Version of 3.8.15

Appendix to Volume 2

Useful Facts

In the course of writing this volume, I have found that a considerable number of concepts and facts from various branches of mathematics are necessary to us. Nearly all of them are embedded in important and well-established theories for which many excellent textbooks are available and which I very much hope that you will one day study in depth. Nevertheless, I am reluctant to send you off immediately to courses in general topology, functional analysis and set theory, as if these were essential prerequisites for our work here, along with real analysis and basic linear algebra. For this reason I have written this Appendix, setting out those results which we actually need at some point in this volume. The great majority of them really are elementary – indeed, some are so elementary that they are not always spelt out in detail in orthodox treatments of their subjects.

While I do not put this book forward as the proper place to learn any of these topics, I have tried to set them out in a way that you will find easy to integrate into regular approaches. I do not expect anybody to read systematically through this work, and I hope that the references given in the main chapters of this volume will be adequate to guide you to the particular items you need.

Version of 20.1.13

2A1 Set theory

Especially for the examples in Chapter 21, we need some non-trivial set theory, which is best approached through the standard theory of cardinals and ordinals; and elsewhere in this volume I make use of Zorn's Lemma. Here I give a very brief outline of the results involved, largely omitting proofs. Most of this material should be in any sound introduction to set theory. The references I give are to books which happen to have come my way and which I can recommend as reasonably suitable for beginners.

I do not discuss axiom systems or logical foundations. The set theory I employ is 'naive' in the sense that I rely on my understanding of the collective experience of the last hundred years, rather than on any attempt at formal description, to distinguish legitimate from unsafe arguments. There are, however, points in Volume 5 at which such a relaxed philosophy becomes inappropriate, and I therefore use arguments which can, I believe, be translated into standard Zermelo-Fraenkel set theory without new ideas being invoked.

Although in this volume I use the axiom of choice without scruple whenever appropriate, I will divide this section into two parts, starting with ideas and results not dependent on the axiom of choice (2A1A-2A1I) and continuing with the remainder (2A1J-2A1P). I believe that even at this level it helps us to understand the nature of the arguments better if we maintain a degree of separation.

2A1A Ordered sets (a) Recall that a **partially ordered set** is a set P together with a relation $\leq$ on P such that

if $p \leq q$ and $q \leq r$ then $p \leq r$,
$p \leq p$ for every $p \in P$,
if $p \leq q$ and $q \leq p$ then $p = q$.

In this context, I will write $p \geq q$ to mean $q \leq p$, and $p < q$ or $q > p$ to mean '$p \leq q$ and $p \neq q$'. $\leq$ is a **partial order** on P.

(b) Let $(P, \leq)$ be a partially ordered set, and $A \subseteq P$. A **maximal** element of A is a $p \in A$ such that $p \not< a$ for any $a \in A$. Note that A may have more than one maximal element. An **upper bound** for A is a $p \in P$ such that $a \leq p$ for every $a \in A$; a **supremum** or **least upper bound** is an upper bound p such that $p \leq q$ for every upper bound q of A. There can be at most one such, because if p, p' are both least upper bounds then $p \leq p'$ and $p' \leq p$. Accordingly we may safely write $p = \sup A$ if p is the least upper bound of A.

Similarly, a **minimal** element of A is a $p \in A$ such that $p \not> a$ for every $a \in A$; a **lower bound** of A is a $p \in P$ such that $p \leq a$ for every $a \in A$; and $\inf A = p$ means that

$$\text{for every } q \in P,\ p \geq q \iff a \geq q\ \forall\, a \in A.$$

A subset A of P is **order-bounded** if it has both an upper bound and a lower bound.

A subset A of P is **upwards-directed** if for any p, $p' \in A$ there is a $q \in A$ such that $p \leq q$ and $p' \leq q$; that is, if any non-empty finite subset of A has an upper bound in A. Similarly, $A \subseteq P$ is **downwards-directed** if for any p, $p' \in A$ there is a $q \in A$ such that $q \leq p$ and $q \leq p'$; that is, if any non-empty finite subset of A has a lower bound in A.

It is sometimes convenient to adapt the notation for closed intervals to arbitrary partially ordered sets: $[p, q]$ will be $\{r : p \leq r \leq q\}$.

(c) A **totally ordered set** is a partially ordered set $(P, \leq)$ such that

for any p, $q \in P$, either $p \leq q$ or $q \leq p$.

$\leq$ is then a **total** or **linear** order on P.

In any totally ordered set we have a **median function**: for p, q, $r \in P$ set

$$\begin{aligned} \operatorname{med}(p, q, r) &= \max(\min(p, q), \min(p, r), \min(q, r)) \\ &= \min(\max(p, q), \max(p, r), \max(q, r)), \end{aligned}$$

so that $\operatorname{med}(p, q, r) = q$ if $p \leq q \leq r$.

(d) A **lattice** is a partially ordered set $(P, \leq)$ such that

for any p, $q \in P$, $p \vee q = \sup\{p, q\}$ and $p \wedge q = \inf\{p, q\}$ are defined in P.

(e) A **well-ordered set** is a totally ordered set $(P, \leq)$ such that $\inf A$ exists and belongs to A for every non-empty set $A \subseteq P$; that is, every non-empty subset of P has a least element. In this case $\leq$ is a **well-ordering** of P.

2A1B Transfinite Recursion: Theorem Let $(P, \leq)$ be a well-ordered set and X any class. For $p \in P$ write L_p for the set $\{q : q \in P,\ q < p\}$ and X^{L_p} for the class of all functions from L_p to X. Let $F : \bigcup_{p \in P} X^{L_p} \to X$ be any function. Then there is a unique function $f : P \to X$ such that $f(p) = F(f{\restriction}L_p)$ for every $p \in P$.

proof There are versions of this result in Enderton 77 (p. 175) and Halmos 60 (§18). Nevertheless I write out a proof, since it seems to me that most elementary books on set theory do not give it its proper place at the very beginning of the theory of well-ordered sets.

(a) Let Φ be the class of all functions ϕ such that

(α) $\operatorname{dom}\phi$ is a subset of P, and $L_p \subseteq \operatorname{dom}\phi$ for every $p \in \operatorname{dom}\phi$;

(β) $\phi(p) \in X$ for every $p \in \operatorname{dom}\phi$, and $\phi(p) = F(\phi{\restriction}L_p)$ for every $p \in \operatorname{dom}\phi$.

(b) If ϕ, $\psi \in \Phi$ then ϕ and ψ agree on $\operatorname{dom}\phi \cap \operatorname{dom}\psi$. **P?** If not, then $A = \{q : q \in \operatorname{dom}\phi \cap \operatorname{dom}\psi,\ \phi(q) \neq \psi(q)\}$ is non-empty. Because P is well-ordered, A has a least element p say. Now $L_p \subseteq \operatorname{dom}\phi \cap \operatorname{dom}\psi$ and $L_p \cap A = \emptyset$, so

$$\phi(p) = F(\phi{\restriction}L_p) = F(\psi{\restriction}L_p) = \psi(p),$$

which is impossible. **XQ**

(c) It follows that Φ is a set, since the function $\phi \mapsto \operatorname{dom}\phi$ is an injective function from Φ to $\mathcal{P}P$, and its inverse is a surjection from a subset of $\mathcal{P}P$ onto Φ. We can therefore, without inhibitions, define a function f by writing

$$\operatorname{dom} f = \bigcup_{\phi \in \Phi} \operatorname{dom}\phi, \quad f(p) = \phi(p) \text{ whenever } \phi \in \Phi \text{ and } p \in \operatorname{dom}\phi.$$

(If you think that a function ϕ is just the set of ordered pairs $\{(p,\phi(p)) : p \in \operatorname{dom}\phi\}$, then f becomes $\bigcup\Phi$.) Then $f \in \Phi$. **P** Of course f is a function from a subset of P to X. If $p \in \operatorname{dom} f$, then there is a $\phi \in \Phi$ such that $p \in \operatorname{dom}\phi$, in which case

$$L_p \subseteq \operatorname{dom}\phi \subseteq \operatorname{dom} f, \quad f(p) = \phi(p) = F(\phi\restriction L_p) = F(f\restriction L_p). \textbf{ Q}$$

(d) f is defined everywhere in P. **P?** Otherwise, $P \setminus \operatorname{dom} f$ is non-empty and has a least element r say. Now $L_r \subseteq \operatorname{dom} f$. Define a function ψ by saying that $\operatorname{dom}\psi = \{r\} \cup \operatorname{dom} f$, $\psi(p) = f(p)$ for $p \in \operatorname{dom} f$ and $\psi(r) = F(f\restriction L_r)$. Then $\psi \in \Phi$, because if $p \in \operatorname{dom}\psi$

either $p \in \operatorname{dom} f$ so $L_p \subseteq \operatorname{dom} f \subseteq \operatorname{dom}\psi$ and

$$\psi(p) = f(p) = F(f\restriction L_p) = F(\psi\restriction L_p)$$

or $p = r$ so $L_p = L_r \subseteq \operatorname{dom} f \subseteq \operatorname{dom}\psi$ and

$$\psi(p) = F(f\restriction L_r) = F(\psi\restriction L_r).$$

Accordingly $\psi \in \Phi$ and $r \in \operatorname{dom}\psi \subseteq \operatorname{dom} f$. **XQ**

(e) Thus $f : P \to X$ is a function such that $f(p) = F(f\restriction L_p)$ for every p. To see that f is unique, observe that any function of this type must belong to Φ, so must agree with f on their common domain, which is the whole of P.

Remark If you have been taught to distinguish between the words 'set' and 'class', you will observe that my naive set theory is a relatively tolerant one in that it is willing to allow class variables in its theorems.

2A1C Ordinals An **ordinal** (sometimes called a 'von Neumann ordinal') is a set ξ such that

if $\eta \in \xi$ then η is a set and $\eta \notin \eta$,

if $\eta \in \zeta \in \xi$ then $\eta \in \xi$,

writing '$\eta \le \zeta$' to mean '$\eta \in \zeta$ or $\eta = \zeta$', $(\xi, \le)$ is well-ordered

(Enderton 77, p. 191; Halmos 60, §19; Henle 86, p. 27; Krivine 71, p. 24; Roitman 90, 3.2.8. Of course many set theories do not allow sets to belong to themselves, and/or take it for granted that every object of discussion is a set, but I prefer not to take a view on such points in general.)

2A1D Basic facts about ordinals (a) If ξ is an ordinal, then every member of ξ is an ordinal. (Enderton 77, p. 192; Henle 86, 6.4; Krivine 71, p. 14; Roitman 90, 3.2.10.)

(b) If ξ, η are ordinals then either $\xi \in \eta$ or $\xi = \eta$ or $\eta \in \xi$ (and no two of these can occur together). (Enderton 77, p. 192; Henle 86, 6.4; Krivine 71, p. 14; Lipschutz 64, 11.12; Roitman 90, 3.2.13.) It is customary, in this case, to write $\eta < \xi$ if $\eta \in \xi$ and $\eta \le \xi$ if either $\eta \in \xi$ or $\eta = \xi$. Note that $\eta \le \xi$ iff $\eta \subseteq \xi$.

(c) If A is any non-empty class of ordinals, then there is an $\alpha \in A$ such that $\alpha \le \xi$ for every $\xi \in A$. (Henle 86, 6.7; Krivine 71, p. 15.)

(d) If ξ is an ordinal, so is $\xi \cup \{\xi\}$; call it '$\xi + 1$'. If $\xi < \eta$ then $\xi + 1 \le \eta$; $\xi + 1$ is the least ordinal greater than ξ. (Enderton 77, p. 193; Henle 86, 6.3; Krivine 71, p. 15.) For any ordinal ξ, either there is a greatest ordinal $\eta < \xi$, in which case $\xi = \eta + 1$ and we call ξ a **successor ordinal**, or $\xi = \bigcup\xi$, in which case we call ξ a **limit ordinal**.

(e) The first few ordinals are $0 = \emptyset$, $1 = 0 + 1 = \{0\} = \{\emptyset\}$, $2 = 1 + 1 = \{0, 1\} = \{\emptyset, \{\emptyset\}\}$, $3 = 2 + 1 = \{0, 1, 2\}, \ldots$. The first infinite ordinal is $\omega = \{0, 1, 2, \ldots\}$, which may be identified with $\mathbb{N}$.

(f) The union of any set of ordinals is an ordinal. (Enderton 77, p. 193; Henle 86, 6.8; Krivine 71, p. 15; Roitman 90, 3.2.19.)

(g) If $(P, \le)$ is any well-ordered set, there is a unique ordinal ξ such that P is order-isomorphic to ξ, and the order-isomorphism is unique. (Enderton 77, pp. 187-189; Henle 86, 6.13; Halmos 60, §20.)

2A1E Initial ordinals An **initial ordinal** is an ordinal κ such that there is no bijection between κ and any member of κ. (Enderton 77, p. 197; Halmos 60, §25; Henle 86, p. 34; Krivine 71, p. 24; Roitman 90, 5.1.10, p. 79).

2A1F Basic facts about initial ordinals (a) All finite ordinals, and the first infinite ordinal ω, are initial ordinals.

(b) For every well-ordered set P there is a unique initial ordinal κ such that there is a bijection between P and κ.

(c) For every ordinal ξ there is a least initial ordinal greater than ξ. (Enderton 77, p. 195; Henle 86, 7.2.1.) If κ is an initial ordinal, write κ^+ for the least initial ordinal greater than κ. We write ω_1 for ω^+, ω_2 for ω_1^+, and so on.

(d) For any initial ordinal $\kappa \ge \omega$ there is a bijection between $\kappa\times\kappa$ and κ; consequently there are bijections between κ and κ^r for every $r \ge 1$.

2A1G Schröder-Bernstein theorem I remind you of the following fundamental result: if X and Y are sets and there are injections $f : X \to Y$, $g : Y \to X$ then there is a bijection $h : X \to Y$. (Enderton 77, p. 147; Halmos 60, §22; Henle 86, 7.4; Lipschutz 64, p. 145; Roitman 90, 5.1.2. It is also a special case of 344D in Volume 3.)

2A1H Countable subsets of $\mathcal{P}\mathbb{N}$ The following results will be needed below.

(a) There is a bijection between $\mathcal{P}\mathbb{N}$ and $\mathbb{R}$. (Enderton 77, p. 149; Lipschutz 64, p. 146.)

(b) Suppose that X is any set such that there is an injection from X into $\mathcal{P}\mathbb{N}$. Let $\mathcal{C}$ be the set of countable subsets of X. Then there is a surjection from $\mathcal{P}\mathbb{N}$ onto $\mathcal{C}$. **P** Let $f : X \to \mathcal{P}\mathbb{N}$ be an injection. Set $f_1(x) = \{0\} \cup \{i+1 : i \in f(x)\}$; then $f_1 : X \to \mathcal{P}\mathbb{N}$ is injective and $f_1(x) \ne \emptyset$ for every $x \in X$. Define $g : \mathcal{P}\mathbb{N} \to \mathcal{P}X$ by setting

$$g(A) = \{x : \exists\, n \in \mathbb{N},\ f_1(x) = \{i : 2^n(2i+1) \in A\}\}$$

for each $A \subseteq \mathbb{N}$. Then $g(A)$ is countable, since we have an injection

$$x \mapsto \min\{n : f_1(x) = \{i : 2^n(2i+1) \in A\}\}$$

from $g(A)$ to $\mathbb{N}$. Thus g is a function from $\mathcal{P}\mathbb{N}$ to $\mathcal{C}$. To see that g is surjective, observe that $\emptyset = g(\emptyset)$, while if $C \subseteq X$ is countable and not empty there is a surjection $h : \mathbb{N} \to C$; now set

$$A = \{2^n(2i+1) : n \in \mathbb{N},\ i \in f_1(h(n))\},$$

and see that $g(A) = C$. **Q**

(c) Again suppose that X is a set such that there is an injection from X to $\mathcal{P}\mathbb{N}$, and write H for the set of functions h such that $\operatorname{dom} h$ is a countable subset of X and h takes values in $\{0,1\}$. Then there is a surjection from $\mathcal{P}\mathbb{N}$ onto H. **P** Let $\mathcal{C}$ be the set of countable subsets of X and let $g : \mathcal{P}\mathbb{N} \to \mathcal{C}$ be a surjection, as in (a). For $A \subseteq \mathbb{N}$ set

$$g_0(A) = g(\{i : 2i \in A\}), \quad g_1(A) = g(\{i : 2i+1 \in A\}),$$

so that $g_0(A)$, $g_1(A)$ are countable subsets of X, and $A \mapsto (g_0(A), g_1(A))$ is a surjection from $\mathcal{P}\mathbb{N}$ onto $\mathcal{C}\times\mathcal{C}$. Let h_A be the function with domain $g_0(A) \cup g_1(A)$ such that $h_A(x) = 1$ if $x \in g_1(A)$, 0 if $x \in g_0(A) \setminus g_1(A)$. Then $A \mapsto h_A$ is a surjection from $\mathcal{P}\mathbb{N}$ onto H. **Q**

2A1I Filters I pause for a moment to discuss a construction which is of great value in investigating topological spaces, but has other uses, and in its nature belongs to elementary set theory (much more elementary, indeed, than the work above).

(a) Let X be a non-empty set. A **filter** on X is a family $\mathcal{F}$ of subsets of X such that

$X \in \mathcal{F}$, $\emptyset \notin \mathcal{F}$,

$E \cap F \in \mathcal{F}$ whenever $E, F \in \mathcal{F}$,

$E \in \mathcal{F}$ whenever $X \supseteq E \supseteq F \in \mathcal{F}$.

The second condition implies (inducing on n) that $F_0 \cap \ldots \cap F_n \in \mathcal{F}$ whenever $F_0, \ldots, F_n \in \mathcal{F}$.

(b) Let X, Y be non-empty sets, $\mathcal{F}$ a filter on X and $f : D \to Y$ a function, where $D \in \mathcal{F}$. Then

$$\{E : E \subseteq Y, f^{-1}[E] \in \mathcal{F}\}$$

is a filter on Y (because $f^{-1}[Y] = D$, $f^{-1}[\emptyset] = \emptyset$, $f^{-1}[E \cap F] = f^{-1}[E] \cap f^{-1}[F]$, $X \supseteq f^{-1}[E] \supseteq f^{-1}[F]$ whenever $Y \supseteq E \supseteq F$); I will call it $f[[\mathcal{F}]]$, the **image filter** of $\mathcal{F}$ under f.

Remark Of course there is a hidden variable in this notation. Ordinarily in this book I regard a function f as being defined by its domain $\operatorname{dom} f$ and its values on its domain; that is, it is determined by its graph $\{(x, f(x)) : x \in \operatorname{dom} f\}$, and indeed I normally do not distinguish between a function and its graph. This means that when I write '$f : D \to Y$ is a function' then the class $D = \operatorname{dom} f$ can be recovered from the function, but the class Y cannot; all I promise is that Y includes the class $f[D]$ of values of f. Now in the notation $f[[\mathcal{F}]]$ above we do actually need to know which set Y it is to be a filter on, even though this cannot be discovered from knowledge of f and $\mathcal{F}$. So you will always have to infer it from the context.

2A1J The Axiom of Choice I come now to the second half of this section, in which I discuss concepts and theorems dependent on the Axiom of Choice. Let me remind you of the statement of this axiom:

(AC) 'whenever I is a set and $\langle X_i \rangle_{i \in I}$ is a family of non-empty sets indexed by I, there is a function f, with domain I, such that $f(i) \in X_i$ for every $i \in I$'.

The function f is a **choice function**; it picks out one member of each of the given family of non-empty sets X_i.

I believe that one's attitude to this principle is a matter for individual choice. It is an indispensable foundation for very large parts of twentieth-century pure mathematics, including a substantial fraction of the present volume; but there are also significant areas in which principles actually contradictory to it can be employed to striking effect, leading – in my view – to equally valid mathematics. (I will describe one of these in §567 of Volume 5.) At present it is the case that more current mathematical activity, by volume, depends on asserting the axiom of choice than on all its rivals put together; but it is a matter of judgement and taste where the most important, or exciting, ideas are to be found. For the present volume I follow standard practice in twentieth-century abstract analysis, using the axiom of choice whenever necessary.

2A1K Zermelo's Well-Ordering Theorem (a) The Axiom of Choice is equiveridical with each of the statements

'for every set X there is a well-ordering of X',

'for every set X there is a bijection between X and some ordinal',

'for every set X there is a unique initial ordinal κ such that there is a bijection between X and κ.'

(Enderton 77, p. 196 et seq.; Halmos 60, §17; Henle 86, 9.1-9.3; Krivine 71, p. 20; Lipschutz 64, 12.1; Roitman 90, 3.6.38.)

(b) When assuming the axiom of choice, as I do nearly everywhere in this treatise, I write $\#(X)$ for that initial ordinal κ such that there is a bijection between κ and X; I call this the **cardinal** of X.

2A1L Fundamental consequences of the Axiom of Choice (a) For any two sets X and Y, there is a bijection between X and Y iff $\#(X) = \#(Y)$. More generally, there is an injection from X to Y iff $\#(X) \le \#(Y)$, and a surjection from X onto Y iff either $\#(X) \ge \#(Y) > 0$ or $\#(X) = \#(Y) = 0$.

(b) In particular, $\#(\mathcal{P}\mathbb{N}) = \#(\mathbb{R})$; write $\mathfrak{c}$ for this common value, the **cardinal of the continuum**. Cantor's theorem that $\mathcal{P}\mathbb{N}$ and $\mathbb{R}$ are uncountable becomes the result $\omega < \mathfrak{c}$, that is, $\omega_1 \le \mathfrak{c}$.

(c) If X is any infinite set, and $r \geq 1$, then there is a bijection between X^r and X. (Enderton 77, p. 162; Halmos 60, §24.) (I note that we need some form of the axiom of choice to prove the result in this generality. But of course for most of the infinite sets arising naturally in mathematics – sets like $\mathbb{N}$ and $\mathcal{P}\mathbb{R}$ – it is easy to prove the result without appeal to the axiom of choice.)

(d) Suppose that κ is an infinite cardinal. If I is a set with cardinal at most κ and $\langle A_i \rangle_{i \in I}$ is a family of sets with $\#(A_i) \leq \kappa$ for every $i \in I$, then $\#(\bigcup_{i \in I} A_i) \leq \kappa$. Consequently $\#(\bigcup \mathcal{A}) \leq \kappa$ whenever $\mathcal{A}$ is a family of sets such that $\#(\mathcal{A}) \leq \kappa$ and $\#(A) \leq \kappa$ for every $A \in \mathcal{A}$. In particular, ω_1 cannot be expressed as a countable union of countable sets, and ω_2 cannot be expressed as a countable union of sets with cardinal at most ω_1.

(e) Now we can rephrase 2A1Hc as: if $\#(X) \leq \mathfrak{c}$, then $\#(H) \leq \mathfrak{c}$, where H is the set of functions from a countable subset of X to $\{0, 1\}$. **P** For we have an injection from X into $\mathcal{P}\mathbb{N}$, and therefore a surjection from $\mathcal{P}\mathbb{N}$ onto H. **Q**

(f) Any non-empty class of cardinals has a least member (by 2A1Dc).

2A1M Zorn's Lemma In 2A1K I described the well-ordering principle. I come now to another proposition which is equiveridical with the axiom of choice:

> 'Let $(P, \leq)$ be a non-empty partially ordered set such that every non-empty totally ordered subset of P has an upper bound in P. Then P has a maximal element.'

This is **Zorn's Lemma**. For the proof that the axiom of choice implies, and is implied by, Zorn's Lemma, see Enderton 77, p. 151; Halmos 60, §16; Henle 86, 9.1-9.3; Roitman 90, 3.6.38.

2A1N Ultrafilters A filter $\mathcal{F}$ on a set X is an **ultrafilter** if for every $A \subseteq X$ either $A \in \mathcal{F}$ or $X \setminus A \in \mathcal{F}$.

If $\mathcal{F}$ is an ultrafilter on X and $f : D \to Y$ is a function, where $D \in \mathcal{F}$, then $f[[\mathcal{F}]]$ is an ultrafilter on Y (because $f^{-1}[Y \setminus A] = D \setminus f^{-1}[A]$ for every $A \subseteq Y$).

One type of ultrafilter can be described easily: if x is any point of a set X, then $\mathcal{F} = \{F : x \in F \subseteq X\}$ is an ultrafilter on X. (You need only read the definitions. Ultrafilters of this type are called **principal ultrafilters**.) But it is not obvious that there are any further ultrafilters, and indeed it is not possible to prove that there are any, without using a strong form of the axiom of choice, as follows.

2A1O The Ultrafilter Theorem As an example of the use of Zorn's lemma which will be of great value in studying compact topological spaces (2A3N *et seq.*, and §247), I give the following result.

Theorem Let X be any non-empty set, and $\mathcal{F}$ a filter on X. Then there is an ultrafilter $\mathcal{H}$ on X such that $\mathcal{F} \subseteq \mathcal{H}$.

proof (Cf. Henle 86, 9.4; Roitman 90, 3.6.37.) Let $\mathfrak{P}$ be the set of all filters on X including $\mathcal{F}$, and order $\mathfrak{P}$ by inclusion, so that, for $\mathcal{G}_1$, $\mathcal{G}_2 \in \mathfrak{P}$, $\mathcal{G}_1 \leq \mathcal{G}_2$ in $\mathfrak{P}$ iff $\mathcal{G}_1 \subseteq \mathcal{G}_2$. It is easy to see that $\mathfrak{P}$ is a partially ordered set, and it is non-empty because $\mathcal{F} \in \mathfrak{P}$. If $\mathfrak{Q}$ is any non-empty totally ordered subset of $\mathfrak{P}$, then $\mathcal{H}_{\mathfrak{Q}} = \bigcup \mathfrak{Q} \in \mathfrak{P}$. **P** Of course $\mathcal{H}_{\mathfrak{Q}}$ is a family of subsets of X. (i) Take any $\mathcal{G}_0 \in \mathfrak{Q}$; then $X \in \mathcal{G}_0 \subseteq \mathcal{H}_{\mathfrak{Q}}$. If $\mathcal{G} \in \mathfrak{Q}$, then $\mathcal{G}$ is a filter, so $\emptyset \notin \mathcal{G}$; accordingly $\emptyset \notin \mathcal{H}_{\mathfrak{Q}}$. (ii) If E, $F \in \mathcal{H}_{\mathfrak{Q}}$, then there are $\mathcal{G}_1$, $\mathcal{G}_2 \in \mathfrak{Q}$ such that $E \in \mathcal{G}_1$ and $F \in \mathcal{G}_2$. Because $\mathfrak{Q}$ is totally ordered, either $\mathcal{G}_1 \subseteq \mathcal{G}_2$ or $\mathcal{G}_2 \subseteq \mathcal{G}_1$. In either case, $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2 \in \mathfrak{Q}$. Now $\mathcal{G}$ is a filter containing both E and F, so it contains $E \cap F$, and $E \cap F \in \mathcal{H}_{\mathfrak{Q}}$. (iii) If $X \supseteq E \supseteq F \in \mathcal{H}_{\mathfrak{Q}}$, there is a $\mathcal{G} \in \mathfrak{Q}$ such that $F \in \mathcal{G}$; and $E \in \mathcal{G} \subseteq \mathcal{H}_{\mathfrak{Q}}$. This shows that $\mathcal{H}_{\mathfrak{Q}}$ is a filter on X. (iv) Finally, $\mathcal{H}_{\mathfrak{Q}} \supseteq \mathcal{G}_0 \supseteq \mathcal{F}$, so $\mathcal{H}_{\mathfrak{Q}} \in \mathfrak{P}$. **Q** Now $\mathcal{H}_{\mathfrak{Q}}$ is evidently an upper bound for $\mathfrak{Q}$ in $\mathfrak{P}$.

We may therefore apply Zorn's Lemma to find a maximal element $\mathcal{H}$ of $\mathfrak{P}$. This $\mathcal{H}$ is surely a filter on X including $\mathcal{F}$.

Now let $A \subseteq X$ be such that $A \notin \mathcal{H}$. Consider

$$\mathcal{H}_1 = \{E : E \subseteq X,\, E \cup A \in \mathcal{H}\}.$$

This is a filter on X. **P** Of course it is a family of subsets of X. (i) $X \cup A = X \in \mathcal{H}$, so $X \in \mathcal{H}_1$. $\emptyset \cup A = A \notin \mathcal{H}$ so $\emptyset \notin \mathcal{H}_1$. (ii) If E, $F \in \mathcal{H}_1$ then

$$(E \cap F) \cup A = (E \cup A) \cap (F \cup A) \in \mathcal{H},$$

so $E \cap F \in \mathcal{H}_1$. (iii) If $X \supseteq E \supseteq F \in \mathcal{H}_1$ then $E \cup A \supseteq F \cup A \in \mathcal{H}$, so $E \cup A \in \mathcal{H}$ and $E \in \mathcal{H}_1$. **Q** Also $\mathcal{H}_1 \supseteq \mathcal{H}$, so $\mathcal{H}_1 \in \mathfrak{P}$. But $\mathcal{H}$ is a maximal element of $\mathfrak{P}$, so $\mathcal{H}_1 = \mathcal{H}$. Since $(X \setminus A) \cup A = X \in \mathcal{H}$, $X \setminus A \in \mathcal{H}_1$ and $X \setminus A \in \mathcal{H}$.

As A is arbitrary, $\mathcal{H}$ is an ultrafilter, as required.

2A1P I come now to a result from infinitary combinatorics for which I give a detailed proof, not because it cannot be found in many textbooks, but because it is usually given in enormously greater generality, to the point indeed that it may be harder to understand why the stated theorem covers the present result than to prove the latter from first principles.

Theorem (a) Let $\langle K_\alpha \rangle_{\alpha \in A}$ be a family of countable sets, with $\#(A)$ strictly greater than $\mathfrak{c}$, the cardinal of the continuum. Then there are a set M, with cardinal at most $\mathfrak{c}$, and a set $B \subseteq A$, with cardinal strictly greater than $\mathfrak{c}$, such that $K_\alpha \cap K_\beta \subseteq M$ whenever α, β are distinct members of B.

(b) Let I be a set, and $\langle f_\alpha \rangle_{\alpha \in A}$ a family in $\{0,1\}^I$, the set of functions from I to $\{0,1\}$, with $\#(A) > \mathfrak{c}$. If $\langle K_\alpha \rangle_{\alpha \in A}$ is any family of countable subsets of I, then there is a set $B \subseteq A$, with cardinal greater than $\mathfrak{c}$, such that f_α and f_β agree on $K_\alpha \cap K_\beta$ for all α, $\beta \in B$.

(c) In particular, under the conditions of (b), there are distinct α, $\beta \in A$ such that f_α and f_β agree on $K_\alpha \cap K_\beta$.

proof (a) Choose inductively a family $\langle M_\xi \rangle_{\xi < \omega_1}$ of sets by the rule

if there is any set N such that

$(*)$ N is disjoint from $\bigcup_{\eta < \xi} M_\eta$, $\#(N) \le \mathfrak{c}$ and $\#(\{\alpha : \alpha \in A, K_\alpha \cap N = \emptyset\}) \le \mathfrak{c}$,

choose such a set for M_ξ;

otherwise set $M_\xi = \emptyset$.

When M_ξ has been chosen for every $\xi < \omega_1$, set $M = \bigcup_{\xi < \omega_1} M_\xi$. The rule ensures that $\langle M_\xi \rangle_{\xi < \omega_1}$ is disjoint and that $\#(M_\xi) \le \mathfrak{c}$ for every $\xi < \omega_1$, while $\omega_1 \le \mathfrak{c}$, so $\#(M) \le \mathfrak{c}$.

Let $\mathfrak{P}$ be the family of sets $P \subseteq A$ such that $K_\alpha \cap K_\beta \subseteq M$ for all distinct α, $\beta \in P$. Order $\mathfrak{P}$ by inclusion, so that it is a partially ordered set. If $\mathfrak{Q} \subseteq \mathfrak{P}$ is totally ordered, then $\bigcup \mathfrak{Q} \in \mathfrak{P}$. **P** If α, β are distinct members of $\bigcup \mathfrak{Q}$, there are Q_1, $Q_2 \in \mathfrak{Q}$ such that $\alpha \in Q_1$, $\beta \in Q_2$; now $P = Q_1 \cup Q_2$ is equal to one of Q_1, Q_2, and in either case belongs to $\mathfrak{P}$ and contains both α and β, so $K_\alpha \cap K_\beta \subseteq M$. **Q** By Zorn's Lemma, $\mathfrak{P}$ has a maximal element B, and we surely have $K_\alpha \cap K_\beta \subseteq M$ for all distinct α, $\beta \in B$.

? Suppose, if possible, that $\#(B) \le \mathfrak{c}$. Set $N = \bigcup_{\alpha \in B} K_\alpha \setminus M$. Then N has cardinal at most $\mathfrak{c}$, being included in a union of at most $\mathfrak{c}$ countable sets. For every $\gamma \in A \setminus B$, $B \cup \{\gamma\} \notin \mathfrak{P}$, so there must be some $\alpha \in B$ such that $K_\alpha \cap K_\gamma \not\subseteq M$; that is, $K_\gamma \cap N \ne \emptyset$. Thus $\{\gamma : K_\gamma \cap N = \emptyset\} \subseteq B$ has cardinal at most $\mathfrak{c}$. But this means that in the rule for choosing M_ξ, there was always an N satisfying the condition $(*)$, and therefore M_ξ also did. Thus $C_\xi = \{\alpha : K_\alpha \cap M_\xi = \emptyset\}$ has cardinal at most $\mathfrak{c}$ for every $\xi < \omega_1$. So $C = \bigcup_{\xi < \omega_1} C_\xi$ also has. But the original hypothesis was that $\#(A) > \mathfrak{c}$, so there is an $\alpha \in A \setminus C$. In this case, $K_\alpha \cap M_\xi \ne \emptyset$ for every $\xi < \omega_1$. But this means that we have a surjection $\phi : K_\alpha \cap M \to \omega_1$ given by setting

$$\phi(i) = \xi \text{ if } i \in K_\alpha \cap M_\xi.$$

Since $\#(K_\alpha) \le \omega < \omega_1$, this is impossible. **X**

Accordingly $\#(B) > \mathfrak{c}$ and we have found a suitable pair M, B.

(b) By (a), we can find a set M, with cardinal at most $\mathfrak{c}$, and a set $B_0 \subseteq A$, with cardinal greater than $\mathfrak{c}$, such that $K_\alpha \cap K_\beta \subseteq M$ for all distinct α, $\beta \in B_0$. Let H be the set of functions from countable subsets of M to $\{0,1\}$; then $f'_\alpha = f_\alpha \restriction (K_\alpha \cap M) \in H$ for each $\alpha \in B_0$. Now $B_0 = \bigcup_{h \in H} \{\alpha : \alpha \in B_0, f'_\alpha = h\}$ has cardinal greater than $\mathfrak{c}$, while $\#(H) \le \mathfrak{c}$ (2A1Le), so there must be some $h \in H$ such that $B = \{\alpha : \alpha \in B_0, f'_\alpha = h\}$ has cardinal greater than $\mathfrak{c}$.

If α, β are distinct members of B, then $K_\alpha \cap K_\beta \subseteq M$, because α, $\beta \in B_0$; but this means that

$$f_\alpha \restriction K_\alpha \cap K_\beta = h \restriction K_\alpha \cap K_\beta = f_\beta \restriction K_\alpha \cap K_\beta.$$

Thus B has the required property. (Of course f_α and f_β agree on $K_\alpha \cap K_\beta$ if $\alpha = \beta$.)

(c) follows at once.

Remark The result we need in this volume (in 216E) is part (c) above. There are other proofs of this, perhaps a little simpler; but the stronger result in part (b) will be useful in Volume 3.

Version of 30.11.09

2A2 The topology of Euclidean space

In the appendix to Volume 1 (§1A2) I discussed open and closed sets in $\mathbb{R}^r$; the chief aim there was to support the idea of 'Borel set', which is vital in the theory of Lebesgue measure, but of course they are also fundamental to the study of continuous functions, and indeed to all aspects of real analysis. I give here a very brief introduction to the further elementary facts about closed and compact sets and continuous functions which we need for this volume. Much of this material can be derived from the generalizations in §2A3, but nevertheless I sketch the proofs, since for the greater part of the volume (most of the exceptions are in Chapter 24) Euclidean space is sufficient for our needs.

2A2A Closures: Definition For any $r \geq 1$ and any $A \subseteq \mathbb{R}^r$, the **closure** of A, $\overline{A}$, is the intersection of all the closed subsets of $\mathbb{R}^r$ including A. This is itself closed (being the intersection of a non-empty family of closed sets, see 1A2Fd), so is the smallest closed set including A. In particular, A is closed iff $\overline{A} = A$.

2A2B Lemma Let $A \subseteq \mathbb{R}^r$ be any set. Then for $x \in \mathbb{R}^r$ the following are equiveridical:
(i) $x \in \overline{A}$, the closure of A;
(ii) $B(x,\delta) \cap A \neq \emptyset$ for every $\delta > 0$, where $B(x,\delta) = \{y : \|y-x\| \leq \delta\}$;
(iii) there is a sequence $\langle x_n\rangle_{n\in\mathbb{N}}$ in A such that $\lim_{n\to\infty} \|x_n - x\| = 0$.

proof (a)(i)⇒(ii) Suppose that $x \in \overline{A}$ and $\delta > 0$. Then $U(x,\delta) = \{y : \|y-x\| < \delta\}$ is an open set (1A2D), so $F = \mathbb{R}^r \setminus U(x,\delta)$ is closed, while $x \notin F$. Now

$$x \in \overline{A} \setminus F \Longrightarrow \overline{A} \not\subseteq F \Longrightarrow A \not\subseteq F \Longrightarrow A \cap U(x,\delta) \neq \emptyset \Longrightarrow A \cap B(x,\delta) \neq \emptyset.$$

As δ is arbitrary, (ii) is true.

(b)(ii)⇒(iii) If (ii) is true, then for each $n \in \mathbb{N}$ we can find an $x_n \in A$ such that $\|x_n - x\| \leq 2^{-n}$, and now $\lim_{n\to\infty} \|x_n - x\| = 0$.

(c)(iii)⇒(i) Assume (iii). **?** Suppose, if possible, that $x \notin \overline{A}$. Then x belongs to the open set $\mathbb{R}^r \setminus \overline{A}$ and there is a $\delta > 0$ such that $U(x,\delta) \subseteq \mathbb{R}^r \setminus \overline{A}$. But now there is an n such that $\|x_n - x\| < \delta$, in which case $x_n \in U(x,\delta) \cap A \subseteq U(x,\delta) \cap \overline{A}$. **X**

2A2C Continuous functions (a) I begin with a characterization of continuous functions in terms of open sets. If r, $s \geq 1$, $D \subseteq \mathbb{R}^r$ and $\phi : D \to \mathbb{R}^s$ is a function, we say that ϕ is **continuous** if for every $x \in D$ and $\epsilon > 0$ there is a $\delta > 0$ such that $\|\phi(y) - \phi(x)\| \leq \epsilon$ whenever $y \in D$ and $\|y - x\| \leq \delta$. Now ϕ is continuous iff for every open set $G \subseteq \mathbb{R}^s$ there is an open set $H \subseteq \mathbb{R}^r$ such that $\phi^{-1}[G] = D \cap H$.

P **(i)** Suppose that ϕ is continuous and that $G \subseteq \mathbb{R}^s$ is open. Set

$$H = \bigcup\{U : U \subseteq \mathbb{R}^r \text{ is open, } \phi[U \cap D] \subseteq G\}.$$

Then H is a union of open sets, therefore open (1A2Bd), and $H \cap D \subseteq \phi^{-1}[G]$. If $x \in \phi^{-1}[G]$, then $\phi(x) \in G$, so there is an $\epsilon > 0$ such that $U(\phi(x),\epsilon) \subseteq G$; now there is a $\delta > 0$ such that $\|\phi(y) - \phi(x)\| \leq \frac{1}{2}\epsilon$ whenever $y \in D$ and $\|y - x\| \leq \delta$, so that

$$\phi[U(x,\delta) \cap D] \subseteq U(\phi(x),\epsilon) \subseteq G$$

and

$$x \in U(x,\delta) \subseteq H.$$

As x is arbitrary, $\phi^{-1}[G] = H \cap D$. As G is arbitrary, ϕ satisfies the condition.

(ii) Now suppose that ϕ satisfies the condition. Take $x \in D$ and $\epsilon > 0$. Then $U(\phi(x),\epsilon)$ is open, so there is an open $H \subseteq \mathbb{R}^r$ such that $H \cap D = \phi^{-1}[U(\phi(x),\epsilon)]$; we see that $x \in H$, so there is a $\delta > 0$ such that $U(x,\delta) \subseteq H$; now if $y \in D$ and $\|y - x\| \leq \frac{1}{2}\delta$ then $y \in D \cap H$, $\phi(y) \in U(\phi(x),\epsilon)$ and $\|\phi(y) - \phi(x)\| \leq \epsilon$. As x and ϵ are arbitrary, ϕ is continuous. **Q**

(b) Using the ϵ-δ definition of continuity, it is easy to see that a function ϕ from a subset D of $\mathbb{R}^r$ to $\mathbb{R}^s$ is continuous iff all its components ϕ_i are continuous, writing $\phi(x) = (\phi_1(x), \dots, \phi_s(x))$ for $x \in D$. **P** **(i)** If ϕ is continuous, $i \le s$, $x \in D$ and $\epsilon > 0$, then there is a $\delta > 0$ such that

$$|\phi_i(y) - \phi_i(x)| \le \|\phi(y) - \phi(x)\| \le \epsilon$$

whenever $y \in D$ and $\|y - x\| \le \delta$. **(ii)** If every ϕ_i is continuous, $x \in D$ and $\epsilon > 0$, then there are $\delta_i > 0$ such that $|\phi_i(y) - \phi_i(x)| \le \epsilon/\sqrt{s}$ whenever $y \in D$ and $\|y - x\| \le \delta_i$; setting $\delta = \min_{1 \le i \le r} \delta_i > 0$, we have $\|\phi(y) - \phi(x)\| \le \epsilon$ whenever $y \in D$ and $\|y - x\| \le \delta$. **Q**

(c) At one or two points, we shall encounter the following strengthening of the notion of 'continuous function'. If r, $s \ge 1$, $D \subseteq \mathbb{R}^r$ and $\phi : D \to \mathbb{R}^s$ is a function, we say that ϕ is **uniformly continuous** if for every $\epsilon > 0$ there is a $\delta > 0$ such that $\|\phi(y) - \phi(x)\| \le \epsilon$ whenever x, $y \in D$ and $\|y - x\| \le \delta$. A uniformly continuous function is of course continuous.

2A2D Compactness in $\mathbb{R}^r$: Definition A subset F of $\mathbb{R}^r$ is called **compact** if whenever $\mathcal{G}$ is a family of open sets covering F then there is a finite subset $\mathcal{G}_0$ of $\mathcal{G}$ still covering F.

2A2E Elementary properties of compact sets Take any $r \ge 1$, and subsets D, F, G and K of $\mathbb{R}^r$.

(a) If K is compact and F is closed, then $K \cap F$ is compact. **P** Let $\mathcal{G}$ be an open cover of $F \cap K$. Then $\mathcal{G} \cup \{\mathbb{R}^r \setminus F\}$ is an open cover of K, so has a finite subcover $\mathcal{G}_0$ say. Now $\mathcal{G}_0 \setminus \{\mathbb{R}^r \setminus F\}$ is a finite subset of $\mathcal{G}$ covering $K \cap F$. As $\mathcal{G}$ is arbitrary, $K \cap F$ is compact. **Q**

(b) If $s \ge 1$, $\phi : D \to \mathbb{R}^s$ is a continuous function, K is compact and $K \subseteq D$, then $\phi[K]$ is compact. **P** Let $\mathcal{V}$ be an open cover of $\phi[K]$. Let $\mathcal{H}$ be

$$\{H : H \subseteq \mathbb{R}^r \text{ is open, } \exists\ V \in \mathcal{V},\ \phi^{-1}[V] = D \cap H\}.$$

If $x \in K$, then $\phi(x) \in \phi[K]$ so there is a $V \in \mathcal{V}$ such that $\phi(x) \in V$; now there is an $H \in \mathcal{H}$ such that $D \cap H\phi^{-1}[V]$ contains x (2A2Ca); as x is arbitrary, $K \subseteq \bigcup \mathcal{H}$. Let $\mathcal{H}_0$ be a finite subset of H covering K. For each $H \in \mathcal{H}_0$, let $V_H \in \mathcal{V}$ be such that $\phi^{-1}[V_H] = D \cap H$; then $\{V_H : H \in \mathcal{H}_0\}$ is a finite subset of $\mathcal{V}$ covering $\phi[K]$. As $\mathcal{V}$ is arbitrary, $\phi[K]$ is compact. **Q**

(c) If K is compact, it is closed. **P** Write $H = \mathbb{R}^r \setminus K$. Take any $x \in H$. Then $G_n = \mathbb{R}^r \setminus B(x, 2^{-n})$ is open for every $n \in \mathbb{N}$ (1A2G). Also

$$\bigcup_{n \in \mathbb{N}} G_n = \{y : y \in \mathbb{R}^r,\ \|y - x\| > 0\} = \mathbb{R}^r \setminus \{x\} \supseteq K.$$

So there is some finite set $\mathcal{G}_0 \subseteq \{G_n : n \in \mathbb{N}\}$ which covers K. There must be an n such that $\mathcal{G}_0 \subseteq \{G_i : i \le n\}$, so that

$$K \subseteq \bigcup \mathcal{G}_0 \subseteq \bigcup_{i \le n} G_i = G_n,$$

and $B(x, 2^{-n}) \subseteq H$. As x is arbitrary, H is open and K is closed. **Q**

(d) If K is compact and G is open and $K \subseteq G$, then there is a $\delta > 0$ such that $K + B(\mathbf{0}, \delta) \subseteq G$. **P** If $K = \emptyset$, this is trivial, as then

$$K + B(\mathbf{0}, 1) = \{x + y : x \in K,\ y \in B(\mathbf{0}, 1)\} = \emptyset.$$

Otherwise, set

$$\mathcal{G} = \{U(x, \delta) : x \in \mathbb{R}^r,\ \delta > 0,\ U(x, 2\delta) \subseteq G\}.$$

Then $\mathcal{G}$ is a family of open sets and $\bigcup \mathcal{G} = G$ (because G is open), so $\mathcal{G}$ is an open cover of K and has a finite subcover $\mathcal{G}_0$. Express $\mathcal{G}_0$ as $\{U(x_0, \delta_0), \dots, U(x_n, \delta_n)\}$ where $U(x_i, 2\delta_i) \subseteq G$ for each i. Set $\delta = \min_{i \le n} \delta_i > 0$. If $x \in K$ and $y \in B(\mathbf{0}, \delta)$, then there is an $i \le n$ such that $x \in U(x_i, \delta_i)$; now

$$\|(x + y) - x_i\| \le \|x - x_i\| + \|y\| < \delta_i + \delta \le 2\delta_i,$$

so $x + y \in U(x_i, 2\delta_i) \subseteq G$. As x and y are arbitrary, $K + B(\mathbf{0}, \delta) \subseteq G$. **Q**

2A2F The value of the concept of 'compactness' is greatly increased by the fact that there is an effective characterization of the compact subsets of $\mathbb{R}^r$.

Theorem For any $r \ge 1$, a subset K of $\mathbb{R}^r$ is compact iff it is closed and bounded.

proof (a) Suppose that K is compact. By 2A2Ec, it is closed. To see that it is bounded, consider $\mathcal{G} = \{U(\mathbf{0}, n) : n \in \mathbb{N}\}$. $\mathcal{G}$ consists entirely of open sets, and $\bigcup \mathcal{G} = \mathbb{R}^r \supseteq K$, so there is a finite $\mathcal{G}_0 \subseteq \mathcal{G}$ covering K. There must be an n such that $\mathcal{G}_0 \subseteq \{G_i : i \le n\}$, so that

$$K \subseteq \bigcup \mathcal{G}_0 \subseteq \bigcup_{i \le n} U(\mathbf{0}, i) = U(\mathbf{0}, n),$$

and K is bounded.

(b) Thus we are left with the converse; I have to show that a closed bounded set is compact. The main part of the argument is a proof by induction on r that the closed interval $[-\mathbf{n}, \mathbf{n}]$ is compact for all $n \in \mathbb{N}$, writing $\mathbf{n} = (n, \dots, n) \in \mathbb{R}^r$.

(i) If $r = 1$ and $n \in \mathbb{N}$ and $\mathcal{G}$ is a family of open sets in $\mathbb{R}$ covering $[-n, n]$, set

$$A = \{x : x \in [-n, n], \text{ there is a finite } \mathcal{G}_0 \subseteq \mathcal{G} \text{ such that } [-n, x] \subseteq \bigcup \mathcal{G}_0\}.$$

Then $-n \in A$, because if $-n \in G \in \mathcal{G}$ then $[-n, -n] \subseteq \bigcup\{G\}$, and A is bounded above by n, so $c = \sup A$ exists and belongs to $[-n, n]$.

Next, $c \in [-n, n] \subseteq \bigcup \mathcal{G}$, so there is a $G \in \mathcal{G}$ containing c. Let $\delta > 0$ be such that $U(c, \delta) \subseteq G$. There is an $x \in A$ such that $x \ge c - \delta$. Let $\mathcal{G}_0$ be a finite subset of $\mathcal{G}$ covering $[-n, x]$. Then $\mathcal{G}_1 = \mathcal{G}_0 \cup \{G\}$ is a finite subset of $\mathcal{G}$ covering $[-n, c + \frac{1}{2}\delta]$. But $c + \frac{1}{2}\delta \notin A$ so $c + \frac{1}{2}\delta > n$ and $\mathcal{G}_1$ is a finite subset of $\mathcal{G}$ covering $[-n, n]$. As $\mathcal{G}$ is arbitrary, $[-n, n]$ is compact and the induction starts.

(ii) For the inductive step to $r + 1$, regard the closed interval $F = [-\mathbf{n}, \mathbf{n}]$, taken in $\mathbb{R}^{r+1}$, as the product of the closed interval $E = [-\mathbf{n}, \mathbf{n}]$, taken in $\mathbb{R}^r$, with the closed interval $[-n, n] \subseteq \mathbb{R}$; by the inductive hypothesis, both E and $[-n, n]$ are compact. Let $\mathcal{G}$ be a family of open subsets of $\mathbb{R}^{r+1}$ covering F. Write $\mathcal{H}$ for the family of open subsets H of $\mathbb{R}^r$ such that $H \times [-n, n]$ is covered by a finite subfamily of $\mathcal{G}$. Then $E \subseteq \bigcup \mathcal{H}$. **P** Take $x \in E$. Set

$$\mathcal{U}_x = \{U : U \subseteq \mathbb{R} \text{ is open}, \exists\ G \in \mathcal{G}, \text{ open } H \subseteq \mathbb{R}^r, x \in H \text{ and } H \times U \subseteq G\}.$$

Then $\mathcal{U}_x$ is a family of open subsets of $\mathbb{R}$. If $\xi \in [-n, n]$, there is a $G \in \mathcal{G}$ containing (x, ξ); there is a $\delta > 0$ such that $U((x, \xi), \delta) \subseteq G$; now $U(x, \frac{1}{2}\delta)$ and $U(\xi, \frac{1}{2}\delta)$ are open sets in $\mathbb{R}^r$, $\mathbb{R}$ respectively and

$$U(x, \tfrac{1}{2}\delta) \times U(\xi, \tfrac{1}{2}\delta) \subseteq U((x, \xi), \delta) \subseteq G,$$

so $U(\xi, \frac{1}{2}\delta) \in \mathcal{U}_x$. As ξ is arbitrary, $\mathcal{U}_x$ is an open cover of $[-n, n]$ in $\mathbb{R}$. By (i), it has a finite subcover $U_0, \dots, U_k$ say. For each $j \le k$ we can find H_j, G_j such that H_j is an open subset of $\mathbb{R}^r$ containing x and $H_j \times U_j \subseteq G_j \in \mathcal{G}$. Now set $H = \bigcap_{j \le k} H_j$. This is an open subset of $\mathbb{R}^r$ containing x, and $H \times [-n, n] \subseteq \bigcup_{j \le n} G_j$ is covered by a finite subfamily of $\mathcal{G}$. So $x \in H \in \mathcal{H}$. As x is arbitrary, $\mathcal{H}$ covers E. **Q**

(iii) Now the inductive hypothesis tells us that E is compact, so there is a finite subfamily $\mathcal{H}_0$ of $\mathcal{H}$ covering E. For each $H \in \mathcal{H}_0$ let $\mathcal{G}_H$ be a finite subfamily of $\mathcal{G}$ covering $H \times [-n, n]$. Then $\bigcup_{H \in \mathcal{H}_0} \mathcal{G}_H$ is a finite subfamily of $\mathcal{G}$ covering $E \times [-n, n] = F$. As $\mathcal{G}$ is arbitrary, F is compact and the induction proceeds.

(iv) Thus the interval $[-\mathbf{n}, \mathbf{n}]$ is compact in $\mathbb{R}^r$ for every r, n. Now suppose that K is a closed bounded set in $\mathbb{R}^r$. Then there is an $n \in \mathbb{N}$ such that $K \subseteq [-\mathbf{n}, \mathbf{n}]$, that is, $K = K \cap [-\mathbf{n}, \mathbf{n}]$. As K is closed and $[-\mathbf{n}, \mathbf{n}]$ is compact, K is compact, by 2A2Ea.

This completes the proof.

2A2G Corollary If $\phi : D \to \mathbb{R}$ is continuous, where $D \subseteq \mathbb{R}^r$, and $K \subseteq D$ is a non-empty compact set, then ϕ is bounded and attains its bounds on K.

proof By 2A2Eb, $\phi[K]$ is compact; by 2A2F it is closed and bounded. To say that $\phi[K]$ is bounded is just to say that ϕ is bounded on K. Because $\phi[K]$ is a non-empty bounded set, it has an infimum a and a supremum b; now both belong to $\overline{\phi[K]}$ (by the criterion 2A2B(ii), or otherwise); because $\phi[K]$ is closed, both belong to $\phi[K]$, that is, ϕ attains its bounds.

2A2H Lim sup and lim inf revisited In §1A3 I briefly discussed $\limsup_{n\to\infty} a_n$, $\liminf_{n\to\infty} a_n$ for real sequences $\langle a_n\rangle_{n\in\mathbb{N}}$. In this volume we need the notion of $\limsup_{\delta\downarrow 0} f(\delta)$, $\liminf_{\delta\downarrow 0} f(\delta)$ for real functions f. I say that $\limsup_{\delta\downarrow 0} f(\delta) = u \in [-\infty,\infty]$ if (i) for every $v > u$ there is an $\eta > 0$ such that $f(\delta)$ is defined and less than or equal to v for every $\delta \in]0,\eta]$ (ii) for every $v < u$ and $\eta > 0$ there is a $\delta \in]0,\eta]$ such that $f(\delta)$ is defined and greater than or equal to v. Similarly, $\liminf_{\delta\downarrow 0} f(\delta) = u \in [-\infty,\infty]$ if (i) for every $v < u$ there is an $\eta > 0$ such that $f(\delta)$ is defined and greater than or equal to v for every $\delta \in]0,\eta]$ (ii) for every $v > u$ and $\eta > 0$ there is an $\delta \in]0,\eta]$ such that $f(\delta)$ is defined and less than or equal to v.

2A2I In the one-dimensional case, we have a particularly simple description of the open sets.

Proposition If $G \subseteq \mathbb{R}$ is any open set, it is expressible as the union of a countable disjoint family of open intervals.

proof For x, $y \in G$ write $x \sim y$ if either $x \le y$ and $[x,y] \subseteq G$ or $y \le x$ and $[y,x] \subseteq G$. It is easy to check that $\sim$ is an equivalence relation on G. Let $\mathcal{C}$ be the set of equivalence classes under $\sim$. Then $\mathcal{C}$ is a partition of G. Now every $C \in \mathcal{C}$ is an open interval. **P** Set $a = \inf C$, $b = \sup C$ (allowing $a = -\infty$ and/or $b = \infty$ if C is unbounded). If $a < x < b$, there are y, $z \in C$ such that $y \le x \le z$, so that $[y,x] \subseteq [y,z] \subseteq G$ and $y \sim x$ and $x \in C$; thus $]a,b[\subseteq C$. If $x \in C$, there is an open interval I containing x and included in G; since $x \sim y$ for every $y \in I$, $I \subseteq C$; so

$$a \le \inf I < x < \sup I \le b$$

and $x \in]a,b[$. Thus $C =]a,b[$ is an open interval. **Q**

To see that $\mathcal{C}$ is countable, observe that every member of $\mathcal{C}$ contains a member of $\mathbb{Q}$, so that we have a surjective function from a subset of $\mathbb{Q}$ onto $\mathcal{C}$, and $\mathcal{C}$ is countable (1A1E).

Version of 25.7.07

2A3 General topology

At various points – principally §§245-247, but also for certain ideas in Chapter 27 – we need to know something about non-metrizable topologies. I must say that you should probably take the time to look at some book on elementary functional analysis which has the phrases 'weak compactness' or 'weakly compact' in the index. But I can list here the concepts actually used in this volume, in a good deal less space than any orthodox, complete treatment would employ.

2A3A Topologies First we need to know what a 'topology' is. If X is any set, a **topology** on X is a family $\mathfrak{T}$ of subsets of X such that (i) $\emptyset$, $X \in \mathfrak{T}$ (ii) if G, $H \in \mathfrak{T}$ then $G \cap H \in \mathfrak{T}$ (iii) if $\mathcal{G} \subseteq \mathfrak{T}$ then $\bigcup \mathcal{G} \in \mathfrak{T}$ (cf. 1A2B). The pair $(X,\mathfrak{T})$ is now a **topological space**. In this context, members of $\mathfrak{T}$ are called **open** and their complements (in X) are called **closed** (cf. 1A2E-1A2F).

2A3B Continuous functions (a) If $(X,\mathfrak{T})$ and $(Y,\mathfrak{S})$ are topological spaces, a function $\phi : X \to Y$ is **continuous** if $\phi^{-1}[G] \in \mathfrak{T}$ for every $G \in \mathfrak{S}$. (By 2A2Ca above, this is consistent with the ϵ-δ definition of continuity for functions from one Euclidean space to another. See also 2A3H below.)

(b) If $(X,\mathfrak{T})$, $(Y,\mathfrak{S})$ and $(Z,\mathfrak{U})$ are topological spaces and $\phi : X \to Y$ and $\psi : Y \to Z$ are continuous, then $\psi\phi : X \to Z$ is continuous. **P** If $G \in \mathfrak{U}$ then $\psi^{-1}[G] \in \mathfrak{S}$ so $(\psi\phi)^{-1}[G] = \phi^{-1}[\psi^{-1}[G]] \in \mathfrak{T}$. **Q**

(c) If $(X,\mathfrak{T})$ is a topological space, a function $f : X \to \mathbb{R}$ is continuous iff $\{x : a < f(x) < b\}$ is open whenever $a < b$ in $\mathbb{R}$. **P (i)** Every interval $]a,b[$ is open in $\mathbb{R}$, so if f is continuous its inverse image $\{x : a < f(x) < b\}$ must be open. **(ii)** Suppose that $f^{-1}[\,]a,b[\,]$ is open whenever $a < b$, and let $H \subseteq \mathbb{R}$ be any open set. By the definition of 'open' set in $\mathbb{R}$ (1A2A),

$$H = \bigcup\{]y-\delta, y+\delta[: y \in \mathbb{R},\ \delta > 0,\ ]y-\delta, y+\delta[\subseteq H\},$$

so

$$f^{-1}[H] = \bigcup\{f^{-1}[\,]y-\delta, y+\delta[\,] : y \in \mathbb{R},\ \delta > 0,\ ]y-\delta, y+\delta[\subseteq H\}$$

is a union of open sets in X, therefore open. **Q**

(d) If $r \ge 1$, $(X, \mathfrak{T})$ is a topological space, and $\phi : X \to \mathbb{R}^r$ is a function, then ϕ is continuous iff $\phi_i : X \to \mathbb{R}$ is continuous for each $i \le r$, where $\phi(x) = (\phi_1(x), \dots, \phi_r(x))$ for each $x \in X$. **P** **(i)** Suppose that ϕ is continuous. For $i \le r$, $y = (\eta_1, \dots, \eta_r) \in \mathbb{R}^r$, set $\pi_i(y) = \eta_i$. Then $|\pi_i(y) - \pi_i(z)| \le \|y - z\|$ for all y, $z \in \mathbb{R}^r$ so $\pi_i : \mathbb{R}^r \to \mathbb{R}$ is continuous. Consequently $\phi_i = \pi_i\phi$ is continuous, by (b) above. **(ii)** Suppose that every ϕ_i is continuous, and that $H \subseteq \mathbb{R}^r$ is open. Set

$$\mathcal{G} = \{G : G \subseteq X \text{ is open, } G \subseteq \phi^{-1}[H]\}.$$

Then $G_0 = \bigcup \mathcal{G}$ is open, and $G_0 \subseteq \phi^{-1}[H]$. But suppose that x_0 is any point of $\phi^{-1}[H]$. Then there is a $\delta > 0$ such that $U(\phi(x_0), \delta) \subseteq H$, because H is open and contains $\phi(x_0)$. For $1 \le i \le r$ set $V_i = \{x : \phi_i(x_0) - \frac{\delta}{\sqrt{r}} < \phi_i(x) < \phi_i(x_0) + \frac{\delta}{\sqrt{r}}\}$; then V_i is the inverse image of an open set under the continuous map ϕ_i, so is open. Set $G = \bigcap_{i \le r} V_i$. Then G is open (using (ii) of the definition 2A3A), $x_0 \in G$, and $\|\phi(x) - \phi(x_0\| < \delta$ for every $x \in G$, so $G \subseteq \phi^{-1}[H]$, $G \in \mathcal{G}$ and $x_0 \in G_0$. This shows that $\phi^{-1}[H] = G_0$ is open. As H is arbitrary, ϕ is continuous. **Q**

(e) If $(X, \mathfrak{T})$ is a topological space, $f_1, \dots, f_r$ are continuous functions from X to $\mathbb{R}$, and $h : \mathbb{R}^r \to \mathbb{R}$ is continuous, then $h(f_1, \dots, f_r) : X \to \mathbb{R}$ is continuous. **P** Set $\phi(x) = (f_1(x), \dots, f_r(x)) \in \mathbb{R}^r$ for $x \in X$. By (d), ϕ is continuous, so by 2A3Bb $h(f_1, \dots, f_r) = h\phi$ is continuous. **Q** In particular, $f + g$, $f \times g$ and $f - g$ are continuous for all continuous functions f, $g : X \to \mathbb{R}$.

(f) If $(X, \mathfrak{T})$ and $(Y, \mathfrak{S})$ are topological spaces and $\phi : X \to Y$ is a continuous function, then $\phi^{-1}[F]$ is closed in X for every closed set $F \subseteq Y$. (For $X \setminus \phi^{-1}[F] = \phi^{-1}[Y \setminus F]$ is open.)

2A3C Subspace topologies If $(X, \mathfrak{T})$ is a topological space and $D \subseteq X$, then $\mathfrak{T}_D = \{G \cap D : G \in \mathfrak{T}\}$ is a topology on D. **P** (i) $\emptyset = \emptyset \cap D$ and $D = X \cap D$ belong to $\mathfrak{T}_D$. (ii) If G, $H \in \mathfrak{T}_D$ there are G', $H' \in \mathfrak{T}$ such that $G = G' \cap D$, $H = H' \cap D$; now $G \cap H = G' \cap H' \cap D \in \mathfrak{T}_D$. (iii) If $\mathcal{G} \subseteq \mathfrak{T}_D$ set $\mathcal{H} = \{H : H \in \mathfrak{T},\ H \cap D \in \mathcal{G}\}$; then $\bigcup \mathcal{G} = (\bigcup \mathcal{H}) \cap D \in \mathfrak{T}_D$. **Q**

$\mathfrak{T}_D$ is called the **subspace topology** on D, or the topology on D **induced** by $\mathfrak{T}$. If $(Y, \mathfrak{S})$ is another topological space, and $\phi : X \to Y$ is $(\mathfrak{T}, \mathfrak{S})$-continuous, then $\phi \restriction D : D \to Y$ is $(\mathfrak{T}_D, \mathfrak{S})$-continuous. (For if $H \in \mathfrak{S}$ then

$$(\phi \restriction D)^{-1}[H] = D \cap \phi^{-1}[H] \in \mathfrak{T}_D.)$$

2A3D Closures and interiors (a) In the proof of 2A3Bd I have already used the following idea. Let $(X, \mathfrak{T})$ be any topological space and A any subset of X. Write

$$\operatorname{int} A = \bigcup\{G : G \in \mathfrak{T},\ G \subseteq A\}.$$

Then $\operatorname{int} A$ is an open set, being a union of open sets, and is of course included in A; it must be the largest open set included in A, and is called the **interior** of A.

(b) Because a set is closed iff its complement is open, we have a complementary notion:

$$\begin{aligned}
\overline{A} &= \bigcap\{F : F \text{ is closed, } A \subseteq F\} \\
&= X \setminus \bigcup\{X \setminus F : F \text{ is closed, } A \subseteq F\} \\
&= X \setminus \bigcup\{G : G \text{ is open, } A \cap G = \emptyset\} \\
&= X \setminus \bigcup\{G : G \text{ is open, } G \subseteq X \setminus A\} = X \setminus \operatorname{int}(X \setminus A).
\end{aligned}$$

$\overline{A}$ is closed (being the complement of an open set) and is the smallest closed set including A; it is called the **closure** of A. (Compare 2A2A.) Because the union of two closed sets is closed (cf. 1A2Fc), $\overline{A \cup B} = \overline{A} \cup \overline{B}$ for all A, $B \subseteq X$.

(c) There are innumerable ways of looking at these concepts; a useful description of the closure of a set is

$$\begin{aligned} x \in \overline{A} &\iff x \notin \operatorname{int}(X \setminus A) \\ &\iff \text{there is no open set containing } x \text{ and included in } X \setminus A \\ &\iff \text{every open set containing } x \text{ meets } A. \end{aligned}$$

2A3E Hausdorff topologies (a) The concept of 'topological space' is so widely drawn, and so widely applicable, that a vast number of different types of topological space have been studied. For this volume we shall not need much of the (very extensive) vocabulary which has been developed to describe this variety. But one useful word (and one of the most important concepts) is that of 'Hausdorff space'; a topological space X is **Hausdorff** if for all distinct x, $y \in X$ there are disjoint open sets G, $H \subseteq X$ such that $x \in G$ and $y \in H$.

(b) In a Hausdorff space X, finite sets are closed. **P** If $z \in X$, then for any $x \in X \setminus \{z\}$ there is an open set containing x but not z, so $X \setminus \{z\}$ is open and $\{z\}$ is closed. So a finite set is a finite union of closed sets and is therefore closed. **Q**

2A3F Pseudometrics Many important topologies (not all!) can be defined by families of pseudometrics; it will be useful to have a certain amount of technical skill with these.

(a) Let X be a set. A **pseudometric** on X is a function $\rho : X \times X \to [0, \infty[$ such that

$\rho(x, z) \le \rho(x, y) + \rho(y, z)$ for all x, y, $z \in X$

(the 'triangle inequality';)

$\rho(x, y) = \rho(y, x)$ for all x, $y \in X$;

$\rho(x, x) = 0$ for all $x \in X$.

A **metric** is a pseudometric ρ satisfying the further condition

if $\rho(x, y) = 0$ then $x = y$.

(b) Examples (i) For x, $y \in \mathbb{R}$, set $\rho(x, y) = |x - y|$; then ρ is a metric on $\mathbb{R}$ (the 'usual metric' on $\mathbb{R}$).

(ii) For x, $y \in \mathbb{R}^r$, where $r \ge 1$, set $\rho(x, y) = \|x - y\|$, defining $\|z\| = \sqrt{\sum_{i=1}^r \zeta_i^2}$, as usual. Then ρ is a metric, the **Euclidean metric** on $\mathbb{R}^r$. (The triangle inequality for ρ comes from Cauchy's inequality in 1A2C: if x, y, $z \in \mathbb{R}^r$, then

$$\rho(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \le \|x - y\| + \|y - z\| = \rho(x, y) + \rho(y, z).$$

The other required properties of ρ are elementary. Compare 2A4Bb below.)

(iii) For an example of a pseudometric which is not a metric, take $r \ge 2$ and define $\rho : \mathbb{R}^r \times \mathbb{R}^r \to [0, \infty[$ by setting $\rho(x, y) = |\xi_1 - \eta_1|$ whenever $x = (\xi_1, \dots, \xi_r)$, $y = (\eta_1, \dots, \eta_r) \in \mathbb{R}^r$.

(c) Now let X be a set and P a non-empty family of pseudometrics on X. Let $\mathfrak{T}$ be the family of those subsets G of X such that for every $x \in G$ there are $\rho_0, \dots, \rho_n \in \mathrm{P}$ and $\delta > 0$ such that

$$U(x; \rho_0, \dots, \rho_n; \delta) = \{y : y \in X, \max_{i \le n} \rho_i(y, x) < \delta\} \subseteq G.$$

Then $\mathfrak{T}$ is a topology on X.

P (Compare 1A2B.) (i) $\emptyset \in \mathfrak{T}$ because the condition is vacuously satisfied. $X \in \mathfrak{T}$ because $U(x; \rho; 1) \subseteq X$ for any $x \in X$, $\rho \in \mathrm{P}$. (ii) If G, $H \in \mathfrak{T}$ and $x \in G \cap H$, take $\rho_0, \dots, \rho_m, \rho_0', \dots, \rho_n' \in \mathrm{P}$, δ, $\delta' > 0$ such that $U(x; \rho_0, \dots, \rho_m; \delta) \subseteq G$, $U(x; \rho_0', \dots, \rho_n'; \delta') \subseteq G$; then

$$U(x; \rho_0, \dots, \rho_m, \rho_0', \dots, \rho_n'; \min(\delta, \delta')) \subseteq G \cap H.$$

As x is arbitrary, $G \cap H \in \mathfrak{T}$. (iii) If $\mathcal{G} \subseteq \mathfrak{T}$ and $x \in \bigcup \mathcal{G}$, there is a $G \in \mathcal{G}$ such that $x \in G$; now there are $\rho_0, \dots, \rho_n \in \mathrm{P}$ and $\delta > 0$ such that

$$U(x; \rho_0, \dots, \rho_n; \delta) \subseteq G \subseteq \bigcup \mathcal{G}.$$

As x is arbitrary, $\bigcup \mathcal{G} \in \mathfrak{T}$. **Q**

$\mathfrak{T}$ is the **topology defined by** P.

(d) You may wish to have a convention to deal with the case in which P is the empty set; in this case the topology on X defined by P is $\{\emptyset, X\}$.

(e) In many important cases, P is upwards-directed in the sense that for any ρ_1, $\rho_2 \in \mathrm{P}$ there is a $\rho \in \mathrm{P}$ such that $\rho_i(x,y) \le \rho(x,y)$ for all x, $y \in X$ and both i. In this case, of course, any set $U(x;\rho_0,\dots,\rho_n;\delta)$, where $\rho_0,\dots,\rho_n \in \mathrm{P}$, includes some set of the form $U(x;\rho;\delta)$, where $\rho \in \mathrm{P}$. Consequently, for instance, a set $G \subseteq X$ is open iff for every $x \in G$ there are $\rho \in \mathrm{P}$, $\delta > 0$ such that $U(x;\rho;\delta) \subseteq G$.

(f) A topology $\mathfrak{T}$ is **metrizable** if it is the topology defined by a family P consisting of a single metric. Thus the **Euclidean topology** on $\mathbb{R}^r$ is the metrizable topology defined by $\{\rho\}$, where ρ is the metric of (b-ii) above.

2A3G Proposition Let X be a set with a topology defined by a non-empty set P of pseudometrics on X. Then $U(x;\rho_0,\dots,\rho_n;\epsilon)$ is open for all $x \in X$, $\rho_0,\dots,\rho_n \in \mathrm{P}$ and $\epsilon > 0$.

proof (Compare 1A2D.) Take $y \in U(x;\rho_0,\dots,\rho_n;\epsilon)$. Set

$$\eta = \max_{i \le n} \rho_i(y,x), \quad \delta = \epsilon - \eta > 0.$$

If $z \in U(y;\rho_0,\dots,\rho_n;\delta)$ then

$$\rho_i(z,x) \le \rho_i(z,y) + \rho_i(y,x) < \delta + \eta = \epsilon$$

for each $i \le n$, so $U(y;\rho_0,\dots,\rho_n;\delta) \subseteq U(x;\rho_0,\dots,\rho_n;\epsilon)$. As y is arbitrary, $U(x;\rho_0,\dots,\rho_n;\epsilon)$ is open.

2A3H Now we have a result corresponding to 2A2Ca, describing continuous functions between topological spaces defined by families of pseudometrics.

Proposition Let X and Y be sets; let P be a non-empty family of pseudometrics on X, and Θ a non-empty family of pseudometrics on Y; let $\mathfrak{T}$ and $\mathfrak{S}$ be the corresponding topologies. Then a function $\phi : X \to Y$ is continuous iff whenever $x \in X$, $\theta \in \Theta$ and $\epsilon > 0$, there are $\rho_0,\dots,\rho_n \in \mathrm{P}$ and $\delta > 0$ such that $\theta(\phi(y),\phi(x)) \le \epsilon$ whenever $y \in X$ and $\max_{i \le n} \rho_i(y,x) \le \delta$.

proof (a) Suppose that ϕ is continuous; take $x \in X$, $\theta \in \Theta$ and $\epsilon > 0$. By 2A3G, $U(\phi(x);\theta;\epsilon) \in \mathfrak{S}$. So $G = \phi^{-1}[U(\phi(x);\theta;\epsilon)] \in \mathfrak{T}$. Now $x \in G$, so there are $\rho_0,\dots,\rho_n \in \mathrm{P}$ and $\delta > 0$ such that $U(x;\rho_0,\dots,\rho_n;\delta) \subseteq G$. In this case $\theta(\phi(y),\phi(x)) \le \epsilon$ whenever $y \in X$ and $\max_{i \le n} \rho_i(y,x) \le \frac{1}{2}\delta$. As x, θ and ϵ are arbitrary, ϕ satisfies the condition.

(b) Suppose ϕ satisfies the condition. Take $H \in \mathfrak{S}$ and consider $G = \phi^{-1}[H]$. If $x \in G$, then $\phi(x) \in H$, so there are $\theta_0,\dots,\theta_n \in \Theta$ and $\epsilon > 0$ such that $U(\phi(x);\theta_0,\dots,\theta_n;\epsilon) \subseteq H$. For each $i \le n$ there are $\rho_{i0},\dots,\rho_{i,m_i} \in \mathrm{P}$ and $\delta_i > 0$ such that $\theta(\phi(y),\phi(x)) \le \frac{1}{2}\epsilon$ whenever $y \in X$ and $\max_{j \le m_i} \rho_{ij}(y,x) \le \delta_i$. Set $\delta = \min_{i \le n} \delta_i > 0$; then

$$U(x;\rho_{00},\dots,\rho_{0,m_0},\dots,\rho_{n0},\dots,\rho_{n,m_n};\delta) \subseteq G.$$

As x is arbitrary, $G \in \mathfrak{T}$. As H is arbitrary, ϕ is continuous.

2A3I Remarks (a) If P is upwards-directed, the condition simplifies to: for every $x \in X$, $\theta \in \Theta$ and $\epsilon > 0$, there are $\rho \in \mathrm{P}$ and $\delta > 0$ such that $\theta(\phi(y),\phi(x)) \le \epsilon$ whenever $y \in X$ and $\rho(y,x) \le \delta$.

(b) Suppose we have a set X and two non-empty families P, Θ of pseudometrics on X, generating topologies $\mathfrak{T}$ and $\mathfrak{S}$ on X. Then $\mathfrak{S} \subseteq \mathfrak{T}$ iff the identity map ϕ from X to itself is a continuous function when regarded as a map from $(X,\mathfrak{T})$ to $(X,\mathfrak{S})$, because this will mean that $G = \phi^{-1}[G]$ belongs to $\mathfrak{T}$ whenever $G \in \mathfrak{S}$. Applying the proposition above to ϕ, we see that this happens iff for every $\theta \in \Theta$, $x \in X$ and $\epsilon > 0$ there are $\rho_0,\dots,\rho_n \in \mathrm{P}$ and $\delta > 0$ such that $\theta(y,x) \le \epsilon$ whenever $y \in X$ and $\max_{i \le n} \rho_i(y,x) \le \delta$. Similarly, reversing the roles of P and Θ, we get a criterion for when $\mathfrak{T} \subseteq \mathfrak{S}$, and putting the two together we obtain a criterion to determine when $\mathfrak{T} = \mathfrak{S}$.

2A3J Subspaces: Proposition If X is a set, P a non-empty family of pseudometrics on X defining a topology $\mathfrak{T}$ on X, and $D \subseteq X$, then

(a) for every $\rho \in \mathrm{P}$, the restriction $\rho^{(D)}$ of ρ to $D \times D$ is a pseudometric on D;

(b) the topology defined by $\mathrm{P}_D = \{\rho^{(D)} : \rho \in \mathrm{P}\}$ on D is precisely the subspace topology $\mathfrak{T}_D$ described in 2A3C.

proof (a) is just a matter of reading through the definition in 2A3Fa. For (b), we have to think for a moment.

(i) Suppose that G belongs to the topology defined by P_D. Set

$$\mathcal{H} = \{H : H \in \mathfrak{T},\, H \cap D \subseteq G\},$$

$$H^* = \bigcup \mathcal{H} \in \mathfrak{T}, \quad G^* = H^* \cap D \in \mathfrak{T}_D;$$

then $G^* \subseteq G$. On the other hand, if $x \in G$, then there are $\rho_0, \dots, \rho_n \in \mathrm{P}$ and $\delta > 0$ such that

$$U(x; \rho_0^{(D)}, \dots, \rho_n^{(D)}; \delta) = \{y : y \in D,\, \max_{i \le n} \rho_i^{(D)}(y, x) < \delta\} \subseteq G.$$

Consider

$$H = U(x; \rho_0, \dots, \rho_n; \delta) = \{y : y \in X,\, \max_{i \le n} \rho_i(y, x) < \delta\} \subseteq X.$$

Evidently

$$H \cap D = U(x; \rho_0^{(D)}, \dots, \rho_n^{(D)}; \delta) \subseteq G.$$

Also $H \in \mathfrak{T}$. So $H \in \mathcal{H}$ and

$$x \in H \cap D \subseteq H^* \cap D = G^*.$$

Thus $G = G^* \in \mathfrak{T}_D$.

(ii) Now suppose that $G \in \mathfrak{T}_D$. Then there is an $H \in \mathfrak{T}$ such that $G = H \cap D$. Consider the identity map $\phi : D \to X$, defined by saying that $\phi(x) = x$ for every $x \in D$. ϕ obviously satisfies the criterion of 2A3H, if we endow D with P_D and X with P, because $\rho(\phi(x), \phi(y)) = \rho^{(D)}(x, y)$ whenever x, $y \in D$ and $\rho \in \mathrm{P}$; so ϕ must be continuous for the associated topologies, and $\phi^{-1}[H]$ must belong to the topology defined by P_D. But $\phi^{-1}[H] = G$. Thus every set in $\mathfrak{T}_D$ belongs to the topology defined by P_D, and the two topologies are the same, as claimed.

2A3K Closures and interiors Let X be a set, P a non-empty family of pseudometrics on X and $\mathfrak{T}$ the topology defined by P.

(a) For any $A \subseteq X$ and $x \in X$,

$$\begin{aligned} x \in \operatorname{int} A &\iff \text{there is an open set included in } A \text{ containing } x \\ &\iff \text{there are } \rho_0, \dots, \rho_n \in \mathrm{P},\ \delta > 0 \text{ such that } U(x; \rho_0, \dots, \rho_n; \delta) \subseteq A. \end{aligned}$$

(b) For any $A \subseteq X$ and $x \in X$, $x \in \overline{A}$ iff $U(x; \rho_0, \dots, \rho_n; \delta) \cap A \ne \emptyset$ for every $\rho_0, \dots, \rho_n \in \mathrm{P}$ and $\delta > 0$. (Compare 2A2B(ii), 2A3Dc.)

2A3L Hausdorff topologies Recall that a topology $\mathfrak{T}$ is Hausdorff if any two points can be separated by open sets (2A3E). Now a topology defined on a set X by a non-empty family P of pseudometrics is Hausdorff iff for any two different points x, y of X there is a $\rho \in \mathrm{P}$ such that $\rho(x, y) > 0$. **P (i)** Suppose that the topology is Hausdorff and that x, y are distinct points in X. Then there is an open set G containing x but not containing y. Now there are $\rho_0, \dots, \rho_n \in \mathrm{P}$ and $\delta > 0$ such that $U(x; \rho_0), \dots, \rho_n; \delta) \subseteq G$, in which case $\rho_i(y, x) \ge \delta > 0$ for some $i \le n$. **(ii)** If P satisfies the condition, and x, y are distinct points of X, take $\rho \in \mathrm{P}$ such that $\rho(x, y) > 0$, and set $\delta = \frac{1}{2}\rho(x, y)$. Then $U(x; \rho; \delta)$ and $U(y; \rho; \delta)$ are disjoint (because if $z \in X$, then

$$\rho(z, x) + \rho(z, y) \ge \rho(x, y) = 2\delta,$$

so at least one of $\rho(z, x)$, $\rho(z, y)$ is greater than or equal to δ), and they are open sets containing x, y respectively. As x and y are arbitrary, the topology is Hausdorff. **Q**

In particular, metrizable topologies are Hausdorff.

2A3M Convergence of sequences (a) If $(X,\mathfrak{T})$ is any topological space, and $\langle x_n\rangle_{n\in\mathbb{N}}$ is a sequence in X, we say that $\langle x_n\rangle_{n\in\mathbb{N}}$ **converges** to $x\in X$, or that x is a **limit** of $\langle x_n\rangle_{n\in\mathbb{N}}$, or $\langle x_n\rangle_{n\in\mathbb{N}}\to x$, if for every open set G containing x there is an $n_0\in\mathbb{N}$ such that $x_n\in G$ for every $n\ge n_0$.

(b) Warning In general topological spaces, it is possible for a sequence to have more than one limit, and we cannot safely write $x=\lim_{n\to\infty}x_n$. But in Hausdorff spaces, this does not occur. **P** If $\mathfrak{T}$ is Hausdorff, and x, y are distinct points of X, there are disjoint open sets G, H such that $x\in G$ and $y\in H$. If now $\langle x_n\rangle_{n\in\mathbb{N}}$ converges to x, there is an n_0 such that $x_n\in G$ for every $n\ge n_0$, so $x_n\notin H$ for every $n\ge n_0$, and $\langle x_n\rangle_{n\in\mathbb{N}}$ cannot converge to y. **Q** In particular, a sequence in a metrizable space can have at most one limit.

(c) Let X be a set, and P a non-empty family of pseudometrics on X, generating a topology $\mathfrak{T}$; let $\langle x_n\rangle_{n\in\mathbb{N}}$ be a sequence in X and $x\in X$. Then $\langle x_n\rangle_{n\in\mathbb{N}}$ converges to x iff $\lim_{n\to\infty}\rho(x_n,x)=0$ for every $\rho\in\mathrm{P}$. **P (i)** Suppose that $\langle x_n\rangle_{n\in\mathbb{N}}\to x$ and that $\rho\in\mathrm{P}$. Then for any $\epsilon>0$ the set $G=U(x;\rho;\epsilon)$ is an open set containing x, so there is an n_0 such that $x_n\in G$ for every $n\ge n_0$, that is, $\rho(x_n,x)<\epsilon$ for every $n\ge n_0$. As ϵ is arbitrary, $\lim_{n\to\infty}\rho(x_n,x)=0$. **(ii)** If the condition is satisfied, take any open set G containing X. Then there are $\rho_0,\dots,\rho_k\in\mathrm{P}$ and $\delta>0$ such that $U(x;\rho_0,\dots,\rho_k;\delta)\subseteq G$. For each $i\le k$ there is an $n_i\in\mathbb{N}$ such that $\rho_i(x_n,x)<\delta$ for every $n\ge n_i$. Set $n^*=\max(n_0,\dots,n_k)$; then $x_n\in U(x;\rho_0,\dots,\rho_k;\delta)\subseteq G$ for every $n\ge n^*$. As G is arbitrary, $\langle x_n\rangle_{n\in\mathbb{N}}\to x$. **Q**

(d) Let (X,ρ) be a metric space, A a subset of X and $x\in X$. Then $x\in\overline{A}$ iff there is a sequence in A converging to x. **P(i)** If $x\in\overline{A}$, then for every $n\in\mathbb{N}$ we can choose a point $x_n\in A\cap U(x;\rho;2^{-n})$ (2A3Kb); now $\langle x_n\rangle_{n\in\mathbb{N}}\to x$. **(ii)** If $\langle x_n\rangle_{n\in\mathbb{N}}$ is a sequence in A converging to x, then for every open set G containing x there is an n such that $x_n\in G$, so that $A\cap G\ne\emptyset$; by 2A3Dc, $x\in\overline{A}$. **Q**

2A3N Compactness The next concept we need is the idea of 'compactness' in general topological spaces.

(a) If $(X,\mathfrak{T})$ is any topological space, a subset K of X is **compact** if whenever $\mathcal{G}$ is a family in $\mathfrak{T}$ covering K, then there is a finite $\mathcal{G}_0\subseteq\mathcal{G}$ covering K. (Cf. 2A2D. A **warning**: many authors reserve the term 'compact' for Hausdorff spaces.) A set $A\subseteq X$ is **relatively compact** in X if there is a compact subset of X including A. (**Warning!** in non-Hausdorff spaces, this is not the same thing as saying that $\overline{A}$ is compact.)

(b) Just as in 2A2E-2A2G (and the proofs are the same in the general case), we have the following results.

(i) If K is compact and E is closed, then $K\cap E$ is compact.

(ii) If $K\subseteq X$ is compact and $\phi:K\to Y$ is continuous, where $(Y,\mathfrak{S})$ is another topological space, then $\phi[K]$ is a compact subset of Y.

(iii) If $K\subseteq X$ is compact and $\phi:K\to\mathbb{R}$ is continuous, then ϕ is bounded and attains its bounds.

2A3O Cluster points (a) If $(X,\mathfrak{T})$ is a topological space, and $\langle x_n\rangle_{n\in\mathbb{N}}$ is a sequence in X, then a **cluster point** of $\langle x_n\rangle_{n\in\mathbb{N}}$ is an $x\in X$ such that whenever G is an open set containing x and $n\in\mathbb{N}$ then there is a $k\ge n$ such that $x_k\in G$.

(b) Now if $(X,\mathfrak{T})$ is a topological space and $A\subseteq X$ is relatively compact, every sequence $\langle x_n\rangle_{n\in\mathbb{N}}$ in A has a cluster point in X. **P** Let K be a compact subset of X including A. Set

$$\mathcal{G}=\{G:G\in\mathfrak{T},\ \{n:x_n\in G\}\text{ is finite}\}.$$

? If $\mathcal{G}$ covers K, then there is a finite $\mathcal{G}_0\subseteq\mathcal{G}$ covering K. Now

$$\mathbb{N}=\{n:x_n\in A\}=\{n:x_n\in\bigcup\mathcal{G}_0\}=\bigcup_{G\in\mathcal{G}_0}\{n:x_n\in G\}$$

is a finite union of finite sets, which is absurd. **X** Thus $\mathcal{G}$ does not cover K. Take any $x\in K\setminus\bigcup\mathcal{G}$. If $G\in\mathfrak{T}$ and $x\in G$ and $n\in\mathbb{N}$, then $G\notin\mathcal{G}$ so $\{k:x_k\in G\}$ is infinite and there is a $k\ge n$ such that $x_k\in G$. Thus x is a cluster point of $\langle x_n\rangle_{n\in\mathbb{N}}$, as required. **Q**

2A3P Filters In $\mathbb{R}^r$, and more generally in all metrizable spaces, topological ideas can be effectively discussed in terms of convergent sequences. (To be sure, this occasionally necessitates the use of a weak form of the axiom of choice, in order to choose a sequence; but as measure theory without such choices is changed utterly – see Chapter 56 in Volume 5 – there is no point in fussing about them here.) For topological spaces in general, however, sequences are quite inadequate, for very interesting reasons which I shall not enlarge upon. Instead we need to use 'nets' or 'filters'. The latter take a moment's more effort at the beginning, but are then (in my view) much easier to work with, so I describe this method now.

2A3Q Convergent filters (a) Let $(X,\mathfrak{T})$ be a topological space, $\mathcal{F}$ a filter on X (see 2A1I) and x a point of X. We say that $\mathcal{F}$ is **convergent** to x, or that x is a **limit** of $\mathcal{F}$, and write $\mathcal{F}\to x$, if every open set containing x belongs to $\mathcal{F}$.

(b) Let $(X,\mathfrak{T})$ and $(Y,\mathfrak{S})$ be topological spaces, $\phi:X\to Y$ a continuous function, $x\in X$ and $\mathcal{F}$ a filter on X converging to x. Then $\phi[[\mathcal{F}]]$ (as defined in 2A1Ib) converges to $\phi(x)$ (because $\phi^{-1}[G]$ is an open set containing x whenever G is an open set containing $\phi(x)$).

2A3R Now we have the following characterization of compactness.

Theorem Let X be a topological space, and K a subset of X. Then K is compact iff every ultrafilter on X containing K has a limit in K.

proof (a) Suppose that K is compact and that $\mathcal{F}$ is an ultrafilter on X containing K. Set

$$\mathcal{G}=\{G:G\subseteq X\text{ is open},\ X\setminus G\in\mathcal{F}\}.$$

Then the union of any two members of $\mathcal{G}$ belongs to $\mathcal{G}$, so the union of any finite number of members of $\mathcal{G}$ belongs to $\mathcal{G}$; also no member of $\mathcal{G}$ can include K, because $X\setminus K\notin\mathcal{F}$. Because K is compact, it follows that $\mathcal{G}$ cannot cover K. Let x be any point of $K\setminus\bigcup\mathcal{G}$. If G is any open set containing x, then $G\notin\mathcal{G}$ so $X\setminus G\notin\mathcal{F}$; but this means that G must belong to $\mathcal{F}$, because $\mathcal{F}$ is an ultrafilter. As G is arbitrary, $\mathcal{F}\to x$. Thus every ultrafilter on X containing K has a limit in K.

(b) Now suppose that every ultrafilter on X containing K has a limit in K. Let $\mathcal{G}$ be a cover of K by open sets in X. **?** Suppose, if possible, that $\mathcal{G}$ has no finite subcover. Set

$$\mathcal{F}=\{F:\text{there is a finite }\mathcal{G}_0\subseteq\mathcal{G},\ F\cup\bigcup\mathcal{G}_0\supseteq K\}.$$

Then $\mathcal{F}$ is a filter on X. **P** (i) $X\cup\bigcup\emptyset\supseteq K$ so $X\in\mathcal{F}$.

$$\emptyset\cup\bigcup\mathcal{G}_0=\bigcup\mathcal{G}_0\not\supseteq K$$

for any finite $\mathcal{G}_0\subseteq\mathcal{G}$, by hypothesis, so $\emptyset\notin\mathcal{F}$. (ii) If E, $F\in\mathcal{F}$ there are finite sets $\mathcal{G}_1$, $\mathcal{G}_2\subseteq\mathcal{G}$ such that $E\cup\bigcup\mathcal{G}_1$ and $F\cup\bigcup\mathcal{G}_2$ both include K; now $(E\cap F)\cup\bigcup(\mathcal{G}_1\cup\mathcal{G}_2)\supseteq K$ so $E\cap F\in\mathcal{F}$. (iii) If $X\supseteq E\supseteq F\in\mathcal{F}$ then there is a finite $\mathcal{G}_0\subseteq\mathcal{G}$ such that $F\cup\mathcal{G}_0\supseteq K$; now $E\cup\bigcup\mathcal{G}_0\supseteq K$ and $E\in\mathcal{F}$. **Q**

By the Ultrafilter Theorem (2A1O), there is an ultrafilter $\mathcal{F}^*$ on X including $\mathcal{F}$. Of course K itself belongs to $\mathcal{F}$, so $K\in\mathcal{F}^*$. By hypothesis, $\mathcal{F}^*$ has a limit $x\in K$. But now there is a set $G\in\mathcal{G}$ containing x, and $(X\setminus G)\cup G\supseteq K$, so $X\setminus G\in\mathcal{F}\subseteq\mathcal{F}^*$; which means that G cannot belong to $\mathcal{F}^*$, and x cannot be a limit of $\mathcal{F}^*$. **X**

So $\mathcal{G}$ has a finite subcover. As $\mathcal{G}$ is arbitrary, K must be compact.

Remark Note that part (b) of the proof of this theorem depends vitally on the Ultrafilter Theorem and therefore on the axiom of choice.

2A3S Further calculations with filters (a) In general, it is possible for a filter to have more than one limit; but in Hausdorff spaces this does not occur. **P** (Compare 2A3Mb.) If $(X,\mathfrak{T})$ is Hausdorff, and x, y are distinct points of X, there are disjoint open sets G, H such that $x\in G$ and $y\in H$. If now a filter $\mathcal{F}$ on X converges to x, $G\in\mathcal{F}$ so $H\notin\mathcal{F}$ and $\mathcal{F}$ does not converge to y. **Q**

Accordingly we can safely write $x=\lim\mathcal{F}$ when $\mathcal{F}\to x$ in a Hausdorff space.

(b) Now suppose that X is a set, $\mathcal{F}$ is a filter on X, $(Y,\mathfrak{S})$ is a Hausdorff space, $D\in\mathcal{F}$ and $\phi:D\to Y$ is a function. Then we write $\lim_{x\to\mathcal{F}}\phi(x)$ for $\lim\phi[[\mathcal{F}]]$ if this is defined in Y; that is, $\lim_{x\to\mathcal{F}}\phi(x)=y$ iff $\phi^{-1}[H]\in\mathcal{F}$ for every open set H containing y.

If Z is another set, $\mathcal{G}$ is a filter on Z, and $\psi : Z \to X$ is such that $\mathcal{F} = \psi[[\mathcal{G}]]$, then the composition $\phi\psi$ is defined on $\psi^{-1}[D] \in \mathcal{G}$, and if one of the limits $\lim_{x\to\mathcal{F}} \phi(x)$, $\lim_{z\to\mathcal{G}} \phi\psi(z)$ is defined in Y so is the other, and they are then equal. **P** Suppose that $y \in Y$ and let $\mathcal{U}$ be the family of open subsets of Y containing y. Then

$$\begin{aligned}\lim_{x\to\mathcal{F}} \phi(x) = y &\iff \phi^{-1}[G] \in \mathcal{F} \text{ for every } G \in \mathcal{U}\\ &\iff \psi^{-1}[\phi^{-1}[G]] \in \mathcal{G} \text{ for every } G \in \mathcal{U}\\ &\iff (\phi\psi)^{-1}[G] \in \mathcal{G} \text{ for every } G \in \mathcal{U} \iff \lim_{z\to\mathcal{G}} \phi\psi(z) = y. \ \mathbf{Q}\end{aligned}$$

In the special case $Y = \mathbb{R}$, $\lim_{x\to\mathcal{F}} \phi(x) = a$ iff $\{x : |\phi(x) - a| \le \epsilon\} \in \mathcal{F}$ for every $\epsilon > 0$ (because every open set containing a includes a set of the form $[a - \epsilon, a + \epsilon]$, which in turn includes the open set $]a - \epsilon, a + \epsilon[$).

(c) Suppose that X and Y are sets, $\mathcal{F}$ is a filter on X, Θ is a non-empty family of pseudometrics on Y defining a topology $\mathfrak{S}$ on Y, and $\phi : X \to Y$ is a function. Then the image filter $\phi[[\mathcal{F}]]$ converges to $y \in Y$ iff $\lim_{x\to\mathcal{F}} \theta(\phi(x), y) = 0$ in $\mathbb{R}$ for every $\theta \in \Theta$. **P** **(i)** Suppose that $\phi[[\mathcal{F}]] \to y$. For every $\theta \in \Theta$ and $\epsilon > 0$, $U(y; \theta; \epsilon) = \{z : \theta(z, y) < \epsilon\}$ is an open set containing y (2A3G), so belongs to $\phi[[\mathcal{F}]]$, and its inverse image $\{x : 0 \le \theta(\phi(x), y) < \epsilon\}$ belongs to $\mathcal{F}$. As ϵ is arbitrary, $\lim_{x\to\mathcal{F}} \theta(\phi(x), y) = 0$. As θ is arbitrary, ϕ satisfies the condition. **(ii)** Now suppose that $\lim_{x\to\mathcal{F}} \theta(\phi(x), y) = 0$ for every $\theta \in \Theta$. Let G be any open set in Y containing y. Then there are $\theta_0, \dots, \theta_n \in \Theta$ and $\epsilon > 0$ such that

$$U(y; \theta_0, \dots, \theta_n; \epsilon) = \bigcap_{i\le n} U(y; \theta_i; \epsilon) \subseteq G.$$

For each $i \le n$,

$$\phi^{-1}[U(y; \theta_i; \epsilon)] = \{x : \theta(\phi(x), y) < \epsilon\}$$

belongs to $\mathcal{F}$; because $\mathcal{F}$ is closed under finite intersections, so do $\phi^{-1}[U(y; \theta_0, \dots, \theta_n; \epsilon)]$ and its superset $\phi^{-1}[G]$. Thus $G \in \phi[[\mathcal{F}]]$. As G is arbitrary, $\phi[[\mathcal{F}]] \to y$. **Q**

(d) In particular, taking $X = Y$ and ϕ the identity map, if X has a topology $\mathfrak{T}$ defined by a non-empty family P of pseudometrics, then a filter $\mathcal{F}$ on X converges to $x \in X$ iff $\lim_{y\to\mathcal{F}} \rho(y, x) = 0$ for every $\rho \in \mathrm{P}$.

(e)(i) If X is any set, $\mathcal{F}$ is an ultrafilter on X, $(Y, \mathfrak{S})$ is a Hausdorff space, and $h : X \to Y$ is a function such that $h[F]$ is relatively compact in Y for some $F \in \mathcal{F}$, then $\lim_{x\to\mathcal{F}} h(x)$ is defined in Y. **P** Let $K \subseteq Y$ be a compact set including $h[F]$. Then $K \in h[[\mathcal{F}]]$, which is an ultrafilter (2A1N), so $h[[\mathcal{F}]]$ has a limit in Y (2A3R), which is $\lim_{x\to\mathcal{F}} h(x)$. **Q**

(ii) If X is any set, $\mathcal{F}$ is an ultrafilter on X, and $h : X \to \mathbb{R}$ is a function such that $h[F]$ is bounded in $\mathbb{R}$ for some set $F \in \mathcal{F}$, then $\lim_{x\to\mathcal{F}} h(x)$ exists in $\mathbb{R}$. **P** $\overline{h[F]}$ is closed and bounded, therefore compact (2A2F), so $h[F]$ is relatively compact and we can use (i). **Q**

(f) The concepts of $\limsup$, $\liminf$ can be applied to filters. Suppose that $\mathcal{F}$ is a filter on a set X, and that $f : X \to [-\infty, \infty]$ is any function. Then

$$\begin{aligned}\limsup_{x\to\mathcal{F}} f(x) &= \inf\{u : u \in [-\infty, \infty],\ \{x : f(x) \le u\} \in \mathcal{F}\}\\ &= \inf_{F\in\mathcal{F}} \sup_{x\in F} f(x) \in [-\infty, \infty],\end{aligned}$$

$$\begin{aligned}\liminf_{x\to\mathcal{F}} f(x) &= \sup\{u : u \in [-\infty, \infty],\ \{x : f(x) \ge u\} \in \mathcal{F}\}\\ &= \sup_{F\in\mathcal{F}} \inf_{x\in F} f(x).\end{aligned}$$

It is easy to see that, for any two functions f, $g : X \to \mathbb{R}$,

$$\lim_{x\to\mathcal{F}} f(x) = a \quad \text{iff} \quad a = \limsup_{x\to\mathcal{F}} f(x) = \liminf_{x\to\mathcal{F}} f(x),$$

and

$$\limsup_{x\to\mathcal{F}} f(x)+g(x) \le \limsup_{x\to\mathcal{F}} f(x)+\limsup_{x\to\mathcal{F}} g(x),$$

$$\liminf_{x\to\mathcal{F}} f(x)+g(x) \ge \liminf_{x\to\mathcal{F}} f(x)+\liminf_{x\to\mathcal{F}} g(x),$$

$$\liminf_{x\to\mathcal{F}}(-f(x)) = -\limsup_{x\to\mathcal{F}} f(x), \quad \limsup_{x\to\mathcal{F}}(-f(x)) = -\liminf_{x\to\mathcal{F}} f(x),$$

$$\liminf_{x\to\mathcal{F}} cf(x) = c\liminf_{x\to\mathcal{F}} f(x), \quad \limsup_{x\to\mathcal{F}} cf(x) = c\limsup_{x\to\mathcal{F}} f(x)$$

whenever the right-hand-sides are defined in $[-\infty,\infty]$ and $c\ge 0$. So if $a=\lim_{x\to\mathcal{F}} f(x)$ and $b=\lim_{x\to\mathcal{F}}(x)$ exist in $\mathbb{R}$, $\lim_{x\to\mathcal{F}} f(x)+g(x)$ exists and is equal to $a+b$ and $\lim_{x\to\mathcal{F}} cf(x)$ exists and is equal to $c\lim_{x\to\mathcal{F}} f(x)$ for every $c\in\mathbb{R}$.

We also see that if $f:X\to\mathbb{R}$ is such that

for every $\epsilon>0$ there is an $F\in\mathcal{F}$ such that $\sup_{x\in F} f(x)\le\epsilon+\inf_{x\in F} f(x)$,

then $\limsup_{x\to\mathcal{F}} f(x)\le\epsilon+\liminf_{x\to\mathcal{F}} f(x)$ for every $\epsilon>0$, so that $\lim_{x\to\mathcal{F}} f(x)$ is defined in $[-\infty,\infty]$.

(g) Note that the standard limits of real analysis can be represented in the form described here. For instance, $\lim_{n\to\infty}$, $\limsup_{n\to\infty}$, $\liminf_{n\to\infty}$ correspond to $\lim_{n\to\mathcal{F}_{\mathrm{Fr}}}$, $\limsup_{n\to\mathcal{F}_{\mathrm{Fr}}}$, $\liminf_{n\to\mathcal{F}_{\mathrm{Fr}}}$ where $\mathcal{F}_{\mathrm{Fr}}$ is the **Fréchet filter** on $\mathbb{N}$, the filter $\{\mathbb{N}\setminus A : A\subseteq\mathbb{N}$ is finite$\}$ of cofinite subsets of $\mathbb{N}$. Similarly, $\lim_{\delta\downarrow a}$, $\limsup_{\delta\downarrow a}$, $\liminf_{\delta\downarrow a}$ correspond to $\lim_{\delta\to\mathcal{F}}$, $\limsup_{\delta\to\mathcal{F}}$, $\liminf_{\delta\to\mathcal{F}}$ where

$$\mathcal{F}=\{A : A\subseteq\mathbb{R},\ \exists\, h>0 \text{ such that }]a,a+h]\subseteq A\}.$$

2A3T Product topologies We need some brief remarks concerning topologies on product spaces.

(a) Let $(X,\mathfrak{T})$ and $(Y,\mathfrak{S})$ be topological spaces. Let $\mathfrak{U}$ be the set of subsets U of $X\times Y$ such that for every $(x,y)\in U$ there are $G\in\mathfrak{T}$, $H\in\mathfrak{S}$ such that $(x,y)\in G\times H\subseteq U$. Then $\mathfrak{U}$ is a topology on $X\times Y$. **P** (i) $\emptyset\in\mathfrak{U}$ because the condition for membership of $\mathfrak{U}$ is vacuously satisfied. $X\times Y\in\mathfrak{U}$ because $X\in\mathfrak{T}$, $Y\in\mathfrak{S}$ and $(x,y)\in X\times Y\subseteq X\times Y$ for every $(x,y)\in X\times Y$. (ii) If U, $V\in\mathfrak{U}$ and $(x,y)\in U\cap V$, then there are G, $G'\in\mathfrak{T}$, H, $H'\in\mathfrak{S}$ such that

$$(x,y)\in G\times H\subseteq U, \quad (x,y)\in G'\times H'\subseteq V;$$

now $G\cap G'\in\mathfrak{T}$, $H\cap H'\in\mathfrak{S}$ and

$$(x,y)\in(G\cap G')\times(H\cap H')\subseteq U\cap V.$$

As (x,y) is arbitrary, $U\cap V\in\mathfrak{U}$. (iii) If $\mathcal{U}\subseteq\mathfrak{U}$ and $(x,y)\in\bigcup\mathcal{U}$, then there is a $U\in\mathcal{U}$ such that $(x,y)\in U$; now there are $G\in\mathfrak{T}$, $H\in\mathfrak{S}$ such that $(x,y)\in G\times H\subseteq U\subseteq\bigcup\mathcal{U}$. As (x,y) is arbitrary, $\bigcup\mathcal{U}\in\mathfrak{U}$. **Q**

$\mathfrak{U}$ is called the **product topology** on $X\times Y$.

(b) Suppose, in (a), that $\mathfrak{T}$ and $\mathfrak{S}$ are defined by non-empty families P, Θ of pseudometrics in the manner of 2A3F. Then $\mathfrak{U}$ is defined by the family $\Upsilon=\{\tilde\rho:\rho\in\mathrm{P}\}\cup\{\bar\theta:\theta\in\Theta\}$ of pseudometrics on $X\times Y$, where

$$\tilde\rho((x,y),(x',y'))=\rho(x,x'), \quad \bar\theta((x,y),(x',y'))=\theta(y,y')$$

whenever x, $x'\in X$, y, $y'\in Y$, $\rho\in\mathrm{P}$ and $\theta\in\Theta$.

P (i) Of course you should check that every $\tilde\rho$, $\bar\theta$ is a pseudometric on $X\times Y$.

(ii) If $U\in\mathfrak{U}$ and $(x,y)\in U$, then there are $G\in\mathfrak{T}$, $H\in\mathfrak{S}$ such that $(x,y)\in G\times H\subseteq U$. There are $\rho_0,\dots,\rho_m\in\mathrm{P}$, $\theta_0,\dots,\theta_n\in\Theta$, δ, $\delta'>0$ such that (in the language of 2A3Fc) $U(x;\rho_0,\dots,\rho_m;\delta)\subseteq G$, $U(x;\theta_0,\dots,\theta_n;\delta)\subseteq H$. Now

$$U((x,y);\tilde\rho_0,\dots,\tilde\rho_m,\bar\theta_0,\dots,\bar\theta_n;\min(\delta,\delta'))\subseteq U.$$

As (x,y) is arbitrary, U is open for the topology generated by Υ.

(iii) If $U\subseteq X\times Y$ is open for the topology defined by Υ, take any $(x,y)\in U$. Then there are $\upsilon_0,\dots,\upsilon_k\in\Upsilon$ and $\delta>0$ such that $U((x,y);\upsilon_0,\dots,\upsilon_k;\delta)\subseteq U$. Take $\rho_0,\dots,\rho_m\in\mathrm{P}$ and $\theta_0,\dots,\theta_n\in\Theta$ such that $\{\upsilon_0,\dots,\upsilon_k\}\subseteq\{\tilde\rho_0,\dots,\tilde\rho_m,\bar\theta_0,\dots,\bar\theta_n\}$; then $G=U(x;\rho_0,\dots,\rho_m;\delta)\in\mathfrak{T}$ (2A3G), $H=U(y;\theta_0,\dots,\theta_n;\delta)\in\mathfrak{S}$, and

$$(x, y) \in G \times H = U((x, y); \tilde{\rho}_0, \dots, \rho_m, \bar{\theta}_0, \dots, \bar{\theta}_n; \delta) \subseteq U((x, y); \upsilon_0, \dots, \upsilon_k; \delta) \subseteq U.$$

As (x, y) is arbitrary, $U \in \mathfrak{U}$. This completes the proof that $\mathfrak{U}$ is the topology defined by Υ. **Q**

(c) In particular, the product topology on $\mathbb{R}^r \times \mathbb{R}^s$ is the Euclidean topology if we identify $\mathbb{R}^r \times \mathbb{R}^s$ with $\mathbb{R}^{r+s}$. **P** The product topology is defined by the two pseudometrics υ_1, υ_2, where for x, $x' \in \mathbb{R}^r$ and y, $y' \in \mathbb{R}^s$ I write

$$\upsilon_1((x, y), (x', y')) = \|x - x'\|, \quad \upsilon_2((x, y), (x', y')) = \|y - y'\|$$

(2A3F(b-ii)). Similarly, the Euclidean topology on $\mathbb{R}^r \times \mathbb{R}^s \cong \mathbb{R}^{r+s}$ is defined by the metric ρ, where

$$\rho((x, y), (x', y')) = \|(x - y) - (x', y')\| = \sqrt{\|x - x'\|^2 + \|y - y'\|^2}.$$

Now if $(x, y) \in \mathbb{R}^r \times \mathbb{R}^s$ and $\epsilon > 0$, then

$$U((x, y); \rho; \epsilon) \subseteq U((x, y); \upsilon_j; \epsilon)$$

for both j, while

$$U((x, y); \upsilon_1, \upsilon_2; \frac{\epsilon}{\sqrt{2}}) \subseteq U((x, y); \rho; \epsilon).$$

Thus, as remarked in 2A3Ib, each topology is included in the other, and they are the same. **Q**

2A3U Dense sets (a) If X is a topological space, a set $D \subseteq X$ is **dense** in X if $\overline{D} = X$, that is, if every non-empty open set meets D. More generally, if $D \subseteq A \subseteq X$, then D is dense in A if it is dense for the subspace topology of A (2A3C), that is, if $A \subseteq \overline{D}$.

(b) If $\mathfrak{T}$ is defined by a non-empty family P of pseudometrics on X, then $D \subseteq X$ is dense iff

$$U(x; \rho_0, \dots, \rho_n; \delta) \cap D \neq \emptyset$$

whenever $x \in X$, $\rho_0, \dots, \rho_n \in \mathrm{P}$ and $\delta > 0$.

(c) If $(X, \mathfrak{T})$, $(Y, \mathfrak{S})$ are topological spaces, of which Y is Hausdorff (in particular, if (X, ρ) and (Y, θ) are metric spaces), and f, $g : X \to Y$ are continuous functions which agree on some dense subset D of X, then $f = g$. **P?** Suppose, if possible, that there is an $x \in X$ such that $f(x) \neq g(x)$. Then there are open sets G, $H \subseteq Y$ such that $f(x) \in G$, $g(x) \in H$ and $G \cap H = \emptyset$. Now $f^{-1}[G] \cap g^{-1}[H]$ is an open set, containing x and therefore not empty, but it cannot meet D, so $x \notin \overline{D}$ and D is not dense. **XQ**

(d) A topological space is called **separable** if it has a countable dense subset. For instance, $\mathbb{R}^r$ is separable for every $r \geq 1$, since $\mathbb{Q}^r$ is dense.

Version of 4.3.14

2A4 Normed spaces

In Chapter 24 I discuss the spaces L^p, for $1 \leq p \leq \infty$, and describe their most basic properties. These spaces form a cluster of fundamental examples for the general theory of 'normed spaces', the basis of functional analysis. This is not the book from which you should learn that theory, but once again it may save you trouble if I briefly outline those parts of the general theory which are essential if you are to make sense of the ideas here.

2A4A The real and complex fields While the most important parts of the theory, from the point of view of measure theory, are most effectively dealt with in terms of *real* linear spaces, there are many applications in which *complex* linear spaces are essential. I will therefore use the phrase

'U is a linear space over $\begin{smallmatrix}\mathbb{R}\\\mathbb{C}\end{smallmatrix}$'

to mean that U is either a linear space over the field $\mathbb{R}$ or a linear space over the field $\mathbb{C}$; it being understood that in any particular context all linear spaces considered will be over the same field. In the same way, I will write '$\alpha \in \begin{smallmatrix}\mathbb{R}\\\mathbb{C}\end{smallmatrix}$' to mean that α belongs to whichever is the current underlying field.

2A4B Definitions (a) A **normed space** is a linear space U over $\frac{\mathbb{R}}{\mathbb{C}}$ together with a **norm**, that is, a functional $\| \, \| : U \to [0, \infty[$ such that

$\|u + v\| \leq \|u\| + \|v\|$ for all u, $v \in U$,

$\|\alpha u\| = |\alpha|\|u\|$ for $u \in U$, $\alpha \in \frac{\mathbb{R}}{\mathbb{C}}$,

$\|u\| = 0$ only when $u = 0$, the zero vector of U.

(Observe that if $u = 0$ (the zero vector) then $0u = u$ (where this 0 is the zero scalar) so that $\|u\| = |0|\|u\| = 0$.)

(b) If U is a normed space, then we have a metric ρ on U defined by saying that $\rho(u, v) = \|u - v\|$ for u, $v \in U$. **P** $\rho(u, v) \in [0, \infty[$ for all u, v because $\|u\| \in [0, \infty[$ for every u. $\rho(u, v) = \rho(v, u)$ for all u, v because $\|v - u\| = |-1|\|u - v\| = \|u - v\|$ for all u, v. If u, v, $w \in U$ then

$$\rho(u, w) = \|u - w\| = \|(u - v) + (v - w)\| \leq \|u - v\| + \|v - w\| = \rho(u, v) + \rho(v, w).$$

If $\rho(u, v) = 0$ then $\|u - v\| = 0$ so $u - v = 0$ and $u = v$. **Q**

We therefore have a corresponding topology, with open and closed sets, closures, convergent sequences and so on.

(c) If U is a normed space, a set $A \subseteq U$ is **bounded** (for the norm) if $\{\|u\| : u \in A\}$ is bounded in $\mathbb{R}$; that is, there is some $M \geq 0$ such that $\|u\| \leq M$ for every $u \in A$.

2A4C Linear subspaces (a) If U is any normed space and V is a linear subspace of U, then V is also a normed space, if we take the norm of V to be just the restriction to V of the norm of U; the verification is trivial.

(b) If V is a linear subspace of U, so is its closure $\overline{V}$. **P** Take u, $u' \in \overline{V}$ and $\alpha \in \frac{\mathbb{R}}{\mathbb{C}}$. If $\epsilon > 0$, set $\delta = \epsilon/(2 + |\alpha|) > 0$; then there are v, $v' \in V$ such that $\|u - v\| \leq \delta$ and $\|u' - v'\| \leq \delta$. Now $v + v'$, $\alpha v \in V$ and

$$\|(u + u') - (v + v')\| \leq \|u - v\| + \|u' - v'\| \leq \epsilon, \quad \|\alpha u - \alpha v\| \leq |\alpha|\|u - v\| \leq \epsilon.$$

As ϵ is arbitrary, $u + u'$ and αu belong to $\overline{V}$; as u, u' and α are arbitrary, and 0 surely belongs to $V \subseteq \overline{V}$, $\overline{V}$ is a linear subspace of U. **Q**

2A4D Banach spaces (a) If U is a normed space, a sequence $\langle u_n \rangle_{n \in \mathbb{N}}$ in U is **Cauchy** if $\|u_m - u_n\| \to 0$ as m, $n \to \infty$, that is, for every $\epsilon > 0$ there is an $n_0 \in \mathbb{N}$ such that $\|u_m - u_n\| \leq \epsilon$ for all m, $n \geq n_0$.

(b) A normed space U is **complete** if every Cauchy sequence has a limit; a complete normed space is called a **Banach space**.

2A4E It is helpful to know the following result.

Lemma Let U be a normed space such that $\langle u_n \rangle_{n \in \mathbb{N}}$ is convergent (that is, has a limit) in U whenever $\langle u_n \rangle_{n \in \mathbb{N}}$ is a sequence in U such that $\|u_{n+1} - u_n\| \leq 4^{-n}$ for every $n \in \mathbb{N}$. Then U is complete.

proof Let $\langle u_n \rangle_{n \in \mathbb{N}}$ be any Cauchy sequence in U. For each $k \in \mathbb{N}$, let $n_k \in \mathbb{N}$ be such that $\|u_m - u_n\| \leq 4^{-k}$ whenever m, $n \geq n_k$. Set $v_k = u_{n_k}$ for each k. Then $\|v_{k+1} - v_k\| \leq 4^{-k}$ (whether $n_k \leq n_{k+1}$ or $n_{k+1} \leq n_k$). So $\langle v_k \rangle_{k \in \mathbb{N}}$ has a limit $v \in U$. I seek to show that v is the required limit of $\langle u_n \rangle_{n \in \mathbb{N}}$. Given $\epsilon > 0$, let $l \in \mathbb{N}$ be such that $\|v_k - v\| \leq \epsilon$ for every $k \geq l$; let $k \geq l$ be such that $4^{-k} \leq \epsilon$; then if $n \geq n_k$,

$$\|u_n - v\| = \|(u_n - v_k) + (v_k - v)\| \leq \|u_n - v_k\| + \|v_k - v\| \leq \|u_n - u_{n_k}\| + \epsilon \leq 2\epsilon.$$

As ϵ is arbitrary, v is a limit of $\langle u_n \rangle_{n \in \mathbb{N}}$. As $\langle u_n \rangle_{n \in \mathbb{N}}$ is arbitrary, U is complete.

2A4F Bounded linear operators (a) Let U, V be two normed spaces. A linear operator $T : U \to V$ is **bounded** if $\{\|Tu\| : u \in U, \|u\| \leq 1\}$ is bounded. (**Warning!** in this context, we do not ask for the whole set of values $T[U]$ to be bounded; a 'bounded linear operator' need not be what we ordinarily call a 'bounded function'.) Write $\mathsf{B}(U; V)$ for the space of all bounded linear operators from U to V, and for $T \in \mathsf{B}(U; V)$ write $\|T\| = \sup\{\|Tu\| : u \in U, \|u\| \leq 1\}$.

(b) A useful fact: $\|Tu\| \le \|T\|\|u\|$ whenever $T \in \mathsf{B}(U;V)$ and $u \in U$. **P** If $|\alpha| > \|u\|$ then

$$\|\frac{1}{\alpha}u\| = \frac{1}{|\alpha|}\|u\| \le 1,$$

so

$$\|Tu\| = \|\alpha T(\frac{1}{\alpha}u)\| = |\alpha|\|T(\frac{1}{\alpha}u)\| \le |\alpha|\|T\|;$$

as α is arbitrary, $\|Tu\| \le \|T\|\|u\|$. **Q**

(c) A linear operator $T : U \to V$ is bounded iff it is continuous for the norm topologies on U and V. **P** (i) If T is bounded, $u_0 \in U$ and $\epsilon > 0$, then

$$\|Tu - Tu_0\| = \|T(u - u_0)\| \le \|T\|\|u - u_0\| \le \epsilon$$

whenever $\|u - u_0\| \le \frac{\epsilon}{1+\|T\|}$; by 2A3H, T is continuous. (ii) If T is continuous, then there is some $\delta > 0$ such that $\|Tu\| = \|Tu - T0\| \le 1$ whenever $\|u\| = \|u - 0\| \le \delta$. If now $\|u\| \le 1$,

$$\|Tu\| = \frac{1}{\delta}\|T(\delta u)\| \le \frac{1}{\delta},$$

so T is a bounded operator. **Q**

(d) If U, V and W are normed spaces, $S \in \mathsf{B}(U;V)$ and $T \in \mathsf{B}(V;W)$ then $TS \in \mathsf{B}(U;W)$ and $\|TS\| \le \|T\|\|S\|$. **P** I am rather supposing that you are aware, but in any case you will find it easy to check, that $TS : U \to W$ is a linear operator. Now if $u \in U$ and $\|u\| \le 1$,

$$\|TSu\| = \|T(Su)\| \le \|T\|\|Su\| \le \|T\|\|S\|$$

(using (b) for the middle inequality), so TS is bounded and $\|TS\| \le \|T\|\|S\|$. **Q**

2A4G Theorem $\mathsf{B}(U;V)$ is a linear space over $\substack{\mathbb{R}\\\mathbb{C}}$, and $\|\ \|$ is a norm on $\mathsf{B}(U;V)$.

proof As in 2A4Fd, it is easy to check, that if $S : U \to V$ and $T : U \to V$ are linear operators, and $\alpha \in \substack{\mathbb{R}\\\mathbb{C}}$, then we have linear operators $S + T$ and αT from U to V defined by the formulae

$$(S + T)(u) = Su + Tu, \quad (\alpha T)(u) = \alpha(Tu)$$

for every $u \in U$; moreover, that under these definitions of addition and scalar multiplication the space of all linear operators from U to V is a linear space. Now we see that whenever S, $T \in \mathsf{B}(U;V)$, $\alpha \in \substack{\mathbb{R}\\\mathbb{C}}$, $u \in U$ and $\|u\| \le 1$,

$$\|(S + T)(u)\| = \|Su + Tu\| \le \|Su\| + \|Tu\| \le \|S\| + \|T\|,$$

$$\|(\alpha T)u\| = \|\alpha(Tu)\| = |\alpha|\|Tu\| \le |\alpha|\|T\|;$$

so that $S + T$ and αT belong to $\mathsf{B}(U;V)$, with $\|S + T\| \le \|S\| + \|T\|$ and $\|\alpha T\| \le |\alpha|\|T\|$. This shows that $\mathsf{B}(U;V)$ is a linear subspace of the space of all linear operators and is therefore a linear space over $\substack{\mathbb{R}\\\mathbb{C}}$ in its own right. To check that the given formula for $\|T\|$ defines a norm, most of the work has just been done; I suppose I should remark, for the sake of form, that $\|T\| \in [0, \infty[$ for every T; if $\alpha = 0$, then of course $\|\alpha T\| = 0 = |\alpha|\|T\|$; for other α,

$$|\alpha|\|T\| = |\alpha|\|\alpha^{-1}\alpha T\| \le |\alpha||\alpha^{-1}|\|\alpha T\| = \|\alpha T\| \le |\alpha|\|T\|,$$

so $\|\alpha T\| = |\alpha|\|T\|$. Finally, if $\|T\| = 0$ then $\|Tu\| \le \|T\|\|u\| = 0$ for every $u \in U$, so $Tu = 0$ for every u and T is the zero operator (in the space of all linear operators, and therefore in its subspace $\mathsf{B}(U;V)$).

2A4H Dual spaces The most important case of $\mathsf{B}(U;V)$ is when V is the scalar field $\substack{\mathbb{R}\\\mathbb{C}}$ itself (of course we can think of $\substack{\mathbb{R}\\\mathbb{C}}$ as a normed space over itself, writing $\|\alpha\| = |\alpha|$ for each scalar α). In this case we call $\mathsf{B}(U; \substack{\mathbb{R}\\\mathbb{C}})$ the **dual** of U; it is commonly denoted U' or U^*; I use the latter.

2A4I Extensions of bounded operators: Theorem Let U be a normed space and $V \subseteq U$ a dense linear subspace. Let W be a Banach space and $T_0 : V \to W$ a bounded linear operator; then there is a unique bounded linear operator $T : U \to W$ extending T_0, and $\|T\| = \|T_0\|$.

proof (a) For any $u \in U$, there is a sequence $\langle v_n\rangle_{n\in\mathbb{N}}$ in V converging to u. Now

$$\|T_0v_m - T_0v_n\| = \|T_0(v_m - v_n)\| \le \|T_0\|\|v_m - v_n\| \le \|T_0\|(\|v_m - u\| + \|u - v_n\|) \to 0$$

as m, $n \to \infty$, so $\langle T_0v_n\rangle_{n\in\mathbb{N}}$ is Cauchy and $w = \lim_{n\to\infty} T_0v_n$ is defined in W. If $\langle v_n'\rangle_{n\in\mathbb{N}}$ is another sequence in V converging to u, then

$$\begin{aligned}\|w - T_0v_n'\| &\le \|w - T_0v_n\| + \|T_0(v_n - v_n')\|\\ &\le \|w - T_0v_n\| + \|T_0\|(\|v_n - u\| + \|u - v_n'\|) \to 0\end{aligned}$$

as $n \to \infty$, so w is also the limit of $\langle T_0v_n'\rangle_{n\in\mathbb{N}}$.

(b) We may therefore define $T : U \to W$ by setting $Tu = \lim_{n\to\infty} T_0v_n$ whenever $\langle v_n\rangle_{n\in\mathbb{N}}$ is a sequence in V converging to u. If $v \in V$, then we can set $v_n = v$ for every n to see that $Tv = T_0v$; thus T extends T_0. If u, $u' \in U$ and $\alpha \in \genfrac{}{}{0pt}{}{\mathbb{R}}{\mathbb{C}}$, take sequences $\langle v_n\rangle_{n\in\mathbb{N}}$, $\langle v_n'\rangle_{n\in\mathbb{N}}$ in V converging to u, u' respectively; in this case

$$\|(u + u') - (v_n + v_n')\| \le \|u - v_n\| + \|u' - v_n'\| \to 0, \quad \|\alpha u - \alpha u_n\| = |\alpha|\|u - u_n\| \to 0$$

as $n \to \infty$, so that $T(u + u') = \lim_{n\to\infty} T_0(v_n + v_n')$, $T(\alpha u) = \lim_{n\to\infty} T_0(\alpha v_n)$, and

$$\begin{aligned}\|T(u + u') - Tu - Tu'\| &\le \|T(u + u') - T_0(v_n + v_n')\| + \|T_0v_n - Tu\| + \|T_0v_n' - Tu'\|\\ &\to 0,\end{aligned}$$

$$\|T(\alpha u) - \alpha Tu\| \le \|T(\alpha u) - T_0(\alpha v_n)\| + |\alpha|\|T_0v_n - Tu\| \to 0$$

as $n \to \infty$. This means that $\|T(u + u') - Tu - Tu'\| = 0$, $\|T(\alpha u) - \alpha Tu\| = 0$ so $T(u + u') = Tu + Tu'$, $T(\alpha u) = \alpha Tu$; as u, u' and α are arbitrary, T is linear.

(c) For any $u \in U$, let $\langle v_n\rangle_{n\in\mathbb{N}}$ be a sequence in V converging to u. Then

$$\begin{aligned}\|Tu\| &\le \|T_0v_n\| + \|Tu - T_0v_n\| \le \|T_0\|\|v_n\| + \|Tu - T_0v_n\|\\ &\le \|T_0\|(\|u\| + \|v_n - u\|) + \|Tu - T_0v_n\| \to \|T_0\|\|u\|\end{aligned}$$

as $n \to \infty$, so $\|Tu\| \le \|T_0\|\|u\|$. As u is arbitrary, T is bounded and $\|T\| \le \|T_0\|$. Of course $\|T\| \ge \|T_0\|$ just because T extends T_0.

(d) Finally, let $\tilde{T}$ be any other bounded linear operator from U to W extending T. If $u \in U$, there is a sequence $\langle v_n\rangle_{n\in\mathbb{N}}$ in V converging to u; now

$$\|\tilde{T}u - Tu\| \le \|\tilde{T}(u - v_n)\| + \|T(v_n - u)\| \le (\|\tilde{T}\| + \|T\|)\|u - v_n\| \to 0$$

as $n \to \infty$, so $\|\tilde{T}u - Tu\| = 0$ and $\tilde{T}u = Tu$. As u is arbitrary, $\tilde{T} = T$. Thus T is unique.

2A4J Normed algebras (a) A **normed algebra** is a normed space $(U, \|\ \|)$ together with a multiplication, a binary operator $\times$ on U, such that

$$u \times (v \times w) = (u \times v) \times w,$$

$$u \times (v + w) = (u \times v) + (u \times w), \quad (u + v) \times w = (u \times w) + (v \times w),$$

$$(\alpha u) \times v = u \times (\alpha v) = \alpha(u \times v),$$

$$\|u \times v\| \le \|u\|\|v\|$$

for all u, v, $w \in U$ and $\alpha \in \genfrac{}{}{0pt}{}{\mathbb{R}}{\mathbb{C}}$.

(b) A **Banach algebra** is a normed algebra which is a Banach space. A normed algebra U is **commutative** if its multiplication is commutative, that is, $u \times v = v \times u$ for all u, $v \in U$.

***2A4K Definition** A normed space U is **uniformly convex** if for every $\epsilon > 0$ there is a $\delta > 0$ such that $\|u + v\| \le 2 - \delta$ whenever u, $v \in U$, $\|u\| = \|v\| = 1$ and $\|u - v\| \ge \epsilon$.

Version of 13.11.07

2A5 Linear topological spaces

The principal objective of §2A3 is in fact the study of certain topologies on the linear spaces of Chapter 24. I give some fragments of the general theory.

2A5A Linear space topologies Something which is not covered in detail by every introduction to functional analysis is the general concept of 'linear topological space'. The ideas needed for the work of §245 are reasonably briefly expressed.

Definition A **linear topological space** or **topological vector space** over $\genfrac{}{}{0pt}{}{\mathbb{R}}{\mathbb{C}}$ (see 2A4A) is a linear space U over $\genfrac{}{}{0pt}{}{\mathbb{R}}{\mathbb{C}}$ together with a topology $\mathfrak{T}$ such that the maps

$$(u, v) \mapsto u + v : U \times U \to U,$$

$$(\alpha, u) \mapsto \alpha u : \genfrac{}{}{0pt}{}{\mathbb{R}}{\mathbb{C}} \times U \to U$$

are both continuous, where the product spaces $U \times U$ and $\genfrac{}{}{0pt}{}{\mathbb{R}}{\mathbb{C}} \times U$ are given their product topologies (2A3T). Given a linear space U, a topology on U satisfying the conditions above is a **linear space topology**. Note that

$$(u, v) \mapsto u - v = u + (-1)v : U \times U \to U$$

will also be continuous.

2A5B All the linear topological spaces we need turn out to be readily presentable in the following terms.

Proposition Suppose that U is a linear space over $\genfrac{}{}{0pt}{}{\mathbb{R}}{\mathbb{C}}$, and T is a family of functionals $\tau : U \to [0, \infty[$ such that

(i) $\tau(u + v) \le \tau(u) + \tau(v)$ for all u, $v \in U$ and $\tau \in \mathrm{T}$;

(ii) $\tau(\alpha u) \le \tau(u)$ if $u \in U$, $|\alpha| \le 1$ and $\tau \in \mathrm{T}$;

(iii) $\lim_{\alpha \to 0} \tau(\alpha u) = 0$ whenever $u \in U$ and $\tau \in \mathrm{T}$.

For $\tau \in \mathrm{T}$, define $\rho_\tau : U \times U \to [0, \infty[$ by setting $\rho_\tau(u, v) = \tau(u - v)$ for all u, $v \in U$. Then each ρ_τ is a pseudometric on U, and the topology defined by $\mathrm{P} = \{\rho_\tau : \tau \in \mathrm{T}\}$ renders U a linear topological space.

proof (a) It is worth noting immediately that

$$\tau(0) = \lim_{\alpha \to 0} \tau(\alpha 0) = 0$$

for every $\tau \in \mathrm{T}$.

(b) To see that every ρ_τ is a pseudometric, argue as follows.

(i) ρ_τ takes values in $[0, \infty[$ because τ does.

(ii) If u, v, $w \in U$ then

$$\begin{aligned}\rho_\tau(u, w) &= \tau(u - w) = \tau((u - v) + (v - w)) \\ &\le \tau(u - v) + \tau(v - w) = \rho_\tau(u, v) + \rho_\tau(v, w).\end{aligned}$$

(iii) If u, $v \in U$, then

$$\rho(v, u) = \tau(v - u) = \tau(-1(u - v)) \le \tau(u, v) = \rho_\tau(u, v),$$

and similarly $\rho_\tau(u, v) \le \rho_\tau(v, u)$, so the two are equal.

(iv) If $u \in U$ then $\rho_\tau(u, u) = \tau(0) = 0$.

(c) Let $\mathfrak{T}$ be the topology on U defined by $\{\rho_\tau : \tau \in \mathrm{T}\}$ (2A3F).

(i) Addition is continuous because, given $\tau \in \mathrm{T}$, we have

$$\begin{aligned}\rho_\tau(u'+v',u+v) &= \tau((u'+v')-(u+v))\\ &\le \tau(u'-u)+\tau(v'-v) \le \rho_\tau(u',u)+\rho_\tau(v',v)\end{aligned}$$

for all u, v, u', $v' \in U$. This means that, given $\epsilon > 0$ and $(u,v) \in U \times U$, we shall have

$$\rho_\tau(u'+v',u+v) \le \epsilon \text{ whenever } (u',v') \in U((u,v);\tilde\rho_\tau,\bar\rho_\tau;\tfrac{\epsilon}{2}),$$

using the language of 2A3Tb. Because $\tilde\rho_\tau$, $\bar\rho_\tau$ are two of the pseudometrics defining the product topology of $U \times U$ (2A3Tb), $(u,v) \mapsto u+v$ is continuous, by the criterion of 2A3H.

(ii) Scalar multiplication is continuous because if $u \in U$ and $n \in \mathbb{N}$ then $\tau(nu) \le n\tau(u)$ for every $\tau \in \mathrm{T}$ (induce on n). Consequently, if $\tau \in \mathrm{T}$,

$$\tau(\alpha u) \le n\tau(\frac{\alpha}{n}u) \le n\tau(u)$$

whenever $|\alpha| < n \in \mathbb{N}$ and $\tau \in \mathrm{T}$. Now, given $(\alpha,u) \in {\mathbb{R} \atop \mathbb{C}} \times U$ and $\epsilon > 0$, take $n > |\alpha|$ and $\delta > 0$ such that $\delta \le \min(n-|\alpha|, \frac{\epsilon}{2n})$ and $\tau(\gamma u) \le \frac{\epsilon}{2}$ whenever $|\gamma| \le \delta$; then

$$\begin{aligned}\rho_\tau(\alpha'u',\alpha u) &= \tau(\alpha'u'-\alpha u) \le \tau(\alpha'(u'-u))+\tau((\alpha'-\alpha)u)\\ &\le n\tau(u'-u)+\tau((\alpha'-\alpha)u)\end{aligned}$$

whenever $u' \in U$ and $\alpha' \in {\mathbb{R} \atop \mathbb{C}}$ and $|\alpha'| < n \in \mathbb{N}$. Accordingly, setting $\theta(\alpha',\alpha) = |\alpha'-\alpha|$ for α', $\alpha \in {\mathbb{R} \atop \mathbb{C}}$,

$$\rho_\tau(\alpha'u',\alpha u) \le n\delta + \frac{\epsilon}{2} \le \epsilon$$

whenever

$$(\alpha',u') \in U((\alpha,u);\tilde\theta,\bar\rho_\tau;\delta).$$

Because $\tilde\theta$ and $\bar\rho_\tau$ are among the pseudometrics defining the topology of ${\mathbb{R} \atop \mathbb{C}} \times U$, the map $(\alpha,u) \mapsto \alpha u$ satisfies the criterion of 2A3H and is continuous.

Thus $\mathfrak{T}$ is a linear space topology on U.

Remark Functionals satisfying the conditions (i)-(iii) above are called **F-seminorms**; an F-seminorm τ such that $\tau(u) \ne 0$ for every non-zero u is an **F-norm**.

***2A5C** We do not need it for Chapter 24, but the following is worth knowing.

Theorem Let U be a linear space and $\mathfrak{T}$ a linear space topology on U.

(a) There is a family T of F-seminorms defining $\mathfrak{T}$ as in 2A5B.

(b) If $\mathfrak{T}$ is metrizable, we can take T to consist of a single functional.

proof (a) Kelley & Namioka 76, p. 50.

(b) Köthe 69, §15.11.

2A5D Definition Let U be a linear space over ${\mathbb{R} \atop \mathbb{C}}$. Then a **seminorm** on U is a functional $\tau : U \to [0,\infty[$ such that

(i) $\tau(u+v) \le \tau(u)+\tau(v)$ for all u, $v \in U$;

(ii) $\tau(\alpha u) = |\alpha|\tau(u)$ if $u \in U$, $\alpha \in {\mathbb{R} \atop \mathbb{C}}$.

Observe that a norm is always a seminorm, and that a seminorm is always an F-seminorm. In particular, the association of a metric with a norm (2A4Bb) is a special case of 2A5B.

2A5E Convex sets (a) Let U be a linear space over ${\mathbb{R} \atop \mathbb{C}}$. A subset C of U is **convex** if $\alpha u+(1-\alpha)v \in C$ whenever u, $v \in C$ and $\alpha \in [0,1]$. The intersection of any family of convex sets is convex, so for every set $A \subseteq U$ there is a smallest convex set including A; this is just the set of vectors expressible as $\sum_{i=0}^{n}\alpha_i u_i$

where $u_0, \dots, u_n \in A$, $\alpha_0, \dots, \alpha_n \in [0,1]$ and $\sum_{i=0}^{n} \alpha_i = 1$ (Bourbaki 87, II.2.3); it is the **convex hull** of A. If C, $C' \subseteq U$ are convex, and $\alpha \in {}^{\mathbb{R}}_{\mathbb{C}}$, then αC and $C + C'$ are convex. If $C \subseteq U$ is convex, V is another linear space over ${}^{\mathbb{R}}_{\mathbb{C}}$, and $T : U \to V$ is a linear operator, then $T[C] \subseteq V$ is convex.

(b) If U is a linear topological space, the closure of any convex set is convex (Bourbaki 87, II.2.6). It follows that, for any $A \subseteq U$, the closure of the convex hull of A is the smallest closed convex set including A; this is the **closed convex hull** of A.

(c) I note for future reference that in a linear topological space, the closure of any linear subspace is a linear subspace. (Bourbaki 87, I.1.3; Köthe 69, §15.2. Compare 2A4Cb.)

2A5F Completeness in linear topological spaces In normed spaces, completeness can be described in terms of Cauchy sequences (2A4D). In general linear topological spaces this is inadequate. The true theory of 'completeness' demands the concept of 'uniform space' (see §3A4 in the next volume, or Kelley 55, chap. 6; Engelking 89, §8.1: Bourbaki 66, chap. II); I shall not describe this here, but will give a version adapted to linear spaces. I mention this only because you will I hope some day come to the general theory (in Volume 3 of this treatise, if not before), and you should be aware that the special case described here gives a misleading emphasis at some points.

Definitions Let U be a linear space over ${}^{\mathbb{R}}_{\mathbb{C}}$, and $\mathfrak{T}$ a linear space topology on U. A filter $\mathcal{F}$ on U is **Cauchy** if for every open set G in U containing 0 there is an $F \in \mathcal{F}$ such that $F - F = \{u - v : u, v \in F\}$ is included in G. U is **complete** if every Cauchy filter on U is convergent.

2A5G Cauchy filters have a simple description when a linear space topology is defined by the method of 2A5B.

Lemma Let U be a linear space over ${}^{\mathbb{R}}_{\mathbb{C}}$, and let T be a family of F-seminorms defining a linear space topology on U, as in 2A5B. Then a filter $\mathcal{F}$ on U is Cauchy iff for every $\tau \in \mathrm{T}$ and $\epsilon > 0$ there is an $F \in \mathcal{F}$ such that $\tau(u - v) \le \epsilon$ for all u, $v \in F$.

proof (a) Suppose that $\mathcal{F}$ is Cauchy, $\tau \in \mathrm{T}$ and $\epsilon > 0$. Then $G = U(0; \rho_\tau; \epsilon)$ is open (using the language of 2A3F-2A3G), so there is an $F \in \mathcal{F}$ such that $F - F \subseteq G$; but this just means that $\tau(u - v) < \epsilon$ for all u, $v \in F$.

(b) Suppose that $\mathcal{F}$ satisfies the criterion, and that G is an open set containing 0. Then there are $\tau_0, \dots, \tau_n \in \mathrm{T}$ and $\epsilon > 0$ such that $U(0; \rho_{\tau_0}, \dots, \rho_{\tau_n}; \epsilon) \subseteq G$. For each $i \le n$ there is an $F_i \in \mathcal{F}$ such that $\tau_i(u, v) < \frac{\epsilon}{2}$ for all u, $v \in F_i$; now $F = \bigcap_{i \le n} F_i \in \mathcal{F}$ and $u - v \in G$ for all u, $v \in F$.

2A5H Normed spaces and sequential completeness I had better point out that for normed spaces the definition of 2A5F agrees with that of 2A4D.

Proposition Let $(U, \| \, \|)$ be a normed space over ${}^{\mathbb{R}}_{\mathbb{C}}$, and let $\mathfrak{T}$ be the linear space topology on U defined by the method of 2A5B from the set $\mathrm{T} = \{\| \, \|\}$. Then U is complete in the sense of 2A5F iff it is complete in the sense of 2A4D.

proof (a) Suppose first that U is complete in the sense of 2A5F. Let $\langle u_n \rangle_{n \in \mathbb{N}}$ be a sequence in U which is Cauchy in the sense of 2A4Da. Set

$$\mathcal{F} = \{F : F \subseteq U, \{n : u_n \notin F\} \text{ is finite}\}.$$

Then it is easy to check that $\mathcal{F}$ is a filter on U, the image of the Fréchet filter under the map $n \mapsto u_n : \mathbb{N} \to U$. If $\epsilon > 0$, take $m \in \mathbb{N}$ such that $\|u_j - u_k\| \le \epsilon$ whenever j, $k \ge m$; then $F = \{u_j : j \ge m\}$ belongs to $\mathcal{F}$, and $\|u - v\| \le \epsilon$ for all u, $v \in F$. So $\mathcal{F}$ is Cauchy in the sense of 2A5F, and has a limit u say. Now, for any $\epsilon > 0$, the set $\{v : \|v - u\| < \epsilon\} = U(u; \rho_{\| \, \|}; \epsilon)$ is an open set containing u, so belongs to $\mathcal{F}$, and $\{n : \|u_n - u\| \ge \epsilon\}$ is finite, that is, there is an $m \in \mathbb{N}$ such that $\|u_m - u\| < \epsilon$ whenever $n \ge m$. As ϵ is arbitrary, $u = \lim_{n \to \infty} u_n$ in the sense of 2A3M. As $\langle u_n \rangle_{n \in \mathbb{N}}$ is arbitrary, U is complete in the sense of 2A4D.

(b) Now suppose that U is complete in the sense of 2A4D. Let $\mathcal{F}$ be a Cauchy filter on U. For each $n \in \mathbb{N}$, choose a set $F_n \in \mathcal{F}$ such that $\|u - v\| \le 2^{-n}$ for all u, $v \in F_n$. For each $n \in \mathbb{N}$, $F'_n = \bigcap_{i \le n} F_i$

belongs to $\mathcal{F}$, so is not empty; choose $u_n \in F'_n$. If $m \in \mathbb{N}$ and $j, k \geq m$, then both u_j and u_k belong to F_m, so $\|u_j - u_k\| \leq 2^{-m}$; thus $\langle u_n \rangle_{n\in\mathbb{N}}$ is a Cauchy sequence in the sense of 2A4Da, and has a limit u say. Now take any $\epsilon > 0$ and $m \in \mathbb{N}$ such that $2^{-m+1} \leq \epsilon$. There is surely a $k \geq m$ such that $\|u_k - u\| \leq 2^{-m}$; now $u_k \in F_m$, so

$$F_m \subseteq \{v : \|v - u_k\| \leq 2^{-m}\} \subseteq \{v : \|v - u\| \leq 2^{-m+1}\} \subseteq \{v : \rho_{\|\,\|}(v, u) \leq \epsilon\},$$

and $\{v : \rho_{\|\,\|}(v, u) \leq \epsilon\} \in \mathcal{F}$. As ϵ is arbitrary, $\mathcal{F}$ converges to u, by 2A3Sd. As $\mathcal{F}$ is arbitrary, U is complete.

(c) Thus the two definitions coincide, provided at least that we allow the countably many simultaneous choices of the u_n in part (b) of the proof.

2A5I Weak topologies I come now to brief notes on 'weak topologies' on normed spaces; from the point of view of this volume, these are in fact the primary examples of linear space topologies. Let U be a normed linear space over ${}^{\mathbb{R}}_{\mathbb{C}}$.

(a) Write U^* for its dual $\mathrm{B}(U; {}^{\mathbb{R}}_{\mathbb{C}})$ (2A4H). If $h \in U^*$, then $|h| : U \to [0, \infty[$ is a seminorm, so $\mathrm{T} = \{|h| : h \in U^*\}$ defines a linear space topology on U, by 2A5B; this is called the **weak topology** of U.

(b) A filter $\mathcal{F}$ on U converges to $u \in U$ for the weak topology of U iff $\lim_{v\to\mathcal{F}} \rho_{|h|}(v, u) = 0$ for every $h \in U^*$ (2A3Sd), that is, iff $\lim_{v\to\mathcal{F}} |h(v - u)| = 0$ for every $h \in U^*$, that is, iff $\lim_{v\to\mathcal{F}} h(v) = h(u)$ for every $h \in U^*$.

(c) A set $C \subseteq U$ is called **weakly compact** if it is compact for the weak topology of U. So (subject to the axiom of choice) a set $C \subseteq U$ is weakly compact iff for every ultrafilter $\mathcal{F}$ on U containing C there is a $u \in C$ such that $\lim_{v\to\mathcal{F}} h(v) = h(u)$ for every $h \in U^*$ (put 2A3R together with (b) above).

(d) A subset A of U is called **relatively weakly compact** if it is a subset of some weakly compact subset of U.

(e) If $h \in U^*$, then $h : U \to {}^{\mathbb{R}}_{\mathbb{C}}$ is continuous for the weak topology on U and the usual topology of ${}^{\mathbb{R}}_{\mathbb{C}}$; this is obvious if we apply the criterion of 2A3H. So if $A \subseteq U$ is relatively weakly compact, $h[A]$ must be bounded in ${}^{\mathbb{R}}_{\mathbb{C}}$. **P** Let $C \supseteq A$ be a weakly compact set. Then $h[C]$ is compact in ${}^{\mathbb{R}}_{\mathbb{C}}$, by 2A3Nb, so is bounded, by 2A2F (noting that if the underlying field is $\mathbb{C}$, then it can be identified, as metric space, with $\mathbb{R}^2$). Accordingly $h[A]$ also is bounded. **Q**

(f) If V is another normed space and $T : U \to V$ is a bounded linear operator, then T is continuous for the respective weak topologies. **P** If $h \in V^*$ then the composition hT belongs to U^*. Now, for any $u, v \in U$,

$$\rho_{|h|}(Tu, Tv) = |h(Tu - Tv)| = |hT(u - v)| = \rho_{|hT|}(u, v),$$

taking $\rho_{|h|}$, $\rho_{|hT|}$ to be the pseudometrics on V, U respectively defined by the formula of 2A5B. By 2A3H, T is continuous. **Q**

(g) Corresponding to the weak topology on a normed space U, we have the **weak*** or **w*-**topology on its dual U^*, defined by the set $\mathrm{T} = \{|\hat{u}| : u \in U\}$, where I write $\hat{u}(f) = f(u)$ for every $f \in U^*$, $u \in U$. As in (a), this is a linear space topology on U^*. (It is essential to distinguish between the 'weak*' topology and the 'weak' topology on U^*. The former depends only on the action of U on U^*, the latter on the action of $U^{**} = (U^*)^*$. You will have no difficulty in checking that $\hat{u} \in U^{**}$ for every $u \in U$, but the point is that there may be members of U^{**} not representable in this way, leading to open sets for the weak topology which are not open for the weak* topology.)

***2A5J Angelic spaces** I do not rely on the following ideas, but they may throw light on some results in §§246-247. First, a topological space X is **regular** if whenever $G \subseteq X$ is open and $x \in G$ then there is an open set H such that $x \in H \subseteq \overline{H} \subseteq G$. Next, a regular Hausdorff space X is **angelic** if whenever $A \subseteq X$ is such that every sequence in A has a cluster point in X, then $\overline{A}$ is compact and every point of $\overline{A}$ is the limit of a sequence in A. What this means is that compactness in X, and the topologies of compact subsets of X, can be effectively described in terms of sequences. Now the theorem (due to Eberlein and Šmulian) is that any normed space is angelic in its weak topology. (462D in Volume 4; Köthe 69, §24; Dunford & Schwartz 57, V.6.1.) In particular, this is true of L^1 spaces, which makes it less surprising that there should be criteria for weak compactness in L^1 spaces which deal only with sequences.

Version of 10.11.14

2A6 Factorization of matrices

I spend a couple of pages on the linear algebra of $\mathbb{R}^r$ required for Chapter 26. I give only one proof, because this is material which can be found in any textbook of elementary linear algebra; but I think it may be helpful to run through the basic ideas in the language which I use for this treatise.

2A6A Determinants We need to know the following things about determinants.

(i) Every $r \times r$ real matrix T has a real determinant $\det T$.

(ii) For any $r \times r$ matrices S and T, $\det ST = \det S \det T$.

(iii) If T is a diagonal matrix, its determinant is just the product of its diagonal entries.

(iv) For any $r \times r$ matrix T, $\det T^\top = \det T$, where $T^\top$ is the transpose of T.

(v) $\det T$ is a continuous function of the coefficients of T.

There are so many routes through this topic that I avoid even a definition of 'determinant'; I invite you to check your memory, or your favourite text, to confirm that you are indeed happy with the facts above.

2A6B Orthonormal families For $x = (\xi_1, \dots, \xi_r)$, $y = (\eta_1, \dots, \eta_r) \in \mathbb{R}^r$, write $x . y = \sum_{i=1}^r \xi_i \eta_i$; of course $\|x\|$, as defined in 1A2A, is $\sqrt{x . x}$. Recall that $x_1, \dots, x_k$ are **orthonormal** if $x_i . x_j = 0$ for $i \ne j$, 1 for $i = j$. The results we need here are:

(i) If $x_1, \dots, x_k$ are orthonormal vectors in $\mathbb{R}^r$, where $k < r$, then there are vectors $x_{k+1}, \dots, x_r$ in $\mathbb{R}^r$ such that $x_1, \dots, x_r$ are orthonormal.

(ii) An $r \times r$ matrix P is **orthogonal** if $P^\top P$ is the identity matrix; equivalently, if the columns of P are orthonormal.

(iii) For an orthogonal matrix P, $\det P$ must be ± 1 (put (ii)-(iv) of 2A6A together).

(iv) If P is orthogonal, then $Px . Py = P^\top Px . y = x . y$ for all x, $y \in \mathbb{R}^r$.

(v) If P is orthogonal, so is $P^\top = P^{-1}$.

(vi) If P and Q are orthogonal, so is PQ.

2A6C I now give a proposition which is not always included in elementary presentations. Of course there are many approaches to this; I offer a direct one.

Proposition Let T be any real $r \times r$ matrix. Then T is expressible as PDQ where P and Q are orthogonal matrices and D is a diagonal matrix with non-negative coefficients.

proof I induce on r.

(a) If $r = 1$, then $T = (\tau_{11})$. Set $D = (|\tau_{11}|)$, $P = (1)$ and $Q = (1)$ if $\tau_{11} \ge 0$, (-1) otherwise.

(b)(i) For the inductive step to $r + 1 \ge 2$, consider the unit ball $B = \{x : x \in \mathbb{R}^{r+1}, \|x\| \le 1\}$. This is a closed bounded set in $\mathbb{R}^{r+1}$, so is compact (2A2F). The maps $x \mapsto Tx : \mathbb{R}^{r+1} \to \mathbb{R}^{r+1}$ and $x \mapsto \|x\| : \mathbb{R}^{r+1} \to \mathbb{R}$ are continuous, so the function $x \mapsto \|Tx\| : B \to \mathbb{R}$ is bounded and attains its bounds (2A2G), and there is a $u \in B$ such that $\|Tu\| \ge \|Tx\|$ for every $x \in B$. Observe that $\|Tu\|$ must be the norm $\|T\|$ of T as defined in 262H. Set $\delta = \|T\| = \|Tu\|$. If $\delta = 0$, then T must be the zero matrix, and the result is trivial; so let us suppose that $\delta > 0$. In this case $\|u\|$ must be exactly 1, since otherwise we should have $u = \|u\|u'$ where $\|u'\| = 1$ and $\|Tu'\| > \|Tu\|$.

(ii) If $x \in \mathbb{R}^{r+1}$ and $x\,.\,u = 0$, then $Tx\,.\,Tu = 0$. **P?** If not, set $\gamma = Tx\,.\,Tu \neq 0$. Consider $y = u + \eta\gamma x$ for small $\eta > 0$. We have

$$\|y\|^2 = y\,.\,y = u\,.\,u + 2\eta\gamma u\,.\,x + \eta^2\gamma^2 x\,.\,x = \|u\|^2 + \eta^2\gamma^2\|x\|^2 = 1 + \eta^2\gamma^2\|x\|^2,$$

while

$$\|Ty\|^2 = Ty\,.\,Ty = Tu\,.\,Tu + 2\eta\gamma Tu\,.\,Tx + \eta^2\gamma^2 Tx\,.\,Tx = \delta^2 + 2\eta\gamma^2 + \eta^2\gamma^2\|Tx\|^2.$$

But also $\|Ty\|^2 \le \delta^2\|y\|^2$ (2A4Fb), so

$$\delta^2 + 2\eta\gamma^2 + \eta^2\gamma^2\|Tx\|^2 \le \delta^2(1 + \eta^2\gamma^2\|x\|^2)$$

and

$$2\eta\gamma^2 \le \delta^2\eta^2\gamma^2\|x\|^2 - \eta^2\gamma^2\|Tx\|^2,$$

that is,

$$2 \le \eta(\delta^2\|x\|^2 - \|Tx\|^2).$$

But this surely cannot be true for all $\eta > 0$, so we have a contradiction. **XQ**

(iii) Set $v = \delta^{-1}Tu$, so that $\|v\| = 1$. Let $u_1, \dots, u_{r+1}$ be orthonormal vectors such that $u_{r+1} = u$, and let Q_0 be the orthogonal $(r+1) \times (r+1)$ matrix with columns $u_1, \dots, u_{r+1}$; then, writing $e_1, \dots, e_{r+1}$ for the standard orthonormal basis of $\mathbb{R}^{r+1}$, we have $Q_0 e_i = u_i$ for each i, and $Q_0 e_{r+1} = u$. Similarly, there is an orthogonal matrix P_0 such that $P_0 e_{r+1} = v$.

Set $T_1 = P_0^{-1} T Q_0$. Then

$$T_1 e_{r+1} = P_0^{-1} Tu = \delta P_0^{-1} v = \delta e_{r+1},$$

while if $x\,.\,e_{r+1} = 0$ then $Q_0 x\,.\,u = 0$ (2A6B(iv)), so that

$$T_1 x\,.\,e_{r+1} = P_0 T_1 x\,.\,P_0 e_{r+1} = TQ_0 x\,.\,v = 0,$$

by (ii). This means that T_1 must be of the form

$$\begin{pmatrix} S & 0 \\ 0 & \delta \end{pmatrix},$$

where S is an $r \times r$ matrix.

(iv) By the inductive hypothesis, S is expressible as $\tilde{P}\tilde{D}\tilde{Q}$, where $\tilde{P}$ and $\tilde{Q}$ are orthogonal $r \times r$ matrices and $\tilde{D}$ is a diagonal $r \times r$ matrix with non-negative coefficients. Set

$$P_1 = \begin{pmatrix} \tilde{P} & 0 \\ 0 & 1 \end{pmatrix}, \quad Q_1 = \begin{pmatrix} \tilde{Q} & 0 \\ 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} \tilde{D} & 0 \\ 0 & \delta \end{pmatrix}.$$

Then P_1 and Q_1 are orthogonal and D is diagonal, with non-negative coefficients, and $P_1 D Q_1 = T_1$. Now set

$$P = P_0 P_1, \quad Q = Q_1 Q_0^{-1},$$

so that P and Q are orthogonal (2A6B(v)-(vi)) and

$$PDQ = P_0 P_1 D Q_1 Q_0^{-1} = P_0 T_1 Q_0^{-1} = T.$$

Thus the induction proceeds.

Version of 17.12.12

References for Volume 2

Alexits G. [78] (ed.) *Fourier Analysis and Approximation Theory.* North-Holland, 1978 (Colloq. Math. Soc. Janos Bolyai 19).

Antonov N.Yu. [96] 'Convergence of Fourier series', East J. Approx. 7 (1996) 187-196. [§286 *notes*.]

Arias de Reyna J. [02] *Pointwise Convergence of Fourier Series.* Springer, 2002 (Lecture Notes in Mathematics 1785). [§286 *notes*.]

Baker R. [04] '"Lebesgue measure" on $\mathbb{R}^\infty$, II', Proc. Amer. Math. Soc. 132 (2004) 2577-2591. [254Yb.]

Bergelson V., March P. & Rosenblatt J. [96] (eds.) *Convergence in Ergodic Theory and Probability.* de Gruyter, 1996.

Bogachev V.I. [07] *Measure theory.* Springer, 2007.

du Bois-Reymond P. [1876] 'Untersuchungen über die Convergenz und Divergenz der Fouriersche Darstellungformeln', Abh. Akad. München 12 (1876) 1-103. [§282 *notes*.]

Bourbaki N. [66] *General Topology.* Hermann/Addison-Wesley, 1968. [2A5F.]

Bourbaki N. [87] *Topological Vector Spaces.* Springer, 1987. [2A5E.]

Carleson L. [66] 'On convergence and growth of partial sums of Fourier series', Acta Math. 116 (1966) 135-157. [§282 *notes*, §286 *intro.*, §286 *notes*.]

Clarkson J.A. [1936] 'Uniformly convex spaces', Trans. Amer. Math. Soc. 40 (1936) 396-414. [244O.]

Defant A. & Floret K. [93] *Tensor Norms and Operator Ideals*, North-Holland, 1993. [§253 *notes*.]

Doob J.L. [53] *Stochastic Processes.* Wiley, 1953.

Dudley R.M. [89] *Real Analysis and Probability.* Wadsworth & Brooks/Cole, 1989. [§282 *notes*.]

Dunford N. & Schwartz J.T. [57] *Linear Operators I.* Wiley, 1957 (reprinted 1988). [§244 *notes*, 2A5J.]

Enderton H.B. [77] *Elements of Set Theory.* Academic, 1977. [§2A1.]

Engelking R. [89] *General Topology.* Heldermann, 1989 (Sigma Series in Pure Mathematics 6). [2A5F.]

Etemadi N. [96] 'On convergence of partial sums of independent random variables', pp. 137-144 in Bergelson March & Rosenblatt 96. [272V.]

Evans L.C. & Gariepy R.F. [92] *Measure Theory and Fine Properties of Functions.* CRC Press, 1992. [263Ec, §265 *notes*.]

Federer H. [69] *Geometric Measure Theory.* Springer, 1969 (reprinted 1996). [262C, 263Ec, §264 *notes*, §265 *notes*, §266 *notes*.]

Feller W. [66] *An Introduction to Probability Theory and its Applications,* vol. II. Wiley, 1966. [Chap. 27 *intro.*, 274H, 275Xc, 285N.]

Fremlin D.H. [74] *Topological Riesz Spaces and Measure Theory.* Cambridge U.P., 1974. [§232 *notes*, 241F, §244 *notes*, §245 *notes*, §247 *notes*.]

Fremlin D.H. [93] 'Real-valued-measurable cardinals', pp. 151-304 in Judah 93. [232H.]

Haimo D.T. [67] (ed.) *Orthogonal Expansions and their Continuous Analogues.* Southern Illinois University Press, 1967.

Hall P. [82] *Rates of Convergence in the Central Limit Theorem.* Pitman, 1982. [274H.]

Halmos P.R. [50] *Measure Theory.* Van Nostrand, 1950. [§251 *notes*, §252 *notes*, 255Yn.]

Halmos P.R. [60] *Naive Set Theory.* Van Nostrand, 1960. [§2A1.]

Hanner O. [56] 'On the uniform convexity of L^p and l^p', Arkiv för Matematik 3 (1956) 239-244. [244O.]

Henle J.M. [86] *An Outline of Set Theory.* Springer, 1986. [§2A1.]

Hoeffding W. [63] 'Probability inequalities for sums of bounded random variables', J. Amer. Statistical Association 58 (1963) 13-30. [272W.]

Hunt R.A. [67] 'On the convergence of Fourier series', pp. 235-255 in Haimo 67. [§286 *notes*.]

Jorsbøe O.G. & Mejlbro L. [82] *The Carleson-Hunt Theorem on Fourier Series.* Springer, 1982 (Lecture Notes in Mathematics 911). [§286 *notes*.]

Judah H. [93] (ed.) Proceedings of the Bar-Ilan Conference on Set Theory and the Reals, 1991. Amer. Math. Soc. (Israel Mathematical Conference Proceedings 6), 1993.

Kelley J.L. [55] *General Topology.* Van Nostrand, 1955. [2A5F.]

Kelley J.L. & Namioka I. [76] *Linear Topological Spaces.* Springer, 1976. [2A5C.]

Kirszbraun M.D. [1934] 'Über die zusammenziehenden und Lipschitzian Transformationen', Fund. Math. 22 (1934) 77-108. [262C.]

Kolmogorov A.N. [1926] 'Une série de Fourier-Lebesgue divergente partout', C. R. Acad. Sci. Paris 183 (1926) 1327-1328. [§282 *notes*.]

Komlós J. [67] 'A generalization of a problem of Steinhaus', Acta Math. Acad. Sci. Hung. 18 (1967) 217-229. [276H.]

Körner T.W. [88] *Fourier Analysis.* Cambridge U.P., 1988. [§282 *notes*.]

Köthe G. [69] *Topological Vector Spaces I.* Springer, 1969. [2A5C, 2A5E, 2A5J.]

Krivine J.-L. [71] *Introduction to Axiomatic Set Theory.* D. Reidel, 1971. [§2A1.]

Lacey M.T. [05] *Carleson's Theorem: Proof, Complements, Variations.* `http://arxiv.org/pdf/math/0307008v4.pdf`. [§286 *notes*.]

Lacey M.T. & Thiele C.M. [00] 'A proof of boundedness of the Carleson operator', Math. Research Letters 7 (2000) 1-10. [§286 *intro.*, 286H.]

Lebesgue H. [72] *Oeuvres Scientifiques.* L'Enseignement Mathématique, Institut de Mathématiques, Univ. de Genève, 1972. [Chap. 27 *intro.*]

Liapounoff A. [1901] 'Nouvelle forme du théorème sur la limite de probabilité', Mém. Acad. Imp. Sci. St-Pétersbourg 12(5) (1901) 1-24. [274Xh.]

Lighthill M.J. [59] *Introduction to Fourier Analysis and Generalised Functions.* Cambridge U.P., 1959. [§284 *notes*.]

Lindeberg J.W. [1922] 'Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung', Math. Zeitschrift 15 (1922) 211-225. [274H, §274 *notes*.]

Lipschutz S. [64] *Set Theory and Related Topics.* McGraw-Hill, 1964 (Schaum's Outline Series). [§2A1.]

Loève M. [77] *Probability Theory I.* Springer, 1977. [Chap. 27 *intro.*, 274H.]

Luxemburg W.A.J. & Zaanen A.C. [71] *Riesz Spaces I.* North-Holland, 1971. [241F.]

Mozzochi C.J. [71] *On the Pointwise Convergence of Fourier Series.* Springer, 1971 (Lecture Notes in Mathematics 199). [§286 *notes*.]

Naor A. [04] 'Proof of the uniform convexity lemma', `http://www.cims.nyu.edu/~naor/homepage files/inequality.pdf`, 26.2.04. [244O.]

Rényi A. [70] *Probability Theory.* North-Holland, 1970. [274H.]

Roitman J. [90] *An Introduction to Set Theory.* Wiley, 1990. [§2A1.]

Roselli P. & Willem M. [02] 'A convexity inequality', Amer. Math. Monthly 109 (2002) 64-70. [244Ym.]

Saks S. [1924] 'Sur les nombres dérivés des fonctions', Fund

Schipp F. [78] 'On Carleson's method', pp. 679-695 in Alexits 78. [§286 *notes*.]

Semadeni Z. [71] *Banach spaces of continuous functions I.* Polish Scientific Publishers, 1971. [§253 *notes*.]

Shiryayev A.N. [84] *Probability.* Springer, 1984. [285N.]

Steele J.M. [86] 'An Efron-Stein inequality of nonsymmetric statistics', Annals of Statistics 14 (1986) 753-758. [274Ya.]

Zaanen A.C. [83] *Riesz Spaces II.* North-Holland, 1983. [241F.]

Zygmund A. [59] *Trigonometric Series.* Cambridge U.P, 1959. [§244 *notes*, §282 *notes*, 284Xk.]