The Tikhonov and Alaoglu–Bourbaki Theorems

The central result of this chapter is the Alaoglu–Bourbaki theorem: Polars of neighbourhoods of zero in a locally convex space E are $\sigma(E', E)$-compact subsets of E'. As a consequence in a dual pair $\langle E, F\rangle$ one concludes that, for a locally convex topology τ on E with $(E, \tau)' = F$, one always has $\sigma(E, F) \subseteq \tau \subseteq \mu(E, F)$, where $\mu(E, F)$ is the Mackey topology on E, corresponding to the collection of absolutely convex $\sigma(F, E)$-compact subsets of F. As a prerequisite we show Tikhonov's theorem, and as a prerequisite to the proof of Tikhonov's theorem we introduce filters describing convergence and continuity of mappings in topological spaces.

Theorem 4.1 (Tikhonov)
Let $(X_\iota)_{\iota\in I}$ be a family of compact topological spaces. Then the product $\prod_{\iota\in I} X_\iota$ is compact.

We will prove this theorem here, even if it is rather part of general topology. However, the proof gives us the opportunity to introduce the notion of filters, which we will need anyway in the further treatment.

We recall that a topological space (X, τ) is called **compact** if every open covering of X (i.e., every collection $\mathcal{S} \subseteq \tau$ satisfying $\bigcup \mathcal{S} = X$) contains a finite subcovering (i.e., a finite collection $\mathcal{F} \subseteq \mathcal{S}$ such that $\bigcup \mathcal{F} = X$). Equivalently, X is compact if every collection $\mathcal{C}$ of closed subsets of X with the **finite intersection property** (i.e., $\bigcap \mathcal{F} \neq \varnothing$ for all finite $\mathcal{F} \subseteq \mathcal{C}$) satisfies $\bigcap \mathcal{C} \neq \varnothing$. Note that we use the notion of compactness in the sense that a compact space need not be Hausdorff.

A subset C of a topological space (X, τ) is called compact if $(C, \tau \cap C)$ is compact. (The topology $\tau \cap C := \{U \cap C;\ U \in \tau\}$ denotes the initial topology on C with respect to the injection $C \hookrightarrow X$, also called the **induced topology**.) If X is a Hausdorff topological

© Springer Nature Switzerland AG 2020
J. Voigt, *A Course on Topological Vector Spaces*, Compact Textbooks in Mathematics,
https://doi.org/10.1007/978-3-030-32945-7_4

space, and C is a compact subset, then it is easy to see that the complement of C is open, i.e., that C is closed.

Let X be a set. A **filter** $\mathcal{F}$ in X is a non-empty collection $\mathcal{F} \subseteq \mathcal{P}(X)$ satisfying the following properties:

$\varnothing \notin \mathcal{F}$;
if $A \in \mathcal{F}$, $A \subseteq B \subseteq X$, then $B \in \mathcal{F}$;
if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

A **filter base** $\mathcal{F}_0$ in X is a non-empty collection $\mathcal{F}_0 \subseteq \mathcal{P}(X)$ with:

$\varnothing \notin \mathcal{F}_0$;
if $A, B \in \mathcal{F}_0$, then there exists $C \in \mathcal{F}_0$ such that $C \subseteq A \cap B$.

If $\mathcal{F}_0$ is a filter base, then

$$\mathrm{fil}(\mathcal{F}_0) := \big\{A \subseteq X;\ \text{there exists } B \in \mathcal{F}_0 \text{ such that } B \subseteq A\big\}$$

is a filter, called the **filter generated by** $\mathcal{F}_0$. A filter $\mathcal{F}$ is called an **ultrafilter** if there is no filter properly containing $\mathcal{F}$.

Let now X be a topological space, $\mathcal{F}$ a filter in X, $x \in X$. Then $\mathcal{F}$ **converges to** x (or x **is a limit of** $\mathcal{F}$), $\mathcal{F} \to x$, if $\mathcal{U}_x \subseteq \mathcal{F}$. If $\mathcal{F}_0$ is a filter base, then one also writes $\mathcal{F}_0 \to x$ if the generated filter $\mathrm{fil}(\mathcal{F}_0)$ converges to x, i.e., if for all $U \in \mathcal{U}_x$ there exists $A \in \mathcal{F}_0$ with $A \subseteq U$. The point x is called a **cluster point** (also 'accumulation point') of a filter $\mathcal{F}$, if for all $U \in \mathcal{U}_x$, $A \in \mathcal{F}$ one has $U \cap A \neq \varnothing$, or equivalently, if $x \in \bigcap_{A \in \mathcal{F}} \overline{A}$.

Examples 4.2
Let X be a set.

(a) If $x \in X$, then $\mathcal{F}_0 := \big\{\{x\}\big\}$ is a filter base. The generated filter is called the **filter fixed at** x.

(b) If (x_n) is a sequence in X, then $\mathcal{F}_0 := \big\{\{x_j;\ j \geqslant n\};\ n \in \mathbb{N}\big\}$ is a filter base. The generated filter is called an **elementary filter**.

If additionally X is a topological space and $x \in X$, then $\mathcal{F}_0 \to x$ if and only if $x_n \to x$ as $n \to \infty$.

(c) Let X be a topological space, $x \in X$. Then $\mathcal{U}_x$ is a filter (the neighbourhood filter of x). △

Remarks 4.3 Let X be a set.

(a) If $\mathcal{F}$ is a filter in X, $A \subseteq X$ such that $A \cap B \neq \varnothing$ for all $B \in \mathcal{F}$, then obviously $\{A \cap B;\ B \in \mathcal{F}\}$ is a filter base, and the generated filter is **finer** than $\mathcal{F}$ (i.e., it contains $\mathcal{F}$).

(b) Let $\mathcal{F}$ be a filter. Then $\mathcal{F}$ is an ultrafilter if and only if for all $A \subseteq X$ one has $A \in \mathcal{F}$ or $X \setminus A \in \mathcal{F}$. (Necessity: If $A \cap B \neq \varnothing$ for all $B \in \mathcal{F}$, then (a) implies that there is a finer filter containing A, and this filter is equal to $\mathcal{F}$ because $\mathcal{F}$ is an ultrafilter; thus $A \in \mathcal{F}$. Otherwise there exists $B \in \mathcal{F}$ such that $A \cap B = \varnothing$, and then $X \setminus A \in \mathcal{F}$. Sufficiency: The condition implies that there is no finer filter.)

(c) For every filter $\mathcal{F}$ in X there exists a finer ultrafilter. This is an immediate consequence of Zorn's lemma. (In the proof that a maximal element is an ultrafilter one uses (a) and (b).)

(d) If X is a topological space, $\mathcal{F}$ is an ultrafilter in X, and $x \in X$ is a cluster point of $\mathcal{F}$, then $\mathcal{F} \to x$. (If $U \in \mathcal{U}_x$, then $U \cap A \neq \varnothing$ for all $A \in \mathcal{F}$, therefore $U \in \mathcal{F}$, because $\mathcal{F}$ is an ultrafilter.) △

Remark 4.4 In our treatment we will use filters to discuss convergence and continuity in topological spaces. Filters generalise sequences – see Example 4.2(b) – which are sufficient for this purpose in metric spaces. (Another generalisation of sequences are 'nets', a notion that we will not need.) The proof of Theorem 4.1 becomes particularly nice with filters, but also for the discussion of completeness (Chapter 9) filters will be convenient. △

Proposition 4.5 *Let X be a topological space. Then the following properties are equivalent:*

(i) *X is compact;*
(ii) *every filter in X possesses a cluster point;*
(iii) *every ultrafilter in X is convergent.*

Proof

(i) ⇒ (ii). Let $\mathcal{F}$ be a filter in X. Then the collection $\{\overline{A};\ A \in \mathcal{F}\}$ has the finite intersection property, and therefore $\bigcap_{A\in\mathcal{F}} \overline{A} \neq \varnothing$, i.e., $\mathcal{F}$ has a cluster point.

(ii) ⇒ (i). Let $\mathcal{C} \subseteq \mathcal{P}(X)$ be a collection of closed sets with the finite intersection property. Then $\mathcal{F}_0 := \big\{\bigcap \mathcal{A};\ \mathcal{A} \subseteq \mathcal{C} \text{ finite}\big\}$ is a filter base. The generated filter $\mathcal{F}$ has a cluster point, i.e., $\varnothing \neq \bigcap_{A\in\mathcal{F}} \overline{A} = \bigcap \mathcal{C}$.

'(ii) ⇒ (iii)' is obvious, in view of Remark 4.3(d).

(iii) ⇒ (ii). If $\mathcal{F}$ is a filter in X, then there exists a finer ultrafilter; see Remark 4.3(c). Every limit of this filter is a cluster point of $\mathcal{F}$. □

Let X, Y be sets, $f : X \to Y$, $\mathcal{F}$ a filter in X. Then $f(\mathcal{F}) := \{f(A);\ A \in \mathcal{F}\}$ is a filter base in Y, and the generated filter $\operatorname{fil}(f(\mathcal{F}))$ is called the **image filter**.

If $\mathcal{F}$ is an ultrafilter, then $f(\mathcal{F})$ is an ultrafilter base. Indeed, for $B \subseteq Y$ one has $f^{-1}(B) \in \mathcal{F}$ or $f^{-1}(Y \setminus B) \in \mathcal{F}$. In the first case one concludes that $f(f^{-1}(B)) \subseteq B \in \operatorname{fil}(f(\mathcal{F}))$, in the second case that $Y \setminus B \in \operatorname{fil}(f(\mathcal{F}))$.

Proposition 4.6

(a) *Let X, Y be topological spaces, $x \in X$, $\mathcal{F}$ a filter in X, $\mathcal{F} \to x$, $f : X \to Y$ continuous at x. Then $f(\mathcal{F}) \to f(x)$.*
(b) *Let X, X_ι $(\iota \in I)$ be topological spaces, $f_\iota : X \to X_\iota$ $(\iota \in I)$, and let the topology on X be the initial topology with respect to $(f_\iota)_{\iota\in I}$. Let $x \in X$, $\mathcal{F}$ a filter in X. Then $\mathcal{F} \to x$ if and only if $f_\iota(\mathcal{F}) \to f_\iota(x)$ for all $\iota \in I$.*

Proof

(a) Let V be a neighbourhood of $f(x)$. Then $f^{-1}(V)$ is a neighbourhood of x, and therefore $f^{-1}(V) \in \mathcal{F}$. From $f(f^{-1}(V)) \subseteq V$ one then obtains $V \in \operatorname{fil}(f(\mathcal{F}))$.

(b) The necessity is clear from (a). For the sufficiency let $U \in \mathcal{U}_x$. Then there exist a finite set $F \subseteq I$ and neighbourhoods U_ι of $f_\iota(x)$ $(\iota \in I)$ such that $\bigcap_{\iota\in F} f_\iota^{-1}(U_\iota) \subseteq U$. There exists $A \in \mathcal{F}$ such that $f_\iota(A) \subseteq U_\iota$ $(\iota \in F)$. Therefore

$$A \subseteq f_\iota^{-1}(f_\iota(A)) \subseteq f_\iota^{-1}(U_\iota) \quad (\iota \in F),$$

$$A \subseteq \bigcap_{\iota\in F} f_\iota^{-1}(U_\iota) \subseteq U.$$

□

Proof of Theorem 4.1

Without restriction all $X_\iota \neq \varnothing$. Let $\mathcal{F}$ be an ultrafilter in $\prod_{\iota\in I} X_\iota$. Then $\operatorname{pr}_\iota(\mathcal{F})$ is an ultrafilter base in X_ι, therefore convergent by Proposition 4.5, $\operatorname{pr}_\iota(\mathcal{F}) \to x_\iota \in X_\iota$ $(\iota \in I)$. Then Proposition 4.6(b) implies that $\mathcal{F} \to (x_\iota)_{\iota\in I}$. □

As in the case of Banach spaces Tikhonov's theorem implies the Banach–Alaoglu theorem, i.e., the closed dual ball is weak*-compact, we now derive the corresponding result for locally convex spaces.

Theorem 4.7 (Alaoglu–Bourbaki)
Let E be a locally convex space, $U \subseteq E$ a neighbourhood of zero. Then $U^\circ \subseteq E'$ is $\sigma(E', E)$-compact.

Lemma 4.8 *Let E be a vector space. Then E^* is closed in $\mathbb{K}^E$ with respect to the product topology.*

Proof

For $\lambda \in \mathbb{K}$, $x, y \in E$ the mapping

$$\varphi_{\lambda,x,y} \colon \mathbb{K}^E \to \mathbb{K}, \ \ f \mapsto f(\lambda x + y) - \lambda f(x) - f(y)$$

is continuous. (Note that, for $x \in E$, the mapping $\mathbb{K}^E \ni f \mapsto f(x) \in \mathbb{K}$ is one of the projections defining the product topology.) Therefore $E^* = \bigcap_{\lambda\in\mathbb{K}, x,y\in E} \varphi_{\lambda,x,y}^{-1}(0)$ is closed. □

Proof of Theorem 4.7

Without loss of generality we may assume that U is absolutely convex. We note that $x' \in U^\circ$ if and only if $x' \in E^*$ and $|\langle x, x'\rangle| \leqslant p_U(x)$ $(x \in E)$. The condition is clearly sufficient. On the other hand, if $x' \in U^\circ$, $x \in E$, $\lambda > p_U(x)$, then $\frac{1}{\lambda}x \in U$, $|\langle \frac{1}{\lambda}x, x'\rangle| \leqslant 1$, $|\langle x, x'\rangle| \leqslant \lambda$;

therefore, $|\langle x, x'\rangle| \leqslant p_U(x)$. This implies that

$$\begin{aligned} U^\circ &= \{x' \in E^*;\ |\langle x, x'\rangle| \leqslant p_U(x)\ (x \in E)\} \\ &= \{f \in \mathbb{K}^E;\ |f(x)| \leqslant p_U(x)\ (x \in E)\} \cap E^* \\ &= \Big(\prod_{x\in E} B_{\mathbb{K}}[0, p_U(x)]\Big) \cap E^*. \end{aligned}$$

Theorem 1.2 implies that the weak topology on E' and the product topology on $\prod_{x\in E} B_{\mathbb{K}}[0, p_U(x)]$ are the restrictions of the product topology on $\mathbb{K}^E = \prod_{x\in E} \mathbb{K}$ to these sets. Because of Lemma 4.8 it therefore follows that U° is a closed subset of the compact set $\prod_{x\in E} B_{\mathbb{K}}[0, p_U(x)]$. □

Let $\langle E, F\rangle$ be a dual pair. Let

$$\mathcal{M}_\mu := \{B \subseteq F;\ B \text{ absolutely convex and } \sigma(F, E)\text{-compact}\}.$$

Obviously one has $\mathcal{M}_\mu \subseteq \mathcal{B}_\sigma(F, E)$. Then the polar topology

$$\mu(E, F) := \tau_{\mathcal{M}_\mu}$$

on E is called the **Mackey topology**. The Mackey topology $\mu(F, E)$ on F is defined correspondingly.

In the following Chapter 5 we will show that $(E, \mu(E, F))' = b_2(F)$, and that $\mu(E, F)$ is the strongest topology with dual $b_2(F)$, in the following sense: If $\langle E, F\rangle$ is a separating dual pair, then a locally convex topology τ on E is compatible with $\langle E, F\rangle$ if and only if $\sigma(E, F) \subseteq \tau \subseteq \mu(E, F)$.

In the last statement, the necessity of the condition is easily obtained from our treatment presented so far. If τ is compatible, the property $\sigma(E, F) \subseteq \tau$ follows from the definition of the topology $\sigma(E, F)$ (and Theorem 1.2), whereas the property $\tau \subseteq \mu(E, F)$ is a consequence of Theorem 4.7, as follows. The space (E, τ) possesses a neighbourhood base of zero $\mathcal{U}$ consisting of closed absolutely convex sets; hence

$$\mathcal{M} := \{U^\circ;\ U \in \mathcal{U}\} \subseteq \mathcal{M}_\mu,$$

by Theorem 4.7, and therefore $\tau = \tau_{\mathcal{M}} \subseteq \tau_{\mathcal{M}_\mu} = \mu(E, F)$.

The definition of $\mathcal{M}_\mu$ suggests the question whether in a locally convex space the closed absolutely convex hull of a compact set is again compact. Example 4.10 given below shows that this is not always the case. We will show in Chapter 11 that it is true if E is quasi-complete (Corollary 11.5). In particular it is true if E is a Banach space (‘Mazur’s theorem’). In Chapter 14 we will show that it is also true for the weak topology in a Banach space (‘Krein’s theorem’). However, it is always true that the closed absolutely convex hull of a compact *convex* set is compact; this is the content of the following lemma. As a consequence one obtains $\mu(E, F) = \tau_{\mathcal{M}'_\mu}$ also for

$$\mathcal{M}'_\mu := \{B \subseteq F;\ B \text{ convex and } \sigma(F, E)\text{-compact}\}.$$

Lemma 4.9 *Let E be a topological vector space, and let $A \subseteq E$ be a compact convex subset. Then $\overline{\operatorname{aco}}\, A$ is compact.*

Proof

(i) If $B \subseteq E$ is a balanced subset, then aco $B =$ co B. This holds because

$$\operatorname{co} B = \left\{\sum_{j=1}^{n} \lambda_j x_j;\ \lambda_1, \ldots, \lambda_n \in [0,1],\ \sum_{j=1}^{n} \lambda_j = 1,\ x_1, \ldots, x_n \in B,\ n \in \mathbb{N}\right\}$$

is easily seen to be balanced.

(ii) If $\mathbb{K} = \mathbb{R}$, then bal $A = [-1, 1] \cdot A \subseteq \operatorname{co}(A \cup (-A))$, and the latter set is compact (as the image of the compact set $\{(\lambda_1, \lambda_2) \in [0, 1]^2;\ \lambda_1 + \lambda_2 = 1\} \times A \times (-A)$ under the continuous mapping $(\lambda_1, \lambda_2, x_1, x_2) \mapsto \lambda_1 x_1 + \lambda_2 x_2$). Hence $\overline{\operatorname{aco}}\, A = \overline{\operatorname{co}(\operatorname{bal} A)} \subseteq \operatorname{co}(A \cup (-A))$ is compact.

(iii) If $\mathbb{K} = \mathbb{C}$, then

$$\operatorname{bal} A = B_{\mathbb{C}}[0, 1] \cdot A \subseteq \sqrt{2}\operatorname{co}\big(A \cup (\mathrm{i}A) \cup (-A) \cup (-\mathrm{i}A)\big),$$

where again the latter set is compact. The remaining argument is as in (ii). □

Example 4.10 (cf. [Kha82, Chap. II, Example 10])

Consider the dual pair $\langle c_c, \ell_1\rangle$, where $c_c := \operatorname{lin}\{e_n;\ n \in \mathbb{N}\}$, with the 'unit vectors' e_n in c_0 (or ℓ_1). The sequence $(2^n e_n)_n$ in ℓ_1 is $\sigma(\ell_1, c_c)$-convergent to 0; therefore $B := \{2^n e_n;\ n \in \mathbb{N}\} \cup \{0\}$ is $\sigma(\ell_1, c_c)$-compact. For $n \in \mathbb{N}$, the element $y^n := \sum_{j=1}^n e_j = \sum_{j=1}^n 2^{-j}(2^j e_j)$ belongs to co B. For a $\sigma(\ell_1, c_c)$-cluster point $y = (y_j)$ of the sequence (y^n), the coordinate y_j would have to be a cluster point of the sequence $(y^n_j)_n$, i.e., $y_j = 1$ $(j \in \mathbb{N})$. However, the element $(1, 1, 1 \ldots)$ does not belong to ℓ_1. This shows that the sequence $(y^n)_n$ does not have a cluster point, and therefore the set co B is not relatively compact with respect to $\sigma(\ell_1, c_c)$. △

We include an additional information on metrisability in the context of the Alaoglu–Bourbaki theorem.

Proposition 4.11 *Let E be a separable locally convex space, $U \subseteq E$ a neighbourhood of zero. Then the topology $\sigma(E', E)$ is metrisable on $U^\circ \subseteq E'$.*

Proof

Let $A \subseteq E$ be a countable dense set, and denote by ρ the initial topology on E' with respect to the family $\big(E' \ni x' \mapsto \langle x, x'\rangle \in \mathbb{K}\big)_{x \in A}$. Then ρ is coarser than $\sigma(E', E)$, and ρ is metrisable, by Proposition 2.17 (where the denseness of A in E implies that ρ is Hausdorff).

As $(U^\circ, \sigma(E', E) \cap U^\circ)$ is compact by the Alaoglu–Bourbaki theorem, one concludes from Lemma 4.12, proved below, that $\rho \cap U^\circ = \sigma(E', E) \cap U^\circ$. □

For completeness we recall (from general topology) the following important basic observation concerning compactness.

Lemma 4.12 *Let X, Y be topological spaces, X compact, Y Hausdorff, $f : X \to Y$ continuous and bijective. Then f is a homeomorphism.*

Proof

We only have to show that f is an open mapping. Let $U \subseteq X$ be an open set. Then $X \setminus U$ is closed, hence compact. This implies that $Y \setminus f(U) = f(X \setminus U)$ is compact, hence closed, i.e., $f(U)$ is open. □

Notes Tikhonov's theorem is one of the basic theorems of topology, in some sense the first result in the development of set theoretic topology which does not come along with a straightforward 'evident' proof. Tikhonov (in early German transcription "Tychonoff") proved the theorem for compact intervals in [Tyc30] and mentioned later that the proof carries over to the general case. The main result of this chapter, the Alaoglu–Bourbaki theorem (Theorem 4.7), uses Tikhonov's theorem. For the case of normed spaces it usually is called the Banach–Alaoglu theorem, proved for the separable case by Banach [Ban32, VIII, § 5, Théorème 3] and for the general case by Bourbaki [Bou38, Corollaire de Théorème 1] (and shortly after by Alaoglu [Ala40, Theorem 1:3]). The first appearance of the general case may be in a paper of Arens [Are47, proof of Theorem 2]. (It is also contained in Bourbaki [Bou64b, Chap. IV, § 2.2, Proposition 2].) The Mackey topology was first defined and used by Arens [Are47]; we use the notation $\mu(E, F)$, for a dual pair $\langle E, F\rangle$, thereby following Floret [Flo80]. (A more traditional notation, used by many authors, would be $\tau(E, F)$, and the author has been told the reason for this notation: $\sigma(E, F)$ is the 'beginning' of the scale of compatible locally convex topologies, and $\tau(E, F)$ is the 'end'; like one often uses $[a, b]$ for intervals in $\mathbb{R}$, the idea is to use the neighbouring letters σ and τ in the Greek alphabet as the ends of the 'interval'. As we use 'τ' quite generally for topologies, we prefer Floret's notation. Anyway, 'σ' in weak topologies probably comes from the 's' in the German "schwach". The earliest place where the author could localise the use of '$\sigma(E, E')$' for the weak topology, is the note [Die40].)

Summarising the previous discussion, if the names given to theorems should indicate their authors, then the Banach–Alaoglu theorem should be called 'Banach–Bourbaki theorem', the Alaoglu–Bourbaki theorem should be called 'Bourbaki–Arens theorem', and the Mackey topology should be called 'Arens topology' (although in the latter case 'Arens–Mackey topology' would be equally justified).

Concerning Lemma 4.9, we refer to [Edw65, Remark 8.13.4(3)].