21. Weak convergence and convexity

Weak convergence is an extremely useful tool in analysis; particular, PDE theory and optimization. Indeed, differential equations are usually solved by first looking for weak solutions (to be discussed later) while functionals are often easier to minimize along weakly convergent sequences as weak limits are easier to achieve due to coarser topology. (For the same reasons, functionals are often less continuous in the weak topology.)

To keep the discussion more specific, consider a functional $\Phi: L^p \to \mathbb{R}$ (not necessarily linear) and suppose that we want to find $f \in L^p$ with $\|f\|_p \leq 1$ that minimizes $\Phi$. An example of such a functional is

$$\Phi(f) := \|f - f_0\|_1$$

(21.1)

where $f_0 \in L^1$, which is well defined (due to $L^p \subseteq L^1$) whenever the underlying measure is finite. The question is: Writing $B_1 := \{f \in L^p: \|f\|_p \leq 1\}$, the question is: What is $\inf_{f \in B_1} \Phi(f)$? And is there a minimizer?

Here is a natural idea: Assuming boundedness of $\Phi$ from below (for otherwise we have nothing to say) permits us to extract a minimizing sequence $\{f_n\}_{n \geq 1} \subseteq L^p$ such that $\Phi(f_n) \to \inf \Phi$. The question is now how to extract a limit point from this sequence. Here the following observation may come handy:

**Theorem 21.1** (Weak sequential compactness) Suppose $1 < p < \infty$, let $q$ be such that $p^{-1} + q^{-1} = 1$ and assume $L^q$ is separable. Then every bounded sequence in $L^p$ contains a weakly convergent subsequence. In particular, the unit ball in $L^p$ is weakly sequentially compact.

**Proof.** Let $\{f_n\}_{n \geq 1} \subseteq L^p$ be a sequence with (WLOG) $\|f_n\|_p \leq 1$ for all $n \geq 1$. Thanks to the Riesz representation theorem, $L^q$ is isometric to $(L^p)^*$ and so $(L^p)^*$ contains a countable dense subset which we enumerate into $\{\phi_i\}_{i \geq 1}$. Since $\|f_n\| \leq 1$ implies $|\phi(f_n)| \leq \|\phi\|$ for all $n \geq 1$, the Cantor diagonal argument permits us to choose a subsequence $\{n_k\}_{k \geq 1}$ such that

$$\forall i \geq 1: \lim_{k \to \infty} \phi_i(f_{n_k}) \text{ exists in } \mathbb{R}. \quad (21.2)$$

Next observe that, by $\|f_{n_k}\| \leq 1$ again, we have $|\phi(f_{n_k}) - \phi_i(f_{n_k})| \leq \|\phi - \phi_i\|$ and so, since $\{\phi_i\}_{i \geq 1}$ is dense in $(L^p)^*$, for each $\phi \in (L^p)^*$ and each $\epsilon > 0$ there is $i \geq 1$ such that $\|\phi - \phi_i\| < \epsilon$. Then $|\phi_i(f_{n_k}) - \phi(f)| < \epsilon$ and so, by a $3\epsilon$-argument

$$\forall \phi \in (L^p)^*: \quad F(\phi) := \lim_{k \to \infty} \phi(f_{n_k}) \text{ exists in } \mathbb{R}. \quad (21.3)$$

The estimate $|\phi(f_{n_k})| \leq \|\phi\|$ also implies

$$\forall \phi \in (L^p)^*: \quad |F(\phi)| \leq \|\phi\| \quad (21.4)$$

and so $F \in (L^p)^{**}$. But $L^p$ is reflexive and so $F(\phi) = \phi(f)$ for some $f \in L^p$. Hence $f_{n_k} \to f$ weakly in $L^p$. \qed

The proof has been written to make the reliance on $L^p$ light. The argument does not work for $p = 1$ because then $L^p$ is not reflexive and $L^\infty$ is, once infinitely dimensional, not separable. This is bolstered by a counterexample: For $L^1(\mathbb{R}, \lambda)$, where $\lambda$ is the Lebesgue measure, let $f_n := n1_{[0,1/n]}$. Then $\|f_n\|_1 = 1$. As $\{f_n\}_{n \geq 1}$ vanishes outside
intervals shrinking to a point, if a subsequence converges weakly then only to 0. But
\( \phi(f) := \int f \, d\lambda \) is continuous and \( \phi(f_n) = 1 \) for all \( n \geq 1 \), a contradiction.

Theorem 21.1 has a general version that is called the \textit{Banach-Alaoglu Theorem} which
we will discuss later. That theorem deals with weak-\( * \) topology and, to get the above,
requires reflexivity of the underlying space. Separability of the dual permits us to restrict
attention to sequential compactness.

Returning to the problem of minimizing a functional \( \Phi \) over a unit ball in \( L^p \), we thus
see that getting a weakly convergent minimizing subsequence is fairly cheap. The next
question is:

\textbf{Why is the weak limit contained in the unit ball?}

This is answered in:

**Lemma 21.2** Let \( 1 < p < \infty \). If \( \{ f_n \}_{n \geq 1} \subseteq L^p \) is weakly convergent to some \( f \in L^p \), then
\[
\| f \|_p \leq \liminf_{n \to \infty} \| f_n \|_p. \tag{21.5}
\]

\textit{Proof.} For each \( \phi \in (L^p)^* \) we have
\[
| \phi(f) | \leq \liminf_{n \to \infty} | \phi(f_n) | \leq \| \phi \| \liminf_{n \to \infty} \| f_n \|_p. \tag{21.6}
\]

Assuming \( f \neq 0 \) (otherwise the result holds trivially) and letting \( \phi := \phi_g \) for \( g := |f|^{p-2}f \)
we get \( \phi_g(f) = \| \phi_g \| \| f \|_p \). Canceling \( \| \phi_g \| = \| f \|^{p-1} \neq 0 \) we get the claim. \( \square \)

Again the statement and the proof apply in full generality of Banach spaces except
that the Hahn-Banach theorem is needed to complete the last step.

Having ensured that the weak limit \( f \) of the subsequence \( \{ f_{n_k} \}_{k \geq 1} \) lies in the domain
of interest, the next question is:

\textbf{Why is} \( \Phi(f) \) \textbf{equal to} \( \lim_{k \to \infty} \Phi(f_{n_k}) \)?

This will work if \( \Phi \) is \textit{weakly continuous}, or at least \textit{weakly lower-continuous}, in the norm
topology, but that is often too much to ask because weaker topology makes continuity
harder to satisfy. One case where weak convergence can be traded for another regularity
property is when \( \Phi \) is convex. Indeed, convex \( \Phi \) is decreased (or at least not increased)
by taking convex combinations and so the following lemma is very useful:

**Lemma 21.3** (Mazur’s lemma) Let \( 1 < p < \infty \) and suppose \( \{ f_n \}_{n \geq 1} \subseteq L^p \) and \( f \) are
such that \( f_n \rightharpoonup f \) weakly in \( L^p \). Then for each \( n \geq 1 \) there are \( \{ a_i^{(n)} : i \geq 1 \} \subseteq [0,1] \)
with \( \{ i \geq 1 : a_i^{(n)} \neq 0 \} \) finite and \( \sum_{i \geq 1} a_i^{(n)} = 1 \) such that
\[
\sum_{i \geq 1} a_i^{(n)} f_i \rightharpoonup f \quad \text{in} \quad L^p. \tag{21.7}
\]

\textit{Proof.} Let \( C \) be the norm-closure in \( L^p \) of the set of all convex combinations of finite
subsets of elements from \( \{ f_n : n \geq 1 \} \). It then suffices to show that \( f \in C \). Suppose this
fails, i.e., \( f \notin C \). Proposition 19.15 yields the existence of \( h_0 \in C \) such that
\[
\| f - h_0 \|_p^p = \inf_{h \in C} \| f - h \|_p^p > 0. \tag{21.8}
\]
As \( h_t := h_0 + t(h - h_0) \in C \) for all \( h \in C \) and all \( t \in [0, 1] \) implies \( \|f - h_t\|_p^p \geq \|f - h_0\|_p^p \), Lemma 20.7 gives

\[
\forall h \in C: \quad \int |f - h_0|^{p-2}(f - h_0)(h - h_0) \, d\mu \leq 0. \tag{21.9}
\]

Denoting \( g := |f - h_0|^{p-2}(f - h_0) \), which lies in \( L^q \), we get \( \phi_g \in (L^p)^\ast \) with

\[
\forall h \in C: \quad \phi_g(h) \leq \phi_g(h_0) \tag{21.10}
\]

while a calculation shows

\[
\phi_g(f) = \phi_g(f - h_0) + \phi_g(h_0) = \|f - h_0\|_p^p + \phi_g(h_0) > \phi_g(h_0). \tag{21.11}
\]

This shows that \( f \) cannot be a weak limit of any sequence from \( C \).

We note that Lemma 21.3 is a special case of the Banach-Saks Theorem, which claims that (21.7) can be improved to Cezaro averages

\[
\frac{1}{n} \sum_{k=1}^{n} f_{n_k} \underset{n \to \infty}{\longrightarrow} f \quad \text{in} \quad L^p, \tag{21.12}
\]

where \( \{f_{n_k}\}_{k \geq 1} \) is a subsequence from \( \{f_n\}_{n \geq 1} \). S. Banach and S. Saks proved this for Hilbert spaces \( (p = 2) \), where the inner product comes to help, but S. Mazur apparently generalized it to all Banach spaces as well. Mazur’s lemma applies in full generality of all Banach spaces as well.

An argument in the proof is worthy a note:

**Corollary 21.4 (Separation from a closed convex set)** Let \( 1 < p < \infty \) and let \( C \subseteq L^p \) be non-empty, closed and convex. Then

\[
\forall f \in L^p \setminus C \exists \phi \in (L^p)^\ast \exists a \in \mathbb{R}: \quad \phi(f) > a \quad \land \quad \phi \leq a \text{ on } C. \tag{21.13}
\]

**Proof.** This is the content of (21.10–21.11) with \( a := \phi_g(h_0) \).

The arguments used in the proof can be boosted further into another separation claim whose proof we will leave to homework:

**Lemma 21.5 (Separation of convex sets by a linear functional)** Let \( 1 < p < \infty \) and let \( C, C' \subseteq L^p \) be two non-empty closed and convex sets such that \( C \cap C' = \emptyset \). Then

\[
\exists \phi \in (L^p)^\ast \exists a < b: \quad \phi < a \text{ on } C \quad \land \quad \phi > b \text{ on } C'. \tag{21.14}
\]

In particular, \( C \) and \( C' \) are separated by the closed linear subspace \( \{f \in L^p: \phi(f) = \frac{a+b}{2}\} \).

Returning to the main line of discussion, next we note the following consequence of Corollary 21.4 and thus also Mazur’s lemma:

**Theorem 21.6 (Mazur’s theorem)** Suppose \( 1 < p < \infty \). Then any norm-closed convex subset of \( L^p \) is weakly closed.

**Proof.** Suppose \( C \subseteq L^p \) is convex, norm-closed (and non-empty because otherwise there is nothing to prove). If \( f \) lies in the weak closure of \( C \), then for every \( \phi \in (L^p)^\ast \) the value
\(\phi(f)\) lies in the closure of \(\{\phi(h) : h \in C\} \subseteq \mathbb{R}\). This rules out that \(\phi(f) > \sup_{h \in C} \phi(h)\) and so we cannot have \(f \notin C\), by contrapositive to Corollary 21.4. Hence \(f \in C\) as well. \(\square\)

While the converse (a subset of \(L^p\) is norm-closed only if it is weakly closed) holds generally, due to the peculiarities of the weak topology, the conclusion fails in general without the convexity assumption. Indeed, non-empty weakly-open sets are generally unbounded which is not true for norm-open sets. So generic norm-closed sets are not weakly closed.

Returning to the example (21.1), where we assume that \(f_0\) does not belong to the underlying unit ball \(B_1\) in \(L^p\) (for otherwise \(f \rightarrow \Phi(f)\) is minimized by \(f := f_0\)), the above now gives the following: Theorem 21.1 gives us a sequence \(\{f_n\}_{n \geq 1} \subseteq B_1\) and \(f \in L^p\) such that

\[
\lim_{n \to \infty} \Phi(f_n) = \inf_{f \in B_1} \Phi(f) \land f_n \overset{w}{\rightharpoonup} f. \tag{21.15}
\]

Lemma 21.2 ensures that \(f \in B_1\) but the same can be achieved in more generality by Mazur’s lemma, by which convex combinations \(h_n := \sum_{i \geq n} \alpha_i^{(n)} f_i\) (which lie in \(B_1\)) converge obey

\[
h_n \rightarrow f \text{ in } L^p. \tag{21.16}
\]

As any finite number of terms in the sequence \(\{f_n\}_{n \geq 1}\) do not matter, we may in fact assume that \(\alpha_i^{(n)} = 0\) for all \(i = 1, \ldots, n\). Since \(\Phi\) is convex by an application of the triangle inequality, we have

\[
\Phi(h_n) \leq \sum_{i \geq n} \alpha_i^{(n)} f_i \tag{21.17}
\]

and so \(\Phi(h_n) \rightarrow \inf_{B_1} \Phi\). The convergence in \(L^p\) implies (under our assumption that \(L^1 \subseteq L^p\)) convergence in \(L^1\) and, since \(\Phi\) is continuous in \(L^1\)-topology, we conclude \(\Phi(f) = \inf_{B_1} \Phi\), i.e., \(\Phi\) admits a minimizer.

We will finish by two observations that are both of independent interest and relevant for the above problem. The first of these is the fact that there is a simple criterion that permits us to "upgrade" weak convergence to norm convergence:

**Theorem 21.7 (Radon-Riesz theorem)** Suppose \(1 < p < \infty\) and let \(\{f_n\}_{n \geq 1} \subseteq L^p\) and \(f \in L^p\). Then

\[
f_n \overset{w}{\rightharpoonup} f \land \|f_n\|_p \to \|f\|_p \Rightarrow f_n \overset{n \to \infty}{\longrightarrow} f \text{ in } L^p \tag{21.18}
\]

**Proof.** By adding an arbitrary function to \(f_n\) and \(f\) we may assume that \(\|f_n\|_p \neq 0\) and \(\|f\|_p \neq 0\). Define \(f'_n := f_n / \|f_n\|_p\) and \(f' := f / \|f\|_p\). Then \(\|f_n\|_p \to \|f\|_p\) ensures that also \(f'_n \rightarrow f'\) weakly in \(L^p\). If \(f'_n\) does not converge to \(f'\) in norm, then (by resorting to a subsequence if necessary), we may assume \(\inf_{n \geq 1} \|f'_n - f'\|_p > 0\). But then the uniform convexity of \(L^p\) forces

\[
\sup_{n \geq 1} \left\| \frac{f'_n + f'}{2} \right\|_p \leq 1 - \delta \tag{21.19}
\]
and so, for any \( f \in P_L \),

\[
|\phi(f')| = \lim_{n \to \infty} \left| \phi \left( \frac{f_n + f'}{2} \right) \right| \leq (1 - \delta) \|\phi\|.
\] (21.20)

Taking \( \phi := \phi^g \) for \( g := |f'|^{p-2}f' \) gives \( \phi^g(f') = \|\phi^g\| \|f'\|_p \) which contradicts (21.20) because \( \|f'\|_p = 1 \).

Theorem 21.7 goes back to F. Riesz and J. Radon; the spaces for which (21.18) holds are called Radon-Riesz spaces or (per Wikipedia page) “spaces with Kadets-Klee property,” after M.I. Kadets’ and V.L. Klee’s work in late 1920s. The original derivation was done in the context of Hilbert spaces; the extension to uniformly convex Banach spaces follows the above lines except that the Hahn-Banach theorem is needed in the last step. Note that Lemma 21.3 and Theorem 21.7 are both based on convexity.

To demonstrate further that convexity is closely tied to weak convergence, we note an important result from variational calculus:

**Theorem 21.8 (Tonelli’s theorem)** For \( O \subseteq \mathbb{R}^d \) open and convex, \( \varphi: O \to [0, \infty) \) continuous and \( p \in (1, \infty) \), define a functional \( \Phi \) on \( L^p(O, \mathcal{B}(O), \lambda) \) by

\[
\Phi(f) := \int_O \varphi \circ f \, d\lambda
\] (21.21)

Then \( \Phi \) is weakly lower-semicontinuous if and only if \( \varphi \) is convex.

We will prove only the easier direction which works for an arbitrary measure on a general normed convex space (which we need to state the convexity of \( \varphi \)):

**Proof of “if part”.** Suppose that \( \varphi \) is convex and \( \{f_n\}_{n \geq 1} \) such that \( f_n \to f \) weakly. There exists a subsequence \( \{f_{k_i}\}_{i \geq 1} \) such that

\[
L := \liminf_{n \to \infty} \Phi(f_n) = \lim_{i \to \infty} \Phi(f_{k_i})
\] (21.22)

exists. Mazur’s lemma yields existence of convex combinations \( h_n := \sum_{i \geq n} \alpha_i^{(n)} f_{k_i} \) such that \( h_n \to f \) in norm. The convexity of \( \varphi \) then gives

\[
\Phi(h_n) \leq \sum_{i \geq n} \alpha_i^{(n)} \Phi(f_{k_i})
\] (21.23)

and so, by (21.22),

\[
\limsup_{n \to \infty} \Phi(h_n) \leq L.
\] (21.24)

As convergence in \( L^p \)-norm implies convergence of a subsequence almost everywhere, Fatou’s lemma gives

\[
\liminf_{n \to \infty} \Phi(h_n) \geq \Phi(f).
\] (21.25)

Combining (21.24–21.25) we get \( \Phi(f) \leq \liminf_{n \to \infty} \Phi(f_n) \) and so \( \Phi \) is indeed weakly lower-semicontinuous.

The above result is another fundamental contribution of L. Tonelli to analysis. While the integral form of the functional is quite typical in applications, the requirement of convexity of \( \varphi \) is often too demanding. To overcome lack thereof one needs to supply additional observations on the regularity of the minimizing sequence.