Riesz-Thorin Interpolation & Applications

Anson Li[1], Mohamed-Amine Azzouz[1]

[1]Department of Electrical and Computer Engineering, McGill University
ting.h.li@mail.mcgill.ca
mohamed-amine.azzouz@mail.mcgill.ca

Contents

1. Preliminaries

Notation and Measure Theory

- $\mathbb{N} = \{0, 1, 2, \ldots\}$ is the set of *natural numbers* (which includes 0),
- $\mathbb{N}^+ = \{1, 2, \ldots\}$ is the set of *counting numbers*,
- $\mathbb{Z} = \{0, -1, +1, \ldots\}$ is the set of *integers*,
- $\mathbb{Q}$ is the set of *rational numbers*,
- $\mathbb{R}$ is the set *real numbers*,
- $\mathbb{C}$ is the set of *complex numbers*, and
- $\mathbb{R}^+ = [0, +\infty)$ are the non-negative reals.

Suppose that X and Y are normed vector spaces over $\mathbb{C}$.

- $\mathcal{L}(X, Y)$ = linear maps $T : X \to Y$.
- $L(X, Y)$ = cont. linear maps $T : X \to Y$.
- X^* = continuous dual of X

Let $(X, \mathcal{M})$ be a measurable space.

- $\mathcal{E}(\mathcal{M})$ = complex-valued meas. functions on X.

Sequences will be indexed by $\mathbb{N}^+$ and will be assumed to take values in a vector space. We say that x_n is *finitely supported* if $x_n = 0$ for all sufficiently large n.

- l_0 = sequences that are finitely supported,
- c_0 = sequences that converge to 0 (range is normed),
- l^+ = real-valued sequences $x_n \geq 0$ for $n \geq 1$.
- l^{++} = real-valued sequences $x_n > 0$ for $n \geq 1$.

We will always assume that all measures are σ-finite. For most of the theorems however, this assumption is not necessary. Let μ be a σ-finite measure on $\mathcal{M}$.

- A measurable function $\phi : X \to \mathbb{C}$ is a *simple function* if its range $\phi(X)$ is a finite subset of $\mathbb{C}$. Every simple function is of the form

  $$\phi = \sum_1^n c_j \chi_{E_j} \quad \text{where} \quad c_j \in \mathbb{C}, E_j \in \mathcal{M},\ j = 1, \ldots, n.$$
- $\Sigma(\mu)$ = v.s. of simple functions,

- $\Sigma_0(\mu)$ = v. subspace of Σ, vanish outside a μ-finite set
$$\Sigma_0(\mu) = \{\phi \in \Sigma,\ \mu(\{\phi \neq 0\}) < \infty\}.$$
- $L^\infty(\mu)$ = B. space of *essentially bdd* functions, with norm
$$\|f\|_\infty = \inf\Big\{a \geq 0, \mu(\{|f| > a\}) = 0\Big\} \quad \forall f \in L^\infty(\mu)$$
- $L^\infty_0(\mu)$ = v. subspace of $L^\infty(\mu)$, vanish outside a μ-finite set.
- $L^+(\mu)$ = meas. functions, values in $[0,\infty]$.[1]

The subsets $\Sigma^+(\mu)$, and $\Sigma_0^+(\mu)$ are defined as positive cones of $\Sigma(\mu)$ and $\Sigma_0(\mu)$, whose elements in $\Sigma^+(\mu)$ and $\Sigma_0^+(\mu)$ take values in $[0,\infty)$. We recall that if $g \in L^+(\mu)$, the integral with respect to μ is the supremum over the integral of its non-negative simple subordinates.
$$\int_X g(x)\,dx = \sup\Big\{\int_X \phi(x)\,dx,\ \phi \leq g,\ \phi \in \Sigma^+\Big\} \tag{1}$$
If μ is σ-finite or $\operatorname{supp}(g) = \{x \in X,\ g(x) \neq 0\}$ is σ-finite, we can shrink the family of lower approximants on the right hand side of the Equation (1).
$$\int_X g(x)\,dx = \sup\Big\{\int_X \phi(x)\,dx,\ \phi \leq g,\ \phi \in \Sigma_0^+\Big\}. \tag{2}$$

Lemma 1.1 [Fol13, Thm 2.10]
Let $(X,\mathcal{M})$ be a measurable space, if $g : X \to \mathbb{C}$ is measurable, there exists a sequence of simple functions $\{\phi_n\} \subseteq \Sigma$, such that
1. $\phi_n \to g$ pointwise,
2. $|\phi_n| \nearrow |g|$ pointwise.
3. If μ is σ-finite, or $\mu(\{|f| > \varepsilon\}) < \infty$ for all ε, we can take $\{\phi_n\} \subseteq \Sigma_0$.

Any sequence satisfying first two properties in Lemma 1.1 is said to be an *increasing sequence of subordinates* of g. If $\{f_n\}$ is any such sequence, then $\int |f_n(x)|\,dx \nearrow \int |g(x)|\,dx$.

Let p be a number in $[1,+\infty)$,
$$\mathcal{L}^p(\mu) = \{f \in \mathcal{M},\ \textstyle\int_X |f(x)|^p\,dx < +\infty\}, \quad \text{and}$$
$$L^p(\mu) = \mathcal{L}^p(X,\mu)/\text{pw a.e.}$$
If $f \in \mathcal{L}^p(\mu)$ or $L^p(\mu)$, we write
$$\|f\|_p = \big(\textstyle\int_X |f(x)|^p\,dx\big)^{1/p}. \tag{3}$$
It is clear that $L^p(\mu)$ is a B. space for $p \in [1,\infty]$.

1.1. Local Integrability

Definition 1.2
A measurable function $g : X \to \mathbb{C}$ is *locally integrable* if
$$\int_E |g(x)|dx < \infty \quad \text{for all} \quad E \in \mathcal{M},\ \mu(E) < \infty.$$
The v.s. of all locally integrable functions (modulo a.e. sets) is denoted by $L^1_{loc}(\mu)$.

If $g \in L^p$ for $p \in (1,\infty]$ then we can identify $g \in L^1_{loc}$ because
$$\int_E |g(x)|dx \leq \int_E (|g(x)|^p + 1)\,dx < \infty. \tag{4}$$
Let $g \in L^1_{loc}$, consider the *scalar product* on L^∞_0
$$\langle g, f\rangle = \int_X g(x)f(x)dx \quad \forall f \in L^\infty_0. \tag{5}$$
The integral in Equation (5) converges absolutely, since
$$|g(x)f(x)| \leq \|f\|_\infty \chi_{\{f\neq 0\}}|g(x)| \in L^1.$$
Given a locally integrable function g, the next proposition shows that its L^q norm (which is possibly equal to ∞) can be computed through its scalar product on L^∞_0.

Proposition 1.3 [Fol13, Thm 6.13, 6.14]
Given a locally integrable g, it is in L^q iff[a]
$$M_q(g) = \sup\Big\{|\langle g, f\rangle|,\ f \in L^\infty_0,\ \|f\|_p = 1\Big\} < \infty, \tag{6}$$
and if this is the case, then $\|g\|_q = M_q(g)$.

[a] implicit $q \in [1,\infty]$, p conjugate to q

Proof of Proposition 1.3. We can break up proposition into two statements and prove them separately.

(i) If $g \in L^q$, then $M_q(g) \leq \|g\|_q$.
(ii) If $g \in L^1_{loc}$ and $M_q(g) < \infty$, then $g \in L^q$ and $\|g\|_q \leq M_q(g)$.

Statement i) says that if $g \in L^q$, its scalar product on $\{f \in L^\infty_0, \|f\|_p = 1\}$ is uniformly bounded by its L^q norm. This follows from Hölder's inequality:
$$|\langle g, f\rangle| \leq \langle |g|, |f|\rangle \leq \|f\|_p\|g\|_q, \quad \text{for all} \quad f \in L^\infty_0.$$
Hence **Statement i)** is proven. We turn to the proof of **Statement ii)**.
1. If $q \in [1,\infty)$, we know that $\operatorname{supp}(g)$ is always σ-finite (in the case where μ is not σ-finite, see Lemma 3.15). Let $\{\phi_n\} \subseteq \Sigma_0$ be an increasing sequence of subordinates of g (justified by Lemma 1.1). Then,
$$\int |g(x)|^q\,dx = \sup_n \int_X |\phi_n(x)|^q\,dx.$$
To show that $\|g\|_q \leq M_q(g)$, it suffices[2] to con-

[1] $[0,\infty]$ has the extended Borel σ-algebra. See [Fol13, Chap 2.1, 2.2, Ex 2.1, 2.2, 2.4]

[2] this is because $t \mapsto t^{1/q}$ ($t \in [0,\infty]$, $q \in [1,\infty)$ satisfies the hypotheses of Lem. 3.8

trol the asymptotics of $\int |\phi_n(x)|^q \, dx$ (meaning $\liminf \int |\phi_n(x)|^q \, dx \le M_q(g)$).

$$\|\phi_n\|_q = \frac{\int_X |\phi_n(x)|^q \, dx}{\|\phi_n\|_q^{q-1}} \tag{7}$$

For each $n = 1, 2, \ldots$, let us define

$$\psi_n(x) = \frac{|\phi_n(x)|^{q-1}}{\|\phi_n\|_q^{q-1}} (\overline{\operatorname{sgn} g}) \in L_0^\infty, \tag{8}$$

An easy calculation will show that $\|\psi_n\|_p = 1$. We can bound $\|\phi_n\|_q$ using the scalar product $\langle g, \psi_n \rangle$ (see Equation (7)) because each ϕ_n is a subordinate of g. This is accomplished by 'stealing' a factor of $|\phi_n(x)|$ under the integral sign.

$$\begin{aligned} \|\phi_n\|_q &= \frac{\int_X |\phi_n(x)|^q \, dx}{\|\phi_n\|_q^{q-1}} \\ &\le \frac{\int_X |\phi_n(x)|^{q-1} (\overline{\operatorname{sgn} g}) g(x) \, dx}{\|\phi_n\|_q^{q-1}} \\ &\le M_q(g) \end{aligned}$$

2. If $q = \infty$, it is fruitful to consider an equivalent characterization of $\|g\|_\infty$.

$$\|g\|_\infty = \sup \Big\{ a \ge 0, \mu(\{x \in X, \ |g(x)| \ge a\}) > 0 \Big\}. \tag{9}$$

If $\|g\|_\infty = 0$, then there is nothing to prove. So in the case that $\|g\|_\infty > 0$, we consider the lower approximants to $\|g\|_\infty$. Let $0 < \varepsilon < \|g\|_\infty$, the set $A_\varepsilon = \{|g| > \varepsilon\}$ must have positive measure (possibly ∞). Our measure μ is semifinite, so we find a measurable subset B of A with

$$0 < \mu(B) < \infty.$$

We **note** that B is a subset of A_ε, so $|g(x)| \ge \varepsilon$ for a.e. $x \in B$. To show that $\varepsilon \le M_q(g)$, we will evaluate g using the averaging operator

$$f_B = \mu(B)^{-1} \chi_B (\overline{\operatorname{sgn} g}).$$

It is easy to see that $f_B \in L_0^\infty$, and $\|f_B\|_1 = 1$. The scalar product of f_B with g gives the expression

$$\langle g, f_B \rangle = \frac{1}{\mu(B)} \int_B |g(x)| \, dx.$$

Our previous **note** tells us that $\varepsilon \le \langle g, f_B \rangle \le M_q(g)$ for all $\varepsilon < \|g\|_\infty$. Sending ε towards $\|g\|_\infty$ proves **Statement ii)**.

■

Few remarks on the proof of Proposition 1.3. Suppose that we have a quantity $b \in \overline{\mathbb{R}}$ that is defined by some kind of limiting process (i.e. $b = \lim b_n$, $b = \sup B^-$, $b = \inf B^+$, and so on). The task of estimating the size of b can always be reduced to the task of bounding or tagging along the terms that lead up to b. More precisely, suppose $\{a_n\}$, $\{b_n\}$, and $\{c_n\}$ are $\overline{\mathbb{R}}$-valued sequences, where $b_n \to b \in \overline{\mathbb{R}}$. If $a_n \le b_n \le c_n$ eventually, then

$$\limsup a_n \le b \le \liminf c_n. \tag{10}$$

[Rud76, Thm 3.19] contains a proof of Equation (10). By giving ourselves some wiggle room, we can pick any sequence $\{b_n\}$ as long as it defines the same limit b. We can also think of the extended reals $\overline{\mathbb{R}}$ as an equivalence class of monotone, $\mathbb{R}$-valued sequences. In this case defining sequence can be taken to be monotone. By passing to a subsequence, we can also take $\{b_n\}$ to be Cauchy with a specified rate of decay and integrable whenever $b \in \mathbb{R}$.

Specializing, elementary integral estimates for $\int g(x) \, dx$, where $g \in L^+$ can be obtained if we control the integrals over the subordinates of g. The m.c.t. tells us that we can choose any sequence of subordinates $\{\phi_n\} \subseteq L^+$, as long as $\phi_n \nearrow g$ pointwise a.e. The construction of a sufficiently well-behaved sequence $\{\phi_n\}$ depends heavily upon the problem you are working on: μ-invariant group actions, properties of g (boundedness, growth), properties of the domain X (topology, σ-finiteness), etc; just to list a few things that one can leverage if available. In the case of $q \in [1, \infty)$ in the proof of Statement i), our choices of b_n and c_n are $b_n = \|\phi_n\|_q \nearrow \|g\|_q$, and $c_n = \langle g, \psi_n \rangle \ge b_n$, respectively.

1.2. Convergence in Measure

A sequence $\{f_n\} \subseteq \mathcal{E}(\mathcal{M})$ is said to *converge in measure* (abbrev. conv. in meas.) to $f \in \mathcal{E}(\mathcal{M})$ if for every ε and $\delta > 0$, there exists $N \in \mathbb{N}^+$ such that

$$\mu(\{x \in X, \ |f_n(x) - f(x)| > \varepsilon\}) \le \delta \quad \text{whenever} \quad n \ge N.$$[3]

If $f_n \to f$ in meas., we think of f as the best possible representative (modulo a.e. sets) of $\{f_n\}$. Specifically, but still on an informal level, the relation between different types of convergences is as follows.

	conv. in norm ($1 \le p \le \infty$)
implies	conv. in meas.,
which implies	subseq. conv. pw a.e.

Proposition 1.4

Let $f_n, f, g : X \to \mathbb{C}$ be measurable functions.

1. For $p \in [1, \infty]$, if $\|f_n - f\|_p \to 0$, then $f_n \to f$ in meas.[a]
2. If $f_n \to f$ in meas., then $f_{n_k} \to f$ a.e. for some subseq.
3. If $f_n \to f$ in meas., and $f_{n_j} \to g$ a.e. for some subseq., then $f = g$ a.e.

[3] I don't like this definition as well. But keep reading for important insights.

[a] f_n, f are not necessarily in L^p.

(Terse) Proof. **1.** The case for $p=\infty$ follows from the definition. If $p \in [1,\infty)$, use Chebyshev inequality, for any $\varepsilon > 0$,

$$\mu(\{|f_n - f| > \varepsilon\}) \le \|f_n - f\|_p^p \varepsilon^{-p} \to 0.$$

2. Let $e_j \in c_0 \cap l^{++}$ be any specified rate of decay. Let n_j be an increasing sequence of numbers such that $\sup_{m \ge n_j} \mu(\{|f_m - f| > e_j\}) \le 2^{-j}$ for every $j \ge 1$. Set $A_j = \{|f_{n_j} - f| > e_j\}$, then $\sum_1^\infty \mu(A_j) < \infty$ must mean that $\mu(\limsup A_j) = 0$. (This comes from Borel-Cantelli, see [Fol13, Thm 10.10a]). We know that $x \in A = \limsup A_j$ iff $|f_{n_j}(x) - f(x)| \ne \mathcal{O}(e_j)$[4]. So that the complement of A is precisely the points x such that the pointwise rate of convergence $|f_{n_j}(x) - f(x)| = \mathcal{O}(e_j)$. So that $f_{n_j} \to f$ pointwise on A^c.

3. If $f_{n_k} \to g$ for some subsequence, we can pick a further subsequence f_{n_j} of f_{n_k}, that converges pointwise to f. This is because taking subsequences also preserves convergence in measure. Let $E_1, E_2 \in \mathcal{M}$ such that $\mu(E_1^c) = \mu(E_2^c) = 0$, and $f_{n_k} \to g$ pointwise on E_1, and $f_{n_j} \to f$ pointwise on E_2. If $x \in E_1 \cap E_2$ (which has null complement) then

$$f(x) = \lim f_{n_j}(x) = \lim f_{n_k}(x) = g(x).$$

■

2. Riesz-Thorin Interpolation

To prepare ourselves for the proof of Riesz-Thorin interpolation theorem, we will recall a few needed results and go through a several warm-up exercises.

Suppose $q, r \in [1,\infty]$, $q \ne \infty$. For any $f \in L^q$, there is a natural way of resizing the magnitude of f so that it is in L^r.

Definition 2.1

Let $q, r \in [1,\infty]$, $q \ne \infty$, $f \in L^q$. If $f = |f|(\operatorname{sgn} f)$[a] is the polar decomposition of f, we define

$$\widetilde{f}_r = \begin{cases} |f|^{q/r}(\operatorname{sgn} f) & r \ne \infty \\ \operatorname{sgn} f & r = \infty. \end{cases} \tag{11}$$

[a] for any complex z, $\operatorname{sgn} z = z/|z|$ whenever $z \ne 0$, and $\operatorname{sgn} 0 = 0$.

We offer a quick proof for the fact that $\widetilde{f}_r$ is in L^r, whose norm is determined by

$$\|\widetilde{f}_r\|_r = \begin{cases} \|f\|_q^{q/r} & r \ne \infty \\ 1 \text{ or } 0 & r = \infty^{5}. \end{cases} \tag{12}$$

[4] we say that $X_j = \mathcal{O}(e_j)$ iff $\limsup |X_j|/e_j < \infty$

- If $r \ne \infty$, then $\|\cdot\|_r$ is computed directly using the integral
$$\|\widetilde{f}_r\|_r^r = \textstyle\int_X |f(x)|^{(q/r)r}\,dx = \int_X |f(x)|^q = \|f\|_q^q.$$
- If instead $r = \infty$, then $\|\operatorname{sgn} f\|_\infty = 0$ iff $f = 0$ pointwise a.e, which proves Eq. (12).

Proposition 2.2 [Fol13, Thm 6.10]

Fix $t \in [0,1]$ and suppose $p_t \in [p_0,\ p_1]$ is given by

$$p_t^{-1} = (1-t)p_0^{-1} + (t)p_1^{-1}, \tag{13}$$

1. For any $f \in L^{p_0} \cap L^{p_1}$, its L^{p_t} norm can be estimated *interpolation inequality*

$$\|f\|_{p_t} \le \|f\|_{p_0}^{(1-t)} \|f\|_{p_1}^{(t)}. \tag{14}$$

For the proof of the interpolation theorem however, a specialized version of Lem. 1.1 is needed.

Lemma 2.3

Let $1 \le p_0 < p_t < p_1 \le \infty$, and $f \in L^{p_t}$. There exists a decomposition $f = f_0 + f_1 \in L^{p_0} + L^{p_1}$, and a sequence $\{\phi_n\} \subseteq \Sigma_0$, such that $\phi_n = \phi_{n0} + \phi_{n1} \in \Sigma_0 + \Sigma_0$. This sequence satisfies

1. $\phi_n \to f$ pointwise a.e., $\phi_{nj} \to f_j$ pointwise a.e ($j = 0,1$),

2. $|\phi_n| \nearrow |f|$ pointwise a.e., $|\phi_{nj}| \nearrow |f_j|$ pointwise a.e. ($j = 0,1$),

3. $\|\phi_n - f\|_{p_t} \to 0$, and $\|\phi_{nj} - f_j\|_{p_j} \to 0$ ($j = 0,1$).

Proof. [Fol13, Thm 6.9, 2.10a, b] ■

For the next proposition, we will assume that $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be positive, σ-finite measure spaces. The statement remains valid if we assume μ,ν are semi-finite measures [Fol13, Thm 6.27].

Proposition 2.4 Riesz-Thorin Interpolation, [Fol13, Thm 6.27]

Let $1 \le p_0, p_1, q_0, q_1 \le \infty$, and constants $M_j \ge 0$ ($j = 0,1$). Suppose we are given a linear mapping T

$$T : \left(L^{p_0}(\mu) + L^{p_1}(\mu)\right) \to \left(L^{q_0}(\nu) + L^{q_1}(\nu)\right)$$

which is (p_j, q_j) *stable and bounded* ($j = 0,1$). That is, $T(L^{p_j}(\mu)) \subseteq L^{q_j}(\nu)$, and

$$\|Tf\|_{q_j} \le M_j \|f\|_{p_j} \quad \text{for every} \quad f \in L^{p_j}(\mu).$$

For any $t \in (0,1)$, let p_t, q_t be exponents defined by the linear equations

$$p_t^{-1} = (1-t)p_0^{-1} + (t)p_1^{-1}$$
$$q_t^{-1} = (1-t)q_0^{-1} + (t)q_1^{-1}.$$

Then, T is (p_t, q_t) stable and bounded, and $\|Tf\|_{q_t} \le M_t \|f\|_{p_t}$ for every $f \in L^{p_t}(\mu)$ where $M_t = M_0^{(1-t)} M_1^{(t)}$.

Proof. Let us consider this problem in the abstract for a little while. By convention, if $p = \infty$, then we write $p^{-1} = 0$. It will prove to be useful to define the exponents r_0, r_1 and r_t to be conjugate to q_0, q_1 and q_t. A straightforward calculation will show that

$$r_t^{-1} = (1-t)r_0^{-1} + (t)r_1^{-1}. \tag{15}$$

If the exponents p_j, q_j are in $[1,\infty]$, their reciprocals must fall inside the interval $[0,1]$. The linear equation that defines p_t^{-1} shows that p_t^{-1} is in the strict interior of a line segment with endpoints p_0^{-1} and p_1^{-1}.

Note 2.5
The *extreme points* of $[0,1]$ are $\{0,1\}$ (see appendix for a review of what an extreme point is). From which we deduce

$$p_t^{-1} \in \{0,1\} \iff p_0^{-1} = p_t^{-1} = p_1^{-1}$$
$$r_t^{-1} \in \{0,1\} \iff q_0^{-1} = q_t^{-1} = q_1^{-1} \in \{0,1\}.$$

There is no typo in the second equivalence, reader should verify with Eq. (15).

The main argument of the proof consists of using Def. 2.1 in conjunction with the three lines lemma (Lem. 3.3) to obtain a uniform estimate for $\|Tf\|_{q_t}$ where f ranges over a dense subset of L^{p_t} with bounded norm. The set Σ_0 is a natural candidate for this. That is to say, we will prove

$$\|Tf\|_{q_t} \le M_t \|f\|_{p_t} \quad \text{for all} \quad f \in \Sigma_0,\ \|f\|_{p_t} = 1. \tag{16}$$

With this, we can define a continuous linear operator $S : L^{p_t} \to L^{q_t}$ that extends $T|_{\Sigma_0}$ by uniform continuity. More explicitly, and at the risk of sounding pedantic, given any $f \in L^{p_t}$, let $\{\phi_n\} \subseteq \Sigma_0$ converge to f in L^{p_t}, and $S(f)$ be the element such that

$$\|S(f) - T(\phi_n)\|_{q_t} \to 0. \tag{17}$$

This uniquely characterizes $S(f)$, because bounded linear operators such as $T|_{\Sigma_0}$ map Cauchy sequences to Cauchy sequences. A similar argument also shows that

$$\|S(f)\|_{q_t} \le M_t \|f\|_{p_t} \quad \text{for every} \quad f \in L^{p_t}. \tag{18}$$

Postponing the proof for the **estimate** in Eq. (16) for the moment, let us verify that $S = T$ in the case where

$$1 \le p_0 < p_t < p_1 \le \infty. \tag{19}$$

For any $f \in L^{p_t}$, we would like to make a special choice of ϕ_n that will be useful in some of the computations later on. The key idea is that we can improve the convergence properties of $\phi_n \to f$ if 1) we take our lower approximants from a smaller class of functions, and 2) we pass to a suitable subsequence.

Lemma 2.3, which we will restate for the convenience of the reader, gives us an approximation of f in L^{p_t} that also 'splits' in $L^{p_t} \subseteq L^{p_0} + L^{p_1}$.

> There exists a decomposition $f = f_0 + f_1 \in L^{p_0} + L^{p_1}$, and a sequence $\{\phi_n\} \subseteq \Sigma_0$, such that $\phi_n = \phi_{n0} + \phi_{n1} \in \Sigma_0 + \Sigma_0$. This sequence satisfies
>
> **1.** $\phi_n \to f$ pointwise a.e., $\phi_{nj} \to f_j$ pointwise a.e $(j = 0,1)$,
> **2.** $|\phi_n| \nearrow |f|$ pointwise a.e., $|\phi_{nj}| \nearrow |f_j|$ pointwise a.e. $(j = 0,1)$,
> **3.** $\|\phi_n - f\|_{p_t} \to 0$, and $\|\phi_{nj} - f_j\|_{p_j} \to 0$ $(j = 0,1)$.

By passing to a subsequence of ϕ_n< we can use the properties of T to our advantage. For $j = 0,1$,

$$\|T(\phi_{nj}) - T(f_j)\|_{q_j} \le M_j \|\phi_{nj} - f_j\|_{p_j}.$$

Convergence in norm ($1 \le q_j \le \infty$) means that we can relabel our sequence ϕ_{nj} and assume that $T(\phi_{nj}) \to T(f_j)$ pw a.e. (This comes from Prop. 1.4 for those that are unaware). Adding the two pieces together $T(\phi_{n0}) + T(\phi_{n1})$ gives us $T(\phi_n) \to T(f)$ pointwise a.e. (Because f is in the domain of T)

Since $1 \le q_t \le \infty$, and $\|S(f) - T(\phi_n)\|_{q_t} \to 0$, the sequence $\{T(\phi_n)\}$ of $\mathcal{N}$-measurable functions converges in measure to $S(f)$. One thinks of $S(f)$ as the best possible representative of the sequence of measurable functions $\{T(\phi_n)\}$. By a.e. uniqueness of this representative (proven in Prop. 1.4), we get $S(f) = T(f)$ pointwise a.e.[6] This proves that S is an extension of T under the assumption in Eq. (19).

We now tackle the **estimate** Eq. (16) (with the assumption $1 \le p_0 < p_t < p_1 \le \infty$ still in place). Given $f \in \Sigma_0$, with $\|f\|_{p_t} = 1$, the norm $\|Tf\|_{q_t}$ can be estimated by looking at the relative largeness of the scalar product (proven in Prop. 1.3[7])

$$\langle Tf, \cdot\rangle : L_0^\infty \to \mathbb{C}.$$

Write $f = (\operatorname{sgn} f)\sum |c_j|\chi_{E_j}$. For any $z \in [1,\infty]$, the rescaled version of f as in Def. 2.1, is given by

$$\tilde{f}_z = \begin{cases} (\operatorname{sgn} f)\sum |c_j|^{p_t/z}\chi_{E_j} & r \ne \infty \\ \operatorname{sgn} f & r = \infty. \end{cases}$$

We further subdivide the condition Eq. (19) into two cases. The reasoning for this is because of the hypothesis of the rescaling trick.

1. If $q_0 = q_1$, to compute $\|Tf\|_{q_t}$, we fix an arbitrary $g \in L_0^\infty(\nu)$, with $\|g\|_{r_t} = 1$. The linearity of T and the evaluation map $\langle g, \cdot\rangle$ allows us to separate the effects of the rescaling from the action of g on $T(\chi_{E_j})$. For any $z \ne \infty$,

$$\begin{aligned} \langle g, T(\tilde{f}_z)\rangle &= \int_Y g(y)T(\tilde{f}_z)(y)\,\mathrm{d}y \\ &= \sum |c_j|^{p_t/z} e^{i\arg c_j}\langle g, T(\chi_{E_j})\rangle \end{aligned} \tag{20}$$

If $z = \infty$, a formula similar to Eq. (20) can be obtained.

[6] the reason why we cannot use positive definiteness of some of the norms involved q_0, q_1, and q_t is because the three sequences $T(\phi_n)$, $T(\phi_{n0})$, and $T(\phi_{n1})$ have limits that lie in three different L^q spaces.

[7] remember that ν is σ-finite (can be relaxed to semifinite)

2. If $q_0 \neq q_1$, then $q_t^{-1} \notin \{0,1\}$. Which rules out the possibility that $r_t = \infty$. For any $g \in \Sigma_0$, $\|g\|_{q_t} = 1$, suppose that $g = \sum |d_k| e^{i \arg d_k} \chi_{F_k}$. We can use the rescaling trick on both sides. Writing $\widetilde{g}_z = |g|^{r_t/z}(\operatorname{sgn} g)$, we see that for any $z \neq \infty$,

$$\begin{aligned}\langle \widetilde{g}_z, T(\widetilde{f}_z)\rangle &= \textstyle\int_Y \widetilde{g}_z(y) T(\widetilde{f}_z)(y)\,\mathrm{d}y \\ &= \sum |d_k|^{r_t/z} |c_j|^{p_t/z} e^{i(\arg c_j + \arg d_k)} \langle \chi_{F_k}, T(\chi_{E_j})\rangle. \end{aligned} \tag{21}$$

The sum in Eq. (21) is over finitely many j,k. A similar equation is obtained if $z = \infty$.

With $g \in L_0^\infty$, or $g \in \Sigma_0$ held fixed, we will show that the scalar product is bounded by M_t. Explicitly this means

$$\begin{cases} |\langle g, T(\widetilde{f}_{p_t})| \le M_t & q_0 = q_1 \\ |\langle \widetilde{g}_{r_t}, T(\widetilde{f}_{p_t})| \le M_t & q_0 \neq q_1. \end{cases} \tag{22}$$

The important part of Eqs. (20) and (21) is in the exponents, i.e. $|c_j|^{p_t/z}$ (resp. $|c_j|^{p_t/z}|d_k|^{r_t/z}$). It is also useful to remember that $\widetilde{f}_{p_t} = f$, and $\widetilde{g}_{r_t} = g$ whenever $r_t \neq \infty$, as the exponent cancels out nicely. For any $\omega \in \mathbb{C}$, we define the function $\varphi : \mathbb{C} \to \mathbb{C}$,

$$\varphi(\omega) = \begin{cases} \langle g, T(\widetilde{f}_{(1-\omega)p_0^{-1} + (\omega)p_1^{-1}})\rangle & q_0 = q_1 \\ \langle \widetilde{g}_{(1-\omega)r_0^{-1} + (\omega)r_1^{-1}}, T(\widetilde{f}_{(1-\omega)p_0^{-1} + (\omega)p_1^{-1}})\rangle & q_0 \neq q_1 \end{cases} \tag{23}$$

In both cases, $\varphi : \mathbb{C} \to \mathbb{C}$ is a holomorphic function because of ω is placed in the exponent. Furthermore, if $\operatorname{Re}\omega \in \{0,1\}$, the relative sizes of the functions

$$\widetilde{g}_{(1-\omega)r_0^{-1} + (\omega)r_1^{-1}} \quad \text{and} \quad \widetilde{f}_{(1-\omega)p_0^{-1} + (\omega)p_1^{-1}}$$

remain unchanged as we perturb ω by a purely imaginary number. That is to say, suppose that $\operatorname{Re}\omega = j \in \{0,1\}$, then

$$\|\widetilde{f}_{(1-\omega)p_0^{-1} + (\omega)p_1^{-1}}\|_{p_j} = \|\widetilde{f}_{p_j}\|_{p_j} = 1, \text{ and} \tag{24}$$

$$\|\widetilde{g}_{(1-\omega)r_j^{-1} + (\omega)r_j^{-1}}\|_{r_j} = \|\widetilde{g}_{r_j}\|_{r_j} = 1 \tag{25}$$

By considering the cases where $\operatorname{Re}\omega = j$ and $\operatorname{Im}\omega = 0$, we obtain

$$\begin{cases} |\langle g, T(\widetilde{f}_{p_j})\rangle| \le M_j & q_0 = q_1 \\ |\langle \widetilde{g}_{r_j}, T(\widetilde{f}_{p_j})\rangle| \le M_j & q_0 \neq q_1. \end{cases}$$

If $\operatorname{Re}\omega = j \in \{0,1\}$, and $\operatorname{Im}\omega \neq 0$. We can use the 'norm-invariance' (Eq. (24)) and see that $|\varphi(\omega)| \le M_j$.

Here is where the complex analysis comes in. By the three lines lemma, Lem. 3.3, this gives us an estimate for $\varphi(t)$

$$|\varphi(t)| \le M_0^{(1-t)} M_1^{(t)} = M_t,$$

where t has the same meaning as before, i.e. $t \in (0,1)$. In both subcases of $1 \le p_0 < p_t < p_1 \le \infty$, we conclude either directly or by a density argument of $\Sigma_0(\nu) \subseteq L_0^\infty(\nu)$, that the number

$$\|Tf\|_{q_t} = \sup\{|\langle g, Tf\rangle|,\ g \in L_0^\infty(\nu),\ \|g\|_{r_t} = 1\}$$

is bounded above by M_t. Since the simple function f is chosen arbitrarily, this completes the proof for the **estimate** under the assumption $1 \le p_0 < p_t < p_1 \le \infty$.

Let us tackle the remaining cases. If $p_0 \neq p_1$, we can always relabel the exponents such that $1 \le p_0 < p_t < p_1 \le \infty$. It remains to prove the **estimate** when $p_0 = p_t = p_1$ (and q_0, q_1, q_t unconstrained). This is straightforward because we can apply the interpolation inequality (proven in Prop. 2.2). For any $f \in L^{p_t}$, the function Tf is in $L^{q_0} \cap L^{q_1}$, which means

$$\|Tf\|_{q_t} \le \|Tf\|_{q_0}^{(1-t)} \|Tf\|_{q_1}^{(t)} \le M_t \|f\|_{p_t}. \tag{26}$$

The extension argument (and the proof of $S = T$) is unnecessary because Eq. (26) holds for every $f \in L^{p_t}$. Therefore T restricts to a (p_t, q_t) stable and bounded linear operator with $\|T\| \le M_t$, and the proof of the interpolation theorem is complete. ■

2.1. Application to the Fourier Transform

Corollary 2.6 Hausdorff-Young, [Fol13, Thm 8.21, 8.30]
Let $1 \le p \le 2$ and $2 \le q \le \infty$ be conjugate exponents. The Fourier transform on $\mathbb{R}^n$, $\mathcal{F} : (L^1 + L^2)(\mathbb{R}^n) \to (L^\infty + L^2)(\mathbb{R}^n)$ is a norm-decreasing (p,q) stable linear operator. A similar result holds for $\mathcal{F} : (L^1 + L^2)(\mathbb{T}^n) \to (l^\infty + l^2)(\mathbb{Z}^n)$ as well.

Proof. This follows almost immediately from the Riesz-Thorin interpolation theorem other than one small thing. If p_j, q_j are exponents in $[1,\infty]$, and p_j, q_j are Holder conjugates ($j = 0,1$), then p_t, q_t is a conjugate pair for any p_t, q_t defined as in Prop. 2.4. ■

2.2. Application to Conditional Expectations

Definition 2.7
Let $(X, \mathcal{M}, \mu)$ be a probability space, and $\mathcal{N} \subseteq \mathcal{M}$ a sub σ-algebra. The *conditional expectation* is a linear map

$$\mathbb{E}(\cdot \mid \mathcal{N}) : L^1(\mathcal{M}, \mu) \to L^1(\mathcal{N}, \mu),\text{[a]}$$

defined by its action on $\mathcal{N}$ measurable sets. That is, for every $A \in \mathcal{N}$ and $f \in L^1(\mu)$,

$$\textstyle\int_A f(x)\,\mathrm{d}\mu(x) = \int_A \mathbb{E}(f \mid \mathcal{N})(x')\,\mathrm{d}\mu(x').\text{[b]} \tag{27}$$

[a] by $(\mathcal{N}, \mu)$ we mean $(\mathcal{N}, \mu|\mathcal{N})$ i.e., the restriction of the measure
[b] for contrast, we use x' on $(X, \mathcal{N}, \mu|\mathcal{N})$, and x on $(X, \mathcal{M}, \mu)$.

Proposition 2.8 [Fol13] Ex 3.17
Let $(X, \mathcal{M}, \mu)$ be a finite measure space ($\mu \ge 0$) and $\mathcal{N}$ a sub σ-algebra of $\mathcal{M}$. Given a real-valued $f \in L^1(X, \mathcal{M}, \mu)$, there exists $g \in L^1(X, \mathcal{N}, \nu)$, where $\nu = \mu|_{\mathcal{N}}$

such that

$$\int_A f d\mu = \int_A g d\nu, \quad \forall A \in \mathcal{N}.$$

The function g is uniquely determined ν a.e. Moreover, if $f \geq 0$ (μ a.e.), then $g \geq 0$ (ν a.e.)

Proof. The uniqueness of g follows from the ability of integration over measurable sets to separate points in L^1. [Fol13, Thm 2.23] Let η_1 be a signed measure on $\mathcal{M}$, defined by $d\eta_1 = f d\mu$. Let $\eta_2 = \eta_1|_{\mathcal{N}}$. We can apply Radon-Nikodym theorem and decompose η_2 with respect to ν. We obtain $g \in L^1(X, \mathcal{N}, \nu)$, finite signed measures (on $\mathcal{N}$) ρ, λ, such that $\lambda \perp \nu$, $\rho \ll \nu$, $d\rho = g d\nu$, and $\eta_2 = \lambda + \rho$. The proof is complete upon showing $\lambda = 0$, and this occurs if and only if $\eta_2 \ll \nu$. Because f is μ-integrable, $\eta_1 \ll \mu$, so $\mathrm{Ker}(\mu) \subseteq \mathrm{Ker}(\eta_1)$.[8] Fix any $A \in \mathrm{Ker}(\nu) = \mathcal{N} \cap \mathrm{Ker}(\mu) \subseteq \mathrm{Ker}(\eta_1)$, then $\eta_2(A) = \mu(A) = 0$. The last claim is obvious. ■

Because μ is a finite measure, Σ_0 is the vector space formed by taking linear combinations of indicator functions. The equation defining the cond. exp. on L^1 extends naturally to a scalar product on Σ_0 as follows. Fix any $f \in L^1(\mathcal{M}, \mu)$ and denote its conditional expectation by g, we can write

$$\langle f, \chi_A \rangle = \int_A f(x)\, \mathrm{d}P(x) = \int_A g(x')\, \mathrm{d}P(x') = \langle g, \chi_A \rangle \quad \forall A \in \mathcal{N}. \tag{28}$$

If $\phi = \sum_1^k c_j \chi_{A_j}$, where $c_j \in \mathbb{R}$ and $A_j \in \mathcal{N}$, by linearity on both sides of Equation (28),

$$\begin{aligned}\langle f, \phi \rangle &= \int_X f(x)\phi(x)\, \mathrm{d}P(x) \\ &= \int_X g(x')\phi(x')\, \mathrm{d}P(x') = \langle g, \phi \rangle\end{aligned}$$

for every $\phi \in \Sigma_0$. Since $\Sigma_0(\mathcal{N}, \mu)$ is dense in $L^\infty(\mathcal{N}, \mu)$ (the ess. bounded, $\mathcal{N}$-measurable equivalence classes, **not pointwise functions**), we arrive at Equation (29), which some people prefer to take as the definition of the cond. exp. (present author included)

$$\langle f, b \rangle = \langle g, b \rangle \quad \text{for every} \quad b \in L^\infty(\mathcal{N}, \mu). \tag{29}$$

This highlights a certain functional analytic significance that is not present in the original definition. Namely, the cond. exp. is the operator adjoint of the inclusion map of $L^\infty(\mathcal{N}, \mu) \to L^\infty(\mathcal{M}, \mu)$. Unfortunately, most introductory texts do not characterize dual space of L^∞ as the space of *finitely additive measures*, or $(L^\infty(\mathcal{M}, \mu))^* = ba(\mathcal{M}, \mu)$.

Proposition 2.9

For $p \in [1, \infty]$, the conditional expectation as defined in Equation (27) is L^p *stable*, meaning it sends $f \in L^p(\mathcal{M}, \mu)$ to $\mathbb{E}(f \mid \mathcal{N}) \in L^p(\mathcal{N}, \mu)$. Moreover,

$$\|\mathbb{E}(f \mid \mathcal{N})\|_{L^p(\mathcal{N},\mu)} \leq \|f\|_{L^p(\mathcal{M},\mu)} \quad \text{for every} \quad f \in L^p(\mathcal{M}, \mu). \tag{30}$$

Proof. We will use Riesz-Thorin's interpolation to show the L^p stability of the cond. exp. First we consider the extreme cases at the endpoints of $[1, \infty]$.

1. If $p = 1$, $f \in L^1(\mathcal{M}, \mu)$, let $g = \mathbb{E}(f \mid \mathcal{N})$ be its conditional expectation. We can compute the $L^1(\mathcal{N}, \mu)$ norm of g by integrating against $\operatorname{sgn} g \in L^\infty(\mathcal{N}, \mu)$:

$$\begin{aligned}\|g\|_{L^1(\mathcal{N},\mu)} &= |\langle f, \operatorname{sgn} g \rangle| \\ &\leq \|f\|_{L^1(\mathcal{M},\mu)} \|g\|_{L^\infty(\mathcal{M},\mu)} \leq \|f\|_{L^1(\mathcal{M},\mu)}\end{aligned}$$

More is true, if we take $f = 1$ (the constant function), then $\|g\|_{L^1(\mathcal{N},\mu)} = \|f\|_{L^1(\mathcal{M},\mu)}$. So that $\mathbb{E}(\cdot \mid \mathcal{N})$ is continuous, and its operator norm is equal to 1.

2. If $p = \infty$, fix $f \in L^\infty(\mathcal{M}, \mu)$, let $c = \|f\|_{L^\infty(\mathcal{M},\mu)}$. The funciton $h = c \pm f$ is non-negative pointwise a.e; and because L^∞ embeds continuously in L^1 for finite measure spaces, we can take its cond. exp. Using linearity of the cond. exp.,

$$\mathbb{E}(h \pm f \mid \mathcal{N}) = c \pm \mathbb{E}(f \mid \mathcal{N}) \geq 0.$$

Using the fact that if $g_1, g_2 \in L^+(\mathcal{N}, \mu)$, $g_1 \leq g_2$ iff $g_1 d\mu \leq g_2 d\mu$ as positive measures, we see that

$$-c \leq \mathbb{E}(f \mid \mathcal{N}) \leq c \quad \text{pointwise a.e..}$$

This proves $\|\mathbb{E}(f \mid \mathcal{N})\|_{L^\infty(\mathcal{N},\mu)} \leq \|f\|_{L^\infty(\mathcal{M},\mu)}$.

Thus, the assumptions of the Riesz-Thorin interpolation theorem are satisfied, because μ and its restriction $\mu|\mathcal{N}$ are probability measures. This gives us L^p stability for $p \in [1, \infty]$. Moreover, the operator norm of $\mathbb{E}(\cdot \mid \mathcal{N}) : L^p(\mathcal{M}, \mu) \to L^p(\mathcal{N}, \mu)$ is at most 1 — a result also from the interpolation theorem. ■

3. Notes and Generalizations

3.1. Characterization of Intervals

Equation (9) deserves some our attention. We restate it for the convenience of the reader.

$$\begin{aligned}\|g\|_\infty &= \inf\Big\{a \geq 0,\ \mu(\{x \in X,\ |g(x)| > a\}) = 0\Big\} \\ &= \sup\Big\{a \geq 0,\ \mu(\{x \in X,\ |g(x)| > a\}) > 0\Big\}.\end{aligned}$$

Suppose we are given a non-empty set $A \subseteq [0, \infty]$ such that

> for any $x \in A$, $[x, \infty] \subseteq A$, then either $A = [c, \infty]$ or $A = (c, \infty]$ where $c \geq 0$; and either $A^c = [0, c)$ or $A^c = [0, c]$.

The number c is also the infimum of A, from which we can deduce that $\inf A = \sup A^c$. We think of the set A (resp. A^c)

[8] I realized that we have not defined the *kernel* of a complex σ-additive measure yet. But we can think of $\mathrm{Ker}\,\mu = \{E \in \mathcal{M},\ |\mu| = 0\}$ where $|\mu|$ is the t.v. measure of μ

as being defined by a single point in $\overline{\mathbb{R}}$. An alternate way of thinking about this is to equip $[0, \infty]$ with the order topology [Fol13, Ex 4.9], and use $\partial A = \partial A^c = \{c\}$.

3.2. Extreme Points

Definition 3.1
The *extreme points* of a convex subset C in a vector space X are the set of points $p \in C$ such that

$$\forall x, y \in C,\ c_t(x, y) \neq p \quad \forall t \in (0, 1).$$

Proposition 3.2 Krein-Milman, [Yos12, Chap XII]
Let X be a l.c.s. and C a precompact convex set in X. Then, $\overline{C}$ is equal to the closed convex hull of its extreme points.

Let $f : C \to \mathbb{R}$ be a convex function defined on a bounded convex domain in $\mathbb{R}^n$, and $S = \{x_1, \dots, x_k\}$ be the set of extreme points of C. (One can show that in finite dimensions, S is finite). If $x \in C$, it can be written as the cv. combination $x = \sum_1^k t_i x_i$, and by convexity of f:

$$\begin{aligned} f(x) = f\left(\sum_1^k t_i x_i\right) &\le \sum_1^k t_i f(x_i) \\ &\le \left(\sum_1^k t_i\right) \sup f(S) = f(S). \end{aligned}$$

In other words, having a control over the extreme points of C gives us a bound on all of C. A complex analytic analogue of this is the following.

Lemma 3.3 Three Lines, [Fol13, Thm 6.26]
Let φ be a bounded continuous function on $\mathrm{Re}(z) \in [0, 1]$ that is holomorphic in $\mathrm{Re}(z) \in (0, 1)$. If $|\varphi(z)| \le M_0$ for $\mathrm{Re}(z) = 0$ and $|\varphi(z)| \le M_1$ for $\mathrm{Re}(z) = 1$, then

$$|\varphi(z)| \le M_0^{1-t} M_1^t$$

for any $t \in (0, 1)$, where $t = \mathrm{Re}(z)$.

3.3. Interpolation Spaces

Proposition 3.4 [Fol13, Ex 6.3, 6.4]
The norms in Eqs. (31) and (32) make i) $(L^{p_0} \cap L^{p_1}, \|\cdot\|_a)$ and ii) $(L^{p_0} + L^{p_1}, \|\cdot\|_b)$ Banach spaces, such that iii) $L^{p_0} \cap L^{p_1} \hookrightarrow L^{p_t} \hookrightarrow L^{p_0} + L^{p_1}$.[a]

$$\|f\|_a = \|f\|_{p_0} + \|f\|_{p_1} \tag{31}$$

$$\|f\|_b = \inf\{\|g_0\|_{p_0} + \|g_1\|_{p_1},\ f = g_0 + g_1 \text{ a.e.}\} \tag{32}$$

(in the right hand side of Eq. (32), we can take $g_j \in L^{p_j}$)

[a] $\hookrightarrow$ = cont. linear embedding

Proof. A bit tedious, maybe I will skip it or put it in the appendix. ■

3.4. Vector Lattices and the Monotone Convergence Theorem

Definition 3.5
Let V be a v.s. over $\mathbb{R}$, a subset $V^+ \subseteq V$ is a *positive cone* if
1. $\phi_1, \phi_2 \in V^+$, so is $\phi_1 + \phi_2$, and
2. if $\phi \in V^+$, then $\alpha\phi \in V^+$ for all $\alpha \ge 0$.

The positive cone V^+ defines a partial order on V, for any $x, y \in V$, we write $x \le y$ if $y - x \in V^+$. We say that V is a *vector lattice* [Yos12] if for every pair $x, y \in V$, there exists elements $z, w \in V$ such that

$$z \le x \le w \quad \text{and} \quad z \le y \le w,$$

and if z', w' are elements that satisfy the same relation, then

$$z' \le z \quad \text{and} \quad w \le w'.$$

lattice we write $z = x \vee y$, and $w = x \wedge y$. It is easy to see that the elements z, w are unique. They are called the *least upper bound* and *greatest lower bound* of x and y.

We use the suggestive notation $\phi_n \nearrow \phi$ to mean that i) $\{\phi_n\} \subseteq \{\psi \in V,\ \psi \le \phi\}$; ii) the sequence $\{\phi_n\}_{n \ge 1}$ directed by increasing n is *monotone*: $\phi_n \le \phi_{n+1}$ at every n; iii) if $\psi \in V$ is an element such that $\phi_n \le \psi$ for all n, then $\phi \le \psi$. This of course leads to a characterization of 'upwards' convergence using the theory of nets (see [Fol13, Ex 4.9]) which is beyond the scope of our current document.

Definition 3.6
Let $\Sigma^+ \subseteq V^+$ be a vector sub-lattice. A functional $I : \Sigma^+ \to [0, \infty]$
1. is *monotonic* if for every $\phi_1, \phi_2 \in \Sigma^+$, $\phi_1 \le \phi_2$ implies $I(\phi_1) \le I(\phi_2)$,
2. is *sequentially l.s.c.* if for every $\phi_n \nearrow \phi$ in Σ^+, then $I(\phi) \le \liminf I(\phi_n)$, and
3. is said to satisfy *the m.c. property* if $\phi_n \nearrow \phi$ in Σ^+ implies $I(\phi_n) \nearrow I(\phi)$ in $[0, \infty]$.

Proposition 3.7 Monotone Convergence Theorem
Let $I : \Sigma^+ \to [0, \infty]$ be a monotone functional. Then I satisfies the m.c. property iff I is sequentially l.s.c. Moreover, if we define

$$I^+ : V^+ \to [0, \infty] \quad \text{where}$$
$$I^+(x) = \sup\{I(\phi),\ \phi \in \Sigma^+,\ \phi \le x\}. \tag{33}$$

then I^+ is a monotone functional, and is sequentially l.s.c. whenever I is (resp. satisfies m.c.).

Lemma 3.8
Let $\varphi : [0, \infty] \to [0, \infty]$ be a non-decreasing function,

and suppose that for every $x \in [0,\infty]$,

$$\varphi(x) = \lim_{y<x,\ y\to x} \varphi(y).$$

For any subset $A \subseteq [0,\infty]$, we have $\varphi(\sup A) = \sup \varphi(A)$.

3.5. Lower semi-continuity

There is a topology on $\mathbb{R}$ that makes it a T_0 but not a T_1 space. Let $\mathcal{E} = \{(a,\ \infty)\}_{a\in\mathbb{R}}$. Since $\mathcal{E}$ is closed under finite intersections, [Fol13, Prop 4.3, 4.4] together tell us it generates a topology on $\mathbb{R}$. Which we can write explicitly as $\mathcal{T}_{\text{lsc}} = \{\varnothing,\ \mathbb{R}\} \cup \mathcal{E}$.

Given two distinct points $t_0 < t_1 \in \mathbb{R}$, take $U = (t_1 - \varepsilon,\ +\infty)$ for sufficiently small ε. This open set separates t_1 from t_0, but any open set containing t_0 must also contain t_1. Therefore $(\mathbb{R},\ \mathcal{T}_{\text{lsc}})$ is T_0 but not T_1.

This topology is actually way more useful than at first glance. It provides a weaker notion of continuity in scenarios which are deemed to be too restrictive; where ordinary continuity refers to continuity w.r.t. $(\mathbb{R},\ \mathcal{T}_{\text{ball}})$. If X is another topological space, a function $I : X \to \mathbb{R}$ is *lower semi-continuous* (abbrev. l.s.c.) whenever it is continuous with respect to $(\mathbb{R},\ \mathcal{T}_{\text{lsc}})$. Moreover, if $x_n \to x \in X$, then

$$I(x) \le \liminf I(x_n).$$

Applications of $(\mathbb{R},\ \mathcal{T}_{\text{lsc}})$ can be found in optimization, the calculus of variations, partial differential equations.

3.6. Approximation Techniques

We construct an abstract approximation framework. Let Ω be an arbitrary set and $\Sigma \subseteq \Omega$ be a subset on which we define a functional $J : \Sigma \to [0,+\infty]$.

Definition 3.9
Suppose for every $x \in \Omega$, we are given a set $\Sigma_x \subseteq \Sigma$. We define the *upper and lower approximations* to J,

$$I^+(x) = \sup_{z\in\Sigma_x} J(x) \quad \text{and} \quad I^-(x) = \inf_{z\in\Sigma_z} J(z). \tag{34}$$

The functional J is said to be *consistent*[a] if

1. for every $x \in \Omega$, the set Σ_x is non-empty,
2. for every $z \in \Sigma$, $z \in \Sigma_z$, and
3. $J(z) = I^+(z) = I^-(z)$ for all $z \in \Sigma$, this makes $I^\pm$ an extension of J.

[a] More generally, a subset $\Sigma' \subseteq \Sigma$ is said to be consistent if $J|_{\Sigma'}$ is a consistent functional.

We always assume that J is consistent. The linear ordering on $[0,\infty]$ induces a relation on Ω through $I = I^\pm$.

Definition 3.10
Given $x, y \in \Omega$, we say that x is *less than* y as measured by I whenever $I(x) \le I(y)$. This is written

$$x \lesssim_I y \quad \text{or} \quad x \le y \text{ (meas. } I)$$

We are ready to discuss three juicy and extremely important proof techniques for estimating the sizes of $I^\pm(x)$.

Technique 1 To show $x \le y$ (meas. I), it is sufficient to show $\Sigma_x \subseteq \Sigma_y$ if $I = I^+$ (or $\Sigma_y \subseteq \Sigma_x$ whenever $I = I^-$).

Let $x \in \Omega$ be arbitrary, we wish to bound $I(x)$ by c (we agree that $c \ge 0$ means $c \in [0,+\infty]$ and 'bounding' means from above unless specified otherwise).

$$I^+(x) \le c \quad \text{or} \quad I^-(x) \le c.$$

Technique 2 On one hand, if $I = I^+$,

$$I^+(x) \le c \quad \text{is implied by} \quad \Sigma_x \le c \text{ (meas. } I),$$

or $\Sigma_x \subseteq \{I \le c\}$. On the other hand,

$$I^-(x) \le c \quad \text{is implied by} \quad \phi \ge c \text{ (meas. } I),\ \exists \phi \in \Sigma_x,$$

or equivalently: $\Sigma_x \cap \{I \ge c\} \ne \varnothing$. An exercise for the reader: apply to [Fol13, Thm 1.10, 1.13].

Technique 3 The most important technique by far is to control the terms leading up to the supremum/infimum.

Since $I^+(x)$ is defined by a supremum, it induces a sequence $(\phi_n) \subseteq \Sigma_x$ that increases to $I^+(x)$ (meas. I). To obtain a bound for $I^+(x)$, it therefore suffices to estimate $(J(\phi_n))_n$ using another sequence (c_n). We end up with the following condition.

Let $I = I^+$ and $x \in \Omega$, let $(\phi_n) \subseteq \Sigma_x$ increase to $I(x)$ (meas. I), suppose further that there exists a sequence $(\psi_n) \subseteq \Omega$, where

$$\phi_n \le \psi_n \text{ (meas. } I) \quad \forall n \ge N$$

and that $\liminf_{n\to\infty} I(\psi_n) \le c$. Then,

$$I(x) \le \liminf_{n\to\infty} I(\psi_n) \le c$$

3.7. Abstract Approximators

Let $\Sigma^{(1)}$ and $\Sigma^{(2)}$ be consistent subsets of Σ, and write $\Sigma_x^{(k)} = \Sigma^{(k)} \cap \Sigma_x$ $(k = 1,2)$. Let us define

$$I^+_{(1)}(x) = \sup_{\phi\in\Sigma_x^{(1)}} J(\phi) \quad \text{and} \quad I^+_{(2)}(x) = \sup_{\phi\in\Sigma_x^{(2)}} J(\phi) \quad \text{for every} \quad x \in \Omega,$$

and similarly for $I^-_k(x)$ $(k = 1,2)$.

Proposition 3.11 Abstract Fatou's Lemma
Let $x \in \Omega$ such that for every $\phi \in \Sigma_x^{(1)}$, there exists a se-

quence $\{\psi_n\} \subseteq \Sigma_x^{(2)}$ satisfying

$$J(\phi) \le \liminf J(\psi_n)^a,$$

then $I^+_{(1)}(x) \le I^+_{(2)}(x)$.[b]

[a] equivalently, for every $\varepsilon > 0$ there exists $\psi_\varepsilon \in \Sigma_x^{(2)}$ such that $\psi_\varepsilon \le \phi + \varepsilon$ (meas. J)

[b] and if $\limsup J(\psi_n) \le J(\phi)$ then $I^-_{(2)}(x) \le I^-_{(1)}(x)$.

To conclude our discussion, we give two sufficient conditions for $I^-_{(1)} \le I^+_{(2)}$. The first of these two is known as an intersection theorem in nonlinear PDE, the second is a weakened version based on a cofinality argument.

Technique 1 Let $\Sigma^{(j)}$ be generating sets on Ω and $J_{(j)} : \Sigma^{(j)} \to \overline{\mathbb{R}}$ be functionals defined. Suppose that for every $x \in \Omega$, the two generating functionals satisfy an 'overlapping' condition:

$$J_{(1)}(\Sigma_x^{(1)}) \cap J_{(2)}(\Sigma_x^{(2)}) \neq \varnothing,$$

then $I^-_{(1)}(x) \le I^+_{(2)}(x)$ and $I^-_{(2)}(x) \le I^+_{(1)}(x)$.

Technique 2 Suppose there exists $\phi_k \in \Sigma_{(k)}(x)$ $(k = 0, 1)$ such that $J_0(\phi_0) \le J_1(\phi_1)$, then $I^-_0(x) \le I^+_{(1)}(x)$.

3.8. Scalar Products

Our treatment of the dual space of L^∞_0 is heavily inspired by distribution theory on $\mathbb{R}^n$. If we assume that μ is σ-finite, we can give L^∞_0 a topology that is very similar to the C^∞ topology of compact convergence of test functions. A net $\langle f_\alpha\rangle_{\alpha\in A} \subseteq L^\infty_0$ converges to $f \in L^\infty_0$ if there exists $E \in \mathcal{M}$, $\mu(E) < \infty$, and $\alpha_0 \in A$ such that

$$\operatorname{supp}(f_\alpha) \subseteq E \quad \forall \alpha \ge \alpha_0 \quad \text{and} \quad \|f_\alpha - f\|_\infty \to 0.$$

If this is the case, we say that $f_\alpha \to f$ *essentially uniformly on μ-finite sets*. This allows for the interchange of limits when integrated against a complex measure ν, which is of course assumed to be absolutely continuous with respect to μ, as f_α is an equivalence class of functions. This has applications in the theory of martingales, for example. We give a quick proof for the sequential continuity of the scalar product with respect to this topology. Fix $g \in L^1_{loc}$ and a sequence $\phi_n \to \phi$ ess. uniformly on μ-finite sets in L^∞_0. Let E be a μ-finite subset such that

$$\operatorname{supp}(\phi) \cup (\cup_1^\infty \operatorname{supp}(\phi_n)) \subseteq E,$$

then the integrands are uniformly dominated as

$$|g(x)\phi_n(x)| \le |g(x)| \sup_n \|\phi_n\|_\infty \chi_E \in L^1(\mu).$$

It is also clear that $g(x)\phi_n(x) \to g(x)\phi(x)$ pointwise a.e. By the dominated convergence theorem,

$$\langle g, \phi_n\rangle = \int_X g(x)\phi_n(x)\,\mathrm{d}x \longrightarrow \int_X g(x)\phi(x)\,\mathrm{d}x = \langle g, \phi\rangle.$$

As for how Σ_0 is related to L^∞_0 (within this new topology), the proof of [Fol13, Thm 2.10] will show that Σ_0 is dense in L^∞_0 w.r.t. the topology of ess. uniform convergence on μ-finite sets. This essentially means that the scalar product is completely characterized by its action on $\{\chi_A,\ \mu(A), \infty\}$.

Last thing we have to say about the scalar product: we can give $L^1_{loc}(\mu)$ the strong operator topology (see [Fol13, Chap 5.4]) with respect to $\langle g, \cdot\rangle$, which is of course the same as the subspace topology inherited from weak-$*$ topology on $(L^\infty_0)^*$. The dual spaces for non-Banachable t.v.s. is a complicated matter, and we will not pursue this line of inquiry any further. We encourage the reader to check out some of the fantastic texts that have been written [Yos12; DS88; HS12].

We now compare the claims and the proofs of Proposition 1.3 with [Fol13, Thm 6.13, 6.14]. We have assumed that μ is σ-finite from the beginning. But this assumption can be removed by using Lemma 3.15. The assertions made in Prop. 1.3, namely Statement i), and ii) loosely correspond to Theorems 6.13, and 6.14 in [Fol13] respectively.

The premise in [Fol13] is that g is a complex-valued, measurable function such that the scalar product $\langle g, \phi\rangle$ converges absolutely for every $\phi \in \Sigma_0$, which is equivalent to $g \in L^1_{loc}(\mu)$ — as the argument below shows.

- Necessity: $\chi_A \in \Sigma_0$ whenever $\mu(A) < \infty$.
- Sufficiency: take linear combinations of χ_{E_j}, where $\mu(E_j) < \infty$.

Given such a g, Folland also defines the quantity

$$M'_q(g) = \sup\Big\{|\langle g, \phi\rangle|,\ \phi \in \Sigma_0,\ \|\phi\|_p = 1\Big\}. \tag{35}$$

It is easy to see that $M'_q(g) = M_q(g)$, but for completeness we will sketch the proof. The estimate $M'_q(g) \le M_q(g)$ is obvious. If the sequence $\{f_n\} \subseteq L^\infty_0$ defines the supremum $M_q(g)$, within any ε tolerance we obtain

$$M_q(g) - \varepsilon \le |\langle g, f_n\rangle|$$

By the density of $\Sigma_0 \subseteq L^\infty_0$: there exists $\{\phi_m\} \subseteq \Sigma_0$ such that $\phi_m \to f_n$ in ess. uniformly on μ-finite sets. Hence

$$M_q(g) - \varepsilon \le |\langle g, \phi_m\rangle|. \tag{36}$$

As in the proof of [Fol13, Thm 2.10], we can take each ϕ_m to be a subordinate of f_n, so $\|\phi_m\|_p \le 1$. Using which we can bound the right hand side of Equation (36) by $M_q(g)$ and conclude $M_q(g) = M'_q(g)$

3.9. Finitely additive measures

We turn our attention to $\mathcal{L}^\infty(\mathcal{M})$ or $L^\infty(\mathcal{M}, \mu)$. It is well known that $\mathcal{L}^\infty(\mathcal{M})$ and $L^\infty(\mathcal{M}, \mu)$ are Banach spaces. The dual space of $\mathcal{L}^\infty(\mathcal{M})$ is the space of *finitely additive measures*, and is denoted by $ba(\mathcal{M})$. For more on this, see [Tol20, Chap 4-6] for an accessible introduction to this subject, [Yos12] has a short section on this, and the classic [DS88, Chap 3].

3.10. Separation of Points

Let Ω_0, Ω_1 be sets. A collection of mappings $\mathcal{F} \subseteq \Omega_1^{\Omega_0}$ *separates* Ω_0 if

$$f(x_0) = f(x_1) \quad \text{for every} \quad f \in \mathcal{F} \quad \text{iff} \quad x_0 = x_1$$

If we fix $x \in \Omega_0$ and look at the correspondence $f \mapsto f(x)$ as f ranges in $\mathcal{F}$, we obtain a mapping $\varphi(x) : \mathcal{F} \mapsto \Omega_1$.

> If $\mathcal{F}$ separates the points of Ω_0, then φ is an injection and we consider Ω_0 as a subset of $\Omega_1^{\mathcal{F}}$.

We give a few examples.

1. [Fol13, Thm 2.23] Let $(X, \mathcal{M}, \mu)$ be a measure space. For any $E \in \mathcal{M}$, we can associate it with $\chi_E \in L^\infty(\mu)$. (This mapping is usually not injective.) For any $E \in \mathcal{M}$, we write
$$\langle E, f\rangle_1 = \int_E f(x)\,d\mu.$$
Which means
$$\langle E, f\rangle_1 = \langle E, g\rangle_1 \quad \forall E \in \mathcal{M} \quad \text{iff} \quad f = g \text{ a.e.}$$
2. [Fol13, Thm 5.8b] Let X be a Banach space. The scalar product
$$\langle f, x\rangle = f(x) \quad \text{for every} \quad f \in X^*, x \in X$$
separates both X and X^*. To see that $\langle f, \cdot\rangle$ separates X^*: pick $f, g \in X^*$, then $f = g$ iff $\|f - g\| = 0$. If $\|f - g\| > 0$, there must exist $x \in X$ where $f(x) \neq g(x)$.
3. [Fol13, Thm 6.15] If $p, q \in (1, \infty)$ are Holder conjugates.
$$\langle f, \phi\rangle_q = \int_X f(x)\phi(x)\,\mathrm{d}x \quad \text{for all} \quad f \in L^p \quad \text{and} \quad \phi \in L^q.$$
The integrals above converge absolutely, and the scalar product separates L^p and L^q.
$$\langle f, \cdot\rangle_q = \langle g, \cdot\rangle_q \quad \text{iff} \quad f = g \text{ a.e.}$$
$$\langle \cdot, \phi\rangle_q = \langle \cdot, \psi\rangle_q \quad \text{iff} \quad \phi = \psi \text{ a.e.}$$
The same thing holds for $p = 1$ whenever μ is semifinite.
4. [Fol13, Thm 7.17] Let X be a LCH space, $M(X)$ the space of complex Radon measures, $C_0(X)$ be the complex-valued continuous vanishing functions from X. The scalar product
$$\langle \mu, f\rangle = \int_X f(x)\,\mathrm{d}x \quad \text{for all} \quad \mu \in M(X), \quad \text{and} \quad f \in C_0(X)$$
separates $M(X)$ and $C_0(X)$, in other words
$$\langle \mu, \cdot\rangle = \langle \nu, \cdot\rangle \quad \text{iff} \quad \mu = \nu$$
$$\langle \cdot, f\rangle = \langle \cdot, g\rangle \quad \text{iff} \quad f = g.$$

3.11. Norming Sets

Definition 3.12
Let X be a normed vector space, a family of functions $\mathcal{F} = \{f_\alpha : X \to [0, \infty)\}$ is said to be *norming for* X if
$$\|x\| = \sup f_\alpha(x) \quad \text{for every} \quad x \in X.^{a}$$

[a] resp. subnorming, supernorming. if $\geq$ or $\leq$

We can rephrase Proposition 1.3 in terms of the following.

> For any semifinite measure space, $\mathcal{S}^p$ norms L^q in L^1_{loc}, where
> $$\mathcal{S}^p = \{f \in L_0^\infty,\ \|f\|_p = 1\}.$$

[Fol13, 5.8b] For any $x \in X$, if we wish to compute $\|x\|$, it suffices to look at the unit sphere of X^*.
$$\|x\| = \max\{f(x),\ \|f\| = 1\}.$$

3.12. Semifinite Measures

Definition 3.13 [Fol13, pp.25]
A measure μ on $\mathcal{M}$ is *semifinite* if every μ-infinite set E admits a measurable subset F such that $0 < \mu(F) < \infty$.[a]

[a] The authors personally think that this definition is not that great, see Equation (37) for a more intuitive characterization.

A semifinite measure on $\mathcal{M}$ can be thought of as something that is completely characterized by its restriction onto its μ-finite subset. This is similar to the Lebesgue measure on $\mathbb{R}$ or any *Radon* measure on a locally compact space [Fol13, Chap 3.4, 7.1, 7.2]. Once we know the values μ takes on its μ-finite subsets, we can approximate every member in the σ-algebra.

Lemma 3.14 [Fol13, Ex 1.13, 1.14, 1.15]
Ex 1.13 Every σ-finite measure is semifinite.
Ex 1.14 If μ is a semifinite measure, then for every $E \in \mathcal{M}$,
$$\mu(E) = \sup\Big\{\mu(F) < \infty,\ F \subseteq E,\ F \in \mathcal{M}\Big\}. \tag{37}$$
Ex 1.15 If μ is a measure on $(X, \mathcal{M})$, define $\mu_0(E)$ to be the right hand side of Equation (37) for every measurable E. Then i) μ_0 is semifinite, ii) $\mu_0 = \mu$ iff μ is semifinite iii) there is another measure ν such that $\nu(E) \in \{0, \infty\}$ such that $\mu + \mu_0 + \nu$.

Proof. Ex 1.13 Let $A_n \nearrow X$ be an increasing σ-finite exhaustion of X, for any $E \in \mathcal{M}$, if $\mu(E) = \infty$, then $\mu(A_n \cap E) \nearrow \infty$ by continuity from below.

Ex 1.14 Fix $E \in \mathcal{M}$, if $\mu(E) < \infty$, then the supremum is attained in Equation (37) and there is nothing to prove. Suppose $\mu(E) = \infty$, let us denote the supremum by c. It is clear that $c > 0$ (by semifiniteness).

The rest of the proof requires some additional effort and is rather technical. Suppose that $c < \infty$, intuitively this means there is a non-trivial 'gap' between the μ-finite subsets and the μ-infinite subsets of E. We **claim** that the supremum is attained: there exists $F \subseteq E$, $F \in \mathcal{M}$, such that $\mu(F) = c$. Postponing the proof for the claim, we will show that this gives us a contradiction.

An easy calculation will show that $\mu(E \setminus F) = \infty$, and by semifiniteness again: we can find $A' \subseteq E \setminus F$ (which is disjoint from F), with $0 < \mu(A) < \infty$. Using additivity, we can improve upon the supremum by at least $\mu(A)$:

$$\mu(F \cup A) = \mu(F) + \mu(A) > c.$$

which contradicts that c is the supremum.

Onto the proof of the **claim**. Let $E_1 = E$, $A_1 \subseteq E_1$, with $\mu(A_1) \in [c2^{-1}, c]$ and then set $E_2 = E_1 \setminus A_1$. Suppose we have $A_1, \ldots, A_{n-1}$ chosen, such that

$$A_j \subseteq E_j = E_{j-1} \setminus A_{j-1} \quad \forall j = 2, \ldots, n-1,$$
$$\mu(A_j) \in [c2^{-j},\ c2^{-(j-1)}] \quad \forall j = 1, \ldots, n-1.$$

This technique is very similar to the decomposition of the unit interval into almost-disjoint intervals

$$[0,1] = \bigcup_1^\infty [2^{-j},\ 2^{-(j-1)}].$$

Set $E_n = E_{n-1} \setminus A_{n-1}$, which has ∞ measure. By additivity, one can argue that

$$\sup\left\{\mu(A) < \infty, A \subseteq E_n,\ A \in \mathcal{M}\right\} \le \sum_n^\infty c2^{-j} = c2^{1-n}.$$

There exists a measurable subset A_n of E_n, with $\mu(A_n) \in [c2^{-n}, c2^{-(n-1)}]$. The principle of induction gives us a disjoint sequence $\{A_n\}_1^\infty$, whose union attains the supremum:

$$\mu(\cup_1^\infty A_n) = \sum_1^\infty \mu(A_n) \ge \sum_1^\infty c2^{-n} = c.$$

This completes the proof.

Ex 1.15 Left as an exercise. ■

For strictly semifinite measure spaces, the proof of Proposition 1.3 requires the following lemma, to make sure that the support of g is σ-finite.

> **Lemma 3.15 [Fol13, Ex 6.17]**
> If $g \in L^1_{loc}$, $q \in [1,\infty)$ and suppose that $M_q(g) = \sup\{|\langle g, f\rangle|,\ f \in L_0^\infty,\ \|f\|_p = 1\} < \infty$. Then,
> **1.** for any $\varepsilon > 0$, $\mu\big(\{x \in X,\ |g(x)| > \varepsilon\}\big) < \infty$, and
> **2.** $\operatorname{supp}(g) = \{x \in X,\ |g(x)| \neq 0\}$ is σ-finite.

Proof of Lemma 3.15. If $\mu(\operatorname{supp}(g)) < \infty$, both of the claims are immediate, so we are free to assume $\mu(\operatorname{supp}(g)) = \infty$. For any $\varepsilon > 0$, let us write

$$A_\varepsilon = \{|g| > \varepsilon\}.$$

The semifiniteness of μ (see Lemma 3.14) means that we can approximate $\mu(A_\varepsilon)$ by its subsets of finite measure.

$$\mu(A_\varepsilon) = \sup\{\mu(B),\ 0 < \mu(B) < \infty,\ B \subseteq A,\ B \in \mathcal{M}\}.$$

For any such B, we can make a clever choice of f that gives us a uniform upper bound on $\mu(B)$. Choose

$$f = \overline{(\operatorname{sgn} g)}\chi_B \in L_0^\infty \quad \text{which gives us} \quad \langle g, f\rangle = \int_B |g(x)|\,\mathrm{d}x. \tag{38}$$

If $p \in [1,\infty)$ the relative largeness of f depends on $\mu(B)$, whereas if $p = \infty$, this 'largeness' equals 1:

$$\|f\|_p = \begin{cases} \mu(B)^{1/p} & p \in [1,\infty) \\ 1 & p = \infty. \end{cases}$$

To obtain a lower bound for the integral on the right of Equation (38), notice that the simple function $\varepsilon\chi_A$ is a *subordinate* of $\chi_A|g|$. By the definition of the integral on L^+, we see that

$$\varepsilon\mu(B) \le \int_B |g(x)|\,\mathrm{d}x \le \|f\|_p M_q(g).$$

Developing this further, we see that

$$\mu(B)^{1/q} \le M_q^q(g)\varepsilon^{-q} \quad \forall q \in [1,\infty).$$

It follows upon taking the supremum over all such B, that

$$\mu(A_\varepsilon) \le M_q^q(g)\varepsilon^{-q} < \infty. \tag{39}$$

■

3.13. Measure Concentration

An estimate such as the one in Eq. (39) is also known as a *Chebyshev-type estimate*. More generally, the *Chebyshev inequality* [Fol13, Thm 6.17] tells us that for any complex-valued measurable function f,

$$\mu\big(\{|f| > \alpha\}\big) \le \|f\|_p^p \alpha^{-p} \quad \text{for all} \quad \alpha > 0,\ p \in (0,\infty). \tag{40}$$

Let $(X,\mathcal{M})$ and $(Y,\mathcal{N})$ be measurable spaces. Estimates such as Eq. (40) help us collect qualitative information about the asymptotic behaviour of a mapping $T : \mathcal{E}(\mathcal{M}) \to \mathcal{E}(\mathcal{N})$, that is not necessarily linear or continuous.

These types of estimates are used in studying the rate of convergence of a stochastic learning algorithm and is also called a *concentration inequality*. The word 'concentration' refers to the concentration of measure, or mass inside the closed f*-balls* $\{x \in X,\ |f(x)| \le c\}$ where c should be thought of as the radius of such a ball. See Boucheron 2013 for a comprehensive overview on this subject.

One should think of conv. in meas. as a natural extension of convergence w.r.t. the L^∞ norm, but for functions which may not even be in L^∞ (see Def. 3.16). It is an easy exercise for the reader to verify that $\|f_n - f\|_\infty \to 0$ if and only if for every $\varepsilon > 0$,

$$\{\mu(\{|f_n - f| > \varepsilon\})\}_1^\infty \in l_0.$$

Definition 3.16
Let $f_n, f \in \mathcal{E}(\mathcal{M})$, we say that $f_n \to f$ *in measure* if for every $\varepsilon > 0$, the sequence formed by the numbers $\mu(\{|f_n - f| > \varepsilon\})$ is in c_0.

The notation c_0 comes from point set topology. If X is a LCH space (see [Fol13, Chap 4.5]), then

$C_0(X)$ = complex-valued, continuous, vanishing functions on X.

If we equip $\mathbb{N}^+$ with the *discrete topology* (see [Fol13, Chap 4.1]), then $c_0 = C_0(\mathbb{N}^+)$ [Fol13, Ex 7.20]. The reason why we say that conv. in meas. is a generalization of L^∞ convergence is because c_0 is the uniform closure of l_0, and a lot of the desirable properties of L^∞ convergence such as i) $\mu(|f_n - f| > \varepsilon) < \infty$, and pointwise a.e. convergence (albeit by passing to a subseq.) are inherited.

Additional theorems relating the three types of convergences: i) pw a.e., ii) in measure, iii) in norm are proven in the remaining sections of this article.

3.14. Convergence of Integrals (given pw a.e.)

This section will answer the question of why we care about pw a.e. convergence at all, since it is relatively weak compared to the other types of convergences. Given $f_n \to f$ pw a.e., and a functional $I : \mathcal{E}(\mathcal{M}) \to \mathbb{C}$, we want to know when does $I(f_n) \to I(f)$. The general case is out of the scope of this document. We refer the reader to the remarks left at the end of this section for references.

Specializing, let $p \in [1,\infty)$, and suppose $f_n, f \in L^p$ and $f_n \to f$ pw a.e. We know that by Fatou's lemma,

$$\textstyle\int_X |f(x)|^p \, \mathrm{d}x \le \liminf \int_X |f_n(x)|^p \, \mathrm{d}x.$$

Natural questions to ask: i) When does $\|f_n - f\|_p \to 0$? ii) How about $\|f_n\|_p \to \|f\|_p$? Proposition 3.18 shows that these two things are equivalent. In other words,

> the diff. of norms vanishes iff the norm of the diff vanishes. (diff. = difference)

To prove this we need a small lemma.

Proposition 3.17 Generalized DCT, [Fol13, Ex 2.20]
Let $f_n, f, g_n, f \in L^1$, suppose that $f_n \to f$ pw a.e, and $g_n \to g$ pw a.e.,

$$|f_n| \le g_n \quad \text{and} \quad \textstyle\int g_n(x)\,\mathrm{d}x \to \int g(x)\,\mathrm{d}x.$$

then $\int f_n(x)\,\mathrm{d}x \to \int f(x)\,\mathrm{d}x$.

Proof. Left as an exercise. Caution: pay attention to which integral you are using, the L^1 integral for non-negative functions or the L^+ integral. The first one allows subtraction under the integral sign, while the second one may not. ■

Proposition 3.18 [Fol13, Ex 2.21, 6.10]
Let $p \in [1,\infty)$, $f_n, f \in L^p$. Suppose $f_n \to f$ a.e., then $\|f_n - f\|_p \to 0$ iff $\|f_n\|_p \to \|f\|_p$.

Proof. The proof requires Prop. 3.17, and is left as an exercise. Hint: use Eq. (41). ■

The *Brezis-Leib lemma*, which is a generalization of Fatou's lemma is used frequently in the calclulus of variations to obtain estimates of functionals just from pw a.e. convergence. As a special case of this, let $p \in (0,\infty)$, a sequence of measurable functions $f_n \to f$ pointwise a.e., such that $\limsup \|f_n\|_p < \infty$ for some $p \in (0,\infty)$, then

$$\int_X |f(x)|^p \, \mathrm{d}x = \lim \left(\int_X |f_n(x)|^p \, \mathrm{d}x - \int_X |f_n(x) - f(x)|^p \, \mathrm{d}x \right).$$

For more, refer to wikipedia, and Brezis-Leib 1983. [LL01] also has a discussion on this in the earlier chapters.

3.15. Convergence of Integrals (given in meas.)

Now we discuss the situation where we only have convergence in measure. The first result is rather elementary, because it does not assume that the sequence f_n has uniformly bounded L^p norm nor that it is uniformly bounded.

Proposition 3.19
Let $f_n \ge 0$, and $f_n \to f$ in measure, then $f \ge 0$ a.e., and $\int f(x)\,\mathrm{d}x \le \liminf \int f_n(x)\,\mathrm{d}x$.

Proof. Let f_k be a subsequence such that $\lim \int f_k(x)\,\mathrm{d}x = \liminf \int f_{n_k}(x)\,\mathrm{d}x$. Taking subseq. preserves conv. in measure, so we can take a further subsequence of f_k (still labelled by f_k) such that $f_k \to f$ pointwise. Since $f_k \ge 0$ a.e., we see that $f \ge 0$ a.e. as well. Using Fatou's Lemma,

$$\begin{aligned} \textstyle\int f(x)\,\mathrm{d}x = \int \liminf f_k(x)\,\mathrm{d}x &\le \liminf \textstyle\int f_k(x)\,\mathrm{d}x \\ &= \liminf \textstyle\int f_n(x)\,\mathrm{d}x. \end{aligned}$$

■

If we assume that $|f_n|$ is controlled by some integrable function, then we obtain much better results. The key trick is to pass to subsequences. It is useful to remember that ([Fol13, Ex 4.30b]) the very overpowered criteria for checking when a sequence converges.

Let $\langle x_\alpha \rangle$ be a net in a topological space X. The net $\langle x_\alpha \rangle \to x \in X$ iff for every cofinal $B \subseteq A$ there is a cofinal $C \subseteq B$ such that $\langle x_\gamma \rangle_{\gamma \in C} \to x$. (Note that every cofinal subset $B \subseteq A$ induces a subnet $\langle x_\beta \rangle_{\beta \in B}$, see [Fol13, Ex 4.30a])

Proposition 3.20

Let $p \in [1,\infty)$, and $f_n \to f$ in measure, suppose that $|f_n| \le g \in L^p$ pointwise a.e., then
1. $f \in L^p$, $\|f_n \to f\|_p \to 0$, and $\|f_n\|_p \to \|f\|_p$.
2. If $p = 1$, then $\int f_n(x)\,\mathrm{d}x \to \int f(x)\,\mathrm{d}x$

Proof. **1.** We will use the subsequential criterion. Fix a subsequence f_{n_j}. By choosing a further subsequence, f_{n_k}, we can assume that $f_{n_k} \to f$ pointwise a.e. For every $k \geq$, we have the uniform dominance

$$|f_{n_k} - f|^p \le 2^p\left(|g|^p + |f|^p\right) \in L^1. \tag{41}$$

This proves that $\|f_{n_k} - f\|_p \to 0$ as $k \to \infty$, and coincidentially also proves $\|f_n - f\|_p \to 0$ as $n \to \infty$. The argument for $\|f_n\|_p \to \|f\|_p$ is similar, we just dominate:

$$|f_{n_j}|^p \le |g|^p \in L^1.$$

2. Same subsequential technique. If f_{n_j} is a subsequence, choose a further subseq such that $f_{n_k} \to f$ a.e. we know that from Part 1 that $f \in L^1$, and because

$$|f_{n_k}| \le |g| \in L^1,$$

by the dominated convergence theorem: $\int f_{n_k}(x)\,\mathrm{d}x \to \int f(x)\,\mathrm{d}x$. This holds for every subsequence f_{n_j}, and we conclude $\int f_n(x)\,\mathrm{d}x \to \int f(x)\,\mathrm{d}x$.

■

two general themes of convergence in measure

1. conv. in measure is preserved under taking subsequences.
2. conv. in measure implies conv. pointwise a.e for some subseqeuence
3. if we want to show that for some functional $I(f_n) \to I(f)$, we can use the subnet convergence trick. Fix any subseq. $I(f_{n_j})$, it suffices to show that there is a further subseq. with $I(f_{n_k}) \to I(f)$.

   In choosing this further subseq. we can assume that $f_{n_k} \to f$ pointwise a.e. If in addition the entire sequence is uniformly dominated, then we can use DCT or GDCT.
4. to compute the lim inf or lim sup of a functional, we can assume that the value is attained by a subsequence. (which inherits conv. in measure).

3.16. Convergence in Measure for Bounded Measures

The topology of conv. in meas. becomes *completely metrizable* if $\mu(X) < \infty$. In this case, it is customary to use the metric

$$\rho(f,g) = \int_X \frac{|f(x) - g(x)|}{1 + |f(x) - g(x)|}\,\mathrm{d}x \quad \text{where} \quad f, g \in \mathcal{E}(X). \tag{42}$$

For proofs and more generalizations, see [Fol13, Ex. 2.32 4.56, 5.50, Thm 2.30]. We remark that $\mu(X) < \infty$ necessary but not a sufficient condition. In the case that μ is a finitely-additive vector measure, there is a notion of integration and a formula similar to Eq. (42) that metrizes conv. in meas., we refer the reader to [DS58, Chap 3].

References

[Fol13] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, June 2013.

[Rud76] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill Publishing Company, 1976. ISBN: 0070856133.

[Yos12] Kosaku Yosida. *Functional Analysis*. Springer Science & Business Media, Dec. 2012. ISBN: 9783642618598.

[DS88] Nelson Dunford and Jacob T. Schwartz. *Linear operators, part 1: General theory*. [Online; accessed 2024-11-03]. John Wiley & Sons, Feb. 1988.

[HS12] E. Hewitt and K. Stromberg. *Real and Abstract Analysis: A modern treatment of the theory of functions of a real variable*. [Online; accessed 2024-11-03]. Springer Science & Business Media, Dec. 2012.

[Tol20] John Toland. *The dual of $L_\infty(X, \mathcal{L}, \lambda)$, finitely additive measures and weak convergence: A primer*. [Online; accessed 2024-11-03]. Springer Nature, Jan. 2020.

[LL01] Elliott H. Lieb and Michael Loss. *Analysis*. American Mathematical Soc., 2001. ISBN: 9780821827833.

[DS58] Nelson Dunford and Jacob T. Schwartz. *Linear Operators: Spectral theory*. [Online; accessed 2024-11-03]. 1958.