Notes for the course *Operatorer i Hilbertrum*

Horia Cornean[1], d. 09/12/2013.

Disclaimer: all theorems presented here are nothing more than very detailed versions of some fundamental results which can be found in a number of excellent books by Walter Rudin (*Principles of mathematical analysis*, *Real and complex analysis* and *Functional analysis*) and Peter Lax (*Functional analysis*).

1 The natural topology of a metric space

Let (X, d) be a metric space. We define the open ball of radius $r > 0$ and center at $a \in X$ the set $B_r(a) := \{x \in X : \; d(x, a) < r\}$.

Given a set $A \subset X$ and $a \in A$, we say that a is an interior point of A if there exists $r > 0$ such that $B_r(a) \subset A$. The set of all interior points of A is denoted by $\text{Int}(A)$. We say that A is an open set if all its points are interior points, i.e. $\text{Int}(A) = A$. By convention, the empty set $\emptyset$ is open.

Lemma 1.1. *Any ball $B_r(a)$ is an open set.*

Proof. Let $x_0 \in B_r(a)$. We have that $d(x_0, a) < r$. Define $r_0 := (r - d(x_0, a))/2 > 0$. Then for all $x \in B_{r_0}(x_0)$ we have that $d(x, x_0) < r_0$ and:

$$d(x, a) \leq d(x, x_0) + d(x_0, a) < (r - d(x_0, a))/2 + d(x_0, a) = (r + d(x_0, a))/2 < r,$$

which shows that $B_{r_0}(x_0) \subset B_r(a)$. Thus $B_r(a)$ has only interior points. $\square$

Lemma 1.2.
(i). Let $\{V_\alpha\}_{\alpha \in \mathcal{F}}$ be an arbitrary collection of open sets. Then $A := \cup_\alpha V_\alpha$ is open.
(ii). Let $\{V_j\}_{j=1}^n$ be a finite collection of open sets. Then $B := \cap_{j=1}^n V_j$ is open.

Proof. We start with (i). Let $a \in \cup_\alpha V_\alpha$. There must exist $\alpha_a \in \mathcal{F}$ such that $a \in V_{\alpha_a}$. Since V_{α_a} is open, there exists $r_a > 0$ such that

$$B_{r_a}(a) \subset V_{\alpha_a} \subset \cup_\alpha V_\alpha = A$$

hence a is an interior point of A.

We continue with (ii). Let $a \in \cap_{j=1}^n V_j$. Thus $a \in V_j$ for all j. Hence there exists $r_j > 0$ such that $B_{r_j}(a) \subset V_j$. Let $r := \min\{r_1, \dots r_n\} > 0$. Thus $B_r(a) \subset B_{r_j}(a) \subset V_j$ for all j, hence $B_r(a) \subset B$ and we are done. $\square$

We say that a set $A \subset X$ is closed if $A^c := \{x \in X : \; x \notin A\}$ is open. Given a set $B \subset X$ and $b \in X$, we say that b is an adherent point of B if there exists a sequence $\{x_n\}_{n \geq 1} \subset B$ such that $x_n \in B_{\frac{1}{n}}(b)$ and $\lim_{n \to \infty} x_n = b$. The set of all adherent points of B is denoted by $\overline{B}$.

Theorem 1.3. *Let $B \subset X$. Then $B \subset \overline{B}$. Moreover, $B = \overline{B}$ if and only if B is closed.*

Proof. If $a \in B$ we can define the constant sequence $x_n = a \in B$ which converges to a, thus $a \in \overline{B}$ and $B \subset \overline{B}$.

Now assume that $B = \overline{B}$. We want to prove that B is closed, i.e. B^c is open. Let $a \in B^c = \overline{B}^c$. Then a is not an adherent point, which means that there exists $\epsilon > 0$ such that no point of B lies in the ball $B_\epsilon(a)$. In other words, $B_\epsilon(a) \subset B^c$, hence B^c is open.

[1]Department of Mathematical Sciences, Aalborg University

Now assume that B is closed. We want to prove that $B = \overline{B}$. Assume that this is not true; it would imply the existence of a point $b \in \overline{B}$ such that $b \in B^c$. Since B^c is open, there exists $\epsilon > 0$ such that $B_\epsilon(b) \subset B^c$, i.e. $B_\epsilon(b) \cap B = \emptyset$. But this is incompatible with $b \in \overline{B}$.

□

2 Compact and sequentially compact sets

Definition 2.1. *Let A be a subset of a metric space (X, d). Let $\mathcal{F}$ be an arbitrary set of indices, and consider the family of sets $\{\mathcal{O}_\alpha\}_{\alpha\in\mathcal{F}}$, where each $\mathcal{O}_\alpha \subseteq X$ is open. This family is called an open covering of A if $A \subseteq \bigcup_{\alpha\in\mathcal{F}} \mathcal{O}_\alpha$.*

Definition 2.2. *Assume that $\{\mathcal{O}_\alpha\}_{\alpha\in\mathcal{F}}$ is an open covering of A. If $\mathcal{F}'$ is a subset of $\mathcal{F}$, we say that $\{\mathcal{O}_\alpha\}_{\alpha\in\mathcal{F}'}$ is a subcovering if we still have the property $A \subseteq \bigcup_{\alpha\in\mathcal{F}'} \mathcal{O}_\alpha$. A subcovering is called finite, if $\mathcal{F}'$ contains finitely many elements.*

Definition 2.3. *Let A be a subset of a metric space (X, d). Then we say that A is covered by a finite ϵ-net if there exists a natural number $N_\epsilon < \infty$ and the points $\{\mathbf{x}_1, ..., \mathbf{x}_{N_\epsilon}\} \subseteq A$ such that $A \subseteq \bigcup_{j=1}^{N_\epsilon} B_\epsilon(\mathbf{x}_j)$.*

Definition 2.4. *A subset $A \subset X$ is called compact, if from* any *open covering of A one can extract a* finite *subcovering.*

Definition 2.5. *$A \subset X$ is called sequentially compact if from any sequence $\{x_n\}_{n\geq 1} \subseteq A$ one can extract a subsequence $\{x_{n_k}\}_{k\geq 1}$ which converges to some point $x_\infty \in A$.*

We will see that in metric spaces the two notions of compactness are equivalent.

2.1 Compact implies sequentially compact

We begin with two lemmas:

Lemma 2.6. *Assume that the sequence $\{x_n\}_{n\geq 1} \subset A$ has a range consisting of finitely many points. Then it admits a convergent subsequence whose limit is one of the elements in the range.*

Proof. Assume that the range of the sequence consists of the distinct points $a_1, a_2, \dots, a_N$. At least one of these points, say a_1, is taken infinitely many times by the sequence elements. Denote by n_k (with $k \geq 1$) the increasing sequence of indices for which $x_{n_k} = a_1$. This defines our convergent subsequence. □

Lemma 2.7. *Assume that the sequence $\{x_n\}_{n\geq 1} \subset A$ has an accumulation point $a \in A$, i.e. for every $\epsilon > 0$ there exists some $x_n \neq a$ such that $x_n \in B_\epsilon(a)$. Then $\{x_n\}_{n\geq 1}$ admits a convergent subsequence whose limit is a.*

Proof. Since a is an accumulation point, there exists an index $j \geq 1$ such that $x_j \neq a$ and $x_j \in B_1(a)$. Denote by n_1 the smallest index for which these two properties hold true. Let $r_1 := d(x_{n_1}, a) > 0$. Define n_2 to be the smallest index j for which $x_j \neq a$ and $x_j \in B_{\min\{r_1, \frac{1}{2}\}}(a)$. We must have $n_2 \geq n_1$ since $x_{n_2} \in B_1(a)$; moreover, because $r_2 := d(x_{n_2}, a) < r_1$, we cannot have $n_1 = n_2$. In general, if $k \geq 2$ we define n_k to be the smallest index j for which $x_j \neq a$ and $x_j \in B_{\min\{r_{k-1}, \frac{1}{k}\}}(a)$; moreover, since $r_k := d(x_{n_k}, a) < r_{k-1} < \cdots < r_1$, we must have $n_k > \cdots > n_1$. Then $\{n_k\}_{k\geq 1}$ is a strictly increasing sequence and $0 < d(x_{n_k}, a) < 1/k$. This shows that $\{x_{n_k}\}_{k\geq 1}$ is a subsequence which converges to a. □

Theorem 2.8. *Let $A \subseteq X$ be compact. Then A is sequentially compact.*

Proof. We will assume the opposite, i.e. there exists a sequence $\{x_n\}_{n\geq 1}$ with no convergent subsequence in A. Such a sequence must have an infinite number of distinct points in the range, due to Lemma 2.6. Moreover, we can assume that $\{x_n\}_{n\geq 1}$ has no accumulation points in A (otherwise such a point would be the limit of a subsequence according to Lemma 2.7).

Since no $x \in A$ can be an accumulation point for $\{x_n\}_{n\geq 1}$, there exists $\epsilon_x > 0$ such that the ball $B_{\epsilon_x}(x)$ contains at most one element of the range of $\{x_n\}_{n\geq 1}$.

Clearly, $\{B_{\epsilon_x}(x)\}_{x\in A}$ is an open covering for A. Because A is compact, we can extract a finite subcovering from it:

$$A \subseteq \bigcup_{j=1}^{N} B_{\epsilon_{y_j}}(y_j), \quad N < \infty, \quad \{y_1, \dots, y_N\} \subset A.$$

Now remember that $\{x_n\}_{n\geq 1} \subseteq A \subseteq \bigcup_{j=1}^{N} B_{\epsilon_{y_j}}(y_j)$ and at the same time, there are at most N distinct points of the range of $\{x_n\}_{n\geq 1}$ in the union $\bigcup_{j=1}^{N} B_{\epsilon_{y_j}}(y_j)$. We conclude that $\{x_n\}_{n\geq 1}$ can only have a finite number of distinct points in its range, thus it must admit a convergent subsequence according to Lemma 2.6. This contradicts our hypothesis. □

2.2 Sequentially compact implies compact

The proof of this fact is slightly more complicated. We need two preparatory results:

Proposition 2.9. *Let A be a sequentially compact set. Then for every $\epsilon > 0$, A can be covered by a finite ϵ-net (see Definition 2.3).*

Proof. If A contains finitely many points, then the proof is obvious, thus we may assume that $\#(A) = \infty$.

Now suppose that there exists some $\epsilon_0 > 0$ such that A cannot be covered by a finite ϵ_0-net. This means that for any N points of A, $\{x_1, ..., x_N\}$, we have:

$$A \not\subset \bigcup_{j=1}^{N} B_{\epsilon_0}(x_j). \tag{2.1}$$

We will now construct a sequence with elements in A which cannot have a convergent subsequence. Choose an arbitrary point $x_1 \in A$. We know from (2.1), for $N = 1$, that we can find $x_2 \in A$ such that $x_2 \in A \setminus B_{\epsilon_0}(x_1)$. This means that $d(x_1, x_2) \geq \epsilon_0$. We use (2.1) again, for $N = 2$, in order to get a point $x_3 \in A \setminus [B_{\epsilon_0}(x_1) \cup B_{\epsilon_0}(x_2)]$. This means that $d(x_3, x_1) \geq \epsilon_0$ and $d(x_3, x_2) \geq \epsilon_0$. Thus we can continue with this procedure and construct a sequence $\{x_n\}_{n\geq 1} \subseteq A$ which obeys

$$d(x_j, x_k) \geq \epsilon_0, \quad j \neq k.$$

In other words, we constructed a sequence in A which cannot have a Cauchy subsequence. This contradicts Definition 2.5. □

The second result states that a compact set is bounded:

Lemma 2.10. *Let A be a (sequentially) compact set. Then there exists a ball which contains A.*

Proof. We know that A can be covered by any finite ϵ-net; choose $\epsilon = 1$. Then here exist N points of A denoted by $\{x_1, ..., x_N\}$ such that $A \subset \bigcup_{j=1}^{N} B_1(x_j)$.

Denote by $R = \max\{1 + d(x_j, x_k) : \ 1 \leq j, k \leq N\}$. Then we have $B_1(x_j) \subset B_R(x_1)$ for every j, thus $A \subset B_R(x_1)$ and we are done. □

Let us now prove the theorem:

Theorem 2.11. *Assume that $A \subseteq X$ is sequentially compact. Then A is compact.*

Proof. Consider an arbitrary open covering of A:

$$A \subseteq \bigcup_{\alpha \in \mathcal{F}} \mathcal{O}_\alpha.$$

We will show that we can extract a finite subcovering from it.

For every $x \in A$, there exists at least one open set $\mathcal{O}_{\alpha(x)}$ such that $x \in \mathcal{O}_{\alpha(x)}$. Because $\mathcal{O}_{\alpha(x)}$ is open, we can find $\epsilon > 0$ such that $B_\epsilon(x) \subseteq \mathcal{O}_{\alpha(x)}$.

For a fixed x, we define the set

$$E_x := \{r > 0 : \text{ there exists } \alpha \in \mathcal{F} \text{ such that } B_r(x) \subseteq \mathcal{O}_\alpha\} \subset \mathbb{R}.$$

From the above argument we conclude that no E_x is empty. Moreover, if $r \in E_x$, then the open interval $(0, r)$ is included in E_x.

If for some x in A we have an unbounded E_x, it follows that for every $r > 0$ we can find some open set $\mathcal{O}_\alpha$ such that $B_r(x) \subseteq \mathcal{O}_\alpha$. But if r is chosen to be large enough, it will contain the ball we constructed in Lemma 2.10, thus $\mathcal{O}_\alpha$ will also contain A. In this case we found our subcovering, which consists of just one open set.

It follows that we may assume that all the sets E_x are bounded intervals admitting a positive and finite supremum $\sup E_x$. Define $0 < \epsilon_x := \frac{1}{2} \sup E_x < \sup E_x$. Note the important thing that $\epsilon_x \in E_x$. Let us also observe that:

$$A \subseteq \bigcup_{x \in A} B_{\epsilon_x}(x) \subseteq \bigcup_{\alpha \in \mathcal{F}} \mathcal{O}_\alpha. \tag{2.2}$$

The first inclusion is obvious, while the second one follows from the above discussion.

We now need to prove a lemma:

Lemma 2.12. *If A is sequentially compact, then*

$$\inf_{x \in A} \epsilon_x =: 2\epsilon_0 > 0.$$

In other words, there exists $\epsilon_0 > 0$ such that $B_{\epsilon_0}(x) \subseteq B_{\epsilon_x}(x)$, for every $x \in A$.

Proof. Assume that $\inf_{x \in A} \epsilon_x = 0$. This implies that there exists a sequence $\{x_n\}_{n \geq 1} \subseteq A$ such that $\epsilon_{x_n} \leq 1/n$ for every $n \geq 1$. Since A is sequentially compact, there exists a convergent subsequence $\{x_{n_k}\}_{k \geq 1}$ which converges to a point $x_0 \in A$, i.e.

$$\lim_{k \to \infty} x_{n_k} = x_0. \tag{2.3}$$

Because x_0 belongs to A, we can find an open set $\mathcal{O}_{\alpha(x_0)}$ which contains x_0, thus we can find $\epsilon_1 > 0$ such that

$$B_{\epsilon_1}(x_0) \subseteq \mathcal{O}_{\alpha(x_0)}. \tag{2.4}$$

Now (2.3) implies that there exists $K > 0$ large enough such that:

$$d(x_{n_k}, x_0) \leq \epsilon_1/4, \quad \text{whenever} \quad k > K. \tag{2.5}$$

If y belongs to $B_{\epsilon_1/4}(x_{n_k})$ (i.e. $d(y, x_{n_k}) < \epsilon_1/4$), then the triangle inequality implies (use also (2.5)):

$$d(y, x_0) \leq d(y, x_{n_k}) + d(x_{n_k}, x_0) < \epsilon_1/2 < \epsilon_1, \quad k > K.$$

But this shows that we must have $y \in B_{\epsilon_1}(x_0)$, or:

$$B_{\epsilon_1/4}(x_{n_k}) \subseteq B_{\epsilon_1}(x_0) \subseteq \mathcal{O}_{\alpha(x_0)}, \quad \forall k > K. \tag{2.6}$$

Thus we got the inclusion

$$B_{\epsilon_1/4}(x_{n_k}) \subseteq \mathcal{O}_{\alpha(x_0)}, \quad \forall k > K,$$

which shows that $\epsilon_1/4$ must be less or equal than $2\epsilon_{x_{n_k}}$, or $\epsilon_1/8 \leq \epsilon_{x_{n_k}}$, for every $k > K$. But this is in contradiction with the fact that $\epsilon_{x_n} \leq 1/n$ for every $n \geq 1$. □

Finishing the proof of Theorem 2.11. We now use Proposition 2.9, and find a finite ϵ_0-net for A. Thus we can choose $\{y_1, ...y_N\} \subseteq A$ such that

$$A \subseteq \bigcup_{n=1}^{N} B_{\epsilon_0}(y_n) \subseteq \bigcup_{n=1}^{N} B_{\epsilon_{y_n}}(y_n) \subseteq \bigcup_{n=1}^{N} \mathcal{O}_n,$$

where $\mathcal{O}_n$ is one of the possibly many open sets which contain $B_{\epsilon_{y_n}}(y_n)$. We have thus extracted our finite subcovering of A and the proof of the theorem is over. □

2.3 The Bolzano-Weierstrass Theorem

We start with the case in which the metric space is $\mathbb{R}$ with the Euclidean distance.

Theorem 2.13. *Let $\{x_n\} \subset \mathbb{R}$ be a bounded real sequence, i.e. there exists $M \geq 0$ such that $|x_n| \leq M$ for all $n \geq 1$. Then there exists a subsequence $\{x_{n_k}\}_{k\geq 1}$ and some $s \in \mathbb{R}$ such that $\lim_{k\to\infty} x_{n_k} = s$.*

Proof. We have that $-M \leq x_n \leq M$ for all n. Define by $a_1 := -M$ and $b_1 := M$. Since either $-M \leq x_n \leq 0$ or $0 \leq x_n \leq M$ for any given n, it follows that at least one of the two intervals $[-M, 0]$ and $[0, M]$ must contain x_n for infinitely many different values of n. If there are infinitely many indices such that $x_n \in [-M, 0]$, then define $a_2 := a_1$ and $b_2 := (a_1 + b_1)/2$. If this is not true, then define $a_2 := (a_1 + b_1)/2$ and $b_2 := b_1$. If the first case holds true, we define n_1 to be the smallest index n for which $-M = a_2 \leq x_n \leq b_2 = 0$, while if the second case is true, we define n_1 to be the smallest index n for which $0 = a_2 \leq x_n \leq b_2 = M$.

In either case, we know that there exist infinitely many indices n such that $a_2 \leq x_n \leq b_2$, and n_1 is the smallest of them. If the interval $[a_2, (a_2 + b_2)/2]$ contains x_n for infinitely many values of n, then define $a_3 := a_2$ and $b_3 := (a_2 + b_2)/2$. If this is not true, then define $a_3 := (a_2 + b_2)/2$ and $b_3 := b_2$; the interval $[a_3, b_3]$ will thus contain x_n infinitely many times. We can thus choose n_2 to be the smallest index $n > n_1$ for which $a_3 \leq x_n \leq b_3$. By induction, for a given $k \geq 1$, we can construct $n_k > n_{k-1} > \cdots > n_1$ such that $a_{k+1} \leq x_{n_k} \leq b_{k+1}$, where either $a_{k+1} := a_k$ and $b_{k+1} := (a_k + b_k)/2$ (if the interval $[a_k, (a_k + b_k)/2]$ contains x_n infinitely many times), or $a_{k+1} := (a_k + b_k)/2$ and $b_{k+1} := b_k$ otherwise. By construction we have that $a_k \leq a_{k+1}$ and $b_{k+1} \leq b_k$ for all k. Moreover, $a_k \leq b_k$ for all k, and in particular $a_k \leq b_1 = M$ and $a_1 = -M \leq b_k$. By induction, we can also prove that $b_k - a_k = (b_1 - a_1)/2^{k-1}$.

Thus $\{a_k\}_{k\geq 1}$ is increasing and bounded from above, hence it converges to $\alpha := \sup_{k\geq 1} a_k$. The sequence $\{b_k\}_{k\geq 1}$ is decreasing and bounded from below, thus it converges to $\beta := \inf_{k\geq 1} b_k$. By taking the limit $k \to \infty$ in the equality $b_k - a_k = (b_1 - a_1)/2^{k-1}$ we conclude that $\alpha = \beta$. Since $a_k \leq x_{n_k} \leq b_k$, by the comparison theorem it follows that $\{x_{n_k}\}_{k\geq 1}$ is convergent and has the limit $s := \alpha = \beta$.

□

We can generalize this result to $\mathbb{R}^d$, with $d \geq 2$. Without loss of generality, assume that $d = 2$; the general case follows by induction. If $\mathbf{x} = [u, v] \in \mathbb{R}^2$, then we define $||\mathbf{x}|| = \sqrt{u^2 + v^2}$. Clearly, $\max\{|u|, |v|\} \leq ||\mathbf{x}|| \leq |u| + |v|$. The Euclidean distance between two vectors $\mathbf{x} = [u_1, v_1]$ and $\mathbf{y} = [u_2, v_2]$ is given by $d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}|| = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}$. It is easy to check that $d(\mathbf{x}, \mathbf{y}) \leq |u_1 - u_2| + |v_1 - v_2|$.

Now assume that the sequence $\{\mathbf{x}_n\}_{n\geq 1} \subset \mathbb{R}^2$ is bounded, i.e. there exists $M \geq 0$ such that $||\mathbf{x}_n|| \leq M$ for all n. We denote the components of $\mathbf{x}_n$ with $[u_n, v_n]$. The real sequence $\{u_n\}_{n\geq 1} \subset \mathbb{R}$ is also bounded by M, thus from Theorem 2.13 it follows that we can find a subsequence $\{u_{n_k}\}_{k\geq 1}$ which is convergent to some $t \in \mathbb{R}$, i.e. $\lim_{k\to\infty} u_{n_k} = t$. Define $z_k := v_{n_k}$; then $\{z_k\}_{k\geq 1}$ is also bounded by M and according to Theorem 2.13 we can find a subsequence $\{z_{k_j}\}_{j\geq 1}$ which is convergent to some $s \in \mathbb{R}$, i.e. $\lim_{j\to\infty} z_{k_j} = s$. Thus we have that $v_{n_{k_j}}$ converges to s while $u_{n_{k_j}}$ still converges to t, as a subsequence of the convergent sequence $\{u_{n_k}\}_{k\geq 1}$.

Define $\mathbf{y} := [t, s]$. We have $0 \leq d(\mathbf{x}_{n_{k_j}}, \mathbf{y}) \leq |u_{n_{k_j}} - t| + |v_{n_{k_j}} - s|$ for all $j \geq 1$, which shows that $\mathbf{y}$ is the limit of $\{\mathbf{x}_{n_{k_j}}\}_{j\geq 1}$.

2.4 The Heine-Borel Theorem

Lemma 2.14. *Let A be a compact set in a metric space (X, d). Then A is bounded and closed.*

Proof. We already know that a compact set A is bounded (see Lemma 2.10). Let us prove that it is closed. Assume it is not. According to Theorem 1.3 it means that there exists an adherent point $a \in \overline{A}$ which does not belong to A. Being an adherent point, there exists a sequence $\{x_n\}_{n\geq 1} \subset A$ which converges to a, thus all of its subsequences must converge to the same limit. Since A is (sequentially) compact, there exists a subsequence $\{x_{n_k}\}_{k\geq 1}$ which converges to some point of A, which has to be a. This contradicts the fact that $a \notin A$.

□

Theorem 2.15. *Consider $\mathbb{R}^d$ with the Euclidean distance. In this metric space, a set A is (sequentially) compact if and only if A is both bounded and closed.*

Proof. The previous lemma showed that a compact set is always bounded and closed; this fact holds for all metric spaces, not just for the Euclidean ones.

If the space is Euclidean, then we can also show the reversed implication. Assume that A is bounded and consider an arbitrary sequence $\{x_n\}_{n\geq 1} \subset A$. The Bolzano-Weierstrass theorem implies the existence of a subsequence $\{x_{n_k}\}_{k\geq 1}$ which converges to some point $a \in \mathbb{R}^d$. Thus $a \in \overline{A}$, and due to Theorem 1.3 we know that $A = \overline{A}$, thus $a \in A$. This proves that A is sequentially compact, therefore compact.

□

3 Continuous functions on metric spaces

Let (X, d) and (Y, ρ) be two metric spaces. If $A \subset X$, the image of A through f is the set

$$f(A) := \{y \in Y : \text{ there exists } x_y \in A \text{ such that } f(x_y) = y\} \subset Y.$$

If $B \subset Y$ the preimage of B through f is the set

$$f^{-1}(B) := \{x \in X : \text{ such that } f(x) \in B\} \subset X.$$

Note that the notation $f^{-1}(B)$ does not imply that f is invertible.

Lemma 3.1. *If $A_1 \subset A_2 \subset X$ and $B_1 \subset B_2 \subset Y$ then $f(A_1) \subset f(A_2)$ and $f^{-1}(B_1) \subset f^{-1}(B_2)$.*

Proof. We only prove the first inclusion. Assume that $y \in f(A_1)$. Then there exists $x_y \in A_1$ such that $f(x_y) = y$. But at the same time $x_y \in A_2$, hence $y \in f(A_2)$. □

A map $f : X \to Y$ is said to be continuous at a point $a \in X$ if for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$B_\delta(a) \subset f^{-1}(B_\epsilon(f(a))), \tag{3.1}$$

which implies that $f(B_\delta(a)) \subset B_\epsilon(f(a))$. The function is continuous on X if it is continuous at all the points of X.

Theorem 3.2. *A function between two metric spaces $f : X \to Y$ is continuous on X if and only if for every nonempty open set $V \subset Y$ we have that $f^{-1}(V)$ is open in X.*

Proof. First we assume that f is continuous on X. Let V a nonempty open set in Y. If $f^{-1}(V)$ is empty then we know that it is open. Otherwise, let $a \in f^{-1}(V)$. Thus $f(a) \in V$. Since V is open, $f(a)$ is an interior point of V, thus there exists $\epsilon > 0$ such that $B_\epsilon(f(a)) \subset V$. Applying Lemma 3.1 we get that $f^{-1}(B_\epsilon(f(a))) \subset f^{-1}(V)$. But from (3.1) it follows that $B_\delta(a) \subset f^{-1}(V)$, thus a is an interior point.

We now assume that f returns any nonempty open set V of Y in an open set $f^{-1}(V)$ of X. Fix $a \in X$. Let $\epsilon > 0$ and consider the ball $B_\epsilon(f(a))$. Lemma 1.1 implies that $V = B_\epsilon(f(a))$ is open in Y. Thus $f^{-1}(B_\epsilon(f(a)))$ must be open in X. Since $a \in f^{-1}(B_\epsilon(f(a)))$, it must be an interior point. Thus there exists $\delta > 0$ such that $B_\delta(a) \subset f^{-1}(B_\epsilon(f(a)))$, which shows that f is continuous at a.

□

Let (X, d) and (Y, ρ) be two metric spaces and consider a subset $A \subset X$. We can organize A as a metric space with the natural distance d_A induced by d. We say that the map $f : A \mapsto Y$ is continuous on A if it is continuous between the metric spaces (A, d_A) and (Y, ρ).

We say that $f : A \mapsto Y$ is sequentially continuous at a point $a \in A$ if for every sequence $\{x_n\}_{n\geq 1} \subset A$ which converges to a we have that $\{f(x_n)\}_{n\geq 1} \subset Y$ converges to $f(a)$. We say that $f : A \mapsto Y$ is sequentially continuous on A if it is sequentially continuous at all points of A.

Theorem 3.3. *With the above notation, consider a map $f : A \mapsto Y$. Then f is continuous on A if and only if it is sequentially continuous on A.*

Proof. First, assume that f is continuous at $a \in A$. Consider any sequence $\{x_n\}_{n\geq 1} \subset A$ which converges to a. From (3.1) we know that for every $\epsilon > 0$ we have that $\rho(f(x_n), f(a)) < \epsilon$ if $d(x_n, a) < \delta$. But the second inequality holds if n is larger than some $N_\delta \geq 1$. Thus $\{f(x_n)\}_{n\geq 1} \subset Y$ converges to $f(a)$.

Second, assume that f is sequentially continuous at $a \in A$. We will show that f must be continuous at a. Suppose this is not true: it means that there exists $\epsilon_0 > 0$ such that for all $\delta > 0$ we have that $B_\delta(a) \not\subset f^{-1}(B_{\epsilon_0}(f(a)))$. By letting $\delta = 1/n$ for all $n \geq 1$, we can find a point $x_n \in B_{\frac{1}{n}}(a)$ such that $f(x_n) \notin B_{\epsilon_0}(f(a))$, or $\rho(f(x_n), f(a)) \geq \epsilon_0$. In this way we constructed a sequence $\{x_n\}_{n\geq 1} \subset A$ which converges to a while $\{f(x_n)\}_{n\geq 1}$ does not converge to $f(a)$, contradiction.

□

Theorem 3.4. *With the above notation, consider a continuous map $f : A \mapsto Y$ where $A \subset X$ is compact. Then $f(A)$ is compact.*

Proof. We show that $f(A)$ is sequentially compact. Let $\{y_n\}_{n\geq 1} \subset f(A)$ be an arbitrary sequence. There exists $\{x_n\}_{n\geq 1} \subset A$ such that $f(x_n) = y_n$. Since A is sequentially compact, there exists a subsequence $\{x_{n_k}\}_{k\geq 1} \subset \{x_n\}_{n\geq 1}$ which converges to some point $a \in A$. But f is sequentially continuous at a, hence $y_{n_k} = f(x_{n_k})$ converges to $f(a) \in f(A)$. Hence $f(A)$ is sequentially compact.

□

The next lemma recalls a general result which says that real continuous functions defined on compact sets attain their extremal values. See also Theorem 10.61 in Wade.

Lemma 3.5. *Let (X, d) be a metric space and let $H \subset X$ be a compact set. Let $f : H \mapsto \mathbb{R}$ be continuous on H. Then there exist x_m and x_M in H such that $f(x_M) = \sup_{x\in H} f(x)$ and $f(x_m) = \inf_{x\in H} f(x)$.*

Proof. We only prove this for $\sup_{x\in H} f(x)$. Let $B := f(H) \subset \mathbb{R}$. Let us show that there exists a sequence $\{x_n\}_{n\geq 1} \subset H$ such that $\lim_{n\to\infty} f(x_n) = \sup_{x\in H} f(x) = \sup(B)$.

Since B is compact, it is bounded. Thus $\sup(B) = \sup_{x\in H} f(x) < \infty$. For every $n \geq 1$ we know that $\sup(B) - 1/n$ is not an upper bound for B, thus there must exist $x_n \in H$ such that $\sup(B) - 1/n < f(x_n) \leq \sup(B)$. Thus $\lim_{n\to\infty} f(x_n) = \sup(B)$.

Because H is compact, we can find a subsequence $\{x_{n_k}\}_{k\geq 1}$ which converges towards some point $a \in H$. Since f is continuous, we have that $\lim_{k\to\infty} f(x_{n_k}) = f(a)$. Since $\{f(x_{n_k})\}_{k\geq 1}$ is a subsequence of the convergent sequence $\{f(x_n)\}_{n\geq 1}$, we must have $f(a) = \sup(B)$. Thus we can choose x_M to be a.

□

We say that $f : A \mapsto Y$ is uniformly continuous on A if for every $\epsilon > 0$ there exists $\delta > 0$ such that $\rho(f(x), f(y)) < \epsilon$ as soon as $x, y \in A$ and $d(x, y) < \delta$. Clearly, if f is uniformly continuous on A then it is also continuous. The next result gives sufficient conditions for the reciprocal statement:

Lemma 3.6. *Let (X, d) and (Y, ρ) be two metric spaces and let $H \subset X$ be a compact set. Let $f : H \mapsto Y$ be continuous on H. Then f is uniformly continuous on H.*

Proof. Assume that the conclusion is false. Then there exists $\epsilon_0 > 0$ such that regardless how large $n \geq 1$ is, we may find two points x_n and y_n in H which obey $d(x_n, y_n) < \frac{1}{n}$ and $\rho(f(x_n), f(y_n)) \geq \epsilon_0$. Since H is sequentially compact, there exists a subsequence $\{x_{n_k}\}_{k\geq 1}$ which converges to some point $a \in H$. Because $d(y_{n_k}, a) \leq \frac{1}{k} + d(x_{n_k}, a)$ for all $k \geq 1$, it follows that y_{n_k} also converges to a. The function f is sequentially compact at a, thus both $f(x_{n_k})$ and $f(y_{n_k})$ converge to $f(a)$. In particular, this contradicts our assumption that $\rho(f(x_{n_k}), f(y_{n_k})) \geq \epsilon_0$ for all k.

□

4 Elementary considerations about measurability

Definition 4.1. *Consider a set X. We call the collection of sets $\mathfrak{S}$ to be a σ-algebra if the following three conditions are fulfilled:*

1. *$X \in \mathfrak{S}$;*
2. *If $A \in \mathfrak{S}$ then $A^c = X \setminus A \in \mathfrak{S}$;*
3. *If each $A_j \in \mathfrak{S}$ for $j \geq 1$, then $A := \cup_{j=1}^{\infty} A_j \in \mathfrak{S}$.*

Note that from the definition of $\mathfrak{S}$ we conclude that $\emptyset = X^c \in \mathfrak{S}$, and moreover, since

$$\Big(\bigcup_{j=1}^{\infty} B_j^c\Big)^c = \bigcap_{j=1}^{\infty} B_j$$

it follows that any countable intersection of elements of $\mathfrak{S}$ is also an element of the σ-algebra. Also, since $A \setminus B = A \cap B^c$, it follows that set differences belong to $\mathfrak{S}$, too.

Proposition 4.2. *Let τ denote the set of all open sets in X (called the topology). Then there exists a σ-algebra $\mathfrak{S}_x$ which contains τ, and $\mathfrak{S}_x$ is also contained in any other σ-algebra which contains τ. We call $\mathfrak{S}_x$ as the Borel σ-algebra, and its elements are called Borel sets.*

Proof. Denote by $\mathcal{P}(X)$ the set of all subsets of X. This is the maximal σ-algebra, which clearly contains the topology τ, hence the set of all σ-algebras S_τ containing τ is not empty. Define $\mathfrak{S}_x := \bigcap S_\tau$ to be the intersection of all these σ-algebras. We have that $\tau \subset \mathfrak{S}_x$ and $\mathfrak{S}_x \subset S_\tau$ for all S_τ, thus we only need to show that $\mathfrak{S}_x$ is a σ-algebra.

We have that $X \in S_\tau$ for all S_τ, hence $X \in \mathfrak{S}_x$. If $A \in \mathfrak{S}_x$ it follows that $A \in S_\tau$ for all S_τ, hence $A^c \in S_\tau$ for all S_τ, thus $A^c \in \mathfrak{S}_x$. Finally, assume that $A_j \in \mathfrak{S}_x$ for $j \geq 1$. Then given any S_τ we have that $A_j \in S_\tau$ for $j \geq 1$. Since S_τ is a σ-algebra, we have that $A = \cup_{j=1}^{\infty} A_j \in S_\tau$ for all S_τ. Thus $A \in \mathfrak{S}_x$ and we are done.

□

Let X and Y be topological spaces with the topologies τ_x and τ_y. Consider the corresponding Borel σ-algebras $\mathfrak{S}_x$ and $\mathfrak{S}_y$.

Definition 4.3. *We say that $f : X \mapsto Y$ is (Borel) measurable if $f^{-1}(V) \in \mathfrak{S}_x$ for every open set $V \subset Y$.*

Clearly, every continuous function is Borel measurable, because in this case $f^{-1}(V)$ is open in X, see Theorem 3.2. The next result shows that if a map is Borel measurable, then $f^{-1}(B) \in \mathfrak{S}_x$ for every Borel set $B \in \mathfrak{S}_y$.

Proposition 4.4. *Let $f : X \mapsto Y$ be Borel measurable. Denote by*

$$\Omega := \{B \subset Y : \ f^{-1}(B) \in \mathfrak{S}_x\}.$$

Then Ω contains all open sets in Y and it is a σ-algebra. Thus $\mathfrak{S}_y \subset \Omega$.

Proof. Clearly, if B is open then $f^{-1}(B)$ is measurable and belongs to $\mathfrak{S}_x$. Thus Ω contains all the open sets. If we can prove that Ω is a σ-algebra, then it must contain $\mathfrak{S}_y$.

First, $Y \in \Omega$ because $X = f^{-1}(Y) \in \mathfrak{S}_x$. Second, if $A \in \Omega$ we have that $f^{-1}(A) \in \mathfrak{S}_x$, hence $f^{-1}(A^c) = X \setminus f^{-1}(A) \in \mathfrak{S}_x$ and $A^c \in \Omega$. Third, if $A_j \in \Omega$ for all j, we have $f^{-1}(\cup_j A_j) = \cup_j f^{-1}(A_j) \in \mathfrak{S}_x$, thus $\cup_j A_j \in \Omega$.

□

We now consider an important particular case, where we allow f to take infinite values. Here $Y = \mathbb{R} \cup \{\pm\infty\} := [-\infty, \infty]$. The set Y is not a metric space but we can see it as a topological space containing all the open sets of $\mathbb{R}$ together with sets of the type $(\alpha, \infty]$ and $[-\infty, \alpha)$ which are by definition open sets containing $\pm\infty$. The next result gives a very useful criterion for when a function is Borel measurable.

Theorem 4.5. *Let $f : X \mapsto [-\infty, \infty]$. If for every $\alpha \in \mathbb{R}$ we have $f^{-1}((\alpha, \infty]) \in \mathfrak{S}_x$, then f is Borel measurable.*

Proof. We have to prove that for every possibile open set $V \subset [-\infty, \infty]$ we have $f^{-1}(V) \in \mathfrak{S}_x$. First, we note that $f^{-1}([-\infty, \beta]) = X \setminus f^{-1}((\beta, \infty]) \in \mathfrak{S}_x$ for all $\beta \in \mathbb{R}$. Moreover, we have:

$$f^{-1}([-\infty, \beta)) = \bigcup_{n \geq 1} f^{-1}([-\infty, \beta - 1/n]) \in \mathfrak{S}_x$$

and $f^{-1}([\alpha, \infty]) = X \setminus f^{-1}([-\infty, \alpha)) \in \mathfrak{S}_x$. Now if $\alpha < \beta$ we have:

$$f^{-1}((\alpha, \beta)) = f^{-1}((\alpha, \infty]) \bigcap f^{-1}([-\infty, \beta)) \in \mathfrak{S}_x$$

which shows that every open interval is returned into a Borel set in X.

Now assume that V is some nonempty open set in $\mathbb{R}$. Let $I := \mathbb{Q} \cap V$ be the set of all rational points which belong to V.

Define the set:

$$W := \bigcup_{q \in \mathbb{Q}_+, a \in I, (a-q, a+q) \subset V} (a - q, a + q).$$

Clearly, the above union is at most countable. Moreover,

$$f^{-1}(W) = \bigcup_{q \in \mathbb{Q}_+, a \in I, (a-q, a+q) \subset V} f^{-1}((a - q, a + q)) \in \mathfrak{S}_x,$$

hence if we can prove that $W = V$ we are done.

One inclusion $W \subset V$ is trivial, because all the intervals $(a-q, a+q)$ are included in V. We show the other inclusion. For every $x \in V$ there exists $r > 0$ such that $(x-r, x+r) \subset V$. The point x belongs to (α, β) and we can find a sequence of rational numbers $a_n \in \mathbb{Q}$ such that $\lim_{n\to\infty} a_n = x$. Moreover, there exists a sequence of positive rational numbers q_j such that $\lim_{j\to\infty} q_j = r$. Thus we can find $a_n \in \mathbb{Q}$ and $q_j \in \mathbb{Q}_+$ such that $|a_n - x| < r/10$ and $|q_j - r| < r/10$. Thus

$$x - r < a_n - q_j/2 < x < a_n + q_j/2 < x + r$$

which shows that $x \in (a_n - q_j/2, a_n + q_j/2) \subset W$ and we are done. □

Theorem 4.6. *Let $f_n : X \mapsto [-\infty, \infty]$, $n \geq 1$, be a sequence of measurable functions. Define $f(x) = \sup_{n\geq 1} f_n(x)$ and $g(x) = \inf_{n\geq 1} f_n(x)$. Then both f and g are Borel measurable. Moreover, the functions defined by $\liminf_{n\to\infty} f_n(x)$ and $\limsup_{n\to\infty} f_n(x)$ are measurable.*

Proof. We start with f. According to Theorem 4.5 we have to prove that

$$A(\alpha) := \{x \in X : \ \alpha < f(x)\}$$

is a Borel set in X for every $\alpha \in \mathbb{R}$. Denote by $A_n(\alpha) := \{x \in X : \ \alpha < f_n(x)\} = f_n^{-1}((\alpha, \infty])$. Because f_n is measurable, then each $A_n(\alpha)$ is Borel. Since $f_n(x) \leq f(x)$ for every n, it follows that $A_n(\alpha) \subset A(\alpha)$ hence $\bigcup_{n\geq 1} A_n(\alpha) \subset A(\alpha)$.

Now assume that $x \in A(\alpha)$. This gives $\alpha < f(x)$. We can find n_x large enough such that $\alpha < f_{n_x}(x)$, hence $x \in A_{n_x}(\alpha) \subset \bigcup_{n\geq 1} A_n(\alpha) \subset A(\alpha)$. Thus $A(\alpha) \subset \bigcup_{n\geq 1} A_n(\alpha)$, hence the two sets are equal. But then $A(\alpha)$ equals a countable union of Borel sets, hence it is a Borel set.

The proof for g is similar: one proves that

$$g^{-1}([\alpha, \infty]) = \bigcap_{n\geq 1} f_n^{-1}([\alpha, \infty]).$$

Indeed, if $x \in g^{-1}([\alpha, \infty])$ then $\alpha \leq g(x) \leq f_n(x)$ for all n, hence $x \in f_n^{-1}([\alpha, \infty])$ for all n. The other way around: let x such that $\alpha \leq f_n(x)$ for all n. Then α is a lower bound and $\alpha \leq g(x)$, hence $x \in g^{-1}([\alpha, \infty])$. Now since each $f_n^{-1}([\alpha, \infty])$ is Borel, it follows that $g^{-1}([\alpha, \infty])$ is Borel. Finally we use:

$$g^{-1}((\alpha, \infty]) = \bigcup_{n\geq 1} g^{-1}([\alpha + 1/n, \infty])$$

and we can apply Theorem 4.5.

For liminf and limsup: we note that

$$\liminf f_n(x) = \sup_{k\geq 1}\{\inf_{n\geq k} f_n(x)\}, \quad \limsup f_n(x) = \inf_{k\geq 1}\{\sup_{n\geq k} f_n(x)\}$$

hence the problem is reduced to the previous two cases. □

The last thing we want to discuss here is the fact that continuous operations with real Borel functions produce Borel functions. We start with a technical result:

Lemma 4.7. *Let $f_1, f_2 : X \mapsto \mathbb{R}$ be two Borel functions. Let $F : X \mapsto \mathbb{R}^2$ with $F(x) = [f_1(x), f_2(x)] \in \mathbb{R}^2$. Then F is measurable. The same results holds if $\mathbb{R}$ is replaced with $[-\infty, \infty]$.*

Proof. Let V be an open set in $\mathbb{R}^2$. Let $I = \mathbb{Q}^2 \cap V$. We denote by $R(q, p, s)$ a rectangle centered at a point $q \in \mathbb{Q}^2$ and with side lengts $p, s \in \mathbb{Q}_+$. Reasoning as in Theorem 4.5 we can prove that

$$V = \bigcup_{q\in I, R(q,p,s)\subset V} R(q, p, s),$$

where the union is at most countable. Each such rectangle can be written as a cartesian product of the type $(a,b)\times(c,d)$. Thus

$$F^{-1}((a,b)\times(c,d)) = f_1^{-1}((a,b))\bigcap f_1^{-1}((c,d)) \in \mathfrak{S}_X,$$

and we conclude by using

$$F^{-1}(V) = \bigcup_{q\in I, R(q,p,s)\subset V} F^{-1}(R(q,p,s)) \in \mathfrak{S}_X.$$

If we work with $[-\infty,\infty]$ instead of $\mathbb{R}$ then we have to allow the above rectangles to include sets like $(a,\infty]\times[-\infty,d)$ and all the other possible combinations. □

Theorem 4.8. *Let $\phi:\mathbb{R}^2\mapsto\mathbb{R}$ be a continuous function, and $f_1, f_2 : X\mapsto\mathbb{R}$ be two Borel functions. Then the function:*

$$X\ni x\mapsto f_3(x) := \phi(f_1(x), f_2(x))\in\mathbb{R}$$

is Borel measurable.

Proof. We note that $f_3=\phi\circ F$, where F was defined in Lemma 4.7. Let V open in $\mathbb{R}$. Then $\phi^{-1}(V)$ is open in $\mathbb{R}^2$ and $f_3^{-1}(V) = F^{-1}(\phi^{-1}(V)$ is Borel. □

This proves in particular that the sum of two measurable functions is measurable.

5 Urysohn's Lemma and partition of unity

We will assume in this section that our metric space X is locally compact, which means that for every $x\in X$ we can find an open set W such that $x\in W$ and $\overline{W}$ is compact. Because of the metric space structure, one can show that W can always be chosen to be an open ball whose closure is compact. Indeed, since x is an interior point of W, we can find $\epsilon>0$ such that

$$\overline{B_{\epsilon/2}(x)}\subset B_\epsilon(x)\subset W.$$

Since $\overline{B_{\epsilon/2}(x)}\subset\overline{W}$ and $\overline{W}$ is (sequentially) compact, then $\overline{B_{\epsilon/2}(x)}$ is also sequentially compact, hence compact. Note that this does not mean that all bounded and closed balls are compact in this space.

Let us start with a technical result:

Proposition 5.1. *Let (X,d) be a locally compact metric space, K a compact set, V open, and $K\subset V$. Then there exists an open set U such that*

$$K\subset U\subset\overline{U}\subset V$$

such that $\overline{U}$ is compact.

Proof. We know that for every $x\in K$ there exists $\epsilon>0$ such that $\overline{B_{\epsilon/2}(x)}$ is compact. Note that for every $\delta<\epsilon/2$ we have $\overline{B_\delta(x)}\subset\overline{B_{\epsilon/2}(x)}$ and $\overline{B_\delta(x)}$ is also compact.

Since the same point $x\in V$ and V is open, we can find $\epsilon'>0$ such that $B_{\epsilon'}(x)\subset V$. Thus by choosing a δ smaller than both $\epsilon/2$ and $\epsilon'/2$ we have that $\overline{B_\delta(x)}$ is compact and $\overline{B_\delta(x)}\subset V$.

Now since K is compact and $K\subset\bigcup_{x\in K}B_\delta(x)$, we can extract a finite subcovering $U := \bigcup_{j=1}^N B_{\delta_j}(x_j)\subset V$, where $\overline{U} = \bigcup_{j=1}^N\overline{B_{\delta_j}(x_j)}\subset V$. Moreover, $\overline{U}$ is compact.

□

We will give here a simplified proof of the Urysohn Lemma, only valid in metric spaces. We say that a function $f : X \mapsto \mathbb{C}$ has compact support if the set

$$\text{supp}(f) := \overline{\{x \in X :\ f(x) \neq 0\}}$$

is compact. We denote the set of continuous and compactly supported functions on X by $C_c(X)$.

Theorem 5.2. *(Urysohn's Lemma). Let (X, d) be a locally compact metric space, K a compact set, and V a bounded open set such that $K \subset V$. Then there exists $u \in C_c(X)$, $0 \leq u \leq 1$, such that $u(x) = 1$ if $x \in K$ and $u(x) = 0$ if $x \in V^c$.*

Proof. We start with a technical result concerning closed sets. Let $A = \overline{A}$ be a closed set in X, and define the real function

$$X \ni x \mapsto f(x) = d(x; A) := \inf_{y \in A} d(x, y) \in \mathbb{R}.$$

Now let us show that f is uniformly continuous, and $f(x) = 0$ iff $x \in A$. Indeed, if $x, x' \in X$ then $f(x') \leq d(x', y) \leq d(x', x) + d(x, y)$ for all $y \in X$, which means that $f(x') - d(x', x)$ is a lower bound for $d(x, y)$ for all y, which gives $f(x') - d(x', x) \leq f(x)$, or $f(x') - f(x) \leq d(x', x)$. By symmetry, we conclude that $|f(x') - f(x)| \leq d(x', x)$ which immediately implies that f is uniformly continuous.

Now if $x \in A$, clearly $f(x) = 0$. The other way around: assume that $f(x) = d(x; A) = 0$. It means that there exists a sequence of points $y_n \in A$ such that $0 = d(x; A) \leq d(x, y_n) \leq 1/n$. But this means that y_n converges to x, i.e. x is an adherent point of A, hence belongs to A.

Getting back to our problem, let us first construct the set U as in Proposition 5.1, which has the property that $K \subset U \subset \overline{U} \subset V$ and $\overline{U}$ is compact. Note that both K and U^c are closed sets. We observe that we have $d(x; K) + d(x; U^c) > 0$ for all x. This is because either $d(x; U^c) > 0$ and we are done, or $d(x; U^c) = 0$ which implies $x \in U^c$ and $x \notin K$, in other words $d(x; K) > 0$.

Finally, the function we are looking for can be defined as:

$$u(x) = \frac{d(x; U^c)}{d(x; U^c) + d(x; K)},$$

where the continuity is obvious, and by construction, $\text{supp}(u) = \overline{U}$ is compact. □

Theorem 5.3. *Let (X, d) be a metric space, K a compact set, and $\{V_j\}_{j=1}^N$ are bounded open sets such that $K \subset \bigcup_{j=1}^N V_j$. Then there exist some continuous non-negative functions $g_j : X \mapsto [0, 1]$, $j \in \{1, \ldots, N\}$, such that $g_j(x) = 0$ if $x \notin V_j$, and if $h(x) := g_1(x) + \cdots + g_N(x)$, then $0 \leq h \leq 1$, h is compactly supported, $h(x) = 1$ if $x \in K$ and $h(x) = 0$ if $x \notin \bigcup_{j=1}^N V_j$.*

Proof. For every $x \in K$, there exists some j such that $x \in V_j$. We can choose $\epsilon > 0$ such that

$$x \in \overline{B_{\epsilon/2}(x)} \subset B_\epsilon(x) \subset V_j$$

and where $\overline{B_{\epsilon/2}(x)}$ is compact. We have the inclusion $K \subset \bigcup_{x \in K} B_{\epsilon/2}(x)$ from which we can find a finite subcovering:

$$K \subset \bigcup_{m=1}^M B_{\epsilon_m/2}(x_m) \subset \bigcup_{m=1}^M \overline{B_{\epsilon_m/2}(x_m)}.$$

Each compact set $\overline{B_{\epsilon_m/2}(x_m)}$ is included in some set V_j. Denote by K_j the union of all closed balls $\overline{B_{\epsilon_m/2}(x_m)}$ which are included in a given V_j. Then K_j is compact and $K_j \subset V_j$. According to Urysohn's lemma, we can find a continuous, compactly supported function $0 \leq u_j \leq 1$ which equals 1 on K_j and $\text{supp}(u_j) \subset V_j$, for all $j \in \{1, \ldots, N\}$.

Let us define $g_1 = u_1$ and $g_2 = u_2(1 - g_1)$. We see that the support of g_2 is included in the support of u_2, thus it must be compact (a closed subset of a compact set is compact). We also have that $0 \leq g_1 + g_2 \leq 1$. Moreover, $g_1 + g_2$ equals $g_1 = u_1 = 1$ on K_1, equals $g_1 + (1 - g_1) = 1$ on K_2, and $g_1 + g_2 = 0$ outside $V_1 \bigcup V_2$. The support of $g_1 + g_2$ is a closed subset of $\mathrm{supp}(g_1) \bigcup \mathrm{supp}(g_2)$ thus it must be compact.

In conclusion, $0 \leq g_1 + g_2 \leq 1$ is compactly supported, equals 1 on $K_1 \bigcup K_2$ and equals 0 outside $V_1 \bigcup V_2$. Now let us assume that we constructed the compactly supported functions g_1,..., g_m such that $0 \leq g_1 + \cdots + g_m \leq 1$ equals 1 on $\bigcup_{j=1}^m K_j$, and equals 0 outside $\bigcup_{j=1}^m V_j$. Then we can construct $g_{m+1} = u_{m+1}(1 - g_1 - \cdots - g_m)$ and check that the induction step is fulfilled. Note that $h = g_1 + \cdots + g_N$ equals 1 on K because $K \subset \bigcup_{j=1}^N K_j$.

Remark. If χ_A denotes the characteristic function of a set A, then we have:

$$0 \leq g_j \leq \chi_{V_j}, \quad \chi_K \leq h \leq \chi_{\bigcup_{j=1}^N V_j}.$$

□

6 The Riesz representation theorem

Again we assume that we work in a locally compact metric space (X, d). A map $\Lambda : C_c(X) \mapsto \mathbb{C}$ is called a linear positive functional if Λ is linear, and $\Lambda(f) \geq 0$ if $f \geq 0$. Note that this implies the monotony property:

$$\Lambda(f) = -\Lambda(g - f) + \Lambda(g) \leq \Lambda(g), \quad \text{if} \quad f \leq g. \tag{6.1}$$

Theorem 6.1. *Let Λ be a positive linear functional on $C_c(X)$. Then there exists a sigma algebra $\mathfrak{S}$ in X which contains the Borel sets in X (in particular the open and compact sets), and there exists a unique positive measure μ defined on $\mathfrak{S}$ such that:*
(a). $\Lambda(f) = \int_X f d\mu$, for all $f \in C_c(X)$;
(b). $\mu(K) < \infty$ if K is compact;
(c). For every $E \in \mathfrak{S}$ we have the identity:

$$\mu(E) = \inf\{\mu(V) : \ E \subset V, \ V \text{ is open}\};$$

(d). If $E \in \mathfrak{S}$ is either open or $\mu(E) < \infty$, we have the identity:

$$\mu(E) = \sup\{\mu(K) : \ K \subset E, \ K \text{ is compact}\};$$

(e). Suppose that $E \in \mathfrak{S}$ with $\mu(E) = 0$. Then every subset $A \subset E$ has the property that $A \in \mathfrak{S}$ and $\mu(A) = 0$.

Remark. Although we will follow quite closely the presentation of Rudin, we will give more details and try to explain things in a more pedagogical manner. The rest of this section is dedicated to the proof of Riesz' theorem, which will be split in a number of technical lemmas.

Lemma 6.2. *If such a measure exists, then it is unique.*

Proof. We assume that there exist two measures μ_1 and μ_2 which obey all the above five conditions. We start by proving that $\mu_1(K) = \mu_2(K)$ for any compact set K.

Fix some compact K. Using (b) and (c) for μ_2 and $E = K$, we know that for every $n \geq 1$, we can find some open set V_n with $K \subset V_n$ such that

$$\mu_2(K) \leq \mu_2(V_n) \leq \mu_2(K) + 1/n.$$

From Urysohn's Lemma we know that we can find some function $u_n \in C_c(X)$ such that $\chi_K \leq u_n \leq \chi_{V_n}$. Then we have:

$$\mu_1(K) \leq \int_X u_n(x) d\mu_1(x) = \Lambda(u_n) = \int_X u_n(x) d\mu_2(x) \leq \mu_2(V_n) \leq \mu_2(K) + 1/n.$$

Taking the limit, it gives $\mu_1(K) \le \mu_2(K)$. By swapping μ_1 with μ_2 in the above argument we can also prove the reversed inequality, thus the two values must be equal.

Knowing that μ_1 and μ_2 coincide on compact sets, then due to (d) we conclude that $\mu_1(V) = \mu_2(V)$ for every open set V. Note the important thing, that if V is open, then $\mu_j(V)$ can be infinite.

Now if E is any set in $\mathfrak{S}$ which is neither open nor compact for which both $\mu_j(E) < \infty$, then again (c) implies that $\mu_1(E) = \mu_2(E)$.

The last possibility is when $\mu_1(E) < \infty$ and $\mu_2(E) = \infty$ for some E which is neither open nor compact. Since $\mu_1(E) < \infty$, from (c) there must exist some open set V which contains E and $\mu_1(V) < \infty$. But then $\mu_2(V) = \mu_1(V) < \infty$ which shows that $\mu_2(E) \le \mu_2(V) < \infty$, contradiction. □

Now we start the construction of both $\mathfrak{S}$ and μ. First we consider a map F defined on the maximal σ-algebra $\mathcal{P}(X)$ (consisting of all possible subsets of X), in the following way: if V is open, then

$$F(V) := \sup\{\Lambda(f) : \; f \in C_c(X), \; 0 \le f \le \chi_V\}, \quad F(\emptyset) := 0, \tag{6.2}$$

and if E is an arbitrary set which is not open then:

$$F(E) := \inf\{F(V) : \; E \subset V, \; V \text{ open}\}. \tag{6.3}$$

We observe that if $U \subset V$ are open sets, then the inequality $f \le \chi_V$ is implied by $f \le \chi_U$, thus we have the inclusion:

$$\{\Lambda(f) : \; f \in C_c(X), \; 0 \le f \le \chi_U\} \subset \{\Lambda(f) : \; f \in C_c(X), \; 0 \le f \le \chi_V\},$$

which imediately implies that $F(U) \le F(V)$. This also shows that (6.3) is compatible with the situation in which E is allowed to be open, since in this case $F(E)$ is a minimum because the infimum is realized for $V = E$.

Lemma 6.3. *Let $F : \mathcal{P}(X) \mapsto \mathbb{R}_+$ defined as above. If $E_1 \subset E_2$, then $F(E_1) \le F(E_2)$.*

Proof. We have already proved the monotony if E_1 and E_2 are open sets. For the general case, we observe that if V is an open set such that $E_2 \subset V$, then we also have $E_1 \subset V$. This shows that we have the inclusion:

$$\{F(V) : \; E_2 \subset V, \; V \text{ open}\} \subset \{F(V) : \; E_1 \subset V, \; V \text{ open}\},$$

thus $F(E_1)$ is a lower bound for the set $\{F(V) : \; E_2 \subset V, \; V \text{ open}\}$, which leads to $F(E_1) \le F(E_2)$. □

Definition 6.4. *We define $\mathcal{M} \subset \mathcal{P}(X)$ as the collection of all sets E for which $F(E) < \infty$ and moreover, the following identity is satisfied:*

$$F(E) = \sup\{F(K) : \; K \subset E, \; K \text{ compact}\}. \tag{6.4}$$

We are now ready to define our σ-algebra $\mathfrak{S}$ and the measure μ.

Definition 6.5. *We denote by $\mathfrak{S}$ the collection of all the sets E which obey the condition that for every compact set K, we have that $E \bigcap K \in \mathcal{M}$. We also denote by μ the restriction of F to $\mathfrak{S}$.*

In the rest of this section we will prove that $\mathfrak{S}$ is a σ-algebra containing all the Borel sets, and that μ is a positive measure obeying (a)-(e).

6.1 $\mathfrak{S}$ is a σ-algebra and μ is a positive measure

Here we will more or less follow the ten steps (numberred from I to X) in Rudin's book.

6.1.1 Step I

We start with a general technical result which is valid for the map F:

Lemma 6.6. *Let $F : \mathcal{P}(X) \mapsto \mathbb{R}_+$ defined as above. If $\{E_j\}_{j=1}^{\infty}$ are arbitrary sets, then*

$$F\Big(\bigcup_{j=1}^{\infty} E_j\Big) \leq \sum_{j=1}^{\infty} F(E_j). \tag{6.5}$$

Proof. Let us start by showing that if V_1 and V_2 are open sets, then $F(V_1 \cup V_2) \leq F(V_1) + F(V_2)$. The set $V = V_1 \cup V_2$ is open, thus according to the definition of $F(V)$ we have to estimate the supremum of $\Lambda(f)$ for every $f \in C_c(X)$ with $f \leq \chi_V$. Denote by K the support of f. Then $K \subset V_1 \cup V_2$, thus using the partition of unity theorem 5.3 we can find two $C_c(X)$ functions g_1 and g_2 such that $g_1 \leq \chi_{V_1}$ and $g_2 \leq \chi_{V_2}$, such that $g_1 + g_2 = 1$ on K. This means that $f = fg_1 + fg_2$, where $fg_1 \leq \chi_{V_1}$ and $fg_2 \leq \chi_{V_2}$. Thus:

$$\Lambda(f) = \Lambda(fg_1) + \Lambda(fg_2) \leq F(V_1) + F(V_2),$$

inequality which shows that $F(V_1) + F(V_2)$ is an upper bound for our set. Since $F(V)$ is the smallest upper bound, we must have $F(V) \leq F(V_1) + F(V_2)$.

Now let us treat the general case. If there exists some i for which $F(E_i) = \infty$, the proof is over. Hence we may assume that all $F(E_j) < \infty$.

Let us fix $\epsilon > 0$. According to formula (6.3), for every $j \geq 1$ we can find an open set V_j such that $E_j \subset V_j$ and

$$F(E_j) \leq F(V_j) \leq F(E_j) + \epsilon 2^{-j}. \tag{6.6}$$

Clearly, $\bigcup_{j=1}^{\infty} E_j \subset \bigcup_{j=1}^{\infty} V_j$, thus using the monotony from Lemma 6.3 we have:

$$F\Big(\bigcup_{j=1}^{\infty} E_j\Big) \leq F\Big(\bigcup_{j=1}^{\infty} V_j\Big).$$

Denote by $V = \bigcup_{j=1}^{\infty} V_j$; from Lemma 1.2 we know that V is open. We have to estimate $F(V)$. As before, choose any $f \in C_c(X)$ with $f \leq \chi_V$. The support K of f is compact and $K \subset V = \bigcup_{j=1}^{\infty} V_j$. Thus we can extract a finite subcovering, i.e. $K \subset \bigcup_{k=1}^{N} V_{j_k}$. Using again the partition of unity theorem 5.3 we can find g_1, ...,g_N such that $g_i \leq \chi_{V_{j_i}}$ and $g_1 + \cdots + g_N = 1$ on K. Reasoning as before, we have:

$$\Lambda(f) = \Lambda(fg_1) + \cdots + \Lambda(fg_N) \leq F(V_{j_1}) + \cdots + F(V_{j_N}) \leq \sum_{j \geq 1} F(V_j).$$

This inequality holds true for every f, thus:

$$F\Big(\bigcup_{j=1}^{\infty} E_j\Big) \leq F(V) \leq \sum_{j \geq 1} F(V_j).$$

Now using the second inequality in (6.6) we have:

$$F\Big(\bigcup_{j=1}^{\infty} E_j\Big) \leq \sum_{j \geq 1} F(E_j) + \epsilon,$$

which finishes the proof.

□

6.1.2 Step II

Step 2 in Rudin's proof corresponds to the next lemma, which coupled with the monotony property and (6.4) it shows that every compact set belongs to $\mathcal{M}$ and obviously to $\mathfrak{S}$.

Lemma 6.7. *Let $F : \mathcal{P}(X) \mapsto \mathbb{R}_+$ defined as above. If K is compact, then $F(K) < \infty$.*

Proof. Every compact set is bounded, thus there exists some ball V such that $K \subset V$. According to Lemma 5.1, we can always find an open set U with $\overline{U}$ compact such that $K \subset U \subset \overline{U} \subset V$. Due to Urysohn's lemma we can find a function $f_u \in C_c(X)$ such that $\chi_{\overline{U}} \leq f_u \leq \chi_V$. Thus for every $f \in C_c(X)$ with $0 \leq f \leq \chi_U$ we must have $f \leq f_u$ and hence $\Lambda(f) \leq \Lambda(f_u)$. Taking the supremum over all such f's, according to (6.2) we must have $F(U) \leq \Lambda(f_u) < \infty$. Then (6.3) finally implies $F(K) \leq F(U) < \infty$. □

6.1.3 Step III

This step is split into two lemmas.

Lemma 6.8. *Let $f \in C_c(X)$ with $f \leq 1$ and let K be the support of f. Then $\Lambda(f) \leq F(K)$.*

Proof. Let V be any open set containing K. Due to Urysohn's Lemma, there exists some non-negative $g \in C_c(X)$ with $\chi_K \leq g \leq \chi_V$, hence $f \leq g$, which implies $\Lambda(f) \leq \Lambda(g)$. Due to the definition of $F(V)$ in (6.2), we have that $\Lambda(g) \leq F(V)$.

Thus we have just shown that $\Lambda(f) \leq F(V)$ for every open V containing K. It means that $\Lambda(f)$ is a lower bound for the set on the right hand side of (6.3) if $E = K$. Hence $\Lambda(f) \leq F(K)$. □

Lemma 6.9. *Every open set satisfies* (6.4)*, hence all open sets V with $F(V) < \infty$ belong to $\mathcal{M}$.*

Proof. Let V be a fixed non-empty open set. The monotony property implies that $F(V)$ is an upper bound for the set containing all the numbers $F(K)$ with $K \subset V$. It remains to show that $F(V)$ is an adherent point for this set. The strategy is to show that for every real number $\alpha < F(V)$, there exists a compact set $K_\alpha \subset V$ such that $\alpha < F(K_\alpha) \leq F(V)$. Note that this also covers the case in which $F(V) = \infty$, because in that case we can take α to be any natural number.

From the definition of $F(V)$ in (6.2), we know that given $\alpha < F(V)$ we can find $f_\alpha \in C_c(X)$ with its support $K_\alpha \subset V$ such that $\alpha < \Lambda(f_\alpha) \leq F(V)$. On one hand, $F(K_\alpha) \leq F(V)$, while on the other hand, due to Lemma 6.8 we know that: $\alpha < \Lambda(f_\alpha) \leq F(K_\alpha)$.

□

6.1.4 Step IV

Here we prove that F is additive on $\mathcal{M}$. We start with a preparatory result.

Lemma 6.10. *Let K_1 and K_2 be two disjoint compact sets. Then $F(K_1 \cup K_2) = F(K_1) + F(K_2)$.*

Proof. The set K_2^c is open and $K_1 \subset K_2^c$. According to Proposition 5.1, we can find an open set U_1 with $\overline{U_1}$ compact, such that $K_1 \subset U_1 \subset \overline{U_1} \subset K_2^c$. In an analogous way, we can construct an open set U_2 with $\overline{U_2}$ compact, such that $K_2 \subset U_2 \subset \overline{U_2} \subset \overline{U_1}^c$. In particular, U_1 and U_2 are disjoint.

We already know from Lemma 6.6 that $F(K_1 \cup K_2) \leq F(K_1) + F(K_2)$, thus we only need to prove the reverse inequality. Fix $\epsilon > 0$.

From (6.3) we know that there exists V open, with $K_1 \cup K_2 \subset V$ such that $F(V) < F(K_1 \cup K_2) + \epsilon$. Denote by $V_1 := V \cap U_1$ and $V_2 := V \cap U_2$. We have $(V_1 \cup V_2) \subset V$ hence

$$F(V_1 \cup V_2) \leq F(V) < F(K_1 \cup K_2) + \epsilon.$$

Now let f_1 and f_2 be two arbitrary non-negative $C_c(X)$ functions with $0 \leq f_j \leq \chi_{V_j}$. Because f_1 and f_2 cannot be simultaneously non-zero (V_1 and V_2 are disjoint), it turns out that $0 \leq f_1 + f_2 \leq \chi_{V_1 \cup V_2}$. Thus $\Lambda(f_1 + f_2) \leq F(V_1 \cup V_2)$ and:

$$\Lambda(f_1) + \Lambda(f_2) = \Lambda(f_1 + f_2) \leq F(V_1 \cup V_2).$$

Now we can separately take the supremum over f_1 and f_2, which gives:

$$F(V_1) + F(V_2) \leq F(V_1 \cup V_2).$$

Since $K_j \subset V_j$, we have $F(K_j) \leq F(V_j)$ and we conclude:

$$F(K_1) + F(K_2) \leq F(V_1) + F(V_2) \leq F(V_1 \cup V_2) \leq F(V) < F(K_1 \cup K_2) + \epsilon.$$

□

Lemma 6.11. *Let E_1 and E_2 be two disjoint sets in $\mathcal{M}$. Then $E_1 \cup E_2 \in \mathcal{M}$ and $F(E_1 \cup E_2) = F(E_1) + F(E_2)$.*

Proof. Denote by $E = E_1 \cup E_2$. Since $E_j \in \mathcal{M}$, according to the definition we know $F(E_j) < \infty$. Hence Step I implies that $F(E) < \infty$.

Moreover, (6.4) is satisfied for E_j. Given $\epsilon > 0$, we can find $K_j \subset E_j$ two compact disjoint sets such that $F(E_j) < F(K_j) + \epsilon/2$. Thus together with Step I we have:

$$F(E) \leq \sum_{j=1}^{2} F(E_j) < \epsilon + \sum_{j=1}^{2} F(K_j) = \epsilon + F(K_1 \cup K_2) \leq \epsilon + F(E)$$

where the equality in the middle is a consequence of the previous lemma, and the last inequality comes from monotony since $K_1 \cup K_2 \subset E$.

□

Lemma 6.12. *Let $\{E_j\}_{j\geq 1}$ be disjoint sets in $\mathcal{M}$ and define $E := \bigcup_{j\geq 1} E_j$. Then $F(E) = \sum_{j\geq 1} F(E_j)$. Moreover, if $F(E) < \infty$ then $E \in \mathcal{M}$.*

Proof. From Step I, we know that

$$F(E) \leq \sum_{j\geq 1} F(E_j),$$

hence if $F(E) = \infty$, the proof is over.

Now let us assume that $F(E) < \infty$. Fix $\epsilon > 0$. As before, for every $j \geq 1$ there exists a compact $K_j \subset E_j$ such that $F(E_j) < F(K_j) + \epsilon 2^{-j}$. Then for every $N \geq 1$ we have:

$$\sum_{j=1}^{N} F(E_j) \leq \epsilon + \sum_{j=1}^{N} F(K_j) = \epsilon + F\big(\bigcup_{j=1}^{N} K_j\big) \leq \epsilon + F(E),$$

where the identity in the middle is a consequence of Lemma 6.10 (all K_j's are disjoint), while the last inequality is due to monotony. The increasing sequence $\sum_{j=1}^{N} F(E_j)$ is bounded from above, thus it converges, and we must have:

$$F(E) \leq \sum_{j\geq 1} F(E_j) \leq \epsilon + F(E),$$

which shows that $F(E) = \sum_{j\geq 1} F(E_j) < \infty$. Moreover, the increasing real sequence $F\big(\bigcup_{j=1}^{N} K_j\big)$ is also convergent and:

$$F(E) \leq \epsilon + \lim_{N\to\infty} F\big(\bigcup_{j=1}^{N} K_j\big)$$

which shows that if N_ϵ is large enough, then

$$F(E) < 2\epsilon + F\big(\bigcup_{j=1}^{N_\epsilon} K_j\big)$$

which shows that $E \in \mathcal{M}$.

□

6.1.5 Step V

Lemma 6.13. *Fix $E \in \mathcal{M}$. For every $\epsilon > 0$ there exists an open set V and a compact set K such that $K \subset E \subset V$, $F(K) \leq F(E) \leq F(V)$ and $F(V \setminus K) < \epsilon$.*

Proof. We have $F(E) < \infty$, hence due to (6.3) there exists V open such that $E \subset V$ and $F(V) < F(E) + \epsilon/2$. From (6.4) we obtain a compact $K \subset E$ such that $F(E) < F(K) + \epsilon/2$. This leads to

$$F(V) < F(K) + \epsilon.$$

The set $V \setminus K = V \cap K^c$ is open, included in V, hence $F(V \setminus K) < \infty$. Due to Lemma 6.9 in Step III, we know that $V \setminus K \in \mathcal{M}$. Then both K and $V \setminus K$ belong to $\mathcal{M}$, they are disjoint and according to Lemma 6.10 we have:

$$F(V \setminus K) = F(V) - F(K) < \epsilon$$

and we are done.

□

6.1.6 Step VI

Lemma 6.14. *Let A and B be two sets in $\mathcal{M}$. Then $A \setminus B$, $A \cup B$ and $A \cap B$ belong to $\mathcal{M}$.*

Proof. We start by showing that $A \setminus B \in \mathcal{M}$. First, $F(A \setminus B) \leq F(A) < \infty$. Second, we need to prove (6.4), i.e. given $\epsilon > 0$ we need to construct a certain compact K included in $A \setminus B$ such that $F(A \setminus B) < F(K) + \epsilon$.

From Step V (Lemma 6.13) we know that there exists an open set V_a and a compact set K_a such that $K_a \subset A \subset V_a$ and $F(V_a \setminus K_a) < \epsilon/2$. In the same way, there exists an open set V_b and a compact set K_b such that $K_b \subset B \subset V_b$ and $F(V_b \setminus K_b) < \epsilon/2$.

The set $K := K_a \setminus V_b = K_a \cap V_b^c$ is compact because V_b^c is closed. We have the inclusion:

$$A \setminus B \subset V_a \setminus K_b \subset (V_a \setminus K_a) \cup K \cup (V_b \setminus K_b).$$

Thus Step I implies:

$$F(A \setminus B) \leq F(V_a \setminus K_a) + F(K) + F(V_b \setminus K_b) < \epsilon + F(K).$$

Thus $A \setminus B \in \mathcal{M}$. Then we can write $A \cup B = (A \setminus B) \cup B$, where $A \setminus B$ and B are disjoint, hence due to Lemma 6.11 we get that $A \cup B \in \mathcal{M}$. Finally, using $A \cap B = A \setminus (A \setminus B)$ the proof is over.

□

6.1.7 Step VII

Lemma 6.15. *The collection of sets $\mathfrak{S}$ is a σ-algebra containing all the Borel sets in X.*

Proof. The collection $\mathfrak{S}$ is defined in Definition 6.5. In order to show that $\mathfrak{S}$ is a σ-algebra, we need to prove three things:

1. $\emptyset \in \mathfrak{S}$; this is trivial. Moreover, since $X \cap K = K$ for every compact K, we have that $X \in \mathfrak{S}$.

2. If $A \in \mathfrak{S}$ then we have to prove that $A^c \in \mathfrak{S}$. In other words, for every compact K we need to show that $A^c \cap K \in \mathcal{M}$. Note that $A^c \cap K = K \setminus (K \cap A)$. Since $A \in \mathfrak{S}$ it means that $K \cap A \in \mathcal{M}$. Then Step VI (Lemma 6.14) shows that $A^c \cap K \in \mathcal{M}$.

3. If $\{A_j\}_{j\geq 1} \subset \mathfrak{S}$ then we have to prove that $A := \cup_{j\geq 1} A_j \in \mathfrak{S}$. Fix some compact K. Denote by $\tilde{A}_j := A_j \cap K \in \mathcal{M}$. Define $B_1 := \tilde{A}_1$ and $B_n := \tilde{A}_n \setminus \tilde{A}_{n-1}$ if $n \geq 2$. The sets B_j are all distinct and belong to $\mathcal{M}$ due to Step VI. We have the identity:

$$A \cap K = \cup_{j\geq 1} B_j.$$

We note that $F(A \cap K) \leq F(K) < \infty$, thus according to Lemma 6.12 we have that $A \cap K \in \mathcal{M}$. Hence $A \in \mathfrak{S}$.

Finally, we need to prove that all open sets are in $\mathfrak{S}$. We will prove instead that all closed sets are in $\mathfrak{S}$; then if V is open, we have that $V = (V^c)^c$ and we are done. Now if C is closed, then for any compact K we have that $C \cap K$ is compact, hence $C \cap K \in \mathcal{M}$ and $C \in \mathfrak{S}$.

□

6.1.8 Step VIII

Remember that μ denotes the restriction of F to $\mathfrak{S}$.

Lemma 6.16. *We have:*

$$\mathcal{M} = \{E \in \mathfrak{S} : \ \mu(E) < \infty\}.$$

Proof. We first prove the inclusion '$\subset$'. Let $A \in \mathcal{M}$. Then for every compact K (which belongs to $\mathcal{M}$ according to Step II) we have that $A \cap K \in \mathcal{M}$ from Step VI. Thus $A \in \mathfrak{S}$, and since $A \in \mathcal{M}$ we have $F(A) = \mu(A) < \infty$.

Now we prove the other inclusion. Let $A \in \mathfrak{S}$ with $\mu(A) < \infty$. From the definition of $F(A) = \mu(A)$ in (6.3), it follows that there exists some open set V with $A \subset V$, $F(V) < \infty$ and $F(V) < \mu(A) + \epsilon/2$. Due to Lemma 6.9 in Step III we conclude that $V \in \mathcal{M}$ (hence in $\mathfrak{S}$).

From Lemma 6.13 in Step V applied to $E = V \in \mathcal{M}$ (there we can put $E = V$ and the result is unchanged), we can construct a compact $H \subset V$ such that $F(V \setminus H) \leq \epsilon/2$.

Since $A \cap H$ belongs to $\mathcal{M}$, there exists a compact $K \subset (A \cap H) \subset A$ such that $F(A \cap H) < F(K) + \epsilon/2$. Since $A \subset (A \cap H) \cup (V \setminus H)$, from Step I we have:

$$F(A) \leq F(A \cap H) + F(V \setminus H) < F(K) + \epsilon.$$

Thus $A \in \mathcal{M}$. □

6.1.9 Step IX

Lemma 6.17. *μ is a measure on $\mathfrak{S}$.*

Proof. The non-negativity of μ comes from the definition of F. We need to prove that μ is countably additive. Let $\{A_j\} \subset \mathfrak{S}$ be disjoint sets and denote by $A := \cup_{j\geq 1} A_j$. From Step VII we know that $\mathfrak{S}$ is a σ-algebra thus $A \in \mathfrak{S}$. Now if $\mu(A) = \infty$, from Step I it follows that $\infty = \mu(A) = \sum_{j\geq 1} \mu(A_j)$.

If $\mu(A) < \infty$, then from Step VIII it follows that $A \in \mathcal{M}$. Since $A_j \subset A$ it follows that $\mu(A_j) \leq \mu(A) < \infty$, thus Step VIII implies again that all A_j's are in $\mathcal{M}$. Finally, from Step IV (Lemma 6.12) it follows that $\sum_{j\geq 1} \mu(A_j) = \mu(A)$ and we are done. □

6.1.10 Step X

Lemma 6.18. *For every $f \in C_c(X)$ we have $\Lambda(f) = \int_X f d\mu$.*

Proof. Every function $f \in C_c(X)$ can be written as $f = u + iv$ where u and v are continuous and real valued. Since Λ is linear, it is enough to prove the lemma for real valued f. In fact it is enough to prove that $\Lambda(f) \leq \int_X f d\mu$ for every real valued function; this is because we would also have:

$$-\Lambda(f) = \Lambda(-f) \leq \int_X (-f) d\mu = -\int_X f d\mu$$

which provides the other inequality.

Now fix $\epsilon > 0$. Let K be the compact support of a given f. Since f is continuous, Theorem 3.4 implies that $f(K) \subset \mathbb{R}$ is compact, thus it must be bounded and closed. Hence there exists an interval $[a, b]$ which contains the range of $f(X) = f(K) \cup \{0\}$ (actually, for continuous functions we have $0 \in f(K)$, or more precisely, $f(\partial K) = \{0\}$).

Choose $n+1$ points $\{y_j\}_{j=0}^n$ such that $y_0 < a < y_1 < \dots y_n = b$ and $\max_{j=1}^n |y_j - y_{j-1}| < \epsilon$. Introduce the sets

$$E_j := K \cap f^{-1}((y_{j-1}, y_j]) = \{x \in K : \ y_{j-1} < f(x) \leq y_j\}, \ j \in \{1, \dots n\}.$$

Because f is continuous, and because all intervals of the type $(\alpha, \beta] \subset \mathbb{R}$ are Borel sets, then the $f^{-1}((y_{j-1}, y_j])$'s are disjoint Borel sets in X and must belong to $\mathfrak{S}$ (see Step VII). Then the E_j's belong to $\mathcal{M}$. Thus we can find some open sets $\tilde{V}_j$ in X such that $E_j \subset \tilde{V}_j$ and $\mu(\tilde{V}_j) < \mu(E_j) + \epsilon/n$.

Since $f^{-1}((y_{j-1}, y_j + \epsilon))$ is open according to Theorem 3.2, and includes E_j, then

$$V_j = \tilde{V}_j \cap f^{-1}((y_{j-1}, y_j + \epsilon))$$

is open, contains E_j and

$$\mu(V_j) \leq \mu(\tilde{V}_j) < \mu(E_j) + \epsilon/n,$$

and at the same time:

$$f(x) \leq y_j + \epsilon, \quad \forall x \in V_j.$$

Clearly, $K \subset \cup_{j=1}^n E_j \subset \cup_{j=1}^n V_j$, and according to the partition of unity theorem we can find some functions h_j such that $\sum_j h_j = 1$ on K and h_j is supported in V_j. Hence:

$$f(x) = \sum_{j=1}^n f(x) h_j(x) \leq \sum_{j=1}^n (y_j + \epsilon) h_j(x).$$

Then using the monotony and linearity of Λ, we have:

$$\Lambda(f) \leq \sum_{j=1}^n (y_j + \epsilon) \Lambda(h_j) = -|a| \sum_{j=1}^n \Lambda(h_j) + \sum_{j=1}^n (y_j + |a| + \epsilon) \Lambda(h_j).$$

Now $\sum_{j=1}^n \Lambda(h_j) = \Lambda(\sum_{j=1}^n h_j) \geq \mu(K)$ right from the definition of μ. Together with the fact that $y_j + |a| > 0$ for all $j \geq 1$, we have:

$$\begin{aligned}
\Lambda(f) &\leq -|a|\mu(K) + \sum_{j=1}^n (y_j + |a| + \epsilon) \Lambda(h_j) \\
&\leq -|a|\mu(K) + \sum_{j=1}^n (y_j + |a| + \epsilon) \mu(V_j) \leq -|a|\mu(K) + \sum_{j=1}^n (y_j + |a| + \epsilon)(\mu(E_j) + \epsilon/n) \\
&\leq -|a|\mu(K) + \sum_{j=1}^n (|a| + 2\epsilon)(\mu(E_j) + \epsilon/n) + \sum_{j=1}^n (y_j - \epsilon)\mu(E_j) + \frac{\epsilon}{n} \sum_{j=1}^n (y_j - \epsilon).
\end{aligned}$$

On each E_j with $j \in \{1, \dots, n\}$ we have that $y_j - \epsilon \leq y_{j-1} < f(x)$ hence

$$\sum_{j=1}^n (y_j - \epsilon) \chi_{E_j}(x) \leq f(x).$$

Thus we have:

$$\sum_{j=1}^{n}(y_j - \epsilon)\mu(E_j) \leq \int_X f d\mu, \quad \sum_{j=1}^{n}\mu(E_j) = \mu(K) < \infty, \quad y_j \leq b.$$

This gives:

$$\Lambda(f) \leq \int_X f d\mu + 2\epsilon\mu(K) + \epsilon(b + |a| + \epsilon).$$

Since ϵ was arbitrary, the proof is over. □

6.2 Summarizing the proof of Riesz' Theorem

We have proved in Lemma 6.2 that if such a measure exists, then it must be unique. Then in Definition 6.5 we introduced a collection of sets $\mathfrak{S}$ and a map $\mu : \mathfrak{S} \mapsto \mathbb{R}_+$ which turn out to be a σ-algebra (proved in Lemma 6.15) and a measure (proved in Lemma 6.17).

The identity claimed in (a) is proved in Step X (see Lemma 6.18). The estimate (b) is proved in Step II (see Lemma 6.7). The identity (c) is fulfilled right from the beggining, in the way μ was defined (through F) in (6.2) and (6.3). Point (d) is shown for open sets in Step III (see Lemma 6.9), while for arbitrary sets $E \in \mathfrak{S}$ with $\mu(E) < \infty$ it is shown in Step VIII (see Lemma 6.16). Finally, point (e) follows from (d): if $A \subset E$ then $F(A) \leq F(E) = \mu(E) = 0$ and every compact $K \subset A$ has $F(K) \leq F(A) = 0$, thus $A \in \mathfrak{S}$ and $\mu(A) = 0$.

7 Spaces of bounded/continuous functions

Proposition 7.1. *Let (X, d) be a metric space, $(Y, ||\cdot||)$ a normed space, and H an arbitrary non-empty subset of X. We define*

$$B(H; Y) := \{f : H \to Y : \sup_{x\in H} ||f(x)|| < \infty\}.$$

Define the map $||\cdot||_\infty : B(H;Y) \to \mathbb{R}_+$, $\quad ||f||_\infty := \sup_{x\in H} ||f(x)||$. Then $(B(H;Y), ||\cdot||_\infty)$ is a normed space.

Proof. Clearly, $||f||_\infty = 0$ if and only if $f(x) = 0$ for all $x \in H$. Moreover,

$$||\lambda f||_\infty = \sup_{x\in H} ||\lambda f(x)|| = |\lambda| \sup_{x\in H} ||f(x)|| = |\lambda| \; ||f||_\infty.$$

Finally, let us prove the triangle inequality. Take $f, g \in B(H;Y)$; then for every $x \in H$ we apply the triangle inequality in $(Y, ||\cdot||)$:

$$||f(x) + g(x)|| \leq ||f(x)|| + ||g(x)|| \leq ||f||_\infty + ||g||_\infty.$$

Thus $||f||_\infty + ||g||_\infty$ is an upper bound for the set $\{||f(x) + g(x)|| : x \in H\}$, hence

$$\sup_{x\in H} ||f(x) + g(x)|| = ||f + g||_\infty \leq ||f||_\infty + ||g||_\infty.$$

□

Proposition 7.2. *We denote by $C(H;Y)$ the set $B(H;Y)$ where the functions are also continuous. Assume that $(Y, ||\cdot||)$ is a Banach space. Then $(C(H;Y), ||\cdot||_\infty)$ is a Banach space, too.*

Proof. We need to prove that every Cauchy sequence is convergent. Assume that $\{f_n\}_{n\geq 1} \subset C(H;Y)$ is Cauchy, i.e. for every $\epsilon > 0$ one can find $N_C(\epsilon) > 0$ such that $||f_p - f_q||_\infty < \epsilon$ if $p, q > N_C(\epsilon)$. We have to show that the sequence has a limit f which belongs to $C(H;Y)$.

We first construct f. For every $x_0 \in H$ we consider the sequence $\{f_n(x_0)\}_{n\geq 1} \subset Y$. Note the difference between $\{f_n(x_0)\}_{n\geq 1}$ (a sequence of vectors from Y) and $\{f_n\}_{n\geq 1}$ (a sequence of functions from $C(H;Y)$). It is easy to see that $\{f_n(x_0)\}_{n\geq 1}$ is Cauchy in Y (exercise), and because Y is complete, then $\{f_n(x_0)\}_{n\geq 1}$ has a limit in Y. We denote it with $f(x_0)$.

Second, we prove the "uniform convergence" part, or the convergence in the norm $||\cdot||_\infty$. More precisely, it means that for every $\epsilon > 0$ we must construct $N_1(\epsilon) > 0$ so that:

$$\sup_{x\in H} ||f(x) - f_n(x)|| < \epsilon \quad \text{whenever} \quad n > N_1(\epsilon). \tag{7.1}$$

In order to do that, take an arbitrary point $x \in H$. For every $p, n \geq 1$ we have

$$\begin{aligned} ||f(x) - f_n(x)|| &\leq ||f(x) - f_p(x)|| + ||f_p(x) - f_n(x)|| \\ &\leq ||f(x) - f_p(x)|| + ||f_p - f_n||_\infty. \end{aligned} \tag{7.2}$$

If we choose $n, p > N_C(\epsilon/2)$, then we have $||f_p - f_n||_\infty < \epsilon/2$ and

$$||f(x) - f_n(x)|| \leq ||f(x) - f_p(x)|| + \epsilon/2, \quad n, p > N_C(\epsilon/2).$$

But the above left hand side does not depend on p, thus if we take $p \to \infty$ on the right hand side, we get:

$$||f(x) - f_n(x)|| \leq \epsilon/2 < \epsilon, \quad n > N_C(\epsilon/2). \tag{7.3}$$

Note that this inequality holds true *for every* x. This means that $\epsilon/2$ is an upper bound for the set $\{||f(x) - f_n(x)|| : x \in H\}$, hence (7.1) holds true with $N_1(\epsilon) = N_C(\epsilon/2)$.

Third, we must prove that f is a continuous function on H. Fix some point $a \in H$. Choose $\epsilon > 0$. Since $\lim_{n\to\infty} f_n(a) = f(a)$, we can find $N_2(\epsilon, a) > 0$ such that $||f_n(a) - f(a)|| < \epsilon/3$ whenever $n > N_2$. We define $n_1 := \max\{N_1(\epsilon/3) + 1, N_C(\epsilon/3) + 1, N_2 + 1\}$. Because f_{n_1} is continuous at a, we can find $\delta(\epsilon, a) > 0$ so that for every $x \in H$ with $d(x, a) < \delta$ we have $||f_{n_1}(x) - f_{n_1}(a)|| < \epsilon/3$. Thus

$$\begin{aligned} ||f(x) - f(a)|| &\leq ||f(x) - f_{n_1}(x)|| + ||f_{n_1}(x) - f_{n_1}(a)|| + ||f_{n_1}(a) - f(a)|| \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon, \end{aligned} \tag{7.4}$$

We used (7.1) in order to replace the first and the third term with $\epsilon/3$, and continuity of f_{n_1} at a for the second term. Since a is arbitrary, we can conclude that f is continuous on H, thus belongs to $C(H;Y)$. Therefore we can rewrite (7.1) as:

$$||f - f_n||_\infty < \epsilon \quad \text{whenever} \quad n > N_1(\epsilon), \tag{7.5}$$

and the proof is over. □

Remark 7.3. *The "ordinary" convergence in the functional space $(C(H;Y), ||\cdot||_\infty)$ (given in (7.5)) is nothing but the uniform convergence of a sequence of functions defined on the set H (see (7.1)). One can find more details in Wade, exercise 10.6.6 in Chapter 10.6 (page 376).*

8 Compactness in $(C(H;\mathbb{R}^n), ||\cdot||_\infty)$

In this section we assume that H is a compact set in (X, d). This extra-condition automatically implies that $||f||_\infty < \infty$ for all continuous functions, because for every continuous map f we have that $H \ni x \to ||f(x)|| \in \mathbb{R}_+$ is a continuous real valued function, defined on a compact set. Then Theorem 10.63 in Wade says that we can find $x_M \in H$ such that $\sup_{x\in H} ||f(x)|| = ||f(x_M)|| < \infty$.

We here are interested in finding some sufficient conditions for a subset of $(C(H;\mathbb{R}^n), ||\cdot||_\infty)$, $n \geq 1$, in order to be compact. (We know that in the Euclidian space $(\mathbb{R}^n, ||\cdot||)$ a set is compact if and only if it is bounded and closed; this is the Heine-Borel theorem. These conditions are not sufficient here.)

Definition 8.1. *We say that $f : H \to Y$ is uniformly continuous if for every $\epsilon > 0$, we can find $\delta(f, \epsilon) > 0$ such that for all points $x, y \in H$ which fulfill $d(x, y) \leq \delta(f, \epsilon)$ we have that $||f(x) - f(y)|| \leq \epsilon$.*

Theorem 9.32 in Wade (Heine's theorem) shows that if H is compact, then $f : H \to \mathbb{R}^n$ is continuous if and only if it is uniformly continuous.

Definition 8.2. *A family of functions $K \subset C(H; \mathbb{R}^n)$ is called* equibounded *if there exists a constant $M_K < \infty$ such that*

$$\sup_{x \in H} ||f(x)|| = ||f||_\infty \leq M_K, \quad \forall f \in K. \tag{8.1}$$

Definition 8.3. *A family of functions $K \subset C(H; \mathbb{R}^n)$ is called* uniformly equicontinuous *if for every $\epsilon > 0$ there exists $\delta(\epsilon) > 0$, such that for every $f \in K$ and for every pair of points $x, y \in H$ which obey $d(x, y) \leq \delta(\epsilon)$, one has that $||f(x) - f(y)|| \leq \epsilon$. In other words, (see Definition 8.1)*

$$\inf_{f \in K} \delta(f, \epsilon) = \delta_K(\epsilon) > 0. \tag{8.2}$$

Definition 8.4. *A subset Z of a metric space (M, d) is called dense in M if every point $x \in M$ is the limit of a sequence $\{x_n\}_{n \geq 1} \subseteq Z$. A set Z is called countable if there exists a map $j : Z \to \mathbb{N}$ which is injective. A metric space is called separable if it contains a countable dense subset.*

Theorem 8.5. *(*Arzela-Ascoli*). Let (X, d) be a metric space, and let H be a compact subset of X. Assume that $Z \subset H$ is countable and dense in H (i.e. (H, d) is separable). Denote by $K \subset C(H; \mathbb{R}^n)$ the family of* all *functions which are equibounded by some M_K and uniformly equicontinuous with some δ_K (ækvibegrænset og uniformt ækvikontinuert). Then K is sequentially compact (følgekompakt) and thus compact. The closure in $C(H; \mathbb{R}^n)$ of any subset of K is also compact.*

Proof. We will show that given an arbitrary sequence of functions $\{f_n\}_{n \geq 1} \subset K$, one can always find a subsequence which converges to a "point" in K (note that a point in K means a function defined on H; we denote this "point" with f). This would prove that K is sequentially compact.

Because the dense set Z is countable, we can represent it in the following way:

$$Z = \{z_1, z_2, z_3, ...\}.$$

The sequence $\{f_n(z_1)\}_{n \geq 1} \subset \mathbb{R}^n$ is bounded because we have $||f_n(z_1)|| \leq M_K$ for every n, see (8.1). The Bolzano-Weierstrass theorem allows us to find a subsequence $\{f_{n_1}(z_1)\}_{n_1 \geq 1} \subset \mathbb{R}^n$, which converges to a point in $\mathbb{R}^n$; we call this point with $f(z_1)$.

Now consider the sequence $\{f_{n_1}(z_2)\}_{n_1 \geq 1} \subset \mathbb{R}^n$. This sequence is also bounded, thus we can find a second subsequence

$$\{f_{n_2}(z_2)\}_{n_2 \geq 1} \subseteq \{f_{n_1}(z_2)\}_{n_1 \geq 1},$$

which converges to a point in $\mathbb{R}^n$; we call this point with $f(z_2)$. Note that the subsequence of *functions* $\{f_{n_2}\}_{n_2 \geq 1} \subseteq \{f_{n_1}\}_{n_1 \geq 1}$ converges pointwise in both z_1 and z_2.

We can continue this procedure and obtain a subsequence of functions $\{f_{n_p}\}_{n_p \geq 1}$ where

$$\{f_{n_p}\}_{n_p \geq 1} \subseteq \{f_{n_{p-1}}\}_{n_{p-1} \geq 1} \subseteq ... \subseteq \{f_n\}_{n \geq 1}$$

and $\{f_{n_p}\}_{n_p \geq 1}$ converges pointwise in all the points $\{z_1, ..., z_p\}$ towards the limits $\{f(z_1), ..., f(z_p)\}$. More precisely, for every $\epsilon > 0$, there exists $N(p, \epsilon) > 1$ such that

$$||f_{n_p}(z_k) - f(z_k)|| < \epsilon, \quad \text{whenever} \quad n_p > N(p, \epsilon), \ k \in \{1, ..., p\}. \tag{8.3}$$

Introduce the notation $N_p := N(p, 1/p) + 1$. Then we have the important estimate:

$$||f_{N_p}(z_k) - f(z_k)|| < 1/p, \quad \text{whenever} \quad k \in \{1, ..., p\}. \tag{8.4}$$

This means that we have constructed a "diagonal subsequence" $\{f_{N_p}\}_{p\geq 1}$ having the property that $\{f_{N_p}(z_k)\}_{p\geq 1} \subset \mathbb{R}^n$ is convergent for every fixed k, and we denote the limits with:

$$\lim_{p\to\infty} f_{N_p}(z_k) = f(z_k), \quad k \text{ fixed}. \tag{8.5}$$

This is the same thing as to say that the sequence $\{f_{N_p}\}_{p\geq 1}$ converges pointwise on Z:

$$\lim_{p\to\infty} f_{N_p}(z) = f(z), \quad \forall z \in Z. \tag{8.6}$$

In the next lemma we will show that the sequence $\{f_{N_p}\}_{p\geq 1}$ is a Cauchy sequence in $C(H;\mathbb{R}^n)$. Let us now assume that this holds true, and let us prove the Arzela-Ascoli theorem.

If this sequence is Cauchy, then according to Proposition 7.2 it will have a limit in $C(H;\mathbb{R}^n)$, which we denote by F. But then F is continuous on H and equal to $f(z)$ for every $z \in Z$. The only thing remained to prove is that $F \in K$, i.e. to verify that F verifies (8.1) and (8.2).

First, (8.1) follows from:

$$||F(x)|| = \lim_{p\to\infty} ||f_{N_p}(x)||, \quad ||f_{N_p}(x)|| \leq M_K, \quad x \in H,$$

and (8.2) from:

$$\begin{aligned} ||F(x) - F(y)|| &= \lim_{p\to\infty} ||f_{N_p}(x) - f_{N_p}(y)||, \\ ||f_{N_p}(x) - f_{N_p}(y)|| &\leq \epsilon \quad \text{whenever} \quad d(x,y) \leq \delta(\epsilon). \end{aligned} \tag{8.7}$$

Thus $F \in K$, and the theorem is proved. Hence the only remaining technical ingredient is the following lemma:

Lemma 8.6. *For every $\epsilon' > 0$, there exists $N_C(\epsilon') > 0$ such that for every $p, q > N_C(\epsilon')$ we have*

$$\sup_{x\in H} ||f_{N_p}(x) - f_{N_q}(x)|| = ||f_{N_p} - f_{N_q}||_\infty < \epsilon'.$$

Proof. Choose $0 < \epsilon < \epsilon'$. Consider $\delta_K(\epsilon/3)$ as defined in (8.2).

Let us now show that

$$\{B_{\delta_K(\epsilon/3)/2}(z_j) : \ z_j \in Z\}$$

is an open covering of H. First, because Z is dense in H, then for every point $x \in H$ there exists a sequence $\{x_m\}_{m\geq 1} \subset Z$ such that $x_m \to x$. Second, we may find $x_M \in Z$ such that $B_{\delta_K(\epsilon/3)/3}(x) \subset B_{\delta_K(\epsilon/3)/2}(x_M)$ provided $d(x, x_M) < \delta_K(\epsilon/3)/6$ (exercise). We can write:

$$H \subset \bigcup_{x\in H} B_{\delta_K(\epsilon/3)/3}(x) \subset \bigcup_{k=1}^{\infty} B_{\delta_K(\epsilon/3)/2}(z_k).$$

Because H is compact, we can extract a finite open subcovering:

$$H \subseteq \bigcup_{l=1}^{m(\epsilon)} B_{\delta_K(\epsilon/3)/2}(z_{k_l}). \tag{8.8}$$

Fix an arbitrary point $x \in H$. We can find some $l \in \{1, ..., m(\epsilon)\}$ such that $x \in B_{\delta_K(\epsilon/3)/2}(z_{k_l})$. We can write:

$$\begin{aligned} & ||f_{N_p}(x) - f_{N_q}(x)|| \\ \leq \ & ||f_{N_p}(x) - f_{N_p}(z_{k_l})|| + ||f_{N_p}(z_{k_l}) - f_{N_q}(z_{k_l})|| + ||f_{N_q}(z_{k_l}) - f_{N_q}(x)||. \end{aligned} \tag{8.9}$$

Because K is uniformly equicontinuous, and because $d(x, z_{k_l}) < \delta_K(\epsilon/3)$, then the first and third term in the right hand side of (8.9) are less than $\epsilon/3$ (see Definition 8.3), uniformly in p and q. Thus

$$||f_{N_p}(x) - f_{N_q}(x)|| \leq 2\epsilon/3 + ||f_{N_p}(z_{k_l}) - f_{N_q}(z_{k_l})||, \quad \forall p, q \geq 1. \tag{8.10}$$

Note the very important thing that there only are a finite number of points of the type z_{k_l}, i.e. $m(\epsilon)$ of them. Hence (8.5) implies that the $m(\epsilon)$ sequences $\{f_{N_r}(z_{k_l})\}_{r\geq 1} \subset \mathbb{R}^n$ are all Cauchy at the same time; we can thus find a large enough index $N_1(\epsilon/3)$ such that if $p, q > N_1(\epsilon/3)$ then

$$||f_{N_p}(z_{k_l}) - f_{N_q}(z_{k_l})|| < \epsilon/3, \quad 1 \leq l \leq m(\epsilon).$$

Use this in (8.10) and obtain:

$$||f_{N_p}(x) - f_{N_q}(x)|| < \epsilon, \quad \text{whenever} \quad p, q \geq N_C(\epsilon') := N_1(\epsilon/3). \tag{8.11}$$

Because $N_C(\epsilon')$ is independent of x, we can write

$$\sup_{x\in H} ||f_{N_p}(x) - f_{N_q}(x)|| \leq \epsilon < \epsilon', \quad p, q \geq N_C(\epsilon')$$

and the lemma is proved, and so is the theorem. □

9 Hausdorff's Maximality Theorem and Zorn's lemma

A set S is *partially ordered* if there exists an order relation $\leq$ which is reflexive ($x \leq x$ for all x), antisymmetric (if $x \leq y$ and $y \leq x$ then $x = y$) and transitive ($x \leq y$ and $y \leq z$ implies $x \leq z$). If $x \leq y$ and $x \neq y$, then we write $x < y$ or $y > x$.

A *chain* in S is a subset C which is totally ordered, i.e. whose any two elements are comparable: For every $x, y \in C$ either $x \leq y$ or $y \leq x$. A chain C has un *upper bound* if there exists $y \in S$ such that $x \leq y$ for all $x \in C$.

An element $m \in S$ is called *maximal* if there is no other $x \in S$ such that $m < x$. This does not mean that m is the largest element, which would be an element $M \in S$ such that $x \leq M$ for every $x \in S$.

9.1 Collection of sets ordered by set inclusion

Let $\mathbb{F}$ be any collection of sets, $\mathbb{F}$ being partially ordered by set inclusion. A *rope* Φ in $\mathbb{F}$ is a subcollection of sets such that for every $A, B \in \Phi$, we either have $A \subset B$ or $B \subset A$. The *hat* of such a rope is by definition the union of all its elements: $\hat{\Phi} = \cup_{A\in\Phi} A$.

Lemma 9.1. *Consider a nonempty partially ordered set $(S, \leq)$. Define $\mathbb{F}_s$ to be the set of chains of S, ordered by set inclusion. Then $\mathbb{F}_s$ is not empty, and the hat of any rope of $\mathbb{F}_s$ is an element of $\mathbb{F}_s$ (i.e. it is a chain in S).*

Proof. $\mathbb{F}_s$ is not empty because for every $x \in S$, the set $\{x\}$ is a chain, thus in fact S can be identified with a subset of F_s. Now let us prove that the hat of any rope of F_s is a chain in S. Consider such a rope Φ. Take any two points x, y in $\hat{\Phi}$. The point x must belong to some chain C_1, while y belongs to some C_2. Moreover, both C_1 and C_2 belong to Φ. From the definition of a rope it follows that either $C_1 \subset C_2$ or $C_2 \subset C_1$. But in that case x and y belong to the same chain, hence they must be comparable. □

We now define the 'succesion function'. For every chain $A \in \mathbb{F}_s$, we consider the set $A^* \subset S \setminus A$ which consist of elements x with the property that $A \cup \{x\}$ is also a chain, hence $A \cup \{x\} \in \mathbb{F}_s$. Assume that A^* is not empty. Using *the axiom of choice*, given A^* we can choose a representative $x^* \in A^*$. Then we define the 'succesion operation' $g_s : \mathbb{F}_s \to \mathbb{F}_s$ given either by $g_s(A) = A$ if A^* is empty, or by $g_s(A) = A \cup \{x^*\}$ if A^* is not empty. A chain for which A^* is empty is called *maximal*. We see that $g_s(A)$ contains at most one extra-element compared with A. We are now able to formulate the Hausdorff maximality theorem:

Theorem 9.2. *Let $(S, \leq)$ be a nonempty partially ordered set. Then there exists a maximal chain A_m, i.e. $g_s(A_m) = A_m$.*

9.2 Proof of Theorem 9.2

We formulate and prove a more general result, given as a proposition:

Proposition 9.3. *Let $\mathbb{F}$ be a nonempty collection of subsets of an arbitrary set S. Suppose that $\mathbb{F}$ is partially ordered by set inclusion, and for every rope $\Phi \in \mathbb{F}$ we have that $\hat{\Phi} \in \mathbb{F}$. Suppose $g : \mathbb{F} \to \mathbb{F}$ is a function such that $A \subset g(A)$ and $g(A) \setminus A$ contains at most one element of S. Then there exists an element A_m of $\mathbb{F}$ such that $g(A_m) = A_m$.*

Proof. We see that the only condition on $\mathbb{F}$ is to contain the hats of all its ropes; if S is a partially ordered set and $\mathbb{F} = \mathbb{F}_s$, then this fact was proved in Lemma 9.1. Thus Proposition 9.3 proves the theorem at once.

Fix a set $A_0 \in \mathbb{F}$. A subcollection of sets $T \in \mathbb{F}$ is called a *tower* if the following three conditions are fulfilled:
P_1: $A_0 \subset A$ if $A \in T$;
P_2: If a rope Φ is included in T, then $\hat{\Phi} \in T$;
P_3: If $A \in T$, then $g(A) \in T$.

Let us first prove that there exist nonempty towers. Consider $T_{max} := \{A \in \mathbb{F} : \ A_0 \subset A\}$. Then clearly P_1 is satisfied. If Φ is a rope in T_{max} then $A_0 \subset B$ for all $B \in \Phi$, thus $A_0 \in \hat{\Phi}$ and $\hat{\Phi} \in T_{max}$, satisfying P_2. Finally, if $A \in T_{max}$ then $A_0 \subset A \subset g(A) \in T_{max}$ and P_3 holds.

Let us define T_{min} to be the intersection of all possible towers; it is easy to see that T_{min} is a tower, nonempty since it contains A_0. It is important to note that if we can prove that a subcollection T' of T_{min} is a tower, then $T' = T_{min}$.

Now let us assume that we can prove that T_{min} is also a rope. Being included in itself (a tower), P_2 implies that $\hat{T}_{min} \in T_{min}$. Then P_3 says that $g(\hat{T}_{min}) \in T_{min}$, thus necessarily $g(\hat{T}_{min}) \subset \hat{T}_{min}$. In this case, $A_m = \hat{T}_{min}$, and the proposition would be proved.

Hence the only thing we miss is to show that T_{min} is a rope, i.e. for every $A, B \in T_{min}$ we either have $A \subset B$ or $B \subset A$. Consider the subcollection $\Gamma \subset T_{min}$ given by

$$\Gamma := \{A \in T_{min} : \ \forall B \in T_{min}, \text{ either } A \subset B \text{ or } B \subset A\}.$$

In other words, Γ is the largest totally ordered subset of T_{min}. The strategy is to show that Γ is a tower, thus it will equal T_{min} and hence T_{min} would be totally ordered, thus a rope.

Let us verify P_1 for Γ. Since A_0 is a subset of all elements of T_{min}, it will be included in all elements of Γ. It also shows that $A_0 \in \Gamma$.

Let us verify P_2 for Γ. Consider a rope $\Phi \subset \Gamma$. Fix an arbitrary $B \in T_{min}$. If for all $A \in \Phi$ we have that $A \subset B$, it follows that $\hat{\Phi} \subset B$. If there exists some $A \in \Phi$ such that $B \subset A$, then $B \subset \hat{\Phi}$. Thus $\hat{\Phi} \in \Gamma$.

Let us verify P_3 for Γ. This is more complicated than for the previous two properties. It boils down to showing that for every $A \in \Gamma$ we have $g(A) \in \Gamma$. In other words, we need to show that for every $B \in T_{min}$ we either have $g(A) \subset B$ or $B \subset g(A)$. Since $A \in \Gamma$, the set A can be compared with B. If $B \subset A$, then clearly $B \subset g(A)$. But if $A \subset B$, we cannot automatically conclude that $g(A) \subset B$. For this we need the next lemma:

Lemma 9.4. *Fix $A \in \Gamma$ and define T_A to be the set of those $B \in T_{min}$ for which either $B \subset A$ or $g(A) \subset B$. Then T_A is a tower and $T_A = T_{min}$. Thus for every $B \in T_{min}$ for which $A \subset B$, we have $g(A) \subset B$.*

Proof. The property P_1 is easily verified. Now consider a rope $\Phi \subset T_A$. If for every $D \in \Phi$ we have that $D \subset A$, then $\hat{\Phi} \subset A$. If there exists some $D \in \Phi$ such that $g(A) \subset D$, then $g(A) \subset \hat{\Phi}$. It follows that $\hat{\Phi} \in T_A$ and P_2 is verified. We need to verify P_3. Take $B \in T_A$ and let us show that either $g(B) \subset A$ or $g(A) \subset g(B)$. Since $B \in T_A$ then either $g(A) \subset B \subset g(B)$, or $B \subset A$; if $B = A$ then again $g(A) \subset g(B)$. Assume that B is a proper subset of A. Because $A \in \Gamma$ it can be compared with $g(B)$. If $g(B) \subset A$ we are done. The other possibility is $A \subset g(B)$. If $g(B) = A$ we are done, thus assume that A is a proper subset of $g(B)$; but this is impossible, because in this case $g(B) \setminus B$ would contain at least two elements. Thus $g(B) \in T_A$ and P_3 is satisfied. Thus T_A is a tower and equals T_{min}.

□

We can now finish the verification of P_3 for Γ. Remember that we had $A \in \Gamma$ and B an arbitrary element of T_{min}. The goal is to show that either $B \subset g(A)$ or $g(A) \subset B$. Since $A \in \Gamma$, then either $B \subset A$ (and then we are done since $B \subset g(A)$), or $A \subset B$. Then Lemma 9.4 implies that $g(A) \subset B$ and P_3 is verified. It follows that Γ is a tower, thus it equals T_{min}. The proof of the proposition is over.

□

9.3 Zorn's Lemma

Theorem 9.5. *Let S be a partially ordered set in which every chain has an upper bound. Then S has at least one maximal element.*

Proof. Assume that S is a partially ordered set, where every chain has an upper bound. According to the Hausdorff maximality theorem, there exists a maximal chain $C_{max} \subseteq S$. The hypothesis implies that C_{max} has an upper bound $x \in S$, and $C_{max} \cup \{x\}$ is another chain in S. But C_{max} is maximal, therefore $x \in C_{max}$ and x must be this chain's largest element. Finally, x is a maximal element in S, because if there exists some $x < y$ we can consider $C_{max} \cup \{y\}$ which would contradict the maximality of C_{max}. The proof of the theorem is over. □

10 The Hahn-Banach Theorem

Let $(V, ||\cdot||_v)$ be a complex normed vector space. We say that $u : V \mapsto \mathbb{R}$ is a real bounded linear functional if u is continuous, $u(x+y) = u(x) + u(y)$ for all $x, y \in V$, $u(\alpha x) = \alpha u(x)$ for every $\alpha \in \mathbb{R}$ and $x \in V$.

Lemma 10.1. *Let $f : V \mapsto \mathbb{C}$ be a bounded complex linear functional. Then $u(y) := \mathrm{Re}(f(y))$ defines a real bounded linear functional. Conversely, if v is a real bounded linear functional, then $\phi(y) := v(y) - iv(iy)$ defines a bounded complex linear functional. In both cases, the norms are preserved.*

Proof. Given f, let us consider $u(y) = \mathrm{Re}(f(y))$ for all $y \in V$. Clearly, u is linear and real homogeneous. Since $f(ix) = if(x)$ we must have $f(ix) = u(ix) + i\mathrm{Im}(f(ix)) = if(x) = iu(x) - \mathrm{Im}(f(x))$, which implies that $\mathrm{Im}(f(x)) = -u(ix)$ for all x. Thus we must have $f(x) = u(x) - iu(ix)$. Moreover, $|u(x)| \leq |f(x)| \leq ||f||$ for all x of norm one, thus $||u|| \leq ||f||$. For a given x of norm one we can write $f(x) = e^{i\arg(f(x))}|f(x)|$, or $|f(x)| = f(e^{-i\arg(f(x))}x) = u(e^{-i\arg(f(x))}x) \leq ||u||$. Thus $||u|| = ||f||$.

Now given v, let us consider $\phi(y) = v(y) - iv(iy)$. CLearly, ϕ is linear and real homogeneous. We only need to check that $\phi(ix) = i\phi(x)$. Indeed, using the definition, we have $\phi(ix) = v(ix) - iv(-x) = iv(x) + v(ix) = i\phi(x)$. □

Theorem 10.2. *(Hahn-Banach) Let $M \subset V$ be a linear subspace and $f : M \mapsto \mathbb{C}$ be a bounded complex linear functional. Then there exists a bounded complex linear functional $F : V \mapsto \mathbb{C}$ such that $F(x) = f(x)$ if $x \in M$ and $||F|| = ||f||$.*

Proof. From the above lemma we see that it is enough to extend the real linear functional $u(x) = \mathrm{Re}(f(x))$ from M to V, i.e. to find $U : V \mapsto \mathbb{R}$ a real linear functional such that $U(x) = u(x)$ if $x \in M$ and $||U|| = ||u||$, and at the end to define $F(x) = U(x) - iU(ix)$. If $f = 0$ on M then we can choose $F = 0$. Otherwise, we may assume without loss of generality that $||f|| = ||u|| = 1$.

Let $\tilde{x}$ be a vector not belonging to M. We can form the set

$$\tilde{M} := \{x + (\lambda + i\mu)\tilde{x} : \ x \in M \text{ and } \lambda, \mu \in \mathbb{R}\}.$$

Clearly, $\tilde{M}$ is a complex subspace of V. When we extend u from M to $\tilde{M}$ we do it in two steps: First we keep $\mu = 0$ and extend u to $\tilde{M}_0 := \{x + \lambda\tilde{x} : \ x \in M \text{ and } \lambda \in \mathbb{R}\}$, and then to $\tilde{M} = \{x + (i\tilde{x})\mu : \ x \in \tilde{M}_0 \text{ and } \mu \in \mathbb{R}\}$. In both cases it boils down to choosing some real values for $\tilde{u}(\tilde{x})$ and $\tilde{u}(i\tilde{x})$, choices which have to preserve linearity, real homogeneity and the unit norm of u.

For $y = x + \lambda\tilde{x} \in \tilde{M}_0$ we define $\tilde{u}(y) := u(x) + \lambda A$, where $A \in \mathbb{R}$ is to be chosen later. First, linearity and real homogeneity are satisfied for any A. Second, we need to insure that if $\lambda \neq 0$ we still have $|\tilde{u}(x + \lambda\tilde{x})| \leq ||x + \lambda\tilde{x}||_v$, or equivalently $|\tilde{u}(x' + \tilde{x})| \leq ||x' + \tilde{x}||_v$, for all $x' \in M$. This is the same as:

$$-||x' + \tilde{x}||_v \leq \tilde{u}(x' + \tilde{x}) = u(x') + A \leq ||x' + \tilde{x}||_v, \quad \forall x' \in M,$$

or:

$$-||x' + \tilde{x}||_v - u(x') \leq A \leq ||x' + \tilde{x}||_v - u(x'), \quad \forall x' \in M.$$

A sufficient condition for the existence of such a number A is to have:

$$-||x' + \tilde{x}||_v - u(x') \leq ||x + \tilde{x}||_v - u(x), \quad \forall x, x' \in M.$$

But this is the same as

$$u(x) - u(x') \,\{= u(x - x') \leq ||x - x'||_v = ||x + \tilde{x} - (x' + \tilde{x})||_v\} \leq ||x + \tilde{x}||_v + ||x' + \tilde{x}||_v$$

and we are done.

In a completely similar way we can extend $\tilde{u}$ from $\tilde{M}_0$ to $\tilde{M}$, where the role of M is played by $\tilde{M}_0$, $\tilde{x}$ is replaced by $i\tilde{x}$ and λ by μ. Then we define $\tilde{f}(x) = \tilde{u}(x) - i\tilde{u}(ix)$ for every $x \in \tilde{M}$, which is a linear, norm preserving extension of f from M to $\tilde{M}$.

Now let us define the set S whose elements are pairs of the form (M', f') where M' is a complex linear subspace of V containing M, while f' is a norm preserving linear extension of f. We can introduce an order relation $\leq$ on S, where $(M_1, f_1) \leq (M_2, f_2)$ means that $M \subset M_1 \subset M_2$, while f_2 extends f_1 and both are extensions of f. Hausdorff's maximality theorem 9.2 implies the existence of a maximal chain A_m. The set of subspaces M' for which $(M', f') \in A_m$ is totally ordered with respect to set inclusion. Define $\widehat{M} := \cup_{(M', f') \in A_m} M'$, which is a linear subspace of V due to the total ordering of M''s. Define $F : \widehat{M} \mapsto \mathbb{C}$ such that $F(x) = f'(x)$ if $x \in M' \subset \widehat{M}$; this is well-defined because if also $x \in M''$, then either $M' \subset M''$ or $M'' \subset M'$, and $f'(x) = f''(x)$ because they coincide on the smaller set.

Thus $\widehat{M}$ is a complex subspace of V, $M \subset \widehat{M}$, and F is a linear, norm preserving extension of f to $\widehat{M}$. But now we claim that $\widehat{M}$ must be equal to V; otherwise, we could find another extension of F by enlarging $\widehat{M}$ in the same way as we did in the beginning of this proof. But this would contradict the maximality of A_m. The proof is over.

□

11 The completion of a normed space

Lemma 11.1. *Let $(V, ||\cdot||)$ be a normed space, and denote by V^* the linear space of all maps $f : V \mapsto \mathbb{C}$ which are linear and continuous (i.e. linear functionals). Then V^* can be organized as a Banach space.*

Proof. The following three statements are equivalent: (a) A linear map is Lipschitz continuous on V, (b) A linear map is continuous at $x = 0$, and (c) A linear map is bounded. Clearly, (a) implies (b). Assuming that (b) holds, then for every $\epsilon_0 > 0$ there exists $\delta_0 > 0$ such that $|f(x)| < \epsilon_0$ if $||x|| < \delta_0$. In order words, if $y \neq 0$:

$$|f(y)| = \frac{2||y||}{\delta_0} \left| f\left(\frac{\delta_0}{2||y||} y\right) \right| \leq \frac{2\epsilon_0}{\delta_0} ||y||,$$

which implies (c). Finally, (c) implies (a) by using the linearity.

For every $f \in V^*$ we set $||f||_* := \inf\{C \geq 0 : \ |f(x)| \leq C||x||, \ \forall x \in V\} = \sup_{||x||=1} |f(x)|$, which defines a norm on V^*. Let us prove that $(V^*, || \cdot ||_*)$ is a Banach space, i.e. every Cauchy sequence is convergent. Let $\{f_n\}_{n\geq 0} \subset V^*$ be such a Cauchy sequence. For a given $x \in V$, the sequence of complex numbers $\{f_n(x)\}_{n\geq 0} \subset \mathbb{C}$ is Cauchy, thus has a limit which we denote with $f(x)$. Using the linearity of each f_n one can prove that f is linear. Moreover, because $\{f_n\}_{n\geq 0} \subset V^*$ is Cauchy, there exists some $C \geq 0$ such that $||f_n||_* \leq C$ for all $n \geq 0$, and for every x with $||x|| = 1$ we have:

$$|f(x)| = \lim_{n\to\infty} |f_n(x)| \leq \limsup_{n\to\infty} ||f_n||_* \leq C.$$

Thus f is bounded, too. Finally, fix some $x \in V$ with $||x|| = 1$. We have:

$$|f(x) - f_q(x)| \leq |f(x) - f_p(x)| + |f_p(x) - f_q(x)| \leq |f(x) - f_p(x)| + ||f_p - f_q||_*$$

which leads to:

$$|f(x) - f_q(x)| \leq \limsup_{p\to\infty} \{|f(x) - f_p(x)| + ||f_p - f_q||_*\} = \limsup_{p\to\infty} ||f_p - f_q||_*,$$

or $||f - f_q||_* \leq \limsup_{p\to\infty} ||f_p - f_q||_*$ for all $q \geq 0$. But because the sequence is Cauchy, the right hand side can be made smaller than any ϵ if q is large enough, and the proof is over. □

Theorem 11.2. *Let $(A, || \cdot ||_a)$ be a normed space. Then there exists a Banach space B with a norm $|| \cdot ||_b$ and a linear mapping $I : A \to B$ such that*

$$||I(x)||_b = ||x||_a, \quad \forall x \in A,$$

and $I(A)$ is dense in B with respect to $|| \cdot ||_b$. We call $(B, || \cdot ||_b)$ the completion of $(A, || \cdot ||_a)$.

Proof. Given A we can construct A^* as in the previous lemma. Then we repeat this construction one more time, where now V is the normed space $(A^*, || \cdot ||_*)$. In this way we obtain $V^* = A^{**}$ as the Banach space containing all linear and bounded maps $g : A^* \mapsto \mathbb{C}$. Let us prove that A can be identified with a proper subspace of A^{**}.

For every any x in A we define the map $g_x : A^* \mapsto \mathbb{C}$ given by $g_x(f) = f(x)$ for every element $f \in A^*$. Then we have

$$g_{\alpha x+\beta y}(f) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha g_x(f) + \beta g_y(f), \quad \forall x, y \in A, \quad \forall \alpha, \beta \in \mathbb{C},$$

which shows that $g_{\alpha x+\beta y} = \alpha g_x + \beta g_y$. Clearly, if $x = 0$ then $g_0 = 0$. The only thing remaining to be proved is that $||g_x||_{**} = ||x||_a$ for every $x \neq 0$.

First, for every $f \in A^*$ we have $|g_x(f)| = |f(x)| \leq ||f||_*||x||_a$, which means that $||g_x||_{**} \leq ||x||_a$.

Second, denote by $M = \{\lambda x : \ \lambda \in \mathbb{C}\}$ the one dimensional linear subspace of A generated by x. For every $y \in M$ there exists a unique $\lambda_y \in \mathbb{C}$ such that $y = \lambda_y x$. Define $f_M(y) = \lambda_y ||x||_a^2$. Clearly, f_M is linear and $||f_M||_* = ||x||_a$. The Hahn-Banach theorem provides us with a norm preserving linear extension F defined on the whole A. Then we have $g_x(F) = F(x) = ||x||_a^2 = ||x||_a||F||_*$, which shows that $||x||_a \leq \sup_{||f||_*=1} |g_x(f)| = ||g_x||_{**}$.

In this way we have identified an injection map $I : A \mapsto A^{**}$ where $I(x) = g_x$, $I(A)$ is a linear subspace of A^{**} and $||I(x)||_{**} = ||x||_a$. Now the Banach space B we are looking for is nothing but the closure $\overline{I(A)}$ in $(A^{**}, || \cdot ||_{**})$.

□

12 Baire's Category Theorem

Denote the open ball of radius ϵ and centred at x by $B_\epsilon(x) := \{y \in B : \ ||y - x|| < \epsilon\}$. The complementary in B of a set $S \subseteq B$ is denoted by S^c.

Theorem 12.1. *Consider a Banach space B, and a sequence of closed sets $\{S_n\}_{n\geq 1}$ such that*

$$B = \bigcup_{n\geq 1} S_n. \tag{12.12}$$

Then there exists at least one set S_n with non-empty interior.

Proof. Assume the contrary, that is each S_n has an empty interior. One can re-state this in a more formal way: for every $x \in S_n$, and for every $\epsilon > 0$, we have:

$$B_\epsilon(x) \cap S_n^c \neq \emptyset, \quad \forall \epsilon > 0. \tag{12.13}$$

We can assume that all sets S_n are non-empty. We also have that $S_n^c \neq \emptyset$, since otherwise $S_n = B$ which would have a non-empty interior.

Let therefore x_1 be a point of S_1^c. Because S_1 is closed, we have that S_1^c is open, therefore there exists $\epsilon_1 > 0$ such that

$$B_{\epsilon_1}(x_1) \subset S_1^c. \tag{12.14}$$

Starting from x_1 and ϵ_1, we will inductively define two sequences $\{x_n\}_{n\geq 1} \subset B$ and $\{\epsilon_n\}_{n\geq 1} \subset \mathbb{R}_+$, having several properties. First, we need:

$$\epsilon_{n+1} < \frac{\epsilon_n}{3}, \quad n \geq 1. \tag{12.15}$$

Second, we need that:

$$B_{\epsilon_n}(x_n) \subset S_n^c, \quad n \geq 1, \tag{12.16}$$

and third:

$$||x_{n+1} - x_n|| < \frac{\epsilon_n}{3}, \quad n \geq 1. \tag{12.17}$$

Let us investigate the consequences of having such sequences, and we will later on prove their existence. First, (12.15) leads us to the estimate:

$$\epsilon_j < \frac{\epsilon_{j-1}}{3} < \ldots < \frac{\epsilon_n}{3^{j-n}}, \quad \forall\, j > n \geq 1. \tag{12.18}$$

In particular, $\epsilon_n < \epsilon_1/3^{n-1} \to 0$ when $n \to \infty$.

Second, we can prove that $\{x_n\}_{n\geq 1}$ is a Cauchy sequence, because for every $p \geq 1$ we can write

$$\begin{aligned} ||x_{n+p} - x_n|| &= ||\sum_{j=n}^{n+p-1} [x_{j+1} - x_j]|| \leq \sum_{j=n}^{n+p-1} ||x_{j+1} - x_j|| \\ &< \sum_{j=n}^{n+p-1} \epsilon_j/3 < \sum_{j=n}^{\infty} \epsilon_j/3 \\ &< \frac{\epsilon_n}{3} \sum_{k\geq 0} 3^{-k} = \frac{\epsilon_n}{2} \to 0, \quad n \to \infty. \end{aligned} \tag{12.19}$$

In the first line we used the triangle inequality, in the second line we used (12.17), and in the third line (12.18).

Because B is a Banach space, $\{x_n\}_{n\geq 1}$ is convergent and has a limit $x \in B$. But then we have (use the triangle inequality)

$$||x - x_n|| \leq ||x - x_{n+p}|| + ||x_{n+p} - x_n|| < ||x - x_{n+p}|| + \frac{\epsilon_n}{2}, \quad \forall\, p \geq 1.$$

Since $\lim_{p\to\infty} ||x - x_{n+p}|| = 0$, taking p to infinity in the above estimate gives us $||x - x_n|| < \epsilon_n$, or $x \in B_{\epsilon_n}(x_n)$, or $x \in S_n^c$ (see (12.16)), or $x \notin S_n$ for all n. But this contradicts (12.12).

Therefore the only remaining thing is the construction of our sequences. Let us first construct x_2 and ϵ_2.

(i). If $x_1 \in S_2^c$, then put $x_2 = x_1$. Then since S_2^c is open, we can find $\epsilon' > 0$ such that $B_{\epsilon'}(x_1) \subset S_2^c$. Now choose ϵ_2 to be the minimum between ϵ' and $\epsilon_1/4$. Clearly, (12.15) and (12.17) hold true for $n = 1$ (we here have $||x_1 - x_2|| = 0$), while (12.16) holds true for $n = 1, 2$.

(ii). If $x_1 \notin S_2^c$, then of course $x_1 \in S_2$. From (12.13) we have that for every $\epsilon' > 0$ we can find $y(\epsilon') \in B_{\epsilon'}(x_1) \cap S_2^c$, that is $||y(\epsilon') - x_1|| < \epsilon'$. Define $x_2 := y(\epsilon_1/4) \in S_2^c$. Because S_2^c is open, we can find $\epsilon'' > 0$ such that $B_{\epsilon''}(x_2) \subset S_2^c$. Finally define ϵ_2 as the minimum between ϵ'' and $\epsilon_1/4$. Then we have $\epsilon_2 < \epsilon_1/3$, $||x_2 - x_1|| < \epsilon' < \epsilon_1/3$, and $B_{\epsilon_2}(x_2) \subset S_2^c$.

The induction step from x_n and ϵ_n to x_{n+1} and ϵ_{n+1} is identical to the one from 1 to 2. The theorem is proved. □

13 The Open Mapping Theorem

Definition 13.1. *Assume that $(X_1, ||\cdot||_1)$ and $(X_2, ||\cdot||_2)$ are normed spaces. We say that the map $F : X_1 \mapsto X_2$ is open if for any open set $U \subset X_1$, the image $F(U) \subset X_2$ is also open.*

Lemma 13.2. *Let $F : X_1 \mapsto X_2$ be a linear, bounded map between two normed spaces. Then for all $r > 0$ we have:*

$$F(\overline{B_r(0_1)}) \subset \overline{F(B_r(0_1))}.$$

Proof. Let $x \in \overline{B_r(0_1)}$. There exists $\{x_n\}_{n\geq 1} \subset B_r(0_1)$ such that $||x_n - x|| \to 0$ when $n \to \infty$. Since F is bounded (thus continuous), we conclude that $F(x)$ is the limit of $y_n := F(x_n) \in F(B_r(0_1))$, hence $F(x) \in \overline{F(B_r(0_1))}$. □

Lemma 13.3. *Assume that $F : X_1 \mapsto X_2$ is a linear map between two normed spaces. If there exists $d > 0$ such that $B_d(0_2) \subset F(B_1(0_1))$, then F is open.*

Proof. Using the homogeneity of F, we can show that $B_d(0_2) \subset F(B_1(0_1))$ implies that for every $c > 0$ we have:

$$B_{cd}(0_2) \subset F(B_c(0_1)), \quad c > 0. \tag{13.20}$$

Indeed, if $y \in B_{cd}(0_2)$ then $c^{-1}y \in B_d(0_2)$. There exists $x \in B_1(0_1)$ such that $c^{-1}y = F(x)$, or $y = cF(x) = F(cx)$. But $cx \in B_c(0_1)$, thus $y \in F(B_c(0_1))$.

Now let U be open in X_1. We have to prove that $V := F(U)$ is open in X_2. Choose an arbitrary $y_0 \in V$. Then there exists $x_0 \in U$ such that $F(x_0) = y_0 \in V$.

If M and N are two subsets of the same vector space X, then we denote:

$$M + N := \{x + y \in X : \quad \forall[x \in M,\ y \in N]\}.$$

With this notation and using the linearity of F we can easily prove that for every $\epsilon > 0$ we have:

$$y_0 + F(B_\epsilon(0_1)) = F(x_0 + B_\epsilon(0_1)) = F(B_\epsilon(x_0)).$$

Because U is open, if ϵ is smaller than some ϵ_0 we have $B_\epsilon(x_0) \subset U$ and

$$y_0 + F(B_\epsilon(0_1)) = F(B_\epsilon(x_0)) \subset F(U), \quad \forall \epsilon < \epsilon_0.$$

Coupling this with (13.20) where we replace c by ϵ we have:

$$B_{\epsilon d}(y_0) = y_0 + B_{\epsilon d}(0_2) \subset y_0 + F(B_\epsilon(0_1)) \subset F(U) = V, \quad \forall \epsilon < \epsilon_0,$$

which shows that y_0 is an interior point of V. □

Theorem 13.4. *Let X_1 and X_2 be two Banach spaces, and let $F : X_1 \mapsto X_2$ be a surjective, bounded linear map. Then F is open.*

Proof. According to Lemma 13.3, it is enough to prove that 0_2 is an interior point of the set $F(B_1(0_1))$, which is the image through F of the unit open ball in X_1.

Let us consider the closed sets $S_n := \overline{F(B_n(0_1))}$, for $n \geq 1$. From Lemma 13.2 we see that $F(\overline{B_n(0_1)}) \subset S_n$ for all n. Clearly:

$$F(X_1) = \bigcup_{n\geq 1} F(\overline{B_n(0_1)}) \subset \bigcup_{n\geq 1} S_n.$$

Because F is surjective, we have $F(X_1) = X_2$. It means that X_2 is a Banach space covered by a countable set of closed sets. Baire's Category Theorem 12.1 implies that there exists some $N \geq 1$ such that S_N has a non-empty interior. Thus we can find some $y_0 \in S_N$ and $\epsilon_0 > 0$ such that $B_{\epsilon_0}(y_0) \subset S_N$. Hence:

$$B_{\epsilon_0}(y_0) = y_0 + B_{\epsilon_0}(0_2) \subset \overline{F(B_N(0_1))}. \tag{13.21}$$

Since F is surjective, there exists $x_0 \in X_1$ such that $y_0 = F(x_0)$. The linearity of F gives:

$$F(B_N(0_1)) = y_0 + F(-x_0 + B_N(0_1)) = y_0 + F(B_N(-x_0)) \subset y_0 + F(B_{N+||x_0||_1}(0_1)),$$

and after taking the closure and coupling it with (13.21) we obtain:

$$y_0 + B_{\epsilon_0}(0_2) \subset y_0 + \overline{F(B_{N+||x_0||_1}(0_1))},$$

which implies:

$$B_{\epsilon_0}(0_2) \subset \overline{F(B_{N+||x_0||_1}(0_1))}. \tag{13.22}$$

Define $r := \epsilon_0(N + ||x_0||_1)^{-1}$. Let $c > 0$ be a fixed constant. Choose any $y \in B_{cr}(0_2)$, i.e. we have $||y||_2 < cr$. Then $(N + ||x_0||_1)c^{-1}y \in B_{\epsilon_0}(0_2)$ and (13.22) implies that there exists a sequence of points $x_n \in B_{N+||x_0||_1}(0_1)$ such that

$$(N + ||x_0||_1)c^{-1}y = \lim_{n\to\infty} F(x_n), \qquad y = \lim_{n\to\infty} F((N + ||x_0||_1)^{-1}cx_n).$$

But $\tilde{x}_n := (N + ||x_0||_1)^{-1}cx_n \in B_c(0_1)$, thus we have that $y = \lim_{n\to\infty} F(\tilde{x}_n) \in \overline{F(B_c(0_1))}$. To conclude, what we proved until now is the existence of a positive number $r > 0$ such that the following inclusion holds true:

$$B_{rc}(0_2) \subset \overline{F(B_c(0_1))}, \quad \forall c > 0. \tag{13.23}$$

We will now show that

$$B_r(0_2) \subset F(B_2(0_1)), \tag{13.24}$$

which according to (13.20) is equivalent with $B_{r/2}(0_2) \subset F(B_1(0_1))$, thus we can take $d = r/2$ in Lemma 13.3 and conclude that F is open. Now let us prove (13.24). Choose any $y \in B_r(0_2)$.

Using (13.23) with $c = 1$ it follows that y is an adherent point of $F(B_1(0_1))$, thus there exists $x_1 \in B_1(0_1)$ such that $||y - F(x_1)||_2 < r/2$.

The vector $y_1 = y - F(x_1) \in B_{r2^{-1}}(0_2)$ is an adherent point of the set $F(B_{2^{-1}}(0_1))$. Using (13.23) with $c = 1/2$ it follows that there exists $x_2 \in B_{2^{-1}}(0_1)$ such that

$$||y_1 - F(x_2)||_2 = ||y - F(x_1 + x_2)||_2 < \frac{r}{2^2}.$$

By induction, for every $n \geq 1$ we can construct the points $x_n \in B_{2^{-n+1}}(0_1)$ such that

$$||y - F(x_1 + x_2 + \cdots + x_n)||_2 < \frac{r}{2^n}, \quad ||x_n||_1 < 2^{-n+1}, \quad \forall n \geq 1.$$

But then $x = \sum_{n\geq 1} x_n$ converges absolutely in X_1 (remember that X_1 is a Banach space), $||x||_1 < 2$ and $y = F(x)$. The proof is over.

□

Corollary 13.5. *Let $F : X_1 \mapsto X_2$ an invertible, bounded linear map between two Banach spaces. Then the inverse map $F^{-1} : X_2 \mapsto X_1$ is also linear and bounded.*

Proof. The linearity of F^{-1} is implied by the equality:

$$F(F^{-1}(\lambda x + \mu y)) = \lambda x + \mu y = \lambda F(F^{-1}(x)) + \mu F(F^{-1}(y)) = F(\lambda F^{-1}(x) + \mu F^{-1}(y))$$

and from the injectivity of F.

Since F is surjective, the open mapping theorem implies the existence of $\delta > 0$ such that $B_\delta(0_2) \subset F(B_1(0_1))$. The invertibility of F implies $F^{-1}(B_\delta(0_2)) \subset B_1(0_1)$. Now for any $y \in X_2$ with $||y||_2 = 1$ we have:

$$2^{-1}\delta F^{-1}(y) = F^{-1}(2^{-1}\delta y) \in B_1(0_1),$$

i.e. $||F^{-1}(y)||_1 \leq 2\delta^{-1}$ if $||y||_2 = 1$. This implies that F^{-1} is bounded. □

14 The Closed Graph Theorem

Let $\mathcal{H}$ be a Hilbert space with its inner product denoted by $\langle x, y \rangle$. We denote by $\mathcal{H}' = \mathcal{H} \oplus \mathcal{H}$ the linear space containing ordered pairs $[x, y]$ with $x, y \in \mathcal{H}$, and

$$\langle [x, y], [u, w] \rangle' := \langle x, u \rangle + \langle y, w \rangle, \quad ||[x, y]||' := \sqrt{||x||^2 + ||y||^2}.$$

It is easy to see that $\mathcal{H}'$ is complete, thus a Hilbert space. Indeed, if $z_n = [x_n, y_n]$ is a Cauchy sequence in $\mathcal{H}'$, then since

$$\max\{||x_p - x_q||, \; ||y_p - y_q||\} \leq ||z_p - z_q||', \quad \forall p, q \geq 1$$

it follows that both x_n and y_n are Cauchy in $\mathcal{H}$ and converge to x and y respectively. Finally we use

$$||z_n - [x, y]||' \leq ||x_n - x|| + ||y_n - y||.$$

Let $A : \mathcal{H} \mapsto \mathcal{H}$ be a linear operator (not necessarily bounded). We define the graph of A to be a linear subspace $G_A \subset \mathcal{H} \oplus \mathcal{H}$ such that

$$G_A := \{[x, Ax] \in \mathcal{H} \oplus \mathcal{H}, \quad \forall x \in \mathcal{H}\}.$$

Theorem 14.1. *The linear map $A : \mathcal{H} \mapsto \mathcal{H}$ is bounded if and only if G_A is closed in $\mathcal{H}'$.*

Proof. 1. Let us first assume that A is bounded. We will prove that every adherent point of G_A belongs to G_A. Indeed, assume that $z_n = [x_n, Ax_n]$ converges to some point $[x, y] \in \mathcal{H}'$. Then x_n must converge to x and Ax_n to y. Since A is bounded (thus continuous), we have that $Ax = y$ and we are done.

2. For the second implication, assume that G_A is closed in $\mathcal{H}'$. We will prove that A is bounded. The linear space $X_1 := (G_A, || \cdot ||')$ is a Banach space because G_A is closed. Denoting by $X_2 := (\mathcal{H}, || \cdot ||)$, define the linear map $F : X_1 \mapsto X_2$ given by $F([x, Ax]) := x$. Clearly, F is invertible with $F^{-1}(y) = [y, Ay]$, and F is bounded:

$$||F([x, Ax])|| = ||x|| \leq ||[x, Ax]||', \quad \forall x \in \mathcal{H}.$$

The Corollary 13.5 implies that F^{-1} is bounded, that is there exists a constant $C > 0$ such that:

$$||Ay|| \leq \sqrt{||y||^2 + ||Ay||^2} = ||F^{-1}(y)||' \leq C||y||, \quad \forall y \in \mathcal{H},$$

i.e. A is bounded. □

The following corrolary is known as the Hellinger-Toeplitz theorem:

Corollary 14.2. *Let $\mathcal{H}$ be a Hilbert space and let $A : \mathcal{H} \mapsto \mathcal{H}$ be any linear and symmetric operator, i.e. it obeys $\langle Ax, y\rangle = \langle x, Ay\rangle$ for every $x, y \in \mathcal{H}$. Then A is bounded.*

Proof. We only need to show that the graph G_A is closed. The fact that G_A contains all its adherent points can be characterized like this: if $\{x_n\}_{n\geq 1} \subset \mathcal{H}$ converges to some $x \in \mathcal{H}$ *and in the same time* $\{Ax_n\}_{n\geq 1}$ converges to some $y \in \mathcal{H}$, then $Ax = y$. Let us prove that this indeed happens when A is also symmetric.

Assume that $x_n \to x$ and $Ax_n \to y$. We need to prove that $Ax = y$. From the equality $\langle Ax_n, z\rangle = \langle x_n, Az\rangle$ which holds for an arbitrary $z \in \mathcal{H}$ and from the continuity of the inner product, we obtain that $\langle y, z\rangle = \langle x, Az\rangle = \langle Ax, z\rangle$. This means that $Ax - y$ belongs to $\mathcal{H}^\perp$, thus $Ax = y$ and we are done. □

15 The spectral theorem for compact and selfadjoint operators

Let $\mathcal{H}$ be a separable Hilbert space, and let $T = T^* \in B(\mathcal{H})$ be a selfadjoint, compact operator. This means that given any bounded sequence $\{x_n\}_{n\geq 1}$, one can always find a convergent subsequence for $\{Tx_n\}_{n\geq 1}$. We assume that the dimension of $\mathcal{H}$ is infinite.

Theorem 15.1. *There exists an orthonormal basis in $\mathcal{H}$ denoted by $\{\psi_j\}_{j\geq 1}$, and a sequence of real numbers $\{\lambda_j\}_{j\geq 1}$ converging to 0 and satisfying $||T|| = |\lambda_1| \geq |\lambda_2| \geq \dots$, such that for every $f \in \mathcal{H}$ we have:*

$$Tf = \sum_{j\geq 1} \lambda_j \psi_j \langle f, \psi_j\rangle. \tag{15.25}$$

15.1 Proof of Theorem 15.1

Lemma 15.2. *Let $z \in \mathbb{C}$. We have $\mathrm{null}(T - z) = \{\mathrm{range}(T - \overline{z})\}^\perp$, and $\mathcal{H} = \mathrm{null}(T - z) \oplus \overline{\{\mathrm{range}(T - \overline{z})\}}$.*

Proof. Let us prove the first equality. We know that T is symmetric, hence $\langle (T - z)f, g\rangle = \langle f, (T - \overline{z})g\rangle$ for all vectors $f, g \in \mathcal{H}$. If $f \in \mathrm{null}(T - z)$, then $0 = \langle f, (T - \overline{z})g\rangle$ for all g, thus $f \in \{\mathrm{range}(T - \overline{z})\}^\perp$. If $f \in \{\mathrm{range}(T - \overline{z})\}^\perp$, then $\langle (T - z)f, g\rangle = 0$ for all $g \in \mathcal{H}$, thus $(T - z)f = 0$ and $f \in \mathrm{null}(T - z)$.

Let us prove the second equality. We know that for any linear subspace M we have $\{M^\perp\}^\perp = \overline{M}$. Thus:

$$\mathcal{H} = \mathrm{null}(T - z) \oplus \{\mathrm{null}(T - z)\}^\perp = \mathrm{null}(T - z) \oplus \overline{\mathrm{range}(T - \overline{z})}. \tag{15.26}$$

□

Lemma 15.3. *Let $z = x + iy$. Then $||(T - z)f|| \geq |y|\ ||f||$ for every $f \in \mathcal{H}$. In particular, if $y \neq 0$, then $\mathrm{null}(T - z) = \{0\}$ and $T - z$ is injective.*

Proof. It is an easy consequence of the fact that $\langle Tf, f\rangle$ is real and:

$$||f||\ ||(T - z)f|| \geq |\langle (T - z)f, f\rangle| = \left|\langle (T - x)f, f\rangle - iy||f||^2\right| \geq |y|\ ||f||^2.$$

□

Lemma 15.4. *Assume that for a given z, there exists $\delta > 0$ such that*

$$||(T - z)f|| \geq \delta||f||, \quad \forall f \in \mathcal{H}. \tag{15.27}$$

Then $T - z$ is both injective and surjective, thus $z \in \rho(T)$.

Proof. Clearly, $T - z$ is injective. Our goal is to prove that $\text{range}(T - z) = \mathcal{H}$.

Let us write $z = x + iy$. If $y \neq 0$, then (15.27) is a consequence of Lemma 15.3. Thus we also have that

$$||(T - \overline{z})f|| \geq \delta||f||, \quad \forall f \in \mathcal{H}. \tag{15.28}$$

When $y = 0$, (15.28) coincides with (15.27).

In both cases, (15.28) implies that $T - \overline{z}$ is injective, thus $\text{null}(T - \overline{z}) = \{0\}$. Using (15.26) with z replaced by $\overline{z}$ we obtain that the range of $T - z$ is dense in $\mathcal{H}$:

$$\overline{\text{range}(T - z)} = \mathcal{H}. \tag{15.29}$$

The only remaining thing in the proof is to show that $\text{range}(T - z)$ is a closed set, which together with (15.29) would show the surjectivity of $T - z$.

Let us do that. Assume that $\{y_n\}_{n\geq 1} \subset \text{range}(T - z)$ converges to $y_\infty \in \mathcal{H}$. We have to show that $y_\infty \in \text{range}(T - z)$. There exists $\{x_n\}_{n\geq 1} \subset H$ such that $y_n = (T - z)x_n$. Using (15.27) we can write:

$$||x_{n+k} - x_n|| \leq \frac{1}{\delta}||(T - z)(x_{n+k} - x_n)|| = \frac{1}{\delta}||y_{n+k} - y_n||, \quad \forall n, k \geq 1. \tag{15.30}$$

Since $\{y_n\}_{n\geq 1}$ is Cauchy, (15.30) implies the same thing for $\{x_n\}_{n\geq 1}$. Thus there exists $x_\infty \in \mathcal{H}$ such that $\lim_{n\to\infty} x_n = x_\infty$. Using this in the equality $Tx_n = zx_n + y_n$ together with the continuity of T, we obtain $Tx_\infty = zx_\infty + y_\infty$ and:

$$y_\infty = (T - z)x_\infty \in \text{range}(T - z).$$

□

Remark 1. The previous lemma shows that if T is a selfadjoint operator and if

$$||(T - z)x|| \geq \delta > 0, \quad \forall ||x|| = 1, \tag{15.31}$$

then $z \in \rho(T)$. Thus if $\lambda \in \sigma(T)$ we must have

$$\inf_{||x||=1} ||(T - \lambda)x|| = 0,$$

or more precisely, there exists a sequence $\{x_n\}_{n\geq 1}$ with $||x_n|| = 1$ such that

$$\lim_{n\to\infty} (T - z)x_n = 0. \tag{15.32}$$

Remark 2. Lemma 15.3 and Lemma 15.4 prove that $\sigma(T) \subset \mathbb{R}$. Moreover, if $|z| > ||T||$ we can write

$$(T - z)^{-1} = -\sum_{n\geq 0} \frac{1}{z^{n+1}} T^n, \tag{15.33}$$

thus $\sigma(T) \subset [-||T||, ||T||]$.

Let us now characterize the structure and nature of the spectrum of T.

Lemma 15.5. *If $\lambda \in \sigma(T)$ and $\lambda \neq 0$, then there exists at least one eigenvector $f \neq 0$ such that $Tf = \lambda f$.*

Proof. Because $\lambda \in \sigma(T)$, we have the bounded sequence $\{x_n\}_{n\geq 1}$ from (15.32). Since T is compact, we can find a subsequence $\{x_{n_k}\}_{k\geq 1}$ such that $\{Tx_{n_k}\}_{k\geq 1}$ is convergent to some y_∞. We can write:

$$x_{n_k} = \frac{1}{\lambda} T x_{n_k} - \frac{1}{\lambda}(T - \lambda)x_{n_k},$$

and since the r.h.s. converges to $\frac{1}{\lambda}y_\infty$ we conclude that $\lim_{k\to\infty} x_{n_k} = \frac{1}{\lambda}y_\infty$. The continuity of T implies that $\lim_{k\to\infty} Tx_{n_k} = \frac{1}{\lambda}Ty_\infty$. Hence:

$$0 = \lim_{k\to\infty}(Tx_{n_k} - \lambda x_{n_k}) = \frac{1}{\lambda}Ty_\infty - y_\infty.$$

Moreover, $||x_{n_k}|| = 1$ implies that $||y_\infty|| = 1$, thus we can choose our eigenvector to be $f = y_\infty$. □

Lemma 15.6. *If $\lambda_1 \neq \lambda_2$ belong to the spectrum, and if f_1 and f_2 are two corresponding eigenvectors, then $\langle f_1, f_2\rangle = 0$.*

Proof. Use the symmetry of T and write $0 = \langle Tf_1, f_2\rangle - \langle f_1, Tf_2\rangle = (\lambda_1 - \lambda_2)\langle f_1, f_2\rangle$. □

Lemma 15.7. *The spectrum of T cannot have other accumulation points outside* 0. *In other words, $\sigma(T) \setminus \{0\}$ is a discrete set consisting from isolated points.*

Proof. Assume that $\lambda \neq 0$ is an accumulation point of $\sigma(T)$. It means that we can find a sequence of points $\{\lambda_n\}_{n\geq 1} \subset \sigma(T)$, all distinct and not zero, such that

$$\lim_{n\to\infty} \lambda_n = \lambda.$$

From Lemma 15.5 we obtain at least an eigenvector x_n, $||x_n|| = 1$, such that $Tx_n = \lambda_n x_n$, or $x_n = \frac{1}{\lambda_n}Tx_n$. Since T is compact, there exists a subsequence x_{n_k} such that Tx_{n_k} converges to some y. Thus

$$\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty} \frac{1}{\lambda_{n_k}}Tx_{n_k} = \frac{1}{\lambda}y.$$

Thus we have just constructed a convergent subsequence of $\{x_n\}_{n\geq 1}$. But since each x_n corresponds to a different λ_n, Lemma 15.6 tells us that $||x_j - x_k|| = \sqrt{2}$ if $j \neq k$, therefore this sequence cannot have Cauchy subsequences. We arrived to a contradiction.

□

Lemma 15.8. *Assume that $\lambda \in \sigma(T) \setminus \{0\}$. Then the dimension of* null$(T - \lambda)$ *is finite.*

Proof. Assume the contrary, i.e. the existence of infinitely many linearly independent vectors in null$(T - \lambda)$. Up to a Gramm-Schmidt procedure, we can consider them to be orthogonal and normalized to one. If $\{x_n\}_{n\geq 1}$ is such a list, then again $||x_j - x_k|| = \sqrt{2}$, thus it cannot have any convergent subsequences. But since $x_n = \frac{1}{\lambda}Tx_n$, the compactness of T would generate a convergent subsequence for $\{x_n\}_{n\geq 1}$, and we arrive to a contradiction.

□

Until now we know that the spectrum of T is contained in the interval $[-||T||, ||T||]$, it consists from isolated points outside 0, and the nullspace associated to each of its nonzero points is finite dimensional. Thus the nonzero spectrum is only composed from eigenvalues with finite geometric multiplicity, and they can only accumulate at 0.

Lemma 15.9. *At least one of the numbers $\pm||T||$ is an eigenvalue for T.*

Proof. Without loss of generality, we may assume that $||T|| > 0$. From the definition of the norm, we have $||T|| = \sup_{||x||=1} ||Tx||$. Thus there exists a sequence $\{x_n\}_{n\geq 1}$, $||x_n|| = 1$ such that $\lim_{n\to\infty} ||Tx_n|| = ||T||$. Since T is compact, we can find a subsequence x_{n_k} such that

$\lim_{k\to\infty} Tx_{n_k} = y$, thus $||y|| = ||T||$. In order to simplify notation, denote $||T||$ by λ. Then we have:

$$\begin{aligned}
&\lim_{k\to\infty} ||(T^2-\lambda^2)x_{n_k}||^2 = \lim_{k\to\infty} \langle (T^2-\lambda^2)x_{n_k}, (T^2-\lambda^2)x_{n_k}\rangle \\
&= \lim_{k\to\infty} \{\langle T^2x_{n_k}, T^2x_{n_k}\rangle - 2\lambda^2\langle T^2x_{n_k}, x_{n_k}\rangle + \lambda^4||x_{n_k}||^2\} = \langle Ty, Ty\rangle - 2\lambda^2\langle y, y\rangle + \lambda^4 \\
&= \langle Ty, Ty\rangle - \lambda^4 \geq 0. \qquad (15.34)
\end{aligned}$$

Thus we get $||Ty|| \geq \lambda^2$. Moreover:

$$\begin{aligned}
0 &\leq ||(T-\lambda)(T+\lambda)y||^2 = ||(T^2-\lambda^2)y||^2 = \langle (T^2-\lambda^2)y, (T^2-\lambda^2)y\rangle \\
&= \langle T^2y, T^2y\rangle - 2\lambda^2\langle T^2y, y\rangle + \lambda^4||y||^2 = \langle T^2y, T^2y\rangle - 2\lambda^2\langle Ty, Ty\rangle + \lambda^4||y||^2 \\
&\leq ||T^2y||^2 - \lambda^6 \leq 0. \qquad (15.35)
\end{aligned}$$

In the last line above we used (15.34). Thus (15.35) implies $(T-\lambda)(T+\lambda)y = 0$. Now if $(T+\lambda)y = 0$, it means that $-\lambda$ is an eigenvalue. If $f = (T+\lambda)y \neq 0$, then necessarily $(T-\lambda)f = 0$ which means that λ is an eigenvalue. □

The previous result together with Lemma 15.8 imply the existence of a finite number of eigenvectors of T which span the subspace $M_\lambda := \text{null}(T-\lambda)$ where λ is one of the values $||T||$ or $-||T||$. Denote by $\{\psi_j(\lambda)\}_{j=1}^{\dim(M_\lambda)}$ an orthonormal basis of M_λ, consisting of eigenvectors of T. Denote by P_λ the orthogonal projection associated to $\text{null}(T-\lambda)$:

$$P_\lambda f := \sum_{j=1}^{\dim(M_\lambda)} \langle f, \psi_j(\lambda)\rangle \psi_j(\lambda). \qquad (15.36)$$

By direct computation, one can show that $P_\lambda^* = P_\lambda = P_\lambda^2$.

By convention, if λ is not in the spectrum of T, then $M_\lambda = \{0\}$ and $P_\lambda = 0$. Denote by

$$M_1 := M_{+||T||} \oplus M_{-||T||}. \qquad (15.37)$$

Lemma 15.10. *The subspace M_1 is a finite dimensional, closed linear subspace, which is left invariant by T (that is $TM_1 \subset M_1$). The same is true for $M_1^\perp$.*

Proof. Every $f \in M_1$ can be written as a finite linear combination of the type $f = \sum_j \langle f, \psi_j\rangle\psi_j$. Since all ψ_j's are eigenvectors of T, then $Tf \in M_1$.

Now let us prove that $M_1^\perp$ is invariant under T. Let $g \in M_1^\perp$. Then for every $f \in M_1$ we have:

$$\langle Tg, f\rangle = \langle g, Tf\rangle = 0,$$

since $Tf \in M_1$. Hence $Tg \in M_1^\perp$. □

Now consider the decomposition $\mathcal{H} = M_1 \oplus M_1^\perp$. The previous invariance result allows us to write our operator T as a direct sum $T = (||T||P_{+||T||} - ||T||P_{-||T||}) \oplus T_1$, where T_1 is simply the restriction of T to $M_1^\perp$. The next technical result is the following:

Lemma 15.11. *The restriction T_1 is also compact and selfadjoint. Moreover, $||T_1|| < ||T||$.*

Proof. The fact that T is compact and selfadjoint follows from

$$T_1 = T(1 - P_{+||T||} - P_{-||T||}) = (1 - P_{+||T||} - P_{-||T||})T.$$

Now let us prove that $||T_1|| < ||T||$. Clearly, $||T_1|| \leq ||T||$, so we only need to prove that the two norms cannot be equal. Assume that they are equal. Then applying Lemma 15.9 to T_1, it would provide an eigenvector $\phi \in M_1^\perp$, $||\phi|| = 1$, for T_1. But ϕ would also be an eigenvector for T corresponding to $||T||$ or $-||T||$, thus $\phi \in M_1$, contradicting $\phi \neq 0$. □

Remark 3. We have the inclusion $\text{null}(T) \subset M_1^{\perp}$; indeed, let $f \in \text{null}(T)$ and let ψ_j one eigenvector of T from M_1 corresponding to the eigenvalue $\lambda \neq 0$. Then

$$0 = \frac{1}{\lambda}\langle Tf, \psi_j\rangle = \frac{1}{\lambda}\langle f, T\psi_j\rangle = \langle f, \psi_j\rangle.$$

Thus f is orthogonal to any linear combination of ψ_j's, thus $f \in M_1^{\perp}$.

Now the proof of Theorem 15.1 is almost over. If $M_1^{\perp} = \text{null}(T)$, then we have $H = M_1 \oplus \text{null}(T)$ and $T = (||T||P_{+||T||} - ||T||P_{-||T||}) \oplus 0$.

Otherwise, define M_2 as the subspace of $M_1^{\perp}$ corresponding to $\text{null}(T_1 + ||T_1||) \oplus \text{null}(T_1 - ||T_1||)$ and decompose $\mathcal{H} = M_1 \oplus (M_2 \oplus M_2^{\perp})$. Here T_1 decomposes as

$$T_1 = (||T_1||P_{+||T_1||} - ||T_1||P_{-||T_1||}) \oplus T_2.$$

By induction, we obtain the decomposition

$$\mathcal{H} = M_1 \oplus M_2 \cdots \oplus (M_n \oplus M_n^{\perp})$$

and

$$T = \oplus_{j=0}^{n-1}(||T_j||P_{+||T_j||} - ||T_j||P_{-||T_j||}) \oplus T_n,$$

where T_n is the restriction of T_{n-1} to $M_n^{\perp}$. By convention, $T_0 = T$. Reasoning as in the proof of Remark 3, we get that $\text{null}(T) \subseteq M_n^{\perp}$. If they are equal, then we stop. Otherwise, we continue the reduction procedure.

Now assume that we never get $\text{null}(T) = M_n^{\perp}$. It follows that $T_n \neq 0$, and also $\lim_{n\to\infty} ||T_n|| = 0$ because Lemma 15.7 forbids the accumulation of eigenvalues outside 0.

Lemma 15.12. *We have* $\oplus_{j\geq 0} M_j = \overline{\text{range}(T)}$.

Proof. Fix $f \in \mathcal{H}$. The vector $\sum_{j=0}^{n-1}(||T_j||P_{+||T_j||}f - ||T_j||P_{-||T_j||}f)$ can be seen as an element of $\oplus_{j\geq 0} M_j$, where all components with an index larger than $n+1$ are zero. We know that $T_n f = Tf - \sum_{j=0}^{n-1}(||T_j||P_{+||T_j||}f - ||T_j||P_{-||T_j||}f)$, and $||T_n f|| \to 0$ when n grows. Thus we can approximate Tf arbitrarily well with elements of $\oplus_{j\geq 0} M_j$. □

Corollary 15.13. *We have the decomposition* $\mathcal{H} = \{\oplus_{j\geq 0} M_j\} \oplus \text{null}(T)$.

Proof. Put $z = 0$ in (15.26) and use Lemma 15.12.

□

We can now conclude the proof of Theorem 15.1. The orthonormal basis consists from the eigenvectors of T corresponding to non-zero eigenvalues, put together with an arbitrary basis in $\text{null}(T)$. The numbers λ_j's are either the nonzero eigenvalues of T or zero. The operator T has a finite rank if a finite number of λ_j's are nonzero.

16 The singular value decomposition of a compact operator

Theorem 16.1. *Let $\mathcal{H}$ be a separable Hilbert space, and let A be a compact operator. Then there exist two orthonormal basis of $\mathcal{H}$, $\{e_j\}_{j\geq 1}$ and $\{f_j\}_{j\geq 1}$, and a nonincreasing sequence of non-negative numbers $s_j \geq 0$ accumulating at zero such that for every $f \in \mathcal{H}$ we have:*

$$Af = \sum_{j\geq 1} s_j \langle f, e_j\rangle f_j.$$

Proof. Let $T := A^*A$. We see that T is compact, selfadjoint and non-negative. Moreover, $\text{null}(A) = \text{null}(T)$; indeed, if $x \in \text{null}(A)$ then $Tx = A^*(Ax) = 0$, thus $x \in \text{null}(T)$. If $x \in \text{null}(T)$, then $0 = \langle x, Tx\rangle = ||Ax||^2$ thus $Ax = 0$ and $x \in \text{null}(A)$.

According to Theorem 15.1, there exists an orthonormal basis $\{e_j\}_{j\geq 1}$ consisting of eigenvectors of T, and let λ_j be their corresponding (non-zero) eigenvalues. We have

$$Af = \sum_{j\geq 1} \langle f, e_j\rangle Ae_j \tag{16.38}$$

In the above sum, only those e_j's appear which are not spanning the null space of T. Denote by $f_j := \frac{1}{||Ae_j||} Ae_j$, if $Ae_j \neq 0$. Clearly, from (16.38) it follows that the f_j's span the closure of the range of A. Now let us prove that the f_j's are orthogonal on each other. If $j \neq k$ we have

$$\langle f_j, f_k\rangle = \frac{1}{||Ae_j||\,||Ae_k||}\langle e_j, Te_k\rangle = 0.$$

We can extend the f_j basis in an arbitrary way to $\text{range}(A)^\perp$. Finally, let us denote by $s_j := ||Ae_j|| = \sqrt{\langle e_j, A^*Ae_j\rangle} = \sqrt{\lambda_j}$. From (16.38) and the definition of f_j's and s_j's, the theorem is proved.

□