(*November 4, 2017*)

# 12. Generalized functions (distributions) on circles

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

[This document is
http://www.math.umn.edu/~garrett/m/real/notes_2017-18/12_Levi-Sobolev.pdf]



The simplest physical object with an interesting function theory is the circle, $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, which inherits *group structure* and *translation-invariant* differential operator $d/dx$ from the real line $\mathbb{R}$. Equivalently, we can consider *periodic* functions on $\mathbb{R}$. This is *not* quite the same as considering functions $f$ on the interval $[0, 2\pi]$, unless we also explicitly require matching of function values and derivatives' values (when they exist) at the endpoints: $f(0) = f(2\pi)$, $f'(0) = f'(2\pi)$, and so on, to the extent applicable.

The exponential functions $x \to e^{inx}$ for $n \in \mathbb{Z}$ are all the continuous group homomorphisms $\mathbb{R}/2\pi\mathbb{Z} \to \mathbb{C}^\times$ *and* are all the eigenfunctions for $d/dx$ on $\mathbb{R}/2\pi\mathbb{Z}$. Finite or infinite linear combinations

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

are *Fourier series*. [1] Conveniently, a function so expressed is a linear combination, although an *infinite* linear combination, of eigenvectors for $d/dx$. That is, on functions *with* Fourier expansions [2] the linear operator of differentiation is *diagonalized*. However, infinite-dimensional linear algebra is subtler than finite-dimensional. Some fundamental questions are [3]

> *In what sense(s) can a function be expressed as a Fourier series?*
>
> *Can a Fourier series be differentiated term-by-term?*
>
> *How cautious must we be in differentiating functions that are only piecewise differentiable?*
>
> *What will derivatives of discontinuous functions be?*

[1] 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.

[2] The notion of *has a Fourier expansion* would need to clarify what *has such an expansion* means. Must it mean that pointwise values can be retrieved from the Fourier series? Less? More?

[3] 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. Abel's result fit the optimistic expectations of the time, but created unreasonable expectations for the behavior of Fourier series.

Several further issues are *implicit*, and the *best* answers need viewpoints created first in 1906 by Beppo Levi, 1907 by G. Frobenius, in the 1930's by Sobolev, and Schwartz post-1949, enabling legitimate discussion of *generalized functions* (a.k.a., *distributions*). [4] There are natural technical questions, like

*Why define generalized functions as dual spaces?*

In brief, Schwartz' 1940's insight to define generalized functions as *dual spaces* is a natural consequence of one natural *relaxation* of the notion of *function*. Rather than demand that functions produce *pointwise values*, which precipitated endless classical discussion of what to do with jump discontinuities, instead declare that *functions* in the broadest sense are merely things that can be *integrated against*. For given $\varphi$, the map that integrates against $\varphi$,

$$f \longrightarrow \int f(x)\,\varphi(x)\,dx$$

is a *functional* (a $\mathbb{C}$-valued linear map), and is, or ought to be, probably *continuous* in a reasonable topology. To consider the collection of *all* continuous linear functionals is a reasonable way to enlarge the collection of functions, as *things to be integrated against*.

From the other side, it might have been that this generalization of *function* is needlessly extravagant, but it turns out that every distribution on the circle $\mathbb{T}$ is a high-order derivative of a continuous function. Thus, since we *do* want to be able to take derivatives indefinitely, there is no waste.

Further, in any of the several natural topologies on distributions, very nice ordinary functions are *dense*, and the space of distributions is *complete* in a sense subsuming that for metric spaces. Thus, taking limits *yields* all distributions, *and* produces no excess.

This discussion is easiest on the circle $\mathbb{T}$, or products $\mathbb{T}^n$ of circles, making use of Fourier series, and clarifying many technical questions about Fourier series. [5] This story is a prototype for more complicated examples.

There is an important auxiliary technical point. Natural spaces of functions *do not* have structures of Hilbert spaces, but typically, of Banach spaces. Nevertheless, the simplicity of Hilbert spaces motivates comparisons of natural function spaces with related Hilbert spaces. Such comparisons are *Levi-Sobolev imbeddings* or *Levi-Sobolev inequalities*.

The present discussion presumes acquaintance with the basics of Fourier series, namely, the Fourier-Dirichlet kernel, the theorem of Fourier-Dirichlet on pointwise convergence for finitely-piecewise-continuous at points with left derivative and right derivative, Féjer's kernel, Féjer's theorem on the density of finite Fourier series in $C^o(\mathbb{T})$, and completeness of exponentials in $L^2(\mathbb{T})$.

We also presume that the notion of *(projective) limit* of Banach spaces is appreciated to some degree, at least in examples such as the *nested intersection*

$$C^\infty(\mathbb{T}) \;=\; \bigcap_k C^k(\mathbb{T}) \;=\; \lim_k C^k(\mathbb{T})$$

We recall this, and introduce *colimits*, especially in the case of *ascending unions* of spaces of *duals* of limits.

---

[4] 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.

[5] The 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. The present discussion neglects many interesting details, *but* is readily adaptible to more complicated situations, so necessarily our treatment is different from Zygmund's.

# 1. *Provocative example*

Let $s(x)$ be the *sawtooth function*[6]

$$s(x) \;=\; x - \pi \qquad (\text{for } 0 \le x < 2\pi)$$

and made *periodic* by demanding $s(x + 2\pi n) = s(x)$ for all $n \in \mathbb{Z}$. In other words, letting $[\![x/2\pi]\!]$ be the greatest integer less than or equal $x/2\pi$,

$$s(x) \quad = \quad x \;-\; 2\pi \cdot \left[\!\left[\frac{x}{2\pi}\right]\!\right] \;-\; \pi \qquad\qquad (\text{for } x \in \mathbb{R})$$

Away from $2\pi\mathbb{Z}$, the sawtooth function is infinitely differentiable, with derivative 1. At $x \in 2\pi\mathbb{Z}$ the sawtooth jumps down from value to $\pi$ to value $-\pi$. There is no reason to worry about defining a value at $x \in 2\pi\mathbb{Z}$.

The exponential functions $\psi_n(x) = e^{inx}$ are not quite an *orthonormal* basis for the Hilbert space $L^2[0, 2\pi]$, but are *orthogonal*:

$$\int_0^{2\pi} \psi_m(x) \cdot \overline{\psi}_n(x)\, dx \;\;=\;\; \begin{cases} 0 & (\text{for } m \neq n) \\ \\ 2\pi & (\text{for } m = n) \end{cases}$$

Anticipating that Fourier coefficients $\widehat{f}(n)$ of $2\pi\mathbb{Z}$-periodic functions $f$ are computed[7] by integrating against $\psi_n(x) = e^{inx}$ (conjugated):

$$\widehat{f}(n) \;\;=\;\; \frac{1}{2\pi} \int_0^{2\pi} f(x)\, e^{-inx}\, dx$$

integration by parts gives

$$\widehat{s}(n) \;\;=\;\; \frac{1}{2\pi} \int_0^{2\pi} s(x) \cdot e^{-inx}\, dx \;=\; \begin{cases} \dfrac{1}{-in} & (\text{for } n \neq 0) \\ \\ 0 & (\text{for } n = 0) \end{cases}$$

Thus, in whatever sense a function *is* its Fourier expansion, we anticipate that

$$s(x) \;\;\sim\;\; \sum_{n \in \mathbb{Z}} \widehat{s}(n) \cdot e^{inx} \;\;=\;\; \sum_{n \neq 0} \frac{1}{-in} \cdot e^{inx}$$

Even though this series does not converge absolutely for *any* value of $x$, we already know (by Fourier-Dirichlet) that it *does* converge to the value of $s(x)$ for $x \notin 2\pi\mathbb{Z}$. Since $s(x)$ has discontinuities at $2\pi\mathbb{Z}$ *anyway*, this is hardly surprising. Nothing disturbing has happened.

Now differentiate. The sawtooth function *is* differentiable away from $2\pi\mathbb{Z}$, with value 1, and with uncertain value at $2\pi\mathbb{Z}$. With exogenous reasons to differentiate the Fourier series term-by-term, with or without confidence in doing so, and the blatant differentiability of $s(x)$ away from $2\pi\mathbb{Z}$ suggests it's not entirely ridiculous to differentiate term-by-term. Then

$$s'(x) \;=\; \begin{cases} 1 & (\text{for } x \notin 2\pi\mathbb{Z}) \\ \\ ? & (\text{for } x \in 2\pi\mathbb{Z}) \end{cases} \qquad \sim \quad -\sum_{n \neq 0} e^{inx}$$

[6] One may also take $s(x) = x$ for $-\pi < x < \pi$ and extend by periodicity. This definition avoids the subtraction of $\pi$, and has the same operational features. In the end, it doesn't matter.

[7] Apparently at first Fourier did not have this expression for the Fourier coefficients!

The right-hand side is hard to interpret, certainly as having pointwise values. On the other hand, reasonably interpreted, it is still ok to integrate against this sum: letting $\widehat{f}(n)$ be the $n^{th}$ Fourier coefficient of a *smooth* function $f$, and not worrying about justifications,

$$\int_0^{2\pi} f(x)\,\Big(-\sum_{n\neq 0} e^{inx}\Big)\,dx \;=\; -\sum_{n\neq 0}\int_0^{2\pi} f(x)\,e^{inx}\;dx \;=\; -2\pi\sum_{n\neq 0}\widehat{f}(-n)$$

$$=\;2\pi\widehat{f}(0) - 2\pi\sum_{n\in\mathbb{Z}}\widehat{f}(n)\,e^{in\cdot 0} \;=\; 2\pi\widehat{f}(0) - 2\pi f(0) \;=\; \int_0^{2\pi} f(x)\,dx \;-\; 2\pi\cdot f(0)$$

The map

$$f \;\longrightarrow\; \int_0^{2\pi} f(x)\,dx \;-\; 2\pi\cdot f(0)$$

has a sense for continuous $f$, and gives a *functional.* That the derivative of the sawtooth is *mostly* 1 gives the integral of $f$ (against 1) over $[0, 2\pi]$. Further, the $-2\pi f(0)$ term forcefully suggests that *the derivative of the discontinuity of the sawtooth function is the (periodic) evaluation-at-*0 *functional* $f \to f(0)$ *multiplied by* $-2\pi$). [8]

**[1.1] Remark:** A truly disastrous choice at this point would be to think that since $s'(x)$ is *almost everywhere* 1 (in a measure-theoretic sense) that its singularities are somehow *removable*, and thus pretend that $s'(x) = 1$. This would give $s''(x) = 0$, and make the following worse than it is, and impossible to explain.

Still, $s'(x)$ *is* differentiable away from $2\pi\mathbb{Z}$, and by repeated differentiation

$$s^{(k+1)}(x) \;=\; \begin{cases} 0 & (\text{for } x \notin 2\pi\mathbb{Z}) \\ \\ ? & (\text{for } x \in 2\pi\mathbb{Z}) \end{cases} \quad\sim\quad -(i)^k\sum_{n\neq 0} n^k\cdot e^{inx}$$

By now the right-hand sides are vividly not convergent. The summands do not go to zero, in fact, are *unbounded.*

One can continue differentiating in this symbolic sense, but the meaning is unclear.

One reaction is to simply object to differentiating a non-differentiable function, even if its discontinuities are mild. This is not productive.

Another unproductive viewpoint is to deny that Fourier series reliably represent the functions that produced their coefficients.

A happier and more useful response is to suspect that the above computation is *correct*, though the question mark needs explanation, *and* that the right-hand side is correct and meaningful, *despite* its divergence in classical senses. The question is *what* meaning to attach. This requires preparation.

We will establish a context in which the derivatives of the sawtooth, and derivatives of other discontinuous functions, are *things to integrate against*, rather than *things to evaluate pointwise*, and see that termwise differentiation of Fourier series *does* capture an extended notion of function and derivative.

---

[8] The *jump* is downward rather than upward.

# 2. *Natural function spaces on the circle* $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$

We review natural families of functions. In all cases, the object is to give the vector space of functions a metric (if possible) which makes it *complete*, to allow *taking limits* inside the same class of functions. For example, *pointwise* limits of continuous functions easily fail to be continuous, but *uniform* pointwise limits of continuous functions *are* continuous. [9]

## [2.1] Continuous functions and sup-norm

First, we care about *continuous* complex-valued functions. Although we have in mind continuous functions on the circle $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, the basic result depends only upon the *compactness* of $\mathbb{R}/2\pi\mathbb{Z}$.

As usual, we give the set $C^o(K)$ of (complex-valued) continuous functions on a compact topological space $K$ the metric

$$d(f,g) \;=\; \sup_{x\in K} |f(x) - g(x)|$$

The sup is *finite* because $K$ is compact and $f - g$ is continuous. The right-hand side of this last equation arises from the *(sup) norm*

$$|f|_\infty \;=\; |f|_{C^o} = \sup_{x\in K} |f(x)|$$

and $d(f,g) = |f - g|_{C^o}$. A main feature of continuous functions is that they have *pointwise values*. Recall the unsurprising but important

**[2.2] Claim:** With the $C^o(K)$ topology, for $x \in K$ the *evaluation functional*[10] $C^o(K) \to \mathbb{C}$ by $f \to f(x)$ is *continuous*.

*Proof:* The inequality

$$|f(x) - g(x)| \;\le\; \sup_{y\in K} |f(y) - g(y)| \qquad\qquad (\text{for } f,g \in C^o(K))$$

proves the continuity of evaluation. ///

Also, recall, yet again, the iconic

**[2.3] Theorem:** The space $C^o(K)$ of (complex-valued) continuous functions on a compact topological space $K$ is *complete*.

**[2.4] Remark:** Thus, being complete with respect to the metric arising in this fashion from a *norm*, by definition $C^o(K)$ is a *Banach space*.

*Proof:* This is a typical three-epsilon argument. The point is the *completeness*, namely that a Cauchy sequence of continuous functions has a *pointwise* limit which is a continuous function. First we observe that a Cauchy sequence $f_i$ does have a pointwise limit. Given $\varepsilon > 0$, choose $N$ large enough such that for $i, j \ge N$ we have $|f_i - f_j| < \varepsilon$. Then, for any $x$ in $K$, $|f_i(x) - f_j(x)| < \varepsilon$. Thus, the sequence of values $f_i(x)$ is a Cauchy sequence of complex numbers, so has a limit $f(x)$. Further, given $\varepsilon' > 0$, choose $j \ge N$ sufficiently large such that $|f_j(x) - f(x)| < \varepsilon'$. Then for all $i \ge N$

$$|f_i(x) - f(x)| \;\le\; |f_i(x) - f_j(x)| + |f_j(x) - f(x)| \;<\; \varepsilon + \varepsilon'$$

[9] Awareness of such possibilities and figuring out how to avoid them was the fruit of embarrassing errors and experimentation throughout the $19^{th}$ century. Unifying abstract notions such as *metric space* and general *topological space* only became available in the early $20^{th}$ century, with the work of Hausdorff, Fréchet, and others.

[10] As usual, a *(continuous) functional* is a (continuous) linear map to $\mathbb{C}$.

Since this is true for every positive $\varepsilon'$

$$|f_i(x) - f(x)| \le \varepsilon \qquad \text{(for all } i \ge N\text{)}$$

This holds for every $x$ in $K$, so the pointwise limit is uniform in $x$.

Now prove that $f(x)$ is continuous. Given $\varepsilon > 0$, let $N$ be large enough so that for $i, j \ge N$ we have $|f_i - f_j| < \varepsilon$. From the previous paragraph

$$|f_i(x) - f(x)| \le \varepsilon \qquad \text{(for every } x \text{ and for } i \ge N\text{)}$$

Fix $i \ge N$ and $x \in K$, and choose a small enough neigborhood $U$ of $x$ such that $|f_i(x) - f_i(y)| < \varepsilon$ for any $y$ in $U$. Then

$$|f(x) - f(y)| \le |f(x) - f_i(x)| + |f_i(x) - f_i(y)| + |f(y) - f_i(y)| < \varepsilon + \varepsilon + \varepsilon$$

Thus, the pointwise limit $f$ is continuous at every $x$ in $U$. ///

## [2.5] Differentiation on $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$

To talk about *differentiability* return to the concrete situation of $\mathbb{R}$ and its quotient $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$.

The continuous quotient map $q : \mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$ yields continuous functions under composition $f \circ q$ for $f \in C^o(\mathbb{T}) = C^o(\mathbb{R}/2\pi\mathbb{Z})$. More is true, namely, that a continuous function $F$ on $\mathbb{R}$ is of the form $f \circ q$ if and only if $F$ is *periodic* in the sense that $F(x + 2\pi n) = F(x)$ for all $x \in \mathbb{R}$ and $n \in \mathbb{Z}$. Indeed, the periodicity gives a *well-defined* function $f$ on $\mathbb{R}/2\pi\mathbb{Z}$. Then the continuity of $f$ follows immediately from the definition of the quotient topology on $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$.

As usual, a real-valued or complex-valued function $f$ on $\mathbb{R}$ is *continuously differentiable* if it has a derivative itself a continuous function. That is, the limit

$$\frac{df}{dx}(x) = f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

is required to exist for all $x$, *and* the function $f'$ is in $C^o(\mathbb{R})$. Let $f^{(1)} = f'$, and inductively define

$$f^{(i)} = \Big(f^{(i-1)}\Big)' \qquad \text{(for } i > 1\text{)}$$

when the corresponding limits exist.

We can make explicit the expectation that *differentiation* on the circle $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ is descended from differentiation on the real line. That is, *characterize* differentiation on $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ in terms of such a compatibility relation. Thus, for $f \in C^k(\mathbb{T})$, require that the differentiation $D$ on $\mathbb{T}$ be related to the differentiation on $\mathbb{R}$ by

$$(Df) \circ q = \frac{d}{dx}(f \circ q)$$

Via the quotient map $q : \mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$, make a *preliminary* definition of the collection of $k$-times continuously differentiable functions on $\mathbb{T}$, with a topology, by

$$C^k(\mathbb{T}) = \{f \text{ on } \mathbb{T} : f \circ q \in C^k(\mathbb{R})\}$$

with the $C^k$-*norm*[11]

$$|f|_{C^k} = \sum_{0 \le i \le k} |(f \circ q)^{(i)}|_\infty = \sum_{0 \le i \le k} \sup_x |(f \circ q)^{(i)}(x)|$$

[11] Granting that the sup norm on *continuous* functions is a norm, verification that the $C^k$-norm is a norm is straightforward.

where $F^{(i)}$ is the (continuous!) $i^{th}$ derivative of $F$ on $\mathbb{R}$. The *associated metric* on $C^k(\mathbb{T})$ is

$$d(f,g) \;=\; |f-g|_{C^k}$$

**[2.6] Remark:** Among other features, the norm on the spaces $C^k$ makes continuity of the *differentiation* map $C^k \to C^{k-1}$ clear.

**[2.7] Remark:** Implicit in this definition is that, viewed as functions on $[0,2\pi]$, the values and derivatives *must agree at the endpoints*: $f(0)=f(2\pi)$ for $f$ continuous on $\mathbb{T}$, $f'(0)=f'(2\pi)$ for $f\in C^1(\mathbb{T})$, and so on. This is not whimsical, but is intrinsic to the structure of $\mathbb{T}$.

An often-seen equivalent version of the norm is

$$|f|^{\text{var}}_{C^k} \;=\; \sup_{0\le i\le k} |(f\circ q)^{(i)}|_\infty \;=\; \sup_{0\le i\le k}\,\sup_x |(f\circ q)^{(i)}(x)|$$

*These two norms give the same topology*, since for complex numbers $a_0,\ldots,a_k$

$$\sup_{0\le i\le k}|a_i| \;\le\; \sum_{0\le i\le k}|a_i| \;\le\; (k+1)\cdot\sup_{0\le i\le k}|a_i|$$

**[2.8] Claim:** There is a unique, well-defined, continuous (differentiation) map $D: C^k(\mathbb{T})\to C^{k-1}(\mathbb{T})$ giving a commutative diagram

$$\begin{tikzcd} C^k(\mathbb{R}) \arrow[r, "d/dx"] & C^{k-1}(\mathbb{R}) \\ C^k(\mathbb{T}) \arrow[u, "-\circ q"] \arrow[r, "D"] & C^{k-1}(\mathbb{T}) \arrow[u, "-\circ q"'] \end{tikzcd}$$

**[2.9] Remark:** One might feel that this proof is needlessly complicated. However, it is worthwhile to do it this way. This approach applies broadly, *and* is as terse as possible without ignoring important details.

*Proof:* The point is that differentiation of periodic functions yields periodic functions. That is, we claim that, for $f\in C^k(\mathbb{T})$, the pullback $f\circ q$ has derivative $\frac{d}{dx}(f\circ q)$ which is the pullback $g\circ q$ of a unique function $g\in \mathrm{C}^{k-1}(\mathbb{T})$. To see this, first recall that, by definition of the quotient topology, a continuous function $F$ on $\mathbb{R}$ descends to a continuous function on $\mathbb{T}=\mathbb{R}/2\pi\mathbb{Z}$ if and only if it is $2\pi\mathbb{Z}$-invariant, that is $F(x+2\pi n)=F(x)$ for all $x\in\mathbb{R}$ and $n\in\mathbb{Z}$. Then, from our definition of $C^k(\mathbb{T})$, a function $F\in C^k(\mathbb{R})$ is a pullback via $q$ from $C^k(\mathbb{R}/2\pi\mathbb{Z})$ exactly when $F^{(i)}(x+2\pi n)=F^{(i)}(x)$ for all $x\in\mathbb{R}$, $n\in\mathbb{Z}$, and $0\le i\le k$, since then these continuous functions descend to the circle. Let

$$(T_yF)(x) \;=\; F(x+y) \qquad\qquad (\text{for } x,y\in\mathbb{R})$$

Since $\frac{d}{dx}$ is a linear, constant-coefficient differential operator, the operations $T_y$ and $\frac{d}{dx}$ commute, that is, $\frac{\partial F}{\partial x}(x+y)=\frac{\partial}{\partial x}(F(x+y))$, which is to say

$$T_y\circ\frac{d}{dx} \;=\; \frac{d}{dx}\circ T_y$$

In particular, for $n\in\mathbb{Z}$,

$$T_{2\pi n}\Big(\frac{d}{dx}(f\circ q)\Big) \;=\; \frac{d}{dx}(T_{2\pi n}(f\circ q)) \;=\; \frac{d}{dx}(f\circ q)$$

This shows that a (continuous) derivative is periodic when the (continuously differentiable) function is periodic.

From the definition of the $C^k$-norm,

$$|Df|_{C^{k-1}} \;\le\; |f|_{C^k}$$

so differentiation is continuous. ///

**[2.10] Remark:** In light of the uniqueness of differentiation on $\mathbb{T}$, from now on write $d/dx$ for the differentiation $D$ on $\mathbb{T}$, and $f^{(k)}$ for $D^k f$, and rewrite the description of $C^k(\mathbb{T})$ more simply, as

$$C^k(\mathbb{T}) \;=\; \{f \;\text{ on } \mathbb{T}: \; f\circ q \in C^k(\mathbb{R})\}$$

with the $C^k$-*norm*

$$|f|_{C^k} \;=\; \sum_{0\le i\le k} |f^{(i)}|_\infty \;=\; \sum_{0\le i\le k} \sup_x |f^{(i)}(x)|$$

where $f^{(i)}$ is the (continuous!) $i^{th}$ derivative of $f$. The *associated metric* on $C^k(\mathbb{T})$ still is

$$d(f,g) \;=\; |f-g|_{C^k}$$

There is the alternative norm

$$|f|^{\mathrm{var}}_{C^k} \;=\; \sup_{0\le i\le k}\,\sup_x |f^{(i)}(x)| \;=\; \sup_{0\le i\le k} |f^{(i)}|_\infty$$

Again, these two norms give the same topology, for the same reason as before.

**[2.11] Claim:** With the topology above, the space $C^k(\mathbb{T})$ is *complete*, so is a *Banach* space.

*Proof:* The case $k=1$ illustrates all the points. For a Cauchy sequence $\{f_n\}$ in $C^1(\mathbb{T})$, both $\{f_n\}$ and $\{f'_n\}$ are Cauchy in $C^o(\mathbb{T})$, so converge uniformly pointwise: let

$$f(x) \;=\; \lim_n f_n(x) \qquad\qquad g(x) \;=\; \lim_n f'_n(x)$$

The convergence is uniformly pointwise, so $f$ and $g$ are $C^o$. *If* we knew that $f$ were pointwise differentiable, then the demonstrated continuity of $\frac{d}{dx}: C^1(\mathbb{T}) \to C^o(\mathbb{T})$ gives the expected conclusion, that $f' = g$.

What could go wrong? One issue is whether $f$ *is differentiable* at all, and why its derivative is $g$.

By the fundamental theorem of calculus, for any index $i$, since $f_i$ is continuous,[12]

$$f_i(x) - f_i(a) \;=\; \int_a^x f'_i(t)\,dt$$

Interchanging limit and integral[13] shows that the limit of the right-hand side is

$$\lim_i \int_a^x f'_i(t)\,dt \;=\; \int_a^x \lim_i f'_i(t)\,dt \;=\; \int_a^x g(t)\,dt$$

Thus, the limit of the left-hand side is

$$f(x) - f(a) \;=\; \int_a^x g(t)\,dt$$

[12] The fundamental theorem of calculus for integrals of *continuous* functions needs only the simplest notion of an integral, for example, Riemann integrals.

[13] For example, interchange of limit and integral is justified by the simplest form of Lebesgue's Dominated Convergence Theorem. Also, for uniform pointwise limits of continuous functions, this can be proven directly.

from which $f' = g$. That the derivative $f'$ of the limit $f$ is the limit of the derivatives is not a surprise, since if $f$ is differentiable, what else could its derivative be? The point is that $f$ *is* differentiable, ascertained by computing its derivative, which happens to be $g$. ///

**[2.12] Remark:** Again, the differentiation map $C^1(\mathbb{T}) \to C^o(\mathbb{T})$ is continuous *by design*. Thus, *if* a limit of $C^1$ functions $f_n$ *is* differentiable, its derivative must be the obvious thing, namely, the limit of the derivatives $f_n'$. The issue was whether the limit of the $f_n$ is differentiable. The proof shows that it is differentiable by computing its derivative via the Mean Value Theorem.

By construction, and from the corresponding result for $C^o$,

**[2.13] Claim:** With the $C^k$-topology, for $x \in \mathbb{T}$ and integer $0 \le i \le k$, the *evaluation functional* $C^k(\mathbb{T}) \to \mathbb{C}$ by

$$f \longrightarrow f^{(i)}(x)$$

is *continuous.* ///

This applies to Fourier series, without any claim about what functions are representable as Fourier series. With $\psi_n(x) = e^{inx}$,

**[2.14] Claim:** For complex numbers $c_n$, when

$$\sum_n |c_n| \cdot |n|^k < +\infty$$

the Fourier series $\sum c_n \psi_n$ converges to a function in $C^k(\mathbb{T})$, and its derivative is computed by termwise differentiation

$$\frac{d}{dx} \sum c_n \psi_n = \sum (in)\, c_n \psi_n \in C^{k-1}(\mathbb{T})$$

*Proof:* The $C^o(\mathbb{T})$ norm of a Fourier series is easily estimated, by

$$\Big| \sum_{|n|\le N} c_n \psi_n(x) \Big| \le \sum_{|n|\le N} |c_n| \qquad \text{(for all } x \in \mathbb{T})$$

The right-hand side is independent of $x \in \mathbb{T}$, so bounds the sup over $x \in \mathbb{T}$. Similarly, estimate derivatives (of partial sums) by

$$\left| \Big( \sum_{|n|\le N} c_n \psi_n \Big)^{(k)} \right| \le \sum_{|n|\le N} |c_n|\, n^k$$

Thus, the hypothesis of the claim implies that the partial sums form a Cauchy sequence in $C^k$. The partial sums of a Fourier series are *finite* sums, so can be differentiated term-by-term. Thus, we have a Cauchy sequence of $C^k$ functions, which converges to a $C^k$ function, by the completeness of $C^k$. That is, the given estimate assures that the Fourier series converges to a $C^k$ function.

Further, since differentiation is a continuous map $C^k \to C^{k-1}$, it maps Cauchy sequences to Cauchy sequences. In particular, the Cauchy sequence of derivatives of partial sums converges to the derivative of the limit of the original Cauchy sequence. ///

We *want* the following to hold. Unsurprisingly, it does:

**[2.15] Claim:** The inclusion $C^k(\mathbb{T}) \subset C^{k-1}(\mathbb{T})$ is continuous. [14]

[14] In fact, the image of $C^k$ in $C^{k-1}$ is *dense*, but, we will prove this later as a side-effect of sharper results.

*Proof:* The point is that, for $f \in C^k(\mathbb{T})$ the obvious inequality

$$|f|_{C^{k-1}} \;\le\; |f|_{C^k}$$

gives an explicit estimate for the continuity. ///

---

# 3. *Topology on* $C^\infty(\mathbb{T})$

Next, we care about infinitely differentiable [15] functions, that is, *smooth* functions, denoted $C^\infty(\mathbb{T})$. At least as *sets* (or vector spaces),

$$C^\infty(\mathbb{T}) \;=\; \bigcap_k C^k(\mathbb{T})$$

However, this space $C^\infty(\mathbb{T})$ of smooth functions provably *does not* have a structure of Banach space. Observing that a descending intersection is a *(projective) limit* we should declare that

$$C^\infty(\mathbb{T}) \;=\; \lim_k C^k(\mathbb{T})$$

That is, for every topological vector space $V$ and compatible [16] family of continuous linear maps $f_k : V \to C^k(\mathbb{T})$, there is a unique $f : V \to C^\infty(\mathbb{T})$ such that all triangles commute in the diagram

$$\begin{tikzcd}
C^\infty(\mathbb{T}) \arrow[rr, bend left] \arrow[rrr, bend left=40] & \dots \arrow[r] & C^1(\mathbb{T}) \arrow[r] & C^o(\mathbb{T}) \\
& V \arrow[ul, dotted, "f"] \arrow[ur, dashed, "f_1"] \arrow[urr, dashed, "f_o"'] & &
\end{tikzcd}$$

Unfortunately, we may be temporarily insufficiently sophisticated about what kind of object this limit might be. In particular, we do not know what kind of auxiliary objects to use in the very definition of *limit*.

Too optimistic speculation about what the limit might be leads to trouble: as it happens, this limit is provably *not* a Banach space (nor Hilbert space). [17] As we have seen, a limit of topological spaces has a unique topology, whatever it may be, by the categorical characterization of this topology.

**[3.1] Remark:** There is also the disquieting question of what *test objects* $V$ we should consider in the diagrammatic characterization, with compatible mappings $V \to C^k(\mathbb{T})$ to *characterize* the limit.

The broadest necessary class of vector spaces with topologies is the following. A *topological vector space* is what one would reasonably imagine, namely, a (complex) vector space $V$ with a topology such that

$$V \times V \to V \quad \text{by} \quad v \times w \to v + w \quad \text{is continuous}$$

---

[15] Use of *infinitely* here is potentially misleading, but is standard. Sometimes the phrase *indefinitely differentiable* is used, but this also offers its own potential for confusion. A better (and standard) contemporary usage is *smooth*.

[16] As earlier, for the maps $f_k$ to be *compatible* means that, naming the inclusion $i_k : C^k(\mathbb{R}) \to C^{k-1}(\mathbb{R})$, $i_k \circ f_k = f_{k-1}$. That is, all the triangles in the relevant diagram commute.

[17] The non-Banach-ness of $C^\infty(\mathbb{T})$ is not the main point, but it is reasonable to wonder how this is proven. Briefly, with a definition of *topological vector space*, we will prove that a topological vector space is *normable* if and only if there is a local basis at 0 consisting of *bounded* opens. This is independent of *completeness*. The relevant sense of bounded *cannot* be the usual metric sense. Instead, a set $E$ in a topological vector space is *bounded* when, for every open neighborhood $U$ of 0, there is $t > 0$ such that $E \subset z \cdot U$ for all complex $z$ with $|z| \ge t$. That is, sufficiently large *dilates* of opens eventually contain $E$. But we will eventually that open balls in $C^k(\mathbb{T})$ are *not* contained in *any* dilate of any open ball in $C^{k+1}(\mathbb{T})$. The definition of the limit topology then shows that $C^\infty(\mathbb{T})$ is not normable. A more detailed discussion will be given later.

and such that

$$\mathbb{C} \times V \to V \quad \text{by} \quad \alpha \times v \to \alpha \cdot v \quad \text{is continuous}$$

and such that the topology is *Hausdorff.* [18] We require that the topological vector spaces be *locally convex* in the sense that there is a local basis at 0 consisting of *convex* sets. [19] It is easy to prove that Hilbert and Banach spaces are locally convex, which is why the issue is invisible in that context. Dismayingly, there are easily constructed complete (invariantly) metrized topological vector spaces which are *not* locally convex. [20]

Returning to the discussion of limits of topological vector spaces: since the continuity requirements for a topological vector space are of the form $A \times B \to C$ (rather than having the arrow going the other direction), there is a *diagrammatic* argument that the continuous algebraic operations on the limitands induce continuous algebraic operations on the limit, in the limit topology (as limit of topological spaces).

**[3.2] Claim:** Products and limits of topological vector spaces exist. Products and limits of locally convex spaces are locally convex. (Proof in appendix.)

**[3.3] Remark:** As usual, if they exist at all, then products and limits are unique up to unique isomorphism.

Thus, $C^\infty(\mathbb{T})$ has a (limit) topology for general reasons. As proven earlier for such spaces on intervals $[a, b]$,

**[3.4] Claim:** Differentiation $f \to f'$ is a *continuous* map $C^\infty(\mathbb{T}) \to C^\infty(\mathbb{T})$.

**[3.5] Remark:** *Of course* differentiation maps the smooth functions to themselves. Continuity of differentiation in the *limit* topology is less clear.

*Proof:* Differentiation $d/dx$ gives a continuous map $C^k(\mathbb{T}) \to C^{k-1}(\mathbb{T})$. Differentiation is compatible with the inclusions among the $C^k(\mathbb{T})$. Thus, we have a commutative diagram

$$\begin{tikzcd}
C^\infty(\mathbb{T}) \arrow[rr, bend left] \arrow[rrr, bend left] & \ldots & C^k(\mathbb{T}) \arrow[r] & C^{k-1}(\mathbb{T}) \arrow[r] & \ldots \\
C^\infty(\mathbb{T}) \arrow[rr, bend right] \arrow[rrr, bend right] & \ldots & C^k(\mathbb{T}) \arrow[r] \arrow[ur, "\frac{d}{dx}"] & C^{k-1}(\mathbb{T}) \arrow[r] \arrow[ur, "\frac{d}{dx}"] & \ldots
\end{tikzcd}$$

Composing the projections with $d/dx$ gives (dashed) induced maps from $C^\infty(\mathbb{T})$ to the limitands, inducing a unique (dotted) map to the limit, as in

$$\begin{tikzcd}
C^\infty(\mathbb{T}) \arrow[rr, bend left] \arrow[rrr, bend left] & \ldots & C^k(\mathbb{T}) \arrow[r] & C^{k-1}(\mathbb{T}) \arrow[r] & \ldots \\
C^\infty(\mathbb{T}) \arrow[u, dotted, "\frac{d}{dx}"] \arrow[urr, dashed] \arrow[urrr, dashed] \arrow[urrrr, dashed] \arrow[rr, bend right] \arrow[rrr, bend right] & \ldots & C^k(\mathbb{T}) \arrow[r] \arrow[ur] & C^{k-1}(\mathbb{T}) \arrow[r] \arrow[ur] & \ldots
\end{tikzcd}$$

---

[18] In fact, soon after giving the definition, one can show that the weaker condition that *points are closed*, implies the Hausdorff condition in topological spaces which are vector spaces with continuous vector addition and scalar multiplication. Indeed, the inverse image of $\{0\}$ under $x \times y \to x - y$ is the diagonal.

[19] This sense of convexity is the usual: a set $X$ in a vector space is convex when, for all tuples $x_1, \ldots, x_n$ of points in $X$ and all tuples $t_1, \ldots, t_n$ of non-negative reals with $\Sigma_i t_i = 1$, the sum $\Sigma_i t_i x_i$ is again in $X$.

[20] The simplest examples of complete metric topological vector spaces which are *not* locally convex are spaces $\ell^p$ with $0 < p < 1$. The metric comes from a norm-like function which is *not* a norm: $|\{c_n\}|_p = \sum_n |c_n|^p$. No, there is no $p^{th}$ root taken, unlike the spaces $\ell^p$ with $p \geq 1$, and this causes the function $|\ |_p$ to *lose* the homogeneity it would need to be a norm. Nevertheless, such a space is *complete.* It is an amusing exercise to prove that it is not locally convex.

This proves the continuity of differentiation, in the limit topology. ///

**[3.6] Corollary:** When a Fourier series $\sum_n c_n \psi_n$ satisfies

$$\sum_m |c_n|\,|n|^N \;<\; +\infty \qquad\qquad (\text{for every } N)$$

the series is a smooth function, which can be differentiated term-by-term, and its derivative is

$$\sum_m c_n \cdot in \cdot \psi_n$$

*Proof:* The hypothesis assures that the Fourier series lies in $C^k$ for every $k$. Differentiation is continuous in the limit topology on $C^\infty$. ///

**[3.7] Remark:** This continuity is necessary to define differentiation of *distributions* below.

---

# 4. *Distributions: generalized functions*

Although much amplification is needed, having an appropriate topology on $C^\infty(\mathbb{T})$ allows the bare definition: a *distribution* or *generalized function*[21] on $\mathbb{T}$ is a *continuous linear functional*[22]

$$u : C^\infty(\mathbb{T}) \longrightarrow \mathbb{C}$$

*Why a dual space?* Unsurprisingly, especially with a precise *intrinsic* notion of *integral* on $\mathbb{T}$ in the next section, a function $\varphi \in C^o(\mathbb{T})$ gives rise to a distribution $u_\varphi$ by *integration against* $\varphi$,

$$u_\varphi(f) \;=\; \int_{\mathbb{T}} f(x)\,\varphi(x)\,dx \qquad\qquad (\text{for } f \in C^\infty(\mathbb{T}))$$

Thus, we relax our notion of *function*, no longer requiring *pointwise values*, but only that a function can be *integrated against.* Then it may make sense to declare functionals in a dual space to be *generalized* functions. The vector space of distributions is denoted

$$\text{distributions} \;=\; \text{continuous dual of } C^\infty(\mathbb{T}) \;=\; \mathrm{Hom}^o_{\mathbb{C}}(C^\infty(\mathbb{T}),\mathbb{C}) \;=\; C^\infty(\mathbb{T})^*$$

That is, given a reasonable notion of integral, we have a continuous imbedding

$$C^o(\mathbb{T}) \;\subset\; C^\infty(\mathbb{T})^* \quad \text{by} \quad \varphi \longrightarrow u_\varphi \quad \text{where (again)} \quad u_\varphi(f) \;=\; \int_{\mathbb{T}} f(x)\,\varphi(x)\,dx \quad (f \in C^\infty(\mathbb{T}))$$

Typically, the dual of a limit of topological vector spaces is not the colimit of the duals of the limitands. Duals of *colimits do* behave well, in the sense that in reasonable situations

$$\mathrm{Hom}(\mathrm{colim}_i X_i, Z) \;\approx\; \lim_i \mathrm{Hom}(X_i, Z)$$

---

[21] What's in a name? In this case, *generalized function* expresses the *intention* to think of distributions as extensions of ordinary functions, not as abstract things in a dual space.

[22] The standard usage is that a *functional* on a complex vector space $V$ is a $\mathbb{C}$-linear map from $V$ to $\mathbb{C}$. Continuity may or may not be required, and the topology in which continuity is required may vary. It is in this sense that there is a subject *functional analysis.*

But $C^\infty(\mathbb{T})$ is a *limit*, not a colimit. Luckily, the dual of a limit of *Banach spaces* is the colimit of the duals:

**[4.1] Theorem:** Let $X = \lim_i B_i$ be a limit of Banach spaces $B_i$ with projections $p_i : X \to B_i$. Any $\lambda \in X^* = \mathrm{Hom}^o_{\mathbb{C}}(X, \mathbb{C})$ *factors through* some $B_i$. That is, there is $\lambda_j : B_j \to \mathbb{C}$ such that

$$\lambda \;=\; \lambda_j \circ p_j : X \to \mathbb{C}$$

Therefore,

$$(\lim_i B_i)^* \;\approx\; \mathrm{colim}_i B_i^*$$

*Proof:* Without loss of generality, each $B_i$ is the closure of the image of $X$, since otherwise replace of each $B_i$ by that closure.

Let $U$ be an open neighborhood of 0 in $X = \lim_i B_i$ such that $\lambda(U)$ is inside the open unit ball at 0 in $\mathbb{C}$, by the continuity at 0. By properties of the limit topology [23] there are finitely-many indices $i_1, \ldots, i_n$ and open neighborhoods $V_{i_t}$ of 0 in $B_{i_t}$ such that

$$\bigcap_{t=1}^n p_{i_t}^{-1} V_{i_t} \;\subset\; U \qquad\qquad \text{(projections } p_i \text{ from the } \textit{limit } X\text{)}$$

To have $\lambda$ factor (continuously) through a limitand $B_j$, we need a *single* condition to replace the conditions from $i_1, \ldots, i_n$. Let $j$ be *any* index [24] with $j \geq i_t$ for all $t$, and

$$V_j' \;=\; \bigcap_{t=1}^n p_{i_t,j}^{-1} V_{i_t} \subset B_j$$

By the compatibility

$$p_{i_t}^{-1} \;=\; p_j^{-1} \circ p_{i_t,j}^{-1}$$

we have a single sufficient condition, namely $p_j^{-1} V_j' \subset U$. By the linearity of $\lambda$, for $\varepsilon > 0$

$$\lambda(\varepsilon \cdot p_j^{-1} V_j) \;=\; \varepsilon \cdot \lambda(p_j^{-1} V_j) \;\subset\; \varepsilon\text{-ball in } \mathbb{C}$$

By continuity [25] of scalar multiplication on $B_j$, $\varepsilon \cdot V_j'$ is an open containing 0 in $B_j$.

We claim that $\lambda$ factors through $p_j X$ with the subspace topology from $B_j$. This makes $p_j X$ a *normed* space, if not Banach. [26] Simplifying notation, let $\lambda : X \to \mathbb{C}$ and $p : X \to N$ be continuous linear to a normed space $N$, with

$$\lambda(p^{-1} V) \;\subset\; \text{unit ball in } \mathbb{C} \qquad\qquad \text{(for some neighborhood } V \text{ of 0 in } N\text{)}$$

---

[23] Recall that $X = \lim_i B_i$ is the closed subspace (with the subspace topology) of the product $Y = \Pi_i B_i$ of all tuples $\{b_i\}$ in which $p_{ij} : b_i \to b_j$ for $i > j$ under the transition maps $p_{ij} : B_i \to B_j$. A local basis at 0 in the product consists of products $V = \Pi_i V_i$ of opens $V_i$ in $B_i$ with $V_i = B_i$ for all but finitely-many $i$, say $i_1, \ldots, i_n$.

[24] The index set need not be the positive integers, but must be a *poset* (partially ordered set), *directed*, in the sense that for any two indices $i, j$ there is an index $k$ such that $k > i$ and $k > j$.

[25] Multiplication by a non-zero scalar is a *homeomorphism*: scalar multiplication by $\varepsilon \neq 0$ is continuous, scalar multiplication by $\varepsilon^{-1}$ is continuous, and these are mutual inverses, so these scalar multiplications are *homeomorphisms*.

[26] Recall that a *normed* space is a topological vector with topology given by a *norm* $|\;\;|$ as in a Banach space, but *without* the requirement that the space is *complete* with respect to the metric $d(x, y) = |x - y|$. This slightly complicated assertion is correct: in most useful situations $p_j X$ is rarely all of $B_j$, even when $B_j$ is a *completion* of $p_j X$.

We claim that $\lambda$ factors through $p : X \to N$ as a (continuous) linear map. Indeed, by the linearity of $\lambda$,

$$\lambda(\frac{1}{n} \cdot p^{-1}V) \;\subset\; \frac{1}{n}\text{-ball in } \mathbb{C}$$

so

$$\lambda\Big(\bigcap_n \frac{1}{n} \cdot p^{-1}V\Big) \;\subset\; \frac{1}{m}\text{-ball} \qquad\qquad (\text{for all } m)$$

Then

$$\lambda\Big(\bigcap_n \frac{1}{n} \cdot p^{-1}V\Big) \;\subset\; \bigcap_m \frac{1}{m}\text{-ball} \;=\; \{0\}$$

Thus,

$$\bigcap_n p^{-1}(\frac{1}{n} \cdot V) \;=\; \bigcap_n \frac{1}{n} \cdot p^{-1}V \;\subset\; \ker\lambda$$

For $x, x'$ in $X$ with $px = px'$, certainly $px - px' \in \frac{1}{n}V$ for all $n = 1, 2, \ldots$. Therefore,

$$x - x' \;\in\; \bigcap_n p^{-1}(\frac{1}{n}V) \;\subset\; \ker\lambda$$

and $\lambda x = \lambda x'$. This proves the subordinate claim that $\lambda$ factors through $p : X \to N$ via a (not necessarily continuous) linear map $\mu : N \to \mathbb{C}$. For the continuity of $\mu$, by its linearity

$$\mu(\varepsilon \cdot V) \;=\; \varepsilon \cdot \mu V \;\subset\; \varepsilon\text{-ball in } \mathbb{C}$$

proving the continuity of $\mu : N \to \mathbb{C}$. [27] This proves the claim.

The claim gives continuous linear $\lambda_j : p_jX \to \mathbb{C}$ through which $\lambda$ factors.

Then $\lambda_j : p_jX \to \mathbb{C}$ extends by continuity [28] to the closure of $p_jX$ in $B_j$, which is $B_j$, giving the desired map. ///

**[4.2] Remark:** The same proof shows that a continuous linear map from a limit of Banach spaces to a *normed* space factors through a limitand, when the images of projections are dense in the limitands.

**[4.3] Corollary:** The space of distributions on $\mathbb{T}$ is the ascending union (colimit)

$$C^\infty(\mathbb{T})^* \;=\; \big(\lim{}_k C^k(\mathbb{T})\big)^* \;=\; \mathrm{colim}_k C^k(\mathbb{T})^* \;=\; \bigcup{}_k C^k(\mathbb{T})^*$$

of duals of the Banach spaces $C^k(\mathbb{T})$. ///

---

[27] Here we need $V$ to be *open*, not merely a *set* containing 0. Continuity at 0 is all that is needed for continuity of *linear* maps, since $|\lambda(x)| < \varepsilon$ for $|x| < \delta$ gives $|\lambda(x - x')| < \varepsilon$ for $|x - x'| < \delta$.

[28] The extension by continuity is unambiguous, since $\lambda_j$ is *linear*. In more detail: for $\lambda$ a continuous linear function on a dense subspace $Y$ of a topological vector space $X$, given $\varepsilon > 0$, take *convex* neighborhood $U$ of 0 in $X$ such that $|\lambda y| < \varepsilon$ for $y \in U$. We may suppose $U = -U$ by replacing $U$ by $-U \cap U$. Let $y_i$ be a Cauchy net approaching $x \in X$. For $y_i$ and $y_j$ inside $x + \frac{1}{2}U$, $|\lambda y_i - \lambda y_j| = |\lambda(y_i - y_j)|$, using the linearity. By the symmetry $U = -U$, since $y_i - y_j \in \frac{1}{2} \cdot 2U = U$, this gives $|\lambda y_i - \lambda y_j| < \varepsilon$. Then unambiguously define $\lambda x$ to be the limit of the $\lambda y_i$.

The *order* of a distribution $u$ is the smallest $k$ such that $u \in C^k(\mathbb{T})^*$. Since for the circle the space of all distributions is exactly this colimit, the order of a distribution is well-defined. [29]

Distributions as generalized *functions* should be *differentiable*, compatibly with the differentiation of *functions*. The idea is that differentiation of distributions should be compatible with *integration by parts* for distributions given by integration against $C^1$ functions. Assuming an integral on $\mathbb{T}$ as in the next section, for *functions* $f, g$, by integration by parts,

$$\int_{\mathbb{T}} f(x)\, g'(x)\, dx \;=\; -\int_{\mathbb{T}} f'(x)\, g(x)\, dx$$

with no boundary terms because $\mathbb{T}$ has empty boundary. Note the negative sign. Motivated by this, define the *distributional derivative* $u'$ of $u \in C^\infty(\mathbb{T})^*$ to be another distribution defined by

$$u'(f) \;=\; -u(f') \qquad\qquad (\text{for any } f \in C^\infty(\mathbb{T}))$$

The continuity of differentiation $\frac{d}{dx} : C^\infty(\mathbb{T}) \to C^\infty(\mathbb{T})$ assures that $u'$ is a distribution, since

$$u' = -(u \circ \frac{d}{dx}) \;:\; C^\infty(\mathbb{T}) \to \mathbb{C}$$

---

# 5. *Invariant integration, periodicization*

We an *(invariant) integral* on the circle $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. The main property required is *translation invariance*, meaning that, for a (for example) continuous function $f$ on $\mathbb{T}$,

$$\int_{\mathbb{T}} f(x+y)\, dx \;=\; \int_{\mathbb{T}} f(x)\, dx \qquad\qquad (\text{for all } y \in \mathbb{T})$$

This invariance is sufficient to prove that various important integrals *vanish*.

For example, let $\psi_m(x) = e^{imx}$. As an instance of an important idea, without explicit calculus-like computations,

**[5.1] Claim:** *(Cancellation Lemma)* For $m \neq n$, for *any* reasonable translation-invariant integral on $\mathbb{T}$

$$\int_{\mathbb{T}} \psi_m(x)\, \overline{\psi}_n(x)\, dx \;=\; 0$$

*Proof:* For $m \neq n$, the function $f(x) = \psi_m(x)\overline{\psi}_n(x)$ is a *non-trivial* (not identically 1) continuous group homomorphism $\mathbb{T} \to \mathbb{C}^\times$, meaning that there is $y \in \mathbb{T}$ such that $f(y) \neq 1$. The *change of variables* $x \to x+y$ in the integral does not change the overall value of the integral, so

$$\int_{\mathbb{T}} f(x)\, dx \;=\; \int_{\mathbb{T}} f(x+y)\, dx \;=\; \int_{\mathbb{T}} f(x)\cdot f(y)\, dx \;=\; f(y) \int_{\mathbb{T}} f(x)\, dx$$

[29] The *Riesz representation theorem* asserts that the dual of $C^o(\mathbb{T})$ is *Borel measures* on $\mathbb{T}$, so order-zero distributions are *Borel measures*. For example, elements $\eta$ of $L^2(\mathbb{T})$ are Borel measures, by giving integrals $f \to \int_{\mathbb{T}} f(x)\,\eta(x)\,dx$ for $f \in C^o(\mathbb{T})$. Thus, integrating continuous functions against Borel measures is a semi-classical instance of *generalizing functions* in our present style, integrating against *measures*. However, the duals of the higher $C^k(\mathbb{T})$'s don't have such a classical interpretation. The fact that $C^1(\mathbb{T})$ can be construed as *distributional derivatives of Borel measures* is not strongly related to Radon-Nikodym derivatives of measures, because, for example, the distributional derivative of a point-mass measure is not a measure.

Thus, the integral $I$ has the property that $I = t \cdot I$ where $t \neq 1$. This gives $(1-t) \cdot I = 0$, so $I = 0$ since $t \neq 1$. ///

**[5.2] Remark:** This *vanishing trick* is impressive, since nothing specific about the continuous group homomorphism $f$ or topological group ($\mathbb{T}$ here) is used, apart from the finiteness of the total measure of the group, which comes from its *compactness*. That is, the same proof would show that *integrals over compact groups of non-trivial group homomorphisms are* 0. However, a notion of *invariant measure*[30] for general groups requires effort. Nevertheless, with an invariant measure, the same argument succeeds.

Less critically than the invariance, we want a *normalization*[31]

$$\int_{\mathbb{T}} 1\, dx \;=\; \mathrm{vol}\,(\mathbb{T}) \;=\; \mathrm{vol}\,(\mathbb{R}/2\pi\mathbb{Z}) \;=\; 2\pi$$

Then

$$\int_{\mathbb{T}} |\psi_n(x)|^2\, dx \;=\; \int_{\mathbb{T}} 1\, dx \;=\; 2\pi$$

Thus, without any explicit presentation of the integral or measure, we have proven that the distinct exponentials are an *orthogonal set* with norms $\sqrt{2\pi}$ with respect to the inner product

$$\langle f, g\rangle \;=\; \int_{\mathbb{T}} f(x)\, \overline{g}(x)\, dx$$

An *integration by parts* formula should be expected, with no *boundary terms* since $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ has empty boundary. Indeed, without *constructing* the invariant integral, we prove what we want from its *properties*:

**[5.3] Claim:** Let $f \to \int_{\mathbb{T}} f(x)\, dx$ be an invariant integral on $\mathbb{T}$, for $f \in C^o(\mathbb{T})$. Then for $f \in C^1(\mathbb{T})$

$$\int_{\mathbb{T}} f'(x)\, dx \;=\; 0$$

and we have the *integration by parts* formula for $f, g \in C^1(\mathbb{T})$

$$\int_{\mathbb{T}} f(x)\, g'(x)\, dx \;=\; -\int_{\mathbb{T}} f(x)'\, g(x)\, dx$$

**[5.4] Remark:** Vanishing of integrals of derivatives does *not* depend on the particulars of the situation. The same argument succeeds on an arbitrary group possessing (translation) invariant differentiation(s) and an invariant integral. Thus, the specific geometry of the circle is *not* needed to argue that $\int_{\mathbb{T}} f'(x)dx = \int_0^{2\pi} f(x)\, dx = f(2\pi) - f(0) = 0$ because $f$ is periodic. The latter classical argument is valid, but fails to show a generally applicable mechanism. The same independence of particulars applies to the integration by parts rule.

*Proof:* The translation invariance of the integral makes the integral of a derivative 0, by direct computation, as follows. We interchange a *differentiation* and an *integral.* [32]

$$\int_{\mathbb{T}} f'(x)\, dx \;=\; \int_{\mathbb{T}} \frac{\partial}{\partial t}\Big|_{t=0}\; f(x+t)\, dx \;=\; \frac{d}{dt}\Big|_{t=0} \int_{\mathbb{T}} f(x+t)\, dx \;=\; \frac{d}{dt}\Big|_{t=0} \int_{\mathbb{T}} f(x)\, dx \;=\; 0$$

---

[30] Translation-invariant measures on topological groups are called *Haar measures*. General proof of their *existence* takes a little work, and invokes the Riesz representation theorem. *Uniqueness* can be made to be an example of a more general argument about uniqueness of invariant functionals.

[31] The measure of the circle need not be normalized to be $2\pi$, but this is natural when presenting it as $\mathbb{R}/2\pi\mathbb{Z}$.

[32] The argument bluntly demands this interchange of limit and differentiation, so *justification* of it is secondary to the act itself. In the near future this and many other necessary interchanges are definitively justified via *Gelfand-Pettis* (also called *weak*) integrals. In the present concrete situation elementary (but opaque) arguments could be invoked, but we do not do this.

by changing variables in the integral. Then apply this to the function $(f \cdot g)' = f'g + fg'$ to obtain

$$\int_{\mathbb{T}} f'(x)\, g(x)\, dx + \int_{\mathbb{T}} f(x)\, g'(x)\, dx \;=\; 0$$

which gives the integration by parts formula. ///

The usual (Lebesgue) integral on the uniformizing $\mathbb{R}$ has the corresponding property of translation invariance. Since we present the circle as a quotient $\mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z} = \mathbb{T}$ of $\mathbb{R}$ we expect a *compatibility*[33]

$$\int_{\mathbb{R}} F(x)\, dx \;=\; \int_{\mathbb{R}/2\pi\mathbb{Z}} \left( \sum_{n \in \mathbb{Z}} F(x + 2\pi n) \right) dx$$

for at least *compactly-supported* continuous functions $F$ on $\mathbb{R}$.

Indeed, we can *define* integrals of functions on $\mathbb{T}$ by this compatibility relation, by expressing a continuous function $f$ on $\mathbb{T}$ as a *periodicization* (or *automorphization*)

$$f(x) \;=\; \sum_{n \in \mathbb{Z}} F(x + 2\pi n)$$

of a compactly supported continuous function $F$ on $\mathbb{R}$, and *define*

$$\int_{\mathbb{T}} f(x)\, dx \;=\; \int_{\mathbb{R}} F(x)\, dx$$

We still need to prove that this value is independent of the choice of $F$ for given $f$.

The properties required of an integral on $\mathbb{T}$ are clear. Sadly, we are not in a good position (yet) either to prove *uniqueness* or to give a *construction* as gracefully as these ideas deserve.

Postponing a systematic approach, we neglect any proof of uniqueness, and for a construction revert to an ugly-but-tangible reduction of the problem to integration on an interval. That is, note that in the quotient $q : \mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z} = \mathbb{T}$ the interval $[0, 2\pi]$ maps surjectively, with the endpoints being identified (and no other points identified). In traditional terminology, $[0, 2\pi]$ is a *fundamental domain*[34] for the action of $\mathbb{Z}$ on $\mathbb{R}$. Then define the integral of $f$ on $\mathbb{T}$ by

$$\int_{\mathbb{T}} f(x)\, dx \;=\; \int_0^{2\pi} (f \circ q)(x)\, dx$$

with usual (Lebesgue) measure on the unit interval. Verification of the *compatibility* with integration on $\mathbb{R}$ is silly, from this viewpoint.

---

[33] In contrast to many sources, this compatibility is *not* about choosing representatives in $[0, 2\pi)$ or anywhere else for $\backslash T$. Rather, this compatibility would be required for a group $G$ (here $\mathbb{R}$), a discrete subgroup $\Gamma$ (here $2\pi\mathbb{Z}$), and the quotient $G/\Gamma$ (here $\mathbb{T}$), whether or not that quotient is otherwise identifiable. This compatibility is a sort of *Fubini theorem*. The usual Fubini theorem applies to products $X \times Y$, whose quotients $(X \times Y)/X \approx Y$ are simply the factors, but another version applies to quotients that are not necessarily factors.

[34] The notion of *fundamental domain* for the action of a group $\Gamma$ on a set $X$ has an obvious appeal, at least that it is more concrete than the notion of *quotient* $\Gamma\backslash X$. However, it is rarely possible to determine an exact fundamental domain, *and* one eventually discovers that the details are seldom useful even if this is possible. Instead, the *quotient* should be treated directly.

This (bad) definition does allow explicit computations, but makes *translation invariance* harder to prove, since the unit interval gets pushed off itself by translation. But we can still manage the verification. [35] Take $y \in \mathbb{R}$, and compute

$$\int_{\mathbb{T}} f(x+y)\,dx \;=\; \int_0^{2\pi} (f\circ q)(x+y)\,dx \;=\; \int_{-y}^{2\pi-y} (f\circ q)(x)\,dx$$

$$= \int_{-y}^{0} (f\circ q)(x)\,dx + \int_0^{2\pi-y} (f\circ q)(x)\,dx \;=\; \int_{-y}^{0} (f\circ q)(x-2\pi)\,dx + \int_0^{2\pi-y} (f\circ q)(x)\,dx$$

since $(f\circ q)(x) = (f\circ q)(x-2\pi)$ by periodicity. Then, replacing $x$ by $x+2\pi$ in the first integral, this is

$$\int_{2\pi-y}^{2\pi} (f\circ q)(x)\,dx + \int_0^{2\pi-y} (f\circ q)(x)\,dx \;=\; \int_0^{2\pi} (f\circ q)(x)\,dx$$

---

# 6. *Levi-Sobolev inequalities, Levi-Sobolev imbeddings*

The simplest $L^2$ theory of Fourier series addresses neither continuity nor differentiability. [36] Yet it *would be* advantageous on general principles to be able to talk about differentiability in the context of Hilbert spaces, since Hilbert spaces have easily understood *dual spaces*. Beppo Levi, Frobenius, and Sobolev made useful comparisons. The idea is to compare $C^k$ norms to norms coming from Hilbert spaces whose inner products refer to derivatives, the Levi-Sobolev spaces.

## [6.1] Levi-Sobolev inequalities

First, we have an easy estimate for a variant $C^k$ norm:

$$\Big|\sum_{|n|\le N} c_n\, e^{inx}\Big|_{C^k} \;=\; \sup_{0\le j\le k}\,\sup_x \Big|\sum_{|n|\le N} c_n\,(in)^j\, e^{inx}\Big|_{\mathbb{C}} \;\le\; \sum_{|n|\le N} |c_n|\cdot(1+n^2)^{k/2}$$

all for elementary reasons. [37] Perhaps surprisingly, rather try to directly obtain a sup norm estimate on this sum, Cauchy-Schwarz-Bunyakowsky is invoked: for any $s \in \mathbb{R}$

$$\Big|\sum_{|n|\le N} c_n\, e^{inx}\Big|_{C^k} \;\le\; \sum_{|n|\le N} |c_n|\cdot(1+n^2)^{s/2}\cdot\frac{1}{(1+n^2)^{(s-k)/2}}$$

$$\le\; \Big(\sum_{|n|\le N} |c_n|^2\cdot(1+n^2)^s\Big)^{1/2}\cdot\Big(\sum_{|n|\le N}\frac{1}{(1+n^2)^{s-k}}\Big)^{1/2}$$

---

[35] While suppressing our disgust.

[36] It was not until the mid-20th century that L. Carleson showed, in L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135-157, that Fourier series of $L^2$ functions do converge pointwise almost everywhere. But this is a fragile sort of result.

[37] The awkward expression $(1+n^2)^{1/2}$ is approximately $n$. However, for $n=0$ we cannot divide by $n$, and replacing $n$ by $(1+n^2)^{1/2}$ is the traditional device stunt to avoid this annoyance.

Convergence of the elementary sum is easy to understand:

$$\sum_{n\in\mathbb{Z}} \frac{1}{(1+n^2)^{s-k}} \;<\; +\infty \qquad\qquad (\text{for } s > k + \tfrac{1}{2})$$

Thus, for any $s > k + \frac{1}{2}$ we have a *Levi-Sobolev inequality*

$$\Big|\sum_{|n|\le N} c_n\,\psi_n\Big|_{C^k} \;\le\; \Big(\sum_{n\in\mathbb{Z}} \frac{1}{(1+n^2)^{s-k}}\Big)^{1/2} \cdot \Big(\sum_{|n|\le N} |c_n|^2\cdot(1+n^2)^s\Big)^{1/2}$$

$$\le\; \Big(\sum_{n\in\mathbb{Z}} \frac{1}{(1+n^2)^{s-k}}\Big)^{1/2} \cdot \Big(\sum_{n\in\mathbb{Z}} |c_n|^2\cdot(1+n^2)^s\Big)^{1/2}$$

which is summarized as

$$\Big|\sum_{n\in\mathbb{Z}} c_n\,\psi_n\Big|_{C^k} \;\le\; \Big(\sum_{n\in\mathbb{Z}} \frac{1}{(1+n^2)^{s-k}}\Big)^{1/2} \cdot \Big(\sum_{n\in\mathbb{Z}} |c_n|^2\cdot(1+n^2)^s\Big)^{1/2} \qquad (\text{for } s > k+\tfrac{1}{2})$$

*Existence of this comparison makes the right side interesting.* Taking away from the right-hand side the uniform constant

$$\omega_{s-k} \;=\; \Big(\sum_{n\in\mathbb{Z}} \frac{1}{(1+n^2)^{s-k}}\Big)^{1/2}$$

gives the $s^{th}$ *Levi-Sobolev norm*

$$s^{th}\text{ Levi-Sobolev norm} \;=\; \Big|\sum_{n\in\mathbb{Z}} c_n\,\psi_n\Big|_{H^s} \;=\; \Big(\sum_{n\in\mathbb{Z}} |c_n|^2\cdot(1+n^2)^s\Big)^{1/2}$$

Paraphrasing, we have proven the dominance relation

$$|\;|_{C^k} \;\le\; \omega_{s-k}\cdot|\;|_{H^s} \qquad\qquad (\text{for any } s > k+\tfrac{1}{2})$$

## [6.2] Levi-Sobolev imbeddings

For $s \ge 0$, the $s^{th}$ *Levi-Sobolev space* is[38]

$$H^s(\mathbb{T}) \;=\; \{f\in L^2(\mathbb{T}) : \sum_n |\widehat{f}(n)|^2\cdot(1+n^2)^s < +\infty\}$$

The inner product on $H^s(\mathbb{T})$ is

$$\Big\langle \sum_n a_n\,\psi_n,\; \sum_n b_n\,\psi_n \Big\rangle \;=\; 2\pi\sum_n a_n\,\overline{b}_n\,(1+n^2)^s$$

[38] This definition is fine for $s \ge 0$, but not sufficient for $s < 0$. We will give the broader definition below. Keep in mind that $L^2(\mathbb{T})$ contains $C^o(\mathbb{T})$ and all the $C^k(\mathbb{T})$'s.

**[6.3] Remark:** This definition of $H^s(\mathbb{T})$ defines a useful space of functions or generalized functions only for $s \ge 0$, since for $s < 0$ the constraint $f \in L^2(\mathbb{T})$ is *stronger* (from the Plancherel theorem) than the condition defining $H^s(\mathbb{T})$ in the previous display.

**[6.4] Remark:** The $0^{th}$ Levi-Sobolev space is just $L^2(\mathbb{T})$.

**[6.5] Corollary:** For $s > k + \frac{1}{2}$ there is a continuous inclusion

$$H^s(\mathbb{T}) \subset C^k(\mathbb{T})$$

*Proof:* For $s > k + \frac{1}{2}$, whenever a Fourier series has a finite Levi-Sobolev norm

$$\Big|\sum_n c_n \, \psi_n\Big|_{H^s} = \left(\sum_{n\in\mathbb{Z}} |c_n|^2 \cdot (1+n^2)^s\right)^{1/2} < +\infty$$

the partial sums of the Fourier series are Cauchy in $H^s$, hence Cauchy in $C^k$, so converge in the Banach space $C^k$:

$$\sum_n c_n \, \psi_n = C^k \text{ function on } \mathbb{T}$$

*Proof:* Apply the Levi-Sobolev inequality $|f|_{C^k} \le \omega \cdot |f|_{H^s}$ to finite linear combinations $f$ of exponentials. Such finite linear combinations are $C^k$, and the inequality implies that an infinite sum of such, convergent in $H^s(\mathbb{T})$, has sequence of partial sums convergent in $C^k(\mathbb{T})$. That is, by the completeness of $C^k(\mathbb{T})$, the limit is still $k$ times continuously differentiable. Thus, we have the containment. Given the containment, the inequality of norms implies the continuity of the inclusion. ///

## [6.6] Levi-Sobolev Hilbert spaces

**[6.7] Claim:** The $s^{th}$ Levi-Sobolev space $H^s(\mathbb{T})$ (with $0 \le s \in \mathbb{R}$) is a Hilbert space. In particular, the sequences of Fourier coefficients of functions in $H^s(\mathbb{T})$ are *all* two-sided sequences $\{c_n : n \in \mathbb{Z}\}$ of complex numbers meeting the condition

$$\sum_n |c_n|^2 \cdot (1+n^2)^s < +\infty$$

**[6.8] Remark:** It is clear that the exponentials $\psi_n$ are an *orthogonal* basis for $H^s(\mathbb{T})$, although their norms depend on the index $s$. In particular, the collection of finite linear combinations of exponentials is *dense* in $H^s(\mathbb{T})$.

**[6.9] Remark:** Again, we do want to *define* these positively-indexed Levi-Sobolev spaces as subspaces of genuine spaces of functions, *not* as sequences of Fourier coefficients meeting the condition, and then *prove* the second assertion of the claim. This does leave open, for the moment, the question of how to define negatively-indexed Levi-Sobolev spaces.

*Proof:* In effect, this is the space of $L^2$ functions on which the $H^s$-norm is finite. If we prove the second assertion of the claim, then invoke the usual proof that $L^2$ spaces are complete to know that $H^s(\mathbb{T})$ is complete, since it is simply a weighted $L^2$-space. Given a two-sided sequence $\{c_n\}$ of complex numbers such that

$$\sum_n |c_n|^2 \cdot (1+n^2)^s < +\infty$$

since $s \ge 0$,

$$\sum_n |c_n|^2 < +\infty$$

and, by Plancherel,

$$\sum_n c_n \, \psi_n \;\in\; L^2(\mathbb{T})$$

This shows that $H^s(\mathbb{T})$ is a Hilbert space for $s \geq 0$. ///

**[6.10] Remark:** Insisting on viewing $L^2(\mathbb{T})$ as *equivalence classes* of functions may mislead us into making a needlessly complicated assertion about Levi-Sobolev imbeddings $H^s(\mathbb{T}) \subset C^k(\mathbb{T})$ for $s > k + \frac{1}{2}$, by insisting that $H^s(\mathbb{T})$ consists of almost-everywhere equivalence classes of $L^2(\mathbb{T})$ functions, only *one* of which is in $C^k(\mathbb{T})$. This is not a genuine issue.

## [6.11] Levi-Sobolev norms in terms of derivatives

**[6.12] Remark:** Apart from having the virtue of giving inner-product structures, the expressions appearing in these Levi-Sobolev norms are *natural* because they have meaning in terms of $L^2$-norms of derivatives. For $f = \sum c_n \, \psi_n \in C^k(\mathbb{T})$, by Plancherel

$$\text{(norm via derivatives)} \;=\; |f|^2 + |f'|^2 + |f''|^2 + \ldots + |f^{(k)}|^2$$

$$=\; \sum_n |c_n|^2 \cdot (1 + n^2 + n^4 + \ldots + n^{2k}) \;\leq\; \sum_n |c_n|^2 \cdot (1 + n^2)^k$$

Conversely,

$$(1+n^2)^k \;\leq\; C_k \cdot (1 + n^2 + n^4 + n^6 + \ldots + n^{2k}) \qquad\qquad \text{(for some constant } C_k\text{)}$$

so

$$\text{(norm via Fourier coefficients)} \;=\; \sum_n |c_n|^2 \cdot (1+n^2)^k \;\leq\; C_k \cdot \Big( |f|^2 + |f'|^2 + |f''|^2 + \ldots + |f^{(k)}|^2 \Big)$$

Thus, the two definitions of Levi-Sobolev norms, in terms of weighted $L^2$ norms of Fourier series, or in terms of $L^2$ norms of derivatives, give comparable Hilbert space structures. In particular, the *topologies* are identical.

**[6.13] Corollary:** For $k \geq 0$,

$$C^k(\mathbb{T}) \;\subset\; H^k(\mathbb{T})$$

*Proof:* For $k = 0$, the assertion is that $C^o(\mathbb{T}) \subset L^2(\mathbb{T})$, which holds because $\mathbb{T}$ is compact. Similarly, the relevant derivatives of $f \in C^k(\mathbb{T})$ are in $L^2(\mathbb{T})$, so $f \in H^k(\mathbb{T})$. ///

**[6.14] Remark:** One can work out the corresponding inequalities for Fourier series in several variables, proving that $(k + \frac{n}{2} + \varepsilon)$-fold $L^2$ differentiability (for any $\varepsilon > 0$) in dimension $n$ is needed to assure $k$-fold continuous differentiability. This is $L^2$ Levi-Sobolev theory.

## [6.15] Uniform pointwise convergence, convergence in $C^k(\mathbb{T})$

At this moment it is very easy to give a *straightforward*, if not *sharp*, result about convergence of $C^k$ functions on $\mathbb{T}$, via the Levi-Sobolev spaces:

**[6.16] Corollary:** The Fourier series of $f \in C^k(\mathbb{T})$ converges to $f$ in $C^{k-1}(\mathbb{T})$.

*Proof:* A function in $C^k(\mathbb{T})$ is in the Hilbert space $H^k(\mathbb{T})$, meaning that the finite partial sums of the Fourier expansion converge to $f$ in $H^k(\mathbb{T})$. The $H^k(\mathbb{T})$ norm dominates that of $C^{k-1}(\mathbb{T})$, so the Fourier series converges to $f$ in $C^{k-1}(\mathbb{T})$. ///

**[6.17] Remark:** It may seem mildly peculiar that the Fourier series of a $C^k$ function can converge to it only in $C^{k-1}$.

### [6.18] $L^2$-differentiation

**[6.19] Claim:** For every $s \geq 0$, the differentiation map

$$\frac{d}{dx} \;:\; \text{finite Fourier series} \;\longrightarrow\; \text{finite Fourier series}$$

is continuous when the source is given the $H^s(\mathbb{T})$ topology and the target is given the $H^{s-1}(\mathbb{T})$ topology.

*Proof:* This continuity is by design:

$$\Big|\frac{d}{dx}\sum_{|n|\le N} c_n\, e^{inx}\Big|^2_{H^{s-1}} \;=\; \Big|\sum_{|n|\le N} c_n\, in\, e^{inx}\Big|^2_{H^{s-1}} \;\le\; \sum_{|n|\le N} |nc_n|^2\cdot(1+n^2)^{s-1}$$

$$\le\; \sum_{|n|\le N} |c_n|^2\cdot(1+n^2)^{s} \;=\; \Big|\sum_{|n|\le N} c_n\, e^{inx}\Big|^2_{H^s}$$

proving the continuity on finite Fourier series. ///

Therefore, we can *extend* $\frac{d}{dx}$ *by continuity* to obtain continuous linear maps

$$(L^2\text{-differentiation}) = (\text{extension by continuity of})\;\frac{d}{dx} \quad:\quad H^s(\mathbb{T}) \;\longrightarrow\; H^{s-1}(\mathbb{T})$$

**[6.20] Remark:** In these terms, *extra* $L^2$-differentiability is needed to assure comparable classical *continuous* differentiability. Specifically, $(k+\frac{1}{2}+\varepsilon)$-fold $L^2$-differentiability (for any $\varepsilon > 0$) suffices for $k$-fold *continuous* differentiability, in this one-dimensional example. *The comparable computations on* $(\mathbb{T})^{\times n}$ *show that the gap widens as the dimension grows.*

---

# 7. $C^\infty = \lim C^k = \lim H^s = H^\infty$

For larger purposes, the specific comparisons of indices in the containments

$$H^s(\mathbb{T}) \quad\subset\quad C^k(\mathbb{T}) \quad (\text{for } s > k+\tfrac{1}{2})$$

$$C^k(\mathbb{T}) \quad\subset\quad H^s(\mathbb{T}) \quad (\text{for } k \ge s)$$

are secondary, since we are more interested in *smooth functions* $C^\infty(\mathbb{T})$ than functions with *limited* continuous differentiability.

Thus, the point is that the Levi-Sobolev spaces and $C^k(\mathbb{T})$ spaces are *cofinal* under taking *descending intersections*. That is, letting $H^\infty(\mathbb{T})$ be the *intersection* of all the $H^s(\mathbb{T})$, as *sets* we have

$$C^\infty(\mathbb{T}) \;=\; \bigcap_k C^k(\mathbb{T}) \;=\; \bigcap_{s\ge 0} H^s(\mathbb{T}) \;=\; H^\infty(\mathbb{T})$$

Since descending nested intersections are *limits*, the topologies behave well for trivial reasons:

**[7.1] Theorem:** As topological vector spaces

$$C^\infty(\mathbb{T}) \;=\; \lim_k C^k(\mathbb{T}) \;=\; \lim_{s\ge 0} H^s(\mathbb{T}) \;=\; H^\infty(\mathbb{T})$$

*Proof:* The cofinality of the $C^k$'s and the $H^s$'s gives a natural isomorphism of the two limits, since they can be combined in a larger limit in which each is cofinal. ///

Again, in general duals of limits are not colimits, but we did show earlier that the dual of a limit of *Banach* spaces is the colimit of the duals of the Banach spaces. Thus,

**[7.2] Corollary:** The space of distributions on $\mathbb{T}$ is

$$C^\infty(\mathbb{T})^* \;=\; \mathrm{colim}_k C^k(\mathbb{T})^* \;=\; \mathrm{colim}_{s\geq 0} H^s(\mathbb{T})^* \;=\; H^\infty(\mathbb{T})^*$$

(and the duals $H^s(\mathbb{T})^*$ admit further explication, below). ///

Expressing $C^\infty(\mathbb{T})$ as a limit of the Hilbert spaces $H^s(\mathbb{T})$, as opposed to its more natural expression as a limit of the Banach spaces $C^k(\mathbb{T})$, is convenient when taking *duals*, since by the *Riesz-Fischer theorem* [39] we have explicit expressions for Hilbert space duals. We exploit this possibility below.

---

# 8. *Distributions, generalized functions, again*

We will see that distributions on $\mathbb{T}$ have Fourier expansions, greatly facilitating their study. [40]

The exponential functions $\psi_n$ are in $C^\infty(\mathbb{T})$, so for any distribution $u$ we can compute Fourier coefficients of $u$ by

$$(n^{th} \text{ Fourier coefficient of } u) \;=\; \widehat{u}(n) \;=\; \frac{1}{2\pi}\cdot u(\psi_{-n})$$

Write

$$u \;\sim\; \sum_n \widehat{u}(n)\cdot\psi_n$$

even though *pointwise* convergence of the indicated sum is certainly not expected. Define Levi-Sobolev spaces for *all* $s\in\mathbb{R}$ by

$$H^s(\mathbb{T}) \;=\; \{u\in C^\infty(\mathbb{T})^* \;:\; \sum_n |u(\psi_{-n})|^2\cdot(1+n^2)^s<\infty\}$$

and the $s^{th}$ *Levi-Sobolev norm* $|u|_{H^s}$ is

$$|u|^2_{H^s} \;=\; \sum_n |u(\psi_{-n})|^2\cdot(1+n^2)^s$$

For $0\leq s\in\mathbb{Z}$, this definition is visibly compatible with the previous definition via derivatives.

**[8.1] Remark:** The formation of the Levi-Sobolev spaces of both positive and negative indices portrays the classical *functions* of various degrees of (continuous) differentiability together with *distributions* of various orders as fitting together as comparable objects. By contrast, thinking only in terms of the spaces $C^k(\mathbb{T})$ does not immediately suggest a comparison with distributions.

---

[39] The Riesz-Fischer theorem asserts that the (continuous) dual $V^*$ of a Hilbert space $V$ is $\mathbb{C}$-conjugate linearly isomorphic to $V$. The isomorphism from $V$ to $V^*$ attaches the linear functional $v\to\langle v,w\rangle$ to an element $w\in V$. Since our hermitian inner products $\langle,\rangle$ are *conjugate*-linear in the second argument, the map $w\to\langle\,,w\rangle$ is conjugate linear.

[40] In contrast, discussion of distributions on the real line $\mathbb{R}$ is more complicated, due to the non-compactness of $\mathbb{R}$. *Not* every distribution on $\mathbb{R}$ is the Fourier transform of a *function*. Distributions which *admit* Fourier transforms, *tempered* distributions, constitue a proper subset of all distributions on $\mathbb{R}$.

For convenience, define a *weighted* version $\ell^{2,s}$ of (a two-sided version of) the classical Hilbert space $\ell^2$ by

$$\ell^{2,s} \;=\; \{\{c_n : n\in\mathbb{Z}\} \;:\; \sum_{n\in\mathbb{Z}} |c_n|^2\cdot(1+n^2)^s < \infty\}$$

with the weighted version of the usual hermitian inner product, namely,

$$\langle\{c_n\},\,\{d_n\}\rangle \;=\; \sum_{n\in\mathbb{Z}} c_n\,\overline{d_n}\cdot(1+n^2)^s$$

**[8.2] Claim:** The complex bilinear *pairing*

$$\langle,\rangle \;:\; \ell^{2,s}\times\ell^{2,-s} \;\longrightarrow\; \mathbb{C}$$

by

$$\langle\{c_n\},\,\{d_n\}\rangle \;=\; \sum_n c_n\,d_{-n}$$

identifies these two Hilbert spaces as mutual duals, where

$$\ell^{2,-s} \;\longrightarrow\; (\ell^{2,s})^* \quad\text{by}\quad \{d_n\}\to\lambda_{\{d_n\}} \quad\text{where}\quad \lambda_{\{d_n\}}(\{c_n\}) = \sum_n c_n\,d_{-n}$$

**[8.3] Remark:** The minus sign in the subscript in the last formula is not the main point, but is a necessary artifact of our change from a *hermitian* form to a *complex bilinear* form. It is (thus) necessary to maintain compatibility with the Plancherel theorem for ordinary functions.

*Proof:* The Cauchy-Schwarz-Bunyakowsky inequality gives the continuity of the functional attached to $\{d_n\}$ in $\ell^{2,-s}$ by

$$\Big|\sum_n c_n\cdot d_{-n}\Big| \;\le\; \sum_n |c_n|\,(1+n^2)^{s/2}\cdot|d_{-n}|\,(1+n^2)^{-s/2}$$

$$\le\; \Big(\sum_n |c_n|^2\,(1+n^2)^s\Big)^{1/2}\cdot\Big(\sum_n |d_n|^2\,(1+n^2)^{-s}\Big)^{1/2} \;=\; |\{c_n\}|_{\ell^{2,s}}\cdot|\{d_n\}|_{\ell^{2,-s}}$$

proving the continuity. To prove the surjectivity we adapt the Riesz-Fischer theorem by a renormalization. That is, given a continuous linear functional $\lambda$ on $\ell^{2,s}$, by Riesz-Fischer there is $\{a_n\}\in\ell^{2,s}$ such that

$$\lambda(\{c_n\}) \;=\; \langle\{c_n\},\,\{a_n\}\rangle_{\ell^{2,s}} \;=\; \sum_n c_n\cdot\overline{a}_n\cdot(1+n^2)^s$$

Take

$$d_n \;=\; \overline{a}_{-n}\cdot(1+n^2)^s$$

Check that this sequence of complex numbers is in $\ell^{2,-s}$, by direct computation, using the fact that $\{a_n\}\in\ell^{2,s}$,

$$\sum_n |d_n|^2\cdot(1+n^2)^{-s} \;=\; \sum_n |\overline{a}_{-n}\cdot(1+n^2)^s|^2\cdot(1+n^2)^{-s} \;=\; \sum_n |a_n|^2\cdot(1+n^2)^s < +\infty$$

Thus, $\ell^{2,-s}$ is (isomorphic to) the dual of $\ell^{2,s}$. ///

**[8.4] Claim:** The map $u\to\{\widehat{u}(n)\}$ on $H^s(\mathbb{T})$ by taking Fourier coefficients is a Hilbert-space isomorphism

$$H^s(\mathbb{T}) \;\approx\; \ell^{2,s}$$

*Proof:* That the two-sided sequence of Fourier coefficients $u(\psi_{-n})$ is in $\ell^{2,s}$ is part of the definition of $H^s(\mathbb{T})$. The more serious question is *surjectivity*.

Let $\{c_n\} \in \ell^{2,s}$. For $s \geq 0$, the $s^{th}$ Levi-Sobolev norm dominates the $0^{th}$, so distributions in $H^s(\mathbb{T})$ are at least $L^2(\mathbb{T})$-functions. The definition of $H^s(\mathbb{T})$ in this case makes $H^s(\mathbb{T})$ a Hilbert space, and we directly invoke the Plancherel theorem, using the orthonormal basis $\frac{\psi_n}{\sqrt{2\pi}} \cdot (1+n^2)^{-s/2}$ for $H^s(\mathbb{T})$. This gives the surjectivity $H^s(\mathbb{T}) \to \ell^{2,s}$ for $s \geq 0$.

For $s < 0$, to prove the surjectivity, for $\{c_n\}$ in $\ell^{2,s}$ we will define a distribution $u$ lying in $H^s(\mathbb{T})$, by

$$u(f) \;=\; \sum_n \widehat{f}(n) \cdot c_{-n} \qquad\qquad (f \in C^\infty(\mathbb{T}))$$

By Cauchy-Schwarz-Bunyakowsky,

$$|\sum_n \widehat{f}(n) \cdot c_{-n}| \;\leq\; \sum_n |\widehat{f}(n)|\,(1+n^2)^{-s/2} \cdot |c_n|\,(1+n^2)^{s/2}$$

$$\leq\; \Big(\sum_n |\widehat{f}(n)|^2\,(1+n^2)^{-s}\Big)^{1/2} \cdot \Big(\sum_n |c_n|^2\,(1+n^2)^{s}\Big)^{1/2} \;=\; |f|_{H^{-s}} \cdot |\{c_n\}|_{\ell^{2,s}}$$

This shows that $u$ is a continuous linear functional on $H^{-s}(\mathbb{T})$. For $s < 0$, the test functions $C^\infty(\mathbb{T})$ imbed continuously into $H^{-s}(\mathbb{T})$, so $u$ gives a continuous functional on $C^\infty(\mathbb{T})$, so is a distribution. This proves that the Fourier coefficient map is a surjection to $\ell^{2,s}$ for $s < 0$. ///

**[8.5] Remark:** After this preparation, the remainder of this section is completely unsurprising. The following corollary is the conceptual point of this story.

**[8.6] Corollary:** For any $s \in \mathbb{R}$, the complex bilinear pairing

$$\langle , \rangle : H^s \times H^{-s} \to \mathbb{C} \quad \text{by} \quad f \times u \to \langle f, u\rangle = \sum_n \widehat{f}(n) \cdot \widehat{u}(-n)$$

gives an isomorphism

$$H^{-s} \;\approx\; (H^s)^*$$

by sending $u \in H^{-s}$ to $\lambda_u \in (H^s)^*$ defined by

$$\lambda_u(f) \;=\; \langle f, u\rangle \qquad\qquad (\text{for } f \in H^s(\mathbb{T}))$$

**[8.7] Remark:** The pairing of this last claim is *unsymmetrical*: the left argument is from $H^s$ while the right argument is from $H^{-s}$.

*Proof:* This pairing via Fourier coefficients is simply the composition of the maps $H^s(\mathbb{T}) \approx \ell^{2,s}$ and $H^{-s}(\mathbb{T}) \approx \ell^{2,-s}$ with the pairing of $\ell^{2,s}$ and $\ell^{2,-s}$ given just above. ///

**[8.8] Corollary:** The space of all distributions on $\mathbb{T}$ is

$$\text{distributions} \;=\; C^\infty(\mathbb{T})^* \;=\; \bigcup_{s\geq 0} H^s(\mathbb{T})^* \;=\; \bigcup_{s\geq 0} H^{-s}(\mathbb{T}) \;=\; \mathrm{colim}_{s\geq 0} H^{-s}(\mathbb{T})$$

thus expressing $C^\infty(\mathbb{T})^*$ as an ascending union of Hilbert spaces. ///

**[8.9] Corollary:** A distribution $u \sim \sum_n c_n \psi_n$ can be evaluated on $f \in C^\infty(\mathbb{T})$ by

$$u(f) \;=\; \sum_n \widehat{f}(n) \cdot \widehat{u}(-n)$$

*Proof:* Since $u$ lies in some $H^{-s}(\mathbb{T})$, it gives a continuous functional on $H^s(\mathbb{T})$, which contains $C^\infty(\mathbb{T})$. The Plancherel-like evaluation formula above gives the equality. ///

A collection of Fourier coefficients $\{c_n\}$ is *of moderate growth* when there is a constant $C$ and an exponent $N$ such that

$$|c_n| \;\le\; C \cdot (1+n^2)^N \qquad\qquad (\text{for all } n \in \mathbb{Z})$$

**[8.10] Corollary:** Let $\{c_n\}$ be a collection of complex numbers of moderate growth. Then there is a distribution $u$ with those as Fourier coefficients, that is, there is $u$ with

$$u(\psi_{-n}) \;=\; c_n$$

*Proof:* For constant $C$ and exponent $N$ such that $|c_n| \le C \cdot (1+n^2)^N$,

$$\sum_n |c_n|^2 \cdot (1+n^2)^{-(2N+1)} \;\le\; \sum_n C^2 \cdot (1+n^2)^{2N} \cdot (1+n^2)^{-(2N+1)} \;=\; C^2 \cdot \sum_n (1+n^2)^{-1} \;<\; \infty$$

That is, from the previous discussion, the sequence gives an element of $H^{-(2N+1)}(\mathbb{T}) \subset C^\infty(\mathbb{T})^*$. ///

**[8.11] Corollary:** For $u \sim \sum_n c_n \psi_n \in H^s(\mathbb{T})$ the derivative (for any $s \in \mathbb{R}$) is

$$u' \;\sim\; \sum_n in \cdot c_n \cdot \psi_n \;\in\; H^{s-1}$$

*Proof:* Invoke the definition (compatible with integration by parts) of the derivative of distributions, and integrating by parts to see that $\widehat{f'}(n) = in \cdot \widehat{f}(n)$ for $f \in C^\infty(\mathbb{T}) = H^\infty(\mathbb{T})$,

$$u'(f) \;=\; -u(f') \;=\; -\sum_n \widehat{f'}(n) \cdot \widehat{u}(-n) \;=\; -\sum_n in\, \widehat{f}(n) \cdot \widehat{u}(-n) \;=\; \sum_n \widehat{f}(n) \cdot -in\, \widehat{u}(-n)$$

as claimed. The Fourier coefficients $-in \cdot \widehat{u}(n)$ do satisfy

$$\sum_n |in\, \widehat{u}(n)|^2 \cdot (1+n^2)^{s-1} \;\le\; \sum_n (1+n^2)\, |\widehat{u}(n)|^2 \cdot (1+n^2)^{s-1} \;=\; \sum_n |\widehat{u}(n)|^2 \cdot (1+n^2)^s \;=\; |u|^2_{H^s} < \infty$$

which proves that the differentiation maps $H^s$ to $H^{s-1}$ continuously. ///

**[8.12] Remark:** In the latter proof the sign in the subscript in the definition of the pairing $\ell^{2,s} \times \ell^{2,-s}$ was essential.

**[8.13] Corollary:** The collection of finite linear combinations of exponentials $\psi_n$ is dense in every $H^s(\mathbb{T})$, for $s \in \mathbb{R}$. In particular, $C^\infty(\mathbb{T})$ is dense in every $H^s(\mathbb{T})$, for $s \in \mathbb{R}$.

*Proof:* The exponentials are an orthogonal basis for every Levi-Sobolev space. ///

**[8.14] Remark:** The topology of colimit of Hilbert spaces is the *finest* of several reasonable topologies on distributions. Density in a finer topology is a stronger assertion than density in a coarser topology.

# 9. *The provocative example explained*

The confusing example of the sawtooth function is clarified in the context we've developed. By now, we know that Fourier series whose coefficients satisfy sufficient decay conditions *are* differentiable. Even when the coefficients do not decay, but only grow *moderately*, the Fourier series is that of a *generalized function*. In other words, *we can (nearly) always differentiate* Fourier series term by term, as long as we can tolerate the outcome being a *generalized* function, rather than necessarily a *classical* function.

Again, $s(x)$ is the *sawtooth function*

$$s(x) \;=\; x - \pi \qquad\qquad (\text{for } 0 \le x < 2\pi)$$

made *periodic* by demanding $s(x + 2\pi n) = s(x)$ for all $n \in \mathbb{Z}$, so

$$s(x) \quad = \quad x - 2\pi \cdot [\![\frac{x}{2\pi}]\!] \;-\; \pi \qquad\qquad (\text{for } x \in \mathbb{R})$$

where $[\![x]\!]$ is the greatest integer less than or equal $x$. Away from $2\pi\mathbb{Z}$, this function is infinitely differentiable, with derivative 1. At integers it jumps down from value to $\pi$ to value $-\pi$. We do not attempt to define a value *at* $2\pi\mathbb{Z}$.

We want to *differentiate* this function compatibly with integration by parts, and compatibly with term-by-term differentiation of Fourier series.

The sawtooth function *is* well-enough behaved to give a *distribution* by integrating against it. Therefore, as we saw above, it *can* be differentiated as a distribution, and be correctly differentiated as (as a distribution) by differentiating its Fourier expansion termwise.

A earlier, Fourier coefficients are computed by integrating against $e^{-inx}$

$$\frac{1}{2\pi}\int_0^1 s(x)\cdot e^{-inx}\,dx \;=\; \begin{cases} \dfrac{1}{-in} & (\text{for } n \neq 0) \\ 0 & (\text{for } n = 0) \end{cases}$$

Thus, *at least* as a distribution, its Fourier expansion is

$$s(x) \;=\; i\sum_{n\neq 0} \frac{1}{n}\cdot e^{inx}$$

The series *does* converge pointwise to $s(x)$ for $x$ away from (images of) integers, as we proved happens at left and right differentiable points for piecewise $C^o$ functions.

We are entitled to differentiate, at worst within the class of distributions, within which we are assured of a reasonable sense to our computations. *Further*, we are entitled (for any distribution) to differentiate the Fourier series term-by-term. That is, as distributions,

$$\begin{aligned} s'(x) \quad &= \quad -\textstyle\sum_{n\neq 0}\, e^{inx} \\ s''(x) \quad &= \quad -\textstyle\sum_{n\neq 0}\, in\, e^{inx} \\ &\;\;\ldots \\ s^{(k)}(x) \quad &= \quad -\textstyle\sum_{n\neq 0}\, (in)^{k-1}\, e^{inx} \end{aligned}$$

and so on, just as successive derivatives of smooth functions $f(x) = \sum_n c_n\, e^{inx}$ are obtained by termwise differentiation

$$f^{(k)}(x) \;=\; \sum_{n\neq 0} (in)^k\, c_n\, e^{inx}$$

The difficulty of interpreting the right-hand side of the Fourier series for $s^{(k)}$ as having pointwise values is irrelevant.

More to the point, these Fourier series are things to integrate smooth functions against, by an extension of the Plancherel formula for inner products of $L^2$ functions. Namely, for any smooth function $f(x) \sim \sum_n c_n \, e^{inx}$, the imagined integral of $f$ against $s^{(k)}$ should be expressible as the sum of products of Fourier coefficients

$$\text{imagined } \langle f, s^{(k)} \rangle \;=\; \sum_{n\neq 0} c_n \cdot \Big(\frac{(in)^k}{-in}\Big)^{\text{conj}}$$

(where $\alpha \to \alpha^{\text{conj}}$ is complex conjugation) and the latter expression should behave well when rewritten in a form that refers to the literal function $s$. Indeed,

$$\sum_{n\neq 0} c_n \cdot \Big(\frac{(in)^k}{-in}\Big)^{\text{conj}} \;=\; (-1)^k \sum_{n\neq 0} (in)^k \, c_n \cdot \Big(\frac{1}{-in}\Big)^{\text{conj}} \;=\; (-1)^k \int_{\mathbb{T}} f^{(k)}(x)\, \overline{s}(x)\, dx$$

by the Plancherel theorem applied to the $L^2$ functions $f^{(k)}$ and $s$. Let $u$ be the distribution given by integration against $s$. Then, by the definition of differentiation of distributions, we have computed that

$$(-1)^k \int_{\mathbb{T}} f^{(k)}(x)\, \overline{s}(x)\, dx \;=\; (-1)^k u(f^{(k)}) \;=\; u^{(k)}(f)$$

It is in this sense that the sum $\sum_{n\neq 0} c_n \cdot \frac{(in)^k}{-in}$ is integration of $s$ against $f$.

Further, for $f$ a smooth function with support away from the discontinuities of $s$, it is true that $u''(f) = 0$, giving $s''$ a vague pointwise sense of being 0 away from the discontinuities of $s$. This was clear at the outset, but now is given precise meaning.

Thus, as claimed at the outset of the discussion of functions on the circle, we can differentiate $s(x)$ legitimately, and the differentiation of the Fourier series of the sawtooth function $s(x)$ correctly represents this differentiation, viewing $s(x)$ and its derivatives as *distributions*.

---

# 10. *Appendix: products and limits of topological vector spaces*

Here we carry out the diagrammatical proof that products and limits of topological vector spaces *exist*, and are locally convex when the factors or limitands are locally convex. Nothing surprising happens.

**[10.1] Claim:** Products and limits of topological vector spaces exist. In particular, limits are *closed* (linear) subspaces of the corresponding products. When the factors or limitands are locally convex, so is the product or limit.

**[10.2] Remark:** Part of the point is that products and limits of locally convex topological vector spaces *in the larger category of not-necessarily locally convex topological vector spaces* are nevertheless locally convex. That is, enlarging the category in which we take test objects does not change the outcome, in this case. By contrast, coproducts and colimits in general are sensitive to local convexity of the test objects. [41]

*Proof:* After we construct products, limits are constructed as closed subspaces of them.

[41] For example, uncountable coproducts do not exist among not-necessarily locally convex topological vector spaces, essentially because the not-locally-convex spaces $\ell^p$ with $0 < p < 1$ exist.

Let $V_i$ be topological vector spaces. We claim that the topological-space product $V = \Pi_i V_i$ (with projections $p_i$) (with the product topology) is a topological vector space product. Let $\alpha_i : V_i \times V_i \to V_i$ be the addition on $V_i$. The family of composites $\alpha_i \circ (p_i \times p_i) : V \times V \to V_i$ induces a map $\alpha : V \times V \to V$ as in

$$\begin{tikzcd} V \times V \arrow[r, dashed, "\alpha"] \arrow[d, "p_i \times p_i"'] & V \arrow[d, "p_i"] \\ V_i \times V_i \arrow[r, "\alpha_i"'] & V_i \end{tikzcd}$$

This defines what we will show to be a *vector addition* on $V$. Similarly, the scalar multiplications $s_i : \mathbb{C} \times V_i \to V_i$ composed with the projections $p_i : V \to V_i$ give a family of maps

$$s_i \circ (1 \times p_i) \ : \ \mathbb{C} \times V \ \longrightarrow \ V_i$$

which induce a map $s : \mathbb{C} \times V \to V$ which we will show to be a *scalar multiplication* on $V$. That these maps are *continuous* is given us by starting with the topological-space product.

That is, we must prove that vector addition is commutative and associative, that scalar multiplication is associative, and that the two have the usual distributivity. All these proofs are the same in form. For commutativity of vector addition, consider the diagram

$$\begin{tikzcd} & V_i \times V_i \arrow[rr, "v\times w \to v+w"] & & V_i \arrow[dd, equal] \\ V \times V \arrow[ur, "p_i \times p_i"] \arrow[rr, dashed, shift left, "v\times w \to v+w"] \arrow[rr, dashed, shift right, "v\times w \to w+v"'] \arrow[d, "v\times w \to w\times v"'] & & V \arrow[ur, "p_i"'] \arrow[dr, "p_i"] & \\ V \times V \arrow[r, "p_i \times p_i"] & V_i \times V_i \arrow[rr, "v\times w \to w+v"] & & V_i \end{tikzcd}$$

The upper half of the diagram is the induced-map definition of vector addition on $V$, and the lower half is the induced map definition of the reversed-order vector addition. The commutativity of addition on each $V_i$ implies that going around the top of the diagram from $V \times V$ to $V_i$ yields the same as going around the bottom. Thus, the two induced maps $V \times V \to V$ must be the same, since induced maps are *unique*.

The proofs of associativity of vector addition, associativity of scalar multiplication, and distributivity, use the same idea. Thus, *products* of topological vector spaces exist.

We should not forget to prove that the product is *Hausdorff*, since we implicitly require this of topological vector spaces. This is immediate, since a (topological space) product of Hausdorff spaces is readily shown to be Hausdorff.

Consider now the case that each $V_i$ is locally convex. By definition of the product topology, every neighborhood of 0 in the product is of the form $\Pi_i U_i$ where $U_i$ is a neighborhood of 0 in $V_i$, and all but finitely many of the $U_i$ are the whole $V_i$. Since $V_i$ is locally convex, we can shrink every $U_i$ that is *not* $V_i$ to be a convex open containing 0, while each *whole* $V_i$ is convex. Thus, the product is locally convex when every factor is.

To construct limits, reduce to the product.

**[10.3] Claim:** Let $V_i$ be topological vector spaces with transition maps $\varphi_i : V_i \to V_{i-1}$. The limit $V = \lim_i V_i$ *exists*, and, in particular, is the closed linear subspace (with subspace topology) of the product $\Pi_i V_i$ (with projections $p_i$) defined by the (closed) conditions

$$\lim_i V_i \ = \ \{v \in \Pi_i V_i : (\varphi_i \circ p_i)(v) = p_{i-1}(v), \ \text{ for all } i\}$$

*Proof:* *(of claim)* Constructing the alleged limit as a closed subspace of the product immediately yields the desired properties of vector addition and scalar multiplication, as well as the Hausdorff-ness. What we must show is that the construction does function as a limit.

Given a compatible family of continuous linear maps $f_i : Z \to V_i$, there is induced a unique continuous linear map $F : Z \to \Pi_i V_i$ to the product, such that $p_i \circ f = f_i$ for all $i$. The *compatibility* requirement on the $f_i$ exactly asserts that $f(Z)$ sits inside the subspace of $\Pi_i V_i$ defined by the conditions $(\varphi_i \circ p_i)(v) = p_{i-1}(v)$. Thus, $f$ maps to this subspace, as desired.

Further, for all limitands locally convex, we have shown that the product is locally convex. The local convexity of a linear subspace (such as the limit) follows immediately. ///

---

# 11. *Appendix: Fréchet spaces and limits of Banach spaces*

A larger class of topological vector spaces arising in practice is the class of *Fréchet spaces*. In the present context, we can give a nice definition: a *Fréchet space* is a *countable* limit of Banach spaces. [42] Thus, for example,

$$C^\infty(\mathbb{T}) \;=\; \bigcap_k C^k(\mathbb{T}) \;=\; \lim_k C^k(\mathbb{T})$$

is a Fréchet space, *by (this) definition.*

Despite its advantages, the present definition is not the usual one. [43] We make a comparison, and elaborate on the features of Fréchet spaces.

A *metric* $d(,)$ on a vector space $V$ is *invariant* (implicitly, under addition), when

$$d(x+z, y+z) \;=\; d(x,y) \qquad\qquad (\text{for all } x,y,z \in V)$$

*All metrics we'll care about on topological vector spaces will be invariant in this sense.*

**[11.1] Claim:** A Fréchet space is locally convex and complete (invariantly) metrizable. [44]

*Proof:* Let $V = \lim_i B_i$ be a countable limit of Banach spaces $B_i$, where $\varphi_i : B_i \to B_{i-1}$ are the transition maps and $p_i : V \to B_i$ are the projections. From the appendix, the limit is a closed linear subspace of the product, and the product is the cartesian product with the product topology and component-wise vector addition. Recall that a product of a *countable* collection of metric spaces is metrizable, and is complete if each factor is complete. A closed subspace of a complete metric space is complete metric. Thus, $\lim_i B_i$ is complete metric.

---

[42] Of course, it suffices that a limit have a countable cofinal subfamily.

[43] A common definition, with superficial appeal, is that a Fréchet space is a complete (invariantly) metrized space that is locally convex. This has the usual disadvantage that there are many different metrics that can give the same topology. This also ignores the manner in which Fréchet spaces usually arise, as countable limits of Banach spaces. There is another common definition that does halfway acknowledge the latter construction, namely, that a Fréchet space is a *complete* topological vector space with topology given by a countable collection of *seminorms*. The latter definition is essentially equivalent to ours, but requires explanation of the suitable notion of *completeness* in a not-necessarily metric situation, as well as explanation of the notion of *seminorm* and how topologies are specified by seminorms. We skirt the latter issues for the moment.

[44] As is necessary to prove the *equivalence* of the various definitions of *Fréchet space*, the converse of this claim is true, namely, that every locally convex and complete (invariantly) metrizable topological vector space is a countable limit of Banach spaces. Proof of the converse requires work, namely, development of ideas about seminorms. Since we don't need this converse at the moment, we do not give the argument.

As proven in the previous appendix, *any* product or limit of locally convex spaces is locally convex, whether or not it has a countable cofinal family. Thus, the limit is Fréchet. ///

Addressing the comparison between local convexity and limits of Banach spaces,

[11.2] Theorem: Every locally convex topological vector space is a *subspace* of a limit of Banach spaces (and vice-versa).

[11.3] Remark: This little theorem encapsulates the construction of *semi-norms* to give a locally convex topology. It can also be used to reduce the general Hahn-Banach theorem for locally convex spaces to the Hahn-Banach theorem for Banach spaces.

*Proof:* In one direction, we already know that a product or limit of Banach spaces is locally convex, since Banach spaces are locally convex.

In the Banach or normed-space situation, the topology comes from a metric $d(v, w) = |v - w|$ defined in terms of a *single* function $v \to |v|$ with the usual properties

$$|\alpha \cdot v| = |\alpha|_{\mathbb{C}} \cdot |v| \qquad \text{(homogeneity)}$$

$$|v + w| \le |v| + |w| \qquad \text{(triangle inequality)}$$

$$|v| \ge 0, \qquad \text{(equality } \textit{only} \text{ for } v = 0) \qquad \text{(definiteness)}$$

By contrast, for more general (but locally convex) situations, we consider a *family* $\Phi$ of functions $p(v)$ for which the definiteness condition is weakened slightly, so we require

$$p(\alpha \cdot v) = |\alpha|_{\mathbb{C}} \cdot p(v) \qquad \text{(homogeneity)}$$

$$p(v + w) \le p(v) + p(w) \qquad \text{(triangle inequality)}$$

$$p(v) \ge 0 \qquad \text{(semi-definiteness)}$$

Such a function $p()$ is a *semi-norm.* For Hausdorff-ness, we further require that the family $\Phi$ is *separating* in the sense that, given $v \ne 0$ in $V$, there is $p \in \Phi$ such that $p(v) > 0$.

A separating family $\Phi$ of semi-norms on a complex vector space $V$ gives a *locally convex* topology by taking as local sub-basis [45] at 0 the sets

$$U_{p,\varepsilon} \;=\; \{v \in V : \; p(v) < \varepsilon\} \qquad \text{(for } \varepsilon > 0 \text{ and } p \in \Phi)$$

Each of these is convex, because of the triangle inequality for the semi-norms.

[11.4] Remark: The topology obtained from a (separating) family of seminorms may appear to be a random or frivolous generalization of the notion of topology obtained from a *norm.* However, it is the correct extension to encompass *all* locally convex topological vector spaces, as we see now. [46]

For a locally convex topological vector space $V$, for every open $U$ in a local basis $B$ at 0 of *convex* opens, try to define a *seminorm*

$$p_U(v) \;=\; \inf\{t > 0 : \; t \cdot U \ni v\}$$

[45] Again, a *sub-basis* for a topology is a set of opens such that finite intersections form a *basis.* In other words, arbitrary unions of finite intersections give all opens.

[46] The semi-norms we construct here are sometimes called *Minkowski functionals*, even though they are not functionals in the sense of being continuous linear maps.

We discover some necessary adjustments, and then verify the semi-norm properties.

First, we show that for any $v \in V$ the set over which the inf is taken is non-empty. Since scalar multiplication $\mathbb{C} \times V \to V$ is (jointly!) continuous, for given $v \in V$, given a neighborhood $U$ of $0 \in V$, there are neighborhoods $W$ of $0 \in \mathbb{C}$ and $U'$ of $v$ such that

$$\alpha \cdot w \in U \quad \text{(for all } \alpha \in W \text{ and } w \in U')$$

In particular, since $W$ contains a disk $\{|\alpha| < \varepsilon\}$ for some $\varepsilon > 0$, we have $t \cdot v \in U$ for all $0 < t < \varepsilon$. That is,

$$v \;\in\; t \cdot U \qquad\qquad \text{(for all } t > \varepsilon^{-1})$$

Semi-definiteness of $p_U$ is built into the definition.

To avoid nagging problems, we should verify that, for convex $U$ containing 0, when $v \in t \cdot U$ then $v \in s \cdot U$ for all $s \geq t$. This follows from the convexity, by

$$s^{-1} \cdot v \;=\; \frac{t}{s} \cdot (t^{-1} \cdot v) \;=\; \frac{t}{s} \cdot (t^{-1} \cdot v) + \frac{s-t}{s} \cdot 0 \;\in\; U$$

since $t^{-1} \cdot v$ and 0 are in $U$.

The homogeneity condition $p(\alpha\, v) = |\alpha|\, p(v)$ already presents a minor issue, since convex sets containing 0 need have no special properties regarding multiplication by complex numbers. That is, the problem is that, given $v \in t \cdot U$, while $\alpha\, v \in \alpha \cdot t \cdot U$, we do *not* know that this implies $\alpha v \in |\alpha| \cdot t \cdot U$. Indeed, in general, it will not. To repair this, to make semi-norms we must use only convex opens $U$ which are *balanced* in the sense that

$$\alpha \cdot U \;=\; U \qquad\qquad \text{(for } \alpha \in \mathbb{C} \text{ with } |\alpha| = 1)$$

Then, given $v \in V$, we have $v \in t \cdot U$ if and only if $\alpha\, v \in t \cdot \alpha U$, and now

$$t\, \alpha U \;=\; t\, |\alpha|\, \frac{\alpha}{|\alpha|}\, U \;=\; t\, |\alpha|\, U$$

by the balanced-ness.

Now we have an obligation to show that there is a local basis (at 0) of convex *balanced* opens. Fortunately, this is easy to see, as follows. Given a convex $U$ containing 0, from the continuity of scalar multiplication, since $0 \cdot v = 0$, there is $\varepsilon > 0$ and a neighborhood $W$ of 0 such that $\alpha \cdot w \in U$ for $|\alpha| < \varepsilon$ and $w \in W$. Let

$$U' \;=\; \{\alpha \cdot w : \, |\alpha| \leq \frac{\varepsilon}{2},\; w \in W\} \;=\; \bigcup_{|\alpha| \leq \varepsilon/2} \alpha \cdot W$$

Being a union of the opens $\alpha \cdot W$, this $U'$ is open. It is inside $U$ by arrangement, and is *balanced* by construction. That is, there is indeed a local basis of convex *balanced* opens at 0.

For the *triangle inequality* for $p_U$, given $v, w \in V$, let $t_1, t_2$ be such that $v \in t \cdot U$ for $t \geq t_1$ and $w \in t \cdot U$ for $t \geq t_2$. Then, using the convexity,

$$v + w \;\in\; t_1 \cdot U + t_2 \cdot U \;=\; (t_1 + t_2) \cdot \left( \frac{t_1}{t_1 + t_2} \cdot U + \frac{t_2}{t_1 + t_2} \cdot U \right) \;\subset\; (t_1 + t_2) \cdot U$$

This gives the triangle inequality

$$p_U(v + w) \;\leq\; p_U(v) + p_U(w)$$

Finally, we check that the semi-norm topology is the original one. This is unsurprising. It suffices to check at 0. On one hand, given an open $W$ containing 0 in $V$, there is a convex, balanced open $U$ contained in $W$, and

$$\{v \in V : \, p_U(v) < 1\} \;\subset\; U \;\subset\; W$$

Thus, the semi-norm topology is at least as fine as the original topology. On the other hand, given convex balanced open $U$ containing 0, and given $\varepsilon > 0$,

$$\{v \in V : p_U(v) < \varepsilon\} \supset \frac{\varepsilon}{2} \cdot U$$

Thus, each sub-basis open for the semi-norm topology contains an open in the original topology. We conclude that the two topologies are the same.

A summary so far: for a locally convex topological vector space, the semi-norms attached to convex balanced neighborhoods of 0 give a topology identical to the original, and *vice-versa*.

Before completing the proof of the theorem, recall that a *completion* of a set with respect to a *pseudo-metric* can be defined much as the completion with respect to a genuine metric. This is relevant because a semi-norm may only give a pseudo-metric, not a genuine metric.

Let $\Phi$ be a (separating) family of seminorms on a vector space $V$. For a *finite subset* $i$ of $\Phi$, let $X_i$ be the *completion* of $V$ with respect to the semi-norm

$$p_i(v) = \sum_{p \in i} p(v)$$

with natural map $f_i : V \to X_i$. Order subsets of $\Phi$ by $i \geq j$ when $i \supset j$. For $i > j$ we have

$$p_i(v) = \sum_{p \in i} p(v) \geq \sum_{p \in j} p(v) = p_j(v)$$

so we have natural continuous (transition) maps

$$\varphi_{ij} : X_i \longrightarrow X_j \qquad \text{(for } i > j\text{)}$$

We claim that each $X_i$ is a *Banach space*, that $V$ with its semi-norm topology has a natural continuous *inclusion* to the limit $X = \lim_i X_i$, and that $V$ has the topology given by the subspace topology inherited from the limit.

The maps $f_i$ form a compatible family of maps to the $X_i$, so there is a unique compatible map $f : V \to X$. By the separating property, given $v \neq 0$, there is $p \in \Phi$ such that $p(v) \neq 0$. Then for all $i$ containing $p$, we have $f_i(v) \neq 0 \in X_i$. The subsets $i$ containing $p$ are *cofinal* in this limit, so $f(v) \neq 0$. Thus, $f$ is an inclusion.

Since the limit is a (closed) subspace of the *product* of the $X_i$, it suffices to prove that the topology on $V$ (imbedded in $\Pi_i X_i$ via $f$) is the subspace topology from $\Pi_i X_i$. Since the topology on $V$ is at *least* this fine (since $f$ is continuous), we need only show that the *subspace* topology is at least as fine as the semi-norm topology. To this end, consider a semi-norm-topology sub-basis set

$$\{v \in V : p_U(v) < \varepsilon\} \qquad \text{(for } \varepsilon > 0 \text{ and convex balanced open } U \text{ containing } 0\text{)}$$

This is simply the intersection of $f(V)$ with the sub-basis set

$$\Pi_{p \neq \{p_U\}} X_i \times \{v \in X_{\{p_U\}} : p_U(v) < \varepsilon\}$$

with the last factor inside $X_{\{p_U\}}$. Thus, by construction, the map $f : V \to X$ is a homeomorphism of $V$ to its image. ///

---