Analysis on Circles: A Modern View of Fourier Series

A Thesis
Presented to
The Division of Mathematics and Natural Sciences
Reed College

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Arts

William B. Rosenbaum

May 2009

Approved for the Division
(Mathematics)

Jerry Shurman

Acknowledgements

First and foremost, I would like to thank my thesis adviser, Jerry Shurman. This thesis would not be what it is without your guidance. You taught me that mathematical exposition can (and should) be both correct and enjoyable for the reader.

I would like to thank my family: Mom, Dad and Liz. Thank you for your constant support and encouragement.

Thank you Devin and Kevin for sharing my belief that the only thing more important than a thesis' content is its typesetting.

Thank you Joel Franklin for your enthusiastic encouragement and your discussions about the interplay of math and physics.

Also, thank you Paul Garrett. Your writing inspired this entire project.

Last, but certainly not least, I want to thank Alivia. You've kept me sane through this whole ordeal. Thank you for reminding me that there are more important things than writing a thesis.

Preface

Harmonic analysis is among the most successful and applicable branches of modern mathematics. It is an indispensable tool in subjects ranging from number theory to partial differential equations and numerical analysis. In this work, we present one of the simplest (and oldest) incarnations of harmonic analysis: Fourier series. Although the idea of Fourier series is not new, we present the subject using the vocabulary of contemporary mathematics.

This exposition is more abstract than many classical descriptions of Fourier series. As a result, the methods used in this text scale up to more general situations. The hope is that this presentation will acquaint the reader with the tools necessary to tackle generalizations of Fourier series such as analysis on spheres, tori and even p-adic analysis.

Of course, the cost of such a presentation is that it requires substantial background in analysis and topology. Throughout the text, the reader is presumed to have a basic understanding of real and functional analysis including point set topology, Banach and Hilbert spaces, Lebesgue integration on $\mathbb{R}$ and group theory. Some cursory knowledge of category theory would also be helpful.

Table of Contents

Abstract

The aim of this work is to develop analysis and distribution theory on the circle S^1. The work culminates with a description of the Sobolev spaces $H_s(S^1)$. Using Sobolev spaces, we show that every distribution on the circle can be represented by Fourier series. Throughout our exposition, we emphasize topology and characterization by mapping properties.

Introduction

Fourier series is a powerful tool in analysis. The theory of Fourier series addresses two basic questions:

1. Given a periodic function
$$f : \mathbb{R} \longrightarrow \mathbb{C}$$
can we represent f as a linear combination of the complex exponential functions $\psi_n(x) = \exp(2\pi i n x)$ for $n \in \mathbb{Z}$?

2. Given a series
$$\sigma = \sum_{n\in\mathbb{Z}} c_n \psi_n \quad \text{with} \quad c_n \in \mathbb{C}$$
to what function (if any) does σ converge?

The short answer to the first question is that we can represent *all* (non-pathological) periodic functions by Fourier series. The caveat is that we have to change our basic notion of what a function is. We will come to view functions as special cases of *distributions* (or *generalized functions*) in order to obtain this elegant result.

Once we have the language of distributions, the answer to the second question also becomes simple. A Fourier series

$$\sigma = \sum_{n\in\mathbb{Z}} c_n \psi_n \quad \text{with} \quad c_n \in \mathbb{C}$$

converges to a distribution if and only if the values $|c_n|$ grow no faster than some polynomial in n.

The answers to these two questions are not nearly so tidy in the context of classical functions.

We warm up with the following example. Let $f : \mathbb{R} \to \mathbb{C}$ be the *square wave* defined by
$$f(x) = \begin{cases} 1 & x - \lfloor x \rfloor \le 1/2 \\ 0 & \text{otherwise.} \end{cases}$$
The graph of f is shown in Figure 0.1. We claim that f is represented by the Fourier series
$$f \sim \frac{1}{2} + \sum_{n\in\mathbb{Z}, n \text{ odd}} \frac{1}{\pi i n} \exp(2\pi i n x).$$
We can see that this Fourier series at least appears to approximate f in Figure 0.2. Once we can show that the Fourier series of f converges to f, a natural next question

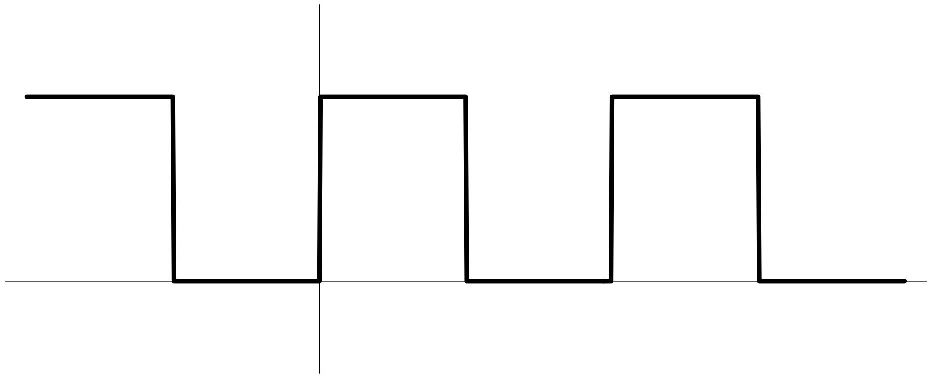


Figure 0.1: The square wave on $\mathbb{R}$ plotted on the region $[-1, 2]$.

is, "Does the derivative of the Fourier series of f converge to f'?" We will see that the answer of this question is a resounding "yes!" in the world of distributions. But in the present example the question is ill-posed; f doesn't even have a well-defined derivative on all of $\mathbb{R}$.

Still, the Fourier series of f can give us some clues as to what is happening with the derivative of f. Assuming that we can differentiate Fourier series term-wise, we compute the derivative

$$f' \sim \frac{d}{dx}\left(\frac{1}{2} + \sum_{n\in\mathbb{Z}, n \text{ odd}} \frac{1}{\pi i n}\exp(2\pi i n x)\right) = \sum_{n\in\mathbb{Z}, n \text{ odd}} 2\exp(2\pi i n x).$$

The second equality holds becuase $d/dx \exp(2\pi i n x) = 2\pi i n \cdot \exp(2\pi i n x)$. Certainly this series does not converge anywhere classically because its terms never tend to 0. For example,

$$\sum_{n\in\mathbb{Z}, n \text{ odd}} 2\exp(2\pi i n x)\big|_{x=0} = 1 + 1 + 1 + \cdots = \infty.$$

Nonetheless, we plot the first few terms in the Fourier series for f' in Figure 0.3. Qualitatively, the derivatives of the truncated series are small away from the half-integers. This agrees with the fact that $f'(x) = 0$ for x not equal to a half integer.

The interesting behavior occurs at the half integers, where the Fourier series diverges. For example, at $x = 0$, the series diverges to $+\infty$. This fact is consistent with the idea that f is "infinitely steep" where it is discontinuous. Eventually, we will see that the derivative of the Fourier series for f quantitatively encodes information about the nature of the discontinuity. We will see that this derivative is an example of a δ distribution. Away from the half integers, this δ distribution acts exactly like the constant function 0. At the half integers, the distributional derivative of f encodes how large the discontinuity in the original function is.

The theory of distributions completes differential calculus in the sense that every distribution has a well-defined derivative. Further, distributional derivatives agree with

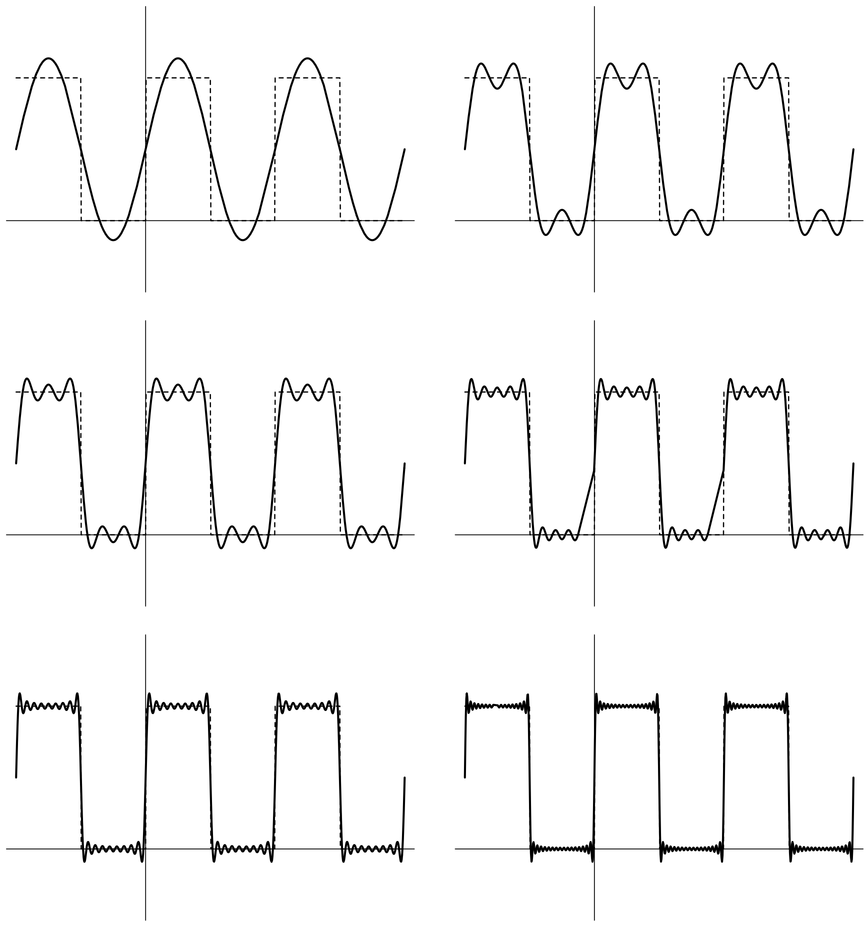


Figure 0.2: Here are the sums for $|n| < N$ of the Fourier series of the square wave for $N = 1, 2, 4, 8, 16, 32$. The series appears to converge to the square wave.

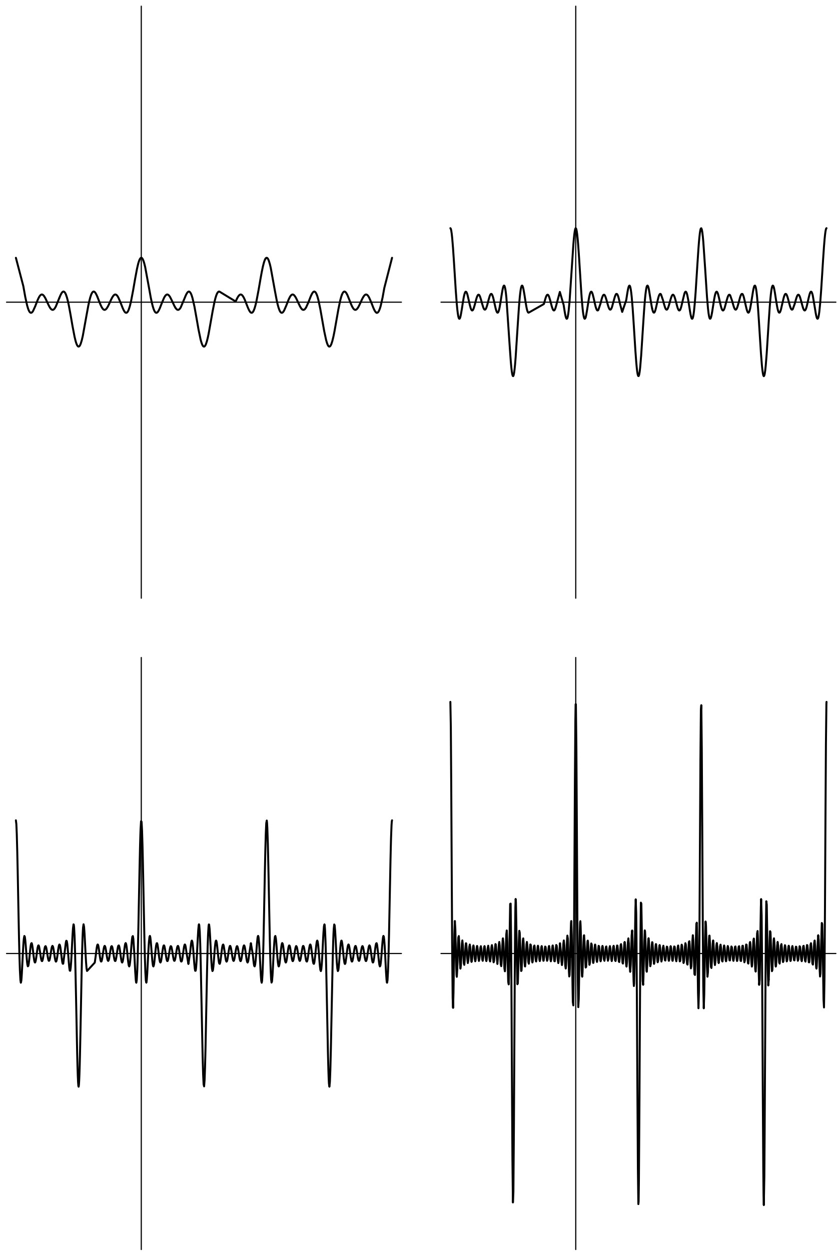


Figure 0.3: Pictured here are the sums for $|n| < N$ of the Fourier series for the derivative of the square wave for $N = 4,\ 8,\ 16,\ 32$. The series does not converge classically.

classical derivatives whenever the classical derivative is well defined, and every distribution arises by differentiating a classical function a finite number of times. Using Fourier series, we will identify the calculus of distributions with arithmetic of Fourier series.

We begin our discussion of Fourier series and distribution theory by developing calculus on the circle S^1. Functions defined on S^1 correspond exactly to periodic functions on $\mathbb{R}$. We will see, however, that the circle is a much more natural context in which to study Fourier series.

1

Differentiation on the Circle

In this chapter we introduce the circle S^1. We define the circle as the quotient of the real numbers by the integers $\mathbb{R}/\mathbb{Z}$. The advantage of defining S^1 as a quotient of $\mathbb{R}$ is that S^1 inherits its algebraic and topological structure from $\mathbb{R}$. Similarly, S^1 inherits a differential operator from $\mathbb{R}$ which allows us to consider differentiable functions on S^1. From differentiable functions, we describe the space of smooth (that is, infinitely differentiable) functions on S^1. In the final section, we introduce Fourier series. Fourier series plays an important role in our description of integration and distributions in the following chapter.

1.1 Function Spaces

In this section, we describe the space of continuous functions defined on a compact space, K. We show that the space $C^0(K)$ has the structure of a Banach space. We then consider the space of k-times continuously differentiable functions on the circle. We show that $C^k(S^1)$ is also a Banach space.

Continuous Functions on Compact Spaces

Let K be a compact space. We consider *continuous* complex-valued functions on K,

$$f : K \longrightarrow \mathbb{C}.$$

The space of continuous functions on K is denoted

$$C^0(K) = \{f : K \to \mathbb{C} \, : \, f \text{ is continuous}\}.$$

The superscript 0 refers to the fact that elements of $C^0(K)$ are zero (or more) times continuously differentiable.

We now explore the algebraic and topological structure of $C^0(K)$. Let $f, g \in C^0(K)$, $z \in \mathbb{C}$. Then,

1. $0 \in C^0(K)$
2. $f + g \in C^0(K)$,
3. $zf \in C^0(K)$.

Thus $C^0(K)$ has the structure of a (complex) vector space. Further, we define a norm[1] on $C^0(K)$,

$$\|f\|_{C^0} = \sup_{x \in K} |f(x)| .$$

This norm is often referred to as the *sup-norm* on K. The sup norm on K gives $C^0(K)$ metric structure via the metric

$$d(f, g) = \|f - g\|_{C^0} .$$

This metric is referred to as the *sup-norm* or C^0 *metric*. Further, this metric gives $C^0(K)$ a (metric) topology (called, unsurprisingly, the *sup-norm topology*). The sets

$$B_r(0) = \left\{ f \in C^0(K) \,:\, \|f\|_{C^0} < r \right\} \quad \text{for} \quad r \in \mathbb{R},\, r > 0.$$

form a local subbasis for the sup-norm topology of $C^0(K)$ around 0. The neighborhoods of arbitrary points $f \in C^0(K)$ are then translates of neighborhoods of 0.

Remark 1. This definition of the topology of $C^0(K)$ is convenient because translation is automatically bicontinuous. Therefore, any translate of an open set is again open. This property is useful to prove continuity of linear operators, because it suffices to prove continuity around 0. Further, the choice of subbasis $\{B_r(0) \,:\, 0 < r\}$ makes it clear that scalar multiplication is continuous, as the subbasis is closed under dilation.

Now that we have endowed $C^0(K)$ with algebraic and topological structure, we are ready to prove some useful properties of $C^0(K)$.

Proposition 2. Let $x \in K$. Then the *evaluation functional*[2] $C^0(K) \to \mathbb{C}$ defined by

$$f \longmapsto f(x)$$

is continuous.

---

[1] Recall that a *norm* on a complex vector space V is a function $\|\cdot\| : V \to \mathbb{R}$ that satisfies

1. $\|v\| \geq 0$ for all $v \in V$.
2. $\|\alpha v\| = |\alpha| \, \|v\|$ for all $v \in V, \alpha \in \mathbb{C}$
3. $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$
4. $\|v\| = 0 \iff v = 0$.

[2] Recall that a *(linear) functional* on a vector space V over a field F is a *linear* mapping $\lambda : V \to F$.

Remark 3. The continuity referred to in the proposition is not the continuity of the functions $f \in C^0(K)$. Rather it is the statement that if f and g are close in $C^0(K)$ (i.e., $d(f,g)$ is small), then $f(x)$ and $g(x)$ are correspondingly close in $\mathbb{C}$.

Proof. Let $f, g \in C^0(K)$. Then

$$|f(x) - g(x)| \leq \sup_{y \in K} |f(y) - g(y)| = \|f - g\|_{C^0} .$$

Fix $\varepsilon > 0$. Then

$$\|f - g\|_{C^0} < \varepsilon \quad \Longrightarrow \quad |f(x) - g(x)| < \varepsilon.$$

Therefore the evaluation functional is continuous. □

Theorem 4. The space $C^0(K)$ is complete with respect to the sup-norm metric.

Remark 5. The proof of this theorem hinges on the fact that convergence with respect to the sup-norm is uniform convergence.[3] Pointwise convergence is weaker than uniform convergence. If we only require pointwise convergence, i.e.,

$$f_i \longrightarrow f \iff f_i(x) \longrightarrow f(x) \quad \text{for all} \quad x \in K$$

then the theorem fails to be true. That is to say, pointwise limits of continuous functions can fail to be continuous. We can write a slogan for the result: *Uniform limits of continuous functions are continuous.*

Proof. We will first show that a Cauchy sequence of functions

$$\{f_0, f_1, f_2, ...\}, \quad f_i \in C^0(K)$$

has a pointwise limit. Given $\varepsilon > 0$ let N be such that $\|f_i - f_j\|_{C^0} < \varepsilon$ for all $i, j > N$. Fix $x \in K$. By the definition of $\|\cdot\|_{C^0}$,

$$|f_i(x) - f_j(x)| < \varepsilon, \quad \text{for all} \quad i, j > N. \tag{1.1}$$

Therefore, the sequence $\{f_i(x)\}$ is a Cauchy sequence in $\mathbb{C}$ which has a limit $f(x)$.

We will now show that the limit is a uniform limit. Fix $i > N$. Taking the limit of (1.1) as $j \to \infty$, for any $x \in K$ we have

$$|f_i(x) - f(x)| \leq \varepsilon, \quad \text{for all} \quad j > N.$$

Since the inequality holds for all $x \in K$,

$$\|f_i - f\|_{C^0} \leq \varepsilon,$$

so $\{f_i\}$ converges *uniformly* to f.

---

[3] Recall that a sequence of functions $\{f_n\}$ converges to a limit function f *uniformly* if for every $\varepsilon > 0$, there exists N such that

$$i > N \implies |f_i(x) - f(x)| < \varepsilon \quad \text{for all} \quad x \in K.$$

Finally, we will show that f is continuous, hence $f \in C^0(K)$. Fix $i > N$ and let $x \in K$. Since f_i is continuous on K, for any $\varepsilon > 0$ there exists an open neighborhood B_x of x such that

$$y \in B_x \implies |f_i(x) - f_i(y)| < \varepsilon.$$

Let $y \in B_x$. Then

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

Therefore f is continuous. □

We have shown that $C^0(K)$ is a complete normed vector space. Such spaces are known as *Banach* spaces. Banach spaces will play an important role in our development of function spaces.

The Circle S^1

In this section, we introduce the circle S^1. Much of the remainder of the chapter will be devoted to developing a rich function theory on the circle.

Definition 6. The circle S^1 is the quotient (of topological groups[4])

$$S^1 = \mathbb{R}/\mathbb{Z}.$$

Classically, we think of the quotient as a set of cosets,

$$S^1 = \{r + \mathbb{Z} \,:\, r \in \mathbb{R}\}.$$

This definition is correct, however we offer a different characterization of the quotient that emphasizes mapping properties. We make use of the universal property of the quotient. In the category of topological groups, the universal property of the quotient amounts to the following.

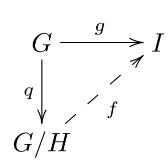
Definition 7. Let G be a topological group, H a closed normal subgroup. The *quotient* of G by H is a topological group G/H along with a continuous homomorphism

$$q : G \longrightarrow G/H$$

with the following property: For any topological group I and continuous homomorphism $g : G \to I$ such that $g(H) = 1_I$ there exists a unique continuous homomorphism

$$f : G/H \longrightarrow I$$

such that $g = f \circ q$. Pictorially, the following diagram commutes:



[4]A *topological group* is a group endowed with a topology such that the group operation and inversion are continuous.

Note that in defining S^1 as the quotient $\mathbb{R}/\mathbb{Z}$, there is a continuous, surjective quotient map

$$q : \mathbb{R} \to S^1.$$

Remark 8. The coset interpretation and the universal property characterization of the quotient are equally valid, and are equivalent for topological groups (although proving this fact is beyond the scope of this work). They differ in their emphasis. The coset interpretation gives explicitly the elements of the quotient, while universal property characterizes how the quotient interacts with other topological groups. We offer both descriptions of the quotient, because in different situations one or the other may be easier or more instructive to work with.

We favor the definition $S^1 = \mathbb{R}/\mathbb{Z}$ because we want to describe differentiable and eventually smooth functions defined on S^1. In defining S^1 as a quotient of $\mathbb{R}$, S^1 inherits a differential operator from to $\frac{d}{dx}$ on $\mathbb{R}$.

Differentiable Functions on S^1

Since we defined

$$S^1 = \mathbb{R}/\mathbb{Z}$$

we have the continuous surjective homomorphism

$$q : \mathbb{R} \longrightarrow \mathbb{R}/\mathbb{Z} = S^1.$$

Let $f : S^1 \to \mathbb{C}$ be a continuous map. Then we can pull back the function f to a function defined on $\mathbb{R}$,

$$q^* f : \mathbb{R} \longrightarrow \mathbb{C}, \quad \text{where} \quad q^* f = f \circ q.$$

The function $q^* f$ is called the *pullback* of f by q. Since f and q are continuous, so is their composition. Therefore

$$g = q^* f : \mathbb{R} \longrightarrow \mathbb{C}$$

is a continuous function. Further, we have

$$g(x) = g(x + n) \quad \text{for all} \quad x \in \mathbb{R},\ n \in \mathbb{Z},$$

so every such pullback g is periodic.

Remark 9. We can generalize the notion of a periodic function in the following manner. Let G be a topological group and let H be a closed normal subgroup, and let $g : G \to \mathbb{C}$. Then we call g an *automorph* of H if

$$g(x + y) = f(x) \quad \text{for all} \quad x \in G,\ y \in H.$$

We will define differentiable functions on S^1 in terms of the derivative d/dx on $\mathbb{R}$. Therefore, we will review differentiation on $\mathbb{R}$ before presenting its analogue on S^1.

Definition 10. Let $g : \mathbb{R} \to \mathbb{C}$ be a continuous function, and let $x \in \mathbb{R}$. If the limit

$$\lim_{h \to 0} \frac{g(x+h) - g(x)}{h}$$

exists, we say g is *differentiable at* x, and write the *derivative* of g

$$g'(x) = \lim_{h \to 0} \frac{g(x+h) - g(x)}{h}.$$

If f is differentiable for all $x \in \mathbb{R}$, we say that f is *differentible*. If the function f' is continuous, we say that f is *continuously differentiable*.

We denote the set of all continuously differentiable functions $g : \mathbb{R} \to \mathbb{C}$ by

$$C^1(\mathbb{R}) = \left\{ g : \mathbb{R} \to \mathbb{C} \,:\, g' \text{ exists and } g' \in C^0(\mathbb{R}) \right\}.$$

We define the $k-$th derivative of a function g (if it exists) recursively by

$$g^{(k)} = (g^{(k-1)})'.$$

Then we define

$$C^k(\mathbb{R}) = \left\{ g : \mathbb{R} \to \mathbb{C} \,:\, g^{(k)} \text{ exists, and } g^{(k)} \in C^0(\mathbb{R}) \right\}.$$

We are now in a position to define k-times differentiable functions on S^1.

Definition 11. Let $f : S^1 \to \mathbb{C}$ be continuous. Then f is *k-times differentiable* if

$$q^* f \in C^k(\mathbb{R}).$$

We denote the i-th derivative of f (if it exists) by $D^{(i)} f$, provisionally determined by the condition

$$q^*(D^{(i)} f) = (q^* f)^{(i)}.$$

The space of all k-times differential functions on S^1 is denoted

$$C^k(S^1) = \left\{ f : S^1 \to \mathbb{C} \,:\, f^{(i)} \text{ exists, and } f^{(i)} \in C^0(S^1) \text{ for all } 1 \leq i \leq k \right\}.$$

We have not demonstrated that $D^{(i)}$ is well-defined. Although $q^* f$ is periodic, we do not know *a priori* that $(q^* f)'$ is also periodic.

Proposition 12. There exists a unique well-defined differentiation map D that makes the following diagram commute:

$$\begin{CD}
C^k(\mathbb{R}) @>{\frac{d}{dx}}>> C^{k-1}(\mathbb{R}) \\
@A{q^*}AA @AA{q^*}A \\
C^k(S^1) @>{D}>> C^{k-1}(S^1)
\end{CD}$$

Proof. Let $g \in C^k(\mathbb{R})$. By the definition of quotient, g factors through q to give a function on S^1 if and only if

$$g(x+n) = g(x) \quad \text{for all} \quad x \in \mathbb{R},\ n \in \mathbb{Z}.$$

Suppose that g does satisfy the above periodicity condition. Let $y \in \mathbb{R}$ and let

$$T_y : C^k(\mathbb{R}) \to C^k(\mathbb{R})$$

be given by

$$(T_y g)(x) = g(x+y) \quad \text{for all} \quad x \in \mathbb{R}.$$

By the chain rule, the operators T_y and $\frac{d}{dx}$ commute:

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

Let $f \in C^k(S^1)$, and let $n \in \mathbb{Z}$. Since $q^* f \in C^k(\mathbb{R})$ and $q^* f$ is periodic,

$$T_n \left(\frac{d}{dx}(q^* f) \right) = \frac{d}{dx}(T_n(q^* f)) = \frac{d}{dx}(q^* f).$$

Hence the derivative $\frac{d}{dx}(q^* f)$ is periodic, and factors through q to give some function

$$Df : S^1 \to \mathbb{C}.$$

The fact that D is unique is a consequence of the injectivity of q^*. If D and $\tilde{D}$ satisfy the above diagram, then

$$q^*(Df) = q^*(\tilde{D} f).$$

Since both sides of the above equation are equal to $d/dx(q * f)$, $Df = \tilde{D} f$ for all $f \in C^k(S^1)$, so $D = \tilde{D}$. □

Remark 13. Since D is uniquely determined on S^1 by $\frac{d}{dx}$ on $\mathbb{R}$, from now on we denote differentiation on S^1 by $\frac{d}{dx}$. We also denote the derivative of particular functions $f : S^1 \to \mathbb{C}$ by f'.

Properties of $C^k(S^1)$

We would like to understand the algebraic and topological structure of $C^k(S^1)$. Let $f, g \in C^k(S^1)$ and $z \in \mathbb{C}$. Then

1. $f + g \in C^k(S^1)$
2. $zf \in C^k(S^1)$
3. $0 \in C^k(S^1)$.

These properties follow immediately from the corresponding properties of differentiation on $\mathbb{R}$. Further, the properties imply that $C^k(S^1)$ has the structure of a complex vector space.

Definition 14. Let $f \in C^k(S^1)$. Then we define the C^k*-norm* by

$$\begin{aligned}\|f\|_{C^k} &= \sum_{i=0}^{k} \left\| f^{(i)} \right\|_{C^0} \\ &= \sum_{i=0}^{k} \sup_x \left| f^{(i)}(x) \right|.\end{aligned}$$

Elementary calculation shows that $\|\cdot\|_{C^k}$ is indeed a norm on $C^k(S^1)$. As before, the norm on $C^k(S^1)$ gives rise to a metric

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

which in turn gives the space a topology. The existence of a norm on $C^k(S^1)$ gives $C^k(S^1)$ the structure of a *pre*-Banach space. Later, we will show that $C^k(S^1)$ is complete, making it a Banach space. First we demonstrate some consequences of the C^k topology on $C^k(S^1)$.

Proposition 15. Let $k \in \mathbb{Z}_{\geq 0}$. Then $C^{k+1}(S^1)$ embeds continuously into $C^k(S^1)$.

Proof. Let $f \in C^{k+1}(S^1)$. Since f is $(k+1)$-times continuously differentiable, it is certainly k-times continuously differentiable. Thus there is a natural inclusion

$$\varphi_{k+1,k} : C^{k+1}(S^1) \longrightarrow C^k(S^1).$$

To see that $\varphi_{k+1\,k}$ is continuous, fix $\varepsilon > 0$. Let $f \in C^{k+1}(S^1)$ with

$$\|f\|_{C^{k+1}} = \sum_{0 \leq i \leq k+1} \left\| (f)^{(i)} \right\|_{C^0} < \varepsilon.$$

Then

$$\|\varphi_{k+1,k}(f)\|_{C^k} = \|f\|_{C^k} = \sum_{0 \leq i \leq k} \left\| (f)^{(i)} \right\|_{C^0} \leq \|f\|_{C^{k+1}} < \varepsilon.$$

□

Due the previous proposition, there is an infinite chain of continuous inclusions

$$\cdots \xrightarrow{\varphi_{k+1\,k}} C^k(S^1) \xrightarrow{\varphi_{k\,k-1}} \cdots \xrightarrow{\varphi_{2\,1}} C^1(S^1) \xrightarrow{\varphi_{1\,0}} C^0(S^1).$$

This chain will be of use in the following section when we describe smooth functions on S^1.

Proposition 16. Let $k \in \mathbb{Z}_{\geq 0}$. Then differentiation

$$\frac{d}{dx} : C^{k+1}(S^1) \longrightarrow C^k(S^1)$$

is continuous.

Proof. Fix $\varepsilon > 0$. Let $f \in C^{k+1}(S^1)$ for some $k \geq 0$ such that

$$\|f\|_{C^{k+1}} < \varepsilon.$$

Then

$$\begin{aligned}
\|f'\|_{C^k} &= \left\| \frac{d}{dx}(f) \right\|_{C^k} \\
&= \sum_{0 \leq i \leq k} \left\| \left(\frac{d}{dx}(f) \right)^{(i)} \right\|_{C^0} \\
&= \sum_{1 \leq i \leq k+1} \left\| (f)^{(i)} \right\|_{C^0} \\
&\leq \sum_{0 \leq i \leq k+1} \left\| (f)^{(i)} \right\|_{C^0} \\
&= \|f\|_{C^{k+1}} \\
&< \varepsilon.
\end{aligned}$$

□

Proposition 17. Let $x \in S^1$, $i \in \mathbb{Z}$, with $0 \leq i \leq k$. Then the evaluation functional defined by

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

is continuous with respect to the C^k topology.

Proof. This is an immediate consequence of continuity of differentiation and continuity of evaluation for continuous functions, as the composition of continuous functions is again continuous. □

We are now ready to prove the main result of this section.

Theorem 18. The metric space $C^k(S^1)$ is a Banach space.

Proof. Since we defined the *k*-th derivative of a function recursively, it suffices to show the case where $k = 1$. The theorem then follows by induction. Let $\{f_n\}$ be a Cauchy sequence in $C^1(S^1)$. Then the pointwise limits

$$f(x) = \lim_n f_n(x), \quad g(x) = \lim_n f_n'(x)$$

are uniform limits, so the limit functions are continuous. We must show that $f' = g$. To this end, we pull the limitands back, via the quotient map q, to (periodic) functions on $\mathbb{R}$. On $\mathbb{R}$ we can comfortably use the fundamental theorem of calculus. Since the $q^* f_i$ are continuous,

$$(q^* f_i)(x) - (q^* f_i)(a) = \int_a^x (q^* f_i)'(t)\, dt \quad \text{for all} \quad x, a \in \mathbb{R}.$$

Taking the limit as $i \to \infty$, we can write

$$\begin{aligned}(q^*f)(x) - (q^*f)(a) &= \lim_i ((q^*f_i)(x) - (q^*f_i)(a)) \\ &= \lim_i \int_a^x (q^*f_i)'(t)\,dt \\ &= \int_a^x \lim_i (q^*f_i)'(t)\,dt \\ &= \int_a^x (q^*g)(t)\,dt.\end{aligned}$$

The interchange of the limit and integral is justified by dominated convergence. Differentiating both sides of the above equation and applying the fundamental theorem of calculus again gives

$$(q^*f)'(x) = (q^*g)(x) \quad \text{for all} \quad x \in \mathbb{R}.$$

By the characterization of differentiation on S^1 given in Definition 11 (and by Proposition 12),

$$f'(x) = g(x) \quad \text{for all} \quad x \in S^1.$$

□

1.2 Smooth Functions on S^1

In this section, we develop the space of smooth functions on S^1, denoted $C^\infty(S^1)$. Smooth functions are functions $f : S^1 \to \mathbb{C}$ such that $f \in C^k(S^1)$ for all $k \in \mathbb{Z}_{\geq 0}$. In order to describe $C^\infty(S^1)$, we introduce topological vector spaces and projective limits.

Topological Vector Spaces

We previously showed that the spaces $C^k(S^1)$ are Banach spaces. The space $C^\infty(S^1)$, however, is not a Banach space. In order to describe the structure of $C^\infty(S^1)$, we must introduce a more general space: the topological vector space.

Definition 19. A *topological vector space* is a complex vector space V with a Hausdorff[5] topology such that vector addition,

$$+ : V \times V \longrightarrow V \qquad v \times w \longmapsto v + w$$

and scalar multiplication,

$$\cdot : \mathbb{C} \times V \longrightarrow V \qquad \alpha \times v \longmapsto \alpha \cdot v$$

are continuous. The topologies on the domains of $+$ and $\cdot$ are taken to be the product topology. Further, we require that V be locally convex. That is, the topology has a basis at 0 consisting of convex sets.[6]

[5] Recall that a topology on a space X is *Hausdorff* if for all $x, y \in X$, $x \neq y$ there exist neighborhoods of x and y that are disjoint.

[6] Recall that a set B is convex if for every $x, y \in B$ and $\alpha, \beta \in \mathbb{R}^+$ with $\alpha + \beta = 1$, $\alpha x + \beta y \in B$.

Remark 20. In order to describe a topology on a topological vector space V, it suffices to give a basis of neighborhoods of the origin. This is because vector addition is continuous. If B_0 is a neighborhood of the origin then define $T_y : V \to V$ by $T_y(x) = x + y$. T_y is bi-continuous because $(T_y)^{-1} = T_{-y}$. Therefore T_y is open, so $T_y(B_0)$ is a neighborhood of y. Therefore, a basis for the topology at the origin generates the topology on all of V by translation.

Topological vector spaces form a category where the morphisms are continuous linear maps. The following proposition shows that the study of topological vector spaces is relevant to our study of functions on circles, as we showed that $C^k(S^1)$ has the structure of a Banach space.

Proposition 21. Let X be a Banach space. Then X is a topological vector space.

Proof. We must show that addition and scalar multiplication are continuous around the origin of X. Since X is a Banach space, it has a metric topology where the metric is defined

$$d(x, y) = \|x - y\|.$$

Therefore, we can use the usual delta-epsilon definition of continuity. Fix $\varepsilon > 0$. Let $x, y \in X$ satisfy $\|x\| < \varepsilon$, $\|y\| < \varepsilon$. Then

$$d(x + y, 0) = \|x + y\| \leq \|x\| + \|y\| < 2\varepsilon.$$

Therefore addition is continuous around 0. Let $\alpha \in \mathbb{C}$. Then

$$d(\alpha \cdot x, 0) = \|\alpha \cdot x\| = |\alpha| \|x\| < |\alpha| \cdot \varepsilon$$

so scalar multiplication is also continuous around 0. Therefore, X is a topological vector space. □

$C^\infty(S^1)$ as a Projective Limit

The space $C^\infty(S^1)$ is the space of functions on S^1 that are infinitely differentiable. That is,

$$C^\infty(S^1) = \left\{ f : S^1 \to \mathbb{C} \, : \, f \in C^k(S^1) \text{ for all } k \in \mathbb{Z}_{\geq 0} \right\}.$$

Since each $C^{k+1}(S^1)$ embeds in $C^k(S^1)$, we can describe $C^\infty(S^1)$ as the intersection

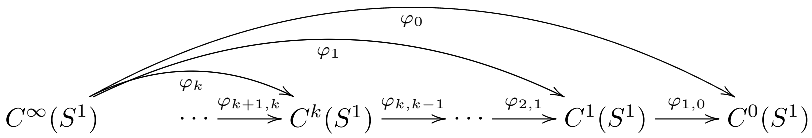
$$C^\infty(S^1) = \bigcap_{k=0}^{\infty} C^k(S^1).$$

This description, however, describes $C^\infty(S^1)$ only as a set. We would like to understand the topological and algebraic structure of $C^\infty(S^1)$, so we give an alternate characterization of $C^\infty(S^1)$ in terms of a *projective limit*.

In order to characterize the projective limit, we form the infinite chain of injections

$$\cdots \xrightarrow{\varphi_{k+1,k}} C^k(S^1) \xrightarrow{\varphi_{k,k-1}} \cdots \xrightarrow{\varphi_{2,1}} C^1(S^1) \xrightarrow{\varphi_{1,0}} C^0(S^1).$$

The *projective limit* of the spaces $C^k(S^1)$ is a topological vector space $C^\infty(S^1)$ together with maps φ_k to each $C^k(S^1)$ such that all triangles commute in the following diagram:



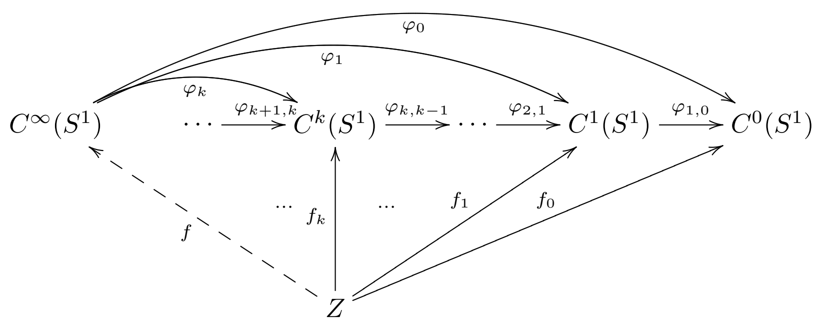
The fact that the diagram *commutes* means

$$\varphi_k = \varphi_{k+1,k} \circ \varphi_{k+1} \quad \text{for all} \quad k \in \mathbb{Z}_{\geq 0}.$$

The projective limit $C^\infty(S^1)$ is characterized by the following universal property. For any topological vector space Z with compatible[7] maps

$$f_k : Z \longrightarrow C^k(S^1) \quad \text{for all} \quad k \in \mathbb{Z}_{\geq 0}$$

there exists a unique map $f : Z \to C^\infty(S^1)$ such that all triangles in the following diagram commute:



Remark 22. The characterization of the (projective) limit described above is not limited to characterizing limits of topological vector spaces. It is a general characterization of the *universal property of the projective limit*. In the general case, all spaces are objects in the same category, and all arrows are morphisms. Although the universal property is the same for every category, the existence of the projective limit must be proven for each category.

Remark 23. In the category of sets, nested intersections are examples of limits. Let $\{U_k\}$ be a family of sets for which

$$U_{k+1} \subset U_k \quad \text{for all} \quad k \in \mathbb{Z}_{\geq 0}.$$

Then

$$\lim_k U_k = \bigcap_k U_k.$$

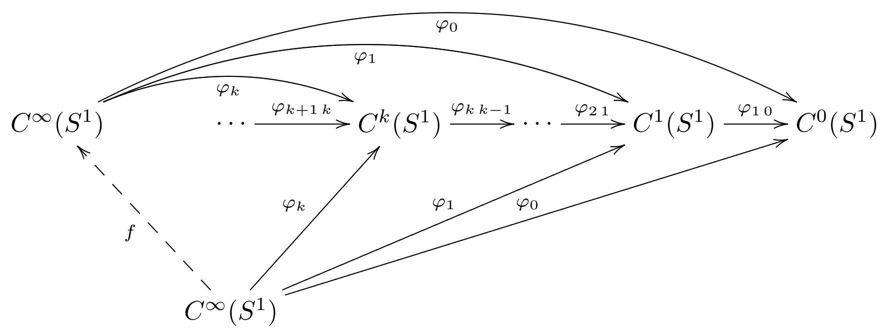
Fact 24. Projective limits of topological vector space exist.

[7] *Compatible* maps are maps such that $\varphi_{k+1,k} \circ f_{k+1} = f_k$ for all $k \in \mathbb{Z}_{\geq 0}$.

The proof of the claim is beyond the scope of the current work; see [6]. Although the existence of the limit is tricky to prove, uniqueness of the limit is easy to prove (although somewhat tedious to typeset).

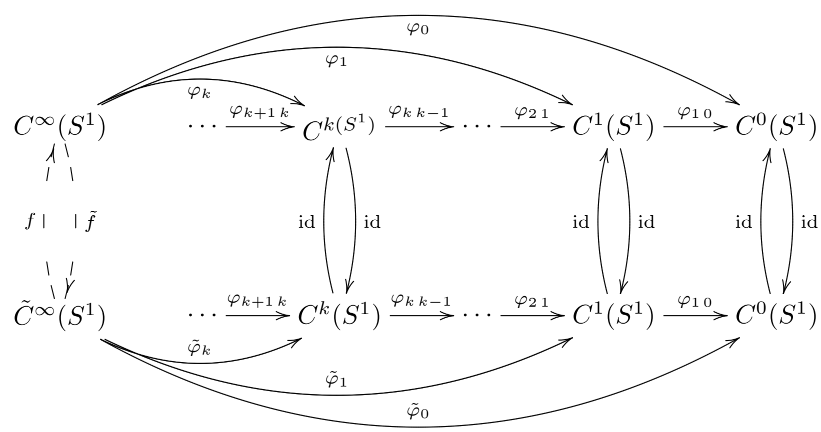
Proposition 25. The projective limit $C^\infty(S^1)$ is unique up to unique isomorphism.

Proof. In the characterization of the limit given above, let $Z = C^\infty(S^1)$. We will show that the only map $f : C^\infty(S^1) \to C^\infty(S^1)$ that makes the following diagram commute is the identity map on $C^\infty(S^1)$.



Certainly, if f is the identity map on $C^\infty(S^1)$, the diagram commutes. By the universal property of the limit, the map f is unique. Therefore, the identity is the only map taking $C^\infty(S^1)$ to itself that is compatible with all of the maps φ_k.

Now let $Z = \tilde{C}^\infty(S^1)$, a limit of the same limitands:



The maps f_k are obtained by the composition

$$f_k = \mathrm{id} \circ \varphi_k.$$

The composition $f \circ \tilde{f}$ is compatible from $C^\infty(S^1)$ to itself, so $f \circ \tilde{f}$ must be the identity on $C^\infty(S^1)$. Similarly $\tilde{f} \circ f$ is the identity on $\tilde{C}^\infty(S^1)$. Therefore f and $\tilde{f}$ are mutual inverses. In particular, each is an isomorphism. Further, by the universal property of the projective limit, the maps f and $\tilde{f}$ are unique. □

Remark 26. At the beginning of the section, we noted that a limit of Banach spaces (such as $C^\infty(S^1)$) is not generally a Banach space. Instead, such limits are *Frechet* spaces. Like Banach spaces, Frechet spaces are complete, metrizable topological vector spaces. Frechet spaces, however, are not normed spaces. See [6].

Remark 27. In defining $C^\infty(S^1)$ as the projective limit of the spaces $C^k(S^1)$, $C^\infty(S^1)$ inherits its algebraic and topological structure from the limitands. Further, the universal property of the projective limit allows us prove theorems about $C^\infty(S^1)$ without reference to its elements. This way we can often avoid messy delta-epsilon type proofs in favor of appealing to the topology on $C^\infty(S^1)$.

Proposition 28. The map

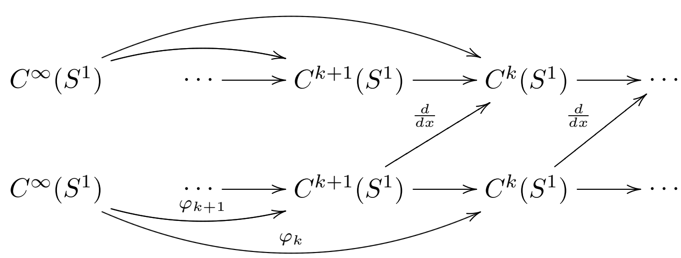
$$\frac{d}{dx} : C^\infty(S^1) \longrightarrow C^\infty(S^1)$$

is continuous with respect to the limit topology.

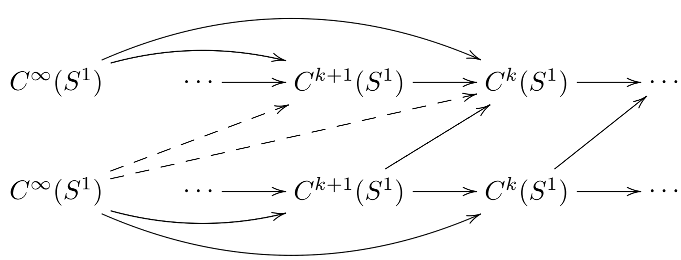
Proof. We previously showed that

$$\frac{d}{dx} : C^{k+1}(S^1) \longrightarrow C^k(S^1) \quad \text{for} \quad k \in \mathbb{Z}_{\geq 0}$$

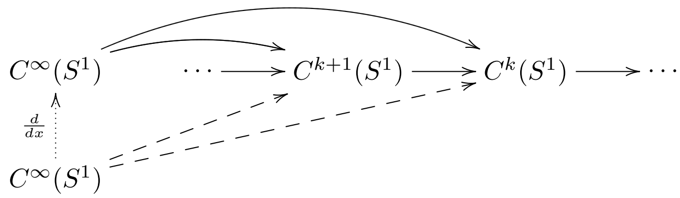
is continuous. We form the diagram



Composing the morphisms φ_k with $\frac{d}{dx}$ gives compatible maps indicated in dashed lines:



By the universal property of of the limit, there exists a unique morphism from $C^\infty(S^1)$ to itself that is the derivative inherited from the limitands:



The map $\frac{d}{dx}$ is a morphism of topological vector spaces, hence it is continuous. □

1.3 Fourier Series

In this section, we introduce Fourier series on the circle S^1. Fourier series is a way of representing functions on the circle as linear combinations of complex exponential functions. We investigate the conditions under which the Fourier series of a function converges to a function that is k-times differentiable or smooth.

The Functions ψ_n

The goal of Fourier series is to represent functions on S^1 as linear combinations of the functions[8]

$$\psi_n(x) = e^{2\pi i n x} \quad \text{for all} \quad n \in \mathbb{Z},\ x \in S^1.$$

The functions ψ_n are infinitely differentiable, so

$$\psi_n \in C^\infty(S^1) \quad \text{for all} \quad n \in \mathbb{Z}.$$

A *Fourier series* is an infinite series of the form

$$f = \sum_{n \in \mathbb{Z}} c_n \psi_n \quad \text{for coefficients} \quad c_n \in \mathbb{C}.$$

Soon, we will give conditions for the coefficients c_n under which the series converges to a function $f \in C^k(S^1)$.

In order to motivate the definition of the functions ψ_n, we introduce the notion of a character. Let G be a locally compact abelian group. Then a (*unitary*) *character* χ of G is a morphism (that is, a continuous group homomorphism)

$$\chi : G \longrightarrow \mathbb{T}$$

where $\mathbb{T}$ is the unit circle in $\mathbb{C}$. The characters of a group themselves form a group, where the group operation comes from pointwise multiplication in $\mathbb{T}$:

$$(\chi\xi)(g) = \chi(g)\xi(g) \quad \text{for all} \quad g \in G.$$

The group of characters is known as the *dual group* (or Pontryagin dual) of G, and is denoted $\widehat{G}$.

[8]Technically, the functions ψ_n are defined on $\mathbb{R}$, but since each is periodic, they correspond to functions on S^1.

The functions ψ_n are characters of S^1, because

$$\psi_n(0) = 1, \quad \psi_n(x+y) = \psi_n(x)\psi_n(y) \quad \text{for all} \quad n \in \mathbb{Z},\ x, y \in S^1.$$

Further, we have the relation

$$(\psi_n\psi_m)(x) = \psi_n(x)\psi_m(x) = \psi_{n+m}(x) \quad \text{for all} \quad n, m \in \mathbb{Z},\ x \in S^1.$$

It turns out that ψ_n are the only characters of S^1. From the previous display, we can see that

$$\widehat{S^1} \cong \mathbb{Z}.$$

Incidentally, the ψ_n are also eigenfunctions of the differential operator d/dx with eigenvalues $2\pi in$. That is

$$\frac{d}{dx}\psi_n(x) = 2\pi in\, \psi_n(x).$$

Therefore, we may suspect that a function $f \in C^k(S^1)$ represented by the Fourier series

$$f(x) = \sum_{n\in\mathbb{Z}} c_n\psi_n(x)$$

has derivative

$$f'(x) = \frac{d}{dx}\sum_{n\in\mathbb{Z}} c_n\psi_n(x) = \sum_{n\in\mathbb{Z}}(2\pi in)c_n\psi_n(x).$$

We will eventually show that we can differentiate every Fourier series term-wise, but we need to develop more machinery before we can prove it.

Properties of Fourier Series

We now give a sufficient condition under which a Fourier series converges to a k-times continuously differentiable function. We will also show that we can differentiate such a series term-wise.

Proposition 29. Consider the Fourier series

$$\sum_{n\in\mathbb{Z}} c_n\psi_n$$

where the sum

$$\sum_{n\in Z} |c_n|\, |n|^{k+1}$$

converges. Then the Fourier series converges to a function $f \in C^k(S^1)$. Further the derivative of f is computed by term-wise differentiation

$$f'(x) = \sum_{n\in\mathbb{Z}}(2\pi in)c_n\psi_n(x).$$

Proof. We will show that the partial sums given by

$$f_N = \sum_{|n|\le N} c_n \psi_n(x)$$

converge to a function $f \in C^k(S^1)$. First, note that

$$|\psi_n(x)| = 1 \quad \text{for all} \quad n \in \mathbb{Z},\ x \in S^1.$$

Therefore

$$|f_N(x)| \le \sum_{n\le N} |c_n| \quad \text{for all} \quad x \in S^1.$$

Since f_N is a finite sum, we can differentiate term-wise, and we find

$$f_N^{(j)} = \sum_{|n|\le N} c_n (2\pi i n)^j \psi_n(x).$$

Taking taking the sup-norm of the above expression gives

$$\left\| f_N^{(j)} \right\|_{C^0} = \left\| \sum_{|n|\le N} c_n (2\pi i)^j \psi_n(x) \right\|_{C^0} \le \sum_{|n|\le N} (2\pi n)^j \left| c_n \right|.$$

Therefore, the C^k norm of f_N is

$$\begin{aligned}
\|f_N\|_{C^k} &= \sum_{0\le j\le k} \left\| f_N^{(j)} \right\|_{C^0} \\
&\le \sum_{0\le j\le k} \sum_{|n|\le N} |c_n| \,(2\pi n)^j \\
&= \sum_{|n|\le N} \sum_{0\le j\le k} |c_n| \,(2\pi n)^j && \text{(sums are finite)} \\
&= \sum_{|n|\le N} |c_n| \frac{|2\pi n|^{k+1} - 1}{|2\pi n| - 1} && \text{(by geometric sum formula)} \\
&\le (2\pi)^{k+1} \sum_{|n|\le N} |c_n| \,|n|^{k+1}.
\end{aligned}$$

By the hypothesis, the final sum converges as $N \to \infty$, therefore $\{f_N\}$ is a Cauchy sequence in $C^k(S^1)$. By Theorem 18, $\{f_N\}$ converges to a function $f \in C^k(S^1)$.

To see that we can differentiate the series term-wise, we note that differentiation maps $C^k(S^1) \to C^{k-1}(S^1)$ continuously. Therefore differentiation takes Cauchy sequences in $C^k(S^1)$ to $C^{k-1}(S^1)$. In particular, d/dx maps the sequence $\{f_N\}$ to $\{(f_N)'\}$, where the latter is a Cauchy sequence in $C^{k-1}(S^1)$. Therefore

$$\frac{d}{dx} f = \frac{d}{dx} \lim_{N\to\infty} f_N = \lim_{N\to\infty} (f_N)'$$

so we can differentiate the series for f term-wise to obtain the derivative of f. □

Proposition 30. Let $f : S^1 \to \mathbb{C}$ be given by the Fourier series

$$f(x) = \sum_{n \in \mathbb{Z}} c_n \psi_n(x) \quad \text{for all} \quad x \in S^1$$

where

$$\sum_{n \in \mathbb{Z}} |c_n| \, n^N \quad \text{converges for every} \quad N \in \mathbb{Z}.$$

Then $f \in C^\infty(S^1)$, and f' is obtained by differentiating the Fourier series for f term-wise.

Proof. By the previous proposition, $f \in C^k(S^1)$ for all $k \in \mathbb{Z}_{\geq 0}$. Therefore $f \in C^\infty(S^1)$. To see that we can differentiate the Fourier series term-wise, we use the fact that differentiation is continuous on $C^\infty(S^1)$. Let

$$f_N = \sum_{|n| \leq N} c_n \psi_n \quad \text{for all} \quad N \in \mathbb{Z}^{\geq 0}.$$

Then

$$f_N \xrightarrow{N \to \infty} f \in C^\infty(S^1).$$

Since the sums f_N are finite, we can differentiate them term-wise. Further, the derivatives $\{f'_N\}$ are a Cauchy sequence in $C^\infty(S^1)$, because differentiation is continuous. Since $C^\infty(S^1)$ is complete, $\{f'_N\}$ converges in $C^\infty(S^1)$. In particular

$$f'_N \xrightarrow{N \to \infty} f'.$$

Therefore, we can differentiate the entire Fourier series for f term-wise. □

So far, we have given conditions under which a Fourier series converges to a k-times differentiable or smooth function. We would still like to answer the question, "Given a function f, is there a Fourier series that converges to f?" Before we can answer this question, we must develop more function theory on the circle. Along the way, we introduce distributions. The theory of distributions gives us a general framework in which we can investigate the representation of distributions (and as a special case, functions) by Fourier series.

2

Integration & Distributions on the Circle

The aim of this chapter is to introduce distributions and integration on S^1. After our discussion of smooth functions in the previous chapter, we are already in a position to define distributions. In order to effectively work with distributions, however, we must develop a suitable theory of integration on S^1. In particular, we will define the Hilbert space $L^2(S^1)$ of square-integrable functions on S^1. We connect the L^2 theory with our previous discussion of Fourier series by showing that every square integrable function can be represented by Fourier series. The space $L^2(S^1)$ will serve as our prototype for Sobolev spaces, which are the focus of the following chapter.

2.1 Introduction to Distributions on S^1

In this section, we introduce distributions (or generalized functions) on the circle. Distributions are linear functionals on the space of smooth functions on S^1. In order to describe the space of distributions we introduce colimits. At the end of the section, we present differentiation of distributions.

Basic Definition

Recall that the *dual space* of a complex vector space V, denoted V^*, is the space of linear functionals

$$\lambda : V \longrightarrow \mathbb{C}.$$

In the case of topological vector spaces, we further require that λ be continuous.

Definition 31. A *distribution* is a continuous linear functional

$$u : C^\infty(S^1) \longrightarrow \mathbb{C}.$$

The space of distributions on the circle is the continuous dual space of $C^\infty(S^1)$, denoted $C^\infty(S^1)^*$.

Remark 32. This definition shows that it may be quite difficult to describe a particular distribution. In order to define a distribution $u \in C^\infty(S^1)^*$, must one specify the value $u(f)$ for every $f \in C^\infty(S^1)$? Fortunately, we will see that the answer to this question is a resounding "no." In fact, it will suffice to define the values $u(\psi_n)$ for all $n \in \mathbb{Z}$. Eventually we will see that every distribution is representable by Fourier series (with some constraints on the exponents).

In the next section, we will develop integration on the circle. Once we develop a suitable integral, we embed continuous functions into $C^\infty(S^1)^*$ in the following manner. Let $f \in C^0(S^1)$. We define the distribution corresponding to f to be

$$u_f(\varphi) = \int_{S^1} f(x)\varphi(x)\,dx \quad \text{for all} \quad \varphi \in C^\infty(S^1).$$

The embedding

$$f \longmapsto u_f$$

is continuous, but in order to establish continuity, we will need to describe the topology on $C^\infty(S^1)^*$.

Colimits

In order to describe the topological structure of $C^\infty(S^1)^*$, we introduce colimits. Like limits, colimits are characterized by a universal property.

Definition 33. Let $\{X_k\}$ be objects with morphisms

$$\varphi_{k,\,k+1} : X_k \longrightarrow X_{k+1} \quad \text{for all} \quad k.$$

The *colimit* of the X_k (if it exists) is an object $\operatorname*{colim}_k X_k$ along with compatible morphisms

$$\varphi_k : X_k \longrightarrow \operatorname*{colim}_k X_k.$$

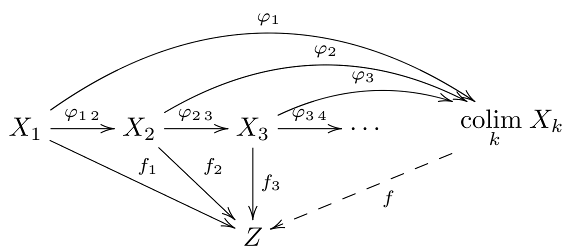
The colimit is characterized by the following universal property. For any object Z with compatible maps

$$f_k : X_k \longrightarrow Z \quad \text{for all} \quad k.$$

there exists a unique map

$$f : \operatorname*{colim}_k X_k \longrightarrow Z$$

that makes all triangles in the following diagram commute:



Remark 34. In our discussion of limits, we noted that nested intersections are examples of limits. In a similar vein, ascending unions are examples of colimits in the category of sets. Let $\{X_k\}$ be a family of sets with $X_k \subset X_{k+1}$ for all k. Then

$$\operatorname*{colim}_k X_k = \bigcup_k X_k.$$

We previously defined the space $C^\infty(S^1)$ as a limit of the $C^k(S^1)$. Similarly, we define the space of distributions, $C^\infty(S^1)^*$, in terms of the dual spaces $C^k(S^1)^*$. As the space $C^{k+1}(S^1)$ embeds in $C^k(S^1)$, the space $C^k(S^1)^*$ embeds in $C^{k+1}(S^1)$. Therefore, we might think that

$$C^\infty(S^1)^* = (\lim_k C^k(S^1))^* = \operatorname*{colim}_k C^k(S^1)^*.$$

This turns out to be true, but the construction does not hold in as much generality as one would like.

We previously claimed that limits of topological vector spaces always exist. However, colimits of topological vector spaces do not always exist. Fortunately the colimits of dual spaces of Banach spaces do exist.

Theorem 35. Let $X = \lim_k B_k$ be the limit of Banach spaces B_k. Then

$$(\lim_k B_k)^* \cong \operatorname*{colim}_k B_k^*.$$

The proof of this theorem requires more background, and is given in Appendix A. We can summarize the result: *The (continuous) dual of the limit of Banach spaces is the colimit of the (continuous) duals.*

Corollary 36. The space of distributions on S^1 is the colimit

$$C^\infty(S^1)^* = (\lim_k C^k(S^1))^* = \operatorname*{colim}_k C^k(S^1)^*.$$

Differentiation of Distributions

In order to define differentiation on the space of distributions, we use the analogy of integration by parts. [1] Let $f, g \in C^k(S^1)$ for some $k \geq 1$. The integration by parts

[1]We will go into depth describing integration on the circle in the following section

formula states that

$$\int_{S^1} f(x)g'(x)\,dx = -\int_{S^1} f'(x)g(x)\,dx.$$

There are no boundary terms because S^1 has empty boundary. By analogy, [2] we define the derivative of a distribution by

$$u'(f) = -u(f') \quad \text{for all} \quad f \in C^\infty(S^1).$$

Since $d/dx : C^\infty(S^1) \to C^\infty(S^1)$ is continuous,

$$u' = -\left(u \circ \frac{d}{dx}\right) : C^\infty(S^1) \longrightarrow \mathbb{C}$$

is well defined and continuous. That is, u' is a distribution.

A consequence of the proof of Theorem 35 (see Appendix A) is that every distribution $u \in C^\infty(S^1)^*$ is an element in $C^k(S^1)^*$ for some $k \in \mathbb{Z}_{\geq 0}$.

Definition 37. If $u \in C^k(S^1)^*$, but $u \notin C^{k-1}(S^1)^*$, we say that u has *order* k.

If u has order k for some $k \geq 1$, then $u(f)$ is well-defined for all $f \in C^k(S^1)$. We evaluate the derivative of u,

$$u'(f) = -u(f') \quad \text{for all} \quad f \in C^k(S^1).$$

However, $f' \in C^{k-1}(S^1)$, so u' is not well defined for all $f \in C^k(S^1)$. We must require that $f \in C^{k+1}(S^1)$ in order for $u'(f)$ to be well defined. Therefore, we anticipate that u' has order $k+1$.

Previously, we suggested that classical functions on S^1 embed in the space of distributions via the relation

$$u_g(f) = \int_{S^1} g(x)f(x)\,dx \quad \text{for all} \quad f \in C^\infty(S^1)$$

where $g : S^1 \to \mathbb{C}$ is a classical function. What is the order of such a distribution? Assuming that g is integrable, the above integral is certainly well-defined for all $f \in C^0(S^1)$, so the order of u_g is apparently 0. If $g \in C^k(S^1)$ for some $k \geq 1$, then

$$(u_g)'(f) = \int_{S^1} -g(x)f'(x)\,dx = \int_{S^1} g'(x)f(x)\,dx = u_{g'}(f).$$

The k-th derivative of g is a classical function, and thus $u_{g^{(k)}} = (u_g)^{(k)}$ has order zero. This suggests that classically differentiable functions correspond to distributions of negative order. The duality of classically differentiable functions and distributions is the subject of Chapter 3.

[2] We can think of evaluation of a distribution as integration against a smooth function. By abuse of notation:

$$u(f) \sim \int_{S^1} u \cdot f.$$

2.2 Integration on S^1

In the previous section, we saw that the theory of distributions on the circle hinges on integration. In this section we characterize a suitable integral on S^1 by the property of translation invariance. Such an integral yields the orthogonality of characters of the circle, and gives the integration by parts formula. At the end of the section, we define an invariant integral on S^1 in terms of the Lebesgue integral on $\mathbb{R}$.

Invariant Integration

We require that our integral on S^1 be *translation invariant*. That is, for any measurable function $f : S^1 \to \mathbb{C}$

$$\int_{S^1} f(x+y)\,dx = \int_{S^1} f(x)\,dx \quad \text{for all} \quad y \in S^1.$$

Remark 38. Translation invariance alone determines a unique integral on compact topological groups (up to scalar multiple). Such an integral arises from an invariant measure on the space, known as *Haar measure*, and the integral is referred to as a *Haar integral*. We will not develop the general machinery of measure theory on compact topological groups, but will instead appeal to more familiar integration on the real line.

With a translation invariant integral, we can already prove a useful result concerning unitary characters on S^1.

Proposition 39. Let $\chi, \xi : S^1 \to \mathbb{C}^\times$ be unitary characters with $\chi \neq \xi$. Then

$$\int_{S^1} \chi(x)\bar{\xi}(x)\,dx = 0.$$

Remark 40. We use the notation $\bar{\xi}$ to denote the complex conjugate of ξ. Note that

$$\bar{\xi} = \xi^{-1}.$$

Proof (of proposition). Since $\chi \neq \xi$, the product $\chi\bar{\xi}$ is not the trivial character. Therefore, for some $y \in S^1$, $(\chi\bar{\xi})(y) \neq 1$. We can then calculate the integral

$$\begin{aligned}\int_{S^1} \chi(x)\bar{\xi}(x)\,dx &= \int_{S^1} \chi(x+y)\bar{\xi}(x+y)\,dx \\ &= \int_{S^1} \chi(x)\chi(y)\bar{\xi}(x)\bar{\xi}(y)\,dx \\ &= (\chi\bar{\xi})(y) \int_{S^1} \chi(x)\bar{\xi}(x)\,dx.\end{aligned}$$

Since $(\chi\bar{\xi})(y) \neq 1$,

$$\int_{S^1} \chi(x)\bar{\xi}(x)\,dx = 0$$

as desired. □

Remark 41. In the previous proof, we made no reference to the particulars of the circle. We only used the fact that χ and ξ were unitary characters and that the integral defined on the space was translation invariant. Therefore, the result holds for all compact topological groups endowed with a translation invariant integral.

We remarked earlier that an invariant integral on S^1 is uniquely determined up to scalar multiple. We would like the integral on S^1 to be normalized, so that

$$\int_{S^1} 1\, dx = 1.$$

The further normalization requirement pins down a unique integral on S^1. Let χ be a unitary character of S^1. For a normalized integral on S^1 we find

$$\int_{S^1} |\chi|^2 \; dx = \int_{S^1} 1\, dx = 1.$$

This result along with the previous proposition suggests an orthonormality relation between characters:

$$\langle \chi, \xi \rangle = \begin{cases} 1 & \chi = \xi \\ 0 & \chi \neq \xi \end{cases}$$

where $\langle \cdot, \cdot \rangle$ is defined by

$$\langle \chi, \xi \rangle = \int_{S^1} \chi(x)\bar{\xi}(x)\, dx \quad \text{for all} \quad \chi, \xi \in \widehat{S^1}.$$

In the following section, we will define the space $L^2(S^1)$ of square integrable functions on S^1. The space $L^2(S^1)$ is an inner product space with the same inner product defined in the preceding display. Therefore, the set of characters of S^1 is orthonormal in $L^2(S^1)$.

The integration by parts formula is another consequence of invariant integration. First, we prove the following lemma.

Lemma 42. Let $\int_{S^1}$ be an invariant integral on S^1 and let $f \in C^1(S^1)$. Then

$$\int_{S^1} f'(x)\, dx = 0.$$

Proof. We prove the lemma by direct computation.

$$\begin{aligned} \int_{S^1} f'(x)\, dx &= \int_{S^1} \frac{\partial}{\partial t}\Big|_{t=0} f(x+t)\, dx \\ &= \frac{d}{dt}\Big|_{t=0} \int_{S^1} f(x+t)\, dx \\ &= \frac{d}{dt}\Big|_{t=0} \int_{S^1} f(x)\, dx \\ &= 0. \end{aligned}$$

Since $f \in C^1(S^1)$, f' is bounded, so the interchange of the derivative and the integral is justified by dominated convergence. □

Remark 43. We could have proved this lemma by appealing to the fact that f is periodic:

$$\int_{S^1} f'(x)\,dx = \int_0^1 f'(x)\,dx = f(1) - f(0) = 0.$$

The previous proof, however, does not require periodicity, and thus holds in a more general context.

Corollary 44 (Integration by parts formula)**.** Let $f, g \in C^1(S^1)$ and let $\int_{S^1}$ be an invariant integral. Then

$$\int_{S^1} f'(x)g(x)\,dx = -\int_{S^1} f(g)g'(x)\,dx.$$

Proof. By the product rule of differentiation

$$(f\,g)' = f'\,g + f\,g'.$$

Applying the lemma

$$0 = \int_{S^1} (f\,g)'(x)\,dx = \int_{S^1} f'(x)g(x)\,dx + \int_{S^1} f(x)g'(x)\,dx.$$

□

Construction of the Integral on S^1

Now that we have seen the advantages of an invariant integral on S^1, we construct an invariant integral inherited from the Lebesgue integral on $\mathbb{R}$. Unfortunately, we are not in a position to show that the invariant integral is unique on S^1, but for our current purposes this construction will suffice.

The space S^1 inherits its invariant integral from the Lebesgue integral on $\mathbb{R}$. We first note that the quotient map

$$q : \mathbb{R} \longrightarrow \mathbb{R}/\mathbb{Z} = S^1$$

maps the closed unit interval $[0, 1] \subset \mathbb{R}$ surjectively onto S^1. Further, q identifies the points[3] 0 and 1, but no other points in the unit interval are identified. Therefore, we call $[0, 1]$ a *fundamental domain* for the action of $\mathbb{Z}$ on $\mathbb{R}$. We then define the integral of $f : S^1 \to \mathbb{R}$ to be

$$\int_{S^1} f(x)\,dx = \int_0^1 (q^* f)(x)\,dx.$$

In order for this integral to be well-defined, we must require $q^* f$ to be measurable. Therefore, we say that f is measurable on S^1 if $q^* f$ is measurable on $[0, 1]$ with respect to Lebesgue measure.

We still must show that $\int_{S^1}$ is translation invariant. We will do so by using the fact that the Lebesgue integral on $\mathbb{R}$ is translation invariant.

[3] That is, $q(0) = q(1)$.

Proposition 45. The integral $\int_{S^1}$ is translation invariant.

Proof. Let $y \in [0,1]$. Then we compute the integral

$$\begin{aligned}
\int_{S^1} f(x+y)\,dx &= \int_0^1 (q^*f)(x+y)\,dx \\
&= \int_{-y}^{1-y} (q^*f)(x)\,dx \\
&= \int_{-y}^{0} (q^*f)(x)\,dx + \int_0^{1-y} (q^*f)(x)\,dx \\
&= \int_{1-y}^{1} (q^*f)(x)\,dx + \int_0^{1-y} (q^*f)(x)\,dx \qquad (q^*f \text{ is periodic}) \\
&= \int_0^1 (q^*f)(x)\,dx \\
&= \int_{S^1} f(x)\,dx.
\end{aligned}$$

□

2.3 Hilbert Space Theory of Fourier Series

In this section we use the integral on S^1 to describe the space of square-integrable functions, $L^2(S^1)$. The space $L^2(S^1)$ is a Hilbert space, and is a natural context in which to study Fourier series. In fact, in the following section we will see that set of functions $\psi_n(x) = \exp(2\pi i n x)$ form an orthonormal basis for $L^2(S^1)$.

Square Integrable Functions on S^1

Let $f : S^1 \to \mathbb{C}$ be a measurable function. Then we say that f is *square integrable* if the integral

$$\int_{S^1} |f(x)|^2 \, dx$$

converges. If f and g are both square integrable functions on the circle, we say that they are *equivalent* if

$$\int_{S^1} |f(x) - g(x)|^2 \, dx = 0.$$

This is to say that $f = g$ *almost everywhere*. The space $L^2(S^1)$ is the space of all square integrable functions on S^1 modulo almost everywhere equality.

Remark 46. Two functions $f, g \in L^2(S^1)$ can differ on any set of measure zero, yet still represent the same function in $L^2(S^1)$. This suggests that pointwise values are not the salient feature of L^2 functions. Rather, L^2 functions are equivalence classes of functions. We can recover concepts such as continuity (which do rely on the point-wise values of a function) if we say that $f \in L^2(S^1)$ is continuous if there exists a continuous function $g : S^1 \to \mathbb{C}$ such that f is equivalent to g.

Since L^2 functions can be added and scaled, $L^2(S^1)$ is a vector space. Further, we define an inner product on $L^2(S^1)$ to be

$$\langle f, g \rangle = \int_{S^1} f(x)\overline{g}(x)\, dx.$$

It is easily verified that this is an inner product[4] on $L^2(S^1)$, so $L^2(S^1)$ is a *pre*-Hilbert (or inner product) space. The space $L^2(S^1)$ has a natural norm

$$\|f\|_{L^2} = \sqrt{\langle f, f \rangle}.$$

As before, this norm gives $L^2(S^1)$ a natural metric and topology. We wish to show that $L^2(S^1)$ is a Hilbert space, which is to say that it is complete with respect to its natural metric.

Proposition 47. The space $L^2(S^1)$ is a Hilbert space.

This is a standard result from measure theory and will not be proven here. See, for example, [10].

Remark 48. So far, we have encountered three important classifications of *complete* topological vector spaces. Frechet spaces are (invariantly) metrizable spaces; Banach spaces are normed spaces; Hilbert spaces are inner product spaces. Further, we have the implications between these spaces

$$\text{Hilbert} \implies \text{Banach} \implies \text{Frechet}.$$

Hilbert and Banach spaces can be given the prefix *pre-* if they are not known to be complete with respect to their natural metrics.

Fourier Series in $L^2(S^1)$

Our work in the previous section showed that the exponential functions

$$\psi_n(x) = \exp(2\pi i n x) \quad \text{for all} \quad n \in \mathbb{Z},\ x \in S^1$$

are an orthonormal set in $L^2(S^1)$. We would like to show that the ψ_n form a an orthonormal *basis* for $L^2(S^1)$. Proving this fact takes some work, and will be the focus of the following section. For now, we can still prove some useful pointwise convergence properties of Fourier series.

[4] One must verify that for all $f, g, h \in L^2(S^1)$ and $c \in \mathbb{C}$:

1. $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$
2. $\langle cf, g \rangle = c \langle f, g \rangle$
3. $\overline{\langle f, g \rangle} = \langle g, f \rangle$
4. $\langle f, f \rangle > 0$ for all $f \nsim 0$.

Definition 49. Let $f \in L^2(S^1)$. Then the n^{th} *Fourier coefficient* of f is

$$\langle f, \psi_n \rangle = \int_{S^1} f(x)\overline{\psi}_n(x)\, dx.$$

The *Fourier expansion* of f is

$$f(x) \sim \sum_{n \in \mathbb{Z}} \langle f, \psi_n \rangle\, \psi_n(x).$$

We do not write equality in the previous display, because it is not clear (or necessarily true) that the Fourier series of f converges pointwise to f.

We will need the following proposition in order to show that the functions ψ_n form a basis for $L^2(S^1)$.

Proposition 50 (Riemann-Lebesgue lemma)**.** Let $f \in L^2(S^1)$. Then

$$\lim_{|n| \to \infty} \langle f, \psi_n \rangle = 0.$$

Proof. We use *Bessel's inequality*[5], which states that for any orthonormal set $S = \{s_n\}$ in a Hilbert space

$$\sum_{n \in \mathbb{Z}} |\langle f, s_n \rangle|^2 \leq \|f\|_{L^2}.$$

In the current context, this means that

$$\sum_{n \in \mathbb{Z}} |\langle f, \psi_n \rangle|^2 \leq \|f\|_{L^2}.$$

Since the right hand side is finite, the sum on the left converges. Therefore

$$\lim_{|n| \to \infty} \langle f, \psi_n \rangle = 0.$$

□

Pointwise Convergence of Fourier Series

In the study of distribution theory, we are most concerned with convergence of Fourier series with respect to the L^2 metric. Nonetheless, it is useful to know when the Fourier series converges pointwise to its limit. We expect that L^2 convergence does not generally imply pointwise convergence, because L^2 functions are equivalence classes of functions. It is less intuitive (but also true) that pointwise convergence does not imply L^2 convergence.

[5]See [10] or [11].

Proposition 51. Let $f : S^1 \to \mathbb{C}$ be piecewise continuous. Let $x_0 \in S^1$ be a point such that f is continuous and has left and right derivatives (even if they do not agree). Then the Fourier series of f evaluated at x_0 converges to $f(x_0)$. That is,

$$f(x_0) = \sum_{n \in \mathbb{Z}} \langle f, \psi_n \rangle \, \psi_n(x_0).$$

Proof. We first make some simplifications in order to make the notation cleaner. By replacing $f(x)$ with $f(x) - f(x_0)$, we may assume without loss of generality that $f(x_0) = 0$. Further, since integration is translation invariant,

$$\begin{aligned} \int_{S^1} f(x + x_0)\overline{\psi}_n(x)\, dx &= \int_{S^1} f(x)\overline{\psi}_n(x - x_0)\, dx \\ &= \psi_n(x_0) \int_{S^1} f(x)\overline{\psi}(x)\, dx \\ &= \psi_n(x_0) \langle f, \psi_n \rangle . \end{aligned}$$

The first expression is the n-th term of the Fourier series for $f(x + x_0)$ evaluated at $x = 0$, while the final expression is the n-th term of the Fourier series for $f(x)$ evaluated at x_0. Since these quantities are equal, we can take $x_0 = 0$ without loss of generality.

We write the partial sum of the Fourier series of f evaluated at zero

$$\begin{aligned} \sum_{-M \leq n < N} \langle f, \psi_n \rangle &= \sum_{-M \leq n < N} \int_{S^1} f(x)\overline{\psi}_n(x)\, dx \\ &= \int_{S^1} f(x) \sum_{-M \leq n < N} \overline{\psi}_n(x)\, dx. \end{aligned}$$

Since the sum is finite, we can interchange the sum and the integral. We can write the inner sum

$$\begin{aligned} \sum_{-M \leq n < N} \overline{\psi}_n(x) &= \sum_{-M \leq n < N} (\psi_{-1})^n \\ &= \frac{\overline{\psi}_N(x) - \overline{\psi}_{-M}(x)}{\psi_{-1}(x) - 1}. \end{aligned}$$

The final equality holds by the geometric sum formula. Therefore

$$\begin{aligned} \sum_{-M \leq n < N} \langle f, \psi_n \rangle &= \int_{S^1} \frac{f(x)}{\psi_{-1}(x) - 1} \left(\overline{\psi}_N(x) - \overline{\psi}_{-M}(x) \right) dx \\ &= \left\langle \frac{f}{\psi_{-1} - 1}, \psi_N \right\rangle - \left\langle \frac{f}{\psi_{-1} - 1}, \psi_{-M} \right\rangle . \end{aligned}$$

If we can show that the function $f/(\psi_{-1} - 1)$ is in $L^2(S^1)$, then the Riemann-Lebesgue lemma will give us

$$\sum_{n \in \mathbb{Z}} \langle f, \psi_n \rangle = \lim_{N \to \infty} \left\langle \frac{f}{\psi_{-1} - 1}, \psi_N \right\rangle - \lim_{M \to \infty} \left\langle \frac{f}{\psi_{-1} - 1}, \psi_{-M} \right\rangle = 0$$

which is exactly what we want to show. Away from the point $x_0 = 0$, the function $f/(\psi_{-1} - 1)$ is well-defined and piecewise continuous, hence locally square-integrable. To show that $f/(\psi_{-1} - 1) \in L^2(S^1)$, we must examine the function's behavior around $x_0 = 0$. Since $x_0 = 0$ and $f(x_0) = 0$, we have

$$\frac{f(x)}{\psi_{-1}(x) - 1} = \frac{f(x) - f(x_0)}{e^{-2\pi i x} - e^{-2\pi i x_0}} \cdot \frac{x - x_0}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} \cdot \frac{x - x_0}{e^{-2\pi i x} - e^{-2\pi i x_0}}.$$

Since the right and left derivatives of $f(x_0)$ exist, $f(x)/(\psi_{-1}(x) - 1)$ is bounded in neighborhoods of $x_0 = 0$. Therefore, $f(x)/\psi_{-1}(x) - 1$ is bounded for all $x \in S^1$ and piecewise continuous except possibly at $x_0 = 0$. By the Riemann-Lebesgue lemma, we then have

$$\sum_{n \in \mathbb{Z}} \langle f, \psi_n \rangle \, \psi_n(x_0) = 0 = f(x_0).$$

□

The following corollary is not the general result that we need in the context of distribution theory, but it is a useful result nonetheless.

Corollary 52. Let $f \in C^1(S^1)$. Then

$$f(x) = \sum_{n \in \mathbb{Z}} \langle f, \psi_n \rangle \, \psi_n(x) \quad \text{for all} \quad x \in S^1.$$

2.4 Completeness of Fourier Series in $L^2(S^1)$

The aim of this section is to prove that the functions $\psi_n(x) = \exp(2\pi i n x)$ for $n \in \mathbb{Z}$ form a basis for $L^2(S^1)$. The proof is somewhat roundabout, so we will preview our reasoning.

Since $\{\psi_n\}$ is an orthonormal set and $L^2(S^1)$ is a Hilbert space, it suffices to show that trigonometric polynomials[6] are dense in $L^2(S^1)$ with respect to the L^2 topology. We will show

1. $C^0(S^1)$ is dense in $L^2(S^1)$ with respect to the L^2 topology
2. The C^0 topology is at least as fine as the L^2 topology
3. Trigonometric polynomials are dense in $C^0(S^1)$ with respect to the C^0 topology.

These three facts together prove that $\{\psi_n\}$ is an orthonormal basis for $L^2(S^1)$. By 1, it suffices to show that trigonometric polynomials are dense in $C^0(S^1)$ with respect to the L^2 topology. By 2, density in $C^0(S^1)$ with respect to the C^0 topology implies density with respect to the L^2 topology. Then 3 implies that trigonometric polynomials are dense in $L^2(S^1)$ with respect to the L^2 topology. Hence $\{\psi_n\}$ is an orthonormal basis for $L^2(S^1)$.

[6] *Trigonometric polynomials* are finite linear combinations of the ψ_n.

Density of Continuous Functions in $L^2(S^1)$

In order to show that $C^0(S^1)$ is dense in $L^2(S^1)$, we will need the following lemma.

Lemma 53 (Urysohn)**.** Let X be a locally compact Hausdorf topological space. Let U be an open subset of X and and K a compact subset of U. Then there exists a continuous function $f : X \to \mathbb{R}$ such that

1. $0 \le f(x) \le 1 \quad$ for all $\quad x \in X$
2. $f(x) = 1 \quad$ for all $\quad x \in K$
3. $f(x) = 0 \quad$ for all $\quad x \notin U$.

The proof of Urysohn's lemma is somewhat technical and will not be given here. See [10] or [11].

The statement and proof of the following corollary require a fair amount of technical language. The result is that measurable functions can be approximated by continuous functions in any Borel measure space.[7]

Corollary 54. Let X be a topological space with a regular[8] Borel measure μ. Then $C^0(X)$ is dense in $L^2(X, \mu)$ with respect to the L^2 topology.

Proof. Let E be a measurable subset of X. Let $K_1 \subset K_2 \subset \cdots$ be a sequence of compact subsets of E such that $\lim_n \mu(K_n) = \mu(E)$, and let $U_1 \supset U_2 \supset \cdots$ be a sequence of open supersets of E such that $\lim_n \mu(U_n) = \mu(E)$. By Urysohn's lemma, define for each n a continuous function $f_n : X \to \mathbb{R}$ such that

$$f_n(x) = \begin{cases} 1 & x \in K_n \\ 0 & x \notin U_n. \end{cases}$$

By Lebesgue's dominated convergence theorem

$$f_n \longrightarrow \chi(E) \quad \text{in} \quad L^2(X, \mu)$$

where $\chi(E)$ is the characteristic function of E.

By the definition of the integral of a measurable function, simple[9] functions are dense in $L^2(X)$. Since finite linear combinations of continuous functions are continuous, $C^0(X)$ is dense in the space of simple functions with respect to the L^2 topology. Therefore, $C^0(X)$ is dense in $L^2(X)$. □

Remark 55. As a special case of this corollary $C^0(S^1)$ is dense in $L^2(S^1)$. Therefore, in order to show that trigonometric polynomials are dense in $L^2(S^1)$, it suffices to show that they are dense in $C^0(S^1)$ with respect to the L^2 topology. Soon we will show that it even suffices to show density in $C^0(S^1)$ with respect to the C^0 topology.

---

[7] Recall that a *Borel measure space* is a measure space that is also a topological space. In a Borel space, the collection of measurable sets is generated by the space's topology.

[8] Recall that a *regular* Borel measure μ is a Borel measure such that the following property holds. For every measurable set M,

$$\mu(M) = \sup_K \mu(K) = \inf_U \mu(U)$$

where K ranges over all compact sets $K \subset M$ and U ranges over all open sets $K \supset M$.

[9] Recall that a *simple function* is a finite linear combination of characteristic functions of measurable sets.

Convolution and Approximate Identities

Definition 56. Let f and g be measurable functions on S^1. Then the *convolution* of f and g (denoted $f * g$) is given by

$$(f * g)(x) = \int_{S_1} f(y)g(x-y)\,dy.$$

The space $L^2(S^1)$ is closed under convolution. Further, convolution is commutative, associative and distributes over addition. Thus $L^2(S^1, *)$ is a complex algebra. A reasonable question is, "Does $L^2(S^1, *)$ have an identity?" The identity would satisfy

$$\mathrm{id} * f = f \quad \text{for all} \quad f \in L^2(S^1).$$

That is,

$$(\mathrm{id} * f)(x) = f(x) \quad \text{almost everywhere.}$$

The identity id does not exist as a classical function, but we will later see that id has a meaningful interpretation as a distribution.

Definition 57. Let $\{\varphi_n\}$ be a sequence of functions $\varphi_n \in L^2(S^1)$ where $\varphi_n(x) \geq 0$ for all $x \in S^1$ and

$$\int_{S^1} \varphi_n(x)\,dx = 1 \quad \text{for all} \quad n \in \mathbb{Z}_{\geq 0}.$$

Then we call $\{\varphi_n\}$ an *approximate identity* if for every $\varepsilon > 0$ and $\delta > 0$ there exists $N \geq 0$ such that

$$\int_{|x|<\delta} (q^* \varphi_n)(x)\,dx > 1 - \varepsilon \quad \text{for all} \quad n > N.$$

Informally, approximate identities are sequences of functions that are increasingly tall and narrow spikes centered at 0, but each function integrates to 1.

Exercise 58. Consider the function $\varphi_1 : \mathbb{R} \to \mathbb{R}$ given by

$$\varphi_1(x) = \begin{cases} 0 & |x| \geq 1/2 \\ 2 - 4\,|x| & |x| < 1/2 \end{cases}$$

Define φ_n for $n > 1$ to be

$$\varphi_n(x) = n\varphi_1(nx) \quad \text{for all} \quad x \in \mathbb{R}.$$

By restricting the domain of the φ_n to the interval $[-1/2, 1/2]$ we can view the φ as functions on S^1. The values of the φ_n are strictly positive and each function integrates to 1. Since the φ_n are nonzero only for $|x| < 1/2n$, the sequence $\{\varphi_n\}$ is an approximate identity (see Figure 2.1).

Proposition 59. Let $f \in C^0(S^1)$ and let $\{\varphi_n\}$ be an approximate identity. Let $f_t(x) = f(x+t)$ for all $x \in S^1$ and let $f_c(x) = f(x) + c$ for all $x \in S^1$. Then

1. $(f_t * \varphi_n)(x) = (f * \varphi_n)(x + t) \quad \text{for all} \quad x \in S^1$

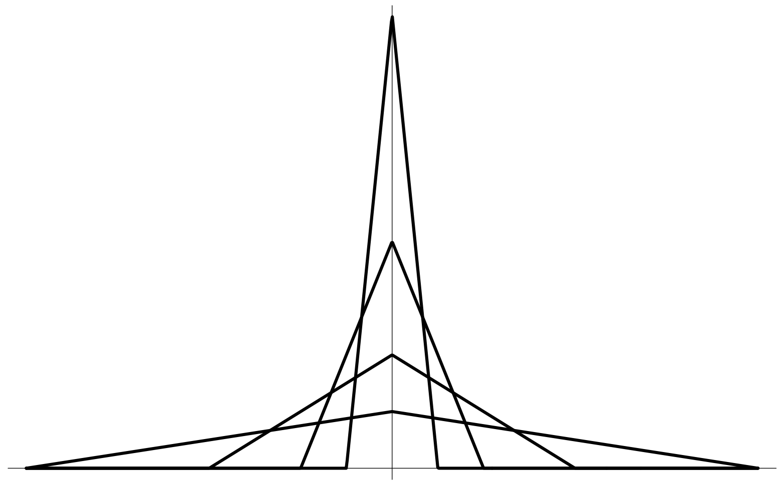


Figure 2.1: Terms of the approximate identity $\{\varphi_n\}$ for $n = 1, 2, 4, 8$. Each function integrates to 1, but the width of the spikes go to zero for large values of n.

2. $(f_c * \varphi_n)(x) = (f * \varphi_n)(x) + c \quad \text{for all} \quad x \in S^1.$

That is, convolution commutes with translation of the origin in S^1.

Proof. We verify the claim by direct computation.

$$\begin{aligned}
(f_t * \varphi_n)(x) &= \int_{S^1} f_t(y)\varphi_n(x-y)\,dy \\
&= \int_{S^1} f(y+t)\varphi_n(x-y)\,dy \\
&= \int_{S^1} f(y)\varphi_n(x+t-y)\,dy \\
&= (f * \varphi_n)(x+t).
\end{aligned}$$

Similarly, adding a constant to f commutes with convolution:

$$\begin{aligned}
(f_c * \varphi_n)(x) &= \int_{S_1} (f(y)-c)\varphi_n(x-y)\,dy \\
&= \int_{S_1} f(y)\varphi_n(x-y)\,dy - c \qquad (\varphi_n \text{ integrates to } 1) \\
&= (f * \varphi_n)(x) - c.
\end{aligned}$$

□

Proposition 60. Let $f \in C^0(S^1)$ and let $\{\varphi_n\}$ be an approximate identity. Then

$$f * \varphi_n \longrightarrow f$$

in the C^0 topology.

Remark 61. Thus the name "approximate identity" for such a sequence of functions is apt, as the functions $\{\varphi_n\}$ approximate the identity id.

Proof. We appeal to the uniform continuity of f, and hence $q^* f$. Fix $\varepsilon > 0$ and let δ be such that

$$|q^* f(x) - q^* f(y)| < \varepsilon \quad \text{for all} \quad |x - y| < \delta.$$

Let N be such that for $n \geq N$

$$\int_{|x|<\delta} (q^*\varphi)_n(x)\, dx > 1 - \varepsilon \quad \text{or equivalently} \quad \int_{\delta<|x|<1/2} (q^*\varphi_n)(x) < \varepsilon.$$

For any $x \in S^1$, consider the difference

$$\begin{aligned}(f * \varphi_n)(x) - f(x) &= \int_{S^1} f(y)\varphi_n(x-y)\, dy - f(x) \\ &= \int_{S^1} (f(y) - f(x))\varphi_n(x-y)\, dy.\end{aligned}$$

The second equality holds because φ_n integrates to 1. By Claim 59, it suffices to consider $x = 0$ and $f(x) = 0$. In order to simplify notation further, we will identify the functions φ_n and f with their pullbacks to the interval $[-1/2, 1/2]$. Then

$$\begin{aligned}\int_{-1/2}^{1/2} f(y)\varphi_n(-y)\, dy = &\int_{|y|<\delta} f(y)\varphi_n(-y)\, dy \\ &+ \int_{\delta<|y|<1/2} f(y)\varphi_n(-y)\, dy.\end{aligned}$$

For the first integral,

$$\int_{|y|<\delta} |f(y)|\, |\varphi_n(-y)|\ dy < \int_{|y|<\delta} \varepsilon\, |\varphi_n(-y)|\ dy \leq \varepsilon$$

because φ_n is real valued, nonnegative and integrates to 1. For the second integral,

$$\begin{aligned}\int_{\delta<|y|<1/2} |f(y)|\, |\varphi_n(-y)|\ dy &\leq \int_{\delta<|y|<1/2} 2\, \|f\|_{C^0}\, |\varphi_n(-y)|\ dy \\ &< 2\, \|f\|_{C^0}\, \varepsilon.\end{aligned}$$

Since the function f is uniformly continuous,

$$(f * \varphi_n)(x) - f(x) = \int_{S^1} (f(y) - f(x))\varphi_n(x-y)\, dy < \varepsilon(1 + 2\, \|f\|_{C^0})$$

for all $x \in S^1$. Therefore

$$\|(f * \varphi_n) - f\|_{C^0} < \varepsilon(1 + 2\, \|f\|_{C^0}).$$

Since the above inequality holds for all $\varepsilon > 0$, $f * \varphi_n$ approaches f in the C^0 topology. □

Proposition 62. The C^0 topology is at least as fine as the L^2 topology on $C^0(S^1)$.

Proof. We must show that for every open set U_{L^2} in the L^2 topology, there is an open set U_{C^0} such that $U_{C^0} \subset U_{L^2}$. Since both topologies are generated by norms, it suffices to show that

$$\|f\|_{L^2} \leq \|f\|_{C^0} \quad \text{for all} \quad f \in C^0(S^1).$$

To this end, let $f \in C^0(S^1)$. Then

$$\|f\|_{L^2} = \left(\int_{S^1} |f(x)|^2 \, dx \right)^{1/2} \leq \sup_x |f(x)| \left(\int_{S^1} dx \right)^{1/2} = \|f\|_{C^0}.$$

□

Remark 63. In the remark on Page 37, we noted that the density of trigonometric polynomials in $C^0(S^1)$ with respect to the L^2 topology implies that such sums are dense in $L^2(S^1)$. Since we have shown that the C^0 topology is at least as fine as the L^2 topology, it suffices to show that finite linear combinations are dense in $C^0(S^1)$ with respect to the C^0 topology.

Exercise 64 (eye candy)**.** We can see (Figure 2.2) that convolving the f_n with a function

$$g(x) = \sin(2\pi x) + \sin(4\pi x) + \cos(6\pi x).$$

does a good job of approximating g even for modest values of n .

Convolution against a particular function is an example of an *integral operator*. Generally, an integral operator is an operator T defined by

$$(Tf)(x) = \int K(x,y) f(y) \, dy$$

where K is known as the *kernel* of the operator. Approximate identities are sequences of kernels φ_n where the sequence of associated integral operators

$$T_{\varphi_n} : f(x) \longmapsto \int_{S^1} f(y) \varphi_n(x-y) \, dy$$

converge to the identity operator.

Completeness

Let $f \in C^0(S^1)$. The truncated Fourier series for f is given by

$$\begin{aligned} \sum_{|n| \leq n} \langle f, \psi_n \rangle \, \psi_n(x) &= \int_{S^1} f(y) \sum_{|n| \leq n} \overline{\psi}_k(y) \psi_k(x) \, dy \\ &= \int_{S^1} f(y) \frac{\psi_{n+1}(x-y) - \psi_{-n}(x-y)}{\psi_1(x-y) - 1} \, dy. \end{aligned}$$

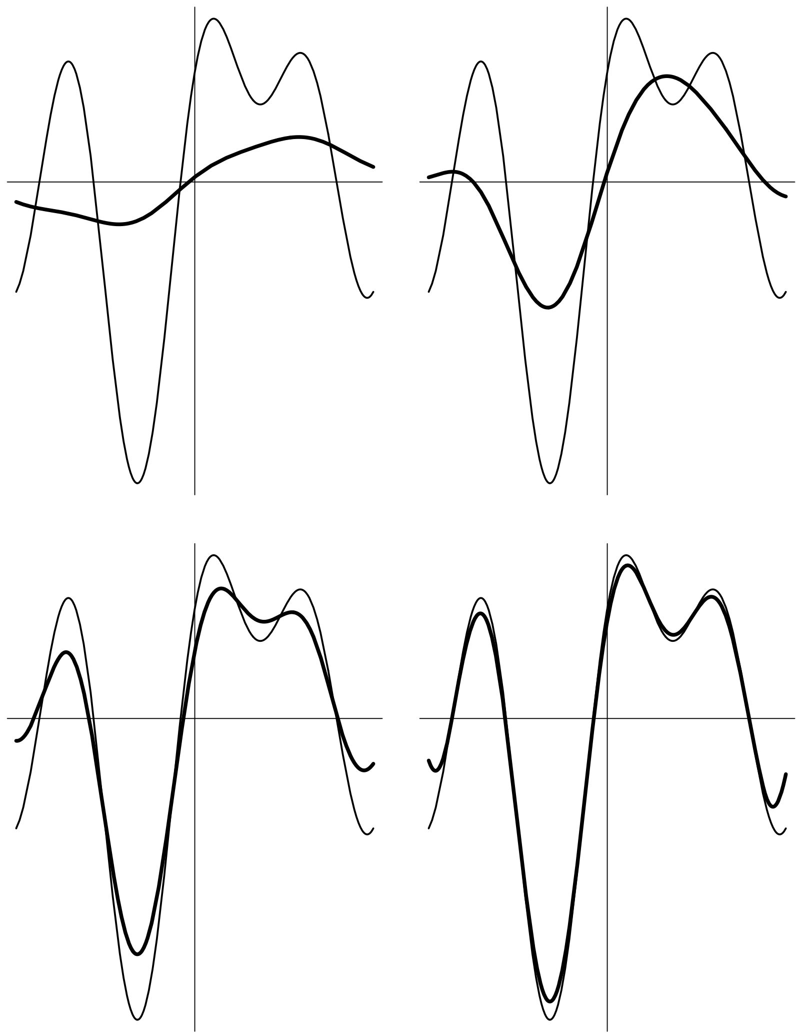


Figure 2.2: Convolutions of the functions φ_n against the function $f(x) = \sin 2\pi x + \sin 4\pi x + \cos 6\pi x$ for $n = 1, 2, 4, 8$. The convolutions are given by the thick lines, while the original function g is indicated with a thin line. Note that the convolutions better approximate f for larger values of n.

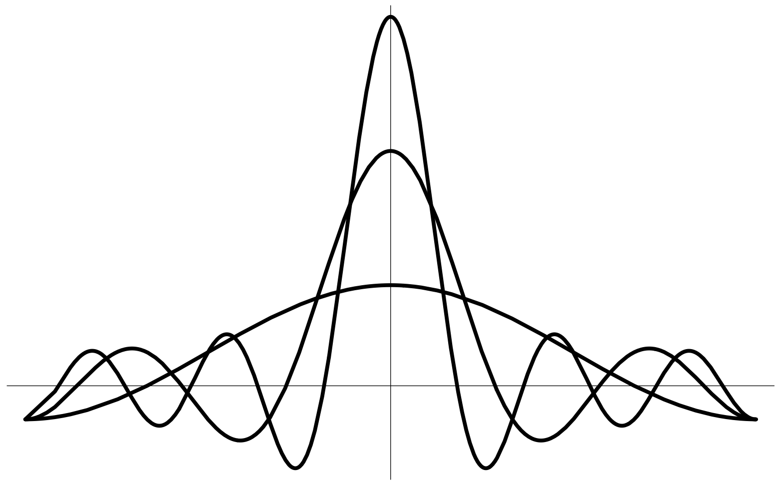


Figure 2.3: The Dirichlet kernels $K_n(x)$ for $n = 1, 3, 5$. Although the kernels bunch up around $x = 0$, they can not constitute an approximate identity because they take on negative values.

where the final equality holds by the geometric sum formula. We define the *Dirichlet kernel* to be

$$K_n(x) = \frac{\psi_{n+1}(x) - \psi_{-n}(x)}{\psi_1(x) - 1} = \frac{\psi_{n+1/2}(x) - \psi_{-n-1/2}(x)}{\psi_{1/2}(x) - \psi_{-1/2}(x)} = \frac{\sin 2\pi(n+1/2)x}{\sin \pi x}.$$

Then the original truncated Fourier series is given by

$$\sum_{|k|\leq n} \langle f, \psi_k \rangle \, \psi_k(x) = \int_{S^1} f(y) \frac{\sin 2\pi(n+1/2)(x-y)}{\sin \pi(x-y)} \, dy = (f * K_n)(x).$$

Certainly, in order to show that the Fourier series of f converges to f, it would suffice to show that K_n is an approximate identity. Unfortunately, K_n is not non-negative, so it cannot be an approximate identity (see Figure 2.4). Nonetheless, the fact that the K_n are real valued and that they obtain an maximum around $x = 0$ is promising. We consider the squares of the Dirichlet kernels K_n^2. The squares of the Dirichlet kernels will certainly be nonnegative, but in order to be approximate identities, they must be normalized.

$$\int_{S^1} K_n^2(x) \, dx = \int_{S^1} (\psi_{-n}(x) + \cdots + \psi_n(x))^2 \, dx.$$

Since the sum $\psi_{-n}(x) + \cdots + \psi_n(x)$ has $2n+1$ terms, ψ_0 will appear $2n+1$ times in the integrand of the above expression. All of the other ψ_i for $i \neq 0$ integrate to zero, so

$$\int_{S^1} K_n^2(x) \, dx = 2n + 1.$$

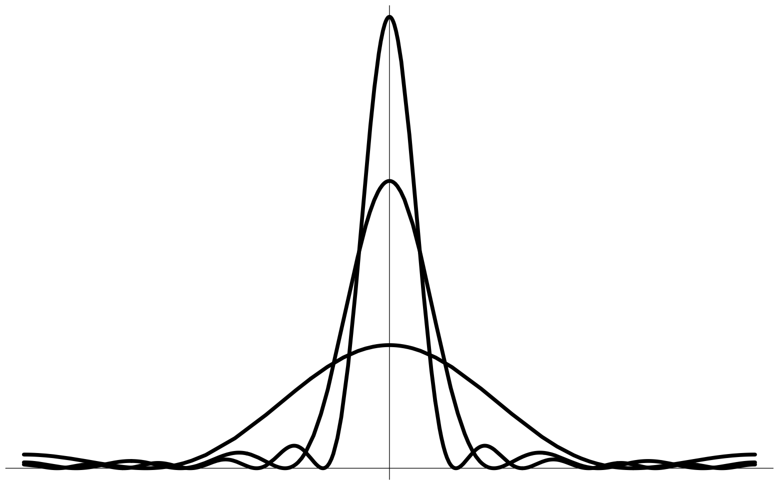


Figure 2.4: The Fejer kernels for $n = 1, 3, 5$. The Fejer kernels are the squares of the Dirichlet kernels, normalized so that they integrate to 1. The Fejer kernels do constitute an approximate identity.

We now define the *Fejer kernels* (see Figure 2.4) to be

$$F_n(x) = \frac{K_n^2(x)}{2n+1},$$

so that each of the Fejer kernels integrates to 1.

We would like to show that the functions F_n are an approximate identity. To these ends, we note that

$$F_n(x) = \frac{\sin^2(\pi(n+1/2)x)}{(2n+1)\sin^2(\pi x)} \leq \frac{1}{(2n+1)x^2}.$$

The inequality holds because $|x| \leq |\sin \pi x|$ for $|x| \leq 1/2$. Then for $|x| \geq n^{-1/3}$,

$$0 \leq F_n(x) \leq \frac{1}{(2n+1)x^2} \leq \frac{1}{(2n+1)n^{-2/3}} \leq n^{-1/3}.$$

For $n^{-1/3} \leq |x| \leq 1/2$ we have

$$\begin{aligned}
\int_{n^{-1/3} \leq |x| \leq 1/2} (q^* F_n)(x)\, dx &\leq \frac{1}{n^{1/3}} \int_{n^{-1/3} \leq |x| \leq 1/2} dx \\
&= 2 \cdot \left(\frac{1}{2} - \frac{1}{n^{1/3}} \right) \cdot \frac{1}{n^{1/3}} \\
&\leq \frac{1}{n^{1/3}}.
\end{aligned}$$

Since $n^{-1/3} \to 0$ for large n, the Fejer kernels evidently are approximate identities: given $\delta, \varepsilon > 0$, take $n > \max(\varepsilon^{-3}, \delta^{-3})$. Then

$$\int_{|x|<\delta} (q^* F_n)(x)\, dx > 1 - \varepsilon.$$

Remark 65. For $f \in \mathbb{C}^0(S^1)$, it is important to note that although

$$f * F_n \longrightarrow f \quad \text{in} \quad C^0(S^1)$$

and each F_n is a trigonometric polynomial, the Fourier series of f may *not* converge to f in the C^0 norm. In particular, $f * F_n$ is not a truncated Fourier series for f.

We now have all all the pieces to prove the main result for this section.

Theorem 66. The exponentials ψ_n form an orthonormal basis for $L^2(S^1)$.

Proof. We have already shown that the ψ_n are orthonormal. What remains to be shown is that the ψ_n span $L^2(S^1)$. To these ends it suffices to show that finite linear combinations of the ψ_n are dense in $L^2(S^1)$. By Corollary 54, $C^0(S^1)$ is dense in $L^2(S^1)$, and by the remarks following Proposition 60 the C^0 topology is at least as fine as the L^2 topology. Therefore it suffices to show that finite linear combinations of the ψ_n are dense in $C^0(S^1)$ with respect to the C^0 topology.

To this end, we use Proposition 60. We showed that the Fejer kernels F_n are an approximate identity in $C^0(S^1)$. The Fejer kernels are finite linear combinations of the ψ_n,

$$F_n(x) = \sum_{|k|\le n} c_{n,k}\, \psi_k(x) \quad \text{for all} \quad x \in S^1.$$

Let $f \in C^0(S^1)$. Then the convolution of f with F_n is given by

$$\begin{aligned}(f * F_n)(x) &= \int_{S^1} f(y) F_n(x-y)\, dy \\ &= \int_{S^1} f(y) \sum_{|k|\le n} c_{n,k}\, \psi_k(x-y)\, dy \\ &= \sum_{|k|\le n} \psi_k(x)\, c_{n,k} \int_{S^1} f(y) \overline{\psi}_k(y)\, dy.\end{aligned}$$

The final sum is a finite linear combination of the ψ_k. Since F_n is an approximate identity, $f * F_n$ converges to f in the C^0 topology. Therefore, the sum

$$\sum_{|k|\le n} \psi_k(x)\, c_{n,k} \int_{S^1} f(y) \overline{\psi}_k(y)\, dy$$

converges to f for large n. Thus finite linear combinations of the ψ_n are dense in $C^0(S^1)$ with respect to the C^0 topology. □

Since the exponentials ψ_n are an orthonormal basis for $L^2(S^1)$, for any $f \in L^2(S^1)$ we have L^2 equality

$$f = \sum_{n\in\mathbb{Z}} \langle f, \psi_n \rangle \, \psi_n \quad \text{in } L^2(S^1).$$

By Plancherel's theorem, [10]

$$\|f\|_{L^2}^2 = \sum_{n\in\mathbb{Z}} |\langle f, \psi_n \rangle|^2 \,.$$

[10]Plancherel's theorem states that for any Hilbert space H with orthonormal basis $\{b_n\}$

$$\|h\|_H^2 = \langle h, h \rangle = \sum_n |\langle h, b_n \rangle|^2 \,.$$

See, for example, [10] or [11].

3

Sobolev Embeddings

The aim of this chapter is to refine our description of the space of distributions on S^1. In order to accomplish this, we use Sobolev spaces. Sobolev spaces are generalizations of the space $L^2(S^1)$. Like $L^2(S^1)$, the Sobolev spaces are Hilbert spaces and $\{\psi_n\}$ is a basis for every Sobolev space. Further, the Sobolev spaces are cofinal *with the spaces $C^k(S^1)$, so we can describe smooth functions and distributions entirely in terms of the Sobolev spaces. This allows us to represent every distribution by a Fourier series.*

3.1 The Sobolev Inequality and Sobolev Embedding

In this section, we introduce the Sobolev spaces $H_s(S^1)$ for $s \geq 0$. Before defining $H_s(S^1)$, we derive the Sobolev inequality from which we obtain the Sobolev norm. The space $H_s(S^1)$ (for $s \geq 0$) is then the space of all functions in $L^2(S^1)$ for which the Sobolev norm is finite. The Sobolev norm generates a topology on $H_s(S^1)$ which is at least as fine as the topology on $L^2(S^1)$. Over the course of the section, we establish some basic facts about $H_s(S^1)$:

- $H_s(S^1)$ is a Hilbert space
- For $k \in \mathbb{Z}_{\geq 0}$ and $s > k + 1/2$, $H_s(S^1)$ embeds continuously into $C^k(S^1)$
- For $k \in \mathbb{Z}_{\geq 0}$ and $k > s + 1/2$, $C^k(S^1)$ embeds continuously into $H_s(S^1)$.

These properties of Sobolev spaces allow us to describe the space of distributions as a colimit of Sobolev spaces. Such a description of distributions on S^1 is the topic of the following section.

The Sobolev Inequality

Let $f \in C^k(S^1)$ for some $k \geq 0$. We previously defined the C^k-norm of f to be

$$\|f\|_{C^k} = \sum_{n=0}^{k} \left\| f^{(n)} \right\|_{C^0} = \sum_{n=0}^{k} \sup_{x \in S^1} \left| f^{(n)}(x) \right|.$$

We give an equivalent[1] formulation of this norm. We define

$$\|f\|'_{C^k} = \sup_{0 \leq n \leq k} \sup_{x \in S^1} \left| f^{(n)}(x) \right|.$$

Then

$$\|f\|'_{C^k} \leq \|f\|_{C^k} \quad \text{for all} \quad f \in C^k(S^1),\ k \in \mathbb{Z}_{\geq 0}.$$

Therefore, the topology generated by $\|\cdot\|_{C^k}$ is at least as fine as the topology generated by $\|\cdot\|'_{C^k}$. However,

$$\|f\|_{C^k} = \sum_{n=0}^{k} \sup_{x \in S^1} \left| f^{(n)}(x) \right| \leq k \sup_{0 \leq n \leq k} \sup_{x \in S^1} \left| f^{(n)}(x) \right| = k \|f\|'_{C^k}$$

for all $f \in C^k(S^1)$. Therefore the topology generated by $\|\cdot\|'_{C^k}$ is at least as fine as the topology generated by $\|\cdot\|_{C^k}$. Consequently, the two topologies are the same. We will therefore use the norms $\|\cdot\|_{C^k}$ and $\|\cdot\|'_{C^k}$ interchangeably, as they generate identical topologies.

We consider the C^k-norm of a trigonometric polynomial

$$\left\| \sum_{|n| \leq N} c_n \psi_n \right\|_{C^k} = \sup_{0 \leq j \leq k} \sup_{x \in S^1} \left| (2\pi i)^j \sum_{|n| \leq N} c_n n^j \psi_n(x) \right| \leq (2\pi)^k \sum_{|n| \leq N} |c_n| \, (1+n^2)^{k/2}.$$

Note that $(1+n^2)^{k/2}$ is asymptotically equal to n^k, although we prefer the former because it is strictly positive for all $n \in \mathbb{Z}$. For any $s \in \mathbb{R}$,

$$\begin{aligned} \left\| \sum_{|n| \leq N} c_n \psi_n \right\|_{C^k} &\leq (2\pi)^k \sum_{|n| \leq N} |c_n| \, (1+n^2)^{k/2} \\ &= (2\pi)^k \sum_{|n| \leq N} |c_n| \, (1+n^2)^{s/2} \frac{1}{(1+n^2)^{(s-k)/2}} \\ &\leq (2\pi)^k \left(\sum_{|n| \leq N} |c_n|^2 \, (1+n^2)^s \right)^{1/2} \cdot \left(\sum_{|n| \leq N} \frac{1}{(1+n^2)^{(s-k)}} \right)^{1/2}. \end{aligned}$$

[1] Two norms are said to be *equivalent* if they generate the same topology.

The final inequality holds by the Cauchy-Schwarz[2] inequality. For $s > k + 1/2$ the sum

$$\sum_{|n|\leq N} \frac{1}{(1+n^2)^{(s-k)}}$$

converges as $N \to \infty$. We therefore define the constant

$$\omega_{s,k} = (2\pi)^k \sum_{n\in\mathbb{Z}} \frac{1}{(1+n^2)^{(s-k)}}.$$

Then we are left with the *Sobolev inequality*

$$\left\| \sum_{|n|\leq N} c_n \psi_n \right\|_{C^k} \leq \omega_{s,k} \left(\sum_{|n|\leq N} |c_n|^2 (1+n^2)^s \right)^{1/2}.$$

Definition 67. For a Fourier series $\sum_{n\in\mathbb{Z}} c_n\psi_n$, we define the s^{th}*Sobolev norm* to be

$$\left\| \sum_{n\in\mathbb{Z}} c_n \psi_n \right\|_{H_s} = \left(\sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^s \right)^{1/2}.$$

At the present time, we restrict our attention to the case where $s \geq 0$. If the H_s norm is finite for $s \geq 0$, then the Fourier series converges in $L^2(S^1)$.

Proposition 68. Let $k \in \mathbb{Z}_{\geq 0}$ and let $s \in \mathbb{R}$ with $s > k + 1/2$. Let f be represented by the Fourier series

$$f = \sum_{n\in\mathbb{Z}} c_n \psi_n$$

with

$$\|f\|_{H_s} < \infty.$$

Then the Fourier series for f converges in $C^k(S^1)$ with respect to the C^k topology.

Proof. By the Sobolev inequality,

$$\left\| \sum_{n\in\mathbb{Z}} c_n \psi_n \right\|_{C^k} \leq \omega_{s,k} \left\| \sum_{n\in\mathbb{Z}} c_n \psi_n \right\|_{H_s} < \infty$$

Therefore the H_s topology is at least as fine as the C^k topology. Since

$$\left\| \sum_{n\in\mathbb{Z}} c_n \psi_n \right\|_{H_s}$$

[2]Recall that the Cauchy-Schwarz inequality states that for elements x and y in an inner product space, $|\langle x, y\rangle| \leq \langle x, x\rangle^{1/2} \cdot \langle y, y\rangle^{1/2}$.

is finite, the partial sums

$$\sum_{|n|\leq N} c_n\psi_n$$

form a Cauchy sequence with respect to the H_s topology as $N \to \infty$. Since the H_s topology is at least as fine as the C^k topology, the partial sums are therefore Cauchy in C^k. Finally, $C^k(S^1)$ is complete, so the series converges in C^k. □

The Sobolev Space H_s

Definition 69. Let $s \in \mathbb{R}_{\geq 0}$. Then we define the s^{th} *Sobolev space* to be

$$H_s(S^1) = \left\{ f \in L^2(S^1) \, : \, \sum_{n\in\mathbb{Z}} |\langle f, \psi_n\rangle_{L^2}|^2 \, (1+n^2)^s < \infty \right\}$$

Remark 70. Eventually, we will define $H_s(S^1)$ for all $s \in \mathbb{R}$. In fact, we will see that $H_{-s}(S^1)$ is the dual space of $H_s(S^1)$. However, for $s < 0$, the requirement that $f \in L^2(S^1)$ is stronger than the requirement that the H_s norm be finite, so we cannot use the same definition as above for negative s.

The space $H_s(S^1)$ is an inner product space with

$$\left\langle \sum_n a_n\psi_n, \sum_n b_n\psi_n \right\rangle_{H_s} = \sum_n a_n\bar{b}_n(1+n^2)^s.$$

It is easily verified that this is indeed an inner product. Further, this inner product is compatible with the H_s norm in the sense that

$$\|f\|_{H_s} = \sqrt{\langle f, f\rangle_{H_s}} \quad \text{for all} \quad f \in H_s(S^1).$$

Remark 71. From the definition of $\langle \cdot, \cdot \rangle_{H_s}$, we can see that the functions ψ_n are orthogonal in $H_s(S^1)$, although not orthonormal unless $s = 0$ (in which case $H_0(S^1) = L^2(S^1)$).

Proposition 72. The space of trigonometric polynomials is dense in $H_s(S^1)$ for $s \geq 0$ with respect to the H_s topology.

Proof. Let $f \in H_s(S^1)$ with $f = \sum_{n\in\mathbb{Z}} c_n\psi_n$. We will show that there exists a sequence of trigonometric polynomials that converge to f. Let

$$f_j = \sum_{|n|\leq j} c_n\psi_n \quad \text{for all} \quad j \in \mathbb{Z}_{\geq 0}.$$

We claim that $f_j \to f$ in the H_s topology. To see this, we consider the square norm of the difference

$$\|f - f_j\|_{H_s}^2 = \sum_{|n|>j} |c_n|^2 \, (1+n^2)^s.$$

The equality holds because f_j is just the truncated Fourier series for f. By the definition of H_s, the sum

$$\sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^s$$

is finite, so

$$\|f - f_j\|_{H_s}^2 = \sum_{|n|>j} |c_n|^2 (1+n^2)^s \xrightarrow{j\to\infty} 0.$$

Therefore f_j converges to f in the H_s topology. Since the choice of f was arbitrary, trigonometric polynomials are dense in H_s. □

Remark 73. Since the functions ψ_n are orthogonal in $H_s(S^1)$ and finite linear combinations of the ψ_n are dense, the ψ_n form a basis for $H_s(S^1)$. This argument is analogous to the proof of Theorem 66.

Proposition 74. The s^{th} Sobolev space $H_s(S^1)$ with $s \geq 0$ is a Hilbert space. Further $H_s(S^1)$ consists of *all* Fourier series $f = \sum_{n\in\mathbb{Z}} c_n\psi_n$ such that the coefficients c_n satisfy $\sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^s < \infty$.

Remark 75. The second claim is subtly different from the definition of $H_s(S^1)$ in that the definition of $H_s(S^1)$ requires that $f \in L^2(S^1)$. Here we make no such assumption. We will show that

$$\sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^s < \infty \implies f \in H_s(S^1).$$

The net result is that we can use the function f and its Fourier series interchangeably.

Proof. We will show the second claim first. Let $c_n \in \mathbb{C}$ for all $n \in \mathbb{Z}$ such that

$$\sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^s < \infty.$$

Since $s \geq 0$, we certainly have

$$\sum_{n\in\mathbb{Z}} |c_n|^2 < \infty$$

which implies that

$$f = \sum_{n\in\mathbb{Z}} c_n\psi_n \in L^2(S^1).$$

By the hypothesis and the definition of $H_s(S^1)$, $f \in H_s(S^1)$, which proves the second claim.

The space $H_s(S^1)$ is an inner product (pre-Hilbert) space, so we must show that it is complete. To this end, let $\{f_j\}$ be a Cauchy sequence of functions in $H_s(S^1)$. That is

$$\|f_j - f_k\|_{H_s} \xrightarrow{j,k\to\infty} 0.$$

Suppose each f_j is represented by the Fourier series

$$f_j = \sum_{n\in\mathbb{Z}} c_{j,n}\psi_n \quad \text{for} \quad c_{j,n} \in \mathbb{C}.$$

We need to show that there exists a function $f \in H_s(S^1)$ such that $\{f_j\} \to f$. Since the sequence $\{f_j\}$ is Cauchy in $H_s(S^1)$, the coefficients $\{c_{j,n}\}$ form a Cauchy sequence for every fixed $n \in \mathbb{Z}$. Then

$$c_{j,n} \xrightarrow{j \to \infty} c_n \quad \text{for some} \quad c_n \in \mathbb{C}.$$

Since this is true for every $n \in \mathbb{Z}$, f_j converges to some formal Fourier series f.

Now we will show that $f \in H_s(S^1)$. Consider the H_s norm of f.

$$\begin{aligned} \|f\|_{H_s} &= \|f - f_j + f_j\|_{H_s} \\ &\leq \|f - f_j\|_{H_s} + \|f_j\|_{H_s} \\ &< 1 + \|f_j\|_{H_s} && \text{(for sufficiently large } j) \\ &< \infty && (f_j \in H_s(S^1), \text{ so } H_s \text{ norm is finite}). \end{aligned}$$

Since the H_s norm of f is finite, $f \in H_s(S^1)$. Therefore every Cauchy sequence of functions $\{f_i\}$ in $H_s(S^1)$ converges, hence $H_s(S^1)$ is a Hilbert space. □

Properties of $H_s(S^1)$

Proposition 76. Let $k \in \mathbb{Z}_{\geq 0}$ and let $s \in \mathbb{R}$ with $s > k + 1/2$. Then there is a continuous inclusion

$$H_s(S^1) \hookrightarrow C^k(S^1).$$

Proof. We will first show the containment. Let $f \in H_s(S^1)$ and let $\{f_n\}$ be a Cauchy sequence of trigonometric polynomials in $H_s(S^1)$ that converges to $f \in H_s(S^1)$. Such a sequence exists by Proposition 72. For every $n \in \mathbb{Z}_{\geq 0}$, we have $f_n \in C^k(S^1)$ because f_n is a finite linear combination of the ψ_n. By the Sobolev inequality

$$\|f_n - f_m\|_{C^k} \leq \omega_{s,k} \|f_n - f_m\|_{H_s} \quad \text{for all} \quad n, m \in \mathbb{Z}_{\geq 0}.$$

Therefore, the sequence $\{f_n\}$ is Cauchy in $C^k(S^1)$. Since $C^k(S^1)$ is complete, $\{f_n\}$ converges in $C^k(S^1)$. Consequently $f \in C^k(S^1)$, which proves the containment.

Since $H_s(S^1)$ and $C^k(S^1)$ are topological vector spaces, it suffices to show the containment is continuous in neighborhoods of the origin. Suppose

$$\|f\|_{H_s} < \varepsilon.$$

Then

$$\|f\|_{C^k} \leq \omega_{s,k} \|f\|_{H_s} < \omega_{s,k}\varepsilon.$$

Therefore, the inclusion $H_s(S^1) \hookrightarrow C^k(S^1)$ is continuous in neighborhoods of 0. □

Proposition 77. Let $s \in \mathbb{R}_{\geq 0}$ and let $k \in \mathbb{Z}_{\geq 0}$ with $k > s + 1/2$. Then there is an inclusion

$$C^k(S^1) \hookrightarrow H_s(S^1).$$

Proof. Let $f \in C^k(S^1)$. Using the integration by parts formula,

$$\left|\left\langle f^{(k)}, \psi_n \right\rangle_{L^2}\right| = \left|\left\langle f, \psi_n^{(k)} \right\rangle_{L^2}\right| = \left|\langle f, (2\pi i n)^k \psi_n \rangle_{L^2}\right| = (2\pi)^k \cdot |n|^k \cdot \left|\langle f, \psi_n \rangle_{L^2}\right|.$$

Since $f^{(k)}$ is continuous, the Riemann-Lebesgue lemma (Proposition 50) states that the Fourier coefficients of $f^{(k)}$ are bounded. Therefore $|n|^k \left|\langle f, \psi_n \rangle_{L^2}\right|$ is bounded as $|n| \to \infty$. Consequently there exists a bound B such that

$$(1+n^2)^k \cdot \left|\langle f, \psi_n \rangle_{L^2}\right|^2 \leq B \quad \text{for all} \quad n \in \mathbb{Z}.$$

Here we used the fact that n^2 and $n^2 + 1$ are asymptotically equal. Then

$$\begin{aligned}
\|f\|_{H_s}^2 &= \sum_{n\in\mathbb{Z}} \left|\langle f, \psi_n \rangle_{L^2}\right|^2 \cdot (1+n^2)^s \\
&= \sum_{n\in\mathbb{Z}} (1+n^2)^k \cdot \left|\langle f, \psi_n \rangle_{L^2}\right|^2 \cdot \frac{1}{(1+n^2)^{k-s}} \\
&\leq B \cdot \sum_{n\in\mathbb{Z}} \frac{1}{(1+n^2)^{k-s}}.
\end{aligned}$$

The final expression is finite because $k - s > 1/2$. Therefore $\|f\|_{H_s}$ is finite, implying that $f \in H_s(S^1)$. □

An Alternate Interpretation of the H_s-norm

At the beginning of the present section, we showed that there are different norms on $C^k(S^1)$ that give the space identical topologies. Similarly we will show a different norm on $H_k(S^1)$ that gives the same topology as $\|\cdot\|_{H_k}$ when $k \in \mathbb{Z}_{\geq 0}$. Let $f \in C^k(S^1)$ be represented by the Fourier series $f = \sum_{n\in\mathbb{Z}} c_n \psi_n$. Then we define the norm of f to be

$$\|f\|_{H_k}' = \left(\|f\|_{L^2}^2 + \|f'\|_{L^2}^2 + \cdots + \left\|f^{(k)}\right\|_{L^2}^2 \right)^{1/2}.$$

To show that $\|\cdot\|_{H_k}'$ and $\|\cdot\|_{H_k}$ are equivalent, we note that

$$\begin{aligned}
\left(\|f\|_{H_k}'\right)^2 &= \sum_{0\leq j\leq k} \left\|f^{(j)}\right\|_{L^2}^2 \\
&= \sum_{0\leq j\leq k} \sum_{n\in\mathbb{Z}} |c_n|^2 (2\pi n)^{2j} \\
&\leq (k+1)(2\pi)^{2k} \sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^k \\
&\leq (k+1)(2\pi)^{2k} \|f\|_{H_k}^2.
\end{aligned}$$

So $\|\cdot\|_{H_k}$ is at least as fine as $\|\cdot\|_{H_k}'$. For the other required inequality, we use the binomial theorem to estimate the term

$$(1+n^2)^k \leq A_k(1+n^2+\cdots+n^{2k}) \quad \text{for some} \quad A_k \in \mathbb{Z}.$$

Then,

$$\begin{aligned} \|f\|_{H_k}^2 &= \sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^k \\ &\leq \sum_{n\in\mathbb{Z}} \sum_{0\leq j\leq k} |c_n|^2 \cdot A_k \cdot n^{2j} \\ &\leq A_k \sum_{n\in\mathbb{Z}} \sum_{0\leq j\leq k} |c_n|^2 (2\pi n)^{2j} \\ &= A_k \sum_{0\leq j\leq k} \left\| f^{(j)} \right\|_{L^2} \\ &= A_k \left(\|f\|_{H_k}' \right)^2 . \end{aligned}$$

Therefore $\|\cdot\|_{H_k}'$ and $\|\cdot\|_{H_k}$ generate identical topologies on $H_k(S^1)$. We can therefore think of the H_s norm (for $s \in \mathbb{Z}_{\geq 0}$) as the sum of L^2 norms of f and its derivatives.

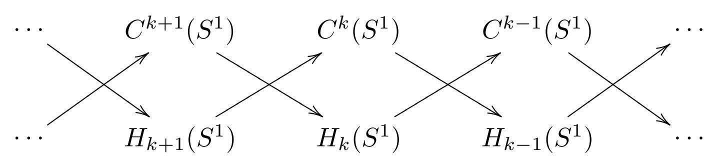
Proposition 78. Let $k \in \mathbb{Z}_{\geq 0}$. Then $C^{k+1}(S^1)$ embeds continuously into $H_k(S^1)$.

Proof. Proposition 77 shows that $C^{k+1}(S^1)$ embeds into $H_k(S^1)$, so all that we must prove is continuity. Let $f \in C^{k+1}(S^1)$. By the above discussion,

$$\|f\|_{H_k} \leq A_k \sum_{0\leq j\leq k} \left\| f^{(j)} \right\|_{L^2} \leq A_k \sum_{0\leq j\leq k+1} \left\| f^{(j)} \right\|_{C^0} = A_k \|f\|_{C^{k+1}}$$

which proves the continuity of the inclusion. □

Remark 79. By Propositions 76 and 78, we have



where all arrows are continuous inclusions. This comparison between $C^k(S^1)$ and $H_s(S^1)$ is interesting because C^k-ness is a statement about the smoothness of a function, while H_s-ness has to do with the asymptotic decay of the Fourier coefficients of a function. Therefore, we can put asymptotic bounds on the Fourier coefficients of a function based on its smoothness, and vice-versa.

3.2 Smooth Functions From $H_s(S^1)$

In this section, we introduce the notion of cofinal sequences of spaces. We show that two cofinal sequences of spaces have the same limit. At the end of the section, we use this fact about cofinal limits to represent the spaces $C^\infty(S^1)$ and $C^\infty(S^1)^*$ in terms of the Sobolev spaces $H_s(S^1)$ for $s \geq 0$. Representing $C^\infty(S^1)^*$ in terms of Hilbert spaces (as opposed to Banach spaces as we did previously) allows us a much cleaner description of distributions on S^1 which we will develop in the following section.

Cofinal Limits

We introduce the notion of cofinal sequences in a general setting.

Definition 80. Let $\{A_k : k \in \mathbb{Z}_{\geq 0}\}$ and $\{B_k : k \in \mathbb{Z}_{\geq 0}\}$ be families of objects in the same category with morphisms

$$A_{k+1} \xrightarrow{\varphi_{k+1,k}} A_k \quad \text{and} \quad B_{k+1} \xrightarrow{\upsilon_{k+1,k}} B_k \quad \text{for all} \quad k \in \mathbb{Z}_{\geq 0}.$$

Then we say the sequences $\{A_k\}$ and $\{B_k\}$ are *cofinal* if for every $k \in \mathbb{Z}_{\geq 0}$ there exist indices $n_k, m_k \in \mathbb{Z}_{\geq 0}$ such that there are compatible morphisms

$$A_{n_k} \xrightarrow{\text{proj.}} B_k \quad \text{and} \quad B_{m_k} \xrightarrow{\text{proj.}} A_k.$$

Remark 81. From our work in the previous section (see the Remark on Page 54) we can see that the sequences $\{C^k(S^1)\}$ and $\{H_k(S^1)\}$ are cofinal. That is, $C^{k+1}(S^1)$ embeds continuously into $H_k(S^1)$, and $H_{k+1}(S^1)$ embeds continuously into $C^{k+1}(S^1)$ for all $k \in \mathbb{Z}_{\geq 0}$.

Given sequences of objects $\{A_k\}$ and $\{B_k\}$, we can form the (projective) limits

$$\lim_k A_k \quad \text{and} \quad \lim_k B_k$$

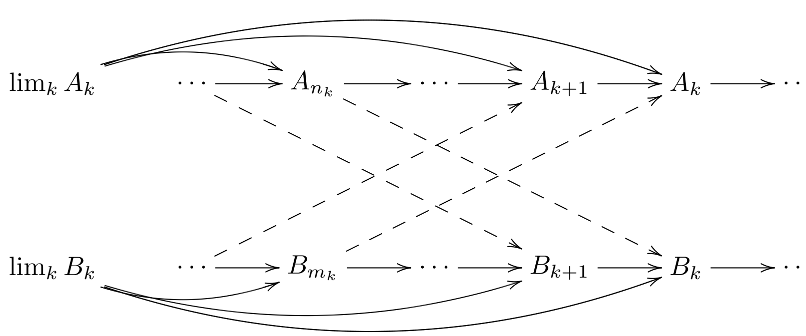
assuming that these limits exist. Since the sequences are cofinal, we might expect that these limits are the same. In fact, this turns out to be the case.

Proposition 82. Let $\{A_k\}$ and $\{B_k\}$ be cofinal sequences. Then

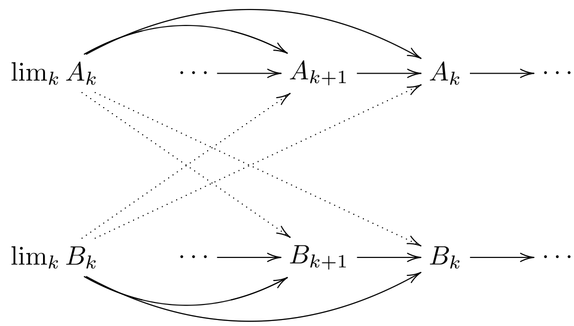
$$\lim_k A_k \cong \lim_k B_k$$

whenever the limits exist.

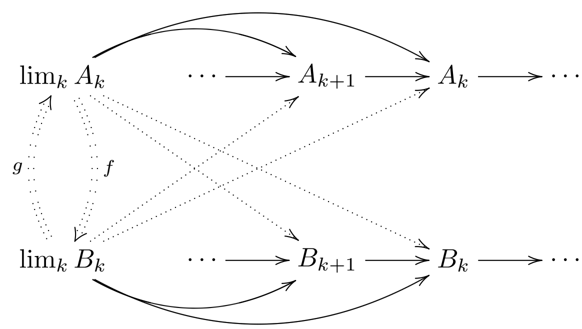
Proof. By Proposition 25 each of the limits $\lim_k A_k$ and $\lim_k B_k$ is unique. Further, in the proof of Proposition 25 we showed that the only map from a limit to itself that is compatible with the projection mappings out of the limit is the identity function. Consider the diagram



The dashed arrows are projection mappings. The existence of the dashed arrows is the statement that the sequences $\{A_k\}$ and $\{B_k\}$ are cofinal. Since the dashed arrows are projection mappings, all triangles in the diagram commute. Further, by composing the projections out of the limits with the dotted arrows we obtain the compatible mappings indicated by dotted arrows:



By the definition of projective limit, the dotted arrows in the previous diagram induce unique compatible maps f and g:



The map $g \circ f$ is a compatible mapping from $\lim_k A_k$ to itself, and is therefore the identity map on $\lim_k A_k$. Similarly, $f \circ g$ is the identity map on $\lim_k B_k$. Since the compositions $g \circ f$ and $f \circ g$ are the identity maps on their respective domains, f and g are mutual inverses, hence each is an isomorphism. Consequently $\lim_k A_k$ and $\lim_k B_k$ are (uniquely) isomorphic. ☐

Remark 83. So far, all of the limits we have encountered have been countable limits of spaces. It is possible to take uncountable limits. For example we can take the limit

$$\lim_{s \in \mathbb{R}_{\geq 0}} H_s.$$

This limit, however, is cofinal with the corresponding limit taken over the nonnegative integers, hence

$$\lim_{s \in \mathbb{R}_{\geq 0}} H_s(S^1) = \lim_{k \in \mathbb{Z}_{\geq 0}} H_k(S^1).$$

The Spaces $C^\infty(S^1)$ and $C^\infty(S^1)^*$ Revisited

We are now ready to define the space of smooth functions on S^1 in terms of the Sobolev spaces.

Theorem 84.

$$C^\infty(S^1) = \lim_k C^k(S^1) = \lim_{s\geq 0} H_s(S^1).$$

Proof. The theorem is an immediate corollary of Proposition 82 and the remark following the proposition. □

Similarly, we can define the space of distributions in terms of the Sobolev spaces.

Corollary 85.

$$C^\infty(S^1)^* = \operatorname*{colim}_k C^k(S^1)^* = \operatorname*{colim}_{s\geq 0} H_s(S^1)^*.$$

Proof. The first equality holds by Theorem 35, and the second by Theorem 84. □

The preceding corollary gives us description of the space of distributions as a colimit of the duals of Sobolev spaces. This description of distributions will give us a much cleaner view of distributions, because the Sobolev spaces are Hilbert spaces. In particular, the dual spaces of Hilbert spaces are again Hilbert spaces. In the following section, we will utilize the structure of $H_s(S^1)$ to show that we can represent all distributions as Fourier series.

3.3 Distributions Revisited

In this section we utilize the Sobolev spaces to show that we can represent all distributions as Fourier series. We first define the Sobolev spaces $H_s(S^1)$ for all $s \in \mathbb{R}$, and show that every element in each Sobolev space is uniquely determined by a Fourier series. Then, we identify the Sobolev space $H_{-s}(S^1)$ as the dual space of $H_s(S^1)$ for all $s \in \mathbb{R}$.

The second half of the section contains a slew of corollaries that relate the Sobolev spaces to calculus on S^1. By the end of the section, we will have completely reduced differential calculus on S^1 to *arithmetic* of Fourier series.

The Fourier Series of a Distribution

Recall that a *distribution* u (on S^1) is a continuous linear functional on the space of smooth functions on S^1. That is, u is a continuous linear map

$$u : C^\infty(S^1) \longrightarrow \mathbb{C}.$$

Now we address the issue of how to represent such objects in terms of Fourier series.

We have shown that every function $f \in C^\infty(S^1)$ can be represented by Fourier series

$$f = \sum_{n\in\mathbb{Z}} c_n \psi_n.$$

In order to evaluate a distribution u at f we suspect that we can write

$$u(f) = u\left(\sum_{n\in\mathbb{Z}} c_n \psi_n\right) = \sum_{n\in\mathbb{Z}} c_n u(\psi_n),$$

where the final equality should hold by the linearity of u. Assuming that the above steps are justified, the value of u on f is entirely determined by the Fourier coefficients of f and the value of u on the functions ψ_n. Thus, it is at least plausible that we can represent distributions as Fourier series, although we have some work to do to make such a claim rigorous.

Definition 86. Let u be a distribution on S^1. We define the *Fourier transform* of u, denoted $\widehat{u}$, to be the mapping

$$\widehat{u} : \mathbb{Z} \longrightarrow \mathbb{C}$$

where

$$\widehat{u}(n) = u(\psi_{-n}) \quad \text{for all} \quad n \in \mathbb{Z}.$$

Then we write the *Fourier series* of u

$$u \sim \sum_{n\in\mathbb{Z}} \widehat{u}(n) \cdot \psi_n.$$

We do not write equality in the definition of the Fourier series of u because we do not know in what sense the Fourier series of u converges to u. We do expect that the Fourier series of u converges to u in the $C^\infty(S^1)^*$ topology.

In order to address the issue of convergence of Fourier series of distributions, we make the following definition.

Definition 87. We define the Sobolev spaces for *all* real s to be

$$H_s(S^1) = \left\{ u \in C^\infty(S^1)^* \,:\, \sum_{n\in\mathbb{Z}} |\widehat{u}(n)|^2 \cdot (1+n^2)^s < \infty \right\}.$$

For $u \in C^\infty(S^1)^*$ we define the s^{th} *Sobolev norm* of u by

$$\|u\|_{H_s}^2 = \sum_{n\in\mathbb{Z}} |\widehat{u}(n)|^2 \cdot (1+n^2)^s.$$

Remark 88. This definition of $H_s(S^1)$ for all $s \in \mathbb{R}$ agrees with the previous definition of $H_s(S^1)$ for $s \geq 0$. To see this, we note that the new definition has the exact same restriction on the Fourier coefficients of elements in $H_s(S^1)$ as the old definition for $s \geq 0$. By Propositions 72 and 74, for $s \geq 0$, we can treat an element $u \in H_s(S^1)$ and its Fourier series interchangeably.

Further, the new definition of $H_s(S^1)$ allows us to compare distributions to functions on S^1. In particular, let $k \in \mathbb{Z}_{\geq 0}$ and let $s > k + 1/2$. Then any distribution $u \in H_s(S^1)$ corresponds to some function $f \in C^k(S^1)$, as the Fourier series of u converges in $C^k(S^1)$.

The Spaces $\ell^{2,s}$

We introduce the spaces $\ell^{2,s}$ (known as weighted ℓ^2 spaces) for convenience of notation. We will show in Proposition 92 that $\ell^{2,s} \cong H_s(S^1)$, but for our present purposes $\ell^{2,s}$ is easier to work with.

Definition 89. We define the space $\ell^{2,s}$ to be the set of two-sided sequences $\{c_n\}$ of complex numbers, subject to the constraint that

$$\sum_{n\in\mathbb{Z}} |c_n|^2 \cdot (1+n^2)^s < \infty.$$

That is,

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

We define the inner product of two sequences $\{c_n\}, \{d_n\} \in \ell^{2,s}$ to be

$$\langle \{c_n\}, \{d_n\} \rangle_{\ell^{2,s}} = \sum_{n\in Z} c_n \overline{d_n} \cdot (1+n^2)^s.$$

The associated norm is given by

$$\|\{c_n\}\|^2_{\ell^{2,s}} = \langle \{c_n\}, \{c_n\} \rangle_{\ell^{2,s}}.$$

Proposition 90. The space $\ell^{2,s}$ is a Hilbert space.

Proof. Since $\ell^{2,s}$ is an inner product space, we must show that it is complete. To this end, let $\{C_j\} = \{\{c_{j,n}\}\}$ be a Cauchy sequence[3] in $\ell^{2,s}$. First, note that

$$|c_{j,n} - c_{k,n}|^2 \cdot (1+n^2)^s \leq \sum_{n\in\mathbb{Z}} |c_{j,n} - c_{k,n}|^2 (1+n^2)^s$$

for all $n \in \mathbb{Z}$ and $j, k \in \mathbb{Z}_{\geq 0}$. Since the right side of the above expression satisfies the Cauchy condition, the sequence $\{c_{j,n}\}$ is a Cauchy sequence of complex numbers for every fixed $n \in Z$. Therefore the sequence $\{C_j\}$ converges to some sequence,

$$C_j \xrightarrow{j\to\infty} C = \{c_n\}.$$

Now we must show that the limit sequence is in $\ell^{2,s}$. To this end, let $j \in \mathbb{Z}_{\geq 0}$ be sufficiently large so that $\|C - C_j\|_{\ell^{2,s}} \leq 1$. Then

$$\begin{aligned} \|C\|_{\ell^{2,s}} &= \|C - C_j + C_j\|_{\ell^{2,s}} \\ &\leq \|C - C_j\|_{\ell^{2,s}} + \|C_j\|_{\ell^{2,s}} \\ &\leq 1 + \|C_j\|_{\ell^{2,s}} \\ &< \infty. \end{aligned}$$

Since $\|C\|_{\ell^{2,s}}$ is finite, $C \in \ell^{2,s}$. Therefore, every Cauchy sequence in $\ell^{2,s}$ converges in $\ell^{2,s}$, so $\ell^{2,s}$ is complete. □

[3] This notation is a little awkward, but $\{C_j\}$ is a sequence of sequences.

Proposition 91. The spaces $\ell^{2,s}$ and $\ell^{2,-s}$ are mutual duals. That is,

$$(\ell^{2,s})^* = \ell^{2,-s} \quad \text{and} \quad (\ell^{2,-s})^* = \ell^{2,s}.$$

Proof. Let

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

be the bilinear form[4] defined by

$$\langle\langle \{c_n\}, \{d_n\} \rangle\rangle = \sum_{n\in Z} c_n d_{-n} \quad \text{for} \quad \{c_n\} \in \ell^{2,s},\ \{d_n\} \in \ell^{2,-s}.$$

Let $\Lambda : \ell^{2,-s} \longrightarrow (\ell^{2,s})^*$ be defined by

$$\Lambda(\{d_n\}) = \lambda_{\{d_n\}} \quad \text{for all} \quad \{d_n\} \in \ell^{2,-s}$$

where

$$\lambda_{\{d_n\}}(\{c_n\}) = \langle\langle \{c_n\}, \{d_n\} \rangle\rangle \quad \text{for all} \quad \{c_n\} \in \ell^{2,s}.$$

We claim that Λ is an isomorphism.

First, we will show that for every $\lambda_{\{d_n\}} \in \ell^{2,-s}$, $\lambda_{\{d_n\}}$ is a continuous linear functional on $\ell^{2,s}$. It is clear from the definition of $\lambda_{\{d_n\}}$ that it is linear, so we will show that $\lambda_{\{d_n\}}$ is continuous. Let $\{c_n\} \in \ell^{2,s}$. Consider the modulus of $\lambda_{\{d_n\}}(\{c_n\})$,

$$\begin{aligned}
\left|\lambda_{\{d_n\}}(\{c_n\})\right| &= \left|\sum_{n\in\mathbb{Z}} c_n d_{-n}\right| \\
&\le \sum_{n\in\mathbb{Z}} |c_n| (1+n^2)^{s/2} |d_{-n}| (1+n^2)^{-s/2} \\
&\le \left(\sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^s\right)^{1/2} \left(\sum_{n\in\mathbb{Z}} |d_n|^2 (1+n^2)^{-s}\right)^{1/2} \\
&= \|\{c_n\}\|_{\ell^{2,s}} \cdot \|\{d_n\}\|_{\ell^{2,-s}}.
\end{aligned}$$

The second inequality in the above display holds by the Cauchy-Schwarz inequality in $\ell^{2,0}$. The inequality $\left|\lambda_{\{d_n\}}(\{c_n\})\right| \le \|\{c_n\}\|_{\ell^{2,s}} \cdot \|\{d_n\}\|_{\ell^{2,-s}}$ proves the continuity [5] of $\lambda_{\{d_n\}}$. Further, the same inequality shows that the map Λ is itself continuous, [6] because

$$\sup\{|\Lambda(\{d_n\})(\{c_n\})| : \|\{c_n\}\|_{\ell^{2,s}} < 1\} \le \|\{d_n\}\|_{\ell^{2,-s}}.$$

[4] Recall that a *bilinear form* is a map that is linear in each of its arguments.

[5] Recall that for a linear map between normed spaces $L : B_1 \to B_2$ is continuous if and only if there exists a positive real number M such that

$$\|L(x)\|_{B_2} \le M \cdot \|x\|_{B_1} \quad \text{for all} \quad x \in B_1.$$

See, for example, [11].

[6] The dual space of a normed complex vector space B is a normed space. For $\lambda \in B^*$, the natural norm of λ is given by

$$\|\lambda\|_{B^*} = \sup\{|\lambda(x)| : \|x\|_B \le 1\}.$$

See [11].

We note that the map Λ is injective. To see this, let $\{d_n\}, \{d'_n\} \in \ell^{2,-s}$ with $\{d_n\} \neq \{d'_n\}$. In particular, we must have $d_j \neq d'_j$ for some $j \in \mathbb{Z}$. Let $c_n = \delta_{nj} \in \ell^{2,s}$, where δ_{nj} is the Kronecker delta function.[7] Then

$$\lambda_{\{d_n\}}(c_n) = d_j \neq d'_j = \lambda_{\{d'_n\}}(c_n).$$

Therefore $\lambda_{\{d_n\}} \neq \lambda_{\{d'_n\}}$.

Now we will show that Λ is surjective. Let $\lambda \in (\ell^{2,s})^*$. Since $\ell^{2,s}$ is a Hilbert space[8], there is some sequence $\{a_n\} \in \ell^{2,s}$ such that

$$\lambda(\{c_n\}) = \langle \{c_n\}, \{a_n\} \rangle_{\ell^{2,s}} = \sum_{n \in \mathbb{Z}} c_n \cdot \overline{a_n}(1+n^2)^s.$$

Therefore $\lambda = \lambda_{\{d_n\}}$ where

$$d_n = \overline{a_{-n}} \cdot (1+n^2)^s.$$

We verify that $\{d_n\} \in \ell^{2,-s}$,

$$\begin{aligned}
\|\{d_n\}\|_{\ell^{2,-s}} &= \sum_{n \in \mathbb{Z}} |d_n|^2 (1+n^2)^{-s} \\
&= \sum_{n \in \mathbb{Z}} \left|\overline{a_{-n}} \cdot (1+n^2)^s\right|^2 (1+n^2)^{-s} \\
&= \sum_{n \in \mathbb{Z}} |a_n|^2 (1+n^2)^s \\
&< \infty.
\end{aligned}$$

Since Λ is linear, bijective and continuous, it is an isomorphism of Hilbert spaces. Thus

$$\ell^{2,-s} \cong (\ell^{2,s})^*$$

as desired. □

Now we identify the spaces $\ell^{2,s}$ with the Sobolev spaces $H_s(S^1)$.

Proposition 92. Let $\mathcal{G} : H_s(S^1) \to \ell^{2,s}$ be defined by

$$\mathcal{G} : u \longmapsto \{\widehat{u}(n)\} \quad \text{for all} \quad u \in H_s(S^1).$$

[7] Recall that the *Kronecker delta function* is defined by the relation

$$\delta_{nj} = \begin{cases} 0 & n \neq j \\ 1 & n = j. \end{cases}$$

[8] Here we quote a general result from Hilbert space theory. Let H be a Hilbert space and let λ be a continuous linear functional on H. Then there is some $y \in H$ such that

$$\lambda(x) = \langle x, y \rangle \quad \text{for all} \quad x \in H.$$

This result is sometimes referred to as the Riesz-Fischer theorem. See, for example, [10].

The map $\mathcal{G}$ is an isomorphism. In particular

$$H_s(S^1) \cong \ell^{2,s}$$

as Hilbert spaces.

Proof. The fact that $\mathcal{G}(u) \in \ell^{2,s}$ is part of the definition of $H_s(S^1)$. We must show that $\mathcal{G}$ is linear, continuous and bijective.

To see that $\mathcal{G}$ is linear let $u, v \in H_s(S^1)$. Then

$$\mathcal{G}(u+v) = \left\{ \widehat{(u+v)}(n) \right\} = \{\widehat{u}(n) + \widehat{v}(n)\} = \mathcal{G}(u) + \mathcal{G}(v)$$

where the second equality holds because Fourier series add term-wise.

To see that $\mathcal{G}$ is continuous, let $u \in H_s(S^1)$. We compute

$$\|\mathcal{G}(u)\|_{\ell^{2,s}} = \sum_{n \in \mathbb{Z}} |\widehat{u}(n)|^2 (1+n^2)^s = \|u\|_{H_s}.$$

Since $\mathcal{G}$ preserves norm, it is continuous.

We claim that $\mathcal{G}$ is injective. To see this, suppose $u \neq v$. Then for some $f \in C^\infty(S^1)$, $u(f) \neq v(f)$. Since we can represent f by a Fourier series, there is some ψ_n such that $u(\psi_n) \neq v(\psi_n)$. Therefore $\widehat{u}(n) \neq \widehat{v}(n)$ for some $n \in \mathbb{Z}$, so $\mathcal{G}(u) \neq \mathcal{G}(v)$.

To show show that $\mathcal{G}$ is surjective, we first consider the case where $s \geq 0$. Let $\{c_n\} \in \ell^{2,s}$. We must show that there is a distribution $u \in H_s(S^1)$ such that $\widehat{u}(n) = c_n$ for all $n \in \mathbb{Z}$. As before (Proposition 74), $H_s(S^1)$ is a Hilbert space with orthogonal basis $\{\psi_n\}$. By scaling the ψ_n appropriately, we form the orthonormal basis $\left\{\psi_n \cdot (1+n^2)^{s/2}\right\}$. Since $H_s(S^1)$ is a Hilbert space, we invoke Plancherel's theorem (see the footnote on page 46) and obtain

$$\|u\|_{H_s}^2 = \sum_{n \in \mathbb{Z}} \left| \left\langle u, \psi_n \cdot (1+n^2)^{s/2} \right\rangle \right|^2 = \sum_{n \in \mathbb{Z}} |\widehat{u}(n)|^2 (1+n^2)^s.$$

If we take $\widehat{u}(n) = c_n$, then

$$\|u\|_{H_s}^2 = \sum_{n \in \mathbb{Z}} |c_n|^2 (1+n^2)^s = \|\{c_n\}\|_{\ell^{2,s}} < \infty.$$

Since the H_s norm of u is finite, $u \in H_s(S^1)$. Therefore $\mathcal{G}$ is surjective for $s \geq 0$.

Now we consider the case where $s < 0$. Let $\{c_n\} \in \ell^{2,-s}$. Let $u \in H_s(S^1)$ be the distribution defined by

$$u(f) = \sum_{n \in \mathbb{Z}} \widehat{f}(n) \cdot c_{-n} \quad \text{for all} \quad f \in C^\infty(S^1).$$

Then

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

We will show that u is a continuous linear functional on $H_{-s}(S^1)$. The linearity of u is an immediate consequence of the linearity of the inner product on $L^2(S^1)$. To see this, we compute

$$u(f+g) = \sum_{n\in\mathbb{Z}} \langle f+g, \psi_n\rangle \cdot c_{-n} = \sum_{n\in\mathbb{Z}} (\langle f, \psi_n\rangle + \langle g, \psi_n\rangle) \cdot c_{-n} = u(f) + u(g).$$

To see that u is continuous, we again use the fact that a linear functional is continuous if and only if it is bounded. Then

$$\begin{aligned}\left|\sum_{n\in\mathbb{Z}} \widehat{f}(n)\cdot c_{-n}\right| &\le \sum_{n\in\mathbb{Z}} \left|\widehat{f}(n)\right| (1+n^2)^{-s/2}\cdot c_{-n}(1+n^2)^{s/2} \\ &\le \left(\sum_{n\in\mathbb{Z}} \left|\widehat{f}(n)\right|^2 (1+n^2)^{-s}\right)^{1/2} \left(\sum_{n\in\mathbb{Z}} |c_n|^2 (1+n^2)^s\right)^{1/2} \\ &= \|\{c_n\}\|_{\ell^{2,s}}\cdot \|f\|_{H_{-s}}.\end{aligned}$$

The second inequality is (once again) the Cauchy-Schwartz inequality. Since the final expression is bounded for all $f \in H_{-s}(S^1)$, u is a continuous linear functional on $H_{-s}(S^1)$. Since $C^\infty(S^1)$ embeds continuously in $H_{-s}(S^1)$, u is a continuous linear functional on $C^\infty(S^1)$. In particular, u is a distribution satisfying

$$\widehat{u}(n) = c_n \quad \text{for all} \quad n\in\mathbb{Z}$$

which proves the surjectivity of the map $\mathcal{G}$. Therefore $\mathcal{G}$ is an isomorphism, and

$$H_s(S^1) \cong \ell^{2,s}.$$

□

The Payoff

At this point all of the pieces are in place to easily prove all of the results that we have anticipated. The reader is encouraged to relax and enjoy the fruits of our labor.

Corollary 93. Let $\langle\langle\cdot,\cdot\rangle\rangle : H_s(S^1)\times H_{-s}(S^1) \to \mathbb{C}$ be the bilinear pairing defined by

$$\langle\langle f, u\rangle\rangle = \sum_{n\in\mathbb{Z}} \widehat{f}(n)\cdot \widehat{u}(-n) \quad \text{for all} \quad f\in H_s(S^1),\ u\in H_{-s}(S^1).$$

Define the map $\Lambda : H_{-s}(S^1) \to (H_s(S^1))^*$ to be

$$\Lambda u = \lambda_u \quad \text{where} \quad \lambda_u(f) = \langle\langle f, u\rangle\rangle \quad \text{for all} \quad u\in H_{-s}(S^1),\ f\in H_s(S^1).$$

The map Λ is an isomorphism. In particular,

$$H_{-s}(S^1) \cong (H_s(S^1))^*.$$

Proof. By Proposition 92 we can identify f and u with the sequences

$$\left\{\widehat{f}(n)\right\} \in \ell^{2,s} \quad \text{and} \quad \{\widehat{u}(n)\} \ \in \ \ell^{2,-s}.$$

The pairing $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is then precisely the same pairing as in the proof of Proposition 91. Combining the results and notation from the two previous propositions,

$$H_{-s}(S^1) \xrightarrow[\cong]{\mathcal{G}} \ell^{2,-s} \xrightarrow[\cong]{\Lambda} (\ell^{2,s})^* \xrightarrow[\cong]{\mathcal{G}^{-1}} (H_s(S^1))^*.$$

□

Corollary 94.

$$C^\infty(S^1)^* \cong \operatorname*{colim}_{s\geq 0} H_{-s}(S^1)$$

Proof.

$$C^\infty(S^1)^* \cong \left(\lim_{s\geq 0} H_s(S^1)\right)^* \cong \operatorname*{colim}_{s\geq 0} H_s(S^1)^* = \operatorname*{colim}_{s\geq 0} H_{-s}(S^1).$$

□

Corollary 95. Let $u \in C^\infty(S^1)^*$, and let $f \in C^\infty(S^1)$. Then

$$u(f) = \sum_{n\in\mathbb{Z}} \widehat{f}(n) \cdot \widehat{u}(-n)$$

Proof. Since $u \in C^\infty(S^1)^*$, $u \in H_{-s}(S^1)$ for some $s \in \mathbb{R}_{\geq 0}$. (This fact is proven in Appendix A.) Therefore u is a continuous linear functional on $H_s(S^1)$. Then for any $f \in H_s(S^1)$, by Corollary 93,

$$u(f) = \langle\!\langle f, u \rangle\!\rangle = \sum_{n\in\mathbb{Z}} \widehat{f}(n) \cdot \widehat{u}(-n).$$

Since $C^\infty(S^1)$ embeds continuously into $H_s(S^1)$, the above equality certainly holds for all $f \in C^\infty(S^1)$. □

Definition 96. Let $\{c_n\}$ be a collection of Fourier coefficients. Then we say that $\{c_n\}$ is of *moderate growth* if there exists constants $C, N \in \mathbb{R}$ such that

$$|c_n| \leq C \cdot (1+n^2)^N \quad \text{for all} \quad n \in \mathbb{Z}.$$

Corollary 97. Let $\{c_n : n \in Z\}$ be a two-sided sequence of moderate growth. Then there is a distribution $u \in C^\infty(S^1)^*$ such that

$$\widehat{u}(n) = c_n \quad \text{for all} \quad n \in \mathbb{Z}.$$

Proof. Let $C, N \in \mathbb{R}$ be such that $|c_n| \leq C \cdot (1+n^2)^N$ for all $n \in \mathbb{Z}$. Then

$$\begin{aligned}\sum_{n\in\mathbb{Z}} |c_n|^2 \cdot (1+n^2)^{-(2N+1)} &\leq \sum_{n\in\mathbb{Z}} C^2 \cdot (1+n^2)^{2N} \cdot (1+n^2)^{-(2N+1)} \\ &= C^2 \sum_{n\in\mathbb{Z}} (1+n^2)^{-1} \\ &< \infty.\end{aligned}$$

Therefore

$$\|u\|_{H_{-(2N+1)}} < \infty$$

which implies that $u \in H_{-(2N+1)}(S^1)$. Therefore $u \in (H_{2N+1}(S^1))^*$, so u is a continuous linear functional on $H_{2N+1}(S^1)$. Since $C^\infty(S^1)$ embeds continuously in $H_{2N+1}(S^1)$, $u \in (C^\infty(S^1))^*$. ☐

Corollary 98. Let $u \in C^\infty(S^1)^*$ with

$$u \sim \sum_{n\in\mathbb{Z}} c_n \cdot \psi_n \in H_s(S^1) \quad \text{for some} \quad s \in \mathbb{R}.$$

Then the derivative of u is given by

$$u' \sim (-2\pi i) \sum_{n\in\mathbb{Z}} n \cdot c_n \cdot \psi_n \in H_{s-1}(S^1).$$

Further, differentiation is a continuous map

$$\frac{d}{dx} H_s(S^1) \longrightarrow H_{s-1}(S^1).$$

Proof. Recall that the derivative of a distribution is defined to be

$$u'(f) = -u(f') \quad \text{for all} \quad f \in C^\infty(S^1).$$

Then we compute

$$\begin{aligned}u'(f) &= -u(f') \\ &= -\sum_{n\in\mathbb{Z}} \widehat{f}'(n) \cdot \widehat{u}(-n) \\ &= -2\pi i \sum_{n\in\mathbb{Z}} \widehat{f}(n) \cdot (-n)\widehat{u}(-n) \qquad\qquad (\psi_n' = 2\pi i n \psi_n).\end{aligned}$$

Therefore u' corresponds to the distribution

$$u' \sim (-2\pi i) \sum_{n\in\mathbb{Z}} n \cdot c_n \cdot \psi_n.$$

To see that $u' \in H_{s-1}(S^1)$, we take the norm of u',

$$\begin{aligned}
\|u'\|^2_{H_{s-1}} &= \sum_{n\in\mathbb{Z}} |(-2\pi i)n\widehat{u}(n)|^2 \cdot (1+n^2)^{s-1} \\
&\leq (2\pi)^2 \sum_{n\in\mathbb{Z}} (1+n^2)\, |\widehat{u}(n)|^2 \cdot (1+n^2)^{s-1} \\
&= (2\pi)^2 \sum_{n\in\mathbb{Z}} |\widehat{u}(n)|^2 \cdot (1+n^2)^{s} \\
&= (2\pi)^2 \|u\|^2_{H_s} \\
&< \infty.
\end{aligned}$$

Therefore, differentiation maps H_s to H_{s-1}. Since $\|u'\|_{H_{s-1}} \leq 2\pi \|u\|_{H_s}$, differentiation is continuous. □

Corollary 99. The space of trigonometric polynomials is dense $H_s(S^1)$ for every $s \in \mathbb{R}$. In particular, $C^\infty(S^1)$ is dense $H_s(S^1)$ for all $s \in \mathbb{R}$.

Proof. The functions ψ_n are an orthogonal basis for $H_s(S^1)$ for all $s \in \mathbb{R}$. Let $s \in \mathbb{R}$ and $u \in H_s(S^1)$. Define the sequence of trigonometric polynomials $\{p_j\}$ to be

$$p_j = \sum_{|n|\leq j} \widehat{u}(n)\psi_n \quad \text{for all} \quad j \in \mathbb{Z}_{\geq}.$$

Then

$$\{p_j\} \xrightarrow{j\to\infty} u \quad \text{in} \quad H_s(S^1).$$

This shows that trigonometric polynomials are dense in $H_s(S^1)$. □

Comments

In this section we have completely reduced the calculus of distributions on S^1 to studying the Fourier series associated to distributions. We have succeeded in showing that the derivative of every distribution is well-defined. Further, every distribution can be represented by a Fourier series, and the Fourier series can be differentiated termwise. Since the functions ψ_n are eigenfunctions for the differential operator d/dx, differentiation of a Fourier series is a completely arithmetic operation.

The Sobolev spaces are a powerful tool for extracting information about a distribution. Let

$$u = \sum_{n\in\mathbb{Z}} c_n\psi_n$$

be a distribution. Recall that the *order* of u is the smallest integer k such that $u \in (C^k(S^1))^*$. Since every distribution is in $(C^k(S^1))^*$ for some k, every distribution has finite order. If u has order k, then certainly $u \in (H_{k+1}(S^1))^* = H_{-k-1}(S^1)$. Therefore, the coefficients c_n obey

$$\sum_{n\in\mathbb{Z}} |c_n|^2 \cdot (1+n^2)^{-k-1} < \infty$$

which implies that

$$\sum_{n\in\mathbb{Z}_{\neq 0}} \left|\frac{c_n}{n^{k+1}}\right|^2 < \infty.$$

Therefore, c_n must grow slower than $n^{k+1/2}$. Thus, we can place an upper bound on the order of a distribution by the growth of the Fourier coefficients of the distribution.

Since we can differentiate Fourier series term-wise, we can also *anti*-differentiate term-wise whenever $c_0 = 0$. Let $u \in (C^\infty(S^1))^*$ with

$$u = \sum_{n\in\mathbb{Z}} c_n\psi_n \quad \text{where} \quad c_0 = 0$$

and let

$$v = \sum_{n\in\mathbb{Z}_{\neq 0}} \frac{1}{2\pi i n} c_n\psi_n.$$

Then $u = v'$, so v is an antiderivative of u. Further, the constant term of v satisfies $\widehat{v}(0) = 0$. Therefore v also has an antiderivative. We can then recursively define the j^{th} antiderivative of u by

$$v_j = \sum_{n\in\mathbb{Z}_{\neq 0}} \frac{c_n}{(2\pi i n)^j}\psi_n.$$

Suppose that u has order k. Then, as previously noted,

$$\sum_{n\in\mathbb{Z}_{\neq 0}} \left|\frac{c_n}{n^{k+1}}\right|^2 < \infty.$$

This, in turn, implies that

$$\sum_{n\in\mathbb{Z}_{\neq 0}} \left|\frac{c_n}{(2\pi i n)^{k+1}}\right|^2 < \infty$$

so we conclude that

$$v_{k+1} \in L^2(S^1).$$

Therefore every distribution u with $\widehat{u}(0) = 0$ arises by differentiating a classical function a finite number of times.

In summary, every classical function corresponds to a distribution and the derivative of every distribution is well-defined. Further, every distribution arises (up to an additive constant) from a finite number of differentiations of a classical function. In this sense, the theory of distributions completes differential calculus on S^1.

4

Examples

In this chapter, we work some examples using the tools we have developed so far. The first section focuses on one particular distribution: the Dirac delta distribution. We will show some elementary properties of the Dirac distribution, and give examples of how it arises. In the second section, we discuss how distribution theory can be used to solve some partial differential equations, especially those that arise in physics. In particular, we examine the heat equation and the Schrödinger wave equation on the circle. This chapter is significantly less formal than the previous chapters.

4.1 The Dirac Delta Distribution

The aim of this section is to examine one particular distribution: the Dirac delta distribution. In order to motivate its definition, we first discuss arithmetic of Fourier series. We previously defined the convolution of two integrable functions, but here we define convolution of any two distributions. We will see that the convolution of two distributions is easily obtained from the Fourier transforms of the distributions. We then define the Dirac delta distribution to be the identity distribution under convolution, and examine some of its properties.

Arithmetic of Fourier Series

Since we have shown that all distributions and differentiable functions can be represented by Fourier series, we would like to see how arithmetic operations affect Fourier series. Of course we may add Fourier series term-wise, but we would like to consider other arithmetic operations. We first consider the multiplication of continuously differentiable functions. Currently, we restrict our attention to C^1 functions because for

$$f = \sum_{n\in\mathbb{Z}} c_n\psi_n \in C^1(S^1)$$

the sum $\sum_{n\in\mathbb{Z}} |c_n|$ converges. Let $f, g \in C^1(S^1)$ with Fourier series

$$f = \sum_{n\in\mathbb{Z}} c_n\psi_n, \quad g = \sum_{n\in\mathbb{Z}} b_n\psi_n.$$

Then we evaluate the product of f and g

$$\begin{aligned} f \cdot g &= \left(\sum_{n\in\mathbb{Z}} c_n\psi_n\right)\left(\sum_{m\in\mathbb{Z}} b_m\psi_m\right) \\ &= \sum_{n,m\in\mathbb{Z}} c_n b_m \psi_{n+m} \\ &= \sum_{n\in Z} \psi_n \sum_{k+\ell=n} c_k b_\ell. \end{aligned}$$

Without too much trouble (using the Cauchy-Schwarz inequality), one can check that the sum

$$\sum_{k+\ell=n} |c_k b_\ell|$$

converges for all $n \in \mathbb{Z}$ because both $\sum_n |c_n|$ and $\sum |b_n|$ converge. The coefficients

$$a_n = \sum_{k+\ell=n} c_k b_\ell$$

are known as the Cauchy product of the sequences $\{c_n\}$ and $\{a_n\}$.

In general, we cannot define the product of two distributions, because the Cauchy product of divergent sequences is not well-defined. However, we can compute the product of a smooth function with a distribution. Let f have the same Fourier series as before, but suppose $f \in C^\infty(S^1)$. Let $u \in (C^\infty(S^1))^*$ with Fourier series

$$u = \sum_{n\in\mathbb{Z}} d_n\psi_n.$$

Then the sum

$$\alpha_n = \sum_{k+\ell} c_k d_\ell$$

converges, so the product $f \cdot u$ is well-defined. Informally, the above sum converges because the coefficients c_n die off faster than any polynomial, yet the coefficients d_n grow at most as fast as some polynomial. To show that the product $f \cdot u$ is a distribution, we must show that the coefficients α_n grow at most as quickly as a polynomial. In fact, the coefficients grow at most polynomially, because (again informally) the coefficients c_k dominate the sum defining α_n for large values of $|n|$.

In Section 2.4, we introduced the *convolution* of two functions defined by

$$(f * g)(x) = \int_{S^1} f(y)g(x-y)\,dy \quad \text{for all} \quad x \in S^1.$$

We will show that convolution is well-defined for distributions as well. To motivate the definition that follows, we first consider the convolution of functions $f, g \in C^\infty(S^1)$ again represented by Fourier series. We evaluate

$$\begin{aligned}
(f * g)(x) &= \left(\sum_{n\in\mathbb{Z}} c_n\psi_n\right) * \left(\sum_{m\in\mathbb{Z}} b_m\psi_m\right)(x) \\
&= \int_{S^1} \left(\sum_{n\in\mathbb{Z}} c_n\psi_n(y)\right)\left(\sum_{m\in\mathbb{Z}} b_m\psi_m(x-y)\right) dy \\
&= \int_{S^1} \left(\sum_{n,m\in\mathbb{Z}} c_n b_m \psi_n(y)\psi_m(x-y)\right) dy \\
&= \int_{S^1} \left(\sum_{n,m\in\mathbb{Z}} c_n b_m \psi_{n-m}(y)\psi_m(x)\right) dy \\
&= \sum_{n\in\mathbb{Z}} c_n b_n \psi_n(x).
\end{aligned}$$

The last equality holds because

$$\int_{S^1} \psi_n(x)\,dx = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0. \end{cases}$$

Motivated by the evaluation of convolution for smooth functions, we make the following definition.

Definition 100. Let $u, v \in (C^\infty(S^1))^*$ with Fourier transforms

$$\widehat{u}(n) = d_n, \quad \widehat{v}(n) = e_n \quad \text{for all} \quad n \in \mathbb{Z}.$$

Then we define the *convolution* of u and v to be the distribution represented by the Fourier series

$$\widehat{(u * v)}(n) = d_n \cdot e_n \quad \text{for all} \quad n \in \mathbb{Z}.$$

The convolution is well-defined for distributions because the sequence $\{d_n \cdot e_n\}$ grows at most polynomially.

The δ-Distribution

In Section 2.4, we introduced *approximate identities*. We saw that approximate identities were sequences of functions that approximated the convolution identity, and noted that the convolution identity did not exist as a classical function. From the definition of convolution for distributions, we can see why no such function exists. Namely, the convolution identity (denoted δ) must be represented by the Fourier series

$$\delta = \sum_{n \in \mathbb{Z}} \psi_n.$$

The distribution δ is called the *Dirac delta distribution*.[1] The Fourier series defining δ diverges classically, but δ is perfectly acceptable as a distribution.

For $f = \sum_{n \in Z} c_n \psi_n \in C^1(S^1)$, we can evaluate

$$\delta(f) = \sum_{n \in \mathbb{Z}} c_n \widehat{\delta}(n) = \sum_{n \in \mathbb{Z}} c_n = \sum_{n \in \mathbb{Z}} c_n \psi_n(0) = f(0)$$

where the third equality holds because $\psi_n(0) = 1$ for all $n \in \mathbb{Z}$. Therefore, δ is the continuous linear functional that simply evaluates a function f at 0. This property is often given as the definition[2] of δ.

The Square Wave

We now readdress the *square wave* described in the introduction. The square wave is defined on the unit interval to be

$$\sigma(x) = \begin{cases} 1 & 0 \leq x \leq 1/2 \\ 0 & 1/2 < x < 1. \end{cases}$$

and is made periodic by requiring that $\sigma(x+n) = x$ for all $x \in \mathbb{R}$ and $n \in \mathbb{Z}$. Therefore σ corresponds to a function defined on S^1. We compute the Fourier series for σ by integration. For $n \neq 0$,

$$\begin{aligned} \widehat{\sigma}(n) &= \int_0^1 \sigma(x)\psi_n(x)\,dx \\ &= \int_0^{1/2} \exp(2\pi i n x)\,dx \\ &= \frac{1}{2\pi i n}\left(\exp(\pi i n x) - 1\right) \\ &= \frac{1}{2\pi i n}((-1)^n - 1). \end{aligned}$$

[1] It is also commonly referred to as the Dirac delta *function*.

[2] Although it is often presented (by abuse of notation) as the statement

$$\int \delta(x) f(x)\,dx = f(0).$$

A more appealing formulation of this statement would be

$$\delta(f) = f(0).$$

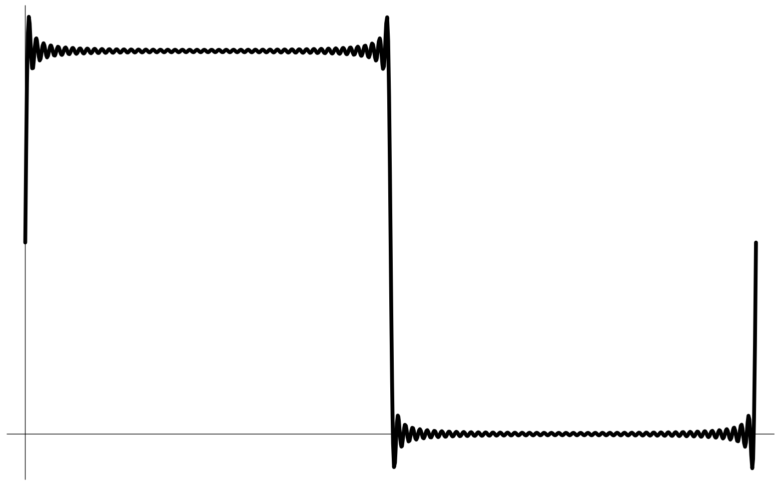


Figure 4.1: Here, the first 101 nonzero terms of the Fourier series ($|n| \leq 101$) for σ are plotted. The series visibly approximates the function σ.

For $n = 0$,

$$\widehat{\sigma}(0) = \int_0^1 \sigma(x)\psi_0(x)\,dx = \int_0^1 \sigma(x) = 1/2.$$

Therefore,

$$\widehat{\sigma}(n) = \begin{cases} 1/2 & n = 0 \\ i/\pi n & n \text{ odd} \\ 0 & \text{otherwise.} \end{cases}$$

In Figure 4.1 we have plotted the sum of the first 101 nonzero terms in the Fourier series of σ.

We would like to make sense of derivative of σ. In accordance with Corollary 98, the derivative of σ has Fourier series

$$\widehat{\sigma'}(n) = -2\pi i n \widehat{\sigma}(n) = \begin{cases} 0 & n \text{ even} \\ 2 & n \text{ odd.} \end{cases}$$

In order to interpret this Fourier series, we again consider the Dirac delta distribution. Recall that

$$\delta(f) = f(0) \quad \text{for all} \quad f \in C^\infty(S^1).$$

Consider the translation of the δ function $\delta_{1/2}$ defined by

$$\delta_{1/2}(f) = f(1/2) \quad \text{for all} \quad f \in C^\infty(S^1).$$

We calculate the Fourier transform of $\delta_{1/2}$ to be

$$\widehat{\delta_{1/2}}(n) = \delta_{1/2}(\psi_n) = \psi_n(1/2) = (-1)^n.$$

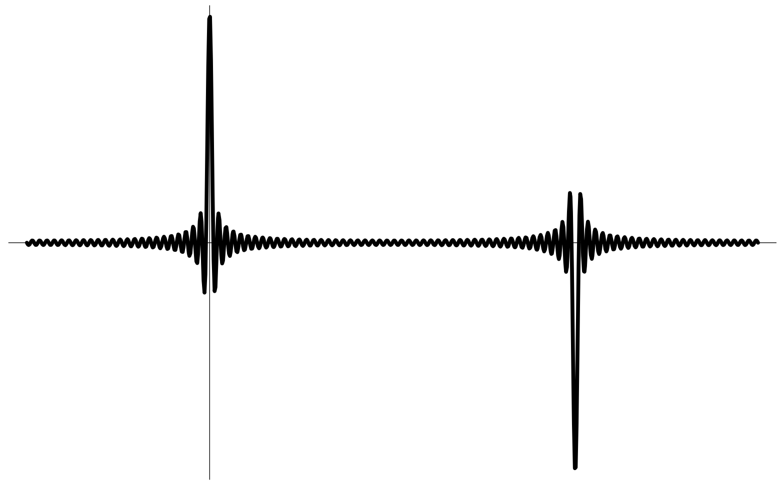


Figure 4.2: This plot shows the sum of the first 101 nonzero terms in the Fourier series of σ' plotted on the domain $[-1/4, 3/4]$. Note that the positive and negative spikes appear at 0 and $1/2$ respectively.

Then the difference $\delta - \delta_{1/2}$ has Fourier transform

$$(\widehat{\delta - \delta_{1/2}})(n) = 1 - (-1)^n = \begin{cases} 0 & n \text{ even} \\ 2 & n \text{ odd.} \end{cases}$$

Therefore

$$\widehat{\sigma'}(n) = (\widehat{\delta - \delta_{1/2}})(n) \quad \text{for all} \quad n \in \mathbb{Z}$$

hence

$$\sigma' = \delta - \delta_{1/2}.$$

The sum of the first 101 nonzero terms in the Fourier series of σ' are plotted in Figure 4.2. The equality $\sigma' = \delta - \delta_{1/2}$ is consistent with our intuition about a would-be classical derivative of σ. Away from the points 0 and $1/2$, σ' is indistinguishable from the classical function 0. This is, of course, consistent with the classical derivative of σ. At the points 0 and $1/2$ where the derivative of σ is not well-defined classically, we find δ distributions. This is consistent with the notion that the function σ is "infinitely steep" where the jump discontinuities occur.

4.2 Distributions in Physics

Spaces of distributions are a natural context in which to study differential equations, because differentiation is always well-defined. Further, distribution theory affords tools such as Fourier series that can simplify the solutions of otherwise difficult differential equations. Since a large portion of the laws of physics are stated in terms of differential equations, it is no surprise that distribution theory is of interest to physicists. In this section we will use Fourier series to solve two differential equations that arise in physics: the heat equation, and the Schrödinger wave equation.

Motivation

Before working the examples, we motivate the use of distributions as objects of study in physics. Classical physics is written in the language of functions. Take for example Newtonian mechanics. The central problem of Newtonian mechanics is this: given a system's initial configuration, find the configuration of the system at a later time. The dynamics in Newtonian mechanics are dictated by Newton's second law. For a single particle system, Newton's second law amounts to

$$\frac{d^2}{dt^2}x(t) = \frac{1}{m}F(t)$$

where x is the position of the particle, F is the force acting on the particle, m is the particles mass, and t is time. Thinking of x and F as classical functions, there are certain assumptions implicit in this description of the dynamics of the system. Most obviously, we are assuming that x is twice differentiable. Is such an assumption reasonable?

Let us step back for a moment. The goal of physics is to describe and predict how nature behaves. That is, a physical theory gives an abstract (often mathematical) description of a physical system. From the abstract model, we then predict how the physical system will evolve. To test the validity of the theory, it must be verified by experiment, which undoubtedly requires some kind of physical measurement. In simplest case of Newtonian mechanics, experimental verification amounts to tracking the trajectory of the particle in question. I argue that even in this conceptually simple classical scenario, a distribution theoretic view is more natural than the classical perspective.

We would *like* to know the position of the particle at every time t. With distributions, it seems that we lose the ability to predict precisely where the particle is at a particular time t, because distributions do not take on point-wise values. This is not a problem, however, because ultimately we must *measure* the position of the particle to verify our theoretical predictions. We cannot even in theory measure the position of the particle to an arbitrary degree of accuracy. Therefore, the precise point-wise values of the function x can never be verified. When we take a physical measurement, we are really finding an average value for the position of the particle around some time t.

If we interpret x as a distribution, instead of evaluating x at a time t, we evaluate x on a smooth function τ. The function τ can be made to encode information about the nature of the measurement to be performed. For example, if a measurement of a position is to be taken at time $t_0 \pm \varepsilon$, τ can be a smooth function with $\tau(t) = 0$ for $t \notin [t_0 - \varepsilon, t_0 + \varepsilon]$, and $\|\tau\|_{L^2} = 1$. Then the value $x(\tau)$ will tell us, within the accuracy of the experiment where we should expect to find the particle around time t_0.

If we allow physical quantities (position, velocity, time, temperature, etc.) to be modeled by distributions rather than classical functions, we allow more general solutions to the equations governing a system's dynamics. That is, we make fewer implicit assumptions about solutions in the statement of a theory. Further, the Sobolev theory allows us to refine distributional solutions after the solutions are found. We can always restrict solutions to be physically reasonable after the solutions have been found. In short, distribution theory may allow solutions that are physically implausible, but classical functions may dismiss solutions that are physically acceptable.

The Heat Equation

The heat equation governs how heat flows through conductive material. In one spatial dimension, the heat equation states (ignoring physical constants)

$$\frac{\partial}{\partial t}u(x,t) = \frac{\partial^2}{\partial x^2}u(x,t).$$

Here, u is the temperature as a function of position and time.

We consider a closed loop of thin wire, so that we can model the wire as the circle S^1. Using distribution theory on the circle, we can represent the temperature at every point on the loop of wire by Fourier series as

$$u(x,t) = \sum_{n\in\mathbb{Z}} c_n(t)\psi_n(x).$$

We allow the coefficients c_n to be functions[3] of time, so that the heat configuration can evolve over time. We take partial derivatives of the Fourier series for u and find

$$\frac{\partial}{\partial t}u(x,t) = \frac{\partial}{\partial t}\sum_{n\in\mathbb{Z}} c_n(t)\psi_n(x) = \sum_{n\in\mathbb{Z}} c_n'(t)\psi_n(x)$$

and

$$\frac{\partial^2}{\partial x^2}u(x,t)\sum_{n\in\mathbb{Z}} c_n(t)\psi_n(x) = \sum_{n\in\mathbb{Z}} c_n(t)\frac{d^2}{dx^2}\psi_n(x) = \sum_{n\in\mathbb{Z}}(-4\pi^2n^2)c_n(t)\psi_n(x).$$

Equating the two sides of the heat equation gives

$$\sum_{n\in\mathbb{Z}} c_n'(t)\psi_n(x) = \sum_{n\in\mathbb{Z}}(-4\pi^2n^2)c_n(t)\psi_n(x).$$

Since the functions ψ_n are orthogonal, this entails that

$$c_n'(t) = -4\pi^2n^2c_n(t) \quad \text{for all} \quad n\in\mathbb{Z}.$$

Given an initial value $c_n(0)$, this differential equation admits the solution

$$c_n(t) = c_n(0)\exp(-4\pi^2n^2t).$$

[3]In the spirit of our previous discussion, we would like to describe the entirety of the system in terms of distributions. However, this would require distribution theory on the real line, not just on the circle.

Therefore, the temperature of each point on the loop of wire is given by

$$u(x,t) = \sum_{n\in\mathbb{Z}} c_n(0)\exp(-4\pi^2 n^2 t)\exp(2\pi i n x). \tag{4.1}$$

We consider the initial heat distribution given by the square wave. Physically, this corresponds to half of the loop being at an initial temperature 1, while the other half of loop starts at a temperature 0. To solve for the heat distribution at a time t we set

$$c_n(0) = \begin{cases} \frac{1}{2} & n = 0 \\ \frac{i}{\pi n} & n \text{ odd} \\ 0 & \text{otherwise.} \end{cases}$$

in Equation 4.1. The resulting progression of temperature configurations is shown in Figure 4.3.

Expressing the solution to the heat equation as a Fourier series allows us to easily identify some qualitative features of solutions without even explicitly calculating solutions. The factor of $\exp(-4\pi^2 n^2 t)$ in the n^{th} Fourier coefficient shows that all Fourier coefficients (except $n = 0$) will decay exponentially in time. Therefore as $t \to \infty$, $u(x,t)$ will tend to the constant $c_0(0)$. Further, the factor of n^2 inside the exponent means that the higher Fourier coefficients will die off more quickly as time passes. Therefore, u becomes smoother as t increases.

The Schrödinger Wave Equation

The Schrödinger wave equation is central to the study of quantum mechanics. It governs the dynamics of quantum particles. The *state* of a quantum system is determined by a wave function, which allows one to compute the probability of finding a particle in a specified region. In one spatial dimension, the (time independent) Schrödinger wave equation is (again ignoring physical constants)

$$i\frac{\partial}{\partial t}\Psi(x,t) = -\frac{\partial^2}{\partial x^2}\Psi(x,t) + V(x)\Psi(x,t).$$

Here, Ψ is the wave function for a given particle and V is the potential energy as a function of position. The probability density for the particle is obtained by computing the square modulus of the wave function Ψ,

$$P(x,t) = (\overline{\Psi}\cdot\Psi)(x,t).$$

For the present example, we will consider a quantum *free particle* constrained to a circle. That is, we will consider the case where $V(x) = 0$ for all $x \in S^1$. Then the Schrödinger equation becomes

$$i\frac{\partial}{\partial t}\Psi(x,t) = -\frac{\partial^2}{\partial x^2}\Psi(x,t).$$

This equation is strikingly similar to the heat equation, but we will see that the factor of i on the left hand side of the equation drastically changes the dynamics of the system.

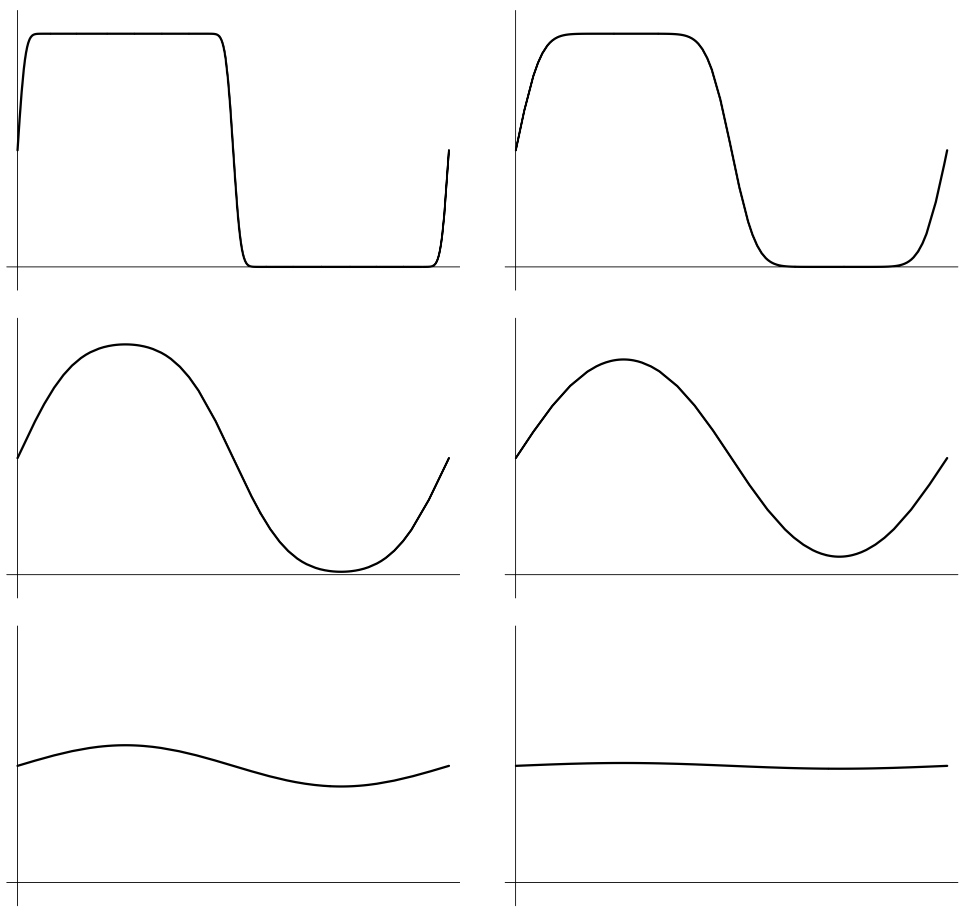


Figure 4.3: Solution to the heat equation for the initial configuration given by the square wave. As t increases the temperature distribution becomes smoother and flatter.

We write Ψ as a Fourier series

$$\Psi(x,t) = \sum_{n\in\mathbb{Z}} c_n(t)\psi_n(x).$$

Taking partial derivatives of Ψ,

$$\frac{\partial}{\partial t}\Psi(x,t) = \sum_{n\in\mathbb{Z}} c_n'(t)\psi_n(x)$$

and

$$\frac{\partial^2}{\partial x^2}\Psi(x,t) = \sum_{n\in\mathbb{Z}}(-4\pi^2 n^2)c_n'(t)\psi_n(x).$$

Then the Schrödinger wave equation states

$$i\sum_{n\in\mathbb{Z}} c_n'(t)\psi_n(x) = -\sum_{n\in\mathbb{Z}}(-4\pi^2 n^2)c_n(t)\psi_n(x).$$

Equating the n^{th} Fourier coefficients gives the differential equation

$$c_n'(t) = -i4\pi^2 n^2 c_n(t)$$

which has solution

$$c_n(t) = c_n(0)\exp(-4i\pi^2 n^2 t).$$

Therefore, we calculate the wave function

$$\begin{aligned}\Psi(x,t) &= \sum_{n\in\mathbb{Z}} c_n(0)\exp(-4i\pi^2 n^2 t)\exp(2\pi i n x)\\ &= \sum_{n\in\mathbb{Z}} c_n(0)\exp(2\pi i n(x - 2\pi n x)).\end{aligned}$$

In Figure 4.4, we calculated the probability density $P = \overline{\Psi}\cdot\Psi$ for a particle with the initial condition

$$\Psi(x,0) = \sum_{|n|\leq 10} \psi_n(x).$$

We can see that the factor of i on the left hand side of the Schrödinger wave equation makes its solutions drastically different from the solutions of the heat equation. The reason for this is the factor of $\exp(-4i\pi^2 n^2 t)$ in the n^{th} Fourier coefficient. Since there is a factor of i in the exponent, the Fourier coefficients oscillate instead of decaying, although each coefficient oscillates at a different frequency.

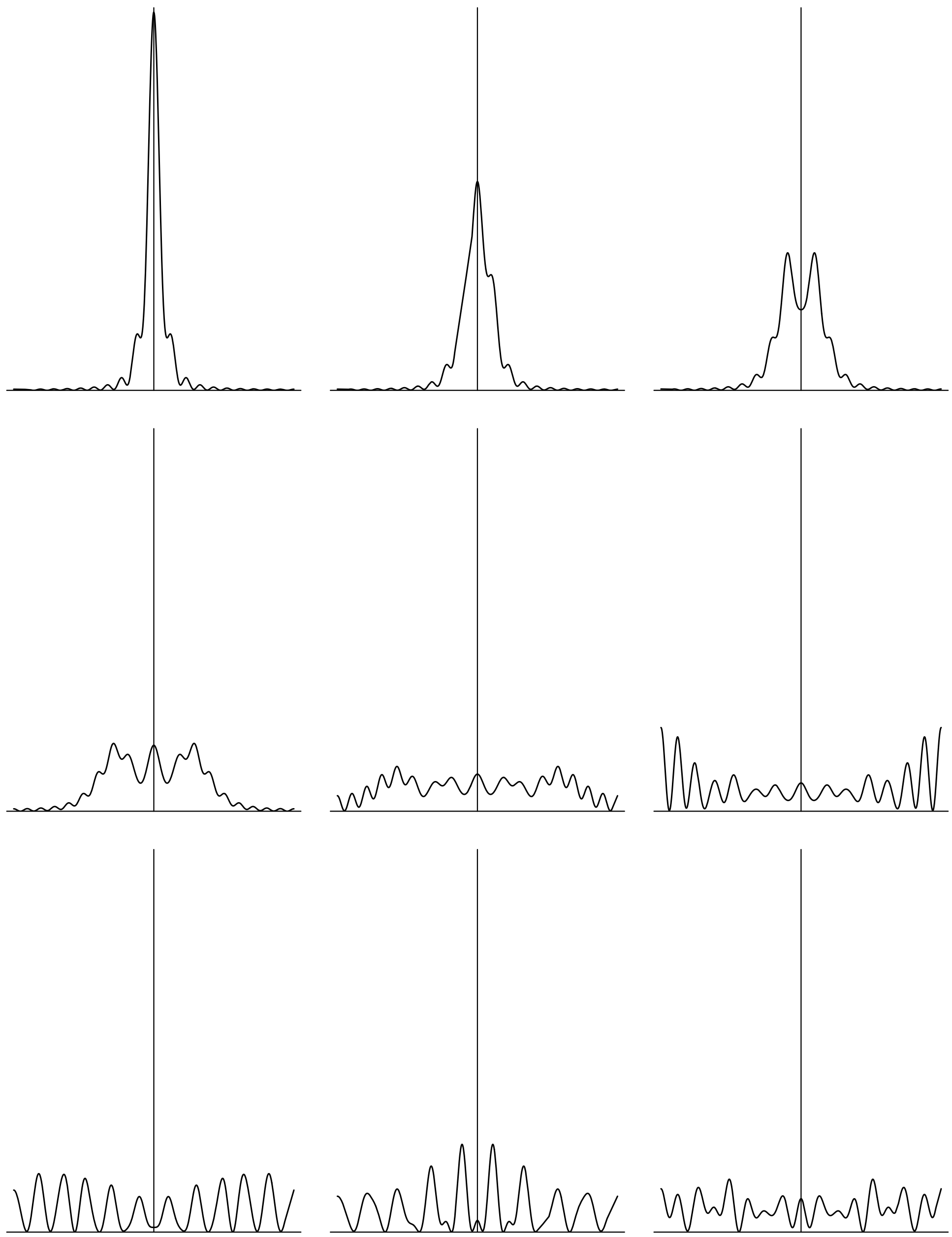


Figure 4.4: These plots are the probability densities associated with a solution to the Schrödinger wave equation with the initial condition $\Psi(x,0) = \sum_{|n|\leq 10} \psi_n(x)$.

A

Dual Spaces of Banach Spaces

The aim of this appendix is to prove the following theorem (Theorem 35):

Theorem. Let $X = \lim_k B_k$ be the limit of Banach spaces B_k. Then

$$(\lim_k B_k)^* \cong \operatorname*{colim}_k B_k^*.$$

In order to prove the theorem, we need a better understanding of the topology of the limit $X = \lim_k B_k$. We get such an understanding by appealing to the *product* $P = \prod_k B_k$.

A.1 Product Spaces

As the reader is probably expecting, we define the product by a mapping property.

Definition 101. Let X_α be objects indexed[1] by the set A. Then the *product* of the objects X_α is an object

$$P = \prod_{\alpha \in A} X_\alpha$$

along with morphisms

$$p_\alpha : P \to X_\alpha \quad \text{for all} \quad \alpha \in A.$$

The product P is characterized by the following universal property. For any object Z with morphisms

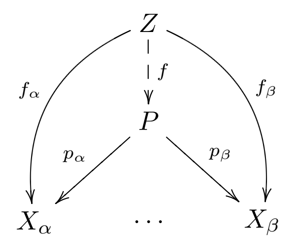
$$f_\alpha : Z \to X_\alpha \quad \text{for all} \quad \alpha \in A$$

[1]It is a standard convention to use greek letters as indices when the index set is not assumed to be countable.

there exists a unique map $f : Z \to P$ such that

$$f_\alpha = p_\alpha \circ f \quad \text{for all} \alpha \in A.$$

Pictorially, the following diagram commutes:



Proposition 102. The product $P = \prod_{\alpha \in A}$ is unique.

We will not prove this proposition here, but we note that the proof is very similar to that of Proposition 25.

For our present purposes, we will consider products only of topological spaces. Of course, we will need the following fact.

Fact 103. Products of topological spaces exist.

We will not prove this rigorously, although we will give some strong evidence that it is true. In the case of topological spaces, the morphisms are continuous maps.

We would like to gain an understanding of the topology of P. Let $U_\beta \subset X_\beta$ be an open set. Since p_β is continuous, the set

$$p_\beta^{-1}(U_\beta) = \prod_\alpha U_\alpha \quad \text{such that} \quad U_\alpha = X_\alpha \quad \text{for all} \quad \alpha \neq \beta$$

must be open in P. Since topologies are closed under finite intersections, every subset of X of the form

$$U = \prod_\alpha U_\alpha \quad \text{such that} \quad U_\alpha = X_\alpha \quad \text{for all but finitely many } \alpha \tag{A.1}$$

is also open in P. Therefore, the topology must be at least as fine as the topology generated by sets of the form A.1.

The mapping property characterizing the product P restricts how fine the topology on P can be. Let Z be a topological space with continuous maps f_α to each of the X_α. The universal property of the product states that there is a unique continuous map $f : Z \to P$. Therefore, the topology on P must be coarse enough that such a continuous map f exists.

We refer to the topology generated by sets of the form A.1 as the *product topology*. We will show that the product topology is sufficiently coarse to allow $f : Z \to P$ to be continuous. Let $U_\alpha \subset X_\alpha$ be open. Since $f_\alpha = p_\alpha \circ f$,

$$f_\alpha^{-1}(U_\alpha) = f^{-1}\left(p_\alpha^{-1}(U_\alpha)\right).$$

This set is open in Z because f_α is continuous. Since inverse mappings respect union and intersection,[2] any open in P (that is, any set of the form A.1) will be open in Z, so f is continuous. Therefore, P with the product topology satisfies the universal property of the product.[3] Since the product is unique, the product topology is the only topology on P that satisfies the universal property.

We now connect our development of product spaces to projective limits.

Proposition 104. Let X_k be topological spaces with inclusion maps

$$\varphi_{k+1,k} : X_{k+1} \longrightarrow X_k.$$

Let $P = \prod_k X_k$ with projection maps

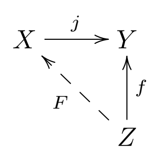
$$p_k : P \longrightarrow X_k.$$

Let X be the subspace of P (with the subspace topology) on which the projections p_k are compatible with the transition maps $\varphi_{k+1,k}$. That is

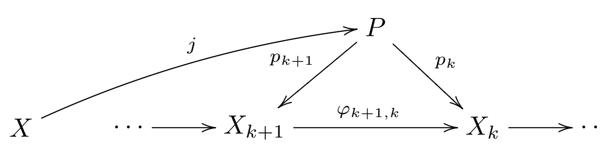
$$X = \{x \in P \,:\, \varphi_{k+1,k}(p_{k+1}(x)) = p_k(x) \text{ for all } k\}.$$

Then $X = \lim_k X_k$.

Remark 105. In order to prove the proposition, it will be useful to have a characterization of the subspace topology through mapping properties. Let Y be a topological space and let $X \subset Y$ with inclusion map $j : X \to Y$. The subspace topology on X is characterized by the following universal property. For any topological space Z and continuous map $f : Z \to Y$ such that $f(Z) \subset X$, f factors through the inclusion j. That is, the following diagram commutes:



Proof (of proposition). Let $j : X \to P$ be the inclusion map. Then we have the following diagram:

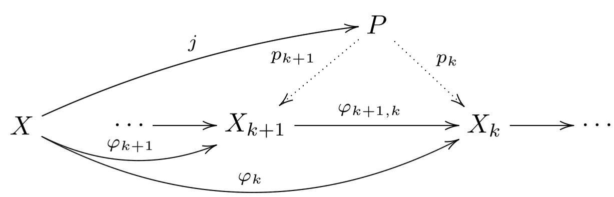


[2]That is, given a map $g : A \to B$, for any subsets $s_1, s_2 \subset B$,

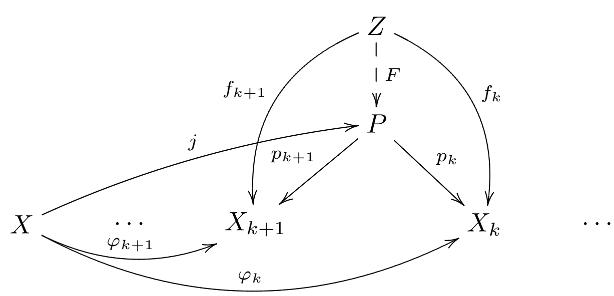
$$f^{-1}(s_1 \cup s_2) = f^{-1}(s_1) \cup f^{-1}(s_2) \quad \text{and} \quad f^{-1}(s_1 \cap s_2) = f^{-1}(s_1) \cap f^{-1}(s_2).$$

[3]It is important to note that we have not proven the existence of the product for topological spaces. We have only shown that if the product exists, it must have the product topology. In order to show that products of topological spaces exist, we could show that products of *sets* exist, since a topological space is merely a set with a topology. We have shown that given a set that satisfies the universal property, there is a topology on the set that is compatible. The products of sets do indeed exist, and the product is the familiar Cartesian product, although verifying this would be tangential to our current aim.

Note that the maps p_k do *not* necessarily commute with the transition maps $\varphi_{k+1,k}$. Define the maps $\varphi_k : X \to X_k$ to be the restrictions of the p_k to X. Then we have the following diagram, where all solid arrows commute:



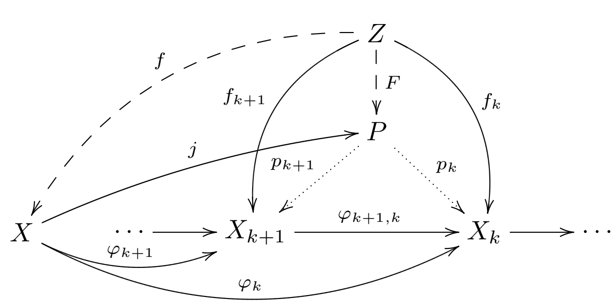
Let Z be any topological space with continuous maps $f_k : Z \to X_k$ that do commute with the transition maps $\varphi_{k+1,k}$. Then, by the universal property of the product, there exists a continuous map $F : Z \to P$ such that all triangles commute in the following diagram:



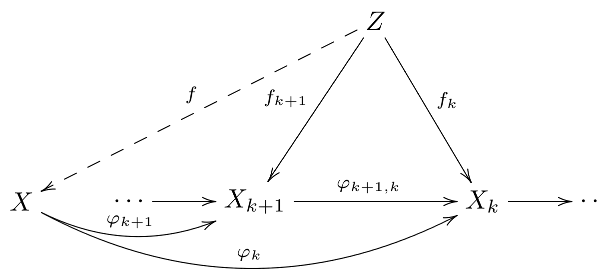
We claim that since the maps f_k are compatible with the $\varphi_{k+1,k}$, the image $F(Z) \subset X$. To see this, we note that for any $z \in Z$,

$$p_k(F(z)) = f(z) = \varphi_{k+1,k}(f_i(z)) = \varphi_{k+1,k}(p_{k+1}(F(z))) \quad \text{for all} \quad k.$$

Therefore, by the definition of X, $F(z) \in X$ so $F(Z) \subset X$. Since X is a subspace of P with the subspace topology, there exists a map $f : Z \to X$ such that F factors through j. Then all of the solid and dashed (but not necessarily dotted) arrows in the following diagram commute:



Now if we remove P from the picture, we obtain the commutative diagram



which we recognize as encoding the universal property of the (projective) limit. Since X satisfies the universal property of the limit,

$$X = \lim_k X_k.$$

In particular, X is a subspace of P with the subspace topology. □

Corollary 106. Let X_n be topological spaces and let $X = \lim_n X_n$. Then every open set $U \subset X$ is of the form

$$\bigcap_{j=1}^{k} \varphi_{n_j}^{-1}(U_{n_j}) \quad \text{for open sets} \quad U_{n_j} \subset X_{n_j}.$$

In particular, the intersection is finite.

Proof. From our previous discussion of the product space $P = \prod_n X_n$, the open sets in P are of the form

$$\bigcap_{j=1}^{k} p_{n_j}^{-1}(U_{n_j}) \quad \text{for open sets} \quad U_{n_j} \subset X_{n_j}.$$

Then the corollary is an immediate consequence of the fact that X is a subspace of P with the subspace topology. □

Remark 107. All of the discussion so far in this section has been concerned with topological spaces. All of the results, however, hold in the context of topological vector spaces as well. In particular, we will use the previous corollary in the context of Banach spaces in the following section. Of course, we will need to know the following fact about topological vector spaces in order to use the corollary: *Products of topological vector spaces exist.*

A.2 Limits and Duals of Banach Spaces

The following lemma will get us most of the way to a proof of the main result in this appendix.

Lemma 108. Let B_k be a sequence of Banach spaces with transition maps

$$\varphi_{k+1,k} : B_{k+1} \longrightarrow B_k$$

and let $X = \lim_k B_k$ with projections $\varphi_k : X \to B_k$. Let $\lambda : X \to \mathbb{C}$ be a continuous linear function on X, $\lambda \in X^*$. Then there exists an index j such that λ factors through B_j. That is,

$$\lambda = \lambda_j \circ \varphi_j \quad \text{for some} \quad \lambda_j \in (B_j)^*.$$

Proof. Let $U \subset X$ be a neighborhood of 0 such that $\lambda(U) \subset \Delta$, where Δ is the unit circle centered at the origin in $\mathbb{C}$. Since U is an open set in X, by Corollary 106, there are finitely many indices $i_1, \ldots, i_n$ and open subsets

$$V_{i_j} \subset X_{i_j} \quad \text{for} \quad 1 \leq j \leq n$$

such that

$$\bigcap_{j=1}^{n} \varphi_{i_j}^{-1}(V_{i_j}) \subset U.$$

Now let $m \in \mathbb{Z}$ with $m \geq \max\{i_1, i_2, \ldots, i_j\}$. Define the open subset of B_m

$$V_m = \bigcap_{j=1}^{n} \varphi_{m,i_j}^{-1}.$$

By the compatibility relations

$$\varphi_{i_j}^{-1} = \varphi_m^{-1} \circ \varphi_{m,i_j}^{-1},$$

we have

$$p_m^{-1}(V_m) \subset U.$$

Note that since λ is linear, for any $\varepsilon > 0$

$$\lambda(\varepsilon \cdot \varphi_m^{-1}(V_m)) = \varepsilon\lambda(p_m^{-1}(V_m)) \subset \varepsilon \cdot \Delta.$$

We claim that λ factors through $\varphi_m(X)$ with the subspace topology from B_m. Before proving the claim, we simplify notation. Let $\varphi : X \to B$ and $V \subset B$ with

$$\lambda(\varphi^{-1}(V)) \subset \Delta.$$

With the new notation, we restate the claim: λ factors through $\varphi : X \to B$ as a continuous linear map.

In order to show that λ factors through φ, it suffices to show that for any $x, x' \in X$ if $\lambda x \neq \lambda(x')$ then $\varphi(x) \neq \varphi(x')$. [4] Since all maps are linear, this is equivalent

[4]To see that this is true, suppose

$$\lambda(x) = \lambda(x') \implies \varphi(x) = \varphi(x') \quad \text{for all} \quad x, x' \in X.$$

Then the map

$$\mu : B \to \mathbb{C}$$

defined by

$$\mu(y) = \lambda(\varphi^{-1}(y)) \quad \text{for all} \quad y \in B$$

is well-defined. Further, μ is linear because λ and φ are linear.

to the statement that [5] $\ker\varphi \subset \ker\lambda$. By the linearity of λ,

$$\lambda\left(\frac{1}{n}\cdot\varphi^{-1}(V)\right) \subset \frac{1}{n}\Delta.$$

Therefore

$$\lambda\left(\bigcap_{n\in\mathbb{Z}^+}\frac{1}{n}\cdot\varphi^{-1}(V)\right) \subset \bigcap_{n\in\mathbb{Z}^+}\frac{1}{n}\Delta = \{0\}.$$

Then

$$\bigcap_{n\in\mathbb{Z}^+}\varphi^{-1}\left(\frac{1}{n}V\right) = \bigcap_{n}^{\infty}\frac{1}{n}\cdot\varphi^{-1}(V) \subset \ker\lambda.$$

Let $x \in \ker\varphi \subset X$. Then

$$\varphi(x) \in \frac{1}{n}V \quad \text{for all} \quad n \in \mathbb{Z}^+.$$

Hence

$$x \in \bigcap_{n\in\mathbb{Z}^+}\varphi^{-1}\left(\frac{1}{n}V\right) \subset \ker\lambda.$$

Therefore, $\ker\varphi \subset \ker\lambda$, so λ factors through $\varphi(X) \subset B$ via a linear map $\mu : \varphi(X) \to \mathbb{C}$.

To see that μ is continuous, fix $\varepsilon > 0$. Then by the linearity of μ,

$$\mu(\varepsilon\cdot V) = \varepsilon\cdot\mu(V) \subset \varepsilon\cdot\Delta$$

which shows that μ is continuous. Hence, $\lambda = \mu\circ\varphi$, so λ factors through $\varphi(X) \subset B$.

What remains to be shown is that we can extend μ to be a continuous linear function on the whole space B. First we show that we can extend μ to be a continuous linear functional on $\overline{\varphi(X)}$, the completion of $\varphi(X)$. In particular, let y_i be a Cauchy sequence in $\varphi(X)$ that converges to $y \in \overline{\varphi(X)}$. Then we define[6] $\mu(y) = \lim_i \mu(y_i)$. Since $\overline{\varphi(X)}$ is a closed subspace of B, we can further extend μ to all of B by the Hahn-Banach theorem.[7] Therefore λ factors through $\mu \in B^*$. $\square$

Theorem. Let $X = \lim_k B_k$ be the limit of Banach spaces B_k. Then

$$(\lim_k B_k)^* \cong \operatorname{colim}_k B_k^*.$$

[5] To see that this is equivalent to the previous statement, note that

$$\lambda(x) = \lambda(x') \iff \lambda(x-x') = 0 \iff x - x' \in \ker\lambda$$

and similarly for φ.

[6] In order to justify this assertion, fix $\varepsilon > 0$. Let $U \subset B$ be a convex neighborhood of 0 such that $U = -U$ and $|\mu(x)| < \varepsilon$ for all $x \in U$. Then for all $y_i, y_j \in y + U$, we have $|\mu(y_i) - \mu(y_j)| = |\mu(y_i - y_j)|$. Since $y_i - y_j \in 2U$, $|\mu(y_i) - \mu(y_j)| < 2\varepsilon$. Therefore, the extension given by $\mu(y) = \lim_i \mu(y_i)$ is continuous. The linearity of the extension is an immediate consequence of the linear of μ restricted to $\varphi(X)$.

[7] The Hahn-Banach theorem states states that given a continuous linear functional μ defined on a closed subspace of a normed linear space B, μ extends to a continuous linear functional on all of B. See [10].

Proof. We consider the dual spaces of the colimitands, B_k^*. Each B_k^* is the space of continuous linear functionals on the Banach spaces B_k. By taking the pullbacks of the continuous linear maps

$$\varphi_{k+1,k} : B_{k+1} \longrightarrow B_k$$

we obtain maps

$$\varphi_{k+1,k}^* : B_k^* \longrightarrow B_{k+1}^*.$$

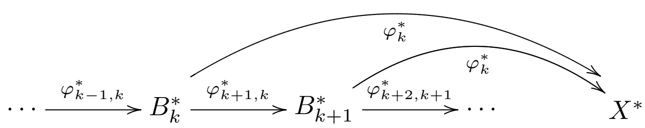
Therefore, we have an infinite chain of continuous linear maps

$$\cdots \xrightarrow{\varphi_{k-1,k}^*} B_k^* \xrightarrow{\varphi_{k+1,k}^*} B_{k+1}^* \xrightarrow{\varphi_{k+2,k+1}^*} \cdots$$

Since $X = \lim_k B_k$, there are continuous linear maps

$$\varphi_k : X \longrightarrow B_k.$$

Taking the pullbacks of the φ_k, we can write the diagram



All triangles in the above diagram commute because

$$(\varphi_{k+1}^* \circ \varphi_{k+1,k}^*) = (\varphi_{k+1,k} \circ \varphi_{k+1})^* = \varphi_k^*.$$

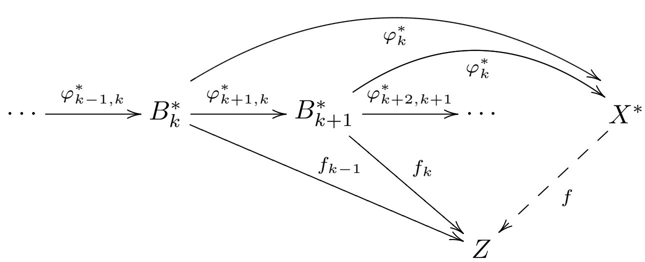
Let Z be any topological vector space with compatible maps

$$f_k : X_k^* : \longrightarrow Z \quad \text{for all } k.$$

We define the map $f : X^* \to Z$ by

$$f(\lambda) = f_k(\lambda_k) \quad \text{where} \quad \lambda = \lambda_k \circ \varphi_k.$$

The definition of f is justified by the lemma. We have the diagram

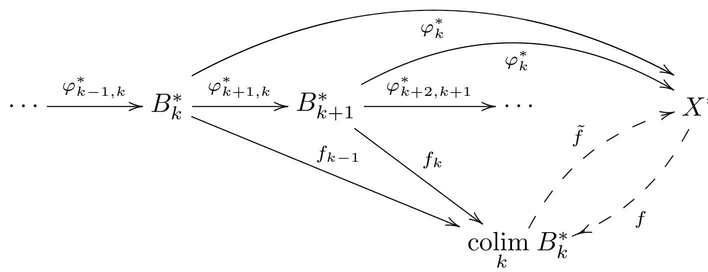


We still must verify that triangles involving the function f commute. To this end,

$$(f \circ \varphi_k^*)(\lambda_k) = f(\varphi^*(\lambda_k)) = f_k(\lambda_k)$$

where the final equality holds by the definition of f. Therefore all triangles in the above diagram commute.

We now consider the case where Z is the colimit of the spaces B_k^*. We then obtain the diagram



The map $\tilde{f}$ is implied by the universal property of the coproduct. We claim that the maps f and $\tilde{f}$ are mutual inverses of one another. This is because the compositions $f \circ \tilde{f}$ and $\tilde{f} \circ f$ are endomorphisms of X^* and $\operatorname{colim}_k B_k^*$ (respectively) such that all triangles in the diagram commute. By the proof of Proposition 25, this implies that $f \circ \tilde{f}$ and $\tilde{f} \circ f$ are the identity maps on their respective domains. Thus, f and $\tilde{f}$ are mutual inverses, so each is an isomorphism. Therefore

$$X^* = \operatorname*{colim}_k B_k^*$$

as desired. $\square$

Bibliography

[1] Anton Deitmar, *A first course in harmonic analysis*, Springer-Verlag, New York, NY, 2002.

[2] Paul Garrett, *Bigger diagrams for solenoids, more automorphisms, colimits*, November 2005, Available at
`http://www.math.umn.edu/$\sim$garrett/m/mfms/notes/03\_more\_autos.pdf`.

[3] ———, *Example of characterization by mapping properties: The product topology*, January 2006, Available at
`http://www.math.umn.edu/$\sim$garrett/m/mfms/notes/01\_product\_top.pdf`.

[4] ———, *Basic categorical constructions*, July 2008, Available at
`http://www.math.umn.edu/$\sim$garrett/m/fun/Notes/06\_categories.pdf`.

[5] ———, *Solenoids*, September 2008, Available at
`http://www.math.umn.edu/$\sim$garrett/m/mfms/notes/02\_solenoids.pdf`.

[6] ———, *Functions on circles*, March 2009, Available at
`http://www.math.umn.edu/$\sim$garrett/m/mfms/notes/09\_sobolev.pdf`.

[7] Lars Hörmander, *The analysis of linear partial differential operators (vol. i)*, Springer-Verlag, Berlin, Germany, 1983.

[8] Ian Richards and Heekyung Youn, *Theory of distributions: A non-technical introduction*, Cambridge University Press, Cambridge, Great Britain, 1990.

[9] Walter Rudin, *Functional analysis*, McGraw-Hill, New York, NY, 1973.

[10] ———, *Real & complex analysis*, McGraw-Hill, New York, NY, 1987.

[11] George Simmons, *Introduction to topology and modern analysis*, Robert E. Krieger Publishing Company, Malabar, FL, 1983.

[12] Elias Stein and Rami Shakarchi, *Fourier analysis: An introduction*, Princeton University Press, Princeton, NJ, 2003.

[13] Robert Strichartz, *A guide to distribution theory and fourier transforms*, CRC Press, Boca Raton, FL, 1994.