University of Stavanger

Faculty of Science and Technology

MASTER'S THESIS

Study program: Lektor i Realfag 8.-13.	Spring semester, 2021 Open access
Author: Anders Rikard Løvoll	Anders Rikard Løvoll …………………………………………
Faculty supervisor: Alexander Ulanovskii	
Thesis title: On the Convergence of Fourier Series in Hilbert Spaces – an introduction to sampling theory.	
Credits (ECTS): **30**	
Key words: Norm Hilbert Space Banach Space Fourier Series Fourier Transformation Convergence (L^2, piecewise, uniform) Shannon sampling theorem Riesz basis Kadets ¼ theorem	Pages: ………**76**……… + enclosure: ……**0**…… Stavanger, 15.06.2021.

On the Convergence of Fourier Series in Hilbert Spaces

An introduction to sampling theory

Anders Rikard Løvoll

A thesis presented for the degree of

Lektor i Realfag 8.-13.

Faculty of Science and Technology

University of Stavanger

Norway

June 15^{th}, 2021

Abstract

Løvoll, A. R. (2021). On the Convergence of Fourier Series in Hilbert Spaces. University of Stavanger.

Classical calculus, which traces its origins back to the 17^{th} century is concerned with the study of continuous change. However, classical calculus lacks the ability to adequately transform natural processes such as audio, radio waves and even images into their digital counterparts. Fast forward to the 18^{th} century, mathematical calculus was gradually encountering problems with some of its assumptions. The work done by Joseph Fourier (1768-1830) marked a turning point, commencing an investigation of such proportions that it ultimately lead to a restructuring into what we now call *Analysis.* Many branches of analysis emerged in the following centuries. Of particular note was the creation of functional analysis by Stefan Banach in the early $20'$s. In it, we find the proper tools for digitization of such natural processes. We provide an introduction to this intriguing theory, together with some examples of Banach- and Hilbert spaces. Including applications such as the reconstruction of band-limited signals through the famous Shannon-sampling theorem. Particularly we take advantage of Kadets 1/4-theorem to improve the sampling process.

Declaration

I declare that this Master thesis entitled ”On the Convergence of Fourier Series in Hilbert Spaces” has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise, by reference or acknowledgement, the work presented is entirely my own.

Acknowledgements

The creation of this thesis would not be possible without the insightful comments and patience received from Professor Alexander Ulanovskii. His expertise and guidance on the subject has been of magnificent help.

Contents

Chapter 1

Introduction

The study of Fourier series is an important field in modern mathematics. Through the course of this thesis we will provide an overview of central topics and use examples to conceptualize them such that a general foundation is reached. Our purpose is to give an account of the processing of signals.

The rest of the thesis is organized as follows. Firstly, we will provide a brief historical account of Joseph Fourier and his mathematical contributions. Following this we will introduce some important concepts from functional analysis - a relatively new branch of mathematics - which will allow us to delve deeper in the following sections. Particularly we will cover the notion of normed vector spaces, along with Banach- and Hilbert spaces. Afterwards we will discuss Fourier series and their convergence before we cover the Fourier transformation. Then we will give some examples of how the theory may be applied and solve some practical examples involving Fourier series. Lastly, we will discuss the famous Shannon-sampling theorem along with some interesting adaptations.

In this thesis much of the necessary background information has been inspired by the detailed book *Functions, Spaces and Expansions* by Ole Christensen, and the Master course *Mathematical Analysis 2* presented by professor Alexander Ulanovskii at the University of Stavanger. The preliminary structure is based on the structure presented there, and several examples and exercises have been found in the book. Our results differ from the work of Christensen by a change of course - away from Wavelet Analysis and Banach Theory - and towards the process of signal sampling and reconstruction.

Chapter 2

History

2.1 Joseph Fourier

The well renowned mathematician Jean-Baptiste Joseph Fourier was born 21 March 1768 in a small town in central France. In the town Auxerre he received a thorough education where he excelled in mathematics (Debnath, 2012). During his youth he contemplated whether to pursue a life of religion in order to become a priest, or to focus on his mathematical prowess. In 1787 he made the decision to pursue priesthood at St.Benedict's Abbey (O'Connor & Robertson, 1997), but did not lose interest in mathematics. A mere two years later at the age of 21, he had a change of heart and left the abbey in pursuit of immortality like many mathematicians before him.

Around the same time, the French Revolution stirred up the mindsets of many Frenchmen, including Fourier. He had an active role in promoting the French Revolution amongst his own surroundings. Though his participation at times led to personal pain, it paved him a path towards a successful research career in mathematics and physics (Debnath, 2012). After the end of the Revolution, he joined the Ècole Normale teacher-training school where he met many prominent mathematicians. More notably Pierre-Simon Laplace, Joseph-Louis Lagrange and Gaspard Monge. Having shown great promise and ability, he became a teacher at the Ècole Polytechnique where he showed his talents as a lecturer (O'Connor & Robertson, 1997). However, during his time as a teacher not much research was performed. In 1798 he was appointed as a scientific advisor in Napoleon's army and headed off to Egypt (O'Connor & Robertson, 1997). During this time he was in charge of the collation of all literary and scientific discoveries of Egyptian documents. His work led to the publication of Description de l'Égypte, a massive collection which sparked the start of Egyptology - the study of ancient Egypt by archaeologists (Egyptology: study of pharanoic Egypt, 2008).

He returned to Paris in 1801 to proceed his position as Professor of Analysis at the Ècole Polytechnique. Not long after he was once again contacted by Napoleon and promoted to Prefect of Isère in Grenoble. In this position he performed several tasks for Napoleon, but his greatest accomplishments were the drainage of $80,000\text{km}^2$ of the Swamp in Burgouin and the planning of a grand highway stretching from Grenoble in France to Turin in Italy. Unfortunately only the French section was ever built. Simultaneously, he began work on his memoir *On the Propagation of Heat in Solid Bodies* which may be considered his most important mathematical contribution. He completed it by 1807 and presented it to the Academy of Sciences of Paris. There were objections to the mathematical rigour of the memoir, which led to a mixed reception in two regards. Firstly, Lagrange and Laplace were not convinced of Fourier's implementation of the (at the time) controversial concept of an infinite sum of trigonometric series to *expand* an arbitrary function into what we now call a Fourier series. They found that there were some instances were the properties presented paradoxical qualities. Secondly, the manner of which Fourier had derived his equations for 'the transfer of heat through a continuous solid' were believed to be lacking physical principles (Debnath, 2012). As a result, Fourier spent the following years defending and revising his work. Though he got it accepted in 1812, it was not officially published until 1822 due to the scepticism and disbelief towards 'Fourier Analysis' in the mathematical community.

Finally, his research was published in 1822, along with the adjusted title *The Analytical Theory of Heat.* Here he presented a great quantity of problems, an example of which was to explain how the heat in a thin-layered material spreads as time passes. By assuming we have measured the temperature at certain locations along the boundary and the exterior at time

0, he then described how to use the subsequent *differential equation* to solve the problem

$$\frac{\partial u}{\partial t} = k\left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right]$$

where u is the temperature at a point (x, y) in the Euclidean plane at any time t, whilst k describes the proportionality of the diffusivity of the material, i.e. k represents how quickly the heat spreads in the material. Thereby, the differential equation presented above can be used to describe the heat conduction in two-dimensional objects (Struik, 2020). Furthermore, the solutions for the one-dimensional case were presented by infinite sums of trigonometric functions such as

$$y = \frac{a_0}{2} + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \cdots ,$$

which reminds us of the well known general form of Fourier series (see Section 8.1). Although he passed away at the age of 62 in 1830, his mathematical contributions lived on and were improved upon. Some of the more notable accomplishments include his work on Fourier series, the Fourier transform and calculating Fourier Integrals. According to Debnath these topics have had major impacts on physical engineering. Additionally, he has been credited with the initial discovery of the Greenhouse effect, where he suggested that the Earth's atmosphere had the property of capturing the heat of sunlight (Debnath, 2012, p. 8).

2.2 Historical development of Fourier series

Today, *Fourier series* is a well established and rigorous tool for mathematical analysis. As infinite series, they were originally introduced and investigated by mathematicians such as Euler, d'Alembert and Bernoulli in relation to describing the vibration of a string (Debnath, 2012). Their mathematical modelling highlighted a false presumption in the contemporary mathematical analysis and concept of a function. When Fourier used the same concepts of infinite series in his work on the propagation of heat, this resulted in the aforementioned mistrust and scepticism where the committee discarded this use of the concept. Namely that there was a problem with the convergence of Fourier's presented series, since the convergence of an infinite series were thought to be impossible. However, through his (and others) continued work and exploration it was eventually accepted. This signified a turning point in mathematical analysis to such a degree that the concept of a *function* and *infinity* would need to be re-evaluated (Bressoud, 2007).

Although the Fourier series expansion is a very useful tool, it has its limitations. Perhaps the most prominent one is that the function must be periodic. Many signals and waves which appear in nature do not follow a specific repeating pattern and are therefore classified as *aperiodic*. An obvious complication which arises for a function which is aperiodic, is that its frequencies may vary throughout the whole domain. Thankfully, another tool can be used in such situations, namely the Fourier transform. One of the main uses for the Fourier transform is that they speed up the calculations of certain Partial Differential Equations. (PDE) Much like how the introduction of $\log x$ by John Napier in 1614 simplified certain algorithmic operations, the Fourier transform may transform the PDE into a more favourable form where its mathematical manipulation becomes easier (Dominguez, 2016). Concretely, there are many applications of the transform within signal- and image processing, as well as interpreting certain events in quantum mechanics. The Fourier transform is currently defined as the generalization of complex Fourier series in the limit where the period L $\to \infty$. This is a (slight) variation to how Fourier defined the Fourier transform, as complex-valued functions were not usual at the time (Dominguez, 2016, p. 6). Although Joseph Fourier and other mathematicians talked about how the functions were transformed, the formal name *Fourier transform* was not used until after 1915 (Dominguez, 2016, p. 5). On the vector space $L^1(\mathbb{R})$ the Fourier transform takes $f \in L^1(\mathbb{R})$ and assigns a new function $\widehat{f} : \mathbb{R} \to \mathbb{C}$ by

$$\widehat{f}(x) := \int_{-\infty}^{\infty} e^{-2\pi i x t} f(t) dt$$

We often consider the above as an *operation*. This will become clear later on.

Chapter 3

Mathematical Background

In this chapter some basic definitions and results from Linear Algebra, as well as Real- and Complex Analysis will be presented. We will focus on a few selected topics and briefly define or recall other concepts as we go.

3.1 Euclidean spaces, $\mathbb{R}^n$ and $\mathbb{C}^n$

Let $\mathbb{R}$ denote the set of all real numbers, and $\mathbb{C}$ the set of all complex numbers. The n-dimensional Euclidean space is the set of all *sequences* of real numbers, often defined as: $\mathbb{R}^n := \{(x_1, x_2, ..., x_n) \,|\, x_k \in \mathbb{R},\, k = 1, ..., n\}$.

In this space we define the norm $\|\cdot\|$ of a vector $\mathbf{v} = (v_1, ..., v_n) \in \mathbb{R}^n$(or $\mathbb{C}^n$) by

$$\|\mathbf{v}\| := \left(\sum_{k=1}^{n} |v_k|^2\right)^{1/2} \tag{3.1}$$

where $|v_k|$ simply represents the absolute value of v_k. The meaning of the norm depends on the context in which it is given. In $\mathbb{R}^2$, the norm $\|(x_1, x_2)\| = \sqrt{x_1^2 + x_2^2}$ describes the length of the vector i.e., its Euclidean distance from the origin. However in other vector spaces (which we formally define later), the norm of an element $\mathbf{v} \in V$ could refer to anything from its size to the largest achievable value the vector attains. Consequently, for two vectors $\mathbf{x} = (x_1, x_2, ..., x_n)$ and $\mathbf{y} = (y_1, y_2, ..., y_n)$ one may prove the triangle inequality

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \tag{3.2}$$

which in $\mathbb{R}^2$ symbolizes that the shortest distance between two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ is a straight line between the points.

Definition 3.1.1. Given a point $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n$, the *ball* centred at $\mathbf{x}$ and with radius $r > 0$ is defined as the set

$$\begin{aligned} B(\mathbf{x}, r) :&= \{\mathbf{y} \in \mathbb{R}^n \big| \ \|\mathbf{y} - \mathbf{x}\| < r\} \\ &= \left\{ \mathbf{y} \in \mathbb{R}^n | \left(\sum_{k-1}^{n} |y_k - x_k|^2 \right)^{1/2} < r \right\} \end{aligned} \tag{3.3}$$

where $B\Big((0,0),1\Big)$ is the unit ball centered at the origin in $\mathbb{R}^2$.

Definition 3.1.2 (Basis)**.** Given $v_1, ..., v_n \in \mathbb{R}^n$ or $\mathbb{C}^n$, we say that $v_1, ..., v_n$ form a basis for $\mathbb{R}^n$ or $\mathbb{C}^n$ if:

$$v_1, ..., v_n \text{ are } \textit{linearly independent}$$
$$\&$$
$$\{v_k\}_{k=1}^{\infty} \textit{ spans } \mathbb{R}^n \text{ or } \mathbb{C}^n.$$

From Linear Algebra we know that the span collects *all* possible linear combinations of vectors in a vector space.

Example 3.1.1. Let $e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)$. Then $\{e_1, e_2, e_3\}$ forms an *orthonormal* basis for $\mathbb{R}^3$. This means that the vectors are all orthogonal, $\perp$, to each other and their norms $\|e_k\| = 1$.

Proof. We need to express any vector $\mathbf{v} \in \mathbb{R}^3$ as a sum of these basis vectors e_k. Observe that $xe_1 + ye_2 + ze_3 = (x, y, z) =: \mathbf{v}$, so $\{e_1, e_2, e_3\}$ spans $\mathbb{R}^3$. □

3.2 Vector spaces

A vector space V is a non-empty set of vector elements $\mathbf{v}$ which are contained within the set under the operations addition $(+)$ and scalar multiplication $(\cdot)$. This means that for the first operation $(+)$, any pair of elements $\mathbf{v}, \mathbf{w} \in V$ generate a third vector $\mathbf{v} + \mathbf{w} = \mathbf{u} \in V$. Furthermore, if we take a complex scalar $\alpha \in \mathbb{C}$, then the length of any vector in V may be scaled through the second operation $(\cdot)$, denoted by $\alpha\mathbf{v}$ or $\mathbf{v}\alpha$. Finally, any vector space must satisfy the so-called *vector axioms* for these operations.

We say that a vector space is finite dimensional whenever $\exists\, n$ linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n \in V$, for $n \in \mathbb{N}$ which span V. The dimension of such a vector space is $\underline{n}$. Otherwise, V is called infinite dimensional. In functional analysis the attention is centered more around such infinite dimensional vector spaces, as we will see later on. An example is the vector space of functions. For this purpose, we present the following definition:

Definition 3.2.1. ($C[a, b]$) Given a closed fixed interval $I = [a, b]$ where $a < b$ we denote $C[a, b]$ as the set of all continuous functions $f : [a, b] \to \mathbb{C}$.

It can be shown that $C[a, b]$ is a vector space of infinite dimension. Furthermore, any function $f \in C[a, b]$ is bounded and attains both its *supremum* and *infimum* on the interval $[a, b]$. It follows from a basic theorem that if f is a continuous function on a closed interval, then $|f|$ attains its max and min values. We can further specify what kind of continuous functions we consider. Denote by $C_0(\mathbb{R})$ the set of all continuous functions on $\mathbb{R}$ which tend towards 0 as $|x| \to \infty$. Denote by $C_c(\mathbb{R})$ the set of all continuous functions which have compact support in $\mathbb{R}$. Let us illustrate this with a short example:

Example 3.2.1. Let the following functions be defined on $\mathbb{R}$.

$$\begin{aligned}
f_1(x) &:= e^{-x^2}, \\
f_2(x) &:= x^3 + 2x + 4, \\
f_3(x) &:= \sin(\pi x)\chi_{[-2,2]}(x), \\
f_4(x) &:= \frac{1}{1 + x^2}.
\end{aligned}$$

Here the functions f_1 and f_4 belong to $C_0(\mathbb{R})$, since they go towards 0 as $x \to \infty$. Whereas $f_3 \in C_c(\mathbb{R})$ since its support $\text{supp} f_3 = [-2, 2]$ is compact, whilst f_2 is simply an element of $C(\mathbb{R})$, see Figure [3.1].

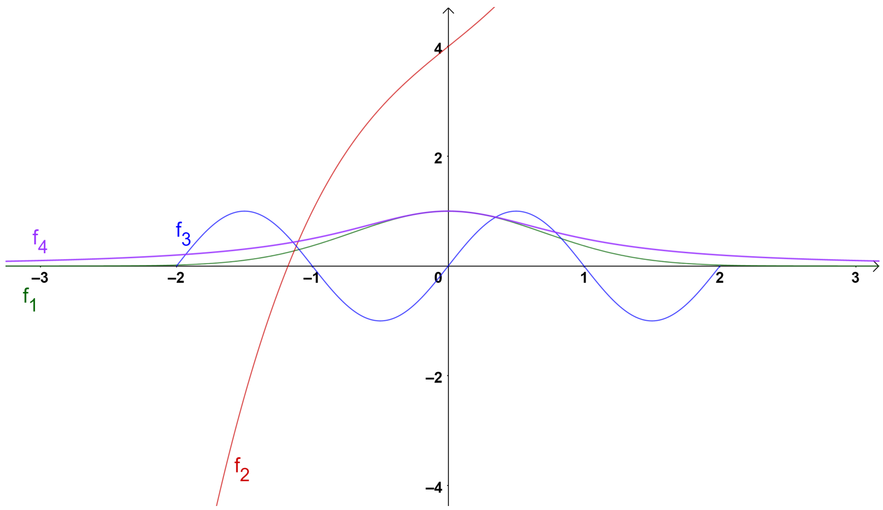

Figure 3.1: A sketch of the functions f_1, f_2, f_3 and f_4.

Definition 3.2.2 (The l^p-space)**.** The set of all complex sequences $\mathbf{v} = \{v_k\}_{k=1}^{\infty}$ where

$$\sum_{k=1}^{\infty} |v_k|^p < \infty, \quad p \geq 1, \tag{3.4}$$

is called the $l^p(\mathbb{N})$-space.

We may sometimes omit the specification of $\mathbb{N}$ as it is often clear from context. The l^p space has infinite dimension, where the more notable l^p-spaces are the l^1, l^2 and l^∞-spaces. We will also explore what happens when the summation is replaced by integrals.
Another operation which may be satisfied in certain vector spaces is the inner product which we will discuss more later on. Finally, for such vector spaces one may easily define the topological notions of open, closed, closure, a boundary, etc. As this is quite trivial, we skip these definitions.

Chapter 4

Normed Vector Spaces

In this chapter we will consider norms and linear operators. Our goal is to eventually establish a theory of vector spaces which allows us to work with vector spaces in a similar manner to $\mathbb{R}^n$. This applied structure will simplify the mathematical modelling in Fourier analysis.

4.1 The norm of a vector space

In the previous chapter we defined the norm of a vector $\mathbf{v} \in \mathbb{R}^n$ (see Definition 3.1). Let us now consider a general vector space V and define how *any* norm applies to elements of V.

Definition 4.1.1 (Norm)**.** A norm in a vector space V is a *function* which sends an element $\mathbf{v} \in V$ into $[0, \infty)$

$$\| \cdot \| : V \to [0, \infty) \tag{4.1}$$

which satisfies the following conditions:

(i) $\|\mathbf{v}\| \geq 0, \quad \forall\, \mathbf{v} \in V, \quad \text{and} \quad \|\mathbf{v}\| = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$

(ii) $\|\alpha \mathbf{v}\| = |\alpha|\, \|\mathbf{v}\|, \quad \forall\, \mathbf{v} \in V,\, \alpha \in \mathbb{C}$

(iii) $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|, \quad \forall\, \mathbf{v}, \mathbf{w} \in V$

If the norm comes from an inner product then $\big|\langle \mathbf{u}, \mathbf{v} \rangle\big|^2 \leq \|\mathbf{u}\|^2 \cdot \|\mathbf{v}\|^2, \quad \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ (or $\mathbb{C}^n$) where $\langle \mathbf{u}, \mathbf{v} \rangle$ represents the inner product between two vector quantities. In particular, if

$V = l^2(\mathbb{N})$ then

$$\langle \mathbf{u}, \mathbf{v} \rangle := \sum_{k=1}^{n} x_k \overline{y_k}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}. \tag{4.2}$$

Now we formally define a *normed* vector space V as a vector space equipped with a norm. This means that the norm may be applied to all the elements $\mathbf{v}$ inside the vector space. Conversely, there are many norms which may be induced upon a vector space V. We will consider several different norms throughout this thesis. Whenever we consider multiple vector spaces with different norms, a symbol in the index of the norm will be used, e.g., $\| \cdot \|_V$. In regards to the above conditions, it is often easy to check whether or not the two first conditions hold, whilst the triangle inequality may require some fidelity for certain norms.

Make special note that the 4^{th} property above is called the Cauchy-Schwartz inequality. It can be derived through the 3^{rd} property and holds in certain other vector spaces as well.

Example 4.1.1. Recall $C[a, b]$ the space of continuous functions, we claim that the supremum norm $\| \cdot \|_\infty$ given by

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)| = \max_{x\in[a,b]} |f(x)| \tag{4.3}$$

is well defined. The last equality holds since any $f \in C[a, b]$ attains its maximum value. Let us verify that expression (4.3) defines a norm in the vector space $C[a, b]$.

Proof. Firstly, we see that $\|f\|_\infty = 0$ if and only if (iff) $f \in C[a, b]$ such that $f = 0$. Clearly $|f(x)| \geq 0 \quad \forall f \in C[a, b]$. Also, if $\|f\| = 0$ then $f = 0$, which means that property (i) is satisfied. Secondly, (ii) is satisfied by checking that $\|\alpha f\|_\infty = \max_{x\in[a,b]} |\alpha f(x)| = |\alpha| \max_{x\in[a,b]} |f(x)| = |\alpha| \, \|f\|_\infty$. Lastly, if we consider two functions $f, g \in C[a, b]$, we should notice that if we pick any $x \in [a, b]$, the absolute value of the sums are equivalent to:

$$\begin{aligned} |f(x) + g(x)| &\leq |f(x)| + |g(x)| \quad \text{(by the triangle inequality)} \\ &\leq \max_{x\in[a,b]} |f(x)| + \max_{x\in[a,b]} |g(x)|. \end{aligned}$$

By using the definition of the supremum norm, we see that

$$\|f + g\|_\infty = \max_{x\in[a,b]} |f(x) + g(x)| \leq \|f\|_\infty + \|g\|_\infty.$$

and the claim holds. □

Example 4.1.2. Now for the vector space $l^p(\mathbb{N})$ from Definition 3.2.2, we claim that we can equip it with the following norm

$$\|\mathbf{v}\|_{l^p} = \left(\sum_{k=1}^{\infty} |v_k|^p\right)^{1/p} \tag{4.4}$$

where the index of $\|\cdot\|$ refers to the specific l^p space, $p \geq 1$.

Proof. It is easy to see that properties (i) and (ii) hold, thus we omit them from the proof. Now for (iii), we do the following. Let $\mathbf{v}, \mathbf{w} \in l^p(\mathbb{N})$, and using the Minkowskii's inequality for sequences observe that the normed sum becomes

$$\|\mathbf{v}+\mathbf{w}\|_{l^p} = \left(\sum_{k=1}^{\infty} |v_k + w_k|^p\right)^{1/p} \leq \left(\sum_{k=1}^{\infty} |v_k|^p\right)^{1/p} + \left(\sum_{k=1}^{\infty} |w_k|^p\right)^{1/p} = \|\mathbf{v}\|_{l^p} + \|\mathbf{w}\|_{l^p}$$

which satisfies (iii). □

Also notice that for $p = 2$, the norm in Equation (4.4) is analogous to the Euclidean norm for infinite dimensions, whilst if $p = \infty$ it can be shown that the supremum norm is well defined for l^∞.

4.2 Linear operators

Consider the structure of $\mathbb{R}^n$ and $\mathbb{C}^n$, notice that they are *linear*, which means that the elements of the sets all satisfy the axioms of linearity (see 1. and 2. below). On finite n-dimensional vector spaces, we often consider linear transformations by $n \times n$ matrix operations. Unfortunately, on infinite dimensional vector spaces the concept of an $\infty \times \infty$ matrix loses its practicality. Therefore we will now see how to establish linear operators in normed vector spaces. Generally, a linear *operator* T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function $T : \mathbb{R}^n \to \mathbb{R}^m$ which satisfies the properties:

1. $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

2. $T(\alpha\mathbf{x}) = \alpha T(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n,\ \alpha \in \mathbb{R}.$

If we combine these two properties and apply them to the topic of normed vector spaces, we can define an operator T which to any element $\mathbf{v} \in V_1$ associates a new element $T\mathbf{v} \in V_2$. Here V_1, V_2 are two complex vector spaces not necessarily equal, i.e., $V_1 \neq V_2$. Any such operator is linear if

$$T(\alpha\mathbf{v} + \beta\mathbf{w}) = \alpha T(\mathbf{v}) + \beta T(\mathbf{w}), \quad \alpha, \beta \in \mathbb{C}, \quad \mathbf{v}, \mathbf{w} \in V_1 \tag{4.5}$$

is satisfied. Consider the following example.

Example 4.2.1. Let $T : C[0,1] \to \mathbb{C}$ where $T_1 f = \big(f(0), f(1)\big)$ and $T_2 f = \big(f(0), f(\frac{1}{2})^2\big)$

1. Is T_1 a linear operator? Choose $f, g \in C[0,1]$ and $\alpha, \beta \in \mathbb{C}$, then

$$\begin{aligned} T_1\big(\alpha f + \beta g\big) &= \big(\alpha f(0) + \beta g(0), \alpha f(1) + \beta g(1)\big) = \\ &= \big(\alpha f(0), \alpha f(1)\big) + \big(\beta g(0), \beta g(1)\big) = \alpha T_1 f + \beta T_1 g \qquad \checkmark \end{aligned}$$

   which satisfies Equation (4.5) $\Rightarrow T_1$ is linear.

2. How about T_2? Following the same procedure as above, we get that

$$\begin{aligned} T_2\big(\alpha f + \beta g\big) &= \big(\alpha f(0) + \beta g(0), \big[\alpha f(\frac{1}{2})\big]^2 + \big[\beta g(\frac{1}{2})\big]^2\big) \\ &= \big(\alpha f(0), \alpha^2 f(\frac{1}{2})^2\big) + \big(\beta g(0), \beta^2 g(\frac{1}{2})^2\big) \neq \alpha T_2 f + \beta T_2 g \qquad \not\checkmark \end{aligned}$$

   which shows that T_2 is not linear.

Since the vector space V is equipped with a norm, the next step is to equip a linear operator with a norm. This may be done in the following manner.

Definition 4.2.1 (The norm of an operator)**.** For two normed vector spaces, V_1, V_2 and a linear operator $T : V_1 \to V_2$, we denote the norm of the operator by $\|T\|$. Then $\|T\|$ is the smallest number $K \geq 0$ which satisfies:

$$\|T\| = \sup_{\mathbf{v} \neq 0} \left(K \frac{\|T\mathbf{v}\|_{V_2}}{\|\mathbf{v}\|_{V_1}} \right) \leq K \frac{\|T\mathbf{v}\|}{\|\mathbf{v}\|} \qquad \forall\, \mathbf{v} \in V_1. \tag{4.6}$$

Observe that any such K bounds the operator. If such a number K does not exist, we say that $\|T\| = \infty$. To better understand this definition, we consider the following example.

Example 4.2.2. What is the norm of T_1 from the previous example?

$$\|T_1\| = \sup_{0\le x\le 1} \frac{\|T_1 f\|_{\mathbb{C}}}{\|f\|_{C[0,1]}} = \sup_{0\le x\le 1} \frac{\|\big(f(0), f(1)\big)\|_{Euclidean\,norm}}{\|f\|_\infty}$$

$$= \sup_{0\le x\le 1} \frac{\sqrt{|f(0)|^2 + |f(1)|^2}}{\sup\limits_{0\le x\le 1} |f(x)|} \qquad \text{where} \begin{cases} |f(0)| & \le \sup\limits_{0\le x\le 1} |f(x)|, \\ & \& \\ |f(1)| & \le \sup\limits_{0\le x\le 1} |f(x)|. \end{cases} \qquad (\star)$$

$$\le \sup_{0\le x\le 1} \frac{\sqrt{\|f(x)\|_\infty^2 + \|f(x)\|_\infty^2}}{\|f(x)\|_\infty} \le \frac{\sqrt{2}\,\|f(x)\|_\infty}{\|f(x)\|_\infty} = \sqrt{2}.$$

So $\|T_1\| \le \sqrt{2}$, and we know that the linear operator T_1 is bounded by this number (the norm cannot achieve a greater value). Now it remains to check if $\exists f \in C[0,1]$ such that $\|T_1\| = \sqrt{2}$. Suppose $f(x) = 1$, then

$$\|T_1\| = \frac{\sqrt{|1|^2 + |1|^2}}{\sup\limits_{0\le x\le 1} |1|} = \frac{\sqrt{2}}{1} = \sqrt{2}. \qquad \checkmark$$

Regarding $(\star)$: Here we are considering <u>all</u> possible functions $f \in C[0,1]$ at the same time.

Remark. In cooperation with the Definition 4.2.1 on the norm of an operator and the above result, we see that $\sqrt{2}$ is the smallest number which guarantees that $\|T_1 f\| \le \sqrt{2}\,\|f\|$ holds $\forall\, f \in C[0,1]$ s.t. $T_1 f \in \mathbb{C}$. However, not all operators are bounded on *every* normed vector space.

Whether the dimension of a vector space V is finite or infinite may have significant impacts on how we need to treat or deal with an operator in the general sense. Yet, if we can show that it is bounded, then we can treat the operator T similarly to how we would in $\mathbb{R}^n$. Like a regular *map* from linear algebra, a linear operator may also be classified as injective and/or surjective, bijective if both are satisfied. Let us apply these definitions to the operator $T_1 : C[0,1] \to \mathbb{C}$ from the previous examples.

Example 4.2.3. Let $T : C[0,1] \to \mathbb{C}, \quad Tf = \big(f(0), f(1)\big)$

1. Is T injective?

   First recall that T injective $\Leftrightarrow Tf = (0,0) \Rightarrow f = \mathbf{0}$. However, we can clearly find a counterexample, e.g.,

$$f_1(x) := x(x-1) \quad \Longrightarrow \quad T\big(x(x-1)\big) = \big(0\cdot(-1), 1\cdot 0\big) = (0,0) = \mathbf{0} \qquad \checkmark$$

2. Is T surjective?

   To satisfy this, we need that $\forall w \in \mathbb{C}, \quad \exists v \in C[0,1]$ such that $T(v) = w$.

   Now if we pick any $w := (\alpha, \beta) \in \mathbb{C}$, notice that for

$$f_2(x) := (1-x)\alpha + x\beta \quad \Longrightarrow \quad T\big(f_2(x)\big) = (\alpha + 0\cdot\beta, 0\cdot\alpha + \beta) = (\alpha, \beta) = w \qquad \checkmark$$

3. T is <u>not</u> bijective since it is not injective **and** surjective.

Example 4.2.4. Let $\mathbf{u} := \{u_1, u_2, ...\}$ be a bounded sequence of complex numbers, and let $T : l^1(\mathbb{N}) \to l^1(\mathbb{N})$ be defined by

$$T\{v_k\}_{k=1}^{\infty} = \{u_k v_k\}_{k=1}^{\infty}.$$

Let us prove that T is a bounded operator, and show that $\|T\| = \sup_{k=1,2,3,...} |u_n|$.

1. To begin, we denote by $\mathbf{v} = \{v_k\}_{k=1}^{\infty}$, and $\mathbf{w} = \{w_k\}_{k=1}^{\infty}$, then

$$\begin{aligned} T(\alpha\mathbf{v} + \beta\mathbf{w}) &= T(\alpha v_1 + \beta w_1, \alpha v_2 + \beta w_2, ...) \\ &= u_1(\alpha v_1 + \beta w_1), u_2(\alpha v_2 + \beta w_2), ... \\ &= \alpha(u_1 v_1, u_2 v_2, ...) + \beta(u_1 w_1, u_2 w_2, ...) = \alpha T\mathbf{v} + \beta T\mathbf{w}, \qquad \checkmark \end{aligned}$$

   which confirms that T is linear.

2. To show that T is bounded, we make note that $|u_k| \le \sup |u_k|$. It follows that

$$\begin{aligned} \|T\mathbf{v}\|_{l^1} = \|u_1 v_1, u_2 v_2, ...\|_{l^1} &= \sum_1^{\infty} |u_k v_k| = \sum_1^{\infty} |u_k| \cdot |v_k| \\ &\le \sup |u_k| \sum_1^{\infty} |v_k| = \sup |u_k| \cdot \|\mathbf{v}\|_{l^1}, \quad k = 1, 2, ... \\ \Longrightarrow \|T\| = \frac{\|T\mathbf{v}\|}{\|\mathbf{v}\|} &\le \sup |u_k| \end{aligned}$$

   which shows that $\|T\|$ is bounded. Now we choose an element $\mathbf{v} = (0, 0, ..., \overbrace{1}^{k}, 0, 0, ...)$ such that $T\mathbf{v} = (0, 0, ..., u_k, 0, 0, ...)$. But this implies that $\|T\mathbf{v}\|_{l^1} = |u_k|$, where the definition of a supremum requires that $|u_k| > \sup |u_n| - \varepsilon$. But this implies that $\|T\| = \frac{|u_k|}{1} \geq \sup |u_n| - \varepsilon$. Consequently, we arrive at the following relation

$$\sup |u_k| \geq \quad \|T\| \quad \leq \sup |u_k| - \varepsilon$$

which squeezes the norm into satisfying $\|T\| = \sup |u_k|$.

The topic of linear operators is one which we will return to and expand upon throughout this thesis.

Chapter 5

Banach Spaces

The concept of a Banach space was first introduced by Stefan Banach in his 1920 dissertation where he formally called them *class* E-spaces. His theories and notions are said to have laid the foundation of what we now call *functional analysis* (O'Connor & Robertson, 2000). As general Banach spaces can be quite complicated to deal with, we will mainly focus on the analytic structures which hold for all Banach spaces.

5.1 A complete vector space

Before we give a formal definition of a Banach space, let us consider some important conditions which need to be fulfilled by the elements of a vector space V.

Definition 5.1.1 (Cauchy sequence)**.** Let $\{v_k\}_{k=1}^{\infty}$ be a sequence in a normed vector space V. Such a sequence is said to be *Cauchy* if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \text{such that} \quad \|v_i - v_j\| \leq \varepsilon \quad \forall i, j \geq N. \tag{5.1}$$

Thus the elements of such a sequence become arbitrarily close to one another. There are many examples where a Cauchy sequence does not converge towards an element within V. However, for such cases it is possible to show (using mathematics beyond the scope of this thesis) that the Cauchy sequence will converge in an extended space V' which expands the initial space V. The next example will illustrate the Cauchy-requirement.

Example 5.1.1. Let us give an example of a sequence which is Cauchy, and one which is not in the normed vector space $\mathbb{R}$ where the norm of an element $\mathbf{x} \in \mathbb{R}$ corresponds to its absolute value, i.e., $\|\mathbf{x}\| = |\mathbf{x}|$.

(i) Consider the sequence $\{x_k\}_{k=1}^{\infty}$ where $x_k := k^2$, $k \in \mathbb{N}$. Then

$$\|x_i - x_j\| = |x_i - x_j| = |i^2 - j^2| \geq \begin{cases} 3, & \text{if} \quad i \neq j. \\ 0, & \text{if} \quad i = j. \end{cases}$$

and we see that the sequence is *not* Cauchy because a simple choice of $\varepsilon = 1$ would not satisfy Equation (5.1).

(ii) Now if $y_k := \frac{k}{k+1}$ we must find N such that $\|y_i - y_j\| < \varepsilon$ for all $i, j \geq N$. We have

$$\|y_i - y_j\| = \left|\frac{i}{i+1} - \frac{j}{j+1}\right| = \left|\frac{(j+1)i - j(i+1)}{(i+1)(j+1)}\right| = \left|\frac{i-j}{(i+1)(j+1)}\right| =: (*)$$

Now assume $i > j$ and notice that $\frac{i-j}{i+1} < 1$, then

$$(*) = \frac{i-j}{(i+1)(j+1)} < \frac{1}{j+1}.$$

Choose N to be any integer s.t. $N > \frac{1}{\varepsilon} \Leftrightarrow \varepsilon > \frac{1}{N}$. Thereby $\frac{1}{j+1} \leq \frac{1}{N} < \varepsilon$, so for every $j \geq (N-1)$ we get that $|y_i - y_j| < \varepsilon$, so y_k is Cauchy.

Notice that the sequence y_k converges towards 1 since $\lim_{k\to\infty} \frac{k}{k+1} = 1$. Then if we were to change the normed vector space $\mathbb{R}$ into $\mathbb{R}/\{1\}$, then the sequence, although Cauchy, would no longer be converging towards an element of the space. To avoid all such cases and thereby be able to properly work with limits, we present the following definition.

Definition 5.1.2 (Completeness)**.** A vector space V is called *complete* whenever all Cauchy sequences $\{v_k\}_{k=1}^{\infty}$ in V converge to an element $\mathbf{v} \in V$.

Now we are able to give a formal definition of a Banach space.

Theorem 5.1.1 (Banach space)**.** *A normed vector space V is a Banach space if it is complete. We denote such spaces by B.*

So the main advantage a Banach space provides, is that it becomes much easier to check whether or not a sequence of vectors converges. If we can show that $\{v_k\}_{k=1}^{\infty} \in B$ is Cauchy, then we immediately know that it converges. Particularly, we know that it converges to an element of the space, and we can give the following corollary.

Corollary 5.1.1.1. *If B is a Banach space $\Leftrightarrow$ every Cauchy sequence $\{v_k\}_{k=1}^{\infty}$ converges to an element of B.*

Hence, we say that the space is *closed* under limiting operations. A final result from completeness is presented in the following lemma.

Lemma 5.1.2. *Let $\{v_k\}_{k=1}^{\infty}$ be a Cauchy sequence in a normed vector space V. Then it is bounded, i.e., $\exists c > 0$ such that $\|v_k\| < c, \quad k \in \mathbb{N}$.*

Proof. Suppose $\{v_k\}_{k=1}^{\infty}$ is a Cauchy sequence, and observe from the triangle inequality that $\|v_k\| = \|v_k - v_j + v_j\| \leq \|v_k - v_j\| + \|v_j\|$. Now fix $\varepsilon := 1$ such that $\|v_i - v_j\| \leq 1, \forall i, j \geq N$. By inserting this into the initial observation, we get that: $\|v_k\| < 1 + \|v_j\|$. Choose $j = N + 1$ such that $\|v_k\| < 1 + \|v_{N+1}\|$ and notice that all terms beyond the N^{th} is bounded, provided $k > N$. So $\{v_k\}_{k=N}^{\infty}$ is bounded. It remains to check the finitely many terms before the N^{th} term, but clearly $\|v_k\| \leq \max(\|v_1\|, ..., \|v_N\|)$. Then we can finally find the constant c which bounds all the terms of $\{v_k\}_{k=1}^{\infty}$. $\square$

To sum up, we see that in a Banach space B, every Cauchy sequence **converges**, and they converge **towards** elements of B.

5.2 Examples of Banach spaces

Some of the vector spaces we have already considered are in fact Banach spaces. $\mathbb{R}^n$ and $\mathbb{C}^n$ induced with the Euclidean norm, $C[a, b]$ with the supremum norm and the sequence spaces $l^p(\mathbb{N})$ with their respective p-norms are all examples of Banach spaces. Actually, it has been proven that all *finite*-dimensional normed vector spaces are *complete* and therefore Banach (Christensen, 2010, p. 49). Let us instead prove that a particular l^p-space is Banach, for example the l^∞ space of all bounded sequences:

$$l^\infty(\mathbb{N}) := \left\{ \{x_k\}_{k=1}^{\infty} \,\middle|\, x_k \in \mathbb{C}, \quad \sup_{k\in\mathbb{N}} |x_k| < \infty \right\}. \tag{5.2}$$

Proof. We wish to prove that $l^\infty(\mathbb{N})$ is a Banach space w.r.t. to the norm $\|\mathbf{x}\|_{l^\infty} = \sup_{k\in\mathbb{N}} |x_k|$. First we must show that $l^\infty(\mathbb{N})$ is a vector space. But clearly $l^\infty(\mathbb{N})$ is a set of sequences where for any pair $\big(\{x_k\}, \{y_k\}\big)$ their sum and products with scalars are contained within the set.

Let us check that the norm is well defined. We leave it to the reader to verify that properties (i) & (ii) are satisfied. Now for (iii): let $\mathbf{x}, \mathbf{y} \in \mathbb{C}$ then by using the triangle inequality along with maximizing the result through the supremum, we see that

$$\|\mathbf{x}+\mathbf{y}\|_{l^\infty} = \sup_{k\in\mathbb{N}} |x_k + y_k| \leq \sup_{k\in\mathbb{N}} \big(|x_k| + |y_k|\big) \leq \sup_{k\in\mathbb{N}} |x_k| + \sup_{k\in\mathbb{N}} |y_k| = \|\mathbf{x}\|_{l^\infty} + \|\mathbf{y}\|_{l^\infty}.$$

Since the result is finite, the norm is well defined. □

We will now consider a more concrete example. Consider the norm: $\|(x_1, x_2)\| := \max_{0\leq x\leq 1} \{|x_1|, |x_2|\}$ in $\mathbb{R}^2$. Notice that this is the supremum norm for the Real-plane. Denote $\mathbf{x} := (x_1, x_2)$. Let us show that it satisfies the conditions for being a norm and provide a geometric representation of its unit ball $B\Big((0,0), 1\Big)$:

(i) $\|\mathbf{x}\| = \max_{0\leq x\leq 1} \{|x_1|, |x_2|\} \geq \mathbf{0} \quad \& \quad \|\mathbf{x}\| = \mathbf{0} \Leftrightarrow x_1 = x_2 = 0$ ✓

(ii)
$$\begin{aligned} \|\alpha\mathbf{x}\| &= \max_{0\leq x\leq 1} \{|\alpha x_1|, |\alpha x_2|\} = \max_{0\leq x\leq 1} \{|\alpha|\,|x_1|, |\alpha|\,|x_2|\} \\ &= |\alpha| \max_{0\leq x\leq 1} \{|x_1|, |x_2|\} = |\alpha|\,\|\mathbf{x}\| \end{aligned}$$ ✓

(iii)
$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\| &= \max_{0\leq x\leq 1} \{|x_1+y_1|, |x_2+y_2|\} \quad \text{using the triangle inequality we get :} \\ &\leq \max_{0\leq x\leq 1} \{|x_1|+|y_1|, |x_2|+|y_2|\} \leq \max_{0\leq x\leq 1} \{|x_1|, |x_2|\} + \max_{0\leq x\leq 1} \{|y_1|, |y_2|\} \\ &= \|\mathbf{x}\| + \|\mathbf{y}\| \end{aligned}$$ ✓

So $\|(x_1, x_2)\|$ is a norm in $\mathbb{R}^2$. Usually we visualize a unit ball in the plane as a circle. However we claim that the following Figure [5.1] represents a unit ball.

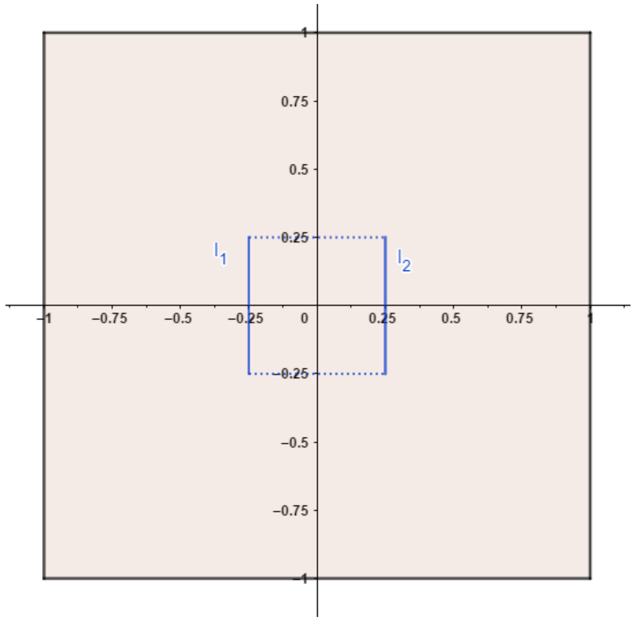

Figure 5.1: A representation of the unit ball in the supremum norm.

Intuitively, our visual sense tells us that this is a square. But if we use the definition given by the norm, we will see that this is in fact the representation of a unit ball (circle)! Think of it as following. Suppose the largest value is $r = 0.25$. Fix $x_1 := -0.25$ and let $x_2 \in (-0.25, 0.25)$. Now the norm $\|(x_1, x_2)\| \leq 0.25$ for all these points, and notice that by marking all the points given by x_2, we can draw the line l_1 shown in Figure 5.1. Then choose $x_1 := 0.25$ and repeat the same procedure for all possible x_2, this should draw the line l_2. If we switch the roles of x_1 and x_2, it should now be clear that we will be able to draw the missing lines of the square centred at $(0, 0)$ with side lengths $\frac{1}{2}$. Consider this as one *iteration*, then by performing more iterations $\forall r \in [0, 1)$ we see that we can mark all values inside the square. Since these are all the points from the origin whose *norm* is strictly less than 1, we see that the figure does represent the unit ball $B\Big((0, 0), 1\Big) = \Big\{\mathbf{v} \in \mathbb{R}^2 \Big| \, \|\mathbf{v}\| < 1\Big\}$.

Example 5.2.1. Another interesting example is the unit ball described by the norm: $\|(x_1, x_2)\| := \sqrt{x_1^2 + 4x_2^2}$ shown in Figure 5.2. We omit the proof that the norm is well defined. In order to find its unit ball, we must consider the equation: $\sqrt{x_1^2 + 4x_2^2} = 1 \Leftrightarrow x_1^2 + (2x_2)^2 = 1^2$. Clearly, this represents a horizontal ellipse with a major axis of 1 on the x_1-axis and semi axis of $\frac{1}{2}$ on the x_2-axis. By following the same idea as in the previous example, notice that a smaller choice of radius in the ball centered at the origin provides us with a smaller horizontal ellipse. E.g., for $r := \frac{1}{4}$, the major- and semi axis are shortened to $\frac{1}{2}$ and $\frac{1}{4}$ respectively. By allowing $r \in (0,1)$ to achieve all these values except $r = 1$, we are able to mark all the points **inside** the vertical ellipse. Thus the unit ball of this norm consists of all the points inside such a vertical ellipse.

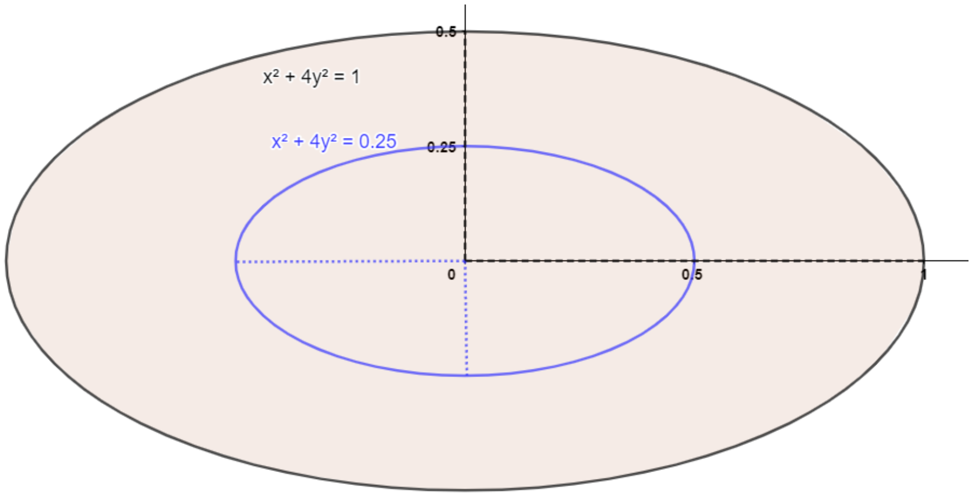

Figure 5.2: Another representation of the unit ball in a norm.

By changing the value of p for the L^p-space of functions, the unit ball can be represented by quite a lot of different figures. Some examples are an astroid, a rhombus and a square with rounded corners, but there are many other possibilities.

5.3 Linear operators on Banach spaces

It is often difficult to define a linear operator directly on an infinite dimensional vector space due to convergence issues (Christensen, 2010, p. 54). We will investigate this further in Section 9.1. For now, we only verify that the properties of such pre-defined linear operators hold. We give the following example based on an exercise from Christensen [p. 59].

Example 5.3.1. Let $\mathbf{x} = (x_1, x_2, ..., x_n, ...)$, then we define
the *right-shift* operator $T : l^1(\mathbb{N}) \to l^1(\mathbb{N}), \quad T(\mathbf{x}) = (0, x_1, x_2, ..., x_n, ...)$ and
the *left-shift* operator $S : l^1(\mathbb{N}) \to l^1(\mathbb{N}), \quad S(\mathbf{x}) = (x_2, x_3, ..., x_n, ...)$

1. What is the norm of these operators? Well, recall that in $l^1(\mathbb{N})$ the norm is given as: $\|\mathbf{x}\|_1 = \sum_{k=1}^{\infty} |x_k|$, for $\mathbf{x} \in l^1(\mathbb{N})$. Before we look at the ratio, let us find the norms $\|T\mathbf{x}\|_1$ & $\|S\mathbf{x}\|_1$:

$$\begin{aligned} \|T\mathbf{x}\|_1 &= |0| + |x_1| + |x_2| + |x_3| + \cdots = \|\mathbf{x}\|_1 \\ \|S\mathbf{x}\|_1 &= |x_2| + |x_3| + |x_4| + \cdots \leq \|\mathbf{x}\|_1 \end{aligned} \tag{5.3}$$

   Then it follows that

$$\|T\| = \frac{\|T\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \frac{\|\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = 1. \qquad \|S\| = \frac{\|\mathbf{Sx}\|_1}{\|\mathbf{x}\|_1} \leq \frac{\|\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = 1.$$

   Since $\|T\mathbf{x}\|_1 = \|\mathbf{x}\|_1$, for all $\mathbf{x} \in l^1(\mathbb{N})$, then T is an *isometry.* However, we claim that this is not the case for S. To prove this we need only find one vector $\mathbf{x}_*$ such that $\|S\mathbf{x}_*\| \neq \|\mathbf{x}_*\|$. Particularly, let $\mathbf{x}_* = (1, 0, 0, ...) \in l^1(\mathbb{N})$ whose norm may be verified as $\|\mathbf{x}_*\|_1 = 1$. When the left-shift operator acts on $\mathbf{x}_*$, we get that $S\mathbf{x}_* = (0, 0, 0, ...)$ which clearly has norm $\|S\mathbf{x}_*\|_1 = 0$. So S is <u>not</u> an isometry.

2. Another question we could investigate is whether $ST = TS = I$. First, let us think of $ST(\mathbf{x})$ similarly to how we do function compositions, i.e., $\big(g \odot f\big)(x) = g\big(f(x)\big)$. Then $ST(\mathbf{x}) = S\big(T(x_1, x_2, ...)\big) = S(0, x_1, x_2, ...) = (x_1, x_2, ...,) = \mathbf{x}$. So $ST = I$. On the contrary, the reversed operation: $TS(\mathbf{x}) = T\big(S(x_1, x_2, ...)\big) = T(x_2, x_3, ...) = (0, x_2, x_3, ...) \neq \mathbf{x}$, and we see that both operators are <u>not</u> invertible.

In this example we see that when the operator T acts on elements of l^1, it actually preserves the norm of l^1. Although S restrains the norm, i.e., keeps it bounded with maximum as $\|\cdot\|_{l^1}$, we were able to construct cases where the norm is smaller compared to elements of l^1. We also discovered that it is necessary to check both the orderings of ST and TS to ensure invertibility of the operators.

Chapter 6

Hilbert Spaces

In this chapter we will discuss a more specialized type of Banach spaces, namely Hilbert spaces. These spaces are often credited to the work performed by the German mathematician David Hilbert in 1902-1912 when he studied integral equations. Hilbert spaces have directly impacted many fields such as integral equations, Fourier analysis, quantum mechanics and many more (Carlson, 2006). The name *Hilbert space* was first coined by John von Neumann in 1929 as a tribute, and was further adopted by others. We will show that for such Banach spaces, the norm $\|\cdot\|$ can be represented by an inner product $\langle\cdot,\cdot\rangle$ which imposes additional structures to the space.

6.1 Inner product spaces and additional structures

An inner product is a *function* $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$ which $\forall\, \alpha, \beta \in \mathbb{C}$ and $\mathbf{v}, \mathbf{u}, \mathbf{w} \in V$ satisfies that:

(i) $\langle \alpha\mathbf{v} + \beta\mathbf{w}, \mathbf{u}\rangle = \alpha\langle \mathbf{v}, \mathbf{u}\rangle + \beta\langle \mathbf{w}, \mathbf{u}\rangle$ (Linearity in its first argument)

(ii) $\langle \mathbf{v}, \mathbf{w}\rangle = \overline{\langle \mathbf{w}, \mathbf{v}\rangle}$

(iii) $\langle \mathbf{v}, \mathbf{v}\rangle \geq \mathbf{0}$ and $\langle \mathbf{v}, \mathbf{v}\rangle = \mathbf{0} \Leftrightarrow \mathbf{v} = \mathbf{0}$

If we equip a vector space V with such an operation, then we call it an inner product space. In fact, some of the norms of the vector spaces we have considered earlier originate from inner products. For such inner product spaces, the norm is defined by

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle},\ \forall\, \mathbf{v} \in V \tag{6.1}$$

Notice that this norm is equivalent to $\|\mathbf{v}\|_{l^2} = \left(\sum_{k=1}^{\infty} |v_k|^2 \right)^{1/2}$ which is the l^2 norm for $p = 2$ which we saw in Equation (4.4). Consequently, Equation (6.1) is generally called the l^2-norm, but may also be identified as the Euclidean norm (Weisstein, 2003).

A vector space whose norm is induced by an inner product achieves multiple additional properties, we will consider the following:

(i) The parallelogram law

$$\|\mathbf{v} + \mathbf{w}\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 = 2\big(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2\big) \tag{6.2}$$

holds for all $\mathbf{v}, \mathbf{w} \in V$.

(ii) The inner product may be related to the norm for any $\mathbf{v} \in V$ by:

$$\|\mathbf{v}\| = \sup \Big\{ \langle |\mathbf{v}, \mathbf{w}| \rangle \Big| \, \mathbf{w} \in V,\ \|\mathbf{w}\| = 1 \Big\}.$$

Beware that if the first property does not hold, then this implies that the norm $\|\cdot\|$ of the vector space is not induced by an inner product. Then we can use this property to check whether or not a norm comes from an inner product. The second property allows us to find the norm of any $\mathbf{v} \in H$ through the inner product between $\mathbf{v}$ and the unit ball surrounding it in H. With this in mind, we can now formally define what a Hilbert space is.

Definition 6.1.1 (Hilbert space)**.** A vector space V with an inner product $\langle \cdot, \cdot \rangle$, which is a Banach space w.r.t. the norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ is called a Hilbert space. Such spaces will be denoted by H.

It follows that an inner product space V is a Hilbert space if every Cauchy sequence $\{v_k\}_1^{\infty}$ converges to an element $\mathbf{v} \in V$ with respect to the norm $\|\mathbf{v}\|$. In short, whenever an inner product space is *complete* it defines a Hilbert space.

Based on this, we can easily check that $\mathbb{R}^n, \mathbb{C}^n, l^2(\mathbb{N})$ are all Hilbert spaces with respect to

$\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{v} \rangle = \sum_{k=1}^{\infty} u_k \overline{v_k}$. However, some of the previously discussed vector spaces do not classify as Hilbert spaces, e.g., l^p for $p \geq 1$, excluding $p = 2$. We present the following example based on an exercise from Christensen's book [p. 85].

Example 6.1.1. We claim that the vector space $C[a, b]$ equipped with the norm $\|\cdot\|_\infty$ does not come from an inner product. Let us give a proof by contradiction. For simplicity, we fix $[a, b] := [0, 2]$ and assume that $\|\cdot\|_\infty$ on $C[0, 2]$ does come from an inner product. Now let

$$f(x) = \begin{cases} x, & \text{if} \quad x \in [0,1], \\ 2 - x, & \text{if} \quad x \in [1,2], \end{cases} \qquad g(x) = \begin{cases} 1 - x, & \text{if} \quad x \in [0,1], \\ x - 1, & \text{if} \quad x \in [1,2]. \end{cases}$$

Where one can verify that:

$$\Big(f + g\Big)(x) = \begin{cases} 1, & \text{if} \quad x \in [0,1], \\ 1, & \text{if} \quad x \in [1,2], \end{cases} \qquad \Big(f - g\Big)(x) = \begin{cases} -1 + 2x, & \text{if} \quad x \in [0,1], \\ 3 - 2x, & \text{if} \quad x \in [1,2]. \end{cases}$$

Let us now use the definition of the $\|\cdot\|_\infty$ norm, which "grabs" the supremum value of each element. Then according to the parallelogram law, the subsequent relation must hold:

$$\begin{aligned} \|f + g\|_\infty^2 + \|f - g\|_\infty^2 &= 2\Big(\|f\|_\infty^2 + \|g\|_\infty^2\Big) \\ 1^2 + 1^2 &= 2\Big(1^2 + 1^2\Big) \\ 2 &= 4. \end{aligned}$$

But clearly this is a contradiction, therefore we conclude that $\|\cdot\|_\infty$ does not come from an inner product. Another interesting and convenient result of an inner product space, is that we can easily withdraw information about the angles between its elements, particularly orthogonality.

Example 6.1.2. Let $\mathbf{v} = (0, 1)$ and $\mathbf{w} = (1, 0)$, then $\langle \mathbf{v}, \mathbf{w} \rangle = (0 \cdot 1, 1 \cdot 0) = \mathbf{0}$, so $\mathbf{v} \perp \mathbf{w}$.

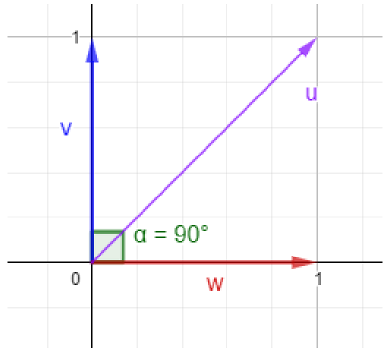

Figure 6.1: A representation of two orthogonal vectors $\mathbf{v} \perp \mathbf{w}$ and their vector sum $\mathbf{u}$.

It is often convenient to collect all vectors $\mathbf{v_1}, \mathbf{v_2}, ... \in H$ which are orthogonal to each other in what we call an *orthogonal system* $\{\mathbf{v}_k\}_{k=1}^{\infty}$, where $\langle \mathbf{v}_k, \mathbf{v}_l \rangle = 0$ for all $k \neq l$. Furthermore, this orthogonal system can be separated into two subspaces U and W of H such that the orthogonal elements are separated from each other. Then these subspaces complement one another in such a way that we may reconstruct the Hilbert space through a *direct sum*.

Definition 6.1.2 (Direct sum)**.** Let $\mathbf{v} \in H$, $\mathbf{u} \in U$ and $\mathbf{w} \in W$. The sum between two subspaces U and W of H is called direct if and only if $\mathbf{v} \in U + W$ is uniquely represented by $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Denoted by $U \bigoplus W$.

Example 6.1.3 (Direct sum)**.** Consider the space $\mathbb{R}^n$. Let

$$W = \text{span}\{w_k\}_{k=1}^{p}, \quad U = \text{span}\{u_j\}_{j=1}^{q}, \qquad \text{where} \quad p, q \in \mathbb{N},\ w_k, u_k \in \mathbb{R}^n.$$

Proposition: $W \bigoplus U$ is direct iff the vectors $w_1, ..., w_p$ and $u_1, ..., u_q$ are linearly independent.

Proof. Assume that $\exists w, w' \in W, u, u' \in U$ s.t. $w + u = v$ and $w' + u' = v$

$$w = c_1 w_1 + \cdots + c_p w_p, \quad w' = c_1' w_1 + \cdots + c_p' w_p$$
$$u = d_1 u_1 + \cdots + d_q u_q, \quad u' = d_1' u_1 + \cdots + d_q' + u_q,$$

then,

$$w + u = w' + u'$$
$$c_1 w_1 + \cdots + c_p w_p + d_1 u_1 + \cdots + d_q u_q = c_1' w_1 + \cdots + c_p' w_p + d_1' u_1 + \cdots + d_q' + u_q$$
$$(c_1 - c_1') w_1 + \cdots + (c_p - c_p') w_p + (d_1 - d_1') u_1 + \cdots + (d_q - d_q') u_q = 0.$$

Now, if we impose that $W \bigoplus U$ is a direct sum i.e., that $\mathbf{v}$ has a *unique* representation, then clearly we need that $c_1 = c_1', ..., c_p = c_p'$ and $d_1 = d_1', ..., d_q = d_q'$ in order to have a non-trivial solution to the above equation. This means that the vectors $w_1, ..., w_p$ and $u_1, ..., u_q$ must be linearly independent, which completes the proof. □

6.2 Functionals and linear operators

A linear operator which sends elements of H into the set of complex numbers $\mathbb{C}$ is called a *functional*. More notably, a functional $\Phi : H \to \mathbb{C}$ may be thought of as a special case of a linear operator, which implies that we can use Definition 4.2.1 to find its norm. Particularly, observe that after Φ has acted on $\mathbf{v} \in H$, then $\Phi\mathbf{v}$ is just a complex number. Thus, we can replace $\|\Phi\mathbf{v}\|$ with $|\Phi\mathbf{v}|$ in the expression of the norm.

$$\|\Phi\mathbf{v}\| = \sup_{\mathbf{v}\neq 0} \frac{|\Phi\mathbf{v}|}{\|\mathbf{v}\|}$$

Interestingly, the set of all continuous linear functionals over H is itself a vector space, since we may add and multiply linear functionals by scalars. Equipped with the norm $\|\cdot\| = \langle\cdot,\cdot\rangle$ we call this the *dual space* of H, commonly denoted by H^*. Moreover, the map defined by $\langle\cdot,\cdot\rangle$ on an inner product space V is said to represent a linear functional on V, see Equation (7.2) in the following chapter. This result leads us towards the famous Riesz Representation theorem.

Theorem 6.2.1 (Riesz' Representation Theorem)**.** *Given any bounded linear functional* $\Phi : H \to \mathbb{C}$ *then there exists a unique* $\mathbf{w} \in H$ *such that* $\Phi\mathbf{v} := \langle\mathbf{v}, \mathbf{w}\rangle, \forall\, \mathbf{v} \in H$.

Consequently a different choice of $\mathbf{w}$ gives root to a different functional. We may also consider linear operators T between two Hilbert spaces H_1, H_2. Note that these are not functionals on their own, since the 'destination' is another Hilbert space. Assume that $T : H_1 \to H_2$ is a bounded linear operator. Then fix any $\mathbf{w} \in H_2$ and let $\Phi : H_1 \to \mathbb{C}, \quad \Phi\mathbf{v} = \langle T\mathbf{v}, \mathbf{w}\rangle$. Make special note that here $\mathbf{v} \in H_1$, whilst both arguments in the inner product $T\mathbf{v}$ and $\mathbf{w} \in H_2$. Now let us give a definition of an *adjoint* to an operator T, often denoted T^*.

Definition 6.2.1. Given a linear operator T which sends elements of a Hilbert space $H_1 \to H_2$, the corresponding adjoint T^* sends elements from $H_2 \to H_1$. Then

$$\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, T^*\mathbf{y}\rangle. \tag{6.3}$$

One may easily show that such an operator is linear and bounded with $(T^*)^* = T$. The norm may also be verified to be equal between T and T^* such that $\|T\| = \|T^*\|$.

Example 6.2.1. Let us investigate the operator $T : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ given by $T(x_1, x_2, ...) := (3x_2, 2x_1, x_3, x_4, ...)$. If $\mathbf{x}, \mathbf{y} \in l^2(\mathbb{N})$ and $\alpha, \beta \in \mathbb{C}$, then we get that

$$\begin{aligned} T(\alpha\mathbf{x} + \beta\mathbf{y}) &= \big(3(\alpha x_2 + \beta y_2), 2(\alpha x_1 + \beta y_1), \alpha x_3 + \beta y_3, ...\big) = \\ &= \big(3\alpha x_2, 2\alpha x_1, \alpha x_3, ...\big) + \big(3\beta y_2, 2\beta y_1, \beta y_3, ...\big) = \alpha T\mathbf{x} + \beta T\mathbf{y} \qquad \checkmark \end{aligned}$$

So T is linear. Next we check if there exists a number K which bounds the operator. Therefore we start by finding the norm of $T\mathbf{x}$ and $\mathbf{x}$. Clearly,

$$\begin{aligned} \|\mathbf{x}\|_{l^2} &= \left(\sum_{k=1}^{\infty} |x_k|^2\right)^{1/2} = \sqrt{|x_1|^2 + |x_2|^2 + |x_3|^2 + |x_4|^2 + \cdots}, \\ \|T\mathbf{x}\|_{l^2} &= \sqrt{|3x_2|^2 + |2x_1|^2 + |x_3|^2 + |x_4|^2 + \cdots} \\ &\leq \sqrt{|3x_1|^2 + |3x_2|^2 + |3x_3|^2 + |3x_4|^2 + \cdots} = 3 \cdot \|\mathbf{x}\|_{l^2}. \end{aligned}$$

We find the norm by using Definition 4.2.1, which simplifies into

$$\|T\| = \frac{\|T\mathbf{x}\|_{l^2}}{\|\mathbf{x}\|_{l^2}} \leq 3,$$

so T is bounded. Moreover, if we choose the vector $\tilde{\mathbf{x}} := (0, 1, 0, 0, ...)$, then

$$\|T\tilde{\mathbf{x}}\|_{l^2} = \sqrt{0 + |3|^2 + 0 + 0 + \cdots} = 3 \quad \Longrightarrow \quad \|T\| = 3. \qquad \checkmark$$

From Definition 6.2.1 we know that $\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, T^*\mathbf{y}\rangle$. We wish to figure out how the adjoint operator T^* acts on a vector $\mathbf{y} \in l^2(\mathbb{N})$, for this purpose let us denote

$$T^*\mathbf{y} := (w_1, w_2, w_3, w_4, ...). \tag{6.4}$$

Consider first the left-hand side, **LHS**, of the equation:

$$\langle T\mathbf{x}, \mathbf{y}\rangle = \big\langle (3x_2, 2x_1, x_3, x_4, ...), (y_1, y_2, y_3, y_4, ...)\big\rangle = 3x_2\bar{y_1} + 2x_1\bar{y_1} + x_3\bar{y_3} + x_4\bar{y_4} + \cdots$$

Then by Equation (6.4) the right-hand side, **RHS**, gives:

$$\langle \mathbf{x}, T^*\mathbf{y} \rangle = \big\langle (x_1, x_2, x_3, x_4, ...), (w_1, w_2, w_3, w_4, ...) \big\rangle = x_1\bar{w}_1 + x_2\bar{w}_2 + x_3\bar{w}_3 + x_4\bar{w}_4 + \cdots$$

Finally, by comparing LHS and RHS notice that we need $w_1 = 2y_2$, $w_2 = 3y_1$, $w_3 = y_3$, $w_4 = y_4, ...$ therefore $T^*\mathbf{y} = (2y_2, 3y_1, y_3, y_4, ...)$ ✓, and we see that the adjoint operator T^* acts rather similarly on a vector $\mathbf{y}$ as T acts on $\mathbf{x}$.

Corollary: In cases where the operators prove to be equal, i.e. $T = T^*$, we say that T is *self-adjoint*.

6.3 Short summary of important aspects

Up until now we have been looking at normed vector spaces with specific properties. Let us now present a short overview along with a figure. A Hilbert space H is always a Banach space, but the opposite need not be true. Similarly, a Banach space is always normed, but a normed vector space is not necessarily Banach. At the same time, a Hilbert space is an inner product space, where an inner product space can have Cauchy sequences that do not converge to an element of the space. In this case it cannot be classified as a Hilbert space as it does not satisfy the Banach properties. Consequently, the Hilbert spaces are found in the intersection between the Banach- and Inner product spaces, see Figure [6.2] below. As such, the Hilbert spaces are clearly the most complex spaces induced with powerful analytical structures. In the following chapter we will introduce a specific Hilbert space which provides a simpler method for checking convergence of Fourier series.

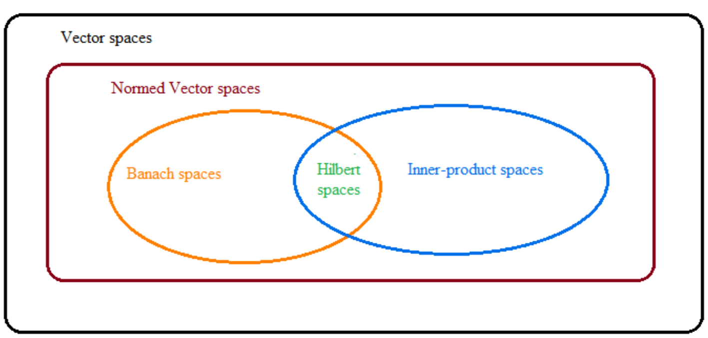

Figure 6.2: A representation of vector spaces.

Chapter 7

The Function spaces

In this chapter we will present our most favourable vector spaces, i.e., the $L^p(\mathbb{R})$-spaces. For a completely rigorous treatment which works for all functions within these spaces, it would be necessary to give an extensive account of measure. To this extent, we will briefly discuss *when* and *how* to apply the relevant measure theory. Afterwards, we investigate how to apply an inner product to a norm.

7.1 L^p spaces

A function $f : \mathbb{R} \to \mathbb{C}$ is said to be absolutely integrable if

$$\int_{-\infty}^{\infty} |f(x)|\, dx < \infty, \tag{7.1}$$

all functions which satisfy Equation (7.1) are found in the $L^1(\mathbb{R})$-space. However, in order for this integral to make sense, we must require the functions to be measurable. Otherwise, we may run into integrability issues whenever the function f fails to be piecewise continuous.

Definition 7.1.1 (Zero measure)**.** A set S is of zero measure, denoted by $m(S) = 0$ if $\forall \varepsilon > 0$, there exist open intervals $I_k = (a_k, b_k)$ such that $S \subset \bigcup_{k=1}^{N}(a_k, b_k)$ and $\sum_{k=1}^{N}(b_k - a_k) < \varepsilon$, where N is either infinite or finite.

Examples of sets with zero measure are the countably infinite sets, e.g., $\{\mathbb{N}, \mathbb{Z}, \mathbb{Q}\}$ and all finite sets. Informally we may think of a *measurable* function as a structure-preserving function between two sets of measure. To provide a more formal definition we require significantly more measure theory which is outside the scope of this thesis. A well known measurable function is the Dirichlet function defined by

$$f(t) = \begin{cases} 1, & \text{if } t \in \mathbb{Q}, \\ 0, & \text{if } t \in \mathbb{R}/\mathbb{Q}. \end{cases}$$

which is discontinuous for all t. Such a function is not Riemann integrable, but has been shown to be Lebesgue integrable. We will briefly discuss what this entails in Section 8.4 along with an example. Generally, we denote the set of all continuous functions f where $|f|^p$ is integrable by:

$$L^p(\mathbb{R}) := \left\{ f : \mathbb{R} \to \mathbb{C} \,\middle|\, \int_{-\infty}^{\infty} |f(x)|^p \, dx < \infty \right\}, \quad 1 \leq p < \infty. \tag{7.2}$$

The "L" stands for Lebesgue. One may easily verify that $\forall p \in [1, \infty)$ these sets make up vector spaces. Furthermore it can be proven that $\forall p$ as described above, the corresponding space satisfies the Banach conditions w.r.t. the norm

$$||f||_{L^p} := \left(\int_{-\infty}^{\infty} |f(x)|^p \, dx \right)^{1/p}, \quad f \in L^p(\mathbb{R})$$

and the proof follows similar procedures as we have shown in earlier examples.

Proposition: When $p \in (0, 1)$ we still have a vector space, but $L^p(\mathbb{R})$ fails to satisfy the triangle inequality.

Proof. Take for example $f(x) = \begin{cases} x, & \text{if } 0 < x < 1 \\ 0, & \text{otherwise.} \end{cases}$ $g(x) = \begin{cases} 1 - x, & \text{if } 0 < x < 1 \\ 0, & \text{otherwise.} \end{cases}$ and fix $p := \frac{1}{2}$. If we assume the norm for $p = \frac{1}{2}$ is well defined, then the norms can be found by:

$$\|f + g\|_{\frac{1}{2}} = \left(\int_{-\infty}^{\infty} |f + g|^{\frac{1}{2}} \right)^{\frac{1}{\frac{1}{2}}} = \left(\int_0^1 \sqrt{|1|} \right)^2 = 1.$$

$$\|f\|_{\frac{1}{2}} = \left(\int_0^1 \sqrt{|x|} \right)^2 = \left(\frac{2}{3} x^{\frac{3}{2}} \Big|_0^1 \right)^2 = \frac{4}{9}.$$

$$\|g\|_{\frac{1}{2}} = \left(\int_0^1 \sqrt{|1 - x|} \right)^2 = \left(\frac{2}{3} (1 - x)^{\frac{3}{2}} \Big|_0^1 \right)^2 = \frac{4}{9}.$$

However, notice that $\|f + g\|_{\frac{1}{2}} = 1 > \frac{4}{9} + \frac{4}{9} = \|f\|_{\frac{1}{2}} + \|g\|_{\frac{1}{2}}$ which contradicts the triangle inequality. □

Similarly we can find examples and show this for the remaining $p \in (0, 1)$. Since the triangle inequality fails, we get that $\|\cdot\|_{L^p}$ is not a norm on L^p for this range of p, ultimately revoking the Banach-title of the vector spaces. We move our focus onto the case $p = 2$ where the functions are said to be square-integrable. In some cases we will also consider $L^p(a, b)$, where we restrict f to a section of $\mathbb{R}$, i.e., $f : [a, b] \to \mathbb{C}$.

7.2 The L^2-space

Our goal is to show that the norm of $L^2(\mathbb{R})$ actually comes from an inner product. For this purpose, we need to briefly explain what an equivalence relation is. Let $f, g \in L^2(\mathbb{R})$

$$f \sim g \Leftrightarrow \int_{-\infty}^{\infty} |f(x) - g(x)|\, dx = 0$$

In such cases, it means that we identify f with g, i.e., $f \approx g$, because f and g are equal *almost everywhere*. We collect all the functions f which are equal to 0 almost everywhere by $\tilde{\mathbf{0}} := \left\{ f \in L^2(\mathbb{R}) \middle| f \sim \mathbf{0} \right\}$. This implies that the elements of $L^p(\mathbb{R})$ are (strictly speaking) equivalence classes of functions which satisfy (7.2). The reason why this is necessary will become clear in the proof of its inner product. Yet, in most cases we can usually pretend that these elements are regular functions $f : \mathbb{R} \to \mathbb{C}$.
Let us now show that the L^2-norm comes from an inner product which in turn implies that L^2 is a Hilbert space.

Theorem 7.2.1. *The vector space $L^2(\mathbb{R})$ is a Hilbert space with respect to the inner product*

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)}\, dx, \quad f, g \in L^2(\mathbb{R}). \tag{7.3}$$

Proof. A key element in this proof will be the usage of the Cauchy-Schwartz inequality, which follows from the Hölders inequality for $p = q = 2$. This allows us the following representation for all $f, g \in L^2(\mathbb{R})$:

$$\begin{aligned}
|\langle f, g\rangle| &= \left| \int_{-\infty}^{\infty} f(x)\overline{g(x)}\,dx \right| \\
&\leq \int_{-\infty}^{\infty} |f(x)\overline{g(x)}|\,dx \\
&\leq \sqrt{\int_{-\infty}^{\infty} |f(x)|^2\,dx} \cdot \sqrt{\int_{-\infty}^{\infty} |g(x)|^2\,dx} \\
&= \|f\|_{L^2} \cdot \|g\|_{L^2} < \infty, \quad \text{whenever } f, g \in L^2(\mathbb{R})
\end{aligned}$$

Hence the expression is *well-defined* (it exists). It remains to check that the function from (7.3) actually represents an inner product. Firstly, for any given $f_1, f_2, g \in L^2(\mathbb{R})$ we show that:

$$\begin{aligned}
\langle \alpha f_1 + \beta f_2, g\rangle &= \int_{-\infty}^{\infty} \Big(\alpha f_1(x) + \beta f_2(x)\Big)\overline{g(x)}\,dx \\
&= \alpha \int_{-\infty}^{\infty} f_1(x)\overline{g(x)}\,dx + \beta \int_{-\infty}^{\infty} f_2(x)\overline{g(x)}\,dx \\
&= \alpha\langle f_1, g\rangle + \beta\langle f_2, g\rangle \qquad \text{(Linearity)} \quad \checkmark
\end{aligned}$$

Secondly, we can show that if we interchange f and g in the inner product, we get:

$$\begin{aligned}
\langle g, f\rangle &= \int_{-\infty}^{\infty} g(x)\overline{f(x)}\,dx \\
&= \overline{\int_{-\infty}^{\infty} f(x)\overline{g(x)}\,dx} = \overline{\langle f, g\rangle} \qquad \checkmark
\end{aligned}$$

Lastly,

$$\begin{aligned}
\langle f, f\rangle &= \int_{-\infty}^{\infty} f(x)\overline{f(x)}\,dx \\
&= \int_{-\infty}^{\infty} |f(x)|^2\,dx \geq 0
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
\langle f, f\rangle &= \mathbf{0} \\
&\Updownarrow \\
f &= \tilde{\mathbf{0}}
\end{aligned}$$

This completes the proof. □

It should now be clear why the elements of $L^2(\mathbb{R})$ must be equivalence classes. Namely that if $f_1(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise,} \end{cases}$ then we have that $f_1 \neq 0$, yet $\langle f_1, f_1\rangle = \mathbf{0}$.

Example 7.2.1. In the previous chapter we mentioned that it is possible to equip $C[a, b]$ with an inner product. We will now show that this is true. For convenience we again fix $[a, b] := [0, 2]$.

Claim: The inner product from Equation (7.3) defines an inner product on $C[0, 2]$, and the proof follows from above, with integral limits from $0 \to 2$. Its associated norm is given by

$$\|f\| = \sqrt{\int_0^2 |f(x)|^2 dx}, \quad f \in C[0, 2].$$

Consider the functions

$$f_n(x) := \min(x^n, 1) \text{ for } n \in \mathbb{N}, \quad \text{and} \quad f(x) := \begin{cases} 0 \text{ if } x \in [0, 1), \\ 1 \text{ if } x \in [1, 2]. \end{cases}$$

Clearly $f_n(x) \in C[0, 2]$, whereas $f(x) \notin C[0, 2]$. Meanwhile, their normed difference shows that

$$\|f_n - f\| = \sqrt{\int_0^1 (x^n - 0)dx + \int_1^2 (1 - 1)dx} = \sqrt{\int_0^1 x^n dx} \to 0 \quad \text{as } n \to \infty.$$

As result, we see that the Cauchy sequence $f_n(x)$ converges towards $f(x)$ which is an element <u>outside</u> $C[0, 2]$. This means that $C[a, b]$ achieves the additional property that we can approximate discontinuous functions through the norm. So even though we were able to apply an inner product to $C[a, b]$, we simultaneously removed the Banach space property of completeness. Consequently, we conclude that $C[a, b]$ is not a Hilbert space w.r.t. the norm given by $\|f\|$.

Chapter 8

Fourier series and convergence

We have now covered the necessary theory in order to properly work with convergence of Fourier series. In this chapter we give two definitions of Fourier series. The structure provided by the $L^p(\mathbb{R})$ spaces allows a more general study on convergence. Although there exists many types of convergence, we will focus on three. Namely, *norm, piecewise* and *uniform* convergence. We end this chapter by briefly discussing an alternative method which is often used to study convergence, the Dirichlet Kernel.

8.1 Fourier series

As mentioned earlier, Fourier series naturally show up in the topic of heat-equations. Other practical areas are in found in electrical engineering, analog signal processing of periodic signals, vibration analysis and much more. Consequently, Fourier series are often used for problems involving the concept of time, and it is therefore natural to replace the unknown variable x by t. When discussing Fourier series, the most common construction is through the trigonometric system, which was the first representation used (Fourier series, 2012). Furthermore, we are interested in how the properties of a function f relates to the properties of its corresponding Fourier series. We often restrict the functions to a 2π-periodic interval, e.g., $f \in L^2(-\pi, \pi)$ or $f \in L^2(0, 2\pi)$ where

$$f(t + 2\pi) = f(t) \quad t \in \mathbb{R}.$$

One reason for considering $L^2(a,b)$, $a,b \in \mathbb{R}$ is that physical models and experiments often are restricted to finite intervals. Examples of periodic systems are 'vibrating strings' and 'musical tones'.

According to Christensen (p. 118) we have that $f \in L^2(-\pi,\pi) \Rightarrow f \in L^1(-\pi,\pi)$, which means that if f is square-integrable on the interval $(-\pi,\pi)$ then it is also an integrable function. This statement fails for $f \in L^2(\mathbb{R})$, as shown in Example 8.2.1. Given that $f \in L^2(-\pi,\pi)$ then its corresponding Fourier series can be found by expanding the function with respect to the trigonometric functions $\cos nt, \sin nt$, for $n \in \mathbb{N}$. Hence the Fourier series of f are represented by

$$f \sim \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_n \cos nt + b_n \sin nt), \tag{8.1}$$

where one can verify that the Fourier coefficients are given by

$$a_n := \frac{1}{\pi}\langle f, \cos nt\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt, \quad n = 0,1,2,...$$
$$b_n := \frac{1}{\pi}\langle f, \sin nt\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt, \quad n = 1,2,3,...$$

This is similar to how Joseph Fourier himself represented the expansions of functions in 1807 (Fourier series, 2012). However, in modern mathematics we also often study complex numbers. By using the well know relation of the complex exponential known as Euler's formula: $\boxed{e^{\pm i\theta} = \cos\theta \pm i\sin\theta}$ we can represent $\cos nt = \frac{e^{int}+e^{-int}}{2}$, and $\sin nt = \frac{e^{int}-e^{-int}}{2i}$. Then by inserting this into Equation (8.1), one can check that if the Fourier coefficients are given by

$$c_k := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}dt, \qquad \text{for } f \in L^2(-\pi,\pi)$$

then the complex (or exponential) form of the Fourier series are:

$$f \sim \sum_{n=-\infty}^{\infty} c_n e^{int}. \tag{8.2}$$

It is often useful to investigate the N-th partial sum of the Fourier series, defined as

$$S_N(t) := \frac{a_0}{2} + \sum_{k=1}^{N}(a_k \cos kt + b_k \sin kt) = \sum_{k=-N}^{N} c_k e^{ikt}. \tag{8.3}$$

Observe that both in Equation (8.1) and (8.2) we use the '$\sim$' symbol rather than the equality sign. This is due to an unclear convergence towards the function f. The "rate" that the Fourier series converges towards the function f is also unknown. Generally, we say that a Fourier series converges slowly if we need N to become sufficiently large, i.e., $N \to \infty$. Before we clarify the convergence of these series, we recall the concept of a basis.

Proposition. The functions $\{e_k\}_1^\infty := \left\{\frac{1}{\sqrt{2\pi}}e^{ikt}\right\}_{k=-\infty}^{\infty}$ form an orthonormal basis for $L^2(-\pi,\pi)$. This means that $\{e_k\}_1^\infty$ spans $L^2(-\pi,\pi)$ where all the $\{e_k\}_1^\infty$ functions are linearly independent. Since the basis is orthonormal, the inner product between the elements is given by $\langle e_i, e_j\rangle = \delta_{ij}$. Where δ_{ij} is the Kronecker delta function, i.e.,

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

It follows that any element $f \in L^2(-\pi,\pi)$ can be represented by $f = \sum_{k=1}^{\infty} \langle f, e_k\rangle e_k$ (Christensen, 2010, p. 2:3). In fact, it is through this basis that the function f has been expanded in Equation (8.2). To show this, we take a function $g \in L^2(\mathbb{R})$ which may be represented as

$$g(t) = \sum_{k=-\infty}^{\infty} \langle g, \frac{1}{\sqrt{2\pi}}e^{ikt}\rangle \frac{1}{\sqrt{2\pi}}e^{ikt}$$

Then we use the norm defined by the inner product on $L^2(\mathbb{R})$ and recall how the exponential coefficients c_k of the Fourier series are given. Lastly, we sum together all the possible representations and arrive at the desired form.

$$\langle g, \frac{1}{\sqrt{2\pi}}e^{ikt}\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(t)e^{-ikt}dt = \sqrt{2\pi}\cdot c_k,$$

$$\text{which implies that } g(t) = \sum_{k=-\infty}^{\infty} (\sqrt{2\pi}\cdot c_k)\frac{1}{\sqrt{2\pi}}e^{ikt} = \sum_{-\infty}^{\infty} c_k e^{ikt}$$

It follows that by expanding a function g with respect to an orthonormal basis, we arrive at its Fourier series. There are other orthonormal bases which may be used to represent Fourier series. However, a change in basis often requires a change of its period through a scaling of the function. E.g., the basis $\{e^{2\pi ikt}\}_{k=-\infty}^{\infty}$ holds on the interval $(-\frac{1}{2},\frac{1}{2})$. As the proof is quite technical, we omit it. Let us instead explain how to change the interval. Given a constant

$L \in \mathbb{R}$, we say that a function is L-periodic if $f(t+L) = f(t)$. So for a function with general period L, its Fourier series can be expressed through

$$f(t) \sim \sum_{-\infty}^{\infty} c_n e^{\frac{2\pi}{L} int},$$

where $\{e^{\frac{2\pi}{L} int}\}_{k=-\infty}^{\infty}$ is an *orthogonal* basis. In the following sections we are interested in expressing when and how Fourier series converge, i.e., that the partial sums $S_N(x)$ converge to $f(x)$. To this extent, we will investigate when the expression $'\lim_{N\to\infty} S_N = f\,'$ is true by considering three types of convergence.

8.2 L^2-convergence

We will begin by discussing the simplest manner in which Fourier series may converge, namely convergence in a norm. For this purpose, we will focus on the norm defined in the L^2-space for the following reasons. Firstly, recall that $\forall\, x = (x_1, x_2, ...) \in \mathbb{R}^n$ we can represent it by $x = \langle x_1, e_1\rangle e_1 + \langle x_2, e_2\rangle e_2 + \cdots$. In comparison, the L^2-space is a generalization of $\mathbb{R}^n$ where the elements are equivalence classes of functions instead of vectors and the coordinates have been extended to infinitely many coordinates. Therefore, one may refer to Hilbert spaces as a natural extension of the Euclidean space (Weisstein, 2005). It follows that many ideal structures and identities are valid in the L^2-space compared to other normed spaces. Due to all the added structure we have been applying through norms, inner products etc., we can see that the functions in our spaces are better equipped to check for convergence. Finally, we say that the Fourier series of a function $f(t) \in L^2(-\pi, \pi)$ converges in the L^2-norm when

$$\Big\| f(t) - S_N(t) \Big\|_{L^2(-\pi,\pi)} \to 0, \quad \text{as } N \to \infty,$$

or more intuitively

$$\int_{-\pi}^{\pi} \left| f(t) - \sum_{k=-N}^{N} c_k e^{ikt} \right|^2 dx \to 0 \quad \text{as } N \to \infty. \tag{8.4}$$

Make note that $S_N(t)$ need not become arbitrarily close towards $f(t)$ **for all** $t \in (-\pi, \pi)$ (Christensen, 2010, p. 118:119). This dates back to 1966 when Carleson proved that given a

L^2-function, its Fourier series converges in the L^2-norm as shown in (8.4) *almost everywhere* (Fourier series, 2012). As before, this means that divergence occurs in a set of points of zero measure. Two years later, Richard Hunt extended this result to cover $f \in L^p(-\pi, \pi),\ \forall p > 1$. So far, the proof given by Hunt is regarded as the most difficult in the theory of Fourier series (Walker, 2003, p. 9). The result from Hunt implies that if we replace the number 2 in Equation (8.4) by any $p \in (1, \infty)$, then we can talk about norm-convergence in general. However, it becomes more difficult to show convergence in these cases, as they are no longer Hilbert spaces.

Additionally, there is no automatic implication of convergence between the $L^p(\mathbb{R})$ spaces. By this we mean that checking the convergence of Fourier series for a function with respect to the $L^2(\mathbb{R})$ norm gives <u>no</u> indication on its convergence in other $L^p(\mathbb{R})$ spaces. We will now consider an example to show that a function of sequences may converge in the L^1 norm, yet diverge in the L^2-norm, and vise-versa.

Example 8.2.1. Let $f_1(x) := \begin{cases} \frac{1}{\sqrt{x}}, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$ and $f_2(x) := \begin{cases} \frac{1}{x}, & 1 < x < \infty, \\ 0, & \text{otherwise.} \end{cases}$.

Then for f_1 we have that

$$\|f_1\|_{L^1(0,1)} = \int_0^1 \left|\frac{1}{\sqrt{x}}\right| dx = 2\sqrt{x}\Big|_0^1 = 2. \qquad \checkmark$$

$$\|f_1\|_{L^2(0,1)} = \int_0^1 \left|\frac{1}{\sqrt{x}}\right|^2 dx = \int_0^1 \frac{1}{x} = \ln x\Big|_0^1 = 0 - (-\infty) = \infty. \qquad ✗$$

Conversely, we see that for f_2

$$\|f_2\|_{L^1(1,\infty)} = \int_1^\infty \left|\frac{1}{x}\right| dx = \ln x\Big|_1^\infty = \infty. \qquad ✗$$

$$\|f_2\|_{L^2(1,\infty)} = \int_1^\infty \left|\frac{1}{x}\right|^2 dx = -\frac{1}{x}\Big|_1^\infty = 0 + 1 = 1. \qquad \checkmark$$

This example shows that $L^1(\mathbb{R})$ is not a subspace of $L^2(\mathbb{R})$ (the same is true for the reversed statement). Similar situations may be found when investigating convergence of Fourier series on the whole of $\mathbb{R}$.

8.3 Pointwise convergence

A 2π-periodic function $f(t)$ is called piecewise continuous on a closed interval $[a,b]$ or on $\mathbb{R}$, if it is continuous except for at a finite number of discontinuous points t_0. At these points we say that $f(t_0)$ jumps such that the right- and left limits of f exists at each point t_0, see Figure [8.1]. More particularly, if both $f(t)$ and its derivative $f'(t)$ are piecewise continuous, then $f(t)$ is called piecewise smooth. This is the case for the function in Figure [8.1], given that we restrict the interval to $[-2,3)$. Furthermore, if $f(t)$ is discontinuous at a point t_0, then its left limit $\lim_{r>0,r\to 0} f(t_0-r) =: f(t_0-0)$ and its right limit $\lim_{r>0,r\to 0} f(t_0+r) =: f(t_0+0)$ are different.

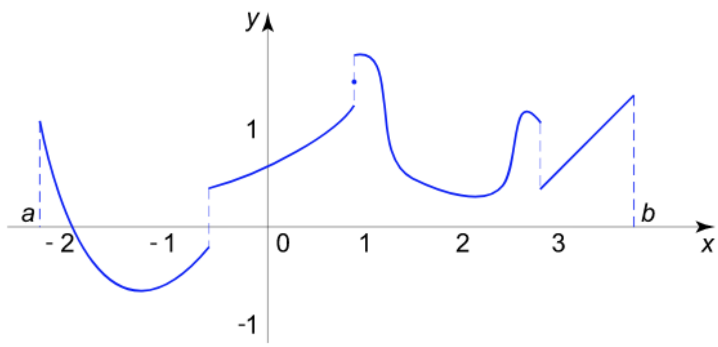

Figure 8.1: A representation of a piecewise continuous function. By Math24, 2005.

The famous mathematician Dirichlet was the first to present a type of convergence for Fourier series in 1829 (Fourier series, 2012). Based on his theorem, we give the following formulation:

Theorem 8.3.1 (Pointwise convergence)**.** *Let f be a 2π-periodic piecewise smooth function with a finite number of maximum and minimum points. Then the Fourier series S_N of a function f converges to*

$$\lim_{N\to\infty} S_N(t_o) = \begin{cases} f(t_0), & \text{if } f(t) \text{ is } \textit{continuous} \text{ at } t_0, \\ \frac{f(t_0-0)+f(t_0+0)}{2}, & \text{if } f(t) \text{ is } \textit{discontinuous} \text{ at } t_0. \end{cases}$$

The proof is quite lengthy and complicated, so we omit it. We end this section by mentioning that the above theorem may fail under certain conditions. This was originally shown by Kolmogorov which constructed an absolutely integrable function, i.e., $f \in L^1(0,2\pi)$, whose

Fourier series diverges at *every* point. Interestingly, whenever there is a discontinuity at t_0 the partial sums of the Fourier series forces *oscillations* to occur on both sides of the discontinuity. Let us represent this by the famous *Square Wave*, denoted by $S_q(t)$ in Figure [8.2]. Clearly the Square Wave function is discontinuous and periodic, where the vertical lines represent the jump-discontinuity of the function.

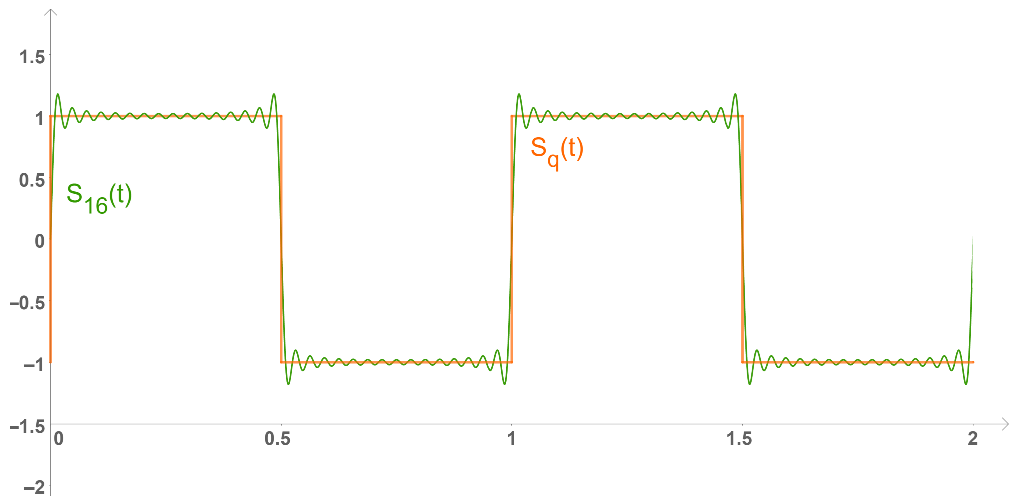

Figure 8.2: An approximation of a square wave (orange) by its first 16 partial sums of Fourier series (green).

One may verify that the partial sum of its Fourier Series is given by

$$S_N(t) = \frac{4}{\pi}\sum_{k=1}^{N}\frac{\sin(2k-1)2\pi t}{2k-1}.$$

Notice that $S_q(t)$ is an *odd* function, which due to the properties of inner products forces all the a_n coefficients in the trigonometric representation to disappear, because $S_q(t)$ is orthogonal to $\cos nt$, for $n = 0, 1, 2, ...$ Through a process called Fourier Synthesis where the N^{th} odd terms in the Fourier series are added together with the average of $S_q(t)$, the resulting Fourier series $S_N(t)$ is a continuous function (Serway & Jewett, 2013, p. 553:554). These N-odd terms represent the number of oscillations (or harmonics) which occur between the smallest periodic interval, e.g., $[0.5, 1.5]$ in $S_N(t)$. Given that the function of interest is continuous, then these oscillations are distributed evenly on the whole period L with frequencies $\frac{n}{L}$. They eventually become unnoticeable as $N \to \infty$, or sooner if the Fourier series converges *quickly* (Walker, 2003, p. 6). Since $S_q(t)$ is discontinuous, notice that we get a so-called *overshoot* value where the partial sum of the Fourier series achieves a greater value than the function $f(t)$, and an *undershoot* where it achieves a smaller value. This is called the *Gibbs phenomenon* and is a direct result from the oscillations being trapped or restricted near the jump discontinuity, coupled with a slow rate of convergence proportional to $|n|^{-1}$(Walker, 2003). Now combine the over- and undershoot value into a *spike*. Then the length, or amplitude, of the spike can be shown to be approximately 18% greater than the jump-discontinuity itself as $N \to \infty$ (Math24, 2005). At the same time, the length of the interval where the spike occurs goes towards $\mathbf{0}$, so the contribution of the spike in the integral becomes arbitrarily small.

8.4 Uniform convergence

Uniform convergence is the *most ideal* and *strongest* type of convergence, which in turn poses it to be difficult to show. In return, uniform convergence implies all other forms of convergence (Math24, 2005). The simplest case to check for uniform convergence is by using the following theorem.

Theorem 8.4.1 (Uniform convergence)**.** *The Fourier series expansion of a 2π-periodic continuous* <u>*and*</u> *piecewise smooth function converges uniformly to the function f.*

We say that a sequence of functions converges uniformly towards another function whenever the speed of convergence is independent of x in the whole domain. The same can be said for Fourier series, which means that when a Fourier series converges uniformly to a function $g(t)$, it does so in the supremum norm, i.e.,

$$\|g(t) - S_N(t)\|_\infty = \max_{x\in[-\pi,\pi]} \Big| g(t) - S_N(t) \Big| \to 0, \quad N \to \infty. \tag{8.5}$$

This means that we can also check that the convergence is independent of t to show that the series expansion converges to the function.

Remark. Fourier series generally converge rather slowly. In a practical sense, this implies that if we wish to approximate a signal through Fourier Synthesis, then we either need to let N become very large, which is computationally expensive, or choose a smaller N which could potentially cut off important information.

Example 8.4.1. Let us determine how the Fourier series S_N^k of the functions f_k below converge to f_k, for $k \in \{1, 2, 3\}$. All the following functions are 2π-periodic.

$$f_1(t) = \begin{cases} 1,\ 0 < t < \pi \\ -1,\ -\pi < t < 0, \end{cases} \qquad f_2(t) = \begin{cases} \pi - |t|,\ t \in [-\pi, \pi], \\ 0, \text{otherwise}, \end{cases} \qquad f_3(t) = \begin{cases} t,\ t \in [-\pi, \pi], \\ 0, \text{otherwise}. \end{cases}$$

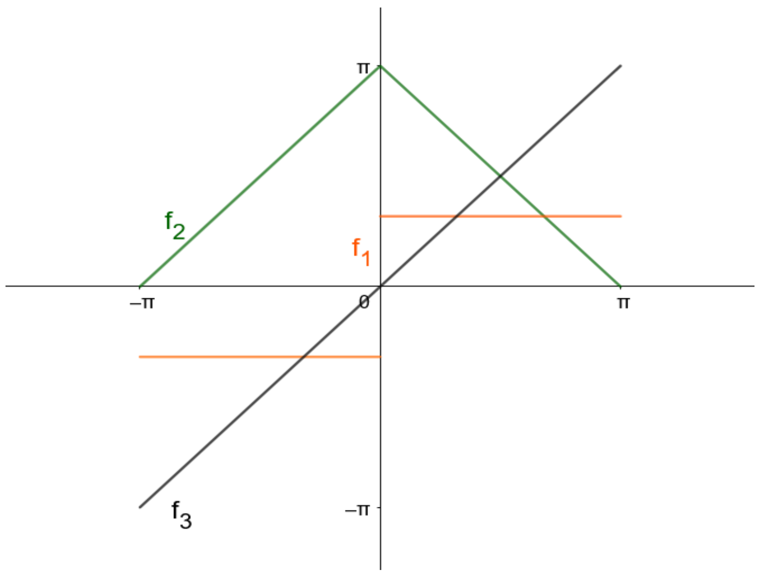

Figure 8.3: A sketch of the functions f_1, f_2 & f_3 on $[-\pi, \pi]$.

Imagine if we extend these functions onto the whole real line, then f_2 becomes the so-called Triangle Wave. Due to the discontinuities which appear for f_1 and f_3, the visualization is more prettier if we draw a vertical line between the edges at each discontinuity. Then f_1 traces the Square Wave whilst f_3 represents the Saw-tooth Wave. If we were to use Fourier series to approximate these three functions then only f_2 would be free of the Gibbs phenomenon (Cheever, 2020). With this in mind, we begin by expanding f_1 into its Fourier series. Observe that $f_1(t)$ is an odd function, which implies that its Fourier series will only have proper contributions from its sine terms. Then the expansion is a lot simpler to perform by using the trigonometric construction. Firstly the coefficients a_n and b_n can be calculated by

$$
\begin{aligned}
a_n &= \frac{1}{\pi}\langle f_1, \cos nt\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt = \frac{1}{\pi}\int_{-\pi}^{\pi}(\text{odd function})\,dt = 0, \quad \forall\, n\in\mathbb{N}.\\
b_n &= \frac{1}{\pi}\langle f_1, \sin nt\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt\\
&= \frac{1}{\pi}\left[\int_{-\pi}^{0} -\sin nt\,dt + \int_{0}^{\pi}\sin nt\,dt\right] = \frac{1}{\pi}\left[\frac{\cos nt}{n}\bigg|_{-\pi}^{0} + \frac{-\cos nt}{n}\bigg|_{0}^{\pi}\right]\\
&= \frac{1}{\pi}\left[\frac{1-(-1)-\big((-1)-1\big)}{n}\right] = \frac{4}{\pi n}, \qquad\qquad \text{where } \pm\cos n\pi = \begin{cases}\pm 1 \text{ if } n \text{ odd},\\ 0 \text{ if } n \text{ even}.\end{cases}
\end{aligned}
$$

Hence, the Fourier series expansion is found as

$$
f_1(t) \sim \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\sin nt}{n} = \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{\sin(2k-1)t}{2k-1}
$$

Now let us check how this Fourier series converges towards the function f_1. Since f_1 is discontinuous at the points $0, \pm\pi, \pm 2\pi, ...$ we immediately know that it cannot converge uniformly. The reason comes from a theorem introduced by Weierstrass which requires that a uniform limit of continuous functions, like S_N, must originate from a continuous function. However, we see that f_1 satisfies the conditions given by Theorem 8.3.1, so it follows that $S_N^1(t)$ converges pointwisely towards f_1. Specifically, the Fourier series converges towards $f_1(t) \quad \forall t \neq 0, \pm\pi, \pm 2\pi, ...$ and it converges to $\frac{(-1)+1}{2} = 0$ at $t = 0, \pm\pi, \pm 2\pi, ...$ Finally, since $f_1 \in L^2(-\pi, \pi)$ then this implies that its Fourier series converges towards f_1 in the L^2-norm as well, based on the definition of L^2 convergence in Equation (8.4).

In contrast, observe from Figure [8.3] that f_2 is both continuous and piecewise smooth (also verified through its derivative), which implies that we can use Theorem 8.4.1 to say that its Fourier series converges *uniformly*.

Lastly, f_3 can be expanded into $S_N^3(x) = \sum_{k=1}^{\infty}\frac{-2(-1)^k}{k}\sin kx$ (see appendix for an outline of the expansion). Then the type of convergence follows similarly to that of f_1. Particularly, the pointwise convergence

$$
\lim_{N\to\infty} S_N^3(t) = \begin{cases} f_3(t), & \text{if } t \neq 0 \pm \pi, \pm 2\pi, ...,\\ \frac{(-\pi)+\pi}{2} = 0, & \text{if } t = 0 \pm \pi, \pm 2\pi, ... \end{cases}
$$

Example 8.4.2. Recall the Dirichlet function from Section 7.1. It can be verified that this function has Fourier coefficients which are 0, yet the series converges towards $f(t)$ in the L^2-norm. As this seems like a rather artificial example, let us instead consider the following function. We define the Modified Dirichlet function on the interval $[-\pi, \pi]$ as the function $f_D(t) \in L^2(-\pi, \pi)$ where

$$f_D(t) = \begin{cases} 1, \text{ if } t \in \mathbb{Q} \cap [-\pi, \pi], & \text{denoted } I_1 \\ t, \text{ if } t \in \mathbb{R}/\mathbb{Q} \cap [-\pi, \pi] & \text{denoted } I_2. \end{cases}$$

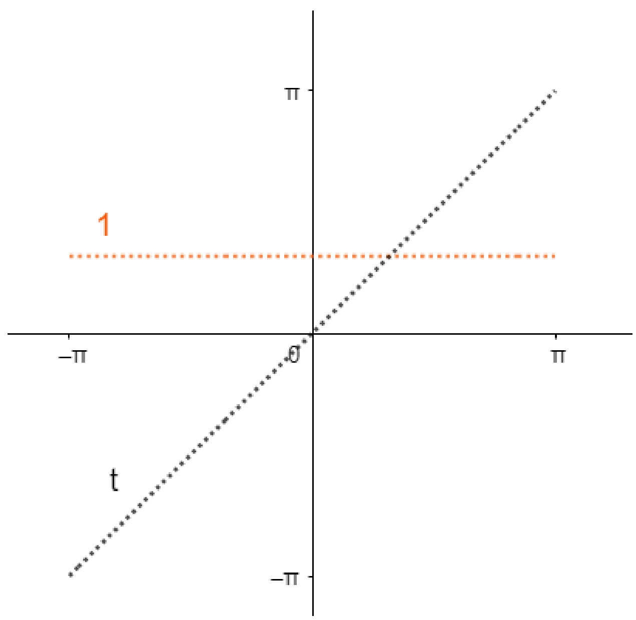

Figure 8.4: A sketch of the Modified Dirichlet function

Before we begin, observe that the Modified Dirichlet function coincides with the function f_3 whenever $t \in [-\pi, \pi]$ is irrational. Similarly to the Dirichlet function, this function is not Riemann integrable. Luckily, it is Lebesgue integrable, where

$$\int_I f_D(t)dt = \int_{I_1} f_D(t)dt + \int_{I_2} f_D(t)dt.$$

Then to find its Fourier coefficients, we simply divide the integral over the two relevant intervals. Since $t \in I_2 \Rightarrow f_D(t) = f_3(t)$ it only remains to find the contributions from $f_D(t) = 1$, when $t \in I_1$. Afterwards we simply add together the respective contributions to find the coefficients. However, since $I_1 \subset \mathbb{Q}$, then I_1 is of zero measure. It follows that the coefficients of $f_D(t) = 1$ all become 0, since we integrate simple functions over a set of zero measure. Finally, we get that the Fourier series expansion of the Modified Dirichlet function is identical to the expansion of $f_3(t)$.

$$S_N^D(x) = \sum_{k=1}^{\infty} \frac{-2(-1)^k}{k} \sin kx.$$

The result means that even though the Modified Dirichlet function is discontinuous at **all** points, its Fourier series converges towards a continuous function.

8.5 Convolution and the Dirichlet Kernel

In Fourier Analysis one often encounters a multitude of functions to study. As such, a useful tool is the ability to convolve pairs of functions.

Definition 8.5.1. Let $f, g \in L^1(\mathbb{R})$, then the convolution of these functions $f * g : \mathbb{R} \to \mathbb{C}$ combines them into a new *continuous* function by

$$(f * g)(t) = \int_{-\infty}^{\infty} f(t-s)g(s)\, ds. \tag{8.6}$$

This integral expresses how much the function f coincides with g as it is being shifted, or moved, across it (Weisstein, 2008). Note that Equation (8.6) may hold under other restrictions on f and g as well.

Example 8.5.1. Let us calculate the following convolution: $\underbrace{\chi_{[0,1]}(t)}_{f} * \underbrace{\chi_{[0,2]}(t)}_{g}$.

We begin by noticing how the convolution will behave at the end-points.

Clearly $f = 0 \begin{cases} t < 0 \\ t > 1 \end{cases}$ and $g = 0 \begin{cases} t < 0 \\ t > 2, \end{cases}$ which implies that $f * g = 0 \begin{cases} t < (0+0) = 0 \\ t > (1+2) = 3. \end{cases}$

Now we can use the definition of the convolution to figure out how the functions combine inside the interval $[0,3]$. Then

$$\left(\chi_{[0,1]} * \chi_{[0,2]}\right)(t) = \int_{-\infty}^{\infty} f(t-s) * g(s)ds, \quad \text{where} \begin{cases} f(t-s) = 1, & \text{if } 0 < t-s < 1 \\ g(s) = 1, & \text{if } 0 < s < 2. \end{cases}$$

$$= \int\limits_{[0,2]\cap[-1+t,t]} 1 \cdot 1 ds = \int\limits_{\max\{0,-1+t\}}^{\min\{2,t\}} 1\, ds = \begin{cases} \mathbf{0}, \text{ if } t < 0, \\ t - 0 = \mathbf{t}, \text{ if } 0 < t < 1 \\ t - (-1+t) = \mathbf{1}, \text{ if } 1 < t < 2, \\ 2 - (-1+t) = \mathbf{3} - \mathbf{t}, \text{ if } 2 < t < 3, \\ \mathbf{0}, \text{ if } t > 3. \end{cases}$$

In Figure [8.5] we provide a sketch of this convolution, and we see that it is indeed a continuous function.

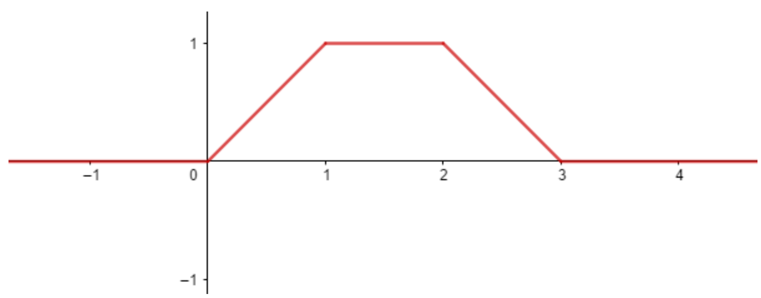

Figure 8.5: The convolution of two functions f and g.

Convolutions can be a very useful tool in regards to Fourier series and Fourier transformation, particularly along with the so-called Dirichlet Kernel.

Definition 8.5.2 (The Dirichlet Kernel)**.** Given any $N \in \mathbb{N}$, the function

$$\mathcal{D}_N(t) := \sum_{k=-N}^{N} e^{ikt} = \frac{\sin(N+\frac{1}{2})t}{\sin(\frac{t}{2})} \tag{8.7}$$

is called the Dirichlet Kernel. Let us present a proof which shows that the representation is valid.

Proof. Let us begin by converting the initial partial sum into a more suitable form. To start of with we use Euler's identity, then recall that $\cos(x)$ is an *even* function, i.e., $\cos(-x) = \cos(x)$, whilst $\sin(x)$ is an *odd* function, i.e., $\sin(-x) = -\sin(x)$. It follows that the sum of sines cancels $\forall N \neq 0$, whereas the sums of cosine add up. It remains to check what happens when $N = 0$:

$$\begin{aligned}\mathcal{D}_N(t) := \sum_{k=-N}^{N} e^{ikt} &= \sum_{k=-N}^{N} \big(\cos(kt) + i\sin(kt)\big) = \sum_{k=-N}^{N} \cos(kt) + i \sum_{k=-N}^{N} \sin(kt) \\ &= 2\sum_{k=1}^{N} \cos(kt) + \cos(0\cdot t) + \sin(0\cdot t) = 1 + 2\sum_{k=1}^{N} \cos(kt).\end{aligned}$$

As $\mathcal{D}_N(t)$ now has a more suitable form, we can multiply and divide the R.H.S. by $\sin(\frac{t}{2})$ and apply the trigonometric identity $\boxed{2\cos\alpha\sin\beta = \sin(\alpha+\beta) - \sin(\alpha-\beta)}$.

$$\begin{aligned}\mathcal{D}_N(t) &= \frac{\sin(\frac{t}{2}) + 2\sum_{k=1}^{N}\cos(kt)\cdot\sin(\frac{t}{2})}{\sin(\frac{t}{2})} &&= \frac{\sin(\frac{t}{2}) + \sum_{k=1}^{N}\Big(\sin(kt+\frac{t}{2}) - \sin(kt-\frac{t}{2})\Big)}{\sin(\frac{t}{2})} \\ &= \frac{\sin(\frac{t}{2}) + \sin(N+\frac{t}{2}) - \sin(\frac{t}{2})}{\sin(\frac{t}{2})} &&\text{where most of the terms cancel, leaving only} \\ &= \frac{\sin(N+\frac{1}{2})t}{\sin(\frac{t}{2})} &&\text{which was the desired result.}\end{aligned}$$

This completes the proof. □

By combining Equation (8.6) and (8.7), we can take the convolution of $\mathcal{D}_N(t)$ with any 2π periodic function f. This provides us with another representation of its Fourier series where the convergence may be studied through the properties of the Dirichlet kernel (Walker, 2003, p. 8).

$$(f * \mathcal{D}_N)(t) = \int_{-\pi}^{\pi} f(t-s)\mathcal{D}_N(s)\,ds = S_N(t).$$

Particularly, this relationship may be used to explain the Gibb's phenomenon and ringing, where a convolution between the Dirichlet Kernel and the square wave, i.e., $(\mathcal{D}_N * S_q)(t)$ can be investigated to verify the 18% increase found in the spike. We end this chapter by mentioning that there are many modifications which may be performed on the partial sums of Fourier series to avoid some of the convergence issues we have encountered. Through a method called *arithmetic means* one may even achieve L^1-convergence for Kolmogorov's function (Walker, 2003, p. 9). In the next chapter we move on to Fourier's other major contribution, the Fourier transform.

Chapter 9

Application to sampling of band-limited signals

In this chapter the Fourier transform for $L^1(\mathbb{R})$ functions, along with important properties are defined. We also discuss how to extend the transform to apply for $L^2(\mathbb{R})$ functions. Afterwards, the Paley-Wiener space and the Shannon-sampling theorem are introduced. Based on this, we try to explain how signal processing works. Lastly an account of Riesz basis is provided along with applications for the Fourier transform.

9.1 The Fourier transform.

Before the introduction of Fourier series, mathematicians would often find it problematic to combine the concept of a function with infinite summations (Bressoud, 2007). Fourier series are a useful tool for investigating periodic functions, yet not all signals are structured this way. We say that a periodic function becomes aperiodic in the limit as its period, $L \to \infty$. Generally for aperiodic functions we do not know which frequencies the function is limited to. Therefore we replace the Fourier series with an integral over all frequencies. The classical way to define the Fourier transform is for absolutely integrable functions.

Definition 9.1.1 (Fourier transform)**.** We say that the Fourier transform operates on a function $f \in L^1(\mathbb{R})$ and assigns a new function $\widehat{f} : \mathbb{R} \to \mathbb{C}$ through

$$\widehat{f}(x) := \mathcal{F}f(x) = \int_{-\infty}^{\infty} e^{-2\pi ixt} f(t)\, dt. \tag{9.1}$$

Observe that the transformed function, $\widehat{f}$ takes x as its input, whereas the original function f takes t as its input. More concretely, we often consider the values of $\widehat{f}(x)$ as the frequency domain, and $f(t)$ as the time domain. This is particularly done in signal processing, where $f(t)$ represents the original signal. The Fourier transform expresses a function or signal as a superposition of sinusoids. Consider the function $f(t) := e^{-t}\chi_{[0,1]}(t)$, then its Fourier transform becomes

$$\begin{aligned}\widehat{f}(x) &= \int_{-\infty}^{\infty} e^{-2\pi ixt} e^{-t} \chi_{[0,1]}(t)\, dt = \int_{0}^{1} e^{t(-2\pi ix-1)}\, dt \\ &= \frac{e^{t(-2\pi ix-1)}}{-2\pi ix - 1}\bigg|_0^1 = \frac{1}{2\pi ix + 1}\Big(1 - e^{-(2\pi ix+1)}\Big).\end{aligned}$$

Common for all such transformations is that they tend towards zero as $x \to \pm\infty$, i.e., $\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$. This property is officially known as the *Riemann Lebesgue Lemma*. Moreover, under certain conditions, the values of $\widehat{f}$ allow us to reconstruct the function f. This is an important result. In other words, the Fourier transform is invertible:

Theorem 9.1.1 (Inverse transform for $L^1(\mathbb{R})$ functions.)**.** *Let $f, \widehat{f} \in L^1(\mathbb{R})$, then*

$$f(t) := \mathcal{F}^{-1}\widehat{f}(x) = \int_{-\infty}^{\infty} e^{2\pi ixt} \widehat{f}(x)\, dx \quad \textit{for almost all } x \in \mathbb{R}. \tag{9.2}$$

Whenever f is continuous, this expression holds pointwise for all $x \in \mathbb{R}$.

When solving Differential Equations through Fourier transformations, the inversion formula allows us to retrieve the solution on the desired form. In signal processing, this is what allows us to represent a signal through its frequencies (Christensen, 2010, p. 136). The Fourier transform can also be defined for functions which are square integrable, i.e., $f \in L^2(\mathbb{R})$. However, due to convergence issues the transformation is no longer an integral in the classical notion. Instead it exists as a limit which converges in the $L^2(\mathbb{R})$ norm. This is achieved through *extending* the linear operator $\mathcal{F}$ from a convenient subspace where it is well-defined. In our case, the $C_c(\mathbb{R})$-space is convenient because if we equip it with the L^2-norm, one may verify that:

(1) $\|\mathcal{F}f\|_2 = \|f\|_2, \quad \forall f \in C_c(\mathbb{R}),$ where $\mathcal{F} : C_c(\mathbb{R}) \to L^2(\mathbb{R})$

(2) The closure $\overline{C_c(\mathbb{R})} = L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$

The fact that $C_c(\mathbb{R})$ is dense in $L^2(\mathbb{R})$ means that for all $\mathbf{v} \in V$, and $\forall \varepsilon > 0$ there exists $\mathbf{v}_1 \in V_1$ such that $\|\mathbf{v} - \mathbf{v}_1\| < \varepsilon$. Furthermore taking (1) and (2) into consideration, it is possible to extend the operator $\mathcal{F}$ to hold for $L^2(\mathbb{R})$ functions. This is analogous to saying $\widehat{f}(x)$ is well defined $\forall f \in L^2(\mathbb{R})$, (Christensen, 2010, p. 55). In summary, we say that the Fourier transform can be extended from $C_c(\mathbb{R})$ to $L^2(\mathbb{R})$ as a unitary operator $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ which satisfies

(i) $\|\widehat{f}\|_{L^2} = \|f\|_{L^2}, \quad \forall f \in L^2(\mathbb{R})$ (Parseval's Equality)

(ii) $\langle \widehat{f}, \widehat{g} \rangle = \langle f, g \rangle, \quad \forall f, g \in L^2(\mathbb{R})$

(iii) $\mathcal{F}^*\mathcal{F} = \mathcal{F}\mathcal{F}^* = I$

(iv) $f(t) = \int_{-\infty}^{\infty} e^{2\pi ixt} \widehat{f}(x)\, dx, \quad \forall f \in L^2(\mathbb{R}), \quad \widehat{f} \in L^2(\mathbb{R})$

where

$$\|f\|_{L^2} = \left(\int_{-\infty}^{\infty} |f(t)|^2\, dt \right)^{1/2}.$$

The third property tells us that he Fourier transform for $L^2(\mathbb{R})$-functions is invertible, and its form is found in (iv). Parseval's Equality confirms that for physical problems involving energy, the total amount of energy contained within a signal $f(t)$ may be calculated from the frequency domain signal through the sum of its expanded Fourier coefficients (Walker, 2003, p. 8). Finally, given a number $R > 0$ and $f(t) \in C_c(\mathbb{R})$ the expression

$$\left\| \int_{-R}^{R} e^{-2\pi ixt} f(t)\, dt - \widehat{f}(x) \right\|_{L^2(\mathbb{R})} \to 0 \quad \text{as } R \to \infty$$

describes the convergence in the L^2-norm.

Let us now introduce an interesting operator which can be shown to be unitary, bounded and linear in the $L^2(\mathbb{R})$-space (Christensen, 2010, p. 120:122). Due to these properties, the following operator (along with *modulation*, and *dilation* operators) is often used in Fourier analysis.

The *Translation* operator moves a function along the x-axis:

$$\big(T_a f\big)(x) := f(x-a), \quad a \in \mathbb{R}$$

Since the operation is linear, bounded and unitary, this implies that the new function is still located inside $L^2(\mathbb{R})$.

$$\begin{aligned}\big(\,\mathcal{F}T_a f\,\big)(x) &= \int_{-\infty}^{\infty} e^{-2\pi i x t}\big(T_a f\big)(t)\,dt = \int_{-\infty}^{\infty} e^{-2\pi i x t} f(t-a)\,dt \qquad \boxed{\text{Let } u = t-a,\, t = u+a} \\ &= \int_{-\infty}^{\infty} e^{-2\pi i x(u+a)} f(u)\,du = e^{-2\pi i x a}\int_{-\infty}^{\infty} e^{-2\pi i x u} f(u)\,du \\ &= e^{-2\pi i x a}\,\widehat{f}(x) = \big(E_{-a}\widehat{f}\,\big)(x).\end{aligned}$$

In the last step of the calculation we used the definition of the *dilation* operator.
Another interesting property of the Fourier transform follows from the Convolution Theorem which was *allegedly* introduced by Percy John Daniell in 1920 (Dominguez, 2015). The theorem relates the Fourier transform of a convolution between functions to the product of their respective transforms, i.e., that $\widehat{f * g}(x) = \widehat{f} \cdot \widehat{g}$. Similar to the Fourier transform there are some restrictions, specifically that it holds for all $x \in \mathbb{R}$ given $f, g \in L^1(\mathbb{R})$, and almost all $x \in \mathbb{R}$ for $f \in L^2(\mathbb{R})$ and $g \in L^1(\mathbb{R})$. We omit the proof. If we consider $f(t)$ and $g(t)$ to be two arbitrary signals, then a direct implication of the Convolution Theorem is that the convolution in the time domain equals multiplication in the frequency domain (DSP Illustrations, 2020). In such cases it is often time-beneficial to transform the signals separately, calculate their product and then find the inverse Fourier transform of the result. This would surely be the case if we wanted to find the Fourier transform of the convolution from example 8.5.1.

We are now nearing a point where we can discuss one of the applications for Fourier transforms and convergence of Fourier series, namely sampling. By this we mean that given a signal, $f(t)$, we wish to record or *sample* enough datapoints in order to reproduce the original signal.

9.2 Sampling theory.

Why do we need sampling? Well, signals found in the real world, such as musical tones, are both continuous in time and amplitude. We say that they are *analog* signals. The ability to repeat these signals at any time and place along with perfect quality has long been yearned for. Although instruments such as the gramophone and the cassette player were introduced between the late 1800s and mid 1900s, they came with their limitations. Large production expenses, storage, the fact that analog signals are prone to noise, and poor durability were some of them. With the rise of computers and digital machines a new possibility emerged. Initially, analog signals posed a problem for digital storage on computers which are discrete by nature. A perfect recording of a continuous signal would require precise measurements and great storage capacity. It was also believed that information would be lost if a continuous measurement was not performed. However, as we will show, this may all be avoided through sampling. This method allows us to convert an analog signal to a *digital* signal which is discrete in both time and amplitude.

Definition 9.2.1. The Paley-Wiener space, often shortened PW is a subspace of $L^2(\mathbb{R})$. It is often defined for a general interval length σ. For simplicity we consider a particular interval given by

$$PW := \left\{ f \in L^2(\mathbb{R}) \,\middle|\, \operatorname{supp}\widehat{f} \subseteq \left[-\frac{1}{2}, \frac{1}{2}\right] \right\}. \tag{9.3}$$

Definition 9.2.2. A signal is band-limited with band $(-\Omega, \Omega)$ if its Fourier transform

$$\widehat{f}(x) = \int_{-\infty}^{\infty} e^{-2\pi ixt} f(t)\, dt$$

vanishes for $|x| > \Omega$. In essence, this is the same as requiring compact support for $\widehat{f}$.

According to these two definitions, we have that the functions belonging to PW are band-limited. In Section 8.1 we mentioned that $\{e^{2\pi ikt}\}_{k=-\infty}^{\infty}$ forms an orthonormal basis for $L^2(-\frac{1}{2}, \frac{1}{2})$. We will now demonstrate an important result by finding its inverse Fourier transform. Thus,

$$\begin{aligned}\mathcal{F}^{-1} e^{2\pi ixt} &= \int_{-\infty}^{\infty} e^{2\pi ikt} \cdot e^{2\pi ixt}\, dt = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i(x+k)t}\, dt \\ &= \frac{e^{2\pi i(x+k)t}}{2\pi i(x+k)}\bigg|_{-\frac{1}{2}}^{\frac{1}{2}} = \frac{e^{\pi i(x+k)} - e^{-\pi i(x+k)}}{2\pi i(x+k)} = \frac{\sin \pi(x+k)}{\pi(x+k)}\end{aligned}$$

which leads us to the cardinal sine function (defined for $t \in (-\frac{1}{2}, \frac{1}{2})$ with $k = 0$):

$$\operatorname{sinc} t := \begin{cases} \frac{\sin \pi t}{\pi t}, & \text{if } t \neq 0, \\ 1, & \text{if } t = 0. \end{cases} \tag{9.4}$$

On this interval, the sinc function is normalized, see also Figure [9.1] where it has been extended to the interval (-6, 6) to show its behaviour.

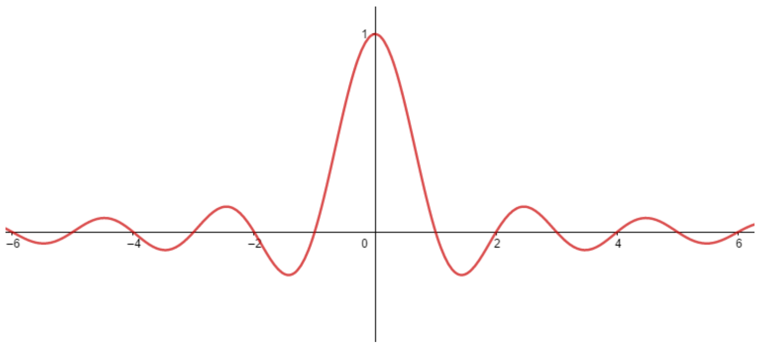

Figure 9.1: The cardinal sine function.

Theorem 9.2.1 (Shannon's sampling theorem)**.** *The functions* $\{\operatorname{sinc}(t+k)\}_{k\in\mathbb{Z}}$ *form an orthonormal basis for PW. If* $f \in PW$ *is continuous, then*

$$f(t) = \sum_{k\in\mathbb{Z}} f(k) \operatorname{sinc}(t-k) \tag{9.5}$$

where $f(k)$ *represents the* k^{th} *sample.*

The series in Equation (9.5) converges towards $f(t)$ both pointwise and in the L^2-norm (Christensen, 2010, p. 151). Notice that this allows us to reconstruct the original function $f(t)$ through the sampled values $f(k)$. Consequently, this theorem is said to have laid the foundation for digitization of images, sound and music in the modern-day-world. The work done by E.T. Whittaker (1915) and Harry Nyquist (1928) contributed towards the discovery of Theorem 9.2.1. Actually, Vladimir Kotelnikov reached the same conclusion as Theorem 9.2.1 in 1933, but he was unable to publish his findings. Therefore it was not until Claude Shannon re-introduced it in two articles in 1948-1949 that the full strength of the theorem was realized within communication theory (Yankin, 2019). Perhaps due to all these different discoveries, it is also referred to as the *cardinal theorem of interpolation* to avoid confusion.

The theorem proves that it is technically possible to perfectly reconstruct any continuous band-limited analog signal without any distortions or errors through discrete sampling of its frequency. For this to be true, it requires that the sampling rate is greater than the so-called *Nyquist rate*, $\frac{1}{2\,dt}$, when sampling at regular time intervals dt (Avantaggiati, Loreti, & Vellucci, 2016, p. 1). In other words, for a given signal $f(t)$ whose largest frequency is ω Hz, the Nyquist criteria allows for a perfect reconstruction when sampling more frequently than 2ω Hz. In certain cases, this may result in large file-sizes, which is undesirable and should therefore be avoided when possible.

Consider a sampling rate exactly equal to the highest frequency, i.e., ω Hz, then it is possible that we are only sampling the peak-value of our signal. Ultimately our result will produce a straight line where almost all information of complexity is lost. Sampling at a different rate (between 0 Hz and 2ω Hz) may result in reproducing a signal with lower frequency and smaller amplitudes than the original. Figure [9.2] illustrates the importance of following the Nyquist rate.

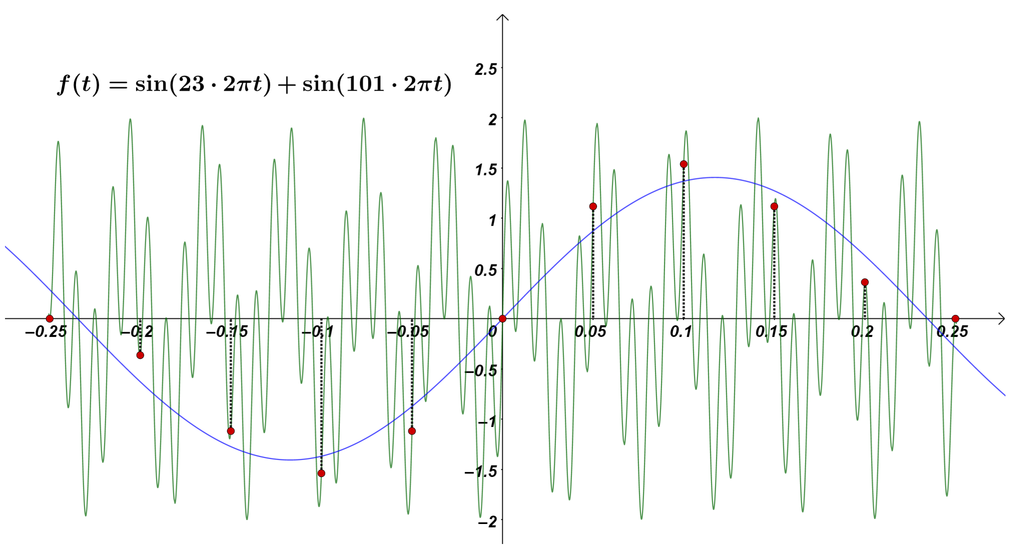

Figure 9.2: A periodic function $f(t)$ with approximate period $L = 0.225$ (green). Sampled uniformly at intervals of $x = 0.05$ and reconstructed into an alias (blue).

In this figure we have sampled uniformly with a low sampling frequency. The reconstructed continuous time signal found in Figure [9.2] is then called an *alias* to the original. It is "indistinguishable" to the original before comparison after reproduction. This phenomena is called aliasing and is a known problem in signal processing (Brunton, 2020, 8:00-9:00). Aliasing occurs due to the presence of unwanted components and/or loss of frequencies found in the original signal. In our example, we see that much of the information stored in $f(t)$ is lost with respect to its alias. Using the same method as above, we would therefore need to increase our sampling rate to get a better approximation. Other issues which may occur when implementing the Shannon sampling theorem are slow convergence rates, sampling at uniformly distributed intervals (which is difficult to realize with physical instruments), large file-sizes, as well as requiring the functions to be band-limited. In the next section we will look at the possibility of allowing the samples to be distributed irregularly to combat the sampling issue.

9.3 Riesz basis

Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis for H. Let $T : H \to H$ be a bounded invertible linear operator. We denote $\mathrm{u}_n = Te_n$ and $\mathrm{v}_n = (T^{-1})^* e_n$.

Definition 9.3.1. A system of elements $\{\mathrm{u}_k\}_{k=1}^{\infty}$ in a Hilbert space is called a Riesz basis if

$$\mathrm{u}_n = Te_n, \quad n = 1, 2, 3, ...$$

where $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis for H and T is a linear bounded invertible operator from H to itself.

If $\{\mathrm{u}_n\}_{n=1}^{\infty}$ is a Riesz basis, then we say that:

1. There is a unique system $\{\mathrm{v}_n\}_{k=1}^{\infty}$ which is biorthogonal to $\{\mathrm{u}_n\}_{k=1}^{\infty}$.
2. Every $g \in H$ admits a unique representation $g = \sum_1^{\infty} c_n \mathrm{u}_n$, where $c_n = \langle g, \mathrm{v}_n \rangle$.
3. There is a constant $c \geq 1$ such that for every sequence $\{c_n\}_{n=1}^{\infty} \in l^2(\mathbb{N})$,

$$\frac{1}{c} \sum_1^{\infty} |c_n|^2 \leq \Big\| \sum_1^{\infty} c_n \mathrm{u}_n \Big\|_H^2 \leq c \sum_1^{\infty} |c_n|^2.$$

We say that a system $\{\mathrm{v}_k\}_{k=1}^{\infty}$ is *biorthogonal* to another system $\{\mathrm{u}_k\}_{k=1}^{\infty}$ if their inner product is reduced to the Kronecker delta function, i.e., $\langle \mathrm{u}_n, \mathrm{v}_k\rangle = \begin{cases} 1 \text{ if } n = k, \\ 0 \text{ otherwise.} \end{cases}$

Lemma 9.3.1. *Let $\{e_n\}_1^{\infty}$ be an orthonormal basis for a Hilbert space H, and $T : H \to H$ a linear bounded invertible operator. Then the systems Te_n and $(T^{-1})^* e_n$ are biorthogonal.*

Proof. From the inner product between these systems and properties of the adjoint operator, we have that

$$\langle Te_n, (T^{-1})^* e_k\rangle = \langle e_n, T^*(T^{-1})^* e_k\rangle = \langle e_n, I \cdot e_k\rangle = \langle e_n, e_k\rangle = \delta_{nk},$$

which completes the proof. □

Lemma 9.3.2. *$\{\mathrm{u}_k\}_1^{\infty}$ and $\{\mathrm{v}_k\}_1^{\infty}$ both form a basis for H.*

Proof. Take any $f \in H$, then we can represent it through the orthonormal basis $\{e_k\}_1^{\infty}$ as $f = \sum_{k=1}^{\infty} \langle f, e_k\rangle e_k$. Let the operator T act on f, and use the definition of u_k. Thereby

$$Tf = T\left(\sum_{k=1}^{\infty} \langle f, e_k\rangle e_k\right) = \sum_1^{\infty} \langle f, e_k\rangle Te_k = \sum_1^{\infty} \langle f, e_k\rangle \mathrm{u}_k,$$

where the second equality is true because $\langle f, e_k\rangle := c_k \in l^2(\mathbb{N})$ and T only acts on elements of H. Choose any $g \in H$ and fix $f = T^{-1}g$ such that $g = Tf$. Using the above expression, properties of the adjoint operator and definition of v_k it follows that

$$g = Tf = \sum_1^{\infty} \langle f, e_k\rangle \mathrm{u}_k = \sum_1^{\infty} \langle T^{-1}g, e_k\rangle \mathrm{u}_k = \sum_1^{\infty} \langle g, (T^{-1})^* e_k\rangle \mathrm{u}_k = \sum_1^{\infty} \langle g, \mathrm{v}_k\rangle \mathrm{u}_k.$$

This proves that $\{\mathrm{u}_k\}_1^{\infty}$ forms a basis for H, and the proof for $\{\mathrm{v}_k\}_1^{\infty}$ follows a similar procedure. □

Observe that $\langle g, \mathrm{v}_k\rangle = c_k \in l^2(\mathbb{N})$. Then if we define another operator $B : l^2(\mathbb{N}) \to H$ where for any sequence $c_k \in l^2(\mathbb{N})$, we have $B\{c_k\}_1^{\infty} = \sum_1^{\infty} c_k \mathrm{u}_k$.

Lemma 9.3.3. *B is a bounded linear operator with $\|B\| \leq \|T\|$.*

Proof. We begin by imposing the norm on the operator acting on c_k. Then from the Cauchy-Schwartz equality and orthonormality of $\{e_k\}_1^\infty$, observe that

$$\begin{aligned}\left\|B\{c_k\}\right\|_H &= \Big\|\sum_1^\infty c_k \mathrm{u}_k\Big\|_H = \Big\|T(\sum_1^\infty c_k e_k)\Big\|_H \\ &\leq \|T\| \Big\|\sum_1^\infty c_k e_k\Big\|_H = \|T\| \sqrt{\sum_1^\infty |\langle c_k, e_k\rangle|} \\ &= \|T\| \sqrt{\sum_1^\infty |c_k|^2} = \|T\| \, \|\{c_k\}\|_{l^2}.\end{aligned}$$

Finally, the norm of B is simply

$$\|B\| = \sup_{c_k \neq 0} \frac{\|B\{c_k\}\|_H}{\|\{c_k\}\|_{l^2}} \leq \frac{\|T\|\|\{c_k\}\|_{l^2}}{\|\{c_k\}\|_{l^2}} = \|T\|, \qquad \checkmark$$

so B is bounded. □

Observe that the facts $\mathrm{u}_n = Te_n$ and B bounded implies the **R.H.S.** of the 3^{rd} Riesz basis property.

Lemma 9.3.4. *The adjoint of B is $B^* : H \to l^2$, given by $B^* f = \{\langle f, \mathrm{u}_k\rangle\}_{k=1}^\infty$*

Proof. By taking advantage of linearity in the first argument along with $\langle \mathrm{v}, \mathrm{w}\rangle = \overline{\langle \mathrm{w}, \mathrm{v}\rangle}$, we get that

$$\begin{aligned}\langle B\{c_k\}, f\rangle = \Big\langle \sum_1^\infty c_k \mathrm{u}_k, f\Big\rangle &= \sum_1^\infty c_k \langle \mathrm{u}_k, f\rangle \\ &= \sum_1^\infty c_k \overline{\langle f, \mathrm{u}_k\rangle} = \langle \{c_k\}, B^* f\rangle.\end{aligned}$$

□

We are now ready to state a well-known theorem made by the Ukrainian mathematician Mikhail Kadets in 1964.

Theorem 9.3.5 (Kadets $\frac{1}{4}$-Theorem.)**.** *Let $\delta_n \in \mathbb{R}$ and assume $0 < L < \frac{1}{4}$. Then for every set $\Lambda = \{n + \delta_n\}_{n=-\infty}^\infty$, $\quad |\delta_n| \leq L$, for $n \in \mathbb{Z}$, the exponential system*

$$\left\{e^{2\pi i (n+\delta_n)t}\right\}_{n=-\infty}^\infty$$

is a Riesz basis for $L^2(-\frac{1}{2}, \frac{1}{2})$.

We regard the set $\Lambda = \{n + \delta_n\}_{n=-\infty}^{\infty}$ as a small "perturbation" of the set of integers $\mathbb{Z} = \{n\}_{-\infty}^{\infty}$. The fact that a perturbation of a known Riesz basis produces another Riesz basis, was originally introduced by Paley and Wiener in 1934 (Avantaggiati et al., 2016, p. 2). The theorem was proven by Kadet to be sharp, which means it fails immediately given that $L \geq \frac{1}{4}$. We call '1/4' an optimal bound when $\Lambda \subset \mathbb{R}$. Thus, every system of functions

$$\mathrm{u}_n(t) := e^{2\pi i(n+\delta_n)t}, \quad n \in \mathbb{Z} \tag{9.6}$$

forms a Riesz basis for $L^2(-\frac{1}{2}, \frac{1}{2})$ provided $|\delta_n| \leq L < \frac{1}{4}$. The system in Equation (9.6) is also called an *exponential* Riesz basis. Its elements are no longer orthogonal to each other in $L^2(-\frac{1}{2}, \frac{1}{2})$. Now it remains to check if we can use the set Λ to reconstruct a signal by sampling at a random rate. Observe that the 1^{st} and 2^{nd} benefits of a Riesz basis allows us to represent a function $g \in L^2(-\frac{1}{2}, \frac{1}{2})$ uniquely by

$$\begin{aligned} g(t) &= \sum_{1}^{\infty} \langle g, \mathrm{v}_n \rangle \, \mathrm{u}_n(t) \\ &\Updownarrow \qquad \text{(since } \mathrm{u}_n \text{ and } \mathrm{v}_n \text{ are biorthogonal)} \\ g(t) &= \sum_{1}^{\infty} \langle g, \mathrm{u}_n \rangle \, \mathrm{v}_n(t). \end{aligned} \tag{9.7}$$

Using the definition of inner products and Equation (9.6), we see that

$$\begin{aligned} \langle g, \mathrm{u}_n \rangle &= \int_{-1/2}^{1/2} g(t) \, \overline{e^{2\pi i(n+\delta_n)t}} \, dt \\ &= \int_{-1/2}^{1/2} g(t) \, e^{-2\pi i(n+\delta_n)t} \, dt = \widehat{g}(n + \delta_n), \end{aligned}$$

where $\widehat{g} \in PW$ since $g \in L^2(-\frac{1}{2}, \frac{1}{2})$. Let $\varphi := \widehat{g}$ and take the Fourier transform of Equation (9.7), then every $\varphi \in PW$ admits a representation

$$\varphi(x) = \sum_{n=-\infty}^{\infty} \varphi(n + \delta_n) \, \widehat{\mathrm{v}_n}(x), \quad x \in \mathbb{R}. \tag{9.8}$$

It remains to find an expression for the biorthogonal system $\widehat{\mathrm{v}_n}$. We begin with the following definition.

Definition 9.3.2. A set $\Lambda \subset \mathbb{R}$ is called uniformly discrete if

$$\delta_n(\lambda) = \sup_{\lambda \neq \lambda'} |\lambda - \lambda'| > 0, \quad \lambda, \lambda' \in \Lambda. \tag{9.9}$$

Which means that the elements of Λ are separated by $|\lambda - \lambda'| \geq \delta(\Lambda)$. We call $\delta(\Lambda)$ the *separation* constant. In the set of integers, $\mathbb{Z}$, one may easily verify that all elements are distanced by $\delta(\mathbb{Z}) = 1$. This implies that $\Lambda = \mathbb{Z}$ is uniformly discrete. In comparison, we can consider the set $\Lambda = \{\sqrt{n} \,|\, n \in \mathbb{N}\}$. Initially, the elements seem to satisfy the condition found in Equation (9.9), but notice that $\delta(\Lambda) = \sqrt{n+1} - \sqrt{n} \to 0$, when $n \to \infty$. Therefore this set is not uniformly discrete.

Since $\{e^{2\pi i(n+\delta_n)t} \,|\, (n+\delta_n) \in \Lambda\}$ forms a Riesz basis for $L^2(-\frac{1}{2}, \frac{1}{2})$, then this implies that $\Lambda = \{n + \delta_n\}_{n=-\infty}^{\infty}$ is uniformly discrete by a well known theorem.

Choose

$$\widehat{\mathrm{v}}(z) = \prod_{n\in\mathbb{Z}} \left(1 - \frac{z}{n+\delta_n}\right), \quad z \in \mathbb{C}.$$

and

$$\widehat{\mathrm{v}_n}(z) = \frac{\widehat{\mathrm{v}}(z)}{\widehat{\mathrm{v}}'(n+\delta_n)\cdot(z-n-\delta_n)}. \tag{9.10}$$

Then another theorem states that the functions $\widehat{\mathrm{v}_n}(z)$ and $\widehat{\mathrm{v}}(z)$, for $n \in \mathbb{Z}$ are well defined and satisfy the following properties

(i) $\widehat{\mathrm{v}_n}(z)$ and $\widehat{\mathrm{v}}(z)$, $n \in \mathbb{Z}$, are *entire* functions.

(ii) $\widehat{\mathrm{v}_n}(x) \in L^2(\mathbb{R}), \quad n \in \mathbb{Z}$.

(iii) $\left\langle e^{2\pi i(n+\delta_n)t}, \mathrm{v}_m(t)\right\rangle = \begin{cases} 1, & n = m, \\ 0, & n \neq m, \end{cases}$ where $\mathrm{v}_m(t)$ is the inverse transform of $\widehat{\mathrm{v}_m}(x)$.

Therefore the system $\{\mathrm{v}_n(t) \,|\, n \in \mathbb{Z}\}$ is indeed biorthogonal to $\{\, e^{2\pi i(n+\delta_n)t} \,|\, n \in \mathbb{Z}\}$, and we can use the representation of v_n given in Equation (9.10). Finally, Equation (9.8) allows us to extend the famous Shannon sampling theorem to the case when sampling points are randomly distributed as $n + \delta_n$, where $|\delta_n| \leq L < \frac{1}{4}$, for $n \in \mathbb{Z}$ and $\delta_n \in \mathbb{R}$. See Figure [9.3] below where we have performed a slight (random) perturbation of the points from Figure [9.2].

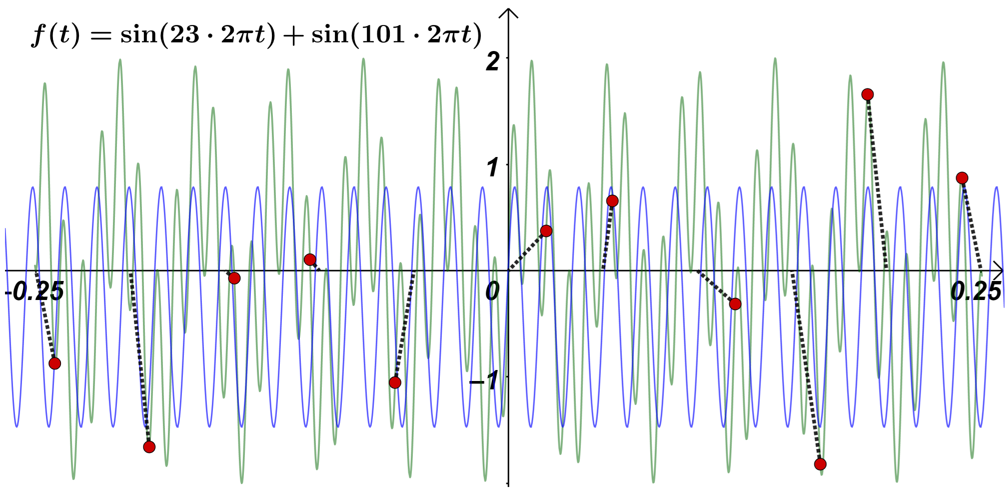

Figure 9.3: A perturbed sampling of the function $f(t)$.

The perturbation allows a discrepancy in measurement tools, but also provides us with a better approximation of the function $f(t)$. The exact perturbation can be seen in Table 9.1.

x	-0.25	-0.2	-0.15	-0.1	-0.05	0.0	0.05	0.1	0.15	0.2	0.25
$x + \delta_n$	-0.24	-0.19	-0.145	-0.11	-0.06	0.02	0.055	0.12	0.17	0.19	0.24

Table 9.1: Perturbed sampling points

We remind that the example shown in Figure [9.2] and [9.3] is only meant to illustrate the effect of perturbation.

The convergence in Equation (9.8) above is in the L^2-norm. It is possible to show that the series

$$\sum_{n=-\infty}^{\infty} \varphi(n + \delta_n)\, \widehat{\mathrm{v}_n}(z)$$

converges to $\varphi(z)$ uniformly on compact sets of the complex plane. Additionally, studies are being performed on situations where $\delta_n \in \mathbb{C}$, and finding the optimal bound (Avantaggiati et al., 2016, p. 3).

9.4 Applications of the Fourier Transform.

Perhaps the most natural example of a Fourier transform process which we can find, is the human ear. Sound is created due to interference in air pressure from vibrations between particles which propagate through air. When sound arrives at the ear drums they resonate in unison with the pressure-waves. The vibration waves are forwarded to the cochlea, located in the inner ear, which transforms these "air pressure signals" into nerve impulses. Lastly, the brain processes this information into hearing. The ear acquires information about volume and pitch. Volume is simply determined by the amplitude of the registered waves. A greater amplitude induces a large impact on the hair cells within the inner ear, which in turn propagates stronger nerve impulses towards the brain, resulting in louder volume. The number of vibrations per unit time, also known as frequency, determines the pitch. For humans, the frequency range is band-limited, covering roughly $20 \to 20\,000$ Hz (Campbell et al., 2018, p.1168:1172). All of this is done naturally by the ear, resulting in everything we hear.

The Fourier transform is a mathematical technique for doing a similar process. When applied to a time-domain function, it produces a frequency spectrum. If we take the Nyquist rate into consideration, we get an intuitive explanation to why most digital audio signals generally are sampled at approximately 44 kHz, a little more than <u>twice</u> the greatest frequency humans may register. Reproducing greater frequencies would not improve the perceived sound for humans and would therefore be wasteful.

In the modern day world we encounter many different signals. A broadband (or dense) signal has many prominent frequencies, i.e., peaks in its frequency spectrum. For such signals it is not possible to beat the Nyquist-rate and Shannon sampling theorem (Brunton, 2020, 14:30-16:54). However, in 2004 mathematicians used the theory based on Riesz bases and applied it to signals whose frequency spectrum was sparse. They discovered that for such *favourable* signals, it was possible to perfectly reconstruct the signal even when sampling significantly below the Nyquist-rate. This resulted in the field of Compressed Sensing, where several advancements have reduced file-sizes for such signals. This field is in rapid development, and may also be used to remove "noise" from signals and images. For broadband signals, other methods must be applied in order to decrease file-size.

Chapter 10

Conclusion

The mathematical contributions of Joseph Fourier have truly ushered the development of many mathematical fields of study. We have seen that the structure provided by Hilbert spaces allow us to properly study convergence of Fourier series. Although it removes limit issues from the equation, there are other issues which may occur. The Gibb's phenomenon and slow convergence rates are examples of this. Depending on the situation, both the rate of convergence and types of convergence may have different implications and uses. Here we focused on sampling theory, yet there are many applications to Differential Equations, Wavelets, physics and much more.

As mentioned before, the Shannon-sampling theorem laid the foundations for digitization for modern communication and signal processing. We have only briefly touched the surface of the mathematical concepts and operations which have blossomed from these roots. Given the recent creation of Compressed Sensing, it is likely that other branches may evolve from sampling theory as well.

Given that advances in technology continues to achieve greater importance, it is reasonable to assume that further improvements in these mathematical topics will be of great significance in the coming years. A profound knowledge about the foundational theory may therefore be useful in forming new ideas and applications for further research.

Appendix A

Additional comments to proofs and results.

A.1 The Dirichlet Kernel

In the proof of representation for the Dirichlet Kernel we took advantage of the fact that almost all the terms in the summation cancel each other out. This is what we call a *telescopic* series, and only the terms marked by red remain.

$$\begin{aligned}
&k = 1 \Rightarrow \sin(t + \frac{t}{2}) - \sin(t - \frac{t}{2}) = \sin(\frac{3t}{2}) - \sin(\frac{t}{2}) \\
&k = 2 \Rightarrow \sin(\frac{5t}{2}) - \sin(\frac{3t}{2}) \\
&\quad \vdots \\
&k = N \Rightarrow \sin(N + \frac{t}{2}) - \sin(N - \frac{t}{2}).
\end{aligned}$$

A.2 Fourier series expansion of f_3

From Section 8.4 we investigated the function $f_3 = \begin{cases} t, & t \in [-\pi, \pi], \\ 0, & otherwise. \end{cases}$ To find its Fourier series expansion, we begin by noticing that f_3 is an odd function. Thereby the a_n terms of its trigonometric representation disappear in the expansion. Oppositely, we get that

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} t \sin nt \, dt = \frac{2\sin \pi n - 2\pi n \cos \pi n}{n^2 \pi} = \frac{-2(-1)^n}{n}.$$

Since $\sin \pi n = 0 \quad \forall n \in \mathbb{Z},$ and $\cos \pi n = (-1)^n \quad \forall n \in \mathbb{Z}$. As a result, the series expansion becomes

$$f \sim \sum_{k=1}^{\infty} b_k \sin kt = \sum_{k=1}^{\infty} \frac{-2(-1)^k}{n} \sin kt.$$

References

Avantaggiati, A., Loreti, P., & Vellucci, P. (2016). Kadec-1/4 Theorem for Sinc Bases. Retrieved from https://arxiv.org/abs/1603.08762

Bressoud, D. (2007). A Radical Approach to Real Analysis. In *The mathematical association of america.* Retrieved from https://www.macalester.edu/aratra/chapt2/chapt2{_}6b.html

Brunton, S. (2020). *Shannon Nyquist Sampling Theorem [Video clip].* Retrieved from https://www.youtube.com/watch?v=FcXZ28BX-xE

Campbell, N. A., Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Jackson, R. B. (2018). *Biology, A Global Approach. 11th edition.* Pearson Education Limited.

Carlson, S. C. (2006). *Hilbert space.* Britannica. Retrieved from https://www.britannica.com/science/Hilbert-space

Cheever, E. (2020). *The Fourier Series.* Retrieved 2021-03-16, from https://lpsa.swarthmore.edu/Fourier/Series/WhyFS.html

Christensen, O. (2010). *Functions, spaces, and expansions: Mathematical tools in physics and engineering* (1st ed.). Springer Science & Business Media.

Debnath, L. (2012). A short biography of Joseph Fourier and historical development of Fourier series and Fourier transforms. *International Journal of Mathematical Education in Science and Technology*, *43*(5), 589–612. Retrieved from https://www.tandfonline.com/doi/full/10.1080/0020739X.2011.633712

Dominguez, A. (2015). A History of the Convolution Operation [Retrospectroscope]. *IEEE Pulse*. Retrieved from https://www.embs.org/pulse/articles/history-convolution-operation/

Dominguez, A. (2016). Highlights in the History of the Fourier Transform [Retrospectroscope]. *IEEE Pulse*, *7*(1), 53–61. Retrieved from https://www.embs.org/pulse/articles/highlights-in-the-history-of-the-fourier-transform/

DSP Illustrations. (2020). *The Convolution Theorem with Application Examples.* Retrieved 2021-04-20, from https://dspillustrations.com/pages/posts/misc/the-convolution-theorem-and-application-examples.html

Egyptology: study of pharanoic Egypt. (2008). In *Encyclopedia Britannica*. Retrieved from https://www.britannica.com/science/Egyptology

Fourier series. (2012). In *Encyclopedia of Mathematics*. Retrieved from https://encyclopediaofmath.org/index.php?title=Fourier{_}series

Math24. (2005). *Convergence of Fourier Series.* Retrieved 2021-03-06, from https://www.math24.net/convergence-fourier-series

O'Connor, J. J., & Robertson, E. (1997). *Joseph Fourier (1768 - 1830).* Retrieved 2021-01-05, from https://mathshistory.st-andrews.ac.uk/Biographies/Fourier/

O'Connor, J. J., & Robertson, E. F. (2000). *Stefan Banach (1892 - 1945).* Retrieved 2021-02-02, from https://mathshistory.st-andrews.ac.uk/Biographies/Banach/

Serway, R., & Jewett, J. (2013). *Physics for Scientists and Engineers, Volume 1, International Edition.* (9th ed.). Physical Sciences: Mary Finch, Physics and Astronomy: Charlie Hartford.

Struik, D. J. (2020). Joseph Fourier. *Encyclopedia Britannica*. Retrieved from https://www.britannica.com/biography/Joseph-Baron-Fourier

Walker, J. S. (2003). Fourier Series. In R. A. Meyers (Ed.), *Encyclopedia of physical science and technology* (Third ed., pp. 167–183). Academic Press. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/B0122274105002581

Weisstein, E. W. (2003). *L^2-Norm.* From MathWorld–A Wolfram Web Resource. Retrieved 2021-03-05, from https://mathworld.wolfram.com/L2-Norm.html

Weisstein, E. W. (2005). *Hilbert Space. From MathWorld–A Wolfram Web Resource.* Retrieved 2021-03-05, from https://mathworld.wolfram.com/HilbertSpace.html

Weisstein, E. W. (2008). *Convolution.* From MathWorld–A Wolfram Web Resource. Retrieved 2021-03-11, from https://mathworld.wolfram.com/Convolution.html

Yankin, S. (2019). *The Origin Story of the Sampling Theorem and Vladimir Kotelnikov [Blog post].* Retrieved 2021-05-07, from https://www.comsol.com/blogs/the-origin-story-of-the-sampling-theorem-and-vladimir-kotelnikov/