

Weak Convergence in Banach Spaces

Joel H. Shapiro

February 28, 2022

These notes accompany my lectures in the Analysis Seminar at Portland State University concerning the notion of “weak convergence” for Banach spaces. The discussion revolves around Issai Schur’s remarkable discovery: *In the classical sequence space ℓ^1 , every weakly convergent sequence is norm-convergent.*

Contents

1	Introduction	1
2	Convergence: weak vs. norm in ℓ^p	2
3	The Weak Topology	3
4	Convergence (The Cosmic Truth)	5
5	Topology via Convergence: Basics	8
6	Topology via Convergence: Compactness	9
7	Schur’s Lemma	12
8	Isomorphisms of ℓ^p Spaces	14
9	References	16

1 Introduction

Let X denote a Banach space and X^* its dual—the space of all continuous linear functionals on X .

Definition 1.1. To say a sequence (x_n) of vectors in X converges weakly to $x \in X$ means: $\varphi(x_n) \rightarrow \varphi(x)$ for every $\varphi \in X^*$ (notation: $x_n \xrightarrow{w} x$).

This definition contrasts markedly with the usual:

“(x_n) converges in norm to x ”,

which simply means that $\|x_n - x\| \rightarrow 0$ (notation: $x_n \xrightarrow{\|\cdot\|} x$).

The norm-continuity of each $\varphi \in X^*$ implies:

Proposition 1.2. $x_n \xrightarrow{\|\cdot\|} x \implies x_n \xrightarrow{w} x$.

EXERCISE. *In euclidean space: weak convergence implies norm convergence.*

All the results we prove for general Banach spaces will hold as well for normed spaces, usually with the same proofs.

This result remains true for all finite dimensional Banach spaces, each one being linearly isomorphic to the euclidean space of the same dimension. See, e.g., [2, pp. 12-13.]

2 Convergence: weak vs. norm in ℓ^p

We'll regard *sequences* to be real-valued functions on the natural numbers \mathbb{N} . For $1 \leq p < \infty$ the space ℓ^p will be the collection of sequences $f: \mathbb{N} \rightarrow \mathbb{R}$ such that

$$(1) \quad \|f\|_p^p := \sum_{j=1}^{\infty} |f(j)|^p < \infty.$$

The functional $\|\cdot\|_p$ is a norm that makes ℓ^p into a Banach space.¹

Definition 2.1. The *standard basis* for ℓ^p is the sequence of vectors $(e_n)_1^\infty$, where for each $n \in \mathbb{N}$:

$$(2) \quad e_n(j) := \begin{cases} 1 & \text{if } j = n \\ 0 & \text{otherwise} \end{cases}$$

The standard basis is a *Schauder basis*² for ℓ^p , in the sense that each $f \in \ell^p$ is represented uniquely by the infinite series $\sum_{j=1}^{\infty} f(j)e_j$, the

series being convergent in ℓ^p (i.e., $\lim_{N \rightarrow \infty} \|f - \sum_{j=1}^N f(j)e_j\|_p = 0$).

Proposition 2.2. For $1 \leq p < \infty$, the standard basis $(e_n)_1^\infty$ of ℓ^p :

- (a) does not converge in norm,
- (b) converges weakly to 0 if $p \neq 1$, but
- (c) does not converge weakly if $p = 1$.

Proof. (a) For $n, m \in \mathbb{N}$ with $m \neq n$ we have $\|e_n - e_m\|_p = 2^{1/p} > 0$. Thus the standard basis is not Cauchy in ℓ^p , so (since ℓ^p is a Banach space) not convergent.*

(b) For $1 < p < \infty$ we wish to show that $\varphi(e_n) \rightarrow 0$ for each $\varphi \in (\ell^p)^*$. Let q denote the index "conjugate" to p , i.e., $\frac{1}{p} + \frac{1}{q} = 1$. Fix $\varphi \in (\ell^p)^*$. The Riesz Representation Theorem provides $g \in \ell^q$ such that

$$\varphi(f) = \sum_{j=1}^{\infty} f(j)g(j) \quad (f \in \ell^p).$$

Thus $\varphi(e_n) = g(n)$ for each $n \in \mathbb{N}$, and $g(n) \rightarrow 0$ since $g \in \ell^q$ with $q < \infty$. Thus $e_n \xrightarrow{w} 0$ in ℓ^p , as claimed.

(c) Define the linear functional φ on ℓ^1 by

$$\varphi(f) = \sum_{j=1}^{\infty} f(j) \quad (f \in \ell^1),$$

Throughout these notes the scalar field will be \mathbb{R} . We could just as well have used \mathbb{C} .

¹ See, e.g., [1, Thm. 7.20, page 204] for this in a more general setting.

² See, e.g., [2, Chapter 3]

* For an argument that doesn't use the completeness of ℓ^p , observe that norm convergence in ℓ^p implies pointwise convergence on \mathbb{N} , and that on \mathbb{N} the standard basis converges pointwise to zero. Thus the only possible norm-limit for the standard basis is the zero-vector, which is manifestly *not* a norm-limit.

and observe that $|\varphi(f)| \leq \|f\|_1$ for each $f \in \ell^1$, so $\varphi \in (\ell^1)^*$.
For $n, m \in \mathbb{N}$ with $n \neq m$ we have

$$|\varphi(e_m) - \varphi(e_n)| = |\varphi(e_m - e_n)| = 2,$$

so the sequence $(\varphi(e_n))_1^\infty$ is not Cauchy in \mathbb{R} , hence not convergent.
Thus the standard basis of ℓ^1 does not converge weakly. \square

Part (c) of Proposition 2.2 is a “baby version” of the following remarkable result:

Theorem 2.3 (Schur’s Lemma). *In ℓ^1 : $f_n \xrightarrow{w} f \implies f_n \xrightarrow{\|\cdot\|} f$.*

Issai Schur, [4, 1920.]

We’ll prove Schur’s Lemma in §7. Until then, we’ll be seeking to understand its “cosmic meaning”.

3 The Weak Topology

In this section we’ll interpret Schur’s Lemma in the setting of the “weak topology”, *the weakest topology on X for which each linear functional in X^* remains continuous*.

To insure that $\varphi \in X^*$ is “weakly continuous” at (say) the origin of X , we must decree that for each $\varepsilon > 0$ the subset

$$(3) \quad V_\varphi(0, \varepsilon) := \varphi^{-1}(-\varepsilon, \varepsilon) = \{x \in X : |\varphi(x)| < \varepsilon\}$$

be “weakly open”. It follows quickly that the collection \mathcal{V}_0 of finite intersections of such sets satisfies the axioms for a base of neighborhoods of 0 in X , and that for each $x \in X$ the translates $\mathcal{V}_x := \mathcal{V} + x$ form a neighborhood base at x , and all these neighborhood bases taken together form a base for a topology on X : the “weak topology.”

The “weak topology” is constructed to preserve the continuity of each “norm-continuous” linear functional. Does it bestow continuity on some linear functionals that are not norm continuous? The answer is NO, but the proof requires a little work; see, e.g., [6, Theorem 2.4.11, pp. 207-8].

Proposition 3.1 (Some weak-topology basics). *The weak topology of a Banach space is Hausdorff, locally convex, and weaker than the norm topology. If a Banach space is infinite dimensional, then its weak topology is strictly weaker than its norm topology.*

Proof. (a) To see that the weak topology is Hausdorff one need only recall that the Hahn-Banach Theorem provides enough continuous linear functionals to separate points. In other words: if x_1 and x_2 are distinct vectors in the Banach space X , then there exists $\varphi \in X^*$ such that $\varphi(x_1) \neq \varphi(x_2)$. Upon choosing $\varepsilon = \frac{1}{4}|\varphi(x_1) - \varphi(x_2)|$ we have $V_\varphi(x_1, \varepsilon)$ and $V_\varphi(x_2, \varepsilon)$ disjoint.

(b) “Locally convex” means that each point has a base of convex neighborhoods. This is obvious since each “sub-basic” neighborhood

$V_\varphi(0, \varepsilon)$ is convex, hence so is each translate $V_\varphi(x, \varepsilon)$, and therefore so is each finite intersection

$$V_\Phi(x, \varepsilon) := \bigcap_{\varphi \in \Phi} V_\varphi(x, \varepsilon)$$

for Φ a finite subset of X^* . Since these sets $V_\Phi(x, \varepsilon)$ form a base for the weak neighborhood system at x , the weak topology of X is locally convex.

(c) Since each linear functional φ in the definition of weak topology is continuous for the norm topology of X , it follows that $V_\varphi(0, \varepsilon) = \varphi^{-1}((0, \varepsilon))$ is norm-open, hence so is every translate $V_\varphi(x, \varepsilon)$, and therefore so is every weakly open subset of X . Thus the weak topology of X is no stronger than the norm topology.

(d) To see that if X is infinite dimensional the weak topology is strictly stronger than the norm topology, just observe that the basic weak 0-neighborhood $V_\Phi(0, \varepsilon)$ contains $\bigcap_{\varphi \in \Phi} \ker \varphi$ which, being a finite intersection of subspaces of codimension one, has finite codimension in X . Since X is infinite dimensional, this intersection must be infinite dimensional; in particular it can't be the zero-subspace.

CONCLUSION: Every weak zero-neighborhood in the infinite dimensional Banach space X contains a nontrivial subspace.

In particular, the open unit ball of X can't contain any weak neighborhood of the origin, so it is not weakly open. Therefore the collection of weakly open subsets of X is strictly contained in the collection of norm-open ones, i.e, the weak topology is strictly weaker than the norm topology. \square

Definition 3.2. To say a sequence (x_n) in a topological space X converges to $x_0 \in X$ means: For every neighborhood V of x_0 there exists $N = N(V) \in \mathbb{N}$ such that $n > N \implies x_n \in V$.

Thus: in a Banach space, a sequence converges weakly iff it converges in the weak topology.

Since weak neighborhoods are all norm-unbounded, it appears that the weak topology must be "much weaker" than the norm topology. Here's a spectacular illustration of this: let B_X denote the open unit ball of the Banach space X , and S_X the unit sphere. That is:

$$S_X := \{x \in X: \|x\| = 1\} \quad \text{and} \quad B_X := \{x \in X: \|x\| < 1\}.$$

Theorem 3.3. *Every point of the open unit ball of a Banach space is a weak limit point of the unit sphere.*³

Proof. Fix $x_0 \in B_X$, and let V be any weak neighborhood of x_0 . We wish to show that V has nonempty intersection with S_X . To this end, note that V contains a basic weak neighborhood

$$V_\Phi(x_0, \varepsilon) := \bigcap_{\varphi \in \Phi} \{x \in X: |\varphi(x) - \varphi(x_0)| < \varepsilon\} = V_\Phi(0, \varepsilon) + x_0$$

for some finite subset Φ of X^* and some $\varepsilon > 0$. For each $\varphi \in \Phi$ the neighborhood $V_\varphi(0, \varepsilon)$ contains $\ker \varphi$, hence $V_\Phi(0, \varepsilon)$ contains $\ker \Phi := \bigcap_{\varphi \in \Phi} \ker \varphi$, a subspace which we have already noted is not the zero-subspace. Thus V is a convex set that contains a vector of norm < 1 (namely x) and vectors of norm > 1 , so it has nontrivial intersection with the unit sphere. \square

According to Proposition 3.3, every point of the open unit ball of ℓ^1 is a limit point of the unit sphere. But according to Schur's Lemma (Theorem 2.3), none of these points is the limit of a *sequence* of unit vectors. Thus, unlike the situation for metric spaces,

Sequences do not suffice to describe the weak topology of ℓ^1 .

Is there a more general mode of convergence that recaptures the friendly connection between topology and convergence afforded to us by metric spaces or, must we: "live in eternal fear of meeting a space for which sequences do not suffice?"⁴

4 Convergence (The Cosmic Truth)

To get a feeling for how sequential convergence might be usefully generalized, we'll focus in this section on topological spaces X which, for the most part, need not even be Hausdorff. Our key observation is that for each $x \in X$ the neighborhood base \mathcal{V}_x of x is "cofinal" in the sense that for any pair of such neighborhoods there is a third one contained in both.

Thus may think of \mathcal{V}_x as being "directed" by *reverse inclusion*: $V \geq W$ iff $V \subset W$. If, for $V \in \mathcal{V}_x$, we choose a point $x_V \in V$, then we can think of the function $\mathcal{V}_x \rightarrow X$ defined by $V \rightarrow x_V$ as a kind of "generalized sequence" that "converges" to x , in the sense that for each fixed neighborhood V of x we have $x_W \in V$ for every $W \in \mathcal{V}_x$ with $W \geq V$.

This example is easily formalized:

³ Thus the closed unit ball $\overline{B_X}$ lies in the weak closure of the unit sphere. In fact, it *equals* the weak closure of the unit sphere. Indeed, if x_0 lies outside $\overline{B_X}$ then the Hahn-Banach theorem supplies $\varphi \in X^*$ such that $\varphi(x_0) > 1$, but $\varphi < 1$ on $\overline{B_X}$. Thus the weakly open neighborhood $\{\varphi > 1\}$ of x_0 does not intersect the unit sphere.

CONCLUSION: No vector in X with norm > 1 can be a weak limit point of S_X . \square

⁴ P. L. Clark [3, pp.158-159], who goes on to remark that many standard introductory topology texts: "... unfortunately seem content to leave their readers in this state of fear."

Definition 4.1 (Directed set). To say a set I is *directed* by a binary relation \geq means that our relation is *reflexive* ($i \geq i \ \forall i \in I$), *transitive* ($i \geq j$ and $j \geq k \implies i \geq k$), and *directible*: for each pair of elements $i, j \in I$ there exists $k \in I$ with $k \geq$ both i and j .

Examples 4.2 (First examples of directed sets). We've already used reverse inclusion to direct the collection of neighborhoods of a point in a topological space. Of course the same is true for any neighborhood base of a point of X . Here are some further examples:

- (a) The natural numbers \mathbb{N} , the integers \mathbb{Z} , and the real line \mathbb{R} , all under the usual ordering \geq .
- (b) The collection of all subsets of a given set, ordered by inclusion.
- (c) The collection \mathcal{T} of "tagged partitions" of the finite interval $[a, b]$ of the real line shows up in any discussion of Riemann integration. An element of \mathcal{T} is a pair (P, T) where P is a "partition": a subset $\{x_k: 1 \leq k \leq n\}$ of $[a, b]$ with $a = x_0 < x_1 < x_2 \dots < x_n = b$, and $T = \{t_k: 1 \leq k \leq n\}$ a "tag set", meaning: $x_{k-1} \leq t_k \leq x_k$ for each $1 \leq k \leq n$.

We order \mathcal{T} by set inclusion on the collection \mathcal{P} of partitions: *

$$(P_2, T_2) \geq (P_1, T_1) \iff P_2 \supset P_1.$$

Definition 4.3 (Net). A *net* in a set X is a function $\mathbf{x}: I \rightarrow X$, where I is a directed set.

Notation: To suggest a the connection with sequences:

For a net $\mathbf{x}: I \rightarrow X$ we will write x_i instead of $\mathbf{x}(i)$, and $(x_i)_{i \in I}$ (or just (x_i)) instead of \mathbf{x} .

Examples 4.4 (of nets). Here \mathbb{N}, \mathbb{Z} , and \mathbb{R} all have their natural order.

- (a) *Sequences* are nets where $I = \mathbb{N}$.
- (b) *Double-sided "sequences"* $(x_n)_{n \in \mathbb{Z}}$ are nets, with $I = \mathbb{Z}$.
- (c) *One-parameter families* $(x_i)_{i \in I}$ with $I = \mathbb{R}$ or $[0, \infty]$ are nets.
- (d) *Riemann integration* can be naturally defined in terms of nets. If f is a real-valued function on a finite closed interval $[a, b]$, then every tagged partition (P, T) , as defined in Examples 4.2(d), gives rise to a *Riemann sum*

$$\mathcal{R}(P, T) := \sum_{k=1}^n f(t_k)(x_k - x_{k-1})$$

Thus, for f a fixed function on $[a, b]$: the function $\mathcal{R}: \mathcal{T} \rightarrow \mathbb{R}$ is a net of real numbers indexed by the directed set \mathcal{T} of tagged partitions.

* Note that this ordering is *not* symmetric: If T_1 and T_2 are different tag sets for $P \in \mathcal{P}$, then the two tagged partitions (P, T_1) and (P, T_2) are both \geq each other, but are not equal. Thus a directed set need not be partially ordered.

Definition 4.5 (Net convergence). To say a net $(x_i)_{i \in I}$ in a topological space X converges to $x \in X$ (notation: $x_i \rightarrow x_0$) means:

For every neighborhood V of x there exists an index $i_V \in I$ such that $j \geq i_V \implies x_j \in V$ (in short: “ (x_i) is eventually in V ”).

Examples 4.6 (of convergent nets). We begin with a result that follows from our initial motivation for the definition of “net”.

- (a) *Nets define limit points.* We know that in a metric space, limit points can be defined using convergent sequences. Suppose X is a topological space, S is a subset of X , and x is a limit point of S . Let \mathcal{B} be any base for the neighborhood system of x . Then each $B \in \mathcal{B}$ intersects S in a point $x_B \neq x$. The map $B \rightarrow x_B$ then gives us a net in $S \setminus \{x\}$ that converges in X to x . Thus: in general topological spaces: *limit points can be defined by convergent nets.*
- (b) A *sequence* in a topological space converges iff it is convergent as a net.
- (c) A real-valued function on a finite real interval is Riemann integrable iff its net of Riemann sums converges.

According to Theorem 3.3, every point of the open unit ball of an infinite dimensional Banach space is a weak limit point of the unit sphere. Example 4.6(a) gives this a net-convergence interpretation:

If X is an infinite dimensional Banach space then each $x \in X$ with norm < 1 , is the weak limit of a net of unit vectors.

If our Banach space is *separable*, we can do better:

Theorem 4.7. *If X is a separable Banach space, then there is a countable set E of unit vectors such that each vector $x \in X$ with $\|x\| < 1$ is the weak limit of a net $(e_i)_{i \in I}$ in E .*

Here the index set I is *uncountable*. Recall that, by Schur’s Lemma, no point of the open unit ball can be the limit of a *sequence* of unit vectors.

Proof. Suppose X is our separable Banach space, and S_X its unit sphere (the set of all unit vectors). Fix a countable norm-dense subset E of S_X . Fix $x \in X$ with $\|x\| < 1$. We’ve already seen that each weak neighborhood V of x has nontrivial intersection with S_X . Since weakly open neighborhoods are norm-open, V contains a vector $x_V \in E$. As before, the net $(x_V)_{V \in \mathcal{V}}$, indexed by the reverse-inclusion-directed set \mathcal{V} of neighborhoods of x , converges to x . \square

5 Topology via Convergence: Basics

In this section we use nets to rephrase many of the topological properties (e.g., Hausdorff-ness, closedness, continuity, compactness) that, for metric spaces, can be characterized in terms of sequences.

Proposition 5.1 (Closed-ness via nets). *For a subset C of a topological space X , the following are equivalent:*

- (a) C is closed in X .
- (b) If a net in C converges to a point $x \in X$, then $x \in C$.

Proof. In view of Example 4.6(a), the Proposition is just a restatement of the fact that a set is closed if and only if it contains all its limit points. \square

We know that in a metric spaces the limit of a convergent sequence must be unique. For nets the corresponding result is:

Proposition 5.2 (Hausdorff-ness via nets). *A topological space is Hausdorff iff no convergent net has more than one limit.*

Proof. (a) Suppose a net (x_i) in X converges to both x and y , with $x \neq y$. Then that net is eventually in every neighborhood of x and eventually in every neighborhood of y . Thus every neighborhood of x intersects every neighborhood of y , so X is not Hausdorff.

Conclusion: Hausdorff implies unique limits.

(b) Conversely, suppose X is not Hausdorff. Then there exist distinct points y and z in X with the property that each neighborhood of one intersects each neighborhood of the other. Let \mathcal{Y} denote the collection of neighborhoods of y , and \mathcal{Z} those of z . For each pair $(Y, Z) \in \mathcal{Y} \times \mathcal{Z}$ there is a point $x_{(Y,Z)} \in Y \cap Z$. We order \mathcal{Y} and \mathcal{Z} by reverse inclusion, and impose the product ordering on their cartesian product:

$$(Y_1, Z_1) \leq (Y_2, Z_2) \iff Y_1 \supset Y_2 \text{ and } Z_1 \supset Z_2$$

One checks easily that, with this ordering, $\mathcal{Y} \times \mathcal{Z}$ is a directed set. These definitions insure that for each $(Y, Z) \in \mathcal{Y} \times \mathcal{Z}$ the net $(x_{(Y,Z)})_{(Y,Z) \in \mathcal{Y} \times \mathcal{Z}}$ is eventually in both Y and Z , and therefore converges to both y and z .

Conclusion: Unique limits implies Hausdorff. \square

Proposition 5.3 (Continuity via nets). *For topological spaces X and Y , and a function $f: X \rightarrow Y$, the following statements are equivalent:*

- (a) f is continuous.
- (b) If a net (x_i) converges in X to x , then $f(x_i) \rightarrow f(x)$ in Y .

Proof. (a) \implies (b). Suppose f is continuous and (x_i) is a net in X that converges in X to x . To show that $f(x_i) \rightarrow f(x)$ in Y , fix an open neighborhood V of $f(x)$. We wish to show that the image net $(f(x_i))$ is eventually in V . By continuity, $f^{-1}(V)$ is open in X , and contains x . Since the net (x_i) converges in X to x , it is eventually in $f^{-1}(V)$, so the image net $(f(x_i))$ is eventually in $f(f^{-1}(V)) = V$, as desired.

(b) \implies (a). Given (b) we wish to show that f is continuous, i.e., that $f^{-1}(V)$ is open in X for every open subset V of Y . Since “inverse functions respect set operations” it’s enough to do the same with “open” replaced by “closed”. Suppose, then, that V is a *closed* subset of Y . We wish to show that $f^{-1}(V)$ is closed in X , i.e., (by Proposition 5.1) that each X -convergent net in $f^{-1}(V)$ has its limit in that set. Let (w_i) be a net in $f^{-1}(V)$ that converges to $x \in X$. To show: $x \in f^{-1}(V)$, i.e., that $f(x) \in V$. Our hypothesis (b) implies that the image net $(f(w_i))$ is a net in V that converges in Y to $f(x)$. Now, as desired, $f(x) \in V$ by Proposition 5.1, since V is closed in Y . \square

This last result brings us full circle, from our original notion of weak convergence of sequences to topological weak convergence:

Corollary 5.4. *A net (x_i) in a Banach space X converges weakly (i.e., in the weak topology) to $x \in X$ iff $\varphi(x_i) \rightarrow \varphi(x)$ for every $\varphi \in X^*$.*

6 Topology via Convergence: Compactness

WE SHOW HERE, for Hausdorff⁵ topological spaces, that nets can be used to characterize the notion of *compactness*. We wish to generalize to arbitrary Hausdorff spaces the fact that a metric space is compact iff every sequence of its points has a convergent subsequence.

For the fundamental idea, recall that if X is a metric space, then to say a point $x \in X$ is a limit point of a sequence (x_n) in X means that *some subsequence converges to x* . Rephrasing this in terms of neighborhoods of x suggests:

Definition 6.1. To say a point of $x \in X$ is a *limit point* of a net $(x_i)_{i \in I}$ in X means that for every neighborhood V of x and every $i \in I$ there exists an index $i_V \geq i$ such that $x_{i_V} \in V$.

In short, “ (x_i) is frequently in V ,” or equivalently:

For every neighborhood V of x , our net is not eventually in $X \setminus V$.

Definition 6.1 suggests that we may test whether or not x is a limit point of a net $(x_i)_{i \in I}$ by:

⁵ The convention seems to be this: A set for which every open cover has a finite subcover is called *compact* whenever the ambient topological space is Hausdorff; otherwise it is called *quasi-compact*.

Quasi-compact sets may not live up to expectations; consider, e.g., the three-point space $X = \{a, b, c\}$ with open sets $\emptyset, X, \{a\}$, and $\{b, c\}$. Here the sets $\{c\}$ and $\{a, b\}$ are both quasi-compact, but not closed. In fact $\{a, b\}$ is dense.

- (a) Making a new index set $\mathcal{I} = I \times \mathcal{V}$, where \mathcal{V} is the collection of neighborhoods of x .
- (b) Directing \mathcal{I} via the *product order*:

$$(i_2, V_2) \geq (i_1, V_1) \iff i_2 \geq i_1 \text{ and } V_2 \subset V_1.$$

- (c) Choosing for $(i, V) \in \mathcal{I}$, an index $i_V \geq i$ such that $x_{i_V} \in V$.

In this way we create a new net $(x_{i_V})_{(i,V) \in \mathcal{I}}$, which makes possible the following rephrasing of Definition 6.1 above:

Proposition 6.2. *For the net $(x_{i_V})_{(i,V) \in \mathcal{I}}$ created above, and a point $x \in X$, the following statements are equivalent:*

- (a) $x_{i_V} \rightarrow x$.
- (b) x is a limit of the original net $(x_i)_{i \in I}$.

IN SUMMARY: Suppose $(x_i)_{i \in I}$ is a net that has x as a limit point. Let's assume that (x_i) does not converge to x . There is, nevertheless, a directed set \mathcal{I} , and a function $h: \mathcal{I} \rightarrow I$ such that:

- (a) For each $i \in I$ there exists $\gamma_i \in \mathcal{I}$ such that $\gamma \geq \gamma_i \implies h(\gamma) \geq i$.
- (b) The net $(x_{h(\gamma)})_{\gamma \in \mathcal{I}}$ converges to x .

The adjective *cofinal* is then attached to the function h .

Definition 6.3 (subnet). For a net $(x_i)_{i \in I}$, any net $(x_{h(j)})_{j \in J}$ with J a directed set and h a cofinal function $J \rightarrow I$, is called a *subnet* of the original net.

Theorem 6.4 (limit points via subnets). *For a net (x_i) in a topological space X , and a point $x \in X$, the following are equivalent:*

- (a) x is a limit point of (x_i)
- (b) (x_i) has a subnet that converges to x .

Proof. (a) \implies (b): This follows from Proposition 6.2 and the definition of "subnet".

(b) \implies (a): Suppose there is a subnet $(x_{h(j)})_{j \in J}$ that converges to x . Fix a neighborhood V of x . The subnet $(x_{h(j)})_{j \in J}$ is eventually in V , i.e., there exists $j_V \in J$ such that $j \geq j_V$ implies $x_{h(j)} \in V$. Since h is cofinal, for each $i \in I$ there exists $j_i \in J$ such that $j \geq j_i$ implies $h(j) \geq i$. Since J is directed, we can choose $j' \in J$ that is \geq both j_i and j_V . Thus $i' := h(j') \in V$ and $i' \geq i$, hence the net (x_i) is not eventually in $X \setminus V$. Since V was an arbitrary neighborhood of x , this shows that x is a limit point of (x_i) . \square

Theorem 6.5 (Compactness via nets). *For a Hausdorff topological space X , the following three statements are equivalent:*

- (a) *Every net in X has a limit point.*
- (b) *Every net in X has a convergent subnet.*
- (c) *X is compact⁶, i.e., every open cover of X has a finite subcover.*

Proof. We've already proved the equivalence of (a) and (b). We'll prove the equivalence of (a) and (c), using the following (equivalent) definition of "compact":

To say a topological space is compact means that: whenever a family \mathcal{F} of closed subsets, has the property that each finite sub-family has nonvoid intersection⁷, then the whole family has nonvoid intersection.

(a) \implies (c). In our topological space X , suppose \mathcal{F} is a family of closed subsets having the FIP. We wish to find a point x_0 that belongs to each member of \mathcal{F} . To this end, note that we may add to \mathcal{F} , without changing $\bigcap \mathcal{F}$, each intersection of a finite subfamily of \mathcal{F} (note that each such added set is closed and nonvoid). This new family, which we'll still denote by \mathcal{F} , is directed by reverse inclusion. For each $F \in \mathcal{F}$, choose a point $x_F \in F$. By our assumption (a), the net $(x_F)_{F \in \mathcal{F}}$ has a limit point x_0 .

To show: $x_0 \in \bigcap \mathcal{F}$.

Suppose this is not the case, i.e., that there exists $F_0 \in \mathcal{F}$ with $x_0 \notin F_0$. Then $V_0 = X \setminus F_0$ is an open neighborhood of x_0 . Since x_0 is a limit point of the net (x_F) , there exists $F \in \mathcal{F}$ with $F \subset F_0$ and $x_F \in V_0$. But $x_F \in F \subset F_0 = X \setminus V_0$. Thus our assumption that $x_0 \notin \bigcap \mathcal{F}$ has led to a contradiction.

(c) \implies (a). Suppose each family of closed subsets of X with the FIP has nonvoid intersection. Let $(x_i)_{i \in I}$ be a net in X . We wish to show that (x_i) has a limit point.

For each $i \in I$ let F_i denote the closure in X of the set $\{x_j : j \geq i\}$, and denote by \mathcal{F} the family of all such F_i . By the definition of "directed set", each finite subfamily of \mathcal{F} has nonvoid intersection.⁸ Thus $\bigcap \mathcal{F} \neq \emptyset$. Choose $x_0 \in \bigcap \mathcal{F}$.

To show. x_0 is a limit point of (x_i) .

Fix a neighborhood V of x_0 and an index $i \in I$. Then $V \cap F_i \neq \emptyset$ since x_0 is in F_i , hence (by the definition of F_i) we know there exists $j \in I$ with $j \geq i$ such that $x_j \in V$. Thus our net (x_i) is "frequently in V ", and since V is an arbitrary neighborhood of x_0 we have shown that x_0 is indeed a limit point of (x_i) . \square

⁶ All the arguments of this section remain true for non-Hausdorff spaces X , provided we replace the word "compact" here by "quasi-compact" (see sidenote 5 above).

⁷ Such a family of closed sets is said to have the *finite intersection property* (FIP). This statement is equivalent, via the boolean algebra of sets, to the more familiar "open cover" definition.

⁸ Given a $\{i_1, \dots, i_n\} \subset I$ there exists $j \in I$ with $j \geq i_k$ ($1 \leq k \leq n$). Thus $x_j \in$ each F_k .

7 Schur's Lemma

Recall that when we say a sequence (x_n) of vectors in a Banach space X converges *weakly* to $x \in X$, we mean that $\lim_n \varphi(x_n) = \varphi(x)$ for each continuous linear functional φ on X . If (x_n) converges to x in norm, then clearly it converges to x weakly, but, as we saw in Proposition 2.2, the converse need not be true.

In this section we return to the Banach space ℓ^1 ; those functions $f: \mathbb{N} \rightarrow \mathbb{R}$ with

$$\|f\|_1 := \sum_{n \in \mathbb{N}} |f(n)| < \infty.$$

Let χ_E denote the “characteristic function of E i.e., the function that takes the value 1 at each point of E , and 0 otherwise.

Our goal is to prove:

Theorem 7.1 (Schur's Lemma). *Every weakly convergent sequence in ℓ^1 is norm-convergent.*

Key to the proof of Schur's Lemma is the following result, which (as we'll see in the next section) has applications far beyond our present concerns.

Lemma 7.2 (The Gliding Hump Lemma). *Suppose (f_n) is a sequence of unit vectors in ℓ^1 that is bounded below in norm, but converges to zero pointwise on \mathbb{N} . Then there exists a subsequence (f_{n_k}) and a sequence (I_k) of contiguous pairwise disjoint subintervals that exhaust \mathbb{N} such that*

$$(4) \quad \lim_{k \rightarrow \infty} \frac{\|f_{n_k} \chi_{I_k}\|_1}{\|f_{n_k}\|_1} = 1.$$

Proof. For $f \in \ell^1$ and $E \subset \mathbb{N}$ it will be convenient to refer to $\|f \chi_E\|_1$ as the *mass* that f places on E .

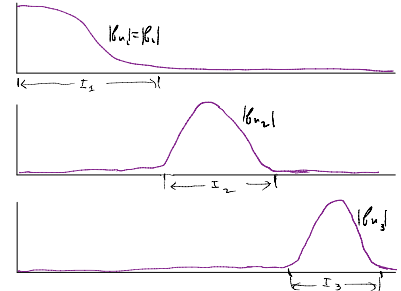
Since we are assuming that $\delta := \inf_n \|f_n\|_1 > 0$, we may replace each function f_n by $f_n / \|f_n\|_1$, and thereby assume without loss of generality that the original sequence (f_n) consists entirely of unit vectors.

Fix a sequence (ε_k) of positive numbers, with $1 > \varepsilon_k \searrow 0$. Let $n_1 = \alpha_1 = 1$, and choose $\alpha_2 > \alpha_1$ so that f_{n_1} places mass $\geq 1 - \varepsilon_1$ on the interval $[\alpha_1, \alpha_2]$.

Proceeding by induction: assume that for $1 \leq j \leq k$, functions f_{n_j} and contiguous intervals $I_j = [\alpha_j, \alpha_{j+1}]$ have been chosen so that $\|f_{n_j} \chi_{I_j}\|_1 \geq 1 - \varepsilon_j$ for $1 \leq j \leq k$. Our goal is to find an appropriate function $f_{n_{k+1}}$ and interval I_{k+1} . To this end we use pointwise convergence (to zero) to choose $n_{k+1} > n_k$ so that $f_{n_{k+1}}$ places mass



Issai Schur 1875–1941



As $k \rightarrow \infty$, the function f_{n_k} places an ever-larger proportion of its mass in a “hump” over I_k .

$< \varepsilon_{k+1}/2$ on $[1, \alpha_{k+1}]$, and then (since $f_{n_{k+1}} \in \ell^1$) we may choose $\alpha_{k+2} > \alpha_{k+1}$ so that $f_{n_{k+1}}$ places mass $< \varepsilon_{k+1}/2$ on (α_{k+2}, ∞) . Thus $f_{n_{k+2}}$ places mass $\geq (1 - \varepsilon_{k+1})$ on $I_{k+1} := [\alpha_{k+1}, \alpha_{k+2})$, as desired. \square

Proof of Schur's Lemma. We wish to show that if $f_n \rightarrow f$ weakly in ℓ^1 , then $\|f_n - f\|_1 \rightarrow 0$. By replacing f_n with $f_n - f$ we may, without loss of generality, assume that $f = 0$ (otherwise, just replace f_n by $f_n - f$).

We'll prove the "contrapositive" result:

If (f_n) is a sequence in ℓ^1 that does not converge in norm to 0, then it does not converge weakly to 0, i.e., $\varphi(f_{n_k}) \not\rightarrow 0$ for some $\varphi \in (\ell^1)^$.*

To this end, we may assume that (f_n) converges pointwise to 0 on \mathbb{N} (for if not, it can't converge weakly to zero,⁹ and we're done). Since we're also assuming that $\|f_n\|_1 \not\rightarrow 0$, we may—upon passing to an appropriate subsequence and re-numbering, if necessary—assume $\delta := \inf_n \|f_n\|_1 > 0$.

⁹ Weak convergence implies pointwise convergence, i.e., pointwise non-convergence implies weak non-convergence.

With these preliminaries out of the way we can invoke the Gliding Hump Lemma to provide a subsequence $n_k \nearrow \infty$ and an exhaustion of \mathbb{N} by a sequence (I_k) of contiguous intervals such that $\lim_n \frac{\|f_{n_k} \chi_{I_k}\|_1}{\|f_{n_k}\|_1} \rightarrow 1$.

Let's write $h_k = f_{n_k} \chi_{I_k}$ for the "hump" that f_{n_k} places over I_k , and $r_k = f_{n_k} - h_k$ for the "remainder" of f_{n_k} . Thus we have the "Gliding Hump decomposition" $f_{n_k} = h_k + r_k$, where

$$(5) \quad \frac{\|h_k\|_1}{\|f_{n_k}\|_1} \rightarrow 1 \quad \text{and} \quad \frac{\|r_k\|_1}{\|f_{n_k}\|_1} \rightarrow 0 \quad \text{as} \quad k \rightarrow \infty.$$

Set $s: \mathbb{N} \rightarrow \{-1, 1\}$ equal to the *signum* of f_{n_k} (i.e., of h_k) on I_k .¹⁰

Then define $\varphi: \ell^1 \rightarrow \mathbb{R}$ by

$$\varphi(f) = \sum_{j=1}^{\infty} f(j) s(j) \quad (f \in \ell^1).$$

Observe that

$$\varphi(h_k) = \sum_{j \in I_k} f_{n_k}(j) s(j) = \sum_{j \in I_k} |h_k(j)| = \|h_k\|_1,$$

while

$$|\varphi(r_k)| = \left| \sum_{j=1}^{\infty} r_k(j) s(j) \right| \leq \sum_{j=1}^{\infty} |r_k(j) s(j)| = \sum_{j=1}^{\infty} |r_k(j)| = \|r_k\|_1.$$

Thus

$$|\varphi(f_{n_k})| = |\varphi(h_k) + \varphi(r_k)| \geq |\varphi(h_k)| - |\varphi(r_k)| \geq \|h_k\|_1 - \|r_k\|_1$$

which, in view of (5) above shows that, as $k \rightarrow \infty$:

$$|\varphi(f_{n_k})| \geq \|f_{n_k}\|_1 [1 - o(1)] - o(\|f_{n_k}\|_1) = \|f_{n_k}\|_1 [1 - o(1)] \geq \delta [1 - o(1)].$$

Thus $\liminf_n |\varphi(f_{n_k})| \geq \delta$, so (f_{n_k}) does not converge weakly to 0. \square

¹⁰ The *signum* of $f: S \rightarrow \mathbb{R}$ is the function that takes the value 1 on the set where f is positive, -1 on the set where f is negative, and 0 elsewhere.

8 Isomorphisms of ℓ^p Spaces

In this section we'll use the Gliding Hump Lemma to show that the different ℓ^p spaces are different in a fundamental way.

Definition 8.1. For Banach spaces X and Y , to say: “ X is isomorphic to Y ” means that there is a continuous linear transformation T mapping X onto Y that has continuous inverse.

We know that ℓ^p is strictly contained in ℓ^q for $1 \leq p \leq q < \infty$, and that their norms are not equivalent. Is it possible, however, that the spaces are isomorphic? According to the next theorem, the answer is NO.

Theorem 8.2. *Suppose $1 \leq p, q < \infty$, with $p \neq q$. Then ℓ^p is not isomorphic to ℓ^q .*

For the proof we'll require a basic result about weak topologies:

Proposition 8.3. *If X and Y are Banach spaces, and $T: X \rightarrow Y$ a linear transformation that is continuous when each space has its norm topology. Then T is also continuous when each space has its weak topology.*

Proof. Let (x_i) be a net in X that converges in the weak topology to $x \in X$. By Corollary 5.4 this means $\varphi(x_i) \rightarrow \varphi(x)$ for each $\varphi \in X^*$. By Proposition 5.3 it's enough to show that $Tx_i \rightarrow Tx$ weakly in Y , i.e., that $\psi(Tx_i) \rightarrow \psi(Tx)$ for each $\psi \in Y^*$. But this is obvious because the linear functional $\psi \circ T$ on X , being the composition of (norm-continuous) functions is also norm-continuous, i.e., belongs to X^* . Thus $(\psi \circ T)(x_i) \rightarrow (\psi \circ T)(x)$, as desired. \square

Proof of Theorem 8.2. Since $\|e_n\|_p = 1$ for each vector e_n of the standard basis for ℓ^p , it's enough to show that:

If $1 \leq q < p < \infty$ and $T: \ell^p \rightarrow \ell^q$ is a continuous linear transformation, then $\|Te_n\|_q \rightarrow 0$.

We will prove this result in contrapositive form:

If $1 \leq p, q < \infty$ and $T: \ell^p \rightarrow \ell^q$ is a continuous linear transformation for which $\|Te_n\|_q \not\rightarrow 0$, then $p \leq q$.

We are assuming that some subsequence of $(\|Te_n\|_q)$ stays bounded away from zero. By replacing the standard basis (e_n) for ℓ^p by the corresponding subsequence and renaming, we may assume that there exists $\delta > 0$ such that

$$(6) \quad \|Te_n\|_q \geq \delta \quad \text{for each } n \in \mathbb{N}.$$

We call T an *isomorphism* of X onto Y . By the Open Mapping Theorem for Banach spaces, the continuity of T^{-1} follows from that of T (see, e.g., [1, Theorems 6.81 & 6.83, pp.186–188]).

N.B. X isomorphic to Y implies Y isomorphic to X .

Theorem 8.2 is trivially true if $p = \infty$ since in this case ℓ^q (and every closed subspace) is separable but ℓ^p is not.

The linear map $Y^* \rightarrow X^*$ given by $\psi \rightarrow \psi \circ T$ is denoted by T^* and called the *adjoint* of T . The adjoint of T is continuous for the norm topologies of these dual spaces, and has operator-norm $= \|T\|$ (see, e.g., [2, page 15]).

In other words: if a continuous linear map squeezes the larger space ℓ^p into the smaller one ℓ^q , then it squeezes the standard unit-vector basis toward the zero-vector.

If $p = 1$ then $p \leq q$ regardless of the behavior of $\|Te_n\|_q$, so we may assume $p > 1$. The standard basis sequence (e_n) converges weakly to 0 in ℓ^p (Proposition 2.2), so $Te_n \rightarrow 0$ weakly in ℓ^q by Proposition 8.3 above. In particular, $Te_n \rightarrow 0$ pointwise on \mathbb{N} .

Consequently the Gliding Hump Lemma (Lemma 7.2) when applied to the sequence $(|Te_n|^q)$, provides:

- (a) A subsequence $n_k \nearrow \infty$, and
- (b) an exhaustion of \mathbb{N} by a sequence (I_k) of contiguous intervals,
- (c) For each $k \in \mathbb{N}$ a ‘‘Gliding Hump’’ decomposition of $f_k := Te_{n_k}$ as

$$(7) \quad Te_{n_k} = h_k + r_k \quad (k \in \mathbb{N})$$

where $h_k = f_k \chi_{I_k}$ is a ‘‘hump function’’, r_k is the ‘‘remainder’’, and

$$(8) \quad \|h_k\|_q = (1 - o(1))\|Te_{n_k}\| \quad \text{with} \quad \sum_k \|r_k\|_q < \infty.$$

To complete the argument, we make two estimates involving the sequence (s_K) of ‘‘test vectors’’ defined by

$$s_K = \sum_{k=1}^K e_{n_k} \quad (K \in \mathbb{N}).$$

The first of these just involves the continuity of our transformation T .

$$(9) \quad \|Ts_K\|_q \leq \|T\| \|s_K\|_p = \|T\| K^{1/p}$$

The second estimate goes in the other direction, and requires our Gliding Hump decomposition (7) of the functions $f_k = Te_{n_k}$. We have for each $K \in \mathbb{N}$:

$$(10) \quad \begin{aligned} \|Ts_K\|_q &= \left\| \sum_{k=1}^K Te_{n_k} \right\|_q = \left\| \sum_{k=1}^K f_k \right\|_q \\ &= \left\| \sum_{k=1}^K h_k + \sum_{k=1}^K r_k \right\|_q \\ &\geq \left\| \sum_{k=1}^K h_k \right\|_q - \left\| \sum_{k=1}^K r_k \right\|_q \\ &\geq \left(\sum_{k=1}^K \|h_k\|_q^q \right)^{1/q} - \sum_{k=1}^K \|r_k\|_q \end{aligned}$$

where the last inequality rests upon the pairwise disjointness of the ‘‘hump function’’ supports.

Upon applying the estimates (8) to inequality (10) we obtain the asymptotic estimate

$$(11) \quad \|Ts_K\|_q \geq \left(\sum_{k=1}^K \|f_k\|_q^q (1 - o(1)) \right)^{1/q} - O(1) \quad (K \rightarrow \infty)$$

Finiteness of $\sum_k \|r_k\|_q$ is insured by by choosing the sequence (ε_k) in the proof of the Gliding Hump Lemma so that $\sum_k \varepsilon_k < \infty$.

which, in view of our hypothesis $\|f_k\|_q > \delta > 0$, yields:

$$\|T\| K^{1/p} \geq \|T s_K\|_q \geq \delta \left(\sum_{k=1}^K 1 - o(1) \right)^{1/q} - O(1).$$

Thus for all K sufficiently large,

$$\|T\| K^{1/p} \geq \frac{\delta}{2} K^{1/q} \quad \text{i.e.,} \quad \frac{2\|T\|}{\delta} \geq K^{1/q-1/p},$$

which requires the exponent of K in the last inequality to be ≤ 0 .

Thus $p \leq q$, as we wished to show. \square

Remarks 8.4. (a) Our argument actually shows that if the continuous linear map T maps ℓ^p into ℓ^q with $p > q$, then T is not an isomorphism onto the closure of its range. Thus the larger space ℓ^p is not isomorphic to *any subspace* of the smaller one ℓ^q .

With a bit more care, this argument can be tweaked to show that if $p \neq q$ then *no infinite dimensional closed subspace of ℓ^p can be isomorphic to a closed subspace of ℓ^q .*

(b) A small modification of the argument shows that $T: \ell^p \rightarrow \ell^q$ must be a *compact operator*.¹¹

¹¹ i.e., it must take bounded sets into relatively compact ones.

9 References

1. Sheldon Axler, *Measure, Integration, and Real Analysis*,¹² Springer 2020. Open access: downloadable at <https://measure.axler.net>
2. Neil L. Carothers, *A Short Course on Banach Space Theory*¹³, London Mathematical Society Student Texts #64, 2005.
3. Pete L. Clark, *Topology*¹⁴. Lecture notes downloadable at: <http://alpha.math.uga.edu/~pete/pointset.pdf>.
4. Harry Dym and Victor Katsnelson, *Contributions of Issai Schur to Analysis*¹⁵, arXiv, June 2007. Preprint downloadable at: <https://arxiv.org/pdf/0706.1868.pdf>
5. John L. Kelley, *General Topology*¹⁶, van Nostrand 1955. Reprinted by Dover 2017.
6. Robert E. Megginson, *An Introduction to Banach Space Theory*, Springer 1998.
7. Issai Schur, *Über lineare Transformationen in der Theorie der unendlichen Reihen*, J. Reine Angew. Math 151 (1920), 79-111.

¹² See especially Chapter 7 for completeness and duality of ℓ^p spaces, done in a more general setting.

¹³ A readable introduction, for non-specialists, to Banach space theory.

¹⁴ Especially Chapter 5: *Convergence*.

¹⁵ A 100 page survey of Schur's many contributions to analysis. Surprisingly, it omits our "Schur's Lemma".

¹⁶ Especially Chapter 2: *Moore-Smith Convergence*.

ACKNOWLEDGMENTS. I wish to thank the participants in the Portland State Analysis Seminar, especially Sheldon Axler, Jim Rulla, and Robert Lyons, for their insightful comments and much-needed corrections, all of which greatly improved these notes.