Intuitive Proofs of Measure & Probability Theorems

Xingjian Wang

January 9, 2026

I will attempt to organize and show several theorems from measure-theoretic probability in the most intuitive manner. This does not mean that the proofs are concise; rather, I expect the exact opposite. The advantage of this approach is that the exposition remains readable and that each step of the argument follows one's normal intuition; our proof ideas are explicitly motivated and traceable, rather than relying on abrupt manipulations.

We expect the reader to bear the burden of summarizing and condensing the material, as we only provide the most stream-of-consciousness proofs, yet no rigor is sacrificed. The goal here is that the proof ideas are completely dissected and understood upon the first close reading and leaves no fuzzy understanding, while we remain completely rigorous throughout the document.

First, we attempt to show the Riesz-Weyl subsequence theorem. We can do this by first revisiting the set theoretic expressions of convergence and using Borel-Cantelli, before relying on a diagonal trick to find a subsequence.

Prop 1. For a sequence of functions (f_n) sending $(\Omega, \mathcal{F}, \mu)$ to $\mathbb{R}$ (matter of fact, any metric space works), we denote the set of points x on which f_n does not converge to f pointwise as $\{x : \exists \epsilon > 0 \text{ s.t. } |f_n(x) - f(x)| \geq \epsilon \text{ i.o.}\}$, or $\{x : \exists \epsilon > 0 \text{ s.t. } \forall N \in \mathbb{N}, \exists n \geq N \text{ s.t. } |f_n(x) - f(x)| \geq \epsilon\}$, or

$$\bigcup_{\epsilon>0} \left\{ \bigcap_{N=1}^{\infty} \bigcup_{n \geq N} \{x : |f_n - f|(x) \geq \epsilon\} \right\}$$

where by definition, inside the braces we have

$$\limsup_n \{x : |f_n - f|(x) \geq \epsilon\}.$$

Or, we can represent the entire thing as

$$\bigcup_{m=1}^{\infty} \left\{ \bigcap_{N=1}^{\infty} \bigcup_{n \geq N} \{x : |f_n - f|(x) \geq \epsilon_m\} \right\}$$

if we take $\epsilon_m \downarrow 0$. These are equal, as one direction is simply inclusion, and the other reduces to the statement "for all $\epsilon > 0$, we can find $\epsilon_m < \epsilon$" from where an inclusion is established by mapping ϵ to some smaller ϵ_m.

Taking complement and flipping everything set theoretically gives points that do converge. Think in the language of $\liminf$.

Lemma 2. (Borel-Cantelli) For a sequence of events (E_n), suppose that $\sum_n \mu(E_n) < \infty$, then $\mu(\limsup_n E_n) = 0$.

Proof. Assume that the summation identity holds, then $\sum_{n \geq N} \mu(E_n) \to 0$ as $N \to \infty$. Since also $\mu\left(\bigcup_{n \geq N} E_n\right) \leq \sum_{n \geq N} \mu(E_n)$, we have

$$\mu\left(\lim_{N\to\infty} \sup_{n \geq N} E_n\right) = \lim_{N\to\infty} \mu\left(\bigcup_{n \geq N} E_n\right) \leq \lim_{N\to\infty} \sum_{n \geq N} \mu(E_n) = 0$$

where the first equality holds by continuity from above, and second equality holds by our previous statement.

□

Remark. In the language of probability, we have that if $\sum_n \mathbb{P}[E_n] < \infty$ then $\mathbb{P}[\limsup_n E_n] = 0$.

These two results allow us to bridge convergence in measure (in probability) and convergence almost everywhere (almost surely).

Theorem 3. (Riesz-Weyl, Weak Version) For any $f_n \xrightarrow{\text{m}} f$, we may find some subsequence $(n_k)_{k=1}^{\infty}$ such that $f_{n_k} \xrightarrow{\text{a.e.}} f$.

Proof. We want to find a subsequence f_{n_k}. We want for all $\epsilon > 0$ that $\mu(\{x : |f_{n_k}(x) - f(x)| \geq \epsilon \text{ i.o.}\}) = 0$ holds.

Since $\{x : |f_{n_k}(x) - f(x)| \geq \epsilon \text{ i.o.}\} = \limsup_k \{x : |f_{n_k}(x) - f(x)| \geq \epsilon\}$, we label this as $\limsup_k E^{\epsilon}_{n_k}$ and try to show $\mu\left(\limsup_k E^{\epsilon}_{n_k}\right) = 0$ by showing $\sum_k \mu(E^{\epsilon}_{n_k}) < \infty$ and using Borel-Cantelli.

We run into a problem; for a fixed ϵ, we can pick n_k accordingly so that $\mu(E^{\epsilon}_{n_k}) < 2^{-k}$ (by the definition of convergence in measure) so the sum is finite. However, for some other $\epsilon' < \epsilon$, this subsequence does not work.

We patch that by the standard diagonal argument where we find $\eta_k \downarrow 0$ and each n_k satisfying $\mu(E^{\eta_k}_{n_k}) < 2^{-k}$. The interpretation is that at n_k, only on a 2^{-k}-sized domain does f_{n_k} get η_k-far from f. Then, for any $\epsilon > 0$, find K such that $\eta_k < \epsilon$ if $k \geq K$, so that the first $\mu(E^{\eta_1}_{n_k})$ to $\mu(E^{\eta_K}_{n_k})$ terms sum to a finite measure, and all $k > K$ has $\mu(E^{\epsilon}_{n_k}) < \mu(E^{\eta_k}_{n_k})$, since suppose

$\eta_k < \epsilon$ then more x enables f_{n_k} to get η_k-far than ϵ-far. This yields the inequality below. Specifically, we have

$$\begin{aligned}
\sum_{k=1}^{\infty} \mu\left(E_{n_k}^{\epsilon}\right) &= \sum_{k=1}^{K} \mu\left(E_{n_k}^{\epsilon}\right) + \sum_{k=K+1}^{\infty} \mu\left(E_{n_k}^{\epsilon}\right) \\
&\leq \sum_{k=1}^{K} \mu\left(E_{n_k}^{\epsilon}\right) + \sum_{k=K+1}^{\infty} \mu\left(E_{n_k}^{\eta_k}\right)
\end{aligned}$$

and the first part is finite while the second part has each term bounded by 2^{-k}. Therefore, $\sum_{k=1}^{\infty} \mu\left(E_{n_k}^{\epsilon}\right)$ is finite, moreover $\mu\left(\limsup_k E_{n_k}^{\epsilon}\right) = 0$ by Borel-Cantelli. Since $\limsup_k E_{n_k}^{\epsilon} = \{x : |f_{n_k}(x) - f(x)| \geq \epsilon \text{ i.o.}\}$, we have for any ϵ the sequence (f_{n_k}) gets ϵ-close almost everywhere. Let $\epsilon_m \downarrow 0$ so $\mu(\{x : |f_{n_k}(x) - f(x)| \geq \epsilon_m \text{ i.o.}\}) \uparrow \mu(\{x : \exists \epsilon \text{ s.t. } |f_{n_k}(x) - f(x)| > \epsilon \text{ i.o.}\})$, by continuity from below. This gives us precisely the nonconvergent domain, where $0 \uparrow 0$, and we are done.

To summarize, let $\eta_k \downarrow 0$ and for each η_k pick n_k so that f_{n_k} is η_k-close to f barring a 2^{-k}-measure set. For each ϵ, "$|f_{n_k} - f|$ exceeding ϵ i.o." is null because $\sum_k \{|f_{n_k} - f| \text{ exceeding } \epsilon\} < \infty$, relying on some $\eta_k < \epsilon$, then Borel-Cantelli. Intersect over ϵ to get $\mu(\text{nonconvergent domain}) = 0$.

□

Remark 1. Note that from "f_{n_k} is η_k-close barring a 2^{-k}-measure set", we can get a.u. easily by switching this to $2^{-k}\delta$ if we require $\mu(E) < \delta$. Then $E = \bigcup_{k=1}^{\infty} \{|f_{n_k} - f| \geq \eta_k\}$. For any ϵ find $\eta_K < \epsilon$ and "unconquered domain starting at K" is exactly $\bigcup_{k \geq K} \{|f_{n_k} - f| \geq \epsilon\}$, which is contained in $\bigcup_{k \geq K} \{|f_{n_k} - f| \geq \epsilon_k\}$, and is hence contained in E.

The bridge from a.u. to a.e. using most textbook-standard proof works exactly the same as the original direction we took, which lets $\limsup$ be expressed and goes down to 0, either by Borel-Cantelli or any one line argument (completely capturing Borel-Cantelli) that looks at the tail (the sup).

Refer to the strong Riesz-Weyl.

Remark 2. Notice that ϵ_m exists only for countability. The statement $\{x : |f_{n_k}(x) - f(x)| \geq \epsilon_m \text{ i.o.}\} \uparrow \{x : \exists \epsilon \text{ s.t. } |f_{n_k}(x) - f(x)| > \epsilon \text{ i.o.}\}$ simply captures $\bigcup_{\epsilon_k} \{\cdot\}$ vs $\bigcup_{\epsilon > 0} \{\cdot\}$'s equality, which we showed in prop 1.

Next we will visit some standard theorems that rely on similar techniques. The ones shown above are quite universal in proving convergence.

One must be familiar with the language of i.o. vs ult., $\limsup$ vs $\liminf$, and $\bigcup_N \bigcap_{n\geq N}$ vs $\bigcap_N \bigcup_{n\geq N}$ before proceeding.

Theorem 4. (Egorov) Assume $\mu(X) < \infty$. Then $f_n \xrightarrow{\text{a.e.}} f$ gives $f_n \xrightarrow{\text{a.u.}} f$.

Remark. Without any restrictions, convergent almost uniformly implies convergent almost everywhere. Adding $\mu(X) < \infty$ establishes that as N grows, the measure of "unconquered x at N" remains finite and goes to zero. This move establishes the reverse implication.

Proof. For any $\eta > 0$, we need to find E such that $\mu(E) < \eta$ and that f_n converges uniformly to f on $X \setminus E$.

Let's examine pointwise convergence more carefully and see what stops us from getting uniform convergence. Fix $\epsilon > 0$, assuming convergence a.e., then for each x in a full set, exists $N = N(x, \epsilon)$ such that $|f_n(x) - f(x)| < \epsilon$ for all $n \geq N$ (definition). Hence, for each N there is a set $X_N^\epsilon \subset X$ on which $|f_n(x) - f(x)|$ stays $< \epsilon$ starting at N. In other words, some x are "conquered" at N and makes $f_n(x)$ remain ϵ-close to $f(x)$ starting at N. Taking union over all N gives a full set in X, as each x must pick an X_N^ϵ.

Let's look at x that remains "unconquered" after N, i.e. points that still allow a ϵ-deviation somewhere after time N. Suppose E_n^ϵ denotes the event $\{x : |f_n(x) - f(x)| > \epsilon\}$. Then $X \setminus X_N^\epsilon = \bigcup_{n\geq N} E_n^\epsilon$, with the interpretation that x "remains unconquered at N" equals "still allows an ϵ-deviation at some $n \geq N$". This is a decreasing sequence that converge to a null set.

Notice the interpretation $\bigcup_{n\geq N} E_n^\epsilon = \sup_{n\geq N} E_n^\epsilon \downarrow \limsup_n E_n^\epsilon$, denoting points on which (f_n) never get fully ϵ-close. This has measure 0 by definition. We try to think about its relationship with the aforementioned "nonconvergent set", which is also represented with $\limsup$. The finiteness of $\mu(X)$ allows for continuity from above, where $\mu(\bigcup_{n\geq N} E_n^\epsilon) < \infty$ works for any N. This is the only use of requiring $\mu(X) < \infty$.

Suppose we want to build E, then for all ϵ, it suffices to pick some N_ϵ, build X_N^ϵ (the conquered set), and throw the remaining unconquered x into E. The only task is to bound $\mu(E) < \eta$. We have two tools; one is the freedom to vary N_ϵ, while the other is countability trick, where instead of including all ϵ, we only include all ϵ_k, where any $(\epsilon_k) \downarrow 0$ works. This way for any ϵ, find $\epsilon_k < \epsilon$ so that given the N_k we pick, the unconquered set $X \setminus X_{N_k}^{\epsilon_k}$ includes $X \setminus X_{N_k}^\epsilon$. The interpretation is that x that allows ϵ-deviation after N_k must also allow ϵ_k-deviation after N_k. Same idea as Riesz-Weyl.

From this construction, if for each ϵ_k, E includes its unconquered set (each built with a corresponding N_k), then every ϵ can rely on some ϵ_k, to borrow its N_k so that if $X \setminus X_{N_k}^{\epsilon_k}$ is in E, then $X \setminus X_{N_k}^\epsilon$ is also in E.

Then our problem comes down to picking each N_k for ϵ_k. Since for each $\epsilon > 0$, its unconquered set $X \setminus X_N^\epsilon = \sup_{n \geq N} E_n^\epsilon$ converges to a null set (by a.e. convergence) as $N \to \infty$, we can find for each ϵ_k some N_k so that $\mu(X \setminus X_{N_k}^{\epsilon_k}) < 2^{-k}\eta$. Letting $E = \bigcup_k (X \setminus X_{N_k}^{\epsilon_k})$, summing over ϵ_k gives a bound of $\mu(E) < \eta$. We are done.

To summarize, given η, let $\epsilon_k \downarrow 0$ and for each k pick N_k so the "un-ϵ_k-conquered x after N_k", $\mu(\{x : \exists n \geq N_k \text{ s.t. } |f_n - f|(x) > \epsilon_k\}) < 2^{-k}\eta$, which is possible as "$..._k$ after N" decreases to a null set as $N \to \infty$, using continuity from above. Make its union over all k precisely E, then for any ϵ pick $\epsilon_k < \epsilon$ and use N_k; we get uniformly $\|f_n - f\|_\infty < \epsilon$ on $X \setminus E$.

□

We employ all the techniques we have used to tackle the strong version of Riesz-Weyl.

Theorem 5. (Riesz-Weyl, Strong Version) If (f_n) is measure-Cauchy, then it has an a.u. (hence a.e.) convergent subsequence $(f_{n_k}) \to f$. Moreover, this would imply that (f_n) itself is measure convergent to f.

Proof. At this point, all the tools were already used. Extending to this result is not a huge jump from the previous two results. It would be a good exercise to attempt the proof freehand.

Find an a.e. Cauchy (hence convergent) subsequence using the same mechanism as the weak Riesz-Weyl, and use the second half of Egorov to establish that since "unconquered set at n_k" goes to null as $k \to \infty$, find each N_k so that the tail sum (bounds the measure of E) is controlled, and this N_k works exactly as the proof for Egorov.

Or, as we remarked above, first show a.u. then show a.e., which requires us to rescale the 2^{-k} but saves the work to restate the Egorov argument (which is not avoided in the proof for a.u. but is more direct given that n_k are already constructed).

The new part is to conclude that f_n is measure-convergent. This only relies on a one-line triangle inequality and can be done freehand.

□

I will briefly go over the work to build $\mathcal{L}^1$, but not before laying out a few propositions, for which the proofs I will omit as they employ the same techniques as covered and/or are of much less difficulty.

Lemma 6. (A.U. Strongest) If (f_n) is a.u. Cauchy, then it is a.u. convergent to some f, and a.e. convergent and measure convergent to this f as well.

Prop 7. (Uniqueness) No matter $(f_n) \xrightarrow{\text{m}} f$ or $(f_n) \xrightarrow{\text{a.u.}} f$ or $(f_n) \xrightarrow{\text{a.e.}} f$, we get that f is a.e. unique.

And now we add the $\|\cdot\|_1$ seminorm onto the space of step functions $\mathrm{St}(X, \mathcal{S}, \mu, B)$, where B is a Banach space with norm $|\cdot|$ (for simplicity's sake, instead of writing $\|\cdot\|_B$). Denote this space's completion w.r.t. $\|\cdot\|_1$ as L^1. Note that this is the space of equivalence classes of mean Cauchy sequences of step maps.

But each element here does not necessarily translate to an actual function on X, because mean Cauchy convergence does not necessarily give pointwise convergence. We want to build the space of integrable functions that rely on a few key lemmas on step functions WHAT?

DEFINE THE INTEGRALS OF STEP FUNCTIONS. ESTABLISH COMPLETENESS and UNIQUENESS AE.

Prop 8. (Uniform Mass Concentration) Suppose $f \in \mathcal{L}^1$, then for any $\epsilon > 0$, we have $\mu(\{|f| \geq \epsilon\}) < \infty$.

This is saying/shown by that if $|f|$ exceeds ϵ on some measure-infinite set, then its $\mathcal{L}^1$-norm is lower bounded by ∞.

Prop 9. ($\mathcal{L}^1$ Mass Concentration) Suppose $f \in \mathcal{L}^1$, then for any $\epsilon > 0$, some finite-measured $X' \subset X$ exists such that $\int_{X \setminus X'} |f| \, d\mu < \epsilon$.

This is shown by the fact that step maps can get arbitrarily $\mathcal{L}^1$-close to f while being nonzero only on a finite-measured set. This deserves the title of a lemma due to later constructions with indefinite integrals.

Lemma 10. (Absolute Continuity w.r.t Domain) Suppose $f \in \mathcal{L}^1$, then for any $\epsilon > 0$, there exists $\delta > 0$ such that if $\mu(E) < \delta$, $\int_E |f| \, d\mu < \epsilon$.

Proof. Pick step function ϕ that is an $\epsilon/2$-approximation of f in the $\mathcal{L}^1$ sense, then $\|\phi\|_\infty < \infty$, and $\int_E |\phi| \, d\mu \leq \|\phi\|_\infty \mu(E)$, which can be made small enough by δ to be under $\epsilon/2$. Use the triangle inequality

$$\|f\|_1 \leq \|g\|_1 + \|f - g\|_1$$

on E so that $\|f|_E\|_1 < \epsilon$. $\square$

Let's focus on some heavy hitters. Dominated Convergence is obtained via the tools above, and its restrictions make it obvious why these tools work. Bounding $|f_n|$ uniformly with an integrable $|g|$ is elegant enough to satisfy all the requirements, while capturing many common scenarios.

Theorem 11. (Dominated Convergence) Suppose $(f_n) \subset \mathcal{L}^1$, $f_n \xrightarrow{\text{a.e.}} f$. If there exists $g \in \mathcal{L}^1$ such that $|f_n| < |g|$ a.e., then $f_n \xrightarrow{\|\cdot\|_1} f$.

Proof. We only need to establish that (f_n) is mean Cauchy. The completeness of $\mathcal{L}^1$ and a.e. uniqueness of f immediately yield mean convergence.

Intuitively, with the mass concentration lemma in mind, if we can get f_n to be arbitrarily close to f on a measure-finite domain X', while controlling the $\mathcal{L}^1$ norm outside of X', getting everything under ϵ, then we are done.

Indeed, the first step is doable via Egorov, and the second step is explicitly enabled with the uniformly-bounding g. What g does here is, for any ϵ, to give a uniform (applicable to all n) finite-measured X'_ϵ so that Egorov is used on X'_ϵ, while $|g|$ and all $|f_n|$ integrate to less than ϵ outside of X'_ϵ.

Let's flesh out the details. First, for any $\epsilon > 0$ there is some X'_ϵ such that for any n, m, we have (since $|g| \geq |f_n|$ for all n) that

$$\begin{aligned}
\|f_n - f_m\|_1 &= \left(\int_{X'_\epsilon} |f_n - f_m| \, d\mu\right) + \left(\int_{X \setminus X'_\epsilon} |f_n - f_m| \, d\mu\right) \\
&\leq \left(\int_{X'_\epsilon} |f_n - f_m| \, d\mu\right) + \left(\int_{X \setminus X'_\epsilon} |f_n| + |f_m| \, d\mu\right) \\
&< \left(\int_{X'_\epsilon} |f_n - f_m| \, d\mu\right) + 2\epsilon.
\end{aligned}$$

Okay. We successfully controlled what happens outside by prop 9 by splitting X into X'_ϵ and $X \setminus X'_\epsilon$. We can now use Egorov on the first term.

Since $f_n \xrightarrow{\text{a.e.}} f$ means that on the finite-measured X'_ϵ, we have $f_n \xrightarrow{\text{a.u.}} f$, meaning that $f_n \rightrightarrows f$ on $X'_\epsilon \setminus G$ for some $G \subset X'_\epsilon$, where $\mu(G)$ can be made arbitrarily small. Let's tackle both the integral on $X'_\epsilon \setminus G$ and on G.

First, note that $|f_n - f_m|$ can get arbitrarily small if n, m is large enough, and $\mu(X'_\epsilon \setminus G) \leq \mu(X'_\epsilon) < \infty$ is fixed before we pick how large n, m gets, so there seems no problem to $\int_{X'_\epsilon \setminus G} |f_n - f_m| \, d\mu$ arbitrarily small.

Meanwhile, we have $\int_G |f_n - f_m| \, d\mu$ upper bounded by $\int_G 2|g| \, d\mu$, which by lemma 10 can also be made arbitrarily small by decreasing $\mu(G)$. We can first pick G to bound this integral first, before setting n, m appropriately large, so $\int_G$ and $\int_{X'_\epsilon \setminus G}$ both get bounded.

So the order is as follows: given ϵ, first pick X'_ϵ to concentrate the mass on a finite domain, making the "leftovers" on $X \setminus X'_\epsilon$ small ($< \epsilon/3$). Then, carve G out of X'_ϵ to use Egorov on X'_ϵ, relying on uniform-Cauchyness to make n, m large enough so $|f_n - f_m|$ integrates to $< \epsilon/3$ on $X'_\epsilon \setminus G$.

However, before that step, make sure that G was selected small enough so $|f_n - f_m| \leq 2|g|$ integrates to another $< \epsilon/3$.

Retrospectively, the condition given in this proof that g uniformly bounds all f seems to be "patching" the errors we would have encountered; specifically, if we didn't have this condition, X'_ϵ being the union of all ϵ-close for each f_n might not have finite measure, which would not have enabled Egorov. The same G would not have worked for all n, m so we cannot pick G before setting how large n, m is. Note that setting n, m to bound the second term does not depend on how large G is, since $X'_\epsilon \supset (X'_\epsilon \setminus G)$. We only need G selected to bound the third term.

Let's proceed with aforementioned strategy. For any $\epsilon > 0$, by prop 9 there must exist X'_ϵ, finite-measured, such that

$$\int_{X \setminus X'_\epsilon} |f_n - f_m| \, d\mu < \epsilon/3.$$

Then, by Egorov, $f_n \xrightarrow{\text{a.u.}} f$ on X'_ϵ. By lemma 10, some δ exists such that if $G \subset X'_\epsilon$ and $\mu(G) < \delta$, then

$$\int_G |f_n - f_m| \, d\mu < \epsilon/3.$$

Pick Egorov's G as such. Now on $X'_\epsilon \setminus G$, pick n, m large enough such that

$$\|f_n - f_m\|_\infty < \frac{\epsilon/3}{\mu(X'_\epsilon)}$$

so

$$\int_{X'_\epsilon \setminus G} |f_n - f_m| \, d\mu < \mu(X'_\epsilon \setminus G) \cdot \frac{\epsilon/3}{\mu(X'_\epsilon)} \leq \epsilon/3.$$

Adding three parts together, we are done.

To summarize, see the paragraph "so the order is...".

□

We first try to use Dominated Convergence to establish a useful conclusion before heading for MCT. Namely, we reduce the need of finding explicit step functions when showing some $f \in \mathcal{L}^1$.

Prop 12. (A.E. Bounded Means Integrable) If f is μ-measurable and some $g \in \mathcal{L}^1$ satisfies $|g| \geq |f|$ a.e., then $f \in \mathcal{L}^1$.

Proof. Being μ-measurable means it is an a.e. limit of simple functions ϕ_n. Each simple function is step when we truncate by setting $\phi'_n(x) = \phi_n(x)$ if $\phi_n(x) \leq 2|g(x)|$ and 0 otherwise. The term 2 exists so $\phi'_n(x)$ ultimately becomes $\phi_n(x)$ (otherwise some $\phi'_n(x)$ potentially never escapes 0). So $\phi'_n \xrightarrow{\text{a.e.}} f$, and g satisfies the criteria for DCT. Use DCT and we are done. $\square$

We also need to collect tools that bridge together measure Cauchy vs mean Cauchy. In probability, this corresponds to $\mathbb{P}$ Cauchy vs $\mathbb{E}\,|\cdot|$ Cauchy.

Lemma 13. (Markov's Inequality) For nonnegative $f \subset \mathcal{L}^1$, we get

$$\mu(\{f \geq \epsilon\}) \leq \frac{1}{\epsilon} \int f \, d\mu.$$

Proof is one line (try freehand). Equivalently, $\mathbb{P}\,[X \geq \epsilon] \leq \epsilon^{-1}\,\mathbb{E}\,[X]$.

Prop 14. (Mean Cauchy Implies Measure Cauchy) If $(f_n) \subset \mathcal{L}^1$ is mean Cauchy, then it is measure Cauchy. Simply use Markov on $|f_n - f_m|$.

So, by strong Riesz-Weyl, for each mean Cauchy sequence we can find an a.u. Cauchy hence convergent subsequence. Since a.u. is the strongest, this subsequence also converges in measure and almost everywhere.

Lemma 15. (Riesz-Weyl, Extended) Any mean Cauchy (hence mean convergent and measure convergent) sequence $(f_n) \subset \mathcal{L}^1$ contain a subsequence (f_{n_k}) that converges a.u., a.e., and in measure to some a.e. unique $f \in \mathcal{L}^1$.

Mean Cauchy and measure Cauchy is bridged together. Riesz-Weyl became eventually useful. Now, we are ready for MCT.

Theorem 16. (Monotone Convergence) Suppose we have an a.e. increasing $(f_n) \subset \mathcal{L}^1$, where some $h \in \mathcal{L}^1$ satisfies $f_n \geq h$ a.e. for all n (common pointwise lower bound h). If the increasing sequence of integrals $(\int f_n \, d\mu)_{n=1}^{\infty}$ is bounded above, then (f_n) is mean and a.u. convergent to some $f \in \mathcal{L}^1$.

Remark. Most formulations use nonnegative functions, where 0 serves as an integrable lower (pointwise a.e.) bound. That is the only usage of nonnegativity, which renders not immediately helpful because in many applications, f_n is shifted down to include negativity. Therefore, we stick to DCT's convention and explicitly say that some h exists as a lower bound. The same story goes for $\liminf$ in Fatou.

Proof. The sequence of integrals must have a supremum $\alpha = \sup_n \int f_n \, d\mu$ as it's bounded above. Our task is to show that (f_n) is mean Cauchy. So,

$$\|f_n - f_m\|_1 = \left| \int f_n \, d\mu - \int f_m \, d\mu \right| \to 0$$

where the first equality is from $f_n \leq f_m$ WLOG, and the increasing integral sequence is bounded hence convergent and Cauchy. So, mean convergence.

From mean convergence to a.e. and a.u. convergence requires one more step. The extended Riesz-Weyl gives an a.e. and a.u. convergent subsequence (f_{n_k}), but since f is a.e. pointwise increasing, each $f_n(x)$ increases to the same limit so f is a.e. convergent. Moreover, since for each f_m there is f_{n_k} such that $n_k \leq m$, we have $\|f_m - f\|_\infty \leq \|f_{n_k} - f\|_\infty$ on any domain by (f_n) being increasing. So, f_n also converges almost uniformly to f.

□

Before we divert our attention to Fatou, notice that if we have h below and g above (f_n), then taking $r(x) = |h| \vee |g|$ gives DCT with r. We use "commonly $\mathcal{L}^1$-bounded" to say some $h \in \mathcal{L}^1$ bounds (a.e.) the sequence.

Lemma 17. (Fatou) For a sequence of commonly $\mathcal{L}^1$-lower bounded, measurable functions $(f_n) : \Omega \to \overline{\mathbb{R}}$, we have

$$\int \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu$$

Proof. Since

$$\int \liminf_n f_n \, d\mu = \int \lim_N \inf_{n\geq N} f_n \, d\mu = \lim_N \int \inf_{n\geq N} f_n \, d\mu$$

where the last inequality is by Monotone Convergence. Since for each N, m, $\int \inf_{n\geq N} f_n \, d\mu \leq \int f_m \, d\mu$ holds, then $\int \inf_{n\geq N} f_n \, d\mu \leq \inf_{m\geq N} \int f_m \, d\mu$. Take the limit (again, MCT) and we are done.

□

Remember that (μ-) measurable functions bounded above by an $\mathcal{L}^1$ function is integrable. That is why we only specified measurable functions.

Lemma 18. (Reverse Fatou) Any commonly $\mathcal{L}^1$-upper bounded measurable sequence (f_n) satisfies

$$\limsup_{n\to\infty} \int f_n \, d\mu \leq \int \limsup_{n\to\infty} \, d\mu.$$