Topics in Analysis
Part III, Autumn 2010

T. W. Körner

April 23, 2016

Small print This is just a first draft for the course. The content of the course will be what I say, not what these notes say. Experience shows that skeleton notes (at least when I write them) are very error prone so use these notes with care. I should **very much** appreciate being told of any corrections or possible improvements and might even part with a small reward to the first finder of particular errors.

Contents

1 Introduction

This course splits into two parts. The first part takes a look at the Baire category theorem, Tychonov's theorem the Hahn Banach theorem together with some of their consequences. There will be two or three lectures of fairly abstract set theory but the the rest of the course is pretty concrete. The second half of the course will look at the theory of distributions. (The general approach is that of [3] but the main application is different.)

I shall therefore assume that you know what is a normed space, and what is a a linear map and that you can do the following exercise.

Exercise 1. *Let $(X, \| \|_X)$ and $(Y, \| \|_Y)$ be normed spaces.*

(i) If $T : X \to Y$ is linear, then T is continuous if and only if there exists a constant K such that

$$\|Tx\|_Y \leq K\|x\|_X$$

for all $x \in X$.

(ii) If $T : X \to Y$ is linear and $x_0 \in X$, then T is continuous at x_0 if and only if there exists a constant K such that

$$\|Tx\|_Y \leq K\|x\|_X$$

for all $x \in X$.

(iii) If we write $\mathcal{L}(X, Y)$ for the space of continuous linear maps from X to Y and write

$$\|T\| = \sup\{\|Tx\|_Y : \|x\|_X = 1, \ x \in X\}$$

then $(\mathcal{L}(X, Y), \| \|)$ is a normed space.

I also assume familiarity with the concept of a metric space and a complete metric space. You should be able to do at least parts (i) and (ii) of the following exercise (part (iii) is a little harder).

Exercise 2. *Let $(X, \| \|_X)$ and $(Y, \| \|_Y)$ be normed spaces.*

(i) If $(Y, \| \|_Y)$ is complete then $(\mathcal{L}(X, Y), \| \|)$ is.

(ii) Consider the set s of sequences $x = (x_1, x_2, \dots)$ in which only finitely many of the x_j are non-zero. Explain briefly how s may be considered as a vector space.

(iii) If $(X, \| \|_X)$ is complete does it follow that $(\mathcal{L}(X, Y), \| \|)$ is? Give a proof or a counter-example.

The reader will notice that I have not distinguished between vector spaces over $\mathbb{R}$ and those over $\mathbb{C}$. I shall try to make the distinction when it matters but, if the two cases are treated in the same way, I shall often proceed as above.

Although I shall stick with metric spaces as much as possible, there will be points where we shall need the notions of a topological space, a compact topological space and a Hausdorff topological space. I would be happy, if requested, to give a supplementary lecture introducing these notions. (Even where I use them, no great depth of understanding is required.)

I shall also use, without proof, the famous Stone-Weierstrass theorem.

Theorem 3. *(A) Let X be a compact space and $C(X)$ the space of real valued continuous functions on X. Suppose A is a subalgebra of $C(X)$ (that is a subspace which is algebraically closed under multiplication) and*

(i) $1 \in A$,

(ii) Given any two distinct points x and y in X there is an $f \in A$ with $f(x) \neq f(y)$.

Then A is uniformly dense in $C(X)$.

(B) Let X be a compact space and $C(X)$ the space of complex valued continuous functions on X. Suppose A is a subalgebra of $C(X)$ and

(i) $1 \in A$,

(ii) Given any two distinct points x and y in X there is an $f \in A$ with $f(x) \neq f(y)$. Then A is uniformly dense in $C(X)$.

The proof will not be examinable, but if you have not met it, you may wish to request a supplementary lecture on the topic. I may mention some measure theory but this is for interest only and will not be examinable[1]. I intend the course to be fully accessible without measure theory.

2 Baire category

If (X, d) is a metric space we say that a set E in X has dense complement[2] if, given $x \in E$ and $\delta > 0$, we can find a $y \notin E$ such that $d(x, y) < \delta$.

Exercise 4. *Consider the space M_n of $n \times n$ complex matrices with an appropriate norm. Show that the set of matrices which do not have n distinct eigenvalues is a closed set with dense complement.*

[1] In this course, as in other Part III courses you should assume that everything in the lectures and nothing outside them is examinable unless you are explicitly told to the contrary. If you are in any doubt, ask the lecturer.

[2] If the lecturer uses the words 'nowhere dense' correct him for using an old fashioned and confusing terminology

Theorem 5 (Baire's theorem)**.** *If (X,d) is a complete metric space and E_1, E_2, ... are closed sets with dense complement then $X \neq \bigcup_{j=1}^{\infty} E_j$.*

Exercise 6. *(If you are happy with general topology.) Show that a result along the same lines holds true for compact Hausdorff spaces.*

We call the countable union of closed sets with dense complement a set of first category. The following observations are trivial but useful.

Lemma 7. *(i) The countable union of first category sets is itself of first category.*

(ii) If (X,d) is a complete metric space, then Baire's theorem asserts that X is not of first category.

Exercise 8. *If (X,d) is a complete metric space and X is countable show that there is an $x \in X$ and a $\delta > 0$ such that the ball $B(x,\delta)$ with centre x and radius δ consists of one point.*

The following exercise is a standard application of Baire's theorem.

Exercise 9. *Consider the space $C([0,1])$ of continuous functions under the uniform norm $\| \ \|$. Let*

$$E_m = \{f \in C([0,1]) : \text{there exists an } x \in [0,1] \text{ with } |f(x+h) - f(x)| \leq m|h| \text{ for all } x+h \in [0,1]\}.$$

(i) Show that E_m is closed in $(C([0,1], \| \ \|_\infty)$.

(ii) If $f \in C([0,1])$ and $\epsilon > 0$ explain why we can find an infinitely differentiable function g such that $\|f-g\|_\infty < \epsilon/2$. By considering the function h given by

$$h(x) = g(x) + \tfrac{\epsilon}{2}\sin Nx$$

with N large show that E_m has dense complement.

(iii) Using Baire's theorem show that there exist continuous nowhere differentiable functions.

Exercise 10. *(This is quite long and not very central.)*

(i) Consider the space $\mathcal{F}$ of non-empty closed sets in $[0,1]$. Show that if we write

$$d_0(x,E) = \inf_{e \in E} |x-e|$$

when $x \in [0,1]$ and $E \in \mathcal{F}$ and write

$$d(E,F) = \sup_{f \in F} d_0(f,E) + \sup_{e \in E} d_0(e,F)$$

then d is a metric on $\mathcal{F}$.

(ii) Suppose E_n is a Cauchy sequence in $(\mathcal{F}, d)$. By considering

$$E = \{x \,:\, \text{there exist } e_n \in E_n \text{ such that } e_n \to x\},$$

or otherwise, show that E_n converges. Thus $(\mathcal{F}, d)$ is complete.

(iii) Show that the set

$$\mathcal{A}_n = \{E \in \mathcal{F} \,:\, \text{there exists an } x \in E \text{ with } (x - 1/n, x + 1/n) \cap E = \{x\}\}$$

is closed with dense complement in $(\mathcal{F}, d)$. Deduce that the set of elements of $\mathcal{F}$ with isolated points is of first category. (A set E has an isolated point e if we can find a $\delta > 0$ such that $(e - \delta, e + \delta) \cap E = \{e\}$.)

(iv) Let $I = [r/n, (r+1)/n]$ with $0 \leq r \leq n - 1$ and r and n integers. Show that the set

$$\mathcal{B}_{r,n} = \{E \in \mathcal{F} \,:\, E \supseteq I\}$$

is closed with dense complement in $(\mathcal{F}, d)$. Deduce that the set of elements of $\mathcal{F}$ containing an open interval is of first category.

(v) Deduce the existence of non-empty closed sets which have no isolated points and contain no intervals.

3 Non-existence of functions of several variables

This course is very much a penny plain rather than tuppence coloured[3]. One exception is the theorem proved in this section.

Theorem 11. *Let λ be irrational We can find increasing continuous functions $\phi_j : [0, 1] \to \mathbb{R}$ $[1 \leq j \leq 5]$ with the following property. Given any continuous function $f : [0, 1]^2 \to \mathbb{R}$ we can find a function $g : \mathbb{R} \to \mathbb{R}$ such that*

$$f(x, y) = \sum_{j=1}^{5} g(\phi_j(x) + \lambda\phi_j(y)).$$

The main point of Theorem 11 may be expressed as follows.

Theorem 12. *Any continuous function of two variables can be written in terms of continuous functions of one variable and addition.*

[3]And thus suitable for those 'who want from books plain cooking made still plainer by plain cooks'.

That is, there are no true functions of two variables! (We shall explain why this statement is slightly less shocking than it seems at the end of this section.)

For the moment we merely observe that the result is due in successively more exact forms to Kolmogorov, Arnol'd and a succession of mathematicians ending with Kahane whose proof we use here. It is, of course, much easier to prove a specific result like Theorem 11 than one like Theorem 12.

Our first step is to observe that Theorem 11 follows from the apparently simpler result that follows.

Lemma 13. *Let λ be irrational. We can find increasing continuous functions $\phi_j : [0,1] \to \mathbb{R}$ $[1 \leq j \leq 5]$ with the following property. Given any continuous function $F : [0,1]^2 \to \mathbb{R}$ we can find a continuous function $G : \mathbb{R} \to \mathbb{R}$ such that $\|G\|_\infty \leq \|F\|_\infty$ and*

$$\sup_{(x,y)\in[0,1]^2} \left| F(x,y) - \sum_{j=1}^{5} G(\phi_j(x) + \lambda\phi_j(y)) \right| \leq \frac{999}{1000}\|F\|_\infty.$$

Next we make the following observation.

Lemma 14. *We can find a sequence of functions $f_n : [0,1]^2 \to \mathbb{R}$ which are uniformly dense in $C([0,1])^2$.*

This enables us to obtain Lemma 13 from a much more specific result.

Lemma 15. *Let λ be irrational and let the f_n be as in Lemma 14. We can find increasing continuous functions $\phi_j : [0,1] \to \mathbb{R}$ $[1 \leq j \leq 5]$ with the following property. We can find continuous functions $g_n : \mathbb{R} \to \mathbb{R}$ such that $\|g_n\|_\infty \leq \|f_n\|_\infty$ and*

$$\sup_{(x,y)\in[0,1]^2} \left| f_n(x,y) - \sum_{j=1}^{5} g_n(\phi_j(x) + \lambda\phi_j(y)) \right| \leq \frac{998}{1000}\|f_n\|_\infty.$$

Now that we have reduced the matter to satisfying a countable set of conditions, we can use a Baire category argument. We need to use the correct metric space.

Lemma 16. *The space Y of continuous functions $\boldsymbol{\phi} : [0,1] \to \mathbb{R}^5$ with norm*

$$\|\boldsymbol{\phi}\|_\infty = \sup_{t\in[0,1]} \|\boldsymbol{\phi}(t)\|$$

is complete. The subset X of Y consisting of those $\boldsymbol{\phi}$ such that each ϕ_j is increasing is a closed subset of Y. Thus if d is the metric on X obtained by restricting the metric on Y derived from $\| \ \|_\infty$ we have (X,d) complete.

Exercise 17. *Prove Lemma 16*

Lemma 18. *Let $f : [0,1]^2 \to \mathbb{R}$ be continuous and let λ be irrational. Consider the set E of $\boldsymbol{\phi} \in X$ such that there exists a continuous $g : \mathbb{R} \to \mathbb{R}$ such that $\|g\|_\infty \leq \|f\|_\infty$*

$$\sup_{(x,y)\in[0,1]^2} \left| f(x,y) - \sum_{j=1}^{5} g(\phi_j(x) + \lambda\phi_j(y)) \right| < \frac{998}{1000}\|f\|_\infty.$$

Then $X \setminus E$ is a closed set with dense complement in (X, d).

(Notice that it is important to take '$<$' rather than '$\leq$' in the displayed formula of Lemma 18.) Lemma 18 is the heart of the proof and once it is proved we can easily retrace our steps and obtain Theorem 11.

By using appropriate notions of information Vitushkin was able to show that we can not replace continuous by continuously differentiable in Theorem 12. Thus Theorem 11 is an 'exotic' rather than a 'central' result.

4 The principle of uniform boundedness

We start with a result which is sometimes useful by itself but which, for us, is merely a stepping stone to Theorem 22.

Lemma 19 (Principle of uniform boundedness)**.** *Suppose that (X, d) is a complete metric space and we have a collection $\mathcal{F}$ of continuous functions $f : X \to \mathbb{R}$ which are pointwise bounded, that is, given any $x \in X$ we can find a $K(x) > 0$ such that*

$$|f(x)| \leq K(x) \text{ for all } f \in \mathcal{F}.$$

Then we can find a ball $B(x_0, \delta)$ and a K such that

$$|f(x)| \leq K \text{ for all } f \in \mathcal{F} \text{ and all } x \in B(x_0, \delta)\ .$$

Exercise 20. *(i) Suppose that (X, d) is a complete metric space and we have a sequence of continuous functions $f_n : X \to \mathbb{R}$ and a function $f : X \to \mathbb{R}$ such that f_n converges pointwise that is*

$$f_n(x) \to f(x) \text{ for all } x \in X.$$

Then we can find a ball $B(x_0, \delta)$ and a K such that

$$|f_n(x)| \leq K \text{ for all } n \text{ and all } x \in B(x_0, \delta)\ .$$

(ii) (This is elementary but acts as a hint for (iii).) Suppose $y \in [0,1]$. Show that we can find a sequence of continuous functions $f_n : [0,1] \to \mathbb{R}$ such that $1 \geq f_n(x) \geq 0$ for all x and n, f_n converges pointwise to 0 everywhere, f_n converges uniformly on $[0,1] \setminus (y-\delta, y+\delta)$ and fails to converge uniformly on $[0,1] \cap (y-\delta, y+\delta)$ for all $\delta > 0$.

(iii) State with reasons whether the following statement is true or false. Under the conditions of (i) we can obtain the stronger conclusion that we can find a ball $B(x_0, \delta)$ such that

$$f_n(x) \to f(x) \text{ uniformly on } B(x_0, \delta).$$

Exercise 21. *Suppose that (X, d) is a complete metric space and Y is a subset of X which is of first category in X. Suppose further that we have a collection $\mathcal{F}$ of continuous functions $f : X \to \mathbb{R}$ which are pointwise bounded on $X \setminus Y$, that is, given any $x \notin Y$, we can find a $K(x) > 0$ such that*

$$|f(x)| \leq K(x) \text{ for all } f \in \mathcal{F}.$$

Show that we can find a ball $B(x_0, \delta)$ and a K such that

$$|f(x)| \leq K \text{ for all } f \in \mathcal{F} \text{ and all } x \in B(x_0, \delta)\ .$$

We now use the principle of uniform boundedness to prove the Banach-Steinhaus theorem[4].

Theorem 22. (Banach-Steinhaus theorem) *Let $(U, \| \ \|_U)$ and $(V, \| \ \|_V)$ be normed spaces and suppose $\| \ \|_U$ is complete. If we have a collection $\mathcal{F}$ of continuous linear maps from U to V which are pointwise bounded then we can find a K such that $\|T\| \leq K$ for all $T \in \mathcal{F}$.*

Here is a typical use of the Banach-Steinhaus theorem.

Theorem 23. *There exists a continuous 2π periodic function $f : \mathbb{R} \to \mathbb{R}$ whose Fourier series fails to converge at a given point.*

The next exercise contains results that most of you will have already met.

Exercise 24. *(i) Show that the set l^∞ of bounded sequences over $\mathbb{F}$ (with $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$)*

$$\mathbf{a} = (a_1,\ a_2,\ \ldots)$$

[4] You should be warned that a lot of people, including the present writer, tend to confuse the names of these two theorems. My research supervisor took the simpler course of referring to all the theorems of functional analysis as 'Banach's theorem'.

can be made into a vector space in a natural manner. Show that $\|\mathbf{a}\|_\infty = \sup_{j\geq 1}|a_j|$ *defines a complete norm on* l^∞*.*

(ii) Show that s*, the set of convergent sequences and* s_0 *the set of sequences convergent to* 0 *are both closed subspaces of* $(l^\infty, \| \ \|_\infty)$*.*

(iii) Show that the set l^1 *of sequences*

$$\mathbf{a} = (a_1,\ a_2,\ \dots) \text{ such that } \sum_{j=1}^{\infty} |a_j| \text{ converges}$$

can be made into vector space in a natural manner. Show that $\|\mathbf{a}\|_1 = \sum_{j=1}^{\infty}|a_j|$ *defines a complete norm on* l^1*.*

(iv) Show that, if $\mathbf{a} \in l^1$*, then*

$$T_{\mathbf{a}}(\mathbf{b}) = \sum_{j=1}^{\infty} a_j b_j$$

defines a continuous linear map from l^∞ *to* $\mathbb{F}$ *and that* $\|T_{\mathbf{a}}\| = \|\mathbf{a}\|_1$*.*

Here is another use of the Banach-Steinhaus theorem.

Lemma 25. *Let* $a_{ij} \in \mathbb{R}$ $[i, j \geq 1]$*. We say that the* a_{ij} *constitute a summation method if whenever* $c_j \to c$ *we have* $\sum_{j=1}^{\infty} a_{ij}c_j$ *convergent for each* i *and*

$$\sum_{j=1}^{\infty} a_{ij}c_j \to c$$

as $i \to \infty$*.*

The following conditions are necessary and sufficient for the a_{ij} *to constitute a summation method:-*

(i) There exists a K *such that*

$$\sum_{j=1}^{\infty} |a_{ij}| \leq K \text{ for all } i.$$

(ii) $\displaystyle\sum_{j=1}^{\infty} a_{ij} \to 1$ *as* $i \to \infty$*.*

(iii) $a_{ij} \to 0$ *as* $i \to \infty$ *for each* j*.*

Exercise 26. *Cesàro's summation method takes a sequence* $c_0,\ c_1,\ c_2, \dots$ *and replaces it with a new sequence whose* n*th term*

$$b_n = \frac{c_1 + c_2 + \cdots + c_n}{n}$$

is the average of the first n terms of the old sequence.

(i) By rewriting the statement above along the lines of Lemma 25 show that if the old sequence converges to c so does the new one.

(ii) Examine what happens when $c_j = (-1)^j$. Examine what happens if $c_j = (-1)^k$ when $2^k \leq j < 2^{k+1}$.

(iii) Show that, in the notation of Lemma 25, taking $a_{n,2n} = 1$, $a_{n,m} = 0$, otherwise, gives a summation method. Show that taking $a_{n,2n+1} = 1$, $a_{n,m} = 0$, otherwise, also gives a summation method. Show that the two methods disagree when presented with the sequence $1, -1, 1, -1, \dots$.

Another important consequence of the Baire category theorem is the open mapping theorem. (Recall that a complete normed space is called a Banach space.)

Theorem 27 (Open mapping theorem)**.** *Let E and F be Banach spaces and $T : E \to F$ be a continuous linear surjection. Then T is an open map (that is to say, if U is open in E we have TU open in F.)*

This has an immediate corollary.

Theorem 28 (Inverse mapping theorem)**.** *Let E and F be Banach spaces and let $T : E \to F$ be a continuous linear bijection. Then T^{-1} is continuous.*

The next exercise is simple, and if you can not do it this reveals a gap in your knowledge (which can be remedied by asking the lecturer) rather than in intelligence.

Exercise 29. *Let (X, d) and (Y, ρ) be metric spaces with associated topologies τ and σ. Then the product topology induced on $X \times Y$ by τ and σ is the same as the topology given by the metric*

$$\triangle((x_1, y_1), (x_2, y_2)) = d(x_1, x_2) + \rho(y_1, y_2).$$

The inverse mapping theorem has the following useful consequence.

Theorem 30 (Closed graph theorem)**.** *Let E and F be Banach spaces and let $T : E \to F$ be linear. Then T is continuous if and only the graph*

$$\{(x, Tx) \,:\, x \in E\}$$

is closed in $E \times F$ with the product topology.

5 Countable choice and Baire's theorem

Most of this course is 'practical' but the next couple of sections deal with 'foundational matters'.

We discuss **The axiom of choice** *Let $\mathcal{A}$ be a non-empty collection of non-empty sets. Then there exists a function*

$$f : \mathcal{A} \to \bigcup_{A \in \mathcal{A}} A$$

such that $f(A) \in A$ for all $A \in \mathcal{A}$.

Note that we do not require a specific axiom if $\mathcal{A}$ only contains one set, or, indeed, a finite set of sets.

However we do need an axiom to make a countable set of choices.

The axiom of countable choice *Let A_1, A_2, ... be non-empty sets. Then there exists a function*

$$f : \mathbb{Z}^{++} \to \bigcup_{j=1}^{\infty} A_j$$

such that $f(j) \in A_j$ for all $j \geq 1$.

The standard proof of the next lemma requires the axiom of countable choice.

Lemma 31. *A closed subset of a separable metric space is separable.*

Note that it may not be necessary to use the axiom of countable choice to prove *specific* cases of Lemma 31 when (X, d) carries other structures.

Lemma 32. *A non-empty closed subset of* $[0, 1]$ *(the closed interval with the usual metric) is separable.*

In fact, logicians have shown that several elementary theorems of analysis require a strictly stronger axiom than the axiom of countable choice. The axiom of countable choice enables us to use arguments of the type 'choose $a_1 \in A_1$, choose $a_2 \in A_2$, choose $a_3 \in A_3$ and so on'. However, sometimes we want to use arguments of the type 'choose $a_1 \in A_1$, choose $a_2 \in A_2$ depending on a_1, choose $a_3 \in A_3$ depending on a_1 and a_2 and so on'. For this we require an axiom which logicians have shown to be strictly stronger than the axiom of countable choice.

The axiom of countable dependent choice *Suppose that X and R are sets with $R \subseteq X \times X$ and such that*

$$\{y \in X \,:\, (x, y) \in R\} \neq \varnothing$$

for all $x \in X$. Then there exists a function $g : \mathbb{Z}^{++} \to X$ such that $\big(g(n), g(n+1)\big) \in R$ for all $n \geq 1$.

We can easily obtain the version of the axiom which we usually use.

Lemma 33. *Let A_1, A_2, ... and Y_1, Y_2, ... be sets with the following properties.*

(i) $A_1 = Y_1$ and A_1 is non-empty.

(ii) $Y_n \subseteq A_1 \times A_2 \times \cdots \times A_n$ for all $n \geq 1$.

(iii) If $a_j \in A_j$ for $1 \leq j \leq n$ then the set

$$\{\mathbf{y} \in Y_{n+1} \,:\, y_j = a_j \text{ for all } 1 \leq j \leq n\} \neq \varnothing.$$

for all $n \geq 1$

Then there exists a function

$$f : \mathbb{Z}^{++} \to \bigcup_{j=1}^{\infty} A_j$$

such that

$$\big(f(1),\, f(2),\, \ldots,\, f(n)\big) \in Y_n$$

for all $n \geq 1$.

If we now look at our proof of the Baire category theorem (Theorem 5) she will see that we made use of the axiom of countable dependent choice.

A clever argument of Blair show that we cannot avoid this and that indeed *the Baire category theorem is equivalent to the axiom!*

Theorem 34. *The following two statements are equivalent.*

(A) If (E, d) is a complete metric space, then E cannot be written as the union of a countable collection of closed sets with empty interior.

(B) Suppose that X and R are sets with $R \subseteq X \times X$ and such that

$$\{y \in X \,:\, (x, y) \in R\} \neq \varnothing$$

for all $x \in X$. Then there exists a function $G : \mathbb{Z}^{++} \to X$ such that $\big(G(n), G(n+1)\big) \in R$ for all $n \geq 1$.

We have seen that use of the axiom of countable dependent choice is deeply embedded in ordinary analysis. Logicians have shown that if the other rules of reasoning and axioms of set theory are free from contradiction then adding this axiom will not produce contradictions. The mode of reasoning expressed by the axiom seems so natural to almost all mathematicians that they are happy to accept it[5].

[5]It is possible to produce a version of analysis which does not use the axiom or any related 'non-constructive rule of reasoning'. Bishop's book [1] shows that the result is of comparable elegance to ordinary analysis but it is a very different theory.

6 Zorn's lemma and Tychonov's theorem

In the previous section we considered 'making a countable number of choices'. In this section we consider the full axiom of choice.

The axiom of choice *Let $\mathcal{A}$ be a non-empty collection of non-empty sets. Then there exists a function*

$$f : \mathcal{A} \to \bigcup_{A \in \mathcal{A}} A$$

such that $f(A) \in A$ for all $A \in \mathcal{A}$.

Most mathematicians are happy to add the axiom of choice to the standard axioms and this is what we shall do. Note that if we prove something using the standard axioms and the axiom of choice then we will be unable to find a counter-example using only the standard axioms. Note also that, when dealing with specific systems we may be able to prove the result for that system without using the axiom of choice.

The axiom of choice is not very easy to use in the form that we have stated it and it is usually more convenient to use an equivalent formulation called Zorn's lemma.

Definition 35. *Suppose X is a non-empty set. We say that $\succeq$ is a partial order on X, that is to say, that $\succeq$ is a relation on X with*

(i) $x \succeq y$, $y \succeq z$ implies $x \succeq z$,

(ii) $x \succeq y$ and $y \succeq x$ implies $x = y$,

(iii) $x \succeq x$

for all x, y, z.

We say that a subset C of X is a chain if, for every x, $y \in C$ at least one of the statements $x \succeq y$, $y \succeq x$ is true.

If Y is a non-empty subset of X we say that $z \in X$ is an upper bound for Y if $z \succeq y$ for all $y \in Y$.

We say that m is a maximal element for $(X, \succeq)$ if $x \succeq m$ implies $x = m$.

You must be able to do the following exercise.

Exercise 36. *(i) Give an example of a partially ordered set which is not a chain.*

(ii) Give an example of a partially ordered set and a chain C such that (a) the chain has an upper bound lying in C, (b) the chain has an upper bound but no upper bound within C, (c) the chain has no upper bound.

(iii) If a chain C has an upper bound lying in C, show that it is unique. Give an example to show that, even in this case C may have infinitely many upper bounds (not lying in C).

(iv) Give examples of partially ordered sets which have (a) no maximal elements, (b) exactly one maximal element, (b) infinitely many maximal elements.

(v) how should a minimal element be defined? Give examples of partially ordered sets which have (a) no maximal or minimal elements, (b) exactly one maximal element and no minimal element, (c) infinitely many maximal elements and infinitely many minimal elements.

Axiom 37 (Zorn's lemma). *Let $(X, \succeq)$ be a partially ordered set. If every chain in X has an upper bound then X contains a maximal element.*

Zorn's lemma is associated with a proof routine which we illustrate in Lemmas 38 and 40

Lemma 38. *Zorn's lemma implies the axiom of choice.*

The converse result is less important to us but we prove it for completeness.

Lemma 39. *The axiom of choice implies Zorn's lemma.*

Proof. (Since the proof we use is non-standard, I give it in detail.) Let X be a non-empty set with a partial order $\succeq$ having no maximal elements. We show that the assumption that every chain has a upper bound leads to a contradiction.

We write $x \succ y$ if $x \succeq y$ and $x \neq y$. If C is a chain we write

$$C_x = \{c \in C \,:\, x \succ c\}.$$

Observe that, if C is a chain in X, we can find an $x \in X$ such that $x \succ c$ for all $c \in C$. (By assumption, C has an upper bound, x', say. Since X has no maximal elements, we can find an $x \in X$ such that $x \succ x'$.) We shall take $\varnothing$ to be a well ordered chain.

We shall look at well ordered chains, that is to say, chains for which every non-empty subset has a minimum. (Formally, if $S \subseteq C$ is non-empty we can find an $s_0 \in C$ such that $s \succeq s_0$ for all $s \in S$. We write $\min C = s_0$.) By the previous paragraph

$$A_C = \{x \,:\, x \succ c \text{ for all } c \in C\} \neq \varnothing.$$

Thus, if we write $\mathcal{W}$ for the set of all well ordered chains, the axiom of choice tells us that there is a function $\kappa : \mathcal{W} \to X$ such that $\kappa(C) \succ c$ for all $c \in C$.

We now consider 'special chains' defined to be well ordered chains C such that

$$\kappa(C_x) = x \text{ for all } x \in C.$$

(Note that 'well ordering' is an important general idea, but 'special chains' are an ad hoc notion for this particular proof. Note also that if C is a special chain and $x \in C$ then C_x is a special chain.)

The key point is that, if K and L are special chains, then either $K = L$ or $K = L_x$ for some $x \in L$ or $L = K_x$ for some $x \in K$.

Subproof If $K = L$, we are done. If not, at least one of $K \setminus L$ and $L \setminus K$ is non-empty. Suppose, without loss of generality, that $K \setminus L \neq \varnothing$. Since K is well ordered, $x = \min K \setminus L$ exists. We observe that $K_x \subseteq L$. If $K_x = L$, we are done.

We show that the remaining possibility $K_x \neq L$ leads to contradiction. In this case, $L \setminus K_x \neq \varnothing$ so $y = \min L \setminus K_x$ exists.

If $L_y = K_x$ then

$$y = \kappa(L_y) = \kappa(K_x) = x$$

so $x = y \in L \cap K$ contradicting the statement that $x \in K \setminus L$.

If $L_y \neq K_x$ let z be the least member of $K_x \setminus L_y$. Observe that, since $K_x \subseteq L$ and so

$$w \in L_y,\ z' \in K_x,\ w \succ z' \Rightarrow z' \in L,\ y \succ w \succ z' \Rightarrow z' \in L_y$$

whence

$$z' \notin L_y,\ z' \in K_x,\ w \succ z' \Rightarrow w \notin L_y.$$

Thus $K_z = L_y$ and

$$y = \kappa(L_y) = \kappa(K_z) = z$$

so $y = z \in K \cap L$ contradicting the definition of y.

End subproof

We now take S to be the union of all special chains. Using the key observation, it is routine to see that:

(i) S is a chain. (If a, $b \in S$, then $a \in L$ and $b \in K$ for some special chains. By our key observation, either $L \supseteq K$ of $K \supseteq L$. Without loss of generality, $K \supseteq L$ so a, $b \in K$ and $a \succeq b$ or $b \succeq a$.)

(ii) If $a \in S$, then S_a is a special chain. (We must have $a \in K$ for some special chain K. Since $K \subseteq S$, we have $K_a \subseteq S_a$. On the other hand, if $b \in S_a$ then $b \in L$ for some special chain L and each of the three possible relationships given in our key observation imply $b \in K_a$. Thus $S_a \subseteq K_a$, so $S_a = K_a$ and S_a is a special chain.)

(iii) S is well ordered. (If E is a non empty subset of S, pick an $x \in E$. If $S_x \cap E = \varnothing$, then x is a minimum for E. If not, then $S_x \cap E$ is a non-empty subset of the special, so well ordered, chain S_x, so $\min S_x \cap E$ exists and is a minimum for E.)

(iv) S is a special chain. (If $x \in S$, we can find a special chain K such that $x \in K$. Let $y = \kappa(K)$. Then $L = K \cup \{y\}$ is a special chain. As in (ii), $S_y = L_y$, so $S_x = L_x$ and $\kappa(S_x) = \kappa(L_x) = x$.)

We can now swiftly obtain a contradiction. Since S is well ordered $\kappa(S)$ exists and does not lie in S. But S is special, so $S \cup \kappa(S)$ is, so $S \cup \kappa(S) \subseteq S$, so $\kappa(S)$ lies in S. The required result follows by reductio ad absurdum[6]. $\square$

Lemma 40 (Hamel basis theorem). *(i) Every vector space has a basis.*

(ii) If U is an infinite dimensional normed space over $\mathbb{F}$ then we can find a linear map $T : U \to \mathbb{F}$.

(iii) If U is an infinite dimensional normed space over $\mathbb{F}$ then we can find a discontinuous linear map $T : U \to \mathbb{F}$.

[Note that we do not claim that T in (ii) is continuous.]

Exercise 41. *(i) Show that, if $f : \mathbb{R} \to \mathbb{R}$ is continuous and satisfies the equation*

$$f(x + y) = f(x) + f(y)$$

for all $x,\ y \in \mathbb{R}$, then there exists a c such that $f(x) = cx$ for all $x \in \mathbb{R}$.

(ii) Show that there exists a discontinuous function $f : \mathbb{R} \to \mathbb{R}$ satisfying the equation

$$f(x + y) = f(x) + f(y)$$

for all $x,\ y \in \mathbb{R}$.
[Hint. Consider $\mathbb{R}$ as a vector space over $\mathbb{Q}$.]

The rest of this section is devoted to a proof of Tychonov's theorem. We recall a definition.

Definition 42. *Let (X_α, τ_α) be a topological space for each $\alpha \in A$. The product (or Tychonov or weak) topology τ on $\prod_{\alpha \in A} X_\alpha$ is the collection of sets U such that if $\mathbf{u} \in U$ we can find $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n \in A$ and $O_{\alpha_j} \in \tau_{\alpha_j}$ such that $u_{\alpha_j} \in O_{\alpha_j}$ for $1 \le j \le n$ and $\mathbf{x} \in U$ whenever $x_{\alpha_j} \in O_{\alpha_j}$ for $1 \le j \le n$.*

Exercise 43. *We retain the notation of Definition 42 and write $\pi_\alpha \mathbf{x} = x_\alpha$.*

(i) Show that τ is indeed a topology and that, with this topology, the maps $\pi_\alpha : X \to X_\alpha$ are continuous.

(ii) Show that τ is the weakest topology for which the π_α are continuous. [Thus, if σ is a topology for which the π_α are continuous, we have $\sigma \supseteq \tau$.]

(iii) Show that if all the τ_α are Hausdorff so is τ.

[6] To the best of my knowledge, this particular proof is due to Jonathon Letwin (*American Mathematical Monthly*, Volume 98, **1991**, pp. 353–4). If you know about transfinite induction, there are more direct proofs.

(iv) Suppose that $A = [0,1]$ and $X_t = \mathbb{R}$ and τ_t is the usual Euclidean topology on X_t. Explain how an $f \in \prod_{t\in[0,1]} X_t$ can be identified in a natural manner with a function $f : [0,1] \to \mathbb{R}$. With this identification show that the sequence of functions $f_n \to f$ pointwise if and only if, given any $U \in \tau$ with $f \in U$ we can find an N such that $f_n \in U$ for all $n \geq N$.

Theorem 44 (Tychonov). *The product of compact spaces is itself compact under the weak topology.*

We follow the presentation in [2]. (The method of proof is due to Bourbaki.)

The following result should be familiar to almost all of my readers.

Lemma 45 (Finite intersection property). *(i) A topological space is compact if and only if whenever a non-empty collection of closed sets $\mathcal{F}$ has the property that $\bigcap_{j=1}^n F_j \neq \varnothing$, for any $F_1, F_2, \ldots, F_n \in \mathcal{F}$ it follows that $\bigcap_{F\in\mathcal{F}} F \neq \varnothing$.*

(ii) A topological space is compact if and only if whenever a non-empty collection of sets $\mathcal{A}$ has the property that $\bigcap_{j=1}^n A_j \neq \varnothing$ for any $A_1, A_2, \ldots, A_n \in \mathcal{A}$ it follows that $\bigcap_{A\in\mathcal{A}} \bar{A} \neq \varnothing$.

Definition 46. *A system $\mathcal{F}$ of subsets of a given set S is said to be of finite character if*

(i) whenever every finite subset of a set $A \subseteq S$ belongs to $\mathcal{F}$ it follows that $A \in \mathcal{F}$ and

(ii) whenever $A \in \mathcal{F}$ every finite subset of A belongs to $\mathcal{F}$.

Lemma 47 (Tukey's lemma). *If a system $\mathcal{F}$ of subsets of a given set S has finite character and $F \in \mathcal{F}$ then $\mathcal{F}$ has a maximal (with respect to inclusion) element containing F.*

We now prove Tychonov's theorem.

Exercise 48. *If (X, τ) is a Hausdorff space and $\mathcal{G}$ is a maximal collection of sets with the finite intersection property explain why $\bigcap_{G\in\mathcal{G}} \bar{G}$ consists of one point.*

If you are interested, examine how the second (but not the first) appeal to the axiom of choice may be avoided in our proof of Tychonov's theorem if all our spaces are Hausdorff.

The reason why Tychonov's theorem demands the axiom of choice is made clear by the final result of this section.

Lemma 49. *Tychonov's theorem implies the axiom of choice.*

7 The Hahn-Banach theorem

A good example of the use of Zorn's lemma occurs when we ask if given a Banach space $(U, \| \ \|)$ (over $\mathbb{C}$, say) there exist any non-trivial continuous linear maps $T : U \to \mathbb{C}$. For any space that we can think of, the answer is obviously yes, but to show that the result is always yes we need Zorn's lemma[7]. Our proof uses the theorem of Hahn-Banach.

One form of this theorem is the following.

Theorem 50. (Hahn–Banach) *Let U be a real vector space. Suppose $p : U \to \mathbb{R}$ is such that*

$$p(u+v) \leq p(u) + p(v) \text{ and } p(au) = ap(u)$$

for all u, $v \in U$ and all real and positive a.

If E is a subspace of U and there exists a linear map $T : E \to \mathbb{R}$ with $Tx \leq p(x)$ for all $x \in E$ then there exists a linear map $\tilde{T} : U \to \mathbb{R}$ with $Tx \leq p(x)$ for all $x \in U$ and $\tilde{T}(x) = Tx$ for all $x \in E$.

[Note that we do not assume that the vector space U is normed but we do assume that the vector space is real.]

Exercise 51. *We say that a function $f : \mathbb{R} \to \mathbb{R}$ is convex if*

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y)$$

for all x, $y \in \mathbb{R}$. Using the ideas of the proof of the Hahn–Banach theorem but not the result itself show that a convex function f is continuous and that given any $a \in \mathbb{R}$ we can find a c such that

$$f(x) \geq f(a) + c(x-a)$$

for all $x \in \mathbb{R}$. Give an example to show that f need not be differentiable.

Exercise 52. *Let X be a real vector space and p, $q : X \to \mathbb{R}$ be functions such that $p(\lambda x) = \lambda p(x)$, $q(\lambda x) = \lambda q(x)$ for all $\lambda \in \mathbb{R}$ with $\lambda \geq 0$ and all $x \in X$, whilst*

$$p(x+y) \leq p(x) + p(y), \ q(x) + q(y) \leq q(x+y)$$

for all x, $y \in X$.

[7] In fact the statement is marginally weaker than Zorn's lemma but you need to be logician either to know or care about this.

(i) Suppose that Y is a subspace of X and $S : Y \to \mathbb{R}$ a linear function such that

$$S(y) \leq p(x+y) - q(x)$$

for all $x \in X$, $y \in Y$. Show that

$$S(y') - p(x' + y' - z) + q(x') \leq -S(y) + p(x + y + z) - q(x)$$

for all x, x', $z \in X$ and y, $y' \in Y$.

(ii) Suppose that Y_0 is a subspace of X and $T_0 : Y \to \mathbb{R}$ a linear function such that

$$T_0(y) \leq p(x+y) - q(x)$$

for all $x \in X$, $y \in Y_0$. Show that there exists a linear function $T_0 : X \to \mathbb{R}$ such that

$$T(y) \leq p(x+y) - q(x)$$

for all x, $y \in X$ and $Tu = Tu_0$ for all $u \in Y_0$. Show that

$$q(x) \leq T(x) \leq p(x)$$

for all $x \in X$.

(iii) Suppose $p(x) \geq q(x)$ for all $x \in X$. Show that there exists a linear function (possibly the zero function) $U : X \to \mathbb{R}$ such that

$$q(x) \leq U(x) \leq p(x)$$

for all $x \in X$.

(iv) Let $X = \mathbb{R}^2$, $\mathbf{n}$ be a unit vector and $p(\mathbf{x}) = |\mathbf{n}.\mathbf{x}|$ (the absolute value of the usual inner product) and $q(\mathbf{x}) = -|\mathbf{n}.\mathbf{x}|$. Show that p and q obey the conditions of the introductory paragraph and part (iii). What can you say about U?

We have the following important corollary to Theorem 50.

Theorem 53. *Let $(U, \| \ \|)$ be a real normed vector space. If E is a subspace of U and there exists a continuous linear map $T : E \to \mathbb{R}$, then there exists a continuous linear map $\tilde{T} : U \to \mathbb{R}$ with $\|\tilde{T}\| = \|T\|$.*

The next result is famous as 'the result that Banach did not prove'.

Theorem 54. *Let $(U, \| \ \|)$ be a complex normed vector space. If E is a subspace of U and there exists a continuous linear map $T : E \to \mathbb{C}$ then there exists a continuous linear map $\tilde{T} : U \to \mathbb{C}$ with $\|\tilde{T}\| = \|T\|$.*

We can now answer the question posed in the first sentence of this section.

Lemma 55. *If $(U, \| \ \|)$ is normed space over the field $\mathbb{F}$ of real or complex numbers and $a \in U$ with $a \neq 0$, then we can find a continuous linear map $T : U \to \mathbb{F}$ with $Ta \neq 0$.*

The importance of this lemma becomes more obvious if we state it in reverse. If $Ta = 0$ for all continuous linear maps then $a = 0$.

Some analysts are consciously or unconsciously unwilling to use the Hahn–Banach theorem because of its association with the axiom of choice. This is a MISTAKE. If we have a 'concretely presented' normed vector space then we can prove the appropriate Hahn–Banach using, at worst, countable dependent choice[8].

Exercise 56. *If Let $(U, \| \ \|)$ be a real separable normed vector space. If E is a subspace of U and there exists a continuous linear map $T : E \to \mathbb{R}$, show using only countable dependent choice that there exists a continuous linear map $\tilde{T} : U \to \mathbb{R}$ with $\|\tilde{T}\| = \|T\|$.*

Here are a couple of results proved by Banach using the full power of his theorem.

Theorem 57 (Generalised limits)**.** *Consider the vector space l^∞ of bounded real sequences. There exists a linear map $L : l^\infty \to \mathbb{R}$ such that*

(i) If $x_n \geq 0$ for all n then $L\mathbf{x} \geq 0$.

(ii) $L((x_1, x_2, x_3, \dots)) = L((x_0, x_1, x_2, \dots))$.

(iii) $L((1, 1, 1, \dots)) = 1$.

The theorem is illustrated by the following lemma.

Lemma 58. *Let L be as in Theorem 57. Then*

$$\limsup_{n\to\infty} x_n \geq L(\mathbf{x}) \geq \liminf_{n\to\infty} x_n.$$

In particular, if $x_n \to x$ then $L(\mathbf{x}) = x$.

Exercise 59. *(i) Show that, even though the sequence $x_n = (-1)^n$ has no limit, $L(\mathbf{x})$ is uniquely defined.*

(ii) Find, with reasons, a sequence $\mathbf{x} \in l^\infty$ for which $L(\mathbf{x})$ is not uniquely defined.

Banach used the same idea to prove the following odd result.

[8]This sentence is a slogan but a good one.

Lemma 60. *Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the unit circle and let $B(\mathbb{T})$ be the vector space of real valued bounded functions. Then we can find a linear map $I : B(\mathbb{T}) \to \mathbb{R}$ obeying the following conditions.*

(i) $I(1) = 1$.

(ii) $If \geq 0$ if f is positive.

(iii) If $f \in B(\mathbb{T})$, $a \in \mathbb{T}$ and we write $f_a(x) = f(x-a)$ then $If_a = If$.

Exercise 61. *Show that if I is as in Lemma 60 and f is Riemann integrable then*

$$If = \int_{\mathbb{T}} f(t)\, dt.$$

However, Lemma 60 is put in context by the following.

Lemma 62. *Let G be the group freely generated by two generators and $B(G)$ be the vector space of real valued bounded functions on G. If $f \in B(G)$ let us write $f_c(x) = f(xc^{-1})$ for all $x,\ c \in G$.*

There exists a function $f \in B(G)$ and $c_1,\ c_2,\ c_3$ such that $f(x) \geq 0$ for all $x \in G$ and

$$f(x) + f_{c_1}(x) - f_{c_2}(x) - f_{c_3}(x) \leq -1$$

for all $x \in G$.

Exercise 63. *If G is as in Lemma 62 then there is no linear map $I : B(G) \to \mathbb{R}$ obeying the following conditions.*

(i) $I(1) = 1$.

(ii) $If \geq 0$ if f is positive.

(iii) $If_c = If$ for all $c \in G$.

It can be shown that there is a finitely additive, congruence respecting integral for $\mathbb{R}$ and $\mathbb{R}^2$ but not $\mathbb{R}^n$ for $n \geq 3$.

8 Three more uses of Hahn-Banach

The following exercise provides background for our first discussion but is not examinable. For the moment $C([a, b])$ will be the set of real valued continuous functions.

Exercise 64. *We say that a function $G : [a, b] \to \mathbb{R}$ is of bounded variation if there exists a K such that whenever we have a dissection*

$$\mathcal{D} = \{x_0,\ x_1,\ x_2,\ \ldots,\ x_n\}$$

$a = x_0 < x_1 < x_2 < \cdots < x_n = b$ *we have*

$$\sum_{j=1}^{n} |G(x_j) - G(x_{j-1})| \leq K.$$

We write

$$\|G\|_{BV} = \sup_{\mathcal{D}} \sum_{j=1}^{n} |G(x_j) - G(x_{j-1})|$$

where the supremum is taken over all possible dissections.

Suppose $f : [a,b] \to \mathbb{R}$ *is continuous. Let us write*

$$S(\mathcal{D}, f, G) = \sum_{j=1}^{n} f(x_j)\big(G(x_j) - G(x_{j-1})\big).$$

If $\mathcal{D} = \{x_0,\ x_1,\ x_2,\ \ldots,\ x_n\}$ *and* $\mathcal{D}' = \{x'_0,\ x'_1,\ x'_2,\ \ldots,\ x'_{n'}\}$ *are such that* $|f(t)-f(s)| < \epsilon$ *for all* $t,\ s \in [x_{j-1}, x_j]$ $[1 \leq j \leq n]$ *and for all* $t,\ s \in [x'_{j-1}, x'_j]$ $[1 \leq j \leq n']$ *show by considering* $\mathcal{D} \cup \mathcal{D}'$, *or otherwise that*

$$|S(\mathcal{D}, f, G) - S(\mathcal{D}', f, G)| \leq 2K\epsilon.$$

Hence, or otherwise, show that there exists a unique $I(f,G)$ *such that, given any* $\epsilon > 0$ *we can find a* $\delta > 0$ *such that, given any*

$$\mathcal{D} = \{x_0,\ x_1,\ x_2,\ \ldots,\ x_n\}$$

with $|x_{j-1} - x_j| < \delta$ $[1 \leq j \leq n]$ *we have*

$$|S(\mathcal{D}, f, G) - I(f, G)| < \epsilon.$$

We write

$$I(f,G) = \int_a^b f(t)\, dG(t).$$

(i) Let $[a,b] = [0,1]$*. Find elementary expressions for* $\int_a^b f(t)\, dG(t)$ *in the three cases when* $G(t) = t$, *when* $G(t) = -t$ *and when* $G(t) = 0$ *for* $t < 1/2$, $G(t) = 1$ *for* $t \geq 1/2$.

(ii) Show that the map $T : (C([a,b]), \|\ \|_\infty) \to \mathbb{R}$ *given by*

$$Tf = \int_a^b f(t)\, dG(t)$$

is linear and continuous with $\|T\| \leq \|G\|_{BV}$.

(In order to obtain the more satisfactory result $\|T\| = \|G\|_{BV}$ we must put an extra condition on G such as left continuity.)

Exercise 65. *(i) Suppose $G : \mathbb{R} \to \mathbb{R}$ is of bounded variation in every interval $[a, b]$. Show that if we write*

$$G_+(t) = G(a) + \sup\left\{\sum_{j=1}^{m} G(b_j) - G(a_j) \ : \ a \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_m \leq b_m \leq t, \ m \geq 1\right\}$$

and

$$G_-(t) = \sup\{\sum_{j=1}^{m} G(a_j) - G(b_j) \ : \ a \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_m \leq b_m \leq t, \ m \geq 1\}$$

then G_+, $G_- : [a, b] \to \mathbb{R}$ are increasing functions with

$$G(t) = G_+(t) - G_-(t)$$

and $\|G\|_{BV} = (G_+(b) - G_+(a)) + (G_-(b) - G_-(a))$.

(ii) By first considering increasing functions, or otherwise, show that if G is a function of bounded variation in every interval then the left and right limits

$$G(t+) = \lim_{h \to 0, h > 0} G(t + h), G(t-) = \lim_{h \to 0, h > 0} G(t - h)$$

exist everywhere.

Theorem 66. *If $T : C([a, b]) \to \mathbb{R}$ is a continuous linear function then we can find a left continuous function $G : \mathbb{R} \to \mathbb{R}$ of bounded variation in every interval such that*

$$Tf = \int_a^b f(t)\, dG(t)$$

for all $f \in C([a, b])$.

If you know a little measure theory you can restate the theorem in more modern language.

Theorem 67. (The Riesz representation theorem.) *The dual of $C([a, b])$ is the space of Borel measures on $[a, b]$.*

The method used can easily be extended to all compact spaces.

Our second result is more abstract. We require Aloaoglu's theorem.

Theorem 68. *The unit ball of the dual of a normed space X is compact in the weak star topology.*

Our proof of the Riesz representation theorem used the Hahn-Banach theorem as a convenience. Our proof of the next result uses it as basic ingredient.

Theorem 69. *Every Banach space is isometrically isomorphic to some subspace of $C(K)$ for some compact space K.*

(In my opinion this result looks more interesting than it is.)

Our third result requires us to recast the Hahn Banach theorem in a geometric form.

Lemma 70. *If V is a real normed space and E is a convex subset of V containing $B(\mathbf{0}, \epsilon)$ for some $\epsilon > 0$, then, given any $\mathbf{x} \notin E$ we can find a continuous linear map $T : V \to \mathbb{R}$ such that $T\mathbf{x} = 1 \geq T\mathbf{e}$ for all $\mathbf{e} \in E$.*

Theorem 71. *If V is a real normed space and F is a closed convex subset of V, then, given any $\mathbf{x} \notin F$ we can find a continuous linear map $T : V \to \mathbb{R}$ and a real α such that $T\mathbf{x} > \alpha > T\mathbf{k}$ for all $\mathbf{k} \in F$.*

Definition 72. *Let V be a real or complex vector space. If K is a non-empty subset of V we say that $E \subseteq K$ is an* extreme set *of K if, whenever $u,\ v \in K$, $1 > \lambda > 0$ and $\lambda u + (1 - \lambda)v \in E$, it follows that $u,\ v \in E$. If $\{e\}$ is an extreme set we call e an extreme point.*

Exercise 73. *Define an extreme point directly.*

Exercise 74. *We work in $\mathbb{R}^2$. Find the extreme points, if any, of the following sets and prove your statements.*

(i) $E_1 = \{\mathbf{x} : \|\mathbf{x}\| < 1\}$.
(ii) $E_2 = \{\mathbf{x} : \|\mathbf{x}\| \leq 1\}$.
(iii) $E_3 = \{(x, 0) : x \in \mathbb{R}\}$.
(iv) $E_4 = \{(x, y) : |x|, |y| \leq 1\}$.

Theorem 75. (Krein–Milman). *A non-empty compact convex subset K of a normed vector space has at least one extreme point.*

Theorem 76. *A non-empty compact convex subset K of a normed vector space is the closed convex hull of its extreme points (that is, is the smallest closed convex set containing its extreme points).*

Our hypotheses in our version of the Krein–Milman theorem are so strong as to make the conclusion practically useless. However the hypotheses can be much weakened as is indicated by the following version.

Theorem 77. (Krein–Milman). *Let E be the dual space of a normed vector space. A non-empty convex subset K which is compact in the weak star topology has at least one extreme point.*

Theorem 78. *Let E be the dual space of a normed vector space. A non-empty convex subset K which is compact in the weak star topology is the weak star closed convex hull of its extreme points.*

The results follow at once from a new version of Lemma 70

Lemma 79. *If V is a real normed space and E is a convex subset of the dual space V' containing an open (in the wek topology) neighbourhood of $\mathbf{0}$, then, given any $\mathbf{x} \notin E$ we can find a continuous (in the weak topology) linear map $T : V' \to \mathbb{R}$ such that $T\mathbf{x} = 1 \geq T\mathbf{e}$ for all $\mathbf{e} \in E$.*

Lemma 80. *The extreme points of the closed unit ball of the dual of $C([0,1])$ are the delta masses δ_a and $-\delta_a$ with $a \in [0,1]$.*

9 The Rivlin-Shapiro formula

In this section we give an elegant use of extreme points due to Rivlin and Shapiro.

Lemma 81. Carathéodory *We work in $\mathbb{R}^n$. Suppose that $\mathbf{x} \in \mathbb{R}^n$ and we are given a finite set of points $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_N$ and positive real numbers λ_1, λ_2, …, λ_N such that*

$$\sum_{j=1}^{N} \lambda_j = 1, \ \sum_{j=1}^{N} \lambda_j \mathbf{e}_j = \mathbf{x}.$$

Then **after renumbering** *the $\mathbf{e}_j$ we can find positive real numbers λ_1', λ_2', …, λ_m' with $m \leq n+1$ such that*

$$\sum_{j=1}^{m} \lambda_j' = 1, \ \sum_{j=1}^{m} \lambda_j' \mathbf{e}_j = \mathbf{x}.$$

Exercise 82. *Show by means of an example that we can not necessarily take $m = n$ in Carathéodory's lemma.*

Lemma 83. *Consider $\mathcal{P}_n$, the subspace of $C([-1,1])$ consisting of real polynomials of degree n or less. If $S : \mathcal{P}_n \to \mathbb{R}$ is linear then we can find an*

$N \leq n+2$ and distinct points $x_1, x_2, \dots, x_N \in [-1,1]$ and non-zero real numbers $\lambda_1, \lambda_2, \dots, \lambda_N$ such that

$$\sum_{j=1}^{N} |\lambda_j| = 1, \ \|S\| \sum_{j=1}^{N} \lambda_j P(x_j) = SP$$

for all $P \in \mathcal{P}_n$.

Lemma 84. *We continue with the hypotheses and notation of Lemma 83 There exists a $P_* \in \mathcal{P}_n$ such that*

$$P_*(x_j) = \|P_*\|_\infty \operatorname{sgn} \lambda_j$$

for all j with $0 \leq j \leq N$. Further, if $P \in \mathcal{P}_n$ satisfies

$$P(x_j) = \|P\|_\infty \operatorname{sgn} \lambda_j$$

then $\|P\|_\infty \|S\| = SP$.

The following results are of considerable interest in view of Lemma 84.

Lemma 85. *We have $\cos n\theta = T_n(\cos\theta)$ where T_n is a real polynomial of degree n. Further*

(i) $|T_n(x)| \leq 1$ for all $x \in [-1,1]$.

(ii) There exist $n+1$ distinct points $x_1, x_2, \dots, x_{n+1} \in [-1,1]$ such that $|T_n(x_j)| = 1$ for all $1 \leq j \leq n+1$.

Lemma 86. *Suppose that P is a real polynomial of degree n or less such that*

(i) $|P(x)| \leq 1$ for all $x \in [-1,1]$ and

(ii) there exist $n+1$ distinct points $x_1, x_2, \dots, x_{n+1} \in [-1,1]$ such that $|P(x_j)| = 1$ for all $1 \leq j \leq n+1$.
Then $P = \pm T_n$.

Note that Lemma 86 tells us that there is no real polynomial of degree n or less which takes its extreme absolute value on $[-1,1]$ at $n+2$ points. (Thus we can replace the condition $N \leq n+2$ in Lemma 83 by $N \leq n+1$.)

Theorem 87. *If P is a real polynomial of degree at most n and $t \notin [-1,1]$, then*

$$|P(t)| \leq \sup_{|x| \leq 1} |P(x)| |T_n(t)|.$$

Exercise 88. *If P is a real polynomial of degree at most n and $t \notin [-1,1]$, then then*

$$|P^{(r)}(t)| \leq |T^{(r)}(t)| \sup_{|x| \leq 1} |P(x)|.$$

Exercise 89. *(This exercise is part of the course.) (i) Show that if $n \geq 1$ the coefficient of t^n in $T_n(t)$ is 2^{n-1}.*

(ii) Show that if $n \geq 1$ and P is a real polynomial of degree n or less with $|P(t)| \leq 1$ then the coefficient of t^n in $P(t)$ has absolute value at most 2^{n-1}.

(iii) Find, with proof, a polynomial P of degree at most $n-1$ which minimises

$$\sup_{t \in [-1,1]} |t^n - P(t)|.$$

Show that P is unique. (Tchebychev introduced his polynomials T_n in this context.)

10 Uniqueness of Fourier series

The subject of distributions has its roots in the study of partial differential equations and the study of trigonometric series. Most of its applications lie in the study of partial differential equations but I shall consider one from harmonic analysis.

Recall that $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, that the Fourier coefficient $\hat{f}$ of of a continuous function $f : \mathbb{T} \to \mathbb{C}$ is given by

$$\hat{f}(n) = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)\chi_{-n}(t)\, dt$$

where $\chi_n(t) = \exp(2\pi i n t)$. We are used to the idea of studying the Fourier sum

$$\sum_{n=-\infty}^{\infty} \hat{f}(n)\chi_n$$

where f is some continuous function.

Lemma 90. *(i) If $f : \mathbb{T} \to \mathbb{C}$ is continuous and $\hat{f}(n) = 0$ for all $n \in \mathbb{Z}$ then $f = 0$.*

(ii) If $f : \mathbb{T} \to \mathbb{C}$ is continuous and $\sum_{n=-\infty}^{\infty} |\hat{f}(n)| < \infty$ then

$$\sum_{n=-N}^{N} a_n \chi_n(t) \to f(t)$$

as $N \to \infty$ for all $t \in \mathbb{T}$

What happens if we study general trigonometric sums

$$\sum_{n=-\infty}^{\infty} a_n \chi_n$$

with $a_n \in \mathbb{C}$? One of the first questions about such sums is the problem of uniqueness. If

$$\sum_{n=-N}^{N} a_n \chi_n(t) \to 0$$

as $N \to \infty$ for all $t \in \mathbb{T}$ does it follow that $a_n = 0$ for all n?

Exercise 91. *(Easy.) Show that, if* $\sum_{n=-N}^{N} a_n \chi_n(t) \to 0$ *as* $N \to \infty$ *for all* $t \in \mathbb{T}$ *then* $a_n \to 0$ *as* $|n| \to \infty$.

Riemann had the happy idea of considering the effect of formally integrating twice to obtain

$$F(t) = A + Bt + \frac{a_0 t^2}{2} - \sum_{n=-\infty}^{\infty} \frac{a_n}{n^2} \chi_n(t).$$

Exercise 92. *(Easy.) Suppose that* $a_n \to 0$ *as* $|n| \to$. *Explain why* F *is a well defined continuous function.*

When $\sum_{n=-N}^{N} a_n \chi_n(t)$ converges to a certain value we can recover that value by looking at the 'generalised second derivative'

$$\lim_{h \to 0} \frac{F(+h) - 2F(t) + F(t-h)}{4h^2}.$$

Exercise 93. *If* $f : \mathbb{R} \to \mathbb{R}$ *is twice differentiable at* 0 *with* $f(0) = f'(0) = f''(0) = 0$ *use the mean value theorem to show that*

$$\frac{f(h) - 2f(0) + f(-h)}{4h^2} \to 0$$

as $h \to 0$.

Deduce that if $g : \mathbb{R} \to \mathbb{R}$ *is twice differentiable at* 0, *then*

$$\frac{g(h) - 2g(0) + g(-h)}{4h^2} \to g''(0)$$

as $h \to 0$.

Exercise 94. *Suppose that* $a_n \in \mathbb{C}$, $a_n \to 0$ *as* $|n| \to \infty$. *If*

$$F(t) = A + Bt + \frac{a_0 t^2}{2} - \sum_{n=-\infty}^{\infty} \frac{a_n}{n^2} \chi_n(t),$$

show that

$$\frac{F(x+h) - 2F(x) + F(x-h)}{4h^2} = a_0 + \sum_{n \neq 0} a_n \chi_n(x) \left(\frac{\sin 2\pi n h}{nh} \right)^2.$$

Lemma 95. *If $\sum_{n=0}^{\infty} b_n$ converges then*

$$b_0 + \sum_{n=1}^{\infty} b_n \left(\frac{\sin nh}{nh}\right)^2 \to \sum_{n=0}^{\infty} b_n$$

as $h \to 0$.

Exercise 96. *Deduce from Lemma 95 that, if*

$$\sum_{n=-N}^{N} a_n \chi_n(t) \to 0$$

as $N \to \infty$ for all $t \in \mathbb{T}$ and we set

$$F(t) = \frac{a_0 t^2}{2} - \sum_{n=-\infty}^{\infty} \frac{a_n}{n^2} \chi_n(t)$$

then

$$\frac{F(t+h) - 2F(t) + F(t-h)}{4h^2} \to 0$$

as $h \to 0$ for all $t \in \mathbb{T}$

Part of the proof of Lemma 95 rests on ideas which are now familiar.

Exercise 97. *(i) Suppose that $\gamma_n(h) \in \mathbb{C}$ satisfies the following two conditions.*

(A) $\gamma_n(h) \to 0$ as $h \to 0$.

(B) There exists a C such that

$$\sum_{n=0}^{\infty} |\gamma_n(h)| \leq C$$

for all h.

Then, if $t_n \to 0$ as $n \to \infty$ it follows that

$$\sum_{n=0}^{\infty} \gamma_n(h) t_n \to 0$$

as $h \to 0$.

(ii) Suppose, in addition, that

(C) $\displaystyle\sum_{n=1}^{\infty} \gamma_n(h) \to 1$ as $h \to 0$.

Then, if $s_n \to t$ as $n \to \infty$, it follows that

$$\sum_{n=0}^{\infty} \gamma_n(h) s_n \to t$$

as $h \to 0$.

We combine the result of Lemma 96 with a very neat result of Schwarz.

Lemma 98. *Let $f : [a, b] \to \mathbb{R}$ be continuous.*
(i) Suppose that $f(a) = f(b)$ and

$$\limsup_{h \to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{4h^2} > 0$$

for all $x \in (a, b)$. Then $f(x) \leq 0$ for all $x \in [a, b]$.
(ii) Suppose that $f(a) = f(b)$ and

$$\limsup_{h \to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{4h^2} \geq 0$$

for all $x \in (a, b)$. Then $f(x) \leq 0$ for all $x \in [a, b]$
(iii) Suppose that $f(a) = f(b)$ and

$$\frac{f(x+h) - 2f(x) + f(x-h)}{4h^2} \to 0$$

as $h \to 0$ for all $x \in (a, b)$. Then $f(x) = 0$ for all $x \in [a, b]$.
(iv) Suppose that

$$\frac{f(x+h) - 2f(x) + f(x-h)}{4h^2} \to 0$$

as $h \to 0$ for all $x \in (a, b)$. Then there exist constants A and B such that $f(x) = Ax + B$ for all $x \in [a, b]$.

Putting our results together we obtain the following uniqueness theorem.

Theorem 99. *If*

$$\sum_{n=-N}^{N} a_n \chi_n(t) \to 0$$

as $N \to \infty$ for all $t \in \mathbb{T}$ then $a_n = 0$ for all n.

If we try to push matters further we arrive at a natural definition.

Definition 100. *A subset E of $\mathbb{T}$ is called a* set of uniqueness *if*

$$\sum_{n=-N}^{N} a_n \chi_n(t) \to 0$$

as $N \to \infty$ for all $t \notin E$ implies that $a_n = 0$ for all n.

We shall use the theory of distributions to show that every countable closed set is a set of uniqueness.

11 A first look at distributions

We will work in $\mathbb{R}^m$ but we will also keep in mind the simpler case of $\mathbb{T}^m$.

Definition 101. *If $f : \mathbb{R}^m \to \mathbb{R}$ is a continuous function we write*

$$\operatorname{supp} f = \operatorname{Cl}\{x \ : \ f(x) \neq 0\}.$$

Definition 102. *We write $\mathcal{D}(\mathbb{R}^m)$ (or $C_c(\mathbb{R}^n)$) for the set of smooth (ie infinitely differentiable) functions $f : \mathbb{R}^n \to \mathbb{R}$ of compact support. If $f_n, f \in \mathcal{D}$ we say that*

$$f_n \underset{\mathcal{D}}{\to} f$$

if and only if there exists a compact set K such that $\operatorname{supp} f_n \subseteq K$ for each n, and

$$sup_{x \in K} |f_n^{(\mathbf{p})}(x) - f^{(\mathbf{p})}(x)| \to 0$$

for every $\mathbf{p} \in \mathbb{Z}_n^+$.

Here

$$f_n^{(\mathbf{p})}(x) = \partial^{\mathbf{p}} f(x) = \frac{\partial^{p_1+p_2+\ldots+p_m}}{\partial^{p_1} x_1 \partial^{p_2} x_2 \ldots \partial^{p_m} x_n}$$

and $\mathbb{Z}_+$ is the set of non-negative integers.

Before going any further we need to establish that $\mathcal{D}$ is reasonably rich.

To do this we go back to a function introduced by Cauchy as a counterexample.

Exercise 103. *(i) Let $F : \mathbb{R} \to \mathbb{R}$ be given by*

$$F(x) = \begin{cases} \exp(-1/x^2) & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

Show inductively that F is n times differentiable with

$$F^{(n)}(x) = \begin{cases} P_n(x)\exp(-1/x^2) & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

for some polynomial P_n.

(ii) By considering functions of the form $F(x-a)F(a-x)$ show that, given any $\delta > 0$ we can find an infinitely differentiable function $G : \mathbb{R} \to \mathbb{R}$ with $\operatorname{supp} g \subseteq [-\epsilon, \epsilon]$, $G(x) \geq 0$ for all x and $G(0) > 0$.

(iii) By considering functions of the form

$$C \int_{-\infty}^{x} G(t+a) - G(t-a)\, dt$$

show that, given any $\eta > 0$, we can find an infinitely differentiable function $K : \mathbb{R} \to \mathbb{R}$ with

$$K(x) = \begin{cases} 1 & \text{if } |x| \leq 1, \\ 0 & \text{if } 1 + \eta \leq |x|. \end{cases}$$

and $0 \leq K(x) \leq 1$ for all x.

(iv) Show that given any $\eta > 0$, we can find an infinitely differentiable function $E : \mathbb{R}^n \to \mathbb{R}$ with

$$K(x) = \begin{cases} 1 & \text{if } \|x\| \leq 1, \\ 0 & \text{if } 1 + \eta \leq \|x\|. \end{cases}$$

and $0 \leq K(x) \leq 1$ for all x.

Exercise 104. *Identify with reasons the set of compactly supported analytic functions $f : \mathbb{C} \to \mathbb{C}$.*

If $\mathbf{p} \in \mathbb{Z}_+^m$ we write $|\mathbf{p}| = p_1 + p_2 + \ldots + p_m$.

Exercise 105. *Let K be a compact set. We write $\mathcal{D}_k$ for the set of $f \in \mathcal{D}$ such that $\operatorname{supp} f \subseteq K$.*

(i) If f_n, $f \in \mathcal{D}_K$ show that

$$f_n \underset{\mathcal{D}}{\rightarrow} f$$

if and only if

$$\sup_{\|x\| \leq q} |\partial^{\mathbf{p}} f_n(x) - \partial^{\mathbf{p}} f(x)| \to 0$$

for all integers $p \in \mathbb{Z}_+^m$, $q \geq 0$.

(ii) Show that

$$d(f,g) = \sum_{q=0}^{\infty} \sum_{\mathbf{p}\in\mathbb{Z}_+^m} |\mathbf{p}|^{-m} 2^{-|\mathbf{p}|-q} \sup_{\|x\|\leq q} |\partial^{\mathbf{p}} f_n(x) - \partial^{\mathbf{p}} f(x)|$$

gives a well defined metric on $\mathcal{D}_K$.

(iii) If $f_n,\ f \in \mathcal{D}_K$ *show that*

$$f_n \underset{\mathcal{D}}{\to} f \Leftrightarrow d(f_n, f) \to 0$$

Exercise 106. *We show that convergence in distribution cannot be derived from convergence in norm even for* $\mathcal{D}(\mathbb{T}$ *(with the obvious modified definitions).*

Suppose that $\|\ \|_D$ *is a norm on* $\mathcal{D}$ *such that*

$$\|f_n - f\|_D \to 0 \Rightarrow f_n \underset{\mathcal{D}}{\to} f.$$

(i) By reductio ad absurdum, or otherwise, show that, for each integer $j \geq 0$, *there exists an* $\epsilon_j > 0$ *such that we have* $\|f\|_D \geq \epsilon_j$ *whenever* $\sup_{x\in[0,1]} |f^{(j)}(x)| \geq 1$.

(ii) Deduce that $\|g\|_D \geq \epsilon_j \sup_{x\in[0,1]} |g^{(j)}(x)|$ *for all* $g \in C^\infty([0,1])$ *and all* $j \geq 0$.

(iii) Show that, by choosing δ_k *and* N_k *appropriately, and setting* $f_k(x) = \delta_k \sin \pi N_k x$, *or otherwise, that we can find* $f_k \in C^\infty([0,1])$ *such that*

$$\sup_{x\in[0,1]} |f_k^{(j)}(x)| \leq 2^{-k} \qquad \text{when } 0 \leq j \leq k-1,$$
$$\sup_{x\in[0,1]} |f_k^{(k)}(x)| \geq 2^j \epsilon_j.$$

(iv) Show that $\|f_n\|_D \to \infty$ *as* $n \to \infty$, *but* $f_n^{(j)}(x) \to 0$ *uniformly on* $[0,1]$ *for each* j. *Thus*

$$f_n \underset{\mathcal{D}}{\to} f \not\Rightarrow \|f_n - f\|_D \to 0.$$

The fact that $\mathcal{D}$ is not a normed space led to investigations of the much more general 'topological vector spaces' but $\mathcal{D}(\mathbb{T}$ only just fails to be normed since its topology is given by a countable collection of norms.

We can now introduce the notion of a distribution.

Definition 107. *We say that a linear map* $T : \mathcal{D} \to \mathbb{C}$ *is a distribution if*

$$\phi_n \underset{\mathcal{D}}{\to} \phi \Rightarrow T\phi_n \to T\phi.$$

The next lemma gives some equivalent formulations of the definition.

Lemma 108. *If $T : \mathcal{D} \to \mathbb{C}$ then the following statements are equivalent.*

(i) $\phi_n \underset{\mathcal{D}}{\to} \phi \Rightarrow T\phi_n \to T\phi$.

(ii) $\phi_n \underset{\mathcal{D}}{\to} 0 \Rightarrow T\phi_n \to 0$.

(iii) If K is a compact set then there exists a constant $C(K)$ and an integer $N(K) \geq 0$ such that whenever $\phi \in \mathcal{D}$ and $\operatorname{supp}\phi \subseteq K$ we have

$$|T\phi| \leq C(K, \mathbf{p}) \sum_{|\mathbf{p}| \leq N(K)} \|\partial^{\mathbf{p}}\phi\|_\infty.$$

We often write $T\phi = \langle T, \phi\rangle$ and say that $\phi \in \mathcal{D}$ is a *test function*.

Lemma 109. *If $a \in \mathbb{R}^m$, then*

$$\langle \delta_a, \phi\rangle = \phi(a)$$

defines a distribution.

We call δ_a the *Dirac delta function* at a

Lemma 110. *If $f \in C(\mathbb{R}^m)$ then*

$$\langle T_f, \phi\rangle = \int_{\mathbb{R}^m} f(t)\phi(t)\, dt$$

defines a distribution.

Exercise 111. *(If you know some measure theory.) We write $f \in L^1_{loc}(\mathbb{R}^m)$ if f is measurable and $f\mathbb{I}_K \in L^1$ (ie $\int_K |f(t)|\, dt < \infty$) for all compact sets, show that*

$$\langle T_f, \phi\rangle = \int_{\mathbb{R}^m} f(t)\phi(t)\, dt$$

defines a distribution.

We often write $T_f = f$ (whence the name expression ‘generalised function for a distribution) and for this reason we must careful to make our definitions consistent when do so.

Lemma 112. *(i) If $T, S \in \mathcal{D}$ and λ, $\mu \in \mathbb{C}$, then*

$$\langle \lambda T + \mu S, \phi\rangle = \lambda\langle T, \phi\rangle + \mu\langle S, \phi\rangle$$

for $\phi \in \mathcal{D}$ defines a distribution $\lambda T + \mu T$.

(ii) $\mathcal{D}'$ is a vector space with these operations.

(iii) If f and g are continuous functions $T_{\lambda f + \mu g} = \lambda T_f + \mu T_g$.

The next result is more interesting.

Lemma 113. *(i) If* $T \in \mathcal{D}$*, and* $\mathbf{p}$ *has all entries zero except one which has value* 1

$$\langle \partial_{\mathbf{p}} T, \phi \rangle = -\langle T, \partial_{\mathbf{p}} f \rangle$$

for $\phi \in \mathcal{D}$ *defines a distribution* T'.

(iii) If f *is a continuous function with continuous partial derivative,then* $T_{\lambda f + \mu g} = \lambda T_f + \mu T_g$.

We call $\partial_{\mathbf{p}} T$ the partial derivative of T.

Exercise 114. *(i) Distributions are infinitely differentiable with*

$$\langle \partial_{\mathbf{p}} T, \phi \rangle = (-1)^{|\mathbf{p}|} \langle T, \partial_{\mathbf{p}} \phi \rangle.$$

(ii) If we work on $\mathbb{R}$, $\langle \delta_0', \phi \rangle = -\phi'(0)$*. (If what you have been told in the past contradicts this, forget it. This is the correct sign.)*

We can also multiply distributions by test functions.

Lemma 115. *(i) If* $T \in \mathcal{D}$*, and* $\psi \in \mathcal{D}$ *then* $\mathbf{p}$ *has all entries zero except one which has value* 1

$$\langle \psi T, \phi \rangle = \langle T, \phi \psi \rangle$$

for $\phi \in \mathcal{D}$ *defines a distribution* ψT.

(ii) If $T \in \mathcal{D}$*, and* $\psi \in \mathcal{D}$ *then we have the following form of Leibniz's rule*

$$\partial_{\mathbf{p}} \phi T = \sum_{\mathbf{r}+\mathbf{s}=\mathbf{p}} \frac{p!}{r!s!} \partial^{\mathbf{r}} \phi \partial^{\mathbf{s}} T.$$

Here and elsewhere we write $(r_1, r_2, \ldots, r_m)! = r_1! r_2! \ldots r_m!$.

Exercise 116. *Show that, in Lemma 115, we can replace* $\psi \in \mathcal{D}$ *by* $\psi \in C^\infty$ *that is to say,* ψ *a smooth function).*

12 The support of distribution

The statement that T is a distribution seems very abstract but in this section we shall see that T actually 'lives on a well defined region' in $\mathbb{R}^m$.

We use the useful idea of a *partition of unity*.

Lemma 117. *(i) Let K be a compact subset and Ω an open subset of $\mathbb{R}^m$ such that $K \subseteq \Omega$. Then we can find a $\phi \in \mathcal{D}$ such that $0 \leq \phi(x) \leq 1$ for all x, $\phi(x) = 1$ for $x \in K$ and $\phi(x) = 0$ for $x \notin \Omega$.*

(ii) Let K be a compact subset and Ω_j open subsets of $\mathbf{R}^m$ $[1 \leq j \leq J]$ such that $K \subseteq \bigcup_{j=1}^{J} \Omega_j$. Then we can find $\phi_j \in \mathcal{D}$ such that $0 \leq \phi_j(x) \leq 1$ for all x, $\phi(x) = 1$ $\phi_j(x) = 0$ for $x \notin \Omega_j$ $[1 \leq j \leq J]$ and

$$\sum_{j=1}^{J} \phi_j(x) = 1$$

for all $x \in K$.

Definition 118. *If $T \in \mathcal{D}$ then $\operatorname{supp} T$ is the complement of the set of points x with the following property. There exists an $\epsilon > 0$ such that if $\phi \in \mathcal{D}$ and $\operatorname{supp} \phi \subseteq B(x, \epsilon)$ (the open ball centre x radius ϵ) then $\langle T, \phi \rangle = 0$.*

Exercise 119. *(i) If T is a distribution $\operatorname{supp} T$ is closed.*

(ii) If f is continuous then

$$\operatorname{supp} T_f = \operatorname{Cl}\{x \,:\, f(x) = 0\}.$$

In other words the support of f as a distribution coincides with the support of f as a function.

Exercise 120. *(i) If T is a distribution, show that $\operatorname{supp} \partial^{\mathbf{p}} T \subseteq \operatorname{supp} T$. Give an example for $\mathbf{R}$ with $p = 1$ to show that the inclusion may be proper.*

(ii) If $T \in \mathcal{D}'$ and $\phi \in \mathcal{D}$, show that

$$\operatorname{supp} \phi T \subseteq \operatorname{supp} T \cap \operatorname{supp} \phi.$$

Lemma 121. *If $T \in \mathcal{D}'$, $\phi \in \mathcal{D}$ $\operatorname{supp} T \cap \operatorname{supp} \phi = \varnothing$, then*

$$\langle T, \phi \rangle = 0.$$

The following example is extremely important.

Example 122. *We work on $\mathbb{R}$. We have $\operatorname{supp} \delta_0' = \{0\}$ but we can find a $\phi \in \mathcal{D}$ such that $\phi(0) = 0$ and $\langle \delta_0', \phi \rangle \neq 0$.*

Thus

$$\phi(x) = 0 \text{ for all } x \in \operatorname{supp} T \not\Rightarrow \langle T, \phi \rangle = 0.$$

Exercise 123. *If $T \in \mathcal{D}'$, $\phi \in \mathcal{D}$ and $\operatorname{supp} T \cap \operatorname{supp} \phi = \varnothing$, show that $\phi T = 0$.*

We now find all distributions with support consisting of a single point. to do this we require a *local* form of Taylor's theorem

Lemma 124. *If $f : \mathbb{R}^m \to \mathbb{R}$ is infinitely differentiable then*

$$f(h) = \sum_{p_1+p_2+\ldots+p_m \leq N} \frac{(p_1+p_2+\ldots+p_m)!}{p_1!p_2!\ldots p_m!}\partial^{\mathbf{p}} f(0) h_1^{p_1} h_2^{p_2} \ldots h_m^{p_m} + A(h)\|h\|^{N+1}$$

where $A(h)$ remains bounded as $\|h\| \to 0$.

Theorem 125. *If $T \in \mathcal{D}$ and $\operatorname{supp} T \subseteq \{a\}$ then we can find an N and $c_{\mathbf{p}}$ such that*

$$T = \sum_{|\mathbf{p}| \leq N} c_{\mathbf{p}} \partial^{\mathbf{p}} \delta_a.$$

13 Distributions on $\mathbb{T}$

In this section we work work on $\mathbb{T}$. To maintain consistency we modify one of earlier definitions (see Lemma 110 and the associated discussion). by introducing a scaling factor.

Definition 126. *If $f \in C(\mathbb{T})$ then*

$$\langle T_f, \phi \rangle = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)\phi(t)\, dt.$$

If we work on $\mathbb{T}$ the space of distributions takes a particularly simple form.

Definition 127. *If $T \in \mathcal{D}(\mathbb{T})$ we write*

$$\hat{T}(r) = \langle T, \chi_{-r} \rangle.$$

Exercise 128. *If $f \in C(\mathbb{T})$ check that the new and old definitions of $\hat{f}$ coincide.*

Theorem 129. *(i) If $T \in \mathcal{D}'(\mathbb{T})$ and then there exists a N such that $|r|^{-N}\hat{T}(r) \to 0$ as $|r| \to \infty$.*

(ii) If $T \in \mathcal{D}'(\mathbb{T})$ and $\phi \in \mathcal{D}(\mathbb{T})$, then

$$\langle T, \phi \rangle = \sum_{r=-\infty}^{\infty} \hat{T}(r)\hat{\phi}(-r)$$

the convergence being uniform.

(iii) If T, $S \in \mathcal{D}'(\mathbb{T})$,

$$T = S \Leftrightarrow \hat{T}(r) = \hat{S}(r) \text{ for all } r \in \mathbb{Z}.$$

(iv) If $a_r \in \mathbb{C}$ and there exists an N such that $|r|^{-N} a_r \to 0$ as $|r| \to \infty$, then

$$\langle T, \phi \rangle = \sum_{r=-\infty}^{\infty} a_r \hat{\phi}(-r)$$

defines a distribution T with $\hat{T}(r) = a_r$.

We could say that we have identified the distributions with sequences of *polynomial growth*.

We link the theory of distributions with the theory of sets of uniqueness via a version of the Riemann localisation theorem.

Theorem 130. *Suppose that T is a distribution on $\mathbb{D}$ with $\hat{T}(n) \to 0$ and $\phi \in \mathcal{D}$. If*

$$\sum_{j=-n}^{n} \hat{T}(j)\chi_j(t) \to 0$$

as $n \to \infty$ on an open interval I with $I \supseteq \operatorname{supp} \phi$ then writing $S = \phi T$ we have

$$\sum_{j=-n}^{n} \hat{S}(j)\chi_j(t) \to 0$$

everywhere.

Lemma 131. *Suppose that $a_n \to 0$ as $|n| \to \infty$. If T is the distribution with $\hat{T}(n) = a_n$ then*

$$\operatorname{supp} T = \operatorname{Cl}\{t : \sum_{j=-n}^{n} \hat{T}(j)\chi_j(t) \text{ fails to converge to } 0\}.$$

We now make the following observation.

Lemma 132. *Any closed countable subset of $\mathbb{T}$ must contain an isolated point.*

Combining our results gives the following classical theorem

Theorem 133. *Every closed countable subset of $\mathbb{T}$ is a set of uniqueness.*

References

[1] E. Bishop and D. S. Bridges *Constructive Analysis* (Springer, 1985. Included for interest but nothing to do with the course.)

[2] B. Bollobás *Linear Analysis: an Introductory Course* (CUP 1991)

[3] F. G. Friedlander and M. S. Joshi *Introduction to Theory of Distributions* (CUP 1998. This is the second edition with extra material. The first edition by Friedlander alone is ample for this course.)

[4] C. Gofman and G. Pedrick *A First Course in Functional Analysis* (Prentice Hall 1965, available as a Chelsea reprint from the AMS)

[5] J. D. Pryce *Basic Methods of Linear Functional Analysis* (Hutchinson 1973)

[6] W. Rudin *Real and Complex Analysis* (McGraw Hill, 2nd Edition, 1974)

[7] W. Rudin *Functional Analysis* (McGraw Hill 1973)