}(f + g + |f - g|) \in L$ and min $(f, g) = \frac{1}{2}(f + g - |f - g|) \in L$.

For a sequence of elements in L (or more generally in any space of real valued functions) we write $f_n \downarrow 0$ to mean that the sequence of functions is monotone decreasing, and at each x we have $f_n(x) \rightarrow 0$.

Theorem 1.9.1 Dini's lemma. *If $f_n \in L$ and $f_n \downarrow 0$ then $\|f_n\|_\infty \rightarrow 0$. In other words, monotone decreasing convergence to 0 implies uniform convergence to zero for elements of L.*

Proof. Given $\epsilon > 0$, let $C_n = \{x | f_n(x) \geq \epsilon\}$. Then the C_n are compact, $C_n \supset C_{n+1}$ and $\bigcap_k C_k = \emptyset$. Hence a finite intersection is already empty, which means that $C_n = \emptyset$ for some n. This means that $\|f_n\|_\infty \leq \epsilon$ for some n, and hence, since the sequence is monotone decreasing, for all subsequent n. QED

1.10 The Lebesgue outer measure of an interval is its length.

For any subset $A \subset \mathbf{R}$ we define its **Lebesgue outer measure** by

$$m^*(A) := \inf \sum \ell(I_n) : \ I_n \text{ are intervals with } A \subset \bigcup I_n. \tag{1.1}$$

Here the length $\ell(I)$ of any interval $I = [a, b]$ is $b - a$ with the same definition for half open intervals $(a, b]$ or $[a, b)$, or open intervals. Of course if $a = -\infty$ and b is finite or $+\infty$, or if a is finite and $b = +\infty$ the length is infinite. So the infimum in (1.1) is taken over all covers of A by intervals. By the usual $\epsilon/2^n$ trick, i.e. by replacing each $I_j = [a_j, b_j]$ by $(a_j - \epsilon/2^{j+1}, b_j + \epsilon/2^{j+1})$ we may assume that the infimum is taken over open intervals. (Equally well, we could use half open intervals of the form $[a, b)$, for example.).

It is clear that if $A \subset B$ then $m^*(A) \leq m^*(B)$ since any cover of B by intervals is a cover of A. Also, if Z is any set of measure zero, then $m^*(A \cup Z) = m^*(A)$. In particular, $m^*(Z) = 0$ if Z has measure zero. Also, if $A = [a, b]$ is an interval, then we can cover it by itself, so

$$m^*([a, b]) \leq b - a,$$

and hence the same is true for $(a, b]$, $[a, b)$, or (a, b). If the interval is infinite, it clearly can not be covered by a set of intervals whose total length is finite, since if we lined them up with end points touching they could not cover an infinite interval. We still must prove that

$$m^*(I) = \ell(I) \tag{1.2}$$

if I is a finite interval. We may assume that $I = [c, d]$ is a closed interval by what we have already said, and that the minimization in (1.1) is with respect to a cover by open intervals. So what we must show is that if

$$[c, d] \subset \bigcup_i (a_i, b_i)$$

then

$$d - c \leq \sum_i (b_i - a_i).$$

We first apply Heine-Borel to replace the countable cover by a finite cover. (This only decreases the right hand side of preceding inequality.) So let n be the number of elements in the cover. We want to prove that if

$$[c, d] \subset \bigcup_{i=1}^{n} (a_i, b_i) \text{ then } d - c \leq \sum_{i=1}^{n} (b_i - a_i).$$

We shall do this by induction on n. If $n = 1$ then $a_1 < c$ and $b_1 > d$ so clearly $b_1 - a_1 > d - c$.

Suppose that $n \geq 2$ and we know the result for all covers (of all intervals $[c, d]$) with at most $n - 1$ intervals in the cover. If some interval (a_i, b_i) is disjoint from $[c, d]$ we may eliminate it from the cover, and then we are in the case of $n - 1$ intervals. So every (a_i, b_i) has non-empty intersection with $[c, d]$. Among the the intervals (a_i, b_i) there will be one for which a_i takes on the minimum possible value. By relabeling, we may assume that this is (a_1, b_1). Since c is covered, we must have $a_1 < c$. If $b_1 > d$ then (a_1, b_1) covers $[c, d]$ and there is nothing further to do. So assume $b_1 \leq d$. We must have $b_1 > c$ since $(a_1, b_1) \cap [c, d] \neq \emptyset$. Since $b_1 \in [c, d]$, at least one of the intervals (a_i, b_i), $i > 1$ contains the point b_1. By relabeling, we may assume that it is (a_2, b_2). But now we have a cover of $[c, d]$ by $n - 1$ intervals:

$$[c, d] \subset (a_1, b_2) \cup \bigcup_{i=3}^{n} (a_i, b_i).$$

So by induction

$$d - c \leq (b_2 - a_1) + \sum_{i=3}^{n} (b_i - a_i).$$

But $b_2 - a_1 \leq (b_2 - a_2) + (b_1 - a_1)$ since $a_2 < b_1$. QED

1.11 Zorn's lemma and the axiom of choice.

In the first few sections we repeatedly used an argument which involved "choosing" this or that element of a set. That we can do so is an axiom known as

The axiom of choice. *If F is a function with domain D such that $F(x)$ is a non-empty set for every $x \in D$ there exists a function f with domain D such that $f(x) \in F(x)$ for every $x \in D$.*

It has been proved by Gödel that if mathematics is consistent without the axiom of choice (a big "if") then mathematics remains consistent with the axiom of choice added.

In fact, it will be convenient for us to take a slightly less intuitive axiom as out starting point:

Zorn's lemma. *Every partially ordered set A has a maximally linearly ordered subset. If every linearly ordered subset of A has an upper bound, then A contains a maximum element.*

The second assertion is a consequence of the first. For let B be a maximum linearly ordered subset of A, and x an upper bound for B. Then x is a maximum element of A, for if $y \succ x$ then we could add y to B to obtain a larger linearly ordered set. Thus there is no element in A which is strictly larger than x which is what we mean when we say that x is a maximum element.

Zorn's lemma implies the axiom of choice.

Indeed, consider the set A of all functions g defined on subsets of D such that $g(x) \in F(x)$. We will let $\mathrm{dom}(g)$ denote the domain of definition of g. The set A is not empty, for if we pick a point $x_0 \in D$ and pick $y_0 \in F(x_0)$, then the function g whose domain consists of the single point x_0 and whose value $g(x_0) = y_0$ gives an element of A. Put a partial order on A by saying that $g \preceq h$ if $\mathrm{dom}(g) \subset \mathrm{dom}(h)$ and the restriction of h to dom g coincides with g. A linearly ordered subset means that we have an increasing family of domains X, with functions h defined consistently with respect to restriction. But this means that there is a function g defined on the union of these domains, $\bigcup X$ whose restriction to each X coincides with the corresponding h. This is clearly an upper bound. So A has a maximal element f. If the domain of f were not all of D we could add a single point x_0 not in the domain of f and $y_0 \in F(x_0)$ contradicting the maximality of f. QED

1.12 The Baire category theorem.

Theorem 1.12.1 *In a complete metric space any countable intersection of dense open sets is dense.*

Proof. Let X be the space, let B be an open ball in X, and let $O_1,\ O_2 \dots$ be a sequence of dense open sets. We must show that

$$B \cap \left(\bigcap_n O_n\right) \neq \emptyset.$$

Since O_1 is dense, $B \cap O_1 \neq \emptyset$, and is open. Thus $B \cap O_1$ contains the closure $\overline{B_1}$ of some open ball B_1. Since B_1 is open and O_2 is dense, $B_1 capO_2$ contains the closure $\overline{B_2}$ of some open ball B_2, and so on. Since X is complete, the intersection of the decreasing sequence of closed balls we have constructed contains some point x which belong both to B and to the intersection of all the O_i. QED

A **Baire space** is defined as a topological space in which every countable intersection of dense open sets is dense. Thus Baire's theorem asserts that every metric space is a Baire space. A set A in a topological space is called **nowhere dense** if its closure contains no open set. Put another way, a set A is nowhere dense if its complement A^c contains an open dense set. A set S is said to be of **firstcategory** if it is a countable union of nowhere dense sets. Then Baire's category theorem can be reformulated as saying that the complement of any set of first category in a complete metric space (or in any Baire space) is dense. A property P of points of a Baire space is said to hold **quasi- surely** or **quasi-everywhere** if it holds on an intersection of countably many dense open sets. In other words, if the set where P does not hold is of first category.

1.13 Tychonoff's theorem.

Let I be a set, serving as an "index set". Suppose that for each $\alpha \in I$ we are given a non-empty topological space S_α. The Cartesian product

$$S := \prod_{\alpha \in I} S_\alpha$$

is defined as the collection of all functions x whose domain in I and such that $x(\alpha) \in S_\alpha$. This space is not empty by the axiom of choice. We frequently write x_α instead of $x(\alpha)$ and called x_α the "α coordinate of x".

The map

$$f_\alpha : \prod_{\alpha \in I} S_\alpha \to S_\alpha, \quad x \mapsto x_\alpha$$

is called the **projection** of S onto S_α. We put on S the weakest topology such that all of these projections are continuous. So the open sets of S are generated by the sets of the form

$$f_\alpha^{-1}(U_\alpha) \text{ where } U_\alpha \subset S_\alpha \text{ is open.}$$

Theorem 1.13.1 [Tychonoff.] *If all the S_α are compact, then so is $S + \prod_{\alpha \in I} S_\alpha$.*

Proof. Let $\mathcal{F}$ be a family of closed subsets of S with the property that the intersection of any finite collection of susbsets from this family is not empty. We must show that the intersection of all the elements of $\mathcal{F}$ is not empty. Using Zorn, extend $\mathcal{F}$ to a maximal familly $\mathcal{F}_0$ of (not necessarily closed) subsets of S with the property that the intersection of any finite collection of elements of $\mathcal{F}_0$ is not empty. For each α, the projection $f_\alpha(\mathcal{F}_0)$ has the property that there is a point $x_\alpha \in S_\alpha$ which isin the closure of all the sets belonging to $f_\alpha(\mathcal{F}_0)$. Let $x \in S$ be the point whose α-th coordinate is x_α. We will show that x is in the closure of every element of $\mathcal{F}_0$ which will complete the proof.

Let U be an open set containing x. By the definition of the product topolgy, there are finitely many α_i and open subsets $U_{\alpha_i} \subset S_{\alpha_i}$ such that

$$x \in \bigcap_{i=1}^{n} f_{\alpha_i}^{-1}(U_{\alpha_i}) \subset U.$$

So for each $i = 1, \ldots, n$, $x_{\alpha_i} \in U_{\alpha_i}$. This means that U_{α_i} intersects every set belonging to $f_{\alpha_i}(\mathcal{F}_0)$. So $f_{\alpha_i}^{-1}(U_{\alpha_i})$ intersects every set belonging to $\mathcal{F}_0$ and hence must belong to $\mathcal{F}_0$ by maximality. Therefore,

$$\bigcap_{i=1}^{n} f_{\alpha_i}^{-1}(U_{\alpha_i}) \in \mathcal{F}_0,$$

again by maximality. This says that U intersects every set of $\mathcal{F}_0$. In other words, any neighborhood of x intersetions every set belonging to $\mathcal{F}_0$, which is just another way of saying x belongs to the closure of every set belonging to $\mathcal{F}_0$. QED

1.14 Urysohn's lemma.

A topological space S is called **normal** if it is Hausdorff, and if for any pair F_1, F_2 of closed sets with $F_1 \cap F_2 = \emptyset$ there are disjoint open sets U_1, U_2 with $F_1 \subset U_1$ and $F_2 \subset U_2$. For example, suppose that S is Hausdorff and compact. For each $p \in F_1$ and $q \in F_2$ there are neighborhoods O_q of p and W_q of q with $O_p \cap W_q = \emptyset$. This is the Hausdorff axiom. A finite number of the W_q cover F_2 since it is compact. Let the intersection of the corresponding O_q be called U_p and the union of the corresponding W_q be called V_p. Thus for each $p \in F_1$ we have found a neighborhood U_p of p and an open set V_p containing F_2 with

$U_p \cap V_p = \emptyset$. Once again, finitely many of the U_p cover F_1. So the union U of these and the intersection V of the corresponding V_p give disjoint open sets U containing F_1 and V containing F_2. So any compact Hausdorff space is normal.

Theorem 1.14.1 **[Urysohn's lemma.]** *If F_0 and F_1 are disjoint closed sets in a normal space S then there is a continuous real valued function $f : S \to \mathbf{R}$ such that $0 \le f \le 1$, $f = 0$ on F_0 and $f = 1$ on F_1.*

Proof. Let
$$V_1 := F_1^c.$$
We can find an open set $V_{\frac{1}{2}}$ containing F_0 and whose closure is contained in V_1, since we can choose $V_{\frac{1}{2}}$ disjoint from an open set containing F_1. So we have

$$F_0 \subset V_{\frac{1}{2}}, \quad \overline{V_{\frac{1}{2}}} \subset V_1.$$

Applying our normality assumption to the sets F_0 and $V_{\frac{1}{2}}^c$ we can find an open set $V_{\frac{1}{4}}$ with $F_0 \subset V_{\frac{1}{4}}$ and $\overline{V_{\frac{1}{4}}} \subset V_{\frac{1}{2}}$. Similarly, we can find an open set $V_{\frac{3}{4}}$ with $\overline{V_{\frac{1}{2}}} \subset V_{\frac{3}{4}}$ and $\overline{V_{\frac{3}{4}}} \subset V_1$. So we have

$$F_0 \subset V_{\frac{1}{4}}, \ \overline{V_{\frac{1}{4}}} \subset V_{\frac{1}{2}}, \ \overline{V_{\frac{1}{2}}} \subset V_{\frac{3}{4}}, \ \overline{V_{\frac{3}{4}}} \subset V_1 = F_1^c.$$

Continuing in this way, for each $0 < r < 1$ where r is a dyadic rational, $r = m/2^k$ we produce an open set V_r with $F_0 \subset V_r$ and $\overline{V_r} \subset V_s$ if $r < s$, including $\overline{V_r} \subset V_1 = F_1^c$. So $f(x) = 1$ for $x \in F_1$. Otherwise, define

$$f(x) = \inf\{r | x \in V_r\}.$$

So $f = 0$ on F_0.

If $0 < b \le 1$, then $f(x) < b$ means that $x \in V_r$ for some $r < b$. Thus

$$f^{-1}([0, b)) = \bigcap_{r<b} V_r.$$

This is a union of open sets, hence open. Similarly, $f(x) > a$ means that there is some $r > a$ such that $x \notin \overline{V_r}$. Thus

$$f^{-1}((a, 1]) = \bigcap_{r>a} (\overline{V_r})^c,$$

also a union of open sets, hence open. So we have shown that

$$f^{-1}([0, b)) \quad \text{and} \quad f^{-1}((a, 1])$$

are open. Hence $f^{-1}((a, b))$ is open. Since the intervals $[0, b), (a, 1]$ and (a, b) form a basis for the open sets on the interval $[0, 1]$, we see that the inverse image of any open set under f is open, which says that f is continuous. QED

We will have several occasions to use this result.

1.15 The Stone-Weierstrass theorem.

This is an important generalization of Weierstrass's theorem which asserted that the polynomials are dense in the space of continuous functions or any compact interval, when we use the uniform topology. We shall have many uses for this theorem.

An algebra A of (real valued) functions on a set S is said to *separate points* if for any $p, q \in S,\ p \neq q$ there is an $f \in A$ with $f(p) \neq f(q)$.

Theorem 1.15.1 [**Stone-Weierstrass.**] *Let S be a compact space and A an algebra of continuous real valued functions on S which separates points. Then the closure of A in the uniform topology is either the algebra of all continuous functions on S, or is the algebra of all continuous functions on S which all vanish at a single point, call it x_∞.*

We will give two different proofs of this important theorem. For our first proof, wee first state and prove some preliminary lemmas:

Lemma 1.15.1 *An algebra A of bounded real valued functions on a set S which is closed in the uniform topology is also closed under the lattice operations $\vee$ and $\wedge$.*

Proof. Since $f \vee g = \frac{1}{2}(f + g + |f - g|)$ and $f \wedge g = \frac{1}{2}(f + g - |f - g|)$ we must show that

$$f \in A \quad \Rightarrow |f| \in A.$$

Replacing f by $f/\|f\|_\infty$ we may assume that

$$|f| \leq 1.$$

The Taylor series for the function $t \mapsto (t + \epsilon^2)^{\frac{1}{2}}$ converges uniformly on $[0, 1]$. So there exists, for any $\epsilon > 0$ there is a polynomial P such that

$$|P(x^2) - (x^2 + \epsilon^2)^{\frac{1}{2}}| < \epsilon \quad \text{on} \quad [-1, 1].$$

Let

$$Q := P - P(0).$$

We have $|P(0) - \epsilon| < \epsilon$ so

$$|P(0)| < 2\epsilon.$$

So $Q(0) = 0$ and

$$|Q(x^2) - (x^2 + \epsilon^2)^{\frac{1}{2}}| < 3\epsilon.$$

But

$$(x^2 + \epsilon^2)^{\frac{1}{2}} - |x| \leq \epsilon$$

for small ϵ. So

$$|Q(x^2) - |x|\,| < 4\epsilon \quad \text{on} \quad [0, 1].$$

As Q does not contain a constant term, and A is an algebra, $Q(f^2) \in A$ for any $f \in A$. Since we are assuming that $|f| \leq 1$ we have

$$Q(f^2) \in A, \quad \text{and} \quad \|Q(f^2) - |f|\,\|_\infty < 4\epsilon.$$

Since we are assuming that A is closed under $\|\cdot\|_\infty$ we conclude that $|f| \in A$ completing the proof of the lemma.

Lemma 1.15.2 *Let A be a set of real valued continuous functions on a compact space S such that*

$$f, g \in A \Rightarrow \quad f \wedge g \in A \quad \text{and} \quad f \vee g \in A.$$

Then the closure of A in the uniform topology contains every continuous function on S which can be approximated at every pair of points by a function belonging to A.

Proof. Suppose that f is a continuous function on S which can be approximated at any pair of points by elements of A. So let $p, q \in S$ and $\epsilon > 0$, and let $f_{p,q,\epsilon} \in A$ be such that

$$|f(p) - f_{p,q,\epsilon}(p)| < \epsilon, \quad |f(q) - f_{p,q,\epsilon}(q)| < \epsilon.$$

Let

$$U_{p,q,\epsilon} := \{x | f_{p,q,\epsilon}(x) < f(x) + \epsilon\}, \quad V_{p,q,\epsilon} := \{x | f_{p,q,\epsilon}(x) > f(x) - \epsilon\}.$$

Fix q and ϵ. The sets $U_{p,q,\epsilon}$ cover S as p varies. Hence a finite number cover S since we are assuming that S is compact. We may take the minimum $f_{q,\epsilon}$ of the corresponding finite collection of $f_{p,q,\epsilon}$. The function $f_{q,\epsilon}$ has the property that

$$f_{q,\epsilon}(x) < f(x) + \epsilon$$

and

$$f_{q,\epsilon}(x) > f(x) - \epsilon$$

for

$$x \in \bigcap_p V_{p,q,\epsilon}$$

where the intersection is again over the same finite set of p's. We have now found a collection of functions $f_{q,\epsilon}$ such that

$$f_{q,\epsilon} < f + \epsilon$$

and $f_{q,\epsilon} > f - \epsilon$ on some neighborhood $V_{q,\epsilon}$ of q. We may choose a finite number of q so that the $V_{q,\epsilon}$ cover all of S. Taking the maximum of the corresponding $f_{q,\epsilon}$ gives a function $f_\epsilon \in A$ with $f - \epsilon < f < f + \epsilon$, i.e.

$$\|f - f_\epsilon\|_\infty < \epsilon.$$

Since we are assuming that A is closed in the uniform topology we conclude that $f \in A$, completing the proof of the lemma.

Proof of the Stone-Weierstrass theorem. Suppose first that for every $x \in S$ there is a $g \in A$ with $g(x) \neq 0$. Let $x \neq y$ and $h \in A$ with $h(y) \neq 0$. Then we may choose real numbers c and d so that $f = cg + dh$ is such that

$$0 \neq f(x) \neq f(y) \neq 0.$$

Then for any real numbers a and b we may find constants A and B such that

$$Af(x) + Bf^2(x) = a \quad \text{and } Af(y) + Bf^2(y) = b.$$

We can therefore approximate (in fact hit exactly on the nose) any function at any two distinct points. We know that the closure of A is closed under $\vee$ and $\wedge$ by the first lemma. By the second lemma we conclude that the closure of A is the algebra of all real valued continuous functions.

The second alternative is that there is a point, call it p_∞ at which all $f \in A$ vanish. We wish to show that the closure of A contains all continuous functions vanishing at p_∞. Let B be the algebra obtained from A by adding the constants. Then B satisfies the hypotheses of the Stone-Weierstrass theorem and contains functions which do not vanish at p_∞. so we can apply the preceding result. If g is a continuous function vanishing at p_∞ we may, for any $\epsilon > 0$ find an $f \in A$ and a constant c so that

$$\|g - (f + c)\|_\infty < \frac{\epsilon}{2}.$$

Evaluating at p_∞ gives $|c| < \epsilon/2$. So

$$\|g - f\|_\infty < \epsilon.$$

QED

The reason for the apparently strange notation p_∞ has to do with the notion of the one point compactification of a locally compact space. A topological space S is called **locally compact** if every point has a closed compact neighborhood. We can make S compact by adding a single point. Indeed, let p_∞ be a point not belonging to S and set

$$S_\infty := S \cup p_\infty.$$

We put a topology on S_∞ by taking as the open sets all the open sets of S together with all sets of the form

$$O \cup p_\infty$$

where O is an open set of S whose complement is compact. The space S_∞ is compact, for if we have an open cover of S_∞, at least one of the open sets in this cover must be of the second type, hence its complement is compact, hence covered by finitely many of the remaining sets. If S itself is compact, then the empty set has compact complement, hence $p\infty$ has an open neighborhood disjoint from S, and all we have done is add a disconnected point to S. The space S_∞ is called the **one-point compactification** of S. In applications of the Stone-Weierstrass theorem, we shall frequently have to do with an algebra of functions on a locally compact space consisting of functions which "vanish at infinity" in the sense that for any $\epsilon > 0$ there is a compact set C such that $|f| < \epsilon$ on the complement of C. We can think of these functions as being defined on S_∞ and all vanishing at p_∞.

We now turn to a second proof of this important theorem.

1.16 Machado's theorem.

Let $\mathcal{M}$ be a compact space and let $C_{\mathbf{R}}(\mathcal{M})$ denote the algebra of continuous real valued functions on $\mathcal{M}$. We let $\|\cdot\| = \|\cdot\|_\infty$ denote the uniform norm on $C_{\mathbf{R}}(\mathcal{M})$. More generally, for any closed set $F \subset \mathcal{M}$, we let

$$\|f\|_F = \sup_{x\in F} |f(x)|$$

so $\|\cdot\| = \|\cdot\|_{\mathcal{M}}$.

If $A \subset C_{\mathbf{R}}(\mathcal{M})$ is a collection of functions, we will say that a subset $E \subset \mathcal{M}$ is a **level set** (for A) if all the elements of A are constant on the set E. Also, for any $f \in C_{\mathbf{R}}(\mathcal{M})$ and any closed set $F \subset \mathcal{M}$, we let

$$d_f(F) := \inf_{g\in A} \|f - g\|_F.$$

So $d_f(F)$ measures how far f is from the elements of A on the set F. (I have suppressed the dependence on A in this notation.) We can look for "small" closed subsets which measure how far f is from A on all of $\mathcal{M}$; that is we look for closed sets with the property that

$$d_f(E) = d_f(\mathcal{M}). \tag{1.3}$$

Let $\mathcal{F}$ denote the collection of all non-empty closed subsets of $\mathcal{M}$ with this property. Clearly $\mathcal{M} \subset \mathcal{F}$ so this collection is not empty. We order $\mathcal{F}$ by the reverse of inclusion: $F_1 \prec F_2$ if $F_1 \supset F_2$. Let $\mathcal{C}$ be a totally ordered subset of $\mathcal{F}$. Since $\mathcal{M}$ is compact, the intersection of any nested family of non-empty closed sets is again non-empty. We claim that the intersection of all the sets in $\mathcal{C}$ belongs to $\mathcal{F}$, i.e. satisfies (1.3). Indeed, since $d_f(F) = d_F(\mathcal{M})$ for any $F \in \mathcal{C}$ this means that for any $g \in A$, the sets

$$\{x \in F \| f(x) - g(x)| \geq d_f(K)\}$$

are non-empty. They are also closed and nested, and hence have a non-empty intersection. So on the set

$$E = \bigcap_{F\in\mathcal{C}} F$$

we have

$$\|f - g\|_E \geq d_f(K).$$

So every chain has an upper bound, and hence by Zorn's lemma, there exists a maximum, i.e. there exists a non-empty closed subset E satisfying (1.3) which has the property that no proper subset of E satisfies (1.3). We shall call such a subset f-**minimal**.

Theorem 1.16.1 [**Machado.**] *Suppose that $A \subset C_{\mathbf{R}}(\mathcal{M})$ is a subalgebra which contains the constants and which is closed in the uniform topology. Then for every $f \in C_{\mathbf{R}}(\mathcal{M})$ there exists an A level set satisfying(1.3). In fact, every f-minimal set is an A level set.*

Proof. Let E be an f-minimal set. Suppose it is not an A level set. This means that there is some $h \in A$ which is not constant on A. Replacing h by $ah + c$ where a and c are constant, we may arrange that

$$\min_{x \in E} h = 0 \quad \text{and} \quad \max_{x \in E} h = 1.$$

Let

$$E_0 := \{x \in E | 0 \leq h(x) \leq]frac23\} \quad \text{and} \quad E_1 : \{x \in E | \frac{1}{3} \leq x \leq 1\}.$$

These are non-empty closed proper subsets of E, and hence the minimality of E implies that there exist $g_0, g_1 \in A$ such that

$$\|f - g_0\|_{E_0} < d_f(K) \quad \text{and} \quad \|f - g_1\|_{E_1} < d_f(K).$$

Define

$$h_n := (1 - h^n)^{2^n} \quad \text{and} \quad k_n := h_n g_0 + (1 - h_n) g_1.$$

Both h_n and k_n belong to A and $0 \leq h_n \leq 1$ on E, with strict inequality on $E_0 \cap E_1$. On this intersection we have

$$\begin{aligned} \|f - k_n\|_{E_0 \cap E_1} &= \|h_n f - h_n g_0 + (1 - h_n) f - (1 - h_n) g_1|_{E_0 \cap E_1} \\ &\leq h_n \|f - g_0|_{E_0 \cap E_1} + (1 - h_n) \|f - g_1|_{E_0 \cap E_1} \\ &\leq h_n \|f - g_)\|_{E_0} + (1 - h_n) \|f - g_1\|_{E_1} < d_f(K). \end{aligned}$$

We will now show that $h_n \to 1$ on $E_0 \setminus E_1$ and $h_n \to 0$ on $E_1 \setminus E_0$. Indeed, on $E_0 \setminus E_1$ we have

$$h^n < \left(\frac{1}{3}\right)^n$$

so

$$h_n = (1 - h^n)^{2^n} \geq 1 - 2^n h^n \geq 1 - \left(\frac{2}{3}\right)^n \to 1$$

since the binomial formula gives an alternating sum with decreasing terms. On the other hand,

$$h_n (1 + h^n)^{2^n} = 1 - h^{2 \cdot 2^n} \leq 1$$

or

$$h_n \leq \frac{1}{(1 + h^n)^{2^n}}.$$

Now the binomial formula implies that for any integer k and any positive number a we have $ka \leq (1+a)^k$ or $(1+a)^-k \leq 1/(ka)$. So we have

$$h_n \leq \frac{1}{2^n h^n}.$$

On $E_0 \setminus E_1$ we have $h^n \geq \left(\frac{2}{3}\right)^n$ so there we have

$$h_n \leq \left(\frac{3}{4}\right)^n \to 0.$$

Thus $k_n \to g_0$ uniformly on $E_0 \setminus E_1$ and $k_n \to g_1$ uniformly on $E_1 \setminus E_0$. We conclude that for n large enough

$$\|f - k_n\|_E < d_f(K)$$

contradicting our assumption that $d_f(K) = d_f(K)$. QED

Corollary 1.16.1 [The Stone-Weierstrass Theorem.] *If A is a uniformly closed subalgebra of $C_{\mathbf{R}}(\mathcal{M})$ which contains the constants and separates points, then $A = C_{\mathbf{R}}(\mathcal{M})$.*

Proof. The only A-level sets are points. But since $\|f - f(a)\|_{\{a\}} = 0$, we conclude that $d_f(\Updownarrow = 0$, i.e. $f \in A$ for any $f \in C_{\mathbf{r}}(\mathcal{M})$. QED

1.17 The Hahn-Banach theorem.

This says:

Theorem 1.17.1 [Hahn-Banach]. *Let M be a subspace of a normed linear space B, and let F be a bounded linear function on M. Then F can be extended so as to be defined on all of B without increasing its norm.*

Proof by Zorn. Suppose that we can prove

Proposition 1.17.1 *Let M be a subspace of a normed linear space B, and let F be a bounded linear function on on M. Let $y \in B, y \notin M$. Then F can be extended to $M + [y]$ without changing its norm.*

Then we could order the extensions of F by inclusion, one extension being $\succeq$ than another if it is defined on a larger space. The extension defined on the union of any family of subspaces ordered by inclusion is again an extension, and so is an upper bound. The proposition implies that a maximal extension must be defined on the whole space, otherwise we can extend it further. So we must prove the proposition.

I was careful in the statement not to specify whether our spaces are over the real or complex numbers. Let us first assume that we are dealing with a real vector space, and then deduce the complex case.

We want to choose a value
$$\alpha = F(y)$$
so that if we then define
$$F(x + \lambda y) := F(x) + \lambda F(y) = F(x) + \lambda\alpha, \quad \forall x \in M, \lambda \in \mathbf{R}$$
we do not increase the norm of F. If $F = 0$ we take $\alpha = 0$. If $F \neq 0$, we may replace F by $F/\|F\|$, extend and then multiply by $\|F\|$ so without loss of generality we may assume that $\|F\| = 1$. We want to choose the extension to have norm 1, which means that we want
$$|F(x) + \lambda\alpha| \leq \|x + \lambda y\| \quad \forall x \in M, \lambda \in \mathbf{R}.$$
If $\lambda = 0$ this is true by hypothesis. If $\lambda \neq 0$ divide this inequality by λ and replace $(1/\lambda)x$ by x. We want
$$|F(x) + \alpha| \leq \|x + y\| \quad \forall\ x \in M.$$
We can write this as two separate conditions:
$$F(x_2) + \alpha \leq |x_2 + y\| \ \ \forall\ x_2 \in M \ \text{ and } \ -F(x_1) - \alpha \leq \|x_1 + y\| \ \ \forall x_1 \in M.$$
Rewriting the second inequality this becomes
$$-F(x_1) - \|x_1 + y\| \leq \alpha \leq -F(x_2) + |x_2 + y\|.$$
The question is whether such a choice is possible In other words, is the supremum of the left hand side (over all $x_1 \in M$) less than or equal to the infimum of the right hand side (over all $x_2 \in M$)? If the answer to this question is yes, we may choose α to be any value between the sup of the left and the inf of the right hand sides of the preceding inequality. So our question is: Is
$$F(x_2) - F(x_1) \leq \|x_2 + y\| + \|x_1 + y\| \quad \forall\ x_1, x_2 \in M?$$
But $x_1 - x_2 = (x_1 + y) - (x_2 + y)$ and so using the fact that $\|F\| = 1$ and the triangle inequality gives
$$|F(x_2) - F(x_1) \leq \|x_2 - x_1\| \leq \|x_2 + y\| + \|x_1 + y\|.$$
This completes the proof of the proposition, and hence of the Hahn-Banach theorem over the real numbers.

We now deal with the complex case. If B is a complex normed vector space, then it is also a real vector space, and the real and imaginary parts of a complex linear function are real linear functions. In other words, we can write any complex linear function F as
$$F(x) = G(x) + iH(x)$$

where G and H are real linear functions. The fact that F is complex linear says that $F(ix) = iF(x)$ or

$$G(ix) = -H(x)$$

or

$$H(x) = -G(ix)$$

or

$$F(x) = G(x) - iG(ix).$$

The fact that $\|F\| = 1$ implies that $\|G\| \leq 1$. So we can adjoin the real one dimensional space spanned by y to M and extend the real linear function to it, keeping the norm ≤ 1. Next adjoin the real one dimensional space spanned by iy and extend G to it. We now have G extended to $M \oplus \mathbf{C}y$ with no increase in norm. Try to define

$$F(z) := G(z) - iG(iz)$$

on $M \oplus \mathbf{C}y$. This map of $M \oplus \mathbf{C}y \to \mathbf{C}$ is $\mathbf{R}$-linear, and coincides with F on M. We must check that it is complex linear and that its norm is ≤ 1: To check that it is complex linear it is enough to observe that

$$F(iz) = G(iz) - iG(-z) = i[G(z) - iG(iz)] = iF(z).$$

To check the norm, we may, for any z choose θ so that $e^{i\theta}F(z)$ is real and is non-negative. Then

$$|F(z)| = |e^{i\theta}F(z)| = |F(e^{i\theta}z)| = G(e^{i\theta}z) \leq \|e^{i\theta}z\| = \|z\|$$

so $\|F\| \leq 1$. QED

Suppose that M is a closed subspace of B and that $y \notin M$. Let d denote the distance of y to M, so that

$$d := \inf_{x \in M} \|y - x\|.$$

Suppose we start with the zero function on M, and extend it first to $M \oplus y$ by

$$F(\lambda y - x) = \lambda d.$$

This is a linear function on $M + \{y\}$ and its norm is ≤ 1. Indeed

$$\|F\| = \sup_{\lambda, x} \frac{|\lambda d|}{\|\lambda y - x\|} = \sup_{x' \in M} \frac{d}{y - x'} = \frac{d}{d} = 1.$$

Let M^0 be the set of all continuous linear functions on B which vanish on M. Then, using the Hahn-Banach theorem we get

Proposition 1.17.2 *If $y \in B$ and $y \notin M$ where M is a closed linear subspace of B, then there is an element $F \in M^0$ with $\|F\| \leq 1$ and $F(y) \neq 0$. In fact we can arrange that $F(y) = d$ where d is the distance from y to M.*

We have an embedding

$$B \to B^{**} \quad x \mapsto x^{**} \text{ where } x^{**}(F) := F(x).$$

The first part of the preceding proposition can be formulated as

$$(M^0)^0 = M$$

if M is a closed subspace of B.

The map $x \mapsto x^{**}$ is clearly linear and

$$|x^{**}(F)| = |F(x)| \leq \|F\|\|x\|.$$

Taking the sup of $|x^{**}(F)|/\|F\|$ shows that

$$\|x^{**}\| \leq \|x\|$$

where the norm on the left is the norm on the space B^{**}. On the other hand, if we take $M = \{0\}$ in the preceding proposition, we can find an $F \in B^*$ with $\|F\| = 1$ and $F(x) = \|x\|$. For this F we have $|x^{**}(F)| = \|x\|$. So

$$\|x^{**}\| \geq \|x\|.$$

We have proved

Theorem 1.17.2 *The map $B \to B^{**}$ given above is a norm preserving injection.*

1.18 The Uniform Boundedness Principle.

Theorem 1.18.1 *Let* **B** *be a Banach space and $\{F_n\}$ be a sequence of elements in B^* such that for every fixed $x \in$* **B** *the sequence of numbers $\{|F_n(x)|\}$ is bounded. Then the sequence of norms $\{\|F_n\|\}$ is bounded.*

Proof. The proof will be by a Baire category style argument. We will prove

Proposition 1.18.1 *There exists some ball $B = B(y,r), r > 0$ and a constant K such that $|F_n(z)| \leq K$ for all $z \in B$.*

Proof that the proposition implies the theorem. For any z with $\|z\| < r$ we have

$$|F_n(z)| \leq |F_n(z-y)| + |F_n(y)| \leq 2K.$$

So

$$\|F_n\| \leq \frac{2K}{r}$$

for all n proving the theorem from the proposition.

Proof of the proposition. If the proposition is false, we can find n_1 such that $|F_{n_1}(x)| > 1$ at some $x \in B(0,1)$ and hence in some ball of radius $\epsilon < \frac{1}{2}$ about x.

Then we can find an n_2 with $|F_{n_2}(z)| > 2$ in some non-empty closed ball of radius $< \frac{1}{3}$ lying inside the first ball. Continuing inductively, we choose a subsequence n_m and a family of nested non-empty balls B_m with $|F_{n_m}(z)| > m$ throughout B_m and the radii of the balls tending to zero. Since B is complete, there is a point x common to all these balls, and $\{|F_n(x)|\}$ is unbounded, contrary to hypothesis. QED

We will have occasion to use this theorem in a "reversed form". Recall that we have the norm preserving injection $B \to B^{**}$ sending $x \mapsto x^{**}$ where $x^{**}(F) = F(x)$. Since B^* is a Banach space (even if B is incomplete) we have

Corollary 1.18.1 *If $\{x_n\}$ is a sequence of elements in a normed linear space such that the numerical sequence $\{|F(x_n)|\}$ is bounded for each fixed $F \in B^*$ then the sequence of norms $\{\|x_n\|\}$ is bounded.*