Egorov's theorem

1 Theorem (Egorov) *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, let $E \subseteq \Omega$ with $E \in \mathcal{A}$ and $\mu(E) < \infty$, and let $f_n, f \colon E \to \mathbb{R}$ be measurable functions so that $f_n \to f$ μ-a.e. on E. Let $\varepsilon > 0$ Then there is some $A \subseteq E$ with $A \in \mathcal{A}$ and $\mu(A) < \varepsilon$ so that $f_n \to f$ uniformly on $E \setminus A$.*

Egorov's theorem is also known as one of *Littlewood's principles*:

Pointwise convergence is almost uniform.

– but note that this principle holds only on sets of finite measure.

Proof: For each $\eta > 0$, let

$$A_\eta = \bigcap_{N=1}^{\infty} \underbrace{\bigcup_{n=N}^{\infty} \{x \in E \colon |f_n(x) - f(x)| \geq \eta\}}_{A_{\eta,N}}.$$

Whenever $f_n(x) \to f(x)$, $x \in A_\eta$. Thus $\mu(A_\eta = 0)$. Since $E \supset A_{\eta,N} \searrow A_\eta$ and $\mu(E) < \infty$, $\mu(A_{\eta,N}) \to 0$ when $N \to \infty$.

For each $k \in \mathbb{N}$, pick $N_k \in \mathbb{N}$ with $\mu(A_{1/k,N_k}) < 2^{-k}\varepsilon$, and let $A = \bigcup_{k=1}^{\infty} A_{1/k,N_k}$.

Clearly, $\mu(A) < \varepsilon$.

I *claim* that $f_n \to f$ uniformly on $E \setminus A$. To see this, let $\eta > 0$ and pick $k > 1/\eta$. If $x \in E \setminus A$ and $n \geq N_k$, then $x \in E \setminus A_{1/k,N_k}$, so $|f_n(x) - f(x)| < 1/k < \eta$. ∎

Lusin's theorem – another one of Littlewood's principles

Littlewood stated is thus:

Every measurable function is almost continuous.

There are several ways to state this as a precise theorem, all (or mostly) bearing Lusin's name. We shall state and prove two of these.

But first, some lemmas.

A short series of lemmas

The following lemma is also one of Littlewood's principles:

A Lebesgue measurable set of finite measure is almost a finite union of intervals.

2 Lemma *Given a Lebesgue measurable set $E \subset \mathbb{R}$ with $\lambda(E) < \infty$ and $\varepsilon > 0$, there exists a finite union K of intervals with $\lambda(E \triangle K) < \varepsilon$.*

Proof: By the definition of outer Lebesgue measure, there is a sequence of intervals I_k with $E \subset \bigcup_{k=1}^{\infty} I_k$ and $\sum_{k=1}^{\infty} \lambda(I_k) < \lambda(E) + \varepsilon$. Since the series converges, we can find n with $\sum_{k=n+1}^{\infty} \lambda(I_k) < \varepsilon$. Let $K = \bigcup_{k=1}^{n} I_k$. Then $E \setminus K \subseteq \bigcup_{k=n+1}^{\infty} I_k$, so $\lambda(E \setminus K) < \varepsilon$. Also, $K \setminus E \subseteq \bigcup_{k=1}^{\infty} I_k \setminus E$, so $\lambda(K \setminus E) \leq \lambda(\bigcup_{k=1}^{\infty} I_k) - \lambda(E) < \varepsilon$ Thus $\lambda(E \triangle K) = \lambda(E \setminus K) + \lambda(K \setminus E) < 2\varepsilon$. ∎

Recall that a *simple function* is a measurable function which takes only a finite number of values, and which vanishes outside a set of finite measure. In particular, it is a finite linear combination of characteristic functions of measurable sets of finite measure.

By a *step function* we shall mean a finite linear combination of characteristic functions of bounded intervals.

3 Lemma *Given a simple function s and $\varepsilon > 0$, there is a step function S so that $\lambda(\{x \in \mathbb{R} \colon S(x) \neq s(x)\}) < \varepsilon$.*

Proof: This is really a corollary to the previous lemma. Let $s = \sum_{k=1}^{n} \chi_{A_k}$, let K_k be a finite union of intervals with $\lambda(A_k \triangle K_k) < \varepsilon/n$, and let $S = \sum_{k=1}^{n} \chi_{K_k}$. Then $S(x) = s(x)$ except for $x \in \bigcup_{k=1}^{n} (A_k \triangle K_k)$, and this set has measure $< \varepsilon$. ∎

4 Lemma *If $f \colon \mathbb{R} \to \mathbb{C}$ is Lebesgue measurable, there is a sequence of step functions converging to f almost everywhere.*

Proof: Take a sequence (s_n) of simple functions so that $s_n \to f$ pointwise. For each n, let S_n be a step function so that $S_n = s_n$ except on a set D_n, where $\lambda(D_n) < 2^{-n}$. Let

$$D = \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} D_n,$$

and note that $\lambda(D) = 0$, because $\lambda(\bigcup_{n=N}^{\infty} D_n) \leq \sum_{n=N}^{\infty} \lambda(D_n) < \sum_{n=N}^{\infty} 2^{-n} = 2^{1-N}$.

If $x \in \mathbb{R} \setminus D$ then for some N, $S_n(x) = s_n(x)$ for all $n \geq N$, so that $S_n(x) \to f(x)$ because $s_n(x) \to f(x)$. ∎

Lusin's theorem, take one

5 Theorem (Lusin) *If $f\colon \mathbb{R} \to \mathbb{C}$ is Lebesgue measurable and $\varepsilon > 0$, there is some Lebesgue measurable set $E \subset \mathbb{R}$ with $\lambda(E) < \varepsilon$ so that $f\big|_{\mathbb{R}\setminus E}$ is continuous.*

Note that this theorem does *not* claim that f is continuous at every $x \in \mathbb{R} \setminus E$. It is the *restriction* of f that is continuous. To illustrate the difference, consider $f = \chi_{\mathbb{Q}}$, which is nowhere continuous. However, its restriction to $\mathbb{R}\setminus\mathbb{Q}$ is continuous: This restriction is constantly zero.

Proof: Take a sequence (S_n) of step functions converging a.e. to f. For each integer N, Egorov's theorem implies the existence of a measurable set $A_N \subseteq (N, N+1)$ with $\lambda(A_N) < 2^{-|N|}\varepsilon$ so that $S_n \to f$ uniformly on $(N, N+1) \setminus A_N$. Let $A = \bigcup_{N\in\mathbb{Z}} A_n$. Then $\lambda(A) < 3\varepsilon$. Further, let D be the set of points where some S_n is discontinuous. Since each S_n is discontinuous at only a finite number of points, D is countable. Let $E = \mathbb{Z} \cup D \cup A$. Then $\lambda(E) = \lambda(A) < 3\varepsilon$. Each S_n is continuous on E, and the convergence to f is uniform on $(N, N+1) \setminus E$, so the restriction of f to $(N, N+1) \setminus E$ is continuous. Therefore, so is the restriction of f to $\mathbb{R} \setminus E$.

That last bit is subtle, so let us do it in more detail: If $x \in \mathbb{R} \setminus E$, then $x \notin \mathbb{Z}$, so $x \in (N, N+1)$ for some $N \in \mathbb{Z}$. Given $\varepsilon > 0$ there is some $\delta > 0$ so that, whenever $y \in (N, N+1) \setminus E$ and $|y - z| < \delta$, then $|f(y) - f(x)| < \varepsilon$. Here we can replace δ by a smaller number, so that the δ-neighbourhood of x is contained in $(N, N+1)$. Then, if $y \in \mathbb{R} \setminus E$ and $|y - x| < \delta$, we also have $y \in (N, N+1) \setminus E$, so that $|f(y) - f(x)| < \varepsilon$. ∎

Lusin's theorem, take two

Theorem 5 is good as far as it goes, but we can improve on it:

6 Theorem (Lusin) *If $f\colon \mathbb{R} \to \mathbb{C}$ is Lebesgue measurable and $\varepsilon > 0$, there is some Lebesgue measurable set $G \subset \mathbb{R}$ with $\lambda(G) < \varepsilon$ and a continuous function $g\colon \mathbb{R} \to \mathbb{C}$ so that $f(x) = g(x)$ for all $x \in \mathbb{R} \setminus G$.*

Proof: Begin by finding a set E as in Theorem 5, so that $\lambda(E) < \varepsilon$, and $f\big|_{\mathbb{R}\setminus E}$ is continuous.

Pick an open set $G \supseteq E$ with $\lambda(G) < 2\varepsilon$. Then also $f\big|_{\mathbb{R}\setminus G}$ is continuous. We shall expand this function to a function $g\colon \mathbb{R} \to \mathbb{C}$. So we begin by defining $g(x) = f(x)$ when $x \in \mathbb{R} \setminus G$.

The set G, being open, is a pairwise disjoint union of open intervals – the components of G. Say,

$$G = \bigsqcup_{k=1}^{\infty} (a_k, b_k).$$

Expand g by making it linear on $[a_k, b_k]$:

$$g(x) = \frac{(b_k - x) f(a_k) + (x - a_k) f(b_k)}{b_k - a_k} \qquad \text{for } x \in (a_k, b_k).$$

Then g is indeed continuous on $\mathbb{R}$ (exercise!), and $g = f$ on $\mathbb{R} \setminus G$. ∎

Approximation in $\mathcal{L}^1$

We don't *need* the full power of Lusin's theorem to prove the following approximation result. But it does seem like a nice application.

7 Theorem *Let $f \in \mathcal{L}^1(\lambda)$ and $\varepsilon > 0$. Then there is a continuous function $g \in \mathcal{L}^1(\lambda)$ with*

$$\int_{\mathbb{R}} |f - g| \, d\lambda < \varepsilon.$$

Proof: We prove this only for $f \colon \mathbb{R} \to \mathbb{R}$. If f is complex valued, we can apply the real result to the real and imaginary parts of f separately.

Let $B_n = \{x \in \mathbb{R} \colon |f(x)| \geq n\}$. Since $\chi_{B_n} f \to 0$ as $n \to \infty$, and $|\chi_{B_n} f| \leq |f|$, we can use the dominated convergence theorem (DCT) to conclude that $\int_{B_n} |f| \, d\lambda \to 0$ when $n \to \infty$. Pick some n so that $\int_{B_n} |f| \, d\lambda < \varepsilon$.

Now define f_n by

$$f_n(x) = \begin{cases} f(x) & \text{if } |f(x)| \leq n, \\ n & \text{if } f(x) > n, \\ -n & \text{if } f(x) < -n. \end{cases}$$

Pick a continuous function $g \colon \mathbb{R} \to \mathbb{R}$ so that $f(x) = g(x)$ for all x except for x in a set A with $\lambda(A) < \varepsilon / n$. We may assume that $|g| \leq n$ everywhere, for otherwise, redefine it by setting $g(x) = n$ where before $g(x) > n$ and $g(x) = -n$ where before $g(x) < -n$. Now

$$\begin{aligned} \int_{\mathbb{R}} |f - g| \, d\lambda &\leq \int_{\mathbb{R}} |f - f_n| \, d\lambda + \int_{\mathbb{R}} |f_n - g| \, d\lambda \\ &= \int_{B_n} |f - f_n| \, d\lambda + \int_A |f_n - g| \, d\lambda \\ &\leq \int_{B_n} |f| \, d\lambda + 2n\lambda(A) \\ &< \varepsilon + 2\varepsilon = 3\varepsilon. \end{aligned}$$

Finally, $g \in \mathcal{L}^1(\lambda)$ because the triangle inequality yields

$$\int_{\mathbb{R}} |g| \, d\lambda \leq \int_{\mathbb{R}} |g - f| \, d\lambda + \int_{\mathbb{R}} |f| \, d\lambda < \infty.$$

The proof is now complete. ∎