(*February 8, 2015*)

Pointwise convergence of Fourier series

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

[This document is
http://www.math.umn.edu/˜garrett/m/complex/notes_2014-15/09a_FourierDirichlet.pdf]

1. Proof of pointwise convergence
2. Elementary Hilbert space theory of Fourier series

The main point here is proof that Fourier series of a periodic function f of one real variable *converge pointwise* to f, under mild hypotheses. Something of this sort was first proven by Dirichlet in 1836.

[0.0.1] Theorem: Let f be piecewise C^o on the circle $S^1 \approx \mathbb{R}/2\pi\mathbb{R}$. Let x_o be a point at which f has both *left and right derivatives* (even if they do not agree), and is *continuous.* Then the Fourier series of f evaluated at x_o converges to $f(x_o)$. That is, $f(x_o)$ is expressed as a convergent series

$$f(x_o) \;=\; \sum_{n\in\mathbb{Z}} \widehat{f}(n)\, e^{inx_o} \qquad\qquad (\text{where } \widehat{f}(n) = \tfrac{1}{2\pi}\int_0^{2\pi} f(\xi)\, e^{-in\xi}\, d\xi)$$

(Proof below.)

Some background: *Fourier series* are finite or infinite linear combinations

$$\sum_{n\in\mathbb{Z}} c_n\, e^{inx}$$

In the early 19^{th} century, J. Fourier was an impassioned advocate of the use of such sums, of course writing sines and cosines rather than complex exponentials. Euler, the Bernouillis, and others had used such sums in similar fashions and for similar ends, but Fourier made a claim extravagant for the time, namely that *all functions* could be expressed in such terms. Unfortunately, in those days there was no clear idea of what a *function* was, no vocabulary to specificy *classes* of functions, and no specification of what it would mean to *represent* a function by such a series. In hindsight, probably issues of *pointwise* and L^2 convergence, unspecified to some degree, were confused with each other.

Conveniently, as exploited by Fourier and many others, a function expressed as a linear combination of exponentials is expressed a linear combination of eigenvectors for the differential operator d/dx. That is, Fourier expansions *diagonalize* the linear operator of differentiation. However, infinite-dimensional linear algebra is subtler than finite-dimensional.

At about the time Fourier was promoting Fourier series, Abel proved that convergent power series *can* be differentiated term-by-term in the interior of their interval (on $\mathbb{R}$) or disk (in $\mathbb{C}$) of convergence, and *are* infinitely-differentiable functions. However, the literal analogue of Abel's theorem for Fourier series cannot be correct: Fourier series of periodic C^1 functions *need not* be term-wise differentiable in an elementary sense, and certainly need not be indefinitely differentiable, all this *despite* having an immediately plausible recipe for that differentiation: *surely*

$$\frac{d}{dx}\sum_{n\in\mathbb{Z}} c_n\, e^{inx} \;=\; \sum_{n\in\mathbb{Z}} c_n\, in\, e^{inx}$$

The difficulty lies in suitable *interpretation* of this obviously-true assertion.

The *best* clarifications and resolutions of such issues needed viewpoints created first in 1906 by Beppo Levi, 1907 by G. Frobenius, in the 1930's by Sobolev, and L. Schwartz post-1949, enabling legitimate discussion of *generalized functions* (a.k.a., *distributions*). K. Friedrichs' important 1934-5 discussions of semi-bounded unbounded operators on Hilbert spaces used norms defined in terms of derivatives, but only internally in proofs, while for Levi, Frobenius, and Sobolev these norms were significant objects themselves.

A classic reference is A. Zygmund, *Trigonometric Series, I, II*, first published in Warsaw in 1935, reprinted several times, including a 1959 Cambridge University Press edition.

---

1. *Proof of pointwise convergence*

The L^2 *inner product* on C^o functions on $S^1 \approx \mathbb{R}/2\pi\mathbb{R}$ is

$$\langle f, g\rangle \;=\; \int_0^{2\pi} f(x)\,\overline{g(x)}\,dx$$

and the L^2 *norm* is $|f| = |f|_{L^2} = \langle f, f\rangle^{\frac{1}{2}}$. Let $\psi_n(x) = e^{inx}$. We can construe $L^2(S^1)$ as being the *completion* of $C^o(S^1)$ with respect to the metric given by $d(f,g) = |f-g|$. In these terms, the Fourier coefficients of f are

$$\widehat{f}(n) \;=\; \frac{\langle f, \psi_n\rangle}{|\psi_n|}$$

[1.0.1] Claim: *(Riemann-Lebesgue)* For $f \in L^2(S^1)$, the Fourier coefficients $\widehat{f}(n)$ of f go to 0.

Proof: The L^2 norm of ψ_n is $\sqrt{2\pi}$. *Bessel's inequality*[1]

$$|f|^2_{L^2} \;\;\geq\;\; \sum_n \Big|\Big\langle f, \frac{e^{inx}}{\sqrt{2\pi}}\Big\rangle\Big|^2$$

from abstract Hilbert-space theory applies to an orthonormal *set*, whether or not it is an orthonormal *basis*. Thus, the sum on the right converges, so by Cauchy's criterion the summands go to 0. ///

A function f on $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ is *(finitely) piecewise* C^o when there are finitely many real numbers $a_0 \leq a_1 \leq \ldots \leq a_{n-1} \leq a_n = a_0 + 1$ and C^0 functions f_i on $[a_i, a_{i+1}]$ such that

$$f_i(x) \;=\; f(x) \quad \text{on } [a_i, a_{i+1}] \;\; \text{(except possibly at the endpoints)}$$

Thus, while $f_i(a_{i+1})$ may differ from $f_{i+1}(a_{i+1})$, and $f(a_{i+1})$ may be different from both of these, the function f is continuous in the interiors of the intervals, and behaves well *near* the endpoints, if not *at* the endpoints.

Proof theorem: First, make reductions to unclutter the notation. By considering $f(x) - f(x_o)$, and observing that *constants* are represented pointwise by their Fourier expansions, assume $f(x_o) = 0$. The Fourier coeffients of *translates* of a function are expressible in terms of the Fourier coefficients of the function itself:

$$\int_0^{2\pi} f(x+x_o)\,\overline{\psi}_n(x)\,dx \;=\; \int_0^{2\pi} f(x)\,\overline{\psi}_n(x-x_o)\,dx \;=\; \psi_n(x_o)\int_0^{2\pi} f(x)\,\overline{\psi}_n(x)\,dx$$

The left-hand side is 2π times the n^{th} Fourier coefficient of $f(x+x_o)$, that is, the n^{th} Fourier *term* of $f(x+x_o)$ evaluated at 0, while the right-hand side is 2π times the n^{th} Fourier *term* of $f(x)$ evaluated at x_o. Thus, $x_o = 0$ without loss of generality.

---

[1] Proof of Bessel's inequality is straightforward: for *finite* orthonormal set $\{e_i\}$ and a vector v in a Hilbert space:

$$0 \;\leq\; |v - \sum_i \langle v, e_i\rangle e_i|^2 \;=\; |v|^2 - 2\sum_i \langle v, e_i\rangle\overline{\langle v, e_i\rangle} + \sum_i |\langle v, e_i\rangle|^2 \;=\; |v|^2 - \sum_i |\langle v, e_i\rangle|^2$$

For an *arbitrary* orthonormal set, the sum is the sup of the finite sub-sums, so the general Bessel inequality follows.

Partial sums of the Fourier expansion evaluated at 0 are

$$\sum_{-M\le n<N} \frac{1}{2\pi}\int_0^{2\pi} f(x)\overline{\psi}_n(x)\,dx \;=\; \frac{1}{2\pi}\int_0^{2\pi} f(x) \sum_{-M\le n<N} \overline{\psi}_n(x)\,dx$$

$$=\; \frac{1}{2\pi}\int_0^{2\pi} \frac{f(x)}{\psi_{-1}(x)-1}\,(\overline{\psi}_N(x)-\overline{\psi}_{-M}(x))\,dx \;=\; \frac{1}{2\pi}\Big\langle \frac{f}{\psi_{-1}-1},\,\psi_N\Big\rangle - \frac{1}{2\pi}\Big\langle \frac{f}{\psi_{-1}-1},\,\psi_{-M}\Big\rangle$$

The latter two terms are Fourier coefficients of $f/(\psi_{-1}-1)$, so go to 0 by the Riemann-Lebesgue lemma for $f(x)/(\psi_{-1}(x)-1)$ in $L^2(S^1)$. Since $x_o = 0$ and $f(x_o) = 0$

$$\frac{f(x)}{\psi_{-1}(x)-1} \;=\; \frac{f(x)}{x}\cdot\frac{x}{\psi_{-1}(x)-1} \;=\; \frac{f(x)-f(x_o)}{x-x_o}\cdot\frac{x-x_o}{e^{-ix}-e^{-ix_o}}$$

Existence of left and right derivatives of f at $x_o = 0$ is exactly the hypothesis that this expression has left and right limits at x_o, even if they do not agree.

At all other points the division by $\psi_{-1}(x)-1$ does not disturb the continuity. Thus, $f/(\psi_{-1}-1)$ is still at least *continuous* on each interval $[a_i, a_{i+1}]$ on which f was essentially a C^o function. Therefore, ignoring the endpoints, which do not contribute to the integrals, $f/(\psi_{-1}-1)$ is continuous on a finite set of closed (finite) intervals, so bounded on each one. Thus, $f/(\psi_{-1}-1)$ is indeed L^2, and we can invoke Riemann-Lebesgue to see that the integral goes to $0 = f(x_o)$. ///

---

2. *Elementary Hilbert space theory of Fourier series*

As observed in the previous section, the exponential functions

$$\psi_n(x) \;=\; e^{inx} \qquad\qquad (\text{for } n\in\mathbb{Z})$$

form an orthogonal *set* in $L^2(S^1)$. It is not clear that they form an orthogonal *basis*. That is, we should show that the finite linear combinations of the exponential functions ψ_n are *dense* in $L^2(S^1)$.

There are many different proofs that the normalized exponentials $\psi_n/\sqrt{2\pi}$ form an orthonormal basis in $L^2(S^1)$. At one extreme, a completely analogous result holds for any *compact abelian* topological group in place of S^1, without further structure. Even the abelian-ness can be dropped without much harm, though with complications. For the moment, we take advantage of the particulars of the circle to give a 19th-century-style argument. We *do* take this opportunity to introduce the notion of *approximate identity* (below).

[2.0.1] Theorem: The exponentials ψ_n are an orthogonal basis for $L^2(S^1)$.

[2.0.2] Remark: The proof does *not* prove that Fourier series converge pointwise, which they often do not. The not-necessarily-uniform pointwise convergence of Fourier series of C^1 functions does *not* immediately yield L^2 convergence, nor *uniform* pointwise convergence, which *would* imply L^2 convergence. Modifying the earlier pointwise convergence idea to achieve these goals motivates introduction of an *approximate identity* in the proof.

Proof: The continuous functions $C^o(S^1)$ are *dense* in $L^2(S^1)$ in the $L^2(S^1)$ topology. [2] Thus, it suffices to prove that C^o functions are approximable in L^2 by finite sums of the exponentials ψ_n, and it suffices to prove that finite sums of exponentials approximate C^o functions in the C^o topology: the total measure of the space S^1 is finite, so the L^2 norm of a continuous function is dominated by its sup norm, and density in sup norm implies density in L^2 norm.

---

[2] We can either take this as a definition, or prove it from some form of *Urysohn's Lemma* (see appendix).

[2.0.3] Remark: Proving that C^o functions are approximable in the sup norm by finite sums of exponentials does *not* require proving that Fourier series of continuous functions converge pointwise, which is not generally true. That is, we are *not* compelled to prove that the partial sums of the Fourier series are the approximating sums, despite these being an obvious candidate sequence. It turns out that this cannot possibly succeed. Examples of continuous functions whose Fourier series diverge were suggested by Riemann, made more rigorous by Weierstraß, and treated carefully in Fejér, L, (1910) *Beispiele stetiger Funktionen mit divergenter Fourierreihe* Journal Reine Angew. Math. 137, pp. 1-5. Existential arguments use the *Baire category theorem.*

We review the situation. As in the proof of not-necessarily-uniform pointwise convergence of Fourier series for piecewise C^1 functions, the N^{th} partial sum of the Fourier series of a function f on $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ is usefully described usefully an *integral operator*[3] by

$$\frac{1}{2\pi}\sum_{|k|\le n} \langle f, \psi_k\rangle\, \psi_k(x) \;=\; \frac{1}{2\pi}\int_{S^1} f(y) \sum_{|k|\le n} \overline{\psi}_k(y)\, \psi_k(x)\, dy \;=\; \frac{1}{2\pi}\int_{S^1} f(y) \sum_{|k|\le n} \psi_k(x-y)\, dy$$

$$= \frac{1}{2\pi}\int_{S^1} f(y)\, \frac{\psi_{n+1}(x-y) - \psi_{-n}(x-y)}{\psi_1(x-y) - 1}\, dy$$

by summing finite geometric series. Let $K_n(x)$ be the summed geometric series

$$K_n(x) \;=\; \frac{\psi_{n+1}(x) - \psi_{-n}(x)}{\psi_1(x) - 1} \;=\; \frac{e^{(n+1)ix} - e^{-inx}}{e^{-ix} - 1} \;=\; \frac{e^{(n+\frac{1}{2})x} - e^{-(n+\frac{1}{2})x}}{e^{ix/2} - e^{-ix/2}} \;=\; \frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}$$

Dropping the $1/2\pi$,

$$\sum_{|k|\le n} \langle f, \psi_k\rangle\, \psi_k(x) \;=\; \int_{S^1} f(y)\, K_n(x-y)\, dy$$

Granting that it is futile to prove that the partial sums converge pointwise for continuous functions, we might try to see what related but different integral operators would work better.

This is our excuse to introduce the generally-useful idea of *approximate identities.* The *rough* idea of *approximate identity* is of a sequence $\{\varphi_n\}$ of functions φ_n approximating a *point-mass measure*[4] at $0 \in S^1$. Precisely, a sequence of continuous functions φ_n on S^1 is an approximate identity when

$$\varphi_n(x) \;\ge\; 0 \quad (\text{for all } n, x) \qquad\qquad \int_{S^1} \varphi_n(x)\, dx \;=\; 1 \quad (\text{for all } n)$$

and if for every $\varepsilon > 0$ and for every $\delta > 0$ there is n_o such that for all $n \ge n_o$

$$\int_{|x|<\delta} \varphi_n(x)\, dx > 1 - \varepsilon \qquad \left(\text{equivalently, } \int_{\delta\le|x|<\frac{1}{2}} \varphi_n(x)\, dx < \varepsilon\right)$$

where we use coordinates in $\mathbb{R}$ for $S^1 = \mathbb{R}/2\pi\mathbb{Z}$. That is, the functions φ_n are non-negative, their integrals are all 1, and their *mass bunches up* at $0 \in S^1$. It is not surprising that the integral operators made from an approximate identity have a useful property:

---

[3] The general notion of *integral operator* T is of the form $Tf(x) = \int K(x,y)\, f(y)\, dy$, the function $K(x,y)$ is the *kernel* of the operator. The class of functions in which the kernel lies, and in which the input and output lie, varies enormously. Schwartz' kernel theorem hugely generalizes this idea.

[4] A *point-mass* measure at a point x_o is a measure which gives the point x_o measure 1 and gives a set not containing x_o measure 0. These are also called *Dirac measures.* A too-naive formulation of the notion of approximate identity fails: the requirement that φ_n form an approximate identity is *strictly stronger* than the condition that φ_n approach a Dirac measure in a *distributional* sense. Specifically, the *non-negativity* condition on an approximate identity is indispensable.

[2.0.4] Claim: For $f \in C^o(S^1)$ on S^1 and for an approximate identity φ_n, in the topology of $C^o(S^1)$,

$$\lim_n \int_{S^1} f(y)\,\varphi_n(x-y)\,dy \;=\; f(x)$$

Granting this claim, making an approximate identity out of finite sums of exponentials will prove that such finite sums are dense in $C^o(S^1)$. [5]

Proof: *(of claim)* Given $f \in C^o(S^1)$, and given $\varepsilon > 0$, by *uniform* continuity of f on the compact S^1, there is $\delta > 0$ such that

$$|f(x) - f(y)| \;<\; \varepsilon \qquad\qquad (\text{for } |x-y| < \delta)$$

Take n large enough so that

$$\int_{|x|<\delta} \varphi_n(x)\,dx \;>\; 1 - \varepsilon$$

Since the total mass of φ_n is 1,

$$\int_{S^1} f(y)\,\varphi_n(x-y)\,dy - f(x) \;=\; \int_{S^1} (f(y) - f(x))\,\varphi_n(x-y)\,dy$$

$$= \int_{|y-x|<\delta} (f(y) - f(x))\,\varphi_n(x-y)\,dy \;+\; \int_{\delta \le |y-x| \le \frac{1}{2}} (f(y) - f(x))\,\varphi_n(x-y)\,dy$$

The second integral is easily estimated by

$$\Big|\int_{\delta \le |y-x| \le \frac{1}{2}} (f(y) - f(x))\,\varphi_n(x-y)\,dy\Big| \;\le\; 2|f|_{C^o} \int_{\delta \le |y-x| \le \frac{1}{2}} \varphi_n(x-y)\,dy \;\le\; 2|f|_{C^o} \cdot \varepsilon$$

Estimation of the integral near 0 uses the *positivity* of the φ_n:

$$\Big|\int_{|y-x|<\delta} (f(y) - f(x))\,\varphi_n(x-y)\,dy\Big| \;\le\; \varepsilon \int_{|y-x|<\delta} \varphi_n(x-y)\,dy \;\le\; \varepsilon \int_{S^1} \varphi_n(x-y)\,dy \;=\; \varepsilon \cdot 1$$

This holds for all $\varepsilon > 0$ and uniformly in x, so the integrals approach $f(x)$ in the C^o topology, proving the claim. ///

As noted above, to prove the completeness, we could exhibit an approximate identity made from finite sums of exponentials. A failing of the *Dirichlet kernels*

$$K_n(x) \;=\; \frac{e^{(n+\frac{1}{2})x} - e^{-(n+\frac{1}{2})x}}{e^{ix/2} - e^{-ix/2}} \;=\; \frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}$$

that yield the partial sums is that these kernels are *not non-negative*, which would entail some complications. Still, the fact (proven above) that a stronger hypothesis of C^1-ness gives a limited result in the direction we want *suggests* that the masses of the K_n *do* bunch up near 0. Indeed, the expression for K_n in terms of sines *does* show that these functions are *real*-valued. Thus, a plausible choice for an approximate identity φ_n is the (essentially) *Fejér kernels*

$$\varphi_n \;=\; K_n^2 \times (\text{constant depending on } n)$$

[5] Similarly, to give S. Bernstein's tangible proof of the Stone-Weierstraß theorem that polynomials are dense in $C^o(K)$ for compact K in $\mathbb{R}^n$, an *approximate identity* is exhibited consisting of *polynomials*.

with constant chosen to give total mass 1. Computing directly in terms of exponentials, expanding the square of the sum,

$$\int_{S^1} K_n(x)^2\,dx \;=\; \int_{S^1} \left(e^{-inx} + \ldots + e^{inx}\right)^2\,dx$$

$$= \int_{S^1} e^{-2nix} + 2e^{-(2n-1)ix} + 3e^{-(2n-2)ix} + 4e^{-(2n-3)ix} + \ldots + (2n+1)\cdot 1 + \ldots + 2e^{(2n-1)ix} + e^{2nix}\,dx$$

$$= \int_{S^1} 2n+1\,dx \;=\; 2\pi\cdot(2n+1)$$

since all the non-trivial exponentials integrate to 0. Thus, take

$$\varphi_n(x) \;=\; \frac{K_n(x)^2}{2\pi\cdot(2n+1)}$$

The discussion so far gives non-negativity and total mass 1. We must show that the mass of the φ_n's bunches up at 0. For this, revert to the expression for K_n in terms of *sines*. For $\delta \le |x| \le \pi$,

$$\varphi_n(x) \;=\; \frac{\sin^2(n+\frac{1}{2})x}{2\pi(2n+1)\sin^2\frac{x}{2}} \;\le\; \frac{1}{2\pi(2n+1)\,x^2} \qquad (\text{since } |\sin\tfrac{x}{2}| \ge |\tfrac{x}{2}| \text{ for } |x|\le\pi)$$

For x bounded away from 0 we get inequalities such as

$$0 \;\le\; \varphi_n(x) \;\le\; \frac{1}{2\pi(2n+1)\,x^2} \;\le\; \frac{1}{2\pi(2n+1)\,n^{-2/3}} \;\le\; \frac{1}{2\pi(2n^{1/3}+n^{-2/3})} \;\le\; \frac{1}{2\pi n^{1/3}} \qquad (\text{for } |x| \ge n^{-1/3})$$

This sup-norm estimate for φ_n on the part of S^1 covered by $n^{-1/3} \le |x| \le \pi$ gives

$$\int_{n^{-1/3}\le|x|\le\pi} \varphi_n(x)\,dx \;\le\; 2\pi\cdot\frac{1}{2\pi\,n^{1/3}} \;\le\; \frac{1}{n^{1/3}}$$

Thus, given $\varepsilon > 0$ and $\delta > 0$, take n large enough so that $n > (2\varepsilon)^3$, and $n > \delta^3$, to meet the mass-concentration criterion for approximate identities. ///

Having shown that the exponential functions ψ_n form an orthogonal basis, for $f \in L^2(S^1)$ there is an L^2-equality

$$f \;=\; \sum_n \langle f, \frac{\psi_n}{|\psi_n|}\rangle\,\frac{\psi_n}{|\psi_n|} \;=\; \sum_n \widehat{f}(n)\,\psi_n \qquad (\text{in } L^2(S^1))$$

The *Plancherel-Parseval theorem* from general Hilbert space theory gives a Plancherel-Parseval theorem here:

$$|f|^2 \;=\; \sum_n |\widehat{f}(n)|^2$$

[2.0.5] Remark: Again, L^2 convergence says *nothing* directly about *pointwise* convergence. Nor is there anything to deny the possibility that a Fourier series *does converge* at a point, but converges to a value *different* from the value of f there.

[2.0.6] Remark: Fourier already worried about pointwise convergence of Fourier series, as did Cantor. From a later viewpoint than theirs, since L^2 functions are only defined *almost everywhere*, pointwise convergence of a Fourier series would distinguish a special function in the equivalence class in $L^2[0,1]$, which might be suspicious. Nevertheless, L. Carleson showed[6] that, given $\sum_n |c_n|^2 < \infty$, the Fourier series

[6] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135-157.

$\sum_n c_n e^{inx}$ converges almost everywhere. Thus, given $f \in L^2[0,2\pi]$, the Fourier series of f *does* converge almost everywhere to f, and *does* distinguish an element in that almost-everywhere-equal equivalence class. Fortunately, this difficult result does not play a role here.

---

3. *Appendix: Urysohn and density of* C^o

[3.0.1] Theorem: *(Urysohn)* In a locally compact Hausdorff topological space X, given a compact subset K contained in an open set U, there is a continuous function $0 \le f \le 1$ which is 1 on K and 0 off U.

Proof: First, we prove that there is an open set V such that

$$K \subset V \subset \overline{V} \subset U$$

For each $x \in K$ let V_x be an open neighborhood of x with compact closure. By compactness of K, some finite subcollection $V_{x_1}, \ldots, V_{x_n}$ of these V_x cover K, so K is contained in the open set $W = \bigcup_i V_{x_i}$ which has compact closure $\bigcup_i \overline{V}_{x_i}$ since the union is *finite*.

Using the compactness again in a similar fashion, for each x in the closed set $X - U$ there is an open W_x containing K and a neighborhood U_x of x such that $W_x \cap U_x = \phi$.

Then

$$\bigcap_{x \in X-U} (X-U) \cap \overline{W} \cap \overline{W}_x = \phi$$

These are compact subsets in a Hausdorff space, so (again from compactness) some *finite* subcollection has empty intersection, say

$$(X-U) \cap \left(\overline{W} \cap \overline{W}_{x_1} \cap \ldots \cap \overline{W}_{x_n}\right) = \phi$$

That is,

$$\overline{W} \cap \overline{W}_{x_1} \cap \ldots \cap \overline{W}_{x_n} \subset U$$

Thus, the open set

$$V = W \cap W_{x_1} \cap \ldots \cap W_{x_n}$$

meets the requirements.

Using the possibility of inserting an open subset and its closure between any $K \subset U$ with K compact and U open, we inductively create opens V_r (with compact closures) indexed by rational numbers r in the interval $0 \le r \le 1$ such that, for $r > s$,

$$K \subset V_r \subset \overline{V}_r \subset V_s \subset \overline{V}_s \subset U$$

From any such configuration of opens we construct the desired continuous function f by

$$f(x) = \sup\{r \text{ rational in } [0,1] : x \in V_r, \} = \inf\{r \text{ rational in } [0,1] : x \in \overline{V}_r, \}$$

It is not immediate that this sup and inf are the same, but if we *grant* their equality then we can prove the *continuity* of this function $f(x)$. Indeed, the sup description expresses f as the supremum of characteristic functions of open sets, so f is at least *lower semi-continuous.* [7] The inf description expresses f as an infimum of characteristic functions of closed sets so is *upper* semi-continuous. Thus, f would be continuous.

---

[7] A (real-valued) function f is *lower* semi-continuous when for all bounds B the set $\{x : f(x) > B\}$ is open. The function f is *upper* semi-continuous when for all bounds B the set $\{x : f(x) < B\}$ is open. It is easy to show that a sup of lower semi-continuous functions is lower semi-continuous, and an inf of upper semi-continuous functions is upper semi-continuous. As expected, a function both upper and lower semi-continuous is continuous.

To finish the argument, we must construct the sets V_r and prove equality of the inf and sup descriptions of the function f.

To construct the sets V_i, start by finding V_0 and V_1 such that

$$K \subset V_1 \subset \overline{V}_1 \subset V_0 \subset \overline{V}_0 \subset U$$

Fix a well-ordering $r_1, r_2, \ldots$ of the rationals in the open interval $(0,1)$. Supposing that $V_{r_1}, \ldots, v_{r_n}$ have been chosen. let i, j be indices in the range $1, \ldots, n$ such that

$$r_j > r_{n+1} > r_i$$

and r_j is the *smallest* among $r_1, \ldots, r_n$ *above* r_{n+1}, while r_i is the *largest* among $r_1, \ldots, r_n$ *below* r_{n+1}. Using the first observation of this argument, find $V_{r_{n+1}}$ such that

$$V_{r_j} \subset \overline{V}_{r_j} \subset V_{r_{n+1}} \subset \overline{V}_{r_{n+1}} \subset V_{r_i} \subset \overline{V}_{r_i}$$

This constructs the nested family of opens.

Let $f(x)$ be the sup and $g(x)$ the inf of the characteristic functions above. If $f(x) > g(x)$ then there are $r > s$ such that $x \in V_r$ and $x \notin \overline{V}_s$. But $r > s$ implies that $V_r \subset \overline{V}_s$, so this cannot happen. If $g(x) > f(x)$, then there are rationals $r > s$ such that

$$g(x) > r > s > f(x)$$

Then $s > f(x)$ implies that $x \notin V_s$, and $r < g(x)$ implies $x \in \overline{V}_r$. But $V_r \subset \overline{V}_s$, contradiction. Thus, $f(x) = g(x)$. ///

[3.0.2] Corollary: For a topological space X with a regular Borel measure μ, $C_c^o(X)$ is dense in $L^2(X, \mu)$.

Proof: The *regularity* of the measure is the property that $\mu(E)$ is both the *sup* of $\mu(K)$ for compacts $K \subset E$, and is the *inf* of $\mu(U)$ for opens $U \supset E$. From Urysohn's lemma, we have a continuous $f_{K,U}(x)$ which is 1 on K and 0 off U. Let $K_1 \subset K_2 \subset ...$ be a sequence of compacts inside E whose measure approaches that of E from *below*, and let $U_1 \supset U_2 \supset \ldots$ be a sequence of opens containing E whose measures approach that of E from *above*. Let f_i be a function as in Urysohn's lemma made to be 1 on K_i and 0 off U_i. Then Lebesgue's Dominated convergence theorem implies that

$$f_i \longrightarrow \text{(characteristic function of } E) \qquad \text{(in } L^2(X, \mu))$$

From the definition of integral of *measurable functions*, finite linear combinations of characteristic functions are dense in L^2 (or any other L^p with $1 \le p < \infty$). Thus, continuous functions are dense. ///
