

Real analysis.

Lecturer Aleksandr Rotkevich, spring 2024.

1 Convergence in Measure and Convergence Almost Everywhere

Now we will define two types of convergence, which play an important role in the theory of integration and probability. Both of them apply to functions defined on a measure space.

We assume that a measure space (X, \mathcal{A}, μ) is fixed. All sets we deal with are assumed measurable, i.e., they belong to the σ -algebra \mathcal{A} . All functions are also assumed measurable, and furthermore we assume that they are finite almost everywhere, i.e., may take infinite values only on sets of zero measure. The class of all such functions on a set E will be denoted by $L^0(E, \mu)$ or merely by $L^0(E)$. Everywhere in this section, we consider functions only from this class.

The pointwise convergence of a sequence $\{f_n\}_n \geq 1$ to a function f will be denoted, as usual, by a simple arrow, $f_n \xrightarrow{n \rightarrow \infty} f$, and the uniform convergence will be denoted by a double arrow: $f_n \rightrightarrows_{n \rightarrow \infty} f$. Recall that χ_E stands for the characteristic function of a set E , and the set $\{x \in E \mid f(x) > a\}$ is also denoted by $E(f > a)$.

We introduce an important new type of convergence of functional sequences.

Definition 1.1. A sequence of functions $f_n \in L^0(E, \mu)$ *converges* to a function $f \in L^0(E, \mu)$ *in measure* if for every $\varepsilon > 0$

$$\mu(E(|f_n - f| > \varepsilon)) \xrightarrow[n \rightarrow \infty]{} 0.$$

Notation:

$$f_n \xrightarrow[n \rightarrow \infty]{\mu} f.$$

Thus $f_n \xrightarrow[n \rightarrow \infty]{\mu} f$ if for sufficiently large n each of the functions f_n is uniformly close to f on the set obtained from E by removing a subset of arbitrarily small measure. It is worth mentioning that, in general, the subset to be removed differs for each n and one cannot generally remove a single set outside of which all functions f_n with sufficiently large indexes are uniformly close to the limit function.

Extending the definition, we say that a sequence $\{f_n\}_{n \geq 1}$ converges in measure on a set \tilde{E} , $\tilde{E} \subset E$, to a function $f \in L^0(\tilde{E})$ if the sequence $\tilde{f}_n = f_n|_{\tilde{E}}$ converges in measure to f . This is obviously equivalent to the condition that the sequence $\{f_n \chi_{\tilde{E}}\}$ converges in measure to the function f extended by zero from \tilde{E} to E . This observation allows us to assume, when discussing convergence in measure, that the functions under consideration are defined on the whole of X , since otherwise we can extend them to X by zero.

Let us discuss how convergence in measure is related to other types of convergence.

Remark 1.2. *Uniform convergence implies convergence in measure.*

Proof. If $f_n \Rightarrow f$ on E . Then for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for every $n > N$ and for every $x \in E$ we have $|f_n(x) - f(x)| < \varepsilon/2$. Hence,

$$E(|f_n(x) - f(x)| > \varepsilon) = \emptyset$$

and $\mu E(|f_n(x) - f(x)| > \varepsilon) = 0$ for $n > N$. □

Remark 1.3. *Pointwise convergence doesn't imply convergence in measure.*

Proof. To obtain a corresponding example, it suffices to consider the real line with the Lebesgue measure and the functions $\chi_{(n,n+1)}$, which converge to zero in \mathbb{R} pointwise, but not in measure since for every $n \in \mathbb{N}$ we have

$$\mu E(|f_n(x) - f(x)| > 1/2) = 1.$$

□

Of course, convergence in measure does not imply pointwise convergence. Indeed, if a sequence of functions f_n converges both pointwise and in measure (as is the case, for example, if the sequence converges uniformly), then we may break the pointwise convergence by modifying the values of f_n on sets of zero measure. However, this does not affect the convergence in measure, as follows from its definition. Hence it is natural to compare convergence in measure with "weakened pointwise convergence", which is insensitive to modifications of functions on sets of zero measure. We make the following definition.

Definition 1.4. *A sequence of measurable functions $f_n : E \rightarrow \overline{\mathbb{R}}$ **converges** to a function f **almost everywhere** on E if there exists a set $e \subset E$ of zero measure such that $f_n \xrightarrow[n \rightarrow \infty]{} f$ pointwise on $E \setminus e$.*

Notation:

$$f_n \xrightarrow[n \rightarrow \infty]{a.e.} f.$$

In this definition (as well as in the previous one), we assume that there is a fixed measure μ . If we also consider other measures, then we speak about convergence μ -almost everywhere (respectively, convergence in measure with respect to μ).

Remark 1.5. *Convergence in measure doesn't imply almost everywhere convergence.*

$$f_n \xrightarrow[n \rightarrow \infty]{\mu} f \not\Rightarrow f_n \xrightarrow[n \rightarrow \infty]{a.e.} f$$

Proof. Let $X = \mathbb{R}$ and $\mu = \lambda$ be the one-dimensional Lebesgue measure. For every positive integer k , consider the partition of the interval $[0, 1)$ into the subintervals $\Delta(k, p) = [\frac{p}{2^k}, \frac{p+1}{2^k})$, where $p = 0, 1, \dots, 2^k - 1$. To define a function f_n we write the index $n > 1$ in the form $n = 2^k + p$, where $0 \leq p < 2^k$ (such a representation is obviously unique, and k is just the integer part of $\log_2 n$, and set $f_n = \chi_{\Delta(k, p)}$. Since

$$X(f_n \neq 0) = \Delta(k, p) \quad \text{and} \quad \lambda(\Delta(k, p)) = \frac{1}{2^k} \leq \frac{2}{n} \xrightarrow[n \rightarrow \infty]{} 0,$$

the constructed sequence converges in measure to zero. However, the numerical sequence $\{f_n(x)\}_{n \geq 1}$ has no limit for any $x \in [0, 1)$, since among the values $f_n(x)$ there are infinitely many ones and zeros. \square

However, the situation changes dramatically if the set under consideration has finite measure.

Theorem 1.6 (Lebesgue). *On a set of finite measure, almost everywhere convergence implies convergence in measure.*

Proof. Let $f_n \xrightarrow[n \rightarrow \infty]{a.e.} f$ on E and $\mu(E) < +\infty$. Suppose that $f_n \xrightarrow[n \rightarrow \infty]{} f$ on $E \setminus e$, where $\mu(e) = 0$. Then redefining on e all functions by 0 we assume that $f_n \xrightarrow[n \rightarrow \infty]{} f$ everywhere on E .

Case 1. For a monotone sequence $\{f_n\}$ that converges pointwise to zero, the desired assertion is almost obvious. Indeed, in this case, for every $\varepsilon > 0$

the sets $E(|f_n| > \varepsilon)$ decrease as n grows and have an empty intersection. Then, by theorem on (upper) continuity of a measure,

$$\mu(E(|f_n| > \varepsilon)) \xrightarrow{n \rightarrow \infty} \mu\emptyset = 0.$$

Notice that here the condition $\mu(E) < +\infty$ is crucial. .

Case 2. Suppose that $f_n(x) \rightarrow f(x)$ for all $x \in E$ and let

$$\varphi_n(x) = \sup_{k \geq n} |f_k(x) - f(x)|.$$

Clearly, $\varphi_n(x) \rightarrow 0$ monotonically and by the first case

$$\mu E(\varphi_n > \varepsilon) \xrightarrow{n \rightarrow \infty} 0.$$

Finally, since $E(|f_k(x) - f(x)| > \varepsilon) \subset E(\varphi_n > \varepsilon)$ we see that

$$\mu E(|f_k(x) - f(x)| > \varepsilon) \leq \mu E(\varphi_n > \varepsilon) \xrightarrow{n \rightarrow \infty} 0.$$

□

Before continuing to discuss the relations between convergence in measure and almost everywhere convergence, we prove a simple but important result often used in probability theory.

Lemma 1.7 (Borel-Cantelli). *Let $\{E_n\}_n \geq 1$ be a sequence of measurable sets and*

$$E = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k = \{x \in X \mid x \in E_n \text{ for infinitely many } n\}.$$

If $\sum_{n=1}^{\infty} \mu(E_n) < +\infty$ then $\mu(E) = 0$.

Proof. Since $E \subset \bigcup_{n=k}^{\infty} E_n$, we have $\mu(E) \leq \sum_{n=k}^{\infty} \mu(E_n) \xrightarrow{k \rightarrow \infty} 0$. \square

This lemma implies a useful criterion for almost everywhere convergence.

Corollary 1.7.1. *Let $\varepsilon_n > 0$, $\varepsilon_n \xrightarrow{n \rightarrow \infty} 0$, $g_n \in L^0(X, \mu)$, and $X_n = X(|g_n| > \varepsilon_n)$. Suppose that $\sum_{n=1}^{\infty} \mu(X_n) < +\infty$. Then*

1. g_n a.e. 0;
2. Furthermore, for every $\varepsilon > 0$ there exists a set $e \subset X$ such that

$$\mu(e) < \varepsilon \quad \text{and} \quad g_n(x) \xrightarrow{n \rightarrow \infty} 0 \text{ on } X \setminus e.$$

Riesz theorem. To prove the almost everywhere convergence, we fix $\varepsilon > 0$ and apply the Borel-Cantelli lemma to the sets $E_n = X(|g_n| > \varepsilon)$, taking into account that $E_n \subset X_n$ for sufficiently large n .

To prove the second claim of the corollary, choose N so large that

$$\sum_{n=N}^{\infty} \mu(X_n) < \varepsilon$$

and put $e = \bigcup_{n=N}^{\infty} X_n$. Then $|g_n(x)| < \varepsilon_n$ for $x \in X \setminus e$ and $n > N$. \square

Theorem 1.8. *Every sequence that converges in measure contains a subsequence that converges almost everywhere to the same limit.*

Proof. Let $f_n \xrightarrow[n \rightarrow \infty]{\mu} f$. Then

$$\mu \left(X \left(|f_n - f| > \frac{1}{k} \right) \right) \xrightarrow[n \rightarrow \infty]{} 0$$

for every $k \in \mathbb{N}$. Hence there exists an increasing sequence of indexes n_k such that

$$\mu \left(X \left(|f_n - f| > \frac{1}{k} \right) \right) < \frac{1}{2^k} \quad \text{for all } n \geq n_k.$$

The sequence $\{f_{n_k}\}_k \geq 1$ has the desired property. Indeed, applying the corollary of the Borel-Cantelli lemma to the functions $g_k = |f_{n_k} - f|$, we see that $g_k \xrightarrow[k \rightarrow \infty]{\text{a.e.}} 0$, i.e., $f_{n_k} \xrightarrow[k \rightarrow \infty]{\text{a.e.}} f$. \square

Remark 1.9. *The subsequence constructed in the proof of Riesz's theorem, besides being almost everywhere convergent, has another useful (and stronger) property. Namely, for every $\varepsilon > 0$ there exists a set e such that*

$$\mu(e) < \varepsilon \quad \text{and} \quad f_{n_k} \xrightarrow[k \rightarrow \infty]{\Rightarrow} f \text{ on } X \setminus e.$$

To prove this, it suffices to apply the definition of the functions f_{n_k} and the corollary of the Borel-Cantelli lemma.

Corollary 1.9.1. *Corollary 1 If a sequence $\{f_n\}_n \geq 1$ converges in measure to functions f and g , then $f(x) = g(x)$ for almost all x .*

Proof. By Riesz's theorem, there exists a subsequence $\{f_{n_k}\}$ that converges to f almost everywhere. Since the subsequence $\{f_{n_k}\}$, along with the original sequence, converges in measure to g , again applying Riesz's

theorem, we can find a subsequence $\{f_{n_{k_j}}\}$ that converges almost everywhere to g . Thus the functions f and g coincide almost everywhere as limits of the almost everywhere convergent sequence $\{f_{n_{k_j}}\}$. \square

Corollary 1.9.2. *If $f_n \xrightarrow[n \rightarrow \infty]{\mu} f$ and $f_n \leq g$ almost everywhere for every n , then $f \leq g$ almost everywhere on E .*

Proof. Let f_{n_k} be a subsequence that converges to f almost everywhere. By our condition, $f_{n_k} \leq g$ outside of some set e_k of zero measure. Putting $e = \bigcup_{k=1}^{\infty} e_k$, we obtain a set of zero measure such that for any $x \notin e$ and $k \in \mathbb{N}$ the inequality $f_{n_k}(x) \leq g(x)$ holds. It remains to pass to the limit as $k \rightarrow \infty$. \square

Almost everywhere convergence is closely related to a stronger type of convergence which we now define.

Definition 1.10. *We say that a sequence $\{f_n\}_n \geq 1$ converges to f almost uniformly on X if for every positive ε there exists a set A_ε such that*

$$\mu(A_\varepsilon) < \varepsilon \quad \text{and} \quad f_n \xrightarrow[n \rightarrow \infty]{} f \text{ on } X \setminus A_\varepsilon.$$

Almost uniform convergence implies almost everywhere convergence. Indeed, the sequence $\{f_n\}$ converges pointwise outside of each set $A_{1/k}$, and hence outside of their intersection $\bigcap_{k=1}^{\infty} A_{1/k}$, which obviously has zero measure. As we observed after Riesz's theorem the sequence constructed in its proof converges not only almost everywhere, but almost uniformly.

Surprisingly, we have the following unexpected result: on a set of finite measure, almost uniform convergence is equivalent to almost everywhere convergence.

Theorem 1.11 (Egorov). Assume that $f_n, f \in L^0(X, \mu)$ and that $f_n \xrightarrow[n \rightarrow \infty]{a.e.} f$. If $\mu(X) < +\infty$ then $f_n \xrightarrow[n \rightarrow \infty]{} f$ almost uniformly on X .

Considering the sequence $\chi_{(n,n+1)}$ shows that this theorem cannot be extended to sets of infinite measure.

Proof. Let $g_n(x) = \sup_{k \geq n} |f_k(x) - f(x)|$. Clearly, $g_n \xrightarrow[n \rightarrow \infty]{a.e.} 0$. By Lebesgue's theorem $g_n \xrightarrow[n \rightarrow \infty]{\mu} 0$ (here the finiteness of μ is crucial). Hence there exists a subsequence $\{g_{n_k}\}$ such that

$$\mu \left(X \left(g_{n_k} > \frac{1}{k} \right) \right) < \frac{1}{2^k}.$$

By the corollary of the Borel-Cantelli lemma, this subsequence converges to zero almost uniformly. Since $|f_n - f| \leq g_{n_k}$ for $n \geq n_k$, the sequence $\{f_n - f\}$ also converges to zero almost uniformly. \square

In conclusion, we establish another useful property of almost everywhere convergence.

Theorem 1.12 (Diagonal sequence.). Let μ be a σ -finite measure, and let $f_k^{(n)} \in L^0(X, \mu)$, $g_n \in L^0(X, \mu)$ for $n, k \in \mathbb{N}$. Suppose that

$$f_k^{(n)} \xrightarrow[k \rightarrow \infty]{a.e.} g_n \text{ for every } n \in \mathbb{N}; \quad g_n \xrightarrow[n \rightarrow \infty]{a.e.} h.$$

Then there exists a strictly increasing sequence of indices k_n such that $f_{k_n}^{(n)} \xrightarrow[n \rightarrow \infty]{a.c.} h$.

Note that h , in contrast to $f_k^{(n)}$ and g_n , may take infinite values on sets of positive measure.

Proof. First assume that the measure is finite. Then $f_k^{(n)} \xrightarrow[k \rightarrow \infty]{\mu} g_n$ by Lebesgue's theorem. This means that

$$\mu \left(X \left(\left| f_k^{(n)} - g_n \right| > \varepsilon \right) \right) \xrightarrow[k \rightarrow \infty]{} 0 \text{ for every } n \in \mathbb{N} \text{ and every } \varepsilon > 0.$$

Hence for every n there exists an index k_n ($k_n > k_{n-1}$) such that

$$\mu \left(X \left(\left| f_{k_n}^{(n)} - g_n \right| > \frac{1}{n} \right) \right) < \frac{1}{2^n}.$$

By the corollary of the Borel-Cantelli lemma, $f_{k_n}^{(n)} - g_n \xrightarrow[n \rightarrow \infty]{\text{a.e.}} 0$. Thus

$$f_{k_n}^{(n)} = \left(f_{k_n}^{(n)} - g_n \right) + g_n \xrightarrow[n \rightarrow \infty]{\text{a.e.}} h,$$

which completes the proof of the theorem for a finite measure. \square

The case of an infinite measure can be reduced to the case of finite measure by the following lemma.

Lemma 1.13. *If μ is a σ -finite measure, then there exists a finite measure ν such that $\nu(E) = 0$ if and only if $\mu(E) = 0$.*

Thus "almost everywhere" assertions for the measures μ and ν hold simultaneously. Hence we may assume without loss of generality that the measure μ in the diagonal sequence theorem is finite.

Proof of the lemma. Let $X = \bigcup_{n=1}^{\infty} X_n$, where $0 < \mu(X_n) < +\infty$. We will obtain a measure with the desired property by putting

$$\nu(E) = \sum_{n \geq 1} \frac{1}{2^n} \frac{\mu(E \cap X_n)}{\mu(X_n)}$$

for every measurable set E .

Then ν is a measure, moreover, μ and ν vanish on the same sets. \square

2 Approximation of Measurable Functions by Continuous Functions. Luzin's Theorem.

In this section, we discuss properties of measurable functions on \mathbb{R}^m . The measurability (of sets and functions) means their measurability with respect to the Lebesgue measure, which we denote by λ .

Definition 2.1. *Let $A \subset \mathbb{R}^m$ and $x \in \mathbb{R}^m$. The value*

$$\text{dist}(x, A) = \inf\{\|x - y\| \mid y \in A\}$$

is called the distance from x to A .

Clearly, $\text{dist}(x, A) = 0$ only for points x lying in the closure of A . In particular, for a closed set A , the inequality $\text{dist}(x, A) > 0$ holds everywhere outside A .

Lemma 2.2. *The function $x \mapsto \text{dist}(x, A)$ is continuous on \mathbb{R}^m .*

Proof. Let $y \in A$ and $x, x_1 \in \mathbb{R}^m$. Then $\|x - y\| \leq \|x_1 - y\| + \|x_1 - x\|$, whence

$$\text{dist}(x, A) \leq \|x_1 - y\| + \|x_1 - x\|.$$

Taking the lower boundary in y of the right-hand side, we see that

$$\text{dist}(x, A) \leq \text{dist}(x_1, A) + \|x - x_1\|,$$

and $\text{dist}(x, A) - \text{dist}(x_1, A) \leq \|x - x_1\|$. Since x and x_1 are interchangeable, it follows that

$$|\text{dist}(x, A) - \text{dist}(x_1, A)| \leq \|x - x_1\|.$$

□

Lemma 2.3. *The characteristic function of a closed set $F \subset \mathbb{R}^m$ is the pointwise limit of a sequence of continuous functions.*

Proof. Obviously, the set-theoretic difference $\mathbb{R}^m \setminus F$ can be exhausted by the closed sets $H_n = \{x \in \mathbb{R}^m \mid \text{dist}(x, F) \geq 1/n\}$. Consider the following smoothings of the characteristic function of F :

$$f_n(x) = \frac{\text{dist}(x, H_n)}{\text{dist}(x, F) + \text{dist}(x, H_n)} \quad (x \in \mathbb{R}^m).$$

These functions are continuous everywhere, since the denominator does not vanish. The reader can easily check that $f_n(x) \xrightarrow[n \rightarrow \infty]{} \chi_F(x)$ for every $x \in \mathbb{R}^m$. □

Theorem 2.4 (Fréchet). *Every (Lebesgue) measurable function f on \mathbb{R}^m is the limit of a sequence of continuous functions converging almost everywhere.*

Proof. The proof will be split into several steps, with the function f becoming more and more complicated.

Step 1. Let f be the characteristic function of a measurable set E . By the regularity of the Lebesgue measure,

$$E = e \cup \bigcup_{n=1}^{\infty} K_n,$$

where $\mu(e) = 0$ and K_n is increasing sequence of compact sets. Then $\chi_{K_n} \rightarrow \chi_E$ almost everywhere. However, by Lemma 2.3, each of the characteristic functions χ_{K_n} is the limit of a sequence of continuous functions. Hence, by the diagonal sequence theorem, χ_E is also the limit of a sequence of continuous functions in the sense of almost everywhere convergence.

Step 2. If f is a simple function, i.e., it can be written in the form

$$f = \sum_{k=1}^N c_k \chi_{E_k},$$

where E_k are measurable sets, then, in order to approximate f by continuous functions, it suffices to approximate the functions χ_{E_k} .

Step 3. In the general case, consider a sequence of simple functions f_n that converges to f pointwise. It remains to approximate each function f_n by continuous functions and then apply the diagonal sequence theorem. \square

Theorem 2.5 (Luzin). *Every Lebesgue measurable function f on \mathbb{R}^m that is finite almost everywhere satisfies the Luzin property, i.e., for every $\delta > 0$ there exists a set $e \subset \mathbb{R}^m$ such that*

$$\mu(e) < \delta \text{ and the restriction of } f \text{ to } \mathbb{R}^m \setminus e \text{ is continuous.}$$

Proof. By the Fréchet theorem, there exists a sequence of continuous functions f_k that converges to f almost everywhere. According to Egorov's theorem, in every spherical layer

$$E_n = \{x \in \mathbb{R}^m \mid n-1 \leq \|x\| < n\}$$

there is a subset e_n such that

$$\mu(e_n) < \delta/2^n \text{ and } f_k \rightrightarrows f \text{ on } E_n \setminus e_n.$$

Clearly, the restriction of f to $E_n \setminus e_n$ is continuous as the uniform limit of continuous functions. Put $e = \bigcup_{n=1}^{\infty} (e_n \cup S_n)$, where S_n is the sphere of radius n centered at the origin. Then, obviously, $\mu(e) < \delta$ and the restriction of f to $\mathbb{R}^m \setminus e$ is continuous. \square

The result we have proved can be slightly strengthened by using the theorem on extension of continuous functions. The latter is formulated as follows.

Theorem 2.6. *Every function continuous on a closed subset F of \mathbb{R}^m is the restriction to F of a function continuous on \mathbb{R}^m .*

The proof of this theorem is omitted. It allows us to state Luzin's theorem in the following form.

Theorem 2.7. *Every Lebesgue measurable function f that is finite almost everywhere on \mathbb{R}^m coincides with a function that is continuous on \mathbb{R}^m except for a set of arbitrarily small measure. In other words, for every $\delta > 0$ there exists a function φ_δ continuous on \mathbb{R}^m such that*

$$\mu(\{x \in \mathbb{R}^m \mid f(x) \neq \varphi_\delta(x)\}) < \delta.$$

Proof. Fix $\delta > 0$ and consider the set e from the statement of Luzin's theorem. By the regularity of the Lebesgue measure, there exists an open set G containing e whose measure is arbitrarily close to the measure of e . Hence we may assume that $\mu(G) < \delta$. Let $F = \mathbb{R}^m \setminus G$, and let f_0 be the restriction of f to F . Now, to obtain φ_δ , it suffices to extend f_0 to a continuous function on \mathbb{R}^m . \square