(*September 15, 2016*)

Preview: Part 00, Real Analysis 2016-17

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/˜garrett/

[This document is http://www.math.umn.edu/˜garrett/m/real/notes_2016-17/real-notes-00.pdf]

(Everything below admits substantial generalization beyond what is literally asserted. Determining the extent of various possible generalizations is often a task in itself, and is often tangential to the main enterprise.)

---

1. *What is the issue? What obstacles to overcome?*

For most purposes, up until 1800 and even afterward, *function* meant *formula*. Also, it was often assumed without comment that decent functions could be represented by *power series*.

[1.1] Euler and the wave equation Many people had considered the (linear) *wave equation*

$$(\Delta_x - \frac{\partial^2}{\partial t^2})\, u \;=\; 0$$

where the *spatial* variable $x \in \mathbb{R}^n$ (mostly $n = 1, 2, 3$) and *time* $t \in \mathbb{R}$, and *Laplace's operator* is

$$\Delta_x \;=\; \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_n^2}$$

For one-dimensional spatial variable, the wave operator *factors*:

$$\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial t^2} \;=\; \Big(\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\Big) \circ \Big(\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\Big) \;=\; \Big(\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\Big) \circ \Big(\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\Big)$$

Thus, apparently, any function u of the form

$$u(x,t) \;=\; f(x-t) + g(x+t)$$

is a solution. The two pieces are *incoming* and *outgoing* components of the solution.

The cognitive dissonance arises when one imagines, as apparently Euler did, that such a formula *makes sense even when f and g are not differentiable (in a classical sense).*

But there did not seem to be any natural or conceptual way to exclude problemmatical functions f, g from this formula, and this heated up the discussion of *what is a function?*

[1.2] A success story: convergent power series By soon after 1800, Abel and others had carefully proven that power series (real or complex) with a positive radius of convergence r really could be differentiated correctly by doing the obvious thing, namely, differentiating term by term:

$$\frac{d}{dz} \sum_{n=0}^{\infty} c_n\, (z-z_o)^n \;=\; \sum_{n=0}^{\infty} \frac{d}{dz} c_n\, (z-z_o)^n \;=\; \sum_{n=0}^{\infty} c_n\, n(z-z_o)^{n-1} \qquad \text{(still convergent in } |z-z_o| < r)$$

This completely justified what people had been doing all along.

[1.3] Fourier 1811-22 and the heat equation

The heat equation is

$$(\Delta_x - \frac{\partial}{\partial t})\, u \;=\; 0 \qquad\qquad \text{(with } \textit{initial condition} \text{ prescribing } u(x,0))$$

Especially in the case of one-dimensional spatial variable x confined to a finite interval such as $[0,2\pi]$, Fourier had the inspiration to express an alleged solution as a superposition of eigenfunctions for Δ_x on $[0,2\pi]$, namely, constants and $\sin(nx)$ and $\cos(nx)$ for $n = 1,2,3,\ldots$:

$$u(x,t) \;=\; c_o(t) + \sum_{n\geq 1} \Big(a_n(t)\cos(nx) + b_n(t)\sin(nx)\Big)$$

This *separated variables*, and if we imagine we can apply the heat operator *termwise*,

$$0 \;=\; (\Delta_x - \frac{\partial}{\partial t})\, u \;=\; -c_o'(t) + \sum_{n\geq 1} \Big(-n^2 a_n - a_n')\cos(nx) + (-n^2 b_n - b_n')\sin(nx)\Big)$$

If we believe in *uniqueness* of such expressions in x, this gives

$$-c_o' = 0 \qquad\quad -n^2 a_n - a_n' = 0 \qquad\quad -n^2 b_n - b_n' = 0 \qquad\quad \text{(for } n = 1,2,3,\ldots)$$

so $c_o(t)$ is a constant, and $a_n(t)$ and $b_n(t)$ are constant multiples of $e^{-n^2 t}$:

$$u(x,t) \;=\; -c_o + \sum_{n\geq 1} e^{-n^2 t}\Big(a_n\cos(nx) + b_n\sin(nx)\Big) \qquad\qquad \text{(with constants } c_o, a_n, b_n)$$

The initial condition at time t presumably determines the constants, by

$$u(x,0) \;=\; -c_o + \sum_{n\geq 1} \Big(a_n\cos(nx) + b_n\sin(nx)\Big)$$

The explicit claim that *every* function $x \to u(x,0)$ could be represented by such a *Fourier series* was appealing, since this device then gave a solution to the heat equation, and would prove uniqueness. *But what is a function?*

Soon after his initial epiphany, Fourier also found the correct formulas determining coefficients:

$$c_o \;=\; \frac{1}{2\pi}\int_0^{2\pi} u(x,t)\, dx \qquad\quad a_n \;=\; \frac{1}{2\pi}\int_0^{2\pi} u(x,t)\cdot\cos(nx)\, dx \qquad\quad b_n \;=\; \frac{1}{2\pi}\int_0^{2\pi} u(x,t)\cdot\sin(nx)\, dx$$

Further, under relatively mild hypotheses [1] on the smoothness-or-not of $x \to u(x,0)$, Fourier proved that the series converges *pointwise* to $u(x,0)$ and $u(x,t)$ for $t > 0$. [2]

However, Fourier made much broader claim about the range of *functions* representable by such series, revivifying the argument over *what is a function?*

More technically, there is the issue of the legitimacy of *termwise differentiation*. Indeed, functions meeting the conditions for pointwise convergence could have derivatives *not* meeting the condition, yet termwise differentiation would still make sense. For example, the periodic *sawtooth function* is

$$(\text{sawtooth}) \;=\; \sum_{n\geq 1} \frac{\sin(nx)}{n}$$

---

[1] For example, if a function is piecewise C^1 except for finitely-many jumps in $[0,2\pi]$, where left and right derivatives exist, then, away from the jumps, the Fourier series converges pointwise to the function.

[2] Apparently what is often called the *Dirichlet kernel* and used to prove this pointwise convergence was in fact used by Fourier prior to Dirichlet's 1829 paper proving convergence of Fourier series.

This converges (not absolutely) to the sawtooth function's values $x - \pi$ for $0 < x < 2\pi$. The sawtooth is differentiable in $(0, 2\pi)$, but termwise differentiation gives

$$\frac{d}{dx}(\text{sawtooth}) \;=\; \sum_{n \geq 1} \sin(nx) \qquad\qquad (???)$$

For most values of $x \in (0, 2\pi)$, the summands do not go to 0. Differentiating again should give 0 for $0 < x < 2\pi$, but

$$\frac{d^2}{dx^2}(\text{sawtooth}) \;=\; \sum_{n \geq 1} n \cdot \sin(nx) \qquad\qquad (???)$$

and so on. These expressions do not converge *pointwise*, and cast reasonable doubt on the legitimacy of this approach.

However, in fact, although these infinite sums of functions do not converge pointwise, they do converge perfectly well in certain topological vector spaces of (generalized) functions, namely, the Sobolev spaces discussed below. But this development would have to wait until the 1930s and 1940s.

Another tension arose when people subsequently discovered that the Fourier series of typical continuous functions would fail to converge pointwise at infinitely many points. (For example, we will prove this via *Baire's Theorem*.)

Yet there is *Parseval's theorem*, that for f such that $\int_0^{2\pi} |f|^2 < \infty$, there is a nice relation between the this integral of f and its Fourier coefficients:

$$\int_0^{2\pi} |f|^2 \;=\; |c_o|^2 + \sum_{n \geq 1} |a_n|^2 + |b_n|^2$$

This implies that, even if the partial sums of the Fourier series of such a function do not converge to the function *pointwise*, they *do* converge to the function in the *mean-square* or L^2 metric

$$d_{L^2}(f, g) \;=\; |f - g|_{L^2[0,2\pi]} \;=\; \Big(\int_0^{2\pi} |f(x) - g(x)|^2 \, dx \Big)^{1/2}$$

That this is a metric on $C^o[0, 2\pi]$ uses the integral form of the *Cauchy-Schwarz-Bunyakowsky inequality*, due to Bunyakowsky. But convergence of a sequence of continuous functions in this L^2-metric does *not* imply pointwise convergence, since the pointwise evaluation maps maps $f \to f(x_o)$ are *not continuous*: there are sequences $\{f_n\}$ of continuous functions that are Cauchy sequences in the L^2 topology, but so that $\{f_n(x_o)\}$ is *not* a Cauchy sequence of real or complex numbers.

Also, simple pointwise convergence does not imply L^2 convergence in general, and simple pointwise convergence does not imply convergence in the sup-norm topology on $C^o[0, 2\pi]$, either.

The seemingly natural notion of pointwise convergence is not all that we had hoped it would be. As a corollary, there are problems if we exclusive think of *functions* as producing pointwise values: there are L^2 limits of Cauchy sequences of continuous functions that lack well-defined pointwise limits.

[1.4] Sturm and Liouville 1830s eigenfunction expansions On the heels of Fourier's ideas, Sturm and Liouville had a similar idea about expressing functions f on $[0, 2\pi]$ in terms of *eigenfunctions* for differential operators of the form

$$Lu \;=\; -(pu')' + q \qquad\qquad (\text{with } p(x) > 0 \text{ on } [0, 2\pi], \text{ real-valued } q)$$

with various possible *boundary conditions* at 0 and 2π. For example, we might require $u(0) = u(2\pi)$ and $u'(0) = u'(2\pi)$ (the periodic case), or we might require $u(0) = 0 = u(2\pi)$ (the Dirichlet condition).

That is, they argued first toward the conclusion that the *eigenfunction equation*

$$Lu \;=\; \lambda \cdot u \qquad\qquad \text{(with the boundary conditions)}$$

should have a list $0 \le \lambda_1 \le \lambda_2 \le \lambda_3, \ldots$ of non-negative real numbers such that there would be non-trivial (real-valued) solutions u_n to the equation $Lu = \lambda_n \cdot u$ and meeting the boundary conditions. Then, when normalized so that $\int_0^2 \pi |u_n(x)|^2\, dx = 1$, an *arbitrary* (real-valued) function f on $[0, 2\pi]$ should be expressible as

$$f(x) \;=\; \sum_{n\ge 1} \Big(\int_0^{2\pi} f(t)\cdot u_n(t)\; dx \Big) \cdot u_n(x)$$

Their difficulty at the time was that various notions of convergence were still unsettled, and the *linear algebra* needed to express things this clearly had not yet been invented. Heuristics were not made into proofs (of some assertions) until Steklov 1898-9, and Bocher 1895-6.

By now we know that for $\int_0^{2\pi} |f(x)|^2\; dx < +\infty$, that expansion *does* converge in the L^2-metric, and we think of the coefficients as being given by *inner products* of f with the exponentials in the space $L^2[0, 2\pi]$:

$$\langle f, g\rangle_{L^2[0,2\pi]} \;=\; \frac{1}{2\pi} \int_0^{2\pi} f(x)\cdot \overline{g(x)}\; dx$$

(with complex conjugation for complex-valued functions). One characterization of the whole space $L^2[0, 2\pi]$ is as the *completion* of $C^o[0, 2\pi]$ with respect to the metric obtained from the L^2-norm.

But pointwise convergence is potentially confusing: with the Dirichlet condition, the eigenfunctions are $u_n(x) = \sin(nx/2)/\sqrt{2\pi}$. But there are many reasonable functions meeting the condition $\int_0^{2\pi} |f(x)|^2\; dx < +\infty$ that do not vanish at 0 and 2π, for example, the constant function 1. So, in an L^2 (mean-square) sense,

$$1 \;=\; \frac{1}{2\pi} \sum_{n\ge 1} \Big(\int_0^{2\pi} 1\cdot \sin(nt/2)\; dx \Big) \cdot \sin(nx/2) \;=\; \frac{1}{2\pi} \sum_{n=1,3,5,\ldots} \frac{\sin(nx/2)}{2n}$$

but this certainly cannot converge pointwise as the endpoints. It *does* provably converge pointwise in the interior.

[1.5] Green's functions 1828 Another approach to solving linear differential equations $Lu = f$ on $\mathbb{R}^n$, not only in one dimension like the Sturm-Liouville equations, was conceived by Green about 1828, and has similar applications to partial differential equations like the heat equation and wave equation.

One way to talk about the method is to refer to a *fundamental solution* or *Green's function* $G(x, y)$ for the given differential operator L, characterized by solving the differential equation $Lu = f$ by

$$u(x) \;=\; \int_{\mathbb{R}^n} G(x,y)\, f(y)\; dy$$

Green's original idea and subsequent applications arose in physically meaningful situations, problems, so the sensibility of solutions to problems obtained by such ideas could be confirmed to some degree by direct observation of physical phenomena.

But, from a mathematical viewpoint, *why* should any such thing exist?

If we already believe from Sturm-Liouville that there is an *orthonormal basis* $\{u_n\}$ for $L^2[0, 2\pi]$ consisting of eigenfunctions u_n for L, in an equation $Lu = f$ expand both u and f in terms of eigenfunctions, computing coefficients by inner products, as in Fourier's case:

$$L\Big(\sum_n \langle u, u_n\rangle \cdot u_n\Big) \;=\; Lu \;=\; f \;=\; \sum_n \langle f, u_n\rangle \cdot u_n$$

Of course, we assume that we can apply L termwise (!), so this gives

$$\sum_n \langle f, u_n\rangle \cdot u_n \;=\; \sum_n \langle u, u_n\rangle \cdot Lu_n \;=\; \sum_n \langle u, u_n\rangle \cdot \lambda_n \cdot u_n$$

Presumably these expansions are unique, so $\langle u, u_n\rangle \cdot \lambda_n = \langle f, u_n\rangle \cdot u_n$ for all n. That is, apparently

$$u(x) \;=\; \sum_n \frac{\langle f, u_n\rangle}{\lambda_n} \cdot u_n(x) \;=\; \Big\langle f(y),\, \sum_n \frac{u_n(y)}{\lambda_n} \cdot u_n(x) \Big\rangle$$

That is, apparently,

$$G(x,y) \;=\; \sum_n \frac{1}{\lambda_n} u_n(x) \cdot u_n(y)$$

For that matter, a more *scandalous* description of $G(x,y)$, but which makes considerable sense in a physical context where Dirac's δ idealizes a *point-mass*, is

$$L_x\, G(x,y) \;=\; \delta(x-y) \qquad\qquad \text{(with a Dirac } \delta\text{-function)}$$

which would have been essentially impossible to make mathematically rigorous until well into the 20th century. Nevertheless, if we apply L_x to the eigenfunction expansion, apparently

$$\delta(x-y) \;=\; L_x G(x,y) \;=\; L_x \sum_n \frac{1}{\lambda_n} u_n(x) \cdot u_n(y) \;=\; \sum_n \frac{1}{\lambda_n} L_x u_n(x) \cdot u_n(y) \;=\; \sum_n u_n(x) \cdot u_n(y)$$

If true, this would be very convenient. But *pointwise* it cannot make sense.

Still, in one dimension, reasonable second-order differential operators L on finite intervals have Green's functions obtained in a straightforward way from two linearly independent solutions, based on the idea that $\frac{d^2}{dx^2}|x| = 2\delta$, as follows. Find one solution u to $Lu = 0$ with $u(0) = 0$, and a solution v to $Lv = 0$ with $v(2\pi) = 0$, and splice them together so that their *values* match at a point $x \in [0, 2\pi]$, but their *derivatives* differ suitably, creating a *corner*.

For example, for the equation $-u'' = f$ on $[0, 2\pi]$ with boundary conditions $u(0) = 0 = u(2\pi)$, solutions of $u'' = 0$ are just linear functions, the solution vanishing at the left edge is x, and the solution vanishing at the right edge is $2\pi - x$. To find the linear combination agreeing at y and derivatives differing by 1, solve for coefficients a, b in

$$\begin{cases} a \cdot y &= \; b \cdot (2\pi - y) \\[2ex] a + 1 &= \; -b \end{cases}$$

and obtain

$$G(x,y) \;=\; \begin{cases} \Big(\frac{y}{2\pi} - 1\Big) \cdot x & (\text{for } 0 \le x \le y) \\[2ex] -\frac{y}{2\pi} \cdot (2\pi - x) & (\text{for } y \le x \le 2\pi) \end{cases}$$

In two dimensions or higher, the geometry is more complicated. Nevertheless, it has been appreciated for a long time, in one way or another, that

$$\begin{cases} \Delta \log |x| &= \; (\text{constant}) \cdot \delta \quad (\text{in } \mathbb{R}^2) \\[2ex] \Delta \frac{1}{|x|^{n-2}} &= \; (\text{constant}) \cdot \delta \quad (\text{in } \mathbb{R}^n,\ n \ge 3) \end{cases}$$

with elementary constants. An elementary computation certainly shows that the Laplacian annihilates those functions *away from* 0, but we are lacking a *persuasive* or *conceptual* argument that *at* 0 we get δ.

Looking at that one-dimensional situation further, apparently

$$\begin{cases} \left(\frac{y}{2\pi} - 1\right) \cdot x & \text{(for } 0 \le x \le y\text{)} \\ -\frac{y}{2\pi} \cdot (2\pi - x) & \text{(for } y \le x \le 2\pi\text{)} \end{cases} \;=\; G(x,y) \;=\; \frac{1}{2\pi} \sum_{n\ge 1} \sin\frac{nx}{2} \cdot \frac{\sin\frac{ny}{2}}{-n^2/4}$$

and applying Δ gives

$$\delta(x-y) \;=\; \frac{1}{2\pi} \sum_{n\ge 1} \sin\frac{nx}{2} \cdot \frac{\sin ny}{2} \qquad (???)$$

But the other eigenfunction expansion similarly apparently gives

$$\delta(x-y) \;=\; \frac{1}{2\pi} \cdot \Big(1 + \sum_{n\ge 1} \sin nx + \cos nx\Big) \qquad (???)$$

and the two expressions are not easily comparable. The heuristic is attractive and useful, but a more refined viewpoint is obviously needed to avoid seeming paradoxes.

[1.6] Heaviside 1880s Also used δ as an idealization of an *impulse* in electrical circuits and similar, with great success. Despite his successes in predicting observable phenomena, mathematicians at the time were apparently disdainful of the mathematics itself, which was unrigorizable at the time.

[1.7] Dirac 1928-9 In nascent quantum physics, Dirac not only used point-masses and point-charges, but geometrically more complicated generalized functions, and did subtle computations that correctly predicted physical phenomena. In contrast to Hilbert's and Schmidt's conversion of *differential* operators to *integral* operators with better continuity properties, Dirac directly manipulated differential operators without apparent concern for their not being everywhere defined or continuous.

Partly in reaction to Dirac's physics success, careful rigorization of unbounded/discontinuous operators, modelling differential operators, was accomplished by Stone and von Neumann by 1930, and more simply in important special situations by Friedrichs in 1934. In 1934 and thereafter, Sobolev created a basic framework adequate to deal with certain generalized functions.

[1.8] Kronig-Penney 1931, Bethe-Peierls 1935 ... but Dirac's success prompted even-more-audacious mathematics: idealizing δ as a very-short-range-acting *potential*, to model nuclear foces (as opposed to electromagnetism or gravity), physicists considered *singular potential* equations

$$(-\Delta + \delta)\, u \;=\; f$$

The intention is fairly clear, but it is not obvious how to be sure one is manipulating such a thing correctly from a mathematical viewpoint. Still, testable physical conclusions were correctly reached, and Nobel prizes were won.

[1.9] Fourier transforms, Plancherel 1910, Wiener 1933, Bochner 1932 In 1910, Plancherel proved the basic fact that Fourier transform on reasonable functions f with $\int_{\mathbb{R}} |f| < \infty$ gave an $L^2(\mathbb{R})$-isometry. That is, with

$$\widehat{f}(\xi) \;=\; \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\xi x}\, f(x)\, dx$$

the L^2 norm of $\widehat{f}$ is equal to that of f. This allows the Fourier transform to be *extended by continuity* to give a map of $L^2(\mathbb{R})$ to itself, although the literal integral does not converge well for general functions in L^2 but not in L^1. Part of the lesson is that maps given by integrals *cannot* be taken literally, but, happily, *need not* be taken literally.

Fourier inversion is that f can be reconstructed from its Fourier transform:

$$f(x) \;=\; \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi x}\; \widehat{f}(\xi)\; d\xi$$

There is a non-trivial issue of the sense of convergence of the integral! A naive but reasonable attempt to prove Fourier inversion is the obvious interchange of the order of integration:

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi x}\; \widehat{f}(\xi)\; d\xi \;=\; f(x) \;=\; \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi x}\Big(\int_{\mathbb{R}} e^{-i\xi u} f(v)\; dv\Big)\; d\xi \;=\; \frac{1}{2\pi} \int_{\mathbb{R}} f(v)\Big(\int_{\mathbb{R}} e^{i\xi(x-v)}\; d\xi\Big)\; dv$$

If we could believe various heuristics that the inner integral is $2\pi\delta(x-v)$, we'd be done. Indeed, this can be justified later, but Fourier inversion is prior.

These examples and others raise basic questions:

What kind of functions can be integrated?

What kind of infinite-sum expansions of functions are legitimate?

What kind of convergence do infinite-sum expansions have?

---

2. *Limits and integrals*

Archetypical issue: integrating on a finite interval $[a, b]$ on the real line,

$$\text{when is } \lim_n \int_a^b f_n \;=\; \int_a^b \lim_n f_n \;\; ???$$

And *limit* in what sense? And what kind of *functions* can be integrated?

As a positive example, if the functions f_n are *continuous*, and if the limit is *uniformly pointwise*, meaning that for every $\varepsilon > 0$ there is n_o such that for every $m, n \geq n_o$ and for every $x \in [a, b]$, the limit $\lim_n f_n$ is itself a continuous function, and, indeed, the integral of the limit is the limit of the integrals. For continuous functions on finite intervals, the Riemann integral behaves well with *uniformly* pointwise limits, and gives us a description of *integral* that allows us to prove the previous assertion.

However, even when the functions f_n are very nice, if the limit is merely *pointwise*, but not *uniformly* so, then the limit function need not be continuous, and the limit of the integrals need not be the integral of the limit.

Also, a pointwise limit of continuous functions need not be continuous! But we can salvage a little, even though the issue will not go away:

[2.1] Theorem: *(Dini)* For a pointwise *monotone* (increasing or decreasing) sequence of real-valued continuous functions f_n on a finite interval $[a, b]$, if the limit is continuous, then the limit is *uniform* pointwise.

The classic example of failure of the integral of the (pointwise) limit to be the limit of the integrals is the sequence of *tent functions* f_n just to the right of 0: $f_n(x) = 0$ on $[\frac{2}{n}, 1]$, and on $[0, \frac{2}{n}]$ is a triangular tent of height n, to make the area under it be 1:

$$f_n(x) \;=\; \begin{cases} 0 & (\text{for } x \leq 0) \\ n^2 \cdot x & (\text{for } 0 \leq x \leq \frac{1}{n}) \\ n - n^2 \cdot (x - \frac{1}{n}) & (\text{for } \frac{1}{n} \leq x \leq \frac{2}{n}) \\ 0 & (\text{for } x \geq \frac{2}{n}) \end{cases}$$

For every individual $x \in \mathbb{R}$, the pointwise limit is $\lim_n f_n(x) = 0$, but the integral of the zero function is not 1.

On the other hand, for $g \in C^o(\mathbb{R})$, while the pointwise limit of these tent functions f_n is 0 everywhere,

$$\lim_n \int_{\mathbb{R}} f_n(x)\, g(x)\, dx \;=\; g(0) \;=\; \delta(g)$$

That is, in a very tangible sense, $f_n \longrightarrow \delta$, where δ is the Dirac delta function at 0, which we imagine produces $g(0)$ when integrated against a continuous function g.

Measure theory can accommodate the Dirac delta, because it is a kind of *measure*. But its *derivative* [3] is not a measure. Nevertheless, using tent-functions, we can make a sequence of continuous functions h_n that go to 0 everywhere pointwise, but so that

$$\lim_n \int_{\mathbb{R}} h_n(x)\, g(x)\, dx \;=\; g'(0)$$

for differentiable g with continuous derivative g'. Specifically, let h_n be a downward-pointing tent to the left together with an upward-pointing tent to the right, with each tent having area $n/2$ (rather than 1):

$$h_n(x) \;=\; \begin{cases} 0 & (\text{for } x \le -\frac{1}{n}) \\ -2n^3 \cdot (x + \frac{1}{n}) & (\text{for } -\frac{1}{n} \le x \le -\frac{1}{2n}) \\ 2n^3 \cdot x & (\text{for } -\frac{1}{2n} \le x \le \frac{1}{2n}) \\ n^2 - 2n^3 \cdot (x - \frac{1}{2n}) & (\text{for } \frac{1}{2n} \le x \le \frac{1}{n}) \\ 0 & (\text{for } x \ge \frac{1}{n}) \end{cases}$$

Among other things, such examples are further evidence for the unfortunate limitations of the notion of *pointwise* values and limits.

---

3. *Measures: one attempt at greater generality*

Motivated by 19th century difficulties related to Fourier series, eigenfunction expansions, and related matters, soon after 1900 several people developed ideas to deal with *pointwise limits* of sequences of somewhat larger classes of functions.

The *Borel subsets* of $\mathbb{R}$ is the smallest collection of subsets of $\mathbb{R}$ closed under taking *countable unions*, under *countable intersections*, under *complements*, and containing all open and closed subsets of $\mathbb{R}$. This is also called the Borel *σ-algebra* in $\mathbb{R}$.

There is traditional terminology for certain simple types of Borel sets. For example a G_δ is a *countable intersection of open sets*, while an F_σ is a *countable union of closed sets*. The notation can be iterated: a $G_{\delta\sigma}$ is a countable union of countable intersections of opens, and so on. We will not need this.

A *Borel measure* μ is a way of assigning (often *positive*) real numbers (measures) to Borel sets, in a fashion that is *countably additive* for disjoint unions:

$$\mu(E_1 \cup E_2 \cup E_3 \cup \ldots) \;=\; \mu(E_1) + \mu(E_2) + \mu(E_3) + \ldots \qquad (\text{for } \textit{disjoint} \text{ Borel sets } E_1, E_2, E_3, \ldots)$$

[3] And we do not mean *derivative* of Dirac delta in the measure-theory context of *Radon-Nikodym* derivative, either.

A prototype is *Lebesgue (outer) measure* of a Borel set $E \subset \mathbb{R}$, described by

$$\mu(E) \;=\; \inf\{\sum_{n=1}^{\infty} |b_n - a_n| \;:\; E \subset \bigcup_{n=1}^{\infty} (a_n, b_n)\}$$

That is, it is the *inf* of the sums of lengths of the intervals in a countable cover of E by open intervals. For example, any countable set has (Lebesgue) measure 0.

We can consider larger classes of real-valued or complex-valued functions than just continuous ones, for example, various classes of *measurable* functions. The simplest useful choice is: A real-valued or complex-valued function f on $\mathbb{R}$ is *Borel-measureable* when the inverse image $f^{-1}(U)$ is a Borel set for every open set U in the target space.

It is occasionally useful to also allow the target space for functions to be the *two-point compactification* $Y = \{-\infty\} \cup \mathbb{R} \cup +\infty$ of the real line, with neighborhood basis $-\infty \cup (-\infty, a)$ at $-\infty$ and $(a, +\infty) \cup \{+\infty\}$ at $+\infty$ when we need to allow functions to blow up in some fashion.

A positive indicator:

[3.1] Theorem: Every pointwise limit of Borel-measurable functions f_n is Borel-measurable.

Verifying that we have not inadvertently needlessly included functions wildly unrelated to continuous functions:

[3.2] Theorem: *(Lusin)* Continuous functions approximate Borel-measurable functions well: given Borel-measurable real-valued or complex-valued f on $\mathbb{R}$, for every $\varepsilon > 0$ and for every Borel subset $\Omega \subset \mathbb{R}$ of finite Lebesgue measure, there is a relative closed $E \subset \Omega$ such that $\mu(\Omega - E) < \varepsilon$, and $f|_E$ is *continuous*.

Not much better can be done than Lusin's theorem says: for example, continuous approximations to the Heaviside step function

$$H(x) \;=\; \begin{cases} 0 & \text{for } x < 0 \\ \\ 1 & \text{for } x \geq 0 \end{cases}$$

have to go from 0 to 1 *somewhere*, by the Intermediate Value Theorem, so will be in $(\frac{1}{4}, \frac{3}{4})$ on an open set of strictly positive measure.

[3.3] Remark: It turns out that the everyday use of measure theory, measurable functions, and so on, does *not* proceed by way of Lusin's theorem or similar direct connections with continuous functions, but, rather, by direct interaction with the more general ideas.

A sequence $\{f_n\}$ of Borel-measurable functions on $\mathbb{R}$ converges (pointwise) *almost everywhere* when there is a Borel set $N \subset \mathbb{R}$ of measure 0 such that $\{f_n\}$ converges pointwise on $\mathbb{R} - N$.

[3.4] Theorem: *(Severini, Egoroff)* Pointwise convergence of sequences of Borel-measurable functions is approximately *uniform* convergence: given a almost-everywhere pointwise-convergent sequence $\{f_n\}$ of Borel-measurable functions on $\mathbb{R}$, for every $\varepsilon > 0$ and for every Borel subset $\Omega \subset \mathbb{R}$ of finite Lebesgue measure, there is a Borel subset $E \subset \Omega$ such that $\{f_n\}$ converges *uniformly* pointwise on E.

[3.5] Remark: Again, despite the connection that the Severini-Egoroff theorem makes between pointwise and *uniform* pointwise convergence, this idea turns out *not* to be the way to understand convergence of measurable functions. Instead, the game becomes ascertaining additional conditions that guarantee convergence of integrals, as just below.

With such notion of *measure*, there is a corresponding *integrability* and *integral*, due to Lebesgue. It amounts to replacing the literal rectangles used in Riemann integration by more general rectangles, with bases not just intervals, but measurable sets, as follows.

The *characteristic function* or *indicator function* ch_E or χ_E of a measurable subset $E \subset \mathbb{R}$ is 1 on E and 0 off. A *simple function* is a finite, positive-coefficiented, linear combination of characteristic functions of bounded measurable sets, that is, is of the form

$$\text{(simple function) } s \;=\; \sum_{i=1}^{n} c_i \cdot \mathrm{ch}_{E_i} \qquad\qquad \text{(with } c_i \geq 0\text{)}$$

The *integral* of s is what one would expect:

$$\int s\, d\mu \;=\; \int \Big(\sum_{i=1}^{n} c_i \cdot \mathrm{ch}_{E_i} \Big)\, d\mu \;=\; \sum_i c_i \cdot \mu(E_i)$$

Next, the measure of a *non-negative* function f is the *sup* of the integrals of all simple functions between f and 0:

$$\int f\, d\mu \;=\; \sup_{0 \leq s \leq f} \int s\, d\mu \qquad\qquad \text{(sup over simple } s \text{ with } 0 \leq s(x) \leq f(x) \text{ for all } x\text{)}$$

After proving that the positive and negative parts f_+ and f_- of Borel measurable real-valued f are again Borel measurable,

$$\int f\, d\mu \;=\; \int f_+\, d\mu - \int (-f_-)\, d\mu$$

Similarly, for complex-valued f, break f into real and imaginary parts.

There are details to be checked:

[3.6] Theorem: Borel-measurable functions f, g taking values in $[0, +\infty]$ are *integrable*, in the sense that the previous prescription yields an assignment $f \to \int_{\mathbb{R}} f \in [0, +\infty]$ such that for positive constants a, b

$$\int_{\mathbb{R}} (af + bg) \;=\; a \int_{\mathbb{R}} f + b \int_{\mathbb{R}} g \qquad\qquad \text{(for all } a, b \geq 0\text{)}$$

For complex-valued Borel-measurable f, g, the absolute values $|f|$ and $|g|$ are Borel-measurable. Assuming $\int_{\mathbb{R}} |f| < \infty$ and $\int_{\mathbb{R}} |g| < \infty$, for any complex a, b

$$\int_{\mathbb{R}} (af + bg) \;=\; a \int_{\mathbb{R}} f + b \int_{\mathbb{R}} g$$

Now we have practical criteria for the integral of a pointwise sequence to be the limit of the integrals:

[3.7] Theorem: *(Lebesgue's dominated convergence)* For Borel-measurable f_n with pointwise limit f, if there is non-negative Borel-measurable real-valued g such that $|f_n(x)| \leq g(x)$ for all x, and if g is *integrable* in the sense that $\int_{\mathbb{R}} g < +\infty$, then the pointwise limit is integrable, and

$$\lim_n \int_{\mathbb{R}} f_n \;=\; \int_{\mathbb{R}} \lim_n f_n$$

[3.8] Theorem: *(Monotone convergence)* For measurable extended-real-valued f_n with (extended-real) pointwise limit f, if $f_n(x) \leq f_{n+1}(x)$ for all x and for all indices n, then

$$\lim_n \int_{\mathbb{R}} f_n \;=\; \int_{\mathbb{R}} \lim_n f_n$$

(although the limit may be $+\infty$).

Less decisive-appearing, but in fact background for the previous two results, is

[3.9] Theorem: *(Fatou's lemma)* For Borel-measurable f_n with values in $[0,+\infty]$, the pointwise $f(x) = \liminf_n f_n(x)$ is Borel-measurable, and

$$\int \liminf_n f_n(x)\, dx \;\le\; \liminf_n \int f_n$$

More interesting, and more useful: after figuring out how to characterize measure on product spaces,

[3.10] Theorem: *(Fubini-Tonelli)* For complex-valued measurable f, g, if any one of $\int_{\mathbb{R}}\int_{\mathbb{R}} |f(x,y)|\, dx\, dy$, $\int_{\mathbb{R}}\int_{\mathbb{R}} |f(x,y)|\, dy\, dx$, or $\int_{\mathbb{R}\times\mathbb{R}} |f(x,y)|\, d\mathrm{vol}$ is finite, then the all are finite, and are equal. For $[0,+\infty]$-valued functions f, we have

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\, dx\, dy \;=\; \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\, dy\, dx \;=\; \int_{\mathbb{R}\times\mathbb{R}} f(x,y)\, d\mathrm{vol}$$

although the values may be $+\infty$.
