REVISED

R

JOHN J. BENEDETTO
University of Maryland

Harmonic Analysis and Applications

CRC PRESS
Boca Raton New York London Tokyo

Contents

Preface

Among the various points of view in harmonic analysis, we shall define and emphasize the *analysis* and *synthesis* of functions in terms of *harmonics*. My *goal* is to present harmonic analysis at a level that exhibits its vitality, intricacy and simplicity, power, elegance, and usefulness. Despite an array of deep theoretical results and important abstract structures associated with harmonic analysis, I *believe* the subject lives because of its essential role in understanding a host of engineering, mathematical, and scientific ideas. Goals and beliefs are not always realities, but this book is my attempt to quantify the preceding aspiration and sentiment.

Harmonic Analysis and Applications is both a textbook and essay.

Textbook

The listing in Prologue I provides the material for an upper level undergraduate course in harmonic analysis and some of its applications. We refer to this course as Course I, and I have taught such a course for many years to engineering, physics, computer science, and mathematics students. The first 30 exercises in each chapter are appropriate for Course I. The exercises range from elementary to difficult and from theoretical to computational and/or computer oriented. Those that are not assigned might provide extracurricular and titillative reading, without the burden of proof (and grading).

Prologue II describes Fourier transforms for various settings such as the real line $\mathbb{R}$, the integers $\mathbb{Z}$, finite intervals, or finite sets of points. Hopefully, it will be useful and provide perspective.

Chapter 1 presents the Fourier analysis of integrable and square integrable (finite energy) functions on $\mathbb{R}$. Chapter 3 presents the Fourier analysis of finite and infinite sequences as well as of functions defined on finite intervals. Chapter 2 presents distribution theory. In my opinion, distribution theory provides a useful vantage point for dealing with many ideas from engineering, physics, and mathematics;

and, in my experience, the calculational presentation herein is assimilable at the undergraduate level.

Notwithstanding the importance of Chapter 2, I have sometimes adjusted Course I by going directly to the Course I material in Chapter 3 after finishing the corresponding material in Chapter 1. This has given me time to present elementary wavelet theory and various algorithms and computer exercises associated with wavelets. The blend of Fourier transforms and Fourier series from Chapters 1 and 3 is an ideal background for discussing wavelets.

Besides wavelet theory, the topics from Chapters 1 and 3 have also served as a natural background for presenting results from the following areas:

The theory of frames and their applications;
Regular and irregular sampling theory, and applications;
Uncertainty principle inequalities;
Fourier analysis and signal processing.

Essay

It is natural to seek the fundamental components of complex phenomena, and then to describe such phenomena in terms of these components. With regard to our opening sentence about the *analysis* and *synthesis* of functions in terms of *harmonics*, the *analysis* is the determination of the harmonics or *components* for a given function, and the *synthesis* is the reconstruction or *decomposition* of this function in terms of its components.

Historically, there are the *atoms* (components) and *atomic decompositions* of the ancient Greek philosophers as well as of physicists and modern harmonic analysts [FJW91]. There are Leibniz' ultimate components or elements of the universe, the so-called *monads* or individual centers of force; and there are Nobel Laureate Dennis Gabor's elementary quanta or components of information, the so-called *logons* [Gabo46]. There are the *basic primary forms* of cubists such as Gris, Braque, and Picasso, whose paintings were created (a veritable inspirational synthesis) in terms of simple, often geometrical, components, e.g., [Gold59]; and there is a comparable musical paradigm in terms of harmonics as originally discovered in antiquity [Pie83]. There is *Riemann's decomposition* of the number theoretic psi function in terms of certain elementary components, which are *nontrigonometric waveforms*; and this decomposition is used to investigate zeros of the Riemann zeta function [Bom92], cf., Example 2.4.6g and Exercise 2.32.

This is a book about *classical harmonic analysis*. The fundamental components are *trigonometric functions*; and we shall deal with these components and the corresponding spectral *trigonometric decompositions* inherent in Fourier's original approach [Zyg59]. Classical harmonic analysis arose naturally in eighteenth and nineteenth century mathematical physics, in studies about the propagation of heat and the decomposition of electromagnetic waves. In our own century, classical harmonic analysis has continued to flourish, with a magnificent array of applications and with profound theoretical developments including a fundamental role in wavelet theory [Dau92], [Mey90].

The general area of harmonic analysis has many other personalities and themes and levels of abstraction than that of classical harmonic analysis. Alas, we shall not deal with most of them in this book. For example, although harmonic analysis at the time of Fourier could be described in terms of the spectral theory of a second-order differential operator, that subject has modern-day group-invariant versions of which we have just said our last word. Similarly, we shall *not* deal with representation theory, Banach algebras, or locally compact groups G, even though these subjects have a significant relationship with harmonic analysis. On the other hand, our treatment of classical harmonic analysis does have a Banach algebraic flavor, and is a substantial part of the harmonic analysis of phenomena defined on G in the case G is commutative.

We shall present a systematic treatment of classical harmonic analysis, and give careful proofs of the basic theorems. A feature of our presentation is that we are also providing expositions and perspectives on many topics. Some of these are extensive, such as our treatment of Wiener's Generalized Harmonic Analysis in Section 2.9. Further, there are several lengthy historical comments, e.g., Section 3.2. Because of the proofs and perspectives and exercises, this book can also serve as a textbook for courses more advanced than Course I.

Our mathematical emphasis has been in the direction of real analysis with very little complex analysis; and our point of view has been to deal thoroughly with central spaces such as L^1 and L^2, as opposed to reporting on some of the important results concerning BMO or Triebel-Lizorkin spaces. We have not stressed the group-theoretical underpinnings of harmonic analysis (even in classical terms), since there are other fundamental (nonalgebraic) characteristics of classical harmonic analysis. Our applications are limited by the usual constraints: author prejudice and author limitation. This is compensated to some extent by a serious bibliography, referenced at appropriate junctures in the text. We have also introduced some applications, not only because of their importance, but because of their intrinsic relationship with theoretical developments, e.g., Sections 3.6 and 3.7.

Notation and Idiosyncracies

The excessive number of references to [Ben xx] should not lead a newcomer to false conclusions about this author's contributions. On the other hand, I can attest that these references will contain thorough bibliographies highlighting the real contributors and will serve to keep our own bibliography for this book more manageable.

The sections labeled, *x.y.z* Definition, will always define a term, and that term will be italicized. However, they may also contain some elementary calculations and examples, and some exposition about the term being defined.

We sometimes use the symbol "≡" to define notation. For example, we write "we first evaluate $a \equiv \int_0^\infty e^{-u^2}\,du$" so that we can deal with "a" instead of the more complicated right side in the ensuing calculation. Generally, we use "=" to define notation in displayed items or when the context is clear.

$\mathbb{C}$ is the set of complex numbers, $\mathbb{R}$ is the set of real numbers, $\mathbb{Q}$ is the set of rational numbers, $\mathbb{Z}$ is the set of integers, and $\mathbb{N}$ is the set of natural numbers, i.e., $\mathbb{N} = \{1, 2, 3, \ldots\}$.

Instead of denoting integration of a function f over $\mathbb{R}$ by $\int_{-\infty}^{\infty} f(t)dt$, we shall often write $\int f(t)dt$, i.e.,

$$\int f(t)dt = \int_{-\infty}^{\infty} f(t)dt.$$

$\varnothing$ designates the empty set. $\operatorname{Re} c$, resp., $\operatorname{Im} c$, is the real part, resp., imaginary part, of $c \in \mathbb{C}$. $\operatorname{sgn} t$ is 1 or -1 depending on whether $t > 0$ or $t < 0$. If $X \subseteq \mathbb{R}$, then $X^c = \mathbb{R}\backslash X$, the complement of X;. and $\mathbf{1}_X$ is the characteristic function of X equal to 1 if $t \in X$ and equal to 0 if $t \in X^c$. Other notation and notions are introduced as needed or in the Appendices, and they are referenced in the list of notation or in the index.

Acknowledgements

First, I wish to express my appreciation to my former graduate students, George Benke, Jean-Pierre Gabardo, Joseph Lakey, David Walnut, and Georg Zimmermann. They carefully read portions of the book and provided valuable comments and invaluable insights.

I have had the good fortune of friendship and learning from some of the modern masters in subjects ranging from spectral synthesis and Banach algebras to Wiener's Generalized Harmonic Analysis to signal processing. I shall not risk an incomplete list except to acknowledge Carl Herz, Stylianos Pichorides, and Harry Pollard, each of whom has passed away in recent years. I want to thank all of my graduate students who have shared some of the intellectual excitement of Fourier analysis with me. Also, my collaborations have been extremely rewarding, and it is a pleasure to express my gratitude to my coauthors, many of whom are referenced in the book.

There was specific assistance, on issues ranging from mathematics to birthplaces of mathematicians to computers to publishing, from Elmar Winkelnkemper, Larry Washington, C. Robert Warner, Sadahiro Saeki, Markus Melenk, Jean Mawhin, Nora Konopka, Raymond Johnson, Andrzej Hulanicki, Christian Houdré, Michael Hernandez, Hans Heinig, Christopher Heil, Melissa Harrison, Israel Gohberg, Ward Evans, David Colella, Christiane Carton-Lebrun, and Gary Benedetto. Finally, I had the extraordinary good fortune to have a great typist; my thanks to Stephanie Smith.

Prologue I

Course I

No proofs of theorems, propositions, or lemmas are required unless indicated.

Definition 1.1.1
Definition 1.1.2
Remark 1.1.3
Remark 1.1.4
Definition 1.1.5
Theorem 1.1.6
Remark 1.1.8
Proposition 1.1.9 and proof
Proposition 1.1.10 and proof
Proposition 1.1.11
Remark 1.1.13

Theorem 1.2.1
Example 1.2.2

Example 1.3.1
Example 1.3.2
Example 1.3.3
Example 1.3.4

Theorem 1.4.1
Remark 1.4.5

Definition 1.5.11
Proposition 1.5.2 and proof
Remark 1.5.3a

Definition 1.6.1
Proposition 1.6.2 and proof
Proposition 1.6.3 and proof
Proposition 1.6.4a
Example 1.6.5a
Example 1.6.6
Theorem 1.6.9 and proofs of parts b,c
Proposition 1.6.11

Remark 1.7.1
Theorem 1.7.6 (= Theorem 1.1.6)
Theorem 1.7.8
Example 1.7.10

Example 1.8.1
Example 1.8.2
Example 1.8.3

Remark 1.9.1
Remark 1.9.2
Remark 1.9.3

Definition 1.10.1
Theorem 1.10.2
Proposition 1.10.3 and proof

Proposition 1.10.4
Example 1.10.5
Remark 1.10.8
Proposition 1.10.9
Example 1.10.13
Example 1.10.14

Proposition 2.1.1 and proof
Remark 2.1.2
Definition 2.1.3
Proposition 2.1.4
Example 2.1.6

Definition 2.2.1
Example 2.2.2 a, b
Definition 2.2.3
Definition 2.2.5
Example 2.2.6
Definition 2.2.7

Definition 2.3.1
Remark 2.3.2
Example 2.3.3
Notation 2.3.4
Proposition 2.3.5 and proof
Example 2.3.8 a, b

Remark 2.4.1
Definition 2.4.2
Definition 2.4.3
Theorem 2.4.4
Definition 2.4.5a
Example 2.4.6 a,b,c,d
Definition 2.4.7a
Example 2.4.8
Example 2.4.9

Definition 2.5.1 a,b
Definition 2.5.3
Proposition 2.5.5 and proof
Definition 2.5.6
Definition 2.5.7a
Definition 2.5.10a
Theorem 2.5.11

Theorem 2.6.1

Theorem 2.10.1 and proof
Example 2.10.2
Theorem 2.10.3a
Example 2.10.5

Definition 3.1.1
Remark 3.1.2
Remark 3.1.3a,b,c
Example 3.1.4
Theorem 3.1.5
Theorem 3.1.6 and proof

Example 3.3.1
Theorem 3.3.2
Corollary 3.3.3
Example 3.3.4a
Definition 3.4.1a
Proposition 3.4.2
Definition 3.4.3
Theorem 3.4.4
Example 3.4.5
Theorem 3.4.6 and proof
Corollary 3.4.7
Remark 3.4.8a
Definition 3.4.10
Theorem 3.4.12
Theorem 3.4.13
Definition 3.5.1
Proposition 3.5.2

May 29, 1996 10:43 c36-pr-i Sheet number 3 Page number 17

Proposition 3.5.4
Example 3.5.8
Theorem 3.5.9

Definition 3.6.2a,b
Theorem 3.6.3
Theorem 3.6.4
Theorem 3.6.6
Example 3.6.7

Definition 3.8.1a,b,c,e
Theorem 3.8.2 and proof
Theorem 3.8.4 and proof
Example 3.8.5a,b,c
Theorem 3.9.1
Example 3.9.2
Example 3.9.3

Definition 3.10.1a,b,c
Example 3.10.2a,b
Definition 3.10.3a
Example 3.10.4a
Theorem 3.10.5
Theorem 3.10.6
Definition 3.10.7
Theorem 3.10.8
Theorem 3.10.10
Definition 3.10.11
Remark 3.10.12

R

Prologue II

Fourier Transforms, Fourier Series, and Discrete Fourier Transforms

Domain of the function f	Definition of the Fourier transform $\hat{f} = F$	Domain of the Fourier transform F
$\mathbb{R}$	$F(\gamma) = \hat{f}(\gamma) = \int f(t)e^{-2\pi it\gamma}\,dt$	$\hat{\mathbb{R}} = \mathbb{R}$
	Inversion Formula: $f(t) = \int F(\gamma)e^{2\pi it\gamma}\,d\gamma$	
	Notes. • Integration is over $\mathbb{R}$. • The domain of the Fourier transform is denoted by $\hat{\mathbb{R}}$.	
$\mathbb{Z}$	$F(\gamma) = \sum f[n]e^{-\pi in\gamma/\Omega}$	$\mathbb{T}_{2\Omega} = \hat{\mathbb{R}}/2\Omega\mathbb{Z}$
	Inversion Formula: $f[n] = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} F(\gamma)e^{\pi in\gamma/\Omega}\,d\gamma$	
	Notes. • $\Omega > 0$ is fixed and summation is over $\mathbb{Z}$. • F is a *Fourier series* with *Fourier coefficients* $f = \{f[n] : n \in \mathbb{Z}\}$. • F is 2Ω-periodic on $\hat{\mathbb{R}}$, and $\mathbb{T}_{2\Omega} = \hat{\mathbb{R}}/2\Omega\mathbb{Z}$ denotes this domain.	
$\mathbb{T}_{2\Omega}$	$F[n] = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} f(\gamma)e^{-\pi in\gamma/\Omega}\,d\gamma$	$\mathbb{Z}$
	Inversion Formula: $f(\gamma) = \sum F[n]e^{\pi in\gamma/\Omega}$	
$\mathbb{Z}_N$	$F[n] = \sum_{m=0}^{N-1} f[m]e^{-2\pi imn/N}$	$\mathbb{Z}_N$
	Inversion Formula: $f[m] = \frac{1}{N}\sum_{n=0}^{N-1} F[n]e^{2\pi imn/N}$	
	Notes. • F is a *Discrete Fourier Transform* (DFT). • F is N-periodic on $\mathbb{Z}$, and $\mathbb{Z}_N$ denotes this domain.	

XX is blank

Chapter 1

Fourier Transforms

1.1 Definitions and formal calculations

1.1.1 Definition. Integrable Functions

Set

$$L^1_{\text{loc}}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{C} : \forall a < b, \quad \int_a^b |f(t)|\,dt < \infty\}$$

and

$$L^1(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{C} : \|f\|_{L^1(\mathbb{R})} = \int_{-\infty}^{\infty} |f(t)|\,dt < \infty\}.$$

$L^1_{\text{loc}}(\mathbb{R})$ is the space of *locally integrable functions* on $\mathbb{R}$ and $L^1(\mathbb{R})$ is the space of *integrable functions* on $\mathbb{R}$.

Let $g \in L^1_{\text{loc}}(\mathbb{R})$. $\int_{-\infty}^{\infty} g(t)\,dt$, which we designate frequently by $\int g(t)\,dt$, is

$$\lim_{S\to\infty, T\to\infty} \int_{-S}^{T} g(t)\,dt.$$

The *Cauchy principal value*, $pv \int g(t)\,dt$, is

$$\lim_{T\to\infty} \int_{-T}^{T} g(t)\,dt.$$

If $\int g(t)\,dt$ exists then $pv \int g(t)\,dt$ exists and the two integrals have the same value. In the opposite direction, $pv \int g(t)\,dt$ may exist while

$\int g(t)\,dt$ does not, e.g., $g(t) = t$. On the other hand, if $pv \int g(t)\,dt$ exists and either g is *even*, i.e., $g(t) = g(-t)$, or $g \geq 0$ on $\mathbb{R}$ then

$$pv \int g(t)\,dt = \int g(t)\,dt,$$

e.g., *Exercise 1.1.*

We have been purposely vague about the definition of $\int_a^b g(t)\,dt$. If you know the Lebesgue integral, then fine! If not, the Riemann integral works for most of our calculations and most of the functions we shall consider.

1.1.2 Definition. FOURIER TRANSFORM

The *Fourier transform* of $f \in L^1(\mathbb{R})$ is the function F defined as

$$F(\gamma) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i t\gamma}dt, \quad \gamma \in \widehat{\mathbb{R}}\,(= \mathbb{R}),$$

cf., *Exercise 1.15.* Notationally, we write the pairing between the function f and F in one of the following ways:

$\check{F}$ or $F^\vee$

$$f \longleftrightarrow F, \quad \hat{f} = F, \quad f = \check{F},$$

The space of Fourier transforms of L^1-functions is denoted by $A(\widehat{\mathbb{R}})$, i.e.,

$$A(\widehat{\mathbb{R}}) = \{F : \widehat{\mathbb{R}} \to \mathbb{C} : \exists f \in L^1(\mathbb{R}) \quad \text{such that} \quad \hat{f} = F\}.$$

1.1.3 Remark. INVERSION FORMULA

Let $f \in L^1(\mathbb{R})$ and let $\hat{f} = F$. The *Fourier transform inversion formula* is

$$f(t) = \int F(\gamma)e^{2\pi i t\gamma}\,d\gamma. \tag{1.1.1}$$

We shall prove (1.1.1) in *Section 1.7*, but let us formally derive it now.

1.1.4 Remark. FORMAL CALCULATION OF THE INVERSION FORMULA

The Dirac δ "function" (actually, it is a probability measure) can be thought of in terms of the "formula"

$$(\delta) \qquad \forall f, \quad f(t) = \int f(u)\,\delta(t-u)\,du;$$

and a model of the uncertainty principle is the "formula"

$$(\text{UP}) \qquad \delta(t) = \int e^{2\pi i t\gamma}\,d\gamma.$$

If the Fourier pair $f \leftrightarrow F$ is given, then (δ) and (UP) allow us to make the following formal calculation:

$$\begin{aligned} \int F(\gamma)e^{2\pi i t\gamma}\,d\gamma &= \iint f(u)e^{-2\pi i u\gamma}e^{2\pi i t\gamma}\,du d\gamma \\ &= \iint f(u)e^{2\pi i(t-u)\gamma}\,d\gamma du \\ &= \int f(u)\,\delta(t-u)\,du = f(t). \end{aligned} \tag{1.1.2}$$

(δ) and (UP) are nonsense as they stand, but do have a sense intuitively. (δ) is easily motivated, e.g., *Section 1.6* and *Section 2.1.* The rationale for (UP) involves the fact, $f'(t) \leftrightarrow (2\pi i\gamma)F(\gamma)$ (which we shall verify shortly), which in turn allows us to use $\widehat{\mathbb{R}}$ as the domain of the momentum and thereby to invoke the usual interpretation from physics of the uncertainty principle, e.g., [vN55, Chapter III.4, esp. page 235], [Wey50a, page 77 and Appendix I]. Mathematically, (UP) can be given a precise meaning in terms of the notion of "oscillatory integral" [Hör83, Volume I, Section 7.8, especially (7.8.5)].

In a more elementary, but also fundamental way, we can think of (UP) in terms of the following idealized piano experiment. The standard for concert pitch is that the A above middle C should have 440 vibrations per second. Thus, the A four octaves down (and the last key on the piano) should have 27.5 vibrations per second. Suppose we could strike this last key for a time interval of 1/30 seconds, i.e., the hammer strikes the string and 1/30 seconds later the damper returns to the string, thereby stopping the sound. In particular, a complete vibration for this key does not occur. We then have very precise time information, represented by $\delta(t)$ on the ~~right~~ left side of (UP), but correspondingly imprecise frequency information since the emitted sound is anything but the desired pure periodic pitch of this low A. In fact, this imprecise frequency information can be thought of as the "noisy sum" of many pure tones, represented by the integral on the right side of (UP).

Even assuming the validity of (something like) (δ) and (UP), the calculation (1.1.2) leaves something to be desired because of the casual switching of order of integration we have made, cf., *Section 1.7*.

1.1.5 Definition. BOUNDED VARIATION

A function $f : \mathbb{R} \to \mathbb{C}$ has *bounded variation* on an interval $I \subseteq \mathbb{R}$ if there is a constant M such that for every finite set $t_0 < t_1 < \cdots < t_n$, $t_j \in I$, we have

$$\sum_{j=1}^{n} |f(t_j) - f(t_{j-1})| \leq M.$$

In this case we write $f \in BV(I)$, and note that we could have $I = \mathbb{R}$. If $f : \mathbb{R} \to \mathbb{C}$ has bounded variation on each interval of finite length then f *locally has bounded variation*, and we write $f \in BV_{\text{loc}}(\mathbb{R})$.

Functions having bounded variation on bounded intervals I have graphs of finite length; and such functions are a natural generalization of continuously differentiable functions, e.g., [Ben76, Chapter 4].

One form of the inversion formula that we shall verify in *Section 1.7* is the *Jordan pointwise inversion formula.*

1.1.6 Theorem. JORDAN THEOREM

Let $f \in L^1(\mathbb{R})$ *and assume* $f \in BV[t-\epsilon, t+\epsilon]$ *for some* $t \in \mathbb{R}$ *and* $\epsilon > 0$. *Then*

$$\frac{f(t+) + f(t-)}{2} = \lim_{\Omega \to \infty} \int_{-\Omega}^{\Omega} F(\gamma) e^{2\pi i t \gamma} \, d\gamma, \tag{1.1.3}$$

where $f \leftrightarrow F$. *If* f *is continuous at* t *then the left side of (1.1.3) can be replaced by* $f(t)$.

Another important pointwise inversion formula, but one which we shall not verify, is the following result due to Pringsheim, e.g., [RL55].

1.1.7 Theorem. PRINGSHEIM THEOREM

Let $f \in BV(\mathbb{R})$ *and assume*

$$\lim_{|t|\to\infty} f(t) = 0.$$

Then

$$F(\gamma) = \int f(t)e^{-2\pi i t\gamma}\, dt$$

exists for all γ *except possibly* $\gamma = 0$, *and*

$$\forall t \in \mathbb{R}, \quad \frac{f(t+) + f(t-)}{2} = \lim_{\Omega\to\infty, \epsilon\to 0} \int_{\epsilon<|\gamma|<\Omega} F(\gamma)e^{2\pi i t\gamma}\, d\gamma.$$

For the remainder of *Section 1.1* our statements and proofs are formal, in the sense that we have not been concerned with correct mathematical hypotheses for asserting the existence of the Fourier transform or using the inversion formula.

1.1.8 Remark. FORMAL CALCULATIONS FOR $f \leftrightarrow F$

Consider the formal pairing $f \leftrightarrow F$, where $f = f_1 + if_2$ and $F = F_1 + iF_2$. Then, ~~using (1.1.1)~~ by the definition of F, we obtain

$$F_1(\gamma) + iF_2(\gamma) = \int (f_1(t) + if_2(t))(\cos 2\pi t\gamma - i \sin 2\pi t\gamma)\, dt,$$

and so

$$F_1(\gamma) = \int (f_1(t) \cos 2\pi t\gamma + f_2(t) \sin 2\pi t\gamma)\, dt$$

and

$$F_2(\gamma) = \int (f_2(t) \cos 2\pi t\gamma - f_1(t) \sin 2\pi t\gamma)\, dt.$$

Similarly, ~~we have~~ by (1.1.1), we have

$$f_1(t) + if_2(t) = \int (F_1(\gamma) + iF_2(\gamma))(\cos 2\pi t\gamma + i \sin 2\pi t\gamma)\, d\gamma$$

and, hence,

$$f_1(t) = \int (F_1(\gamma) \cos 2\pi t\gamma - F_2(\gamma) \sin 2\pi t\gamma)\, d\gamma$$

and

$$f_2(t) = \int (F_2(\gamma) \cos 2\pi t\gamma + F_1(\gamma) \sin 2\pi t\gamma)\, d\gamma.$$

1.1.9 Proposition.

Consider the formal pairing $f \leftrightarrow F$. f is real and even if and only if F is real and even. In this case

$$F(\gamma) = 2\int_0^\infty f(t)\cos 2\pi t\gamma\, dt$$

and

$$f(t) = 2\int_0^\infty F(\gamma)\cos 2\pi t\gamma\, d\gamma.$$

Proof. Suppose f is real and even. Then $f_2 = 0$ and $f = f_1$; and we apply the calculations of *Remark 1.1.8* to obtain

$$F(\gamma) = \int f(t)(\cos 2\pi t\gamma - i\sin 2\pi t\gamma)\, dt$$
$$= \int f(t)\cos 2\pi t\gamma\, dt = 2\int_0^\infty f(t)\cos 2\pi t\gamma\, dt.$$

In particular, F is real and even. A similar calculation works for the opposite direction. □

1.1.10 Proposition.

Consider the formal pairing $f \leftrightarrow F$. f is real if and only if $F(\gamma) = \overline{F(-\gamma)}$.

Proof. If f is real then $f = f_1$; and we apply the calculations of *Remark 1.1.8* to obtain

$$F_1(\gamma) = \int f_1(t)\cos 2\pi t\gamma\, dt$$

and

$$F_2(\gamma) = -\int f_1(t)\sin 2\pi t\gamma\, dt.$$

Thus, we have $F_1(-\gamma) = F_1(\gamma)$ and $F_2(-\gamma) = -F_2(\gamma)$, and, hence,

$$\overline{F(-\gamma)} = \overline{F_1(-\gamma) + iF_2(-\gamma)} = F_1(-\gamma) - iF_2(-\gamma)$$
$$= F_1(\gamma) + iF_2(\gamma) = F(\gamma).$$

For the converse, suppose $\overline{F(\gamma)} = F(-\gamma)$. Therefore,

$$F_1(\gamma) - iF_2(\gamma) = F_1(-\gamma) + iF_2(-\gamma),$$

i.e., $F_1(\gamma) = F_1(-\gamma)$ and $F_2(\gamma) = -F_2(-\gamma)$; hence, F_1 is even and F_2 is odd. Using these facts and the calculations of *Remark 1.1.8*, we calculate

$$f_2(t) = \int (F_1(\gamma) \sin 2\pi t\gamma + F_2(\gamma) \cos 2\pi t\gamma)\, dt = 0;$$

from which we conclude that f is real. □

1.1.11 Proposition.

Consider the formal pairing $f \leftrightarrow F$ and write

$$F(\gamma) = A(\gamma)e^{i\phi(\gamma)},$$

where $A(\gamma) \geq 0$ and $\phi(\gamma) \in \mathbb{R}$. If f is real then

$$f(t) = \int A(\gamma) \cos(2\pi t\gamma + \phi(\gamma))\, d\gamma.$$

Proof. We calculate

$$\cos(2\pi t\gamma + \phi(\gamma)) = \cos 2\pi t\gamma \cos \phi(\gamma) - \sin 2\pi t\gamma \sin \phi(\gamma)$$

and so

$$\begin{aligned} &A(\gamma) \cos(2\pi t\gamma + \phi(\gamma)) \\ &= A(\gamma) \cos \phi(\gamma) \cos 2\pi t\gamma - A(\gamma) \sin \phi(\gamma) \sin 2\pi t\gamma \\ &= F_1(\gamma) \cos 2\pi t\gamma - F_2(\gamma) \sin 2\pi t\gamma \end{aligned}$$

since

$$F(\gamma) = F_1(\gamma) + iF_2(\gamma) = A(\gamma)(\cos \phi(\gamma) + i \sin \phi(\gamma)).$$

Using this fact and the hypothesis that f is real, we see from the calculations of *Remark 1.1.8* that

$$f(t) = \int A(\gamma) \cos(2\pi t\gamma + \phi(\gamma))\, d\gamma.$$

□

1.1.12 Remark. AMPLITUDE AND PHASE

Consider the formal pairing $f \leftrightarrow F$ and write $F(\gamma) = A(\gamma)e^{i\phi(\gamma)}$. $A(\gamma)$ is the *amplitude* and $\phi(\gamma)$ is the *phase angle* of $F(\gamma)$.

The inversion formula (1.1.1) allows us to think of a signal f as a "sum" (integral)

$$\text{“}f(t) = \sum_{\gamma} (A(\gamma)e^{i\varphi(\gamma)})\, e^{2\pi i t\gamma}\text{”}$$

of exponentials $e^{2\pi i t\gamma}$ with complex coefficients $A(\gamma)\, e^{i\phi(\gamma)}$. Different phase angles φ can produce quite different looking signals f even if $A(\gamma)$ remains the same. As an elementary example, note that if $\varphi(\gamma)$ is replaced by $\varphi(\gamma) - 2\pi u\gamma$ for some fixed $u \in \mathbb{R}$, then $f(t)$ is replaced by the translate $f(t-u)$, cf., *Theorem 1.2.1d*.

The amplitude squared $|A(\gamma)|^2$ can be thought of as the amount of energy of f in the frequency band about a small neighborhood of γ; and $|A|^2$ is often a measurable quantity in signal processing, spectroscopy, fluid mechanics, etc. The physical measurement is based on ideas about correlations and translations of f, e.g., *Exercise 1.33*, *Example 2.8.10*, and Michelson's invention of the interferometer [Loe66], [Mic62]. The resulting 0-phase information or some windowed form of it is called the *spectrogram or periodogram* or *power spectrum* of f, e.g., *Section 2.8* and *Definition 2.9.5*. A basic methodology in engineering and the sciences is to approximate a reconstruction of f from its spectrogram. This is a first step in harmonic analysis signal reconstruction technology.

or (roman)

1.1.13 Remark. A TABLE

Consider the formal pairing $f \leftrightarrow F$, where $F = F_1 + iF_2$.

$$\begin{aligned} f \text{ real if and only if } \overline{F(\gamma)} &= F(-\gamma), \\ F(\gamma) &= \int f(t)\cos 2\pi t\gamma\, dt - i\int f(t)\sin 2\pi t\gamma\, dt, \\ f(t) &= 2\int_0^\infty (F_1(\gamma)\cos 2\pi t\gamma - F_2(\gamma)\sin 2\pi t\gamma)\, d\gamma, \\ &= 2Re\int_0^\infty F(\gamma)e^{2\pi i t\gamma} d\gamma. \end{aligned}$$

f real and even if and only if F real and even,

$$F(\gamma) = 2\int_0^\infty f(t)\cos 2\pi t\gamma\, dt,$$
$$f(t) = 2\int_0^\infty F(\gamma)\cos 2\pi t\gamma\, d\gamma.$$

f real and odd if and only if F odd and F imaginary,

$$F(\gamma) = -2i\int_0^\infty f(t)\sin 2\pi t\gamma\, dt,$$
$$f(t) = 2i\int_0^\infty F(\gamma)\sin 2\pi t\gamma\, d\gamma.$$

f imaginary if and only if $\overline{F(\gamma)} = -F(-\gamma)$,

e.g., *Exercise 1.2.*

In light of *Remark 1.1.13* we ask:

Question. Can we characterize the case when f is not only real but is also non-negative?

Answer. Yes, but the answer is not simple, and involves the notion of *positive definite functions* which play a key role in theoretical considerations associated with spectral estimation, e.g., *Example 2.7.9b*, *Theorem 2.7.10*, and *Sections 3.6* and *3.7.*

1.2 Algebraic properties of Fourier transforms

Notationally, for a fixed γ, we set

$$e_\gamma(t) = e^{2\pi it\gamma};$$

and, for a fixed u and a given function f, we set

$$(\tau_u f)(t) = f(t-u).$$

$\tau_u f$ is the *translation* of f by u.

1.2.1 Theorem. ALGEBRAIC PROPERTIES OF FOURIER TRANSFORMS

a. Linearity. Consider $f_j \leftrightarrow F_j$ *and let* $c_j \in \mathbb{C}$, $j = 1, 2$, *where* $f_j \in L^1(\mathbb{R})$. *Then* $(c_1 f_1 + c_2 f_2)^\wedge(\gamma) = (c_1 \hat{f}_1 + c_2 \hat{f}_2)(\gamma)$, *i.e.,*

$$c_1 f_1 + c_2 f_2 \longleftrightarrow c_1 F_1 + c_2 F_2.$$

b. Symmetry. Consider $f \leftrightarrow F$, *where* $f \in L^1(\mathbb{R})$ *and* $F \in L^1(\widehat{\mathbb{R}})$. *Then* $\widehat{F}(\gamma) = f(-\gamma)$, *i.e.,*

$$F(t) \longleftrightarrow f(-\gamma).$$

c. Conjugation. Consider $f \leftrightarrow F$, *where* $f \in L^1(\mathbb{R})$. *Then* $(\overline{f})^\wedge(\gamma) = \overline{\hat{f}(-\gamma)}$, *i.e.,*

$$\overline{f(t)} \longleftrightarrow \overline{F(-\gamma)},$$

cf., Proposition 1.1.10.

d. Translation (time shifting). Consider $f \leftrightarrow F$, *where* $f \in L^1(\mathbb{R})$, *and take* $u \in \mathbb{R}$. *Then* $(\tau_u f)^\wedge(\gamma) = e^{-2\pi i u \gamma} \hat{f}(\gamma)$, *i.e.,*

$$(\tau_u f)(t) \longleftrightarrow e_{-u}(\gamma)\, F(\gamma).$$

e. Modulation (frequency shifting). Consider $f \leftrightarrow F$, *where* $f \in L^1(\mathbb{R})$, *and take* $\lambda \in \widehat{\mathbb{R}}$. *Then* $(e^{2\pi i t \lambda} f(t))^\wedge(\gamma) = \hat{f}(\gamma - \lambda)$, *i.e.,*

$$e_\lambda(t) f(t) \longleftrightarrow F(\gamma - \lambda) = \tau_\lambda F(\gamma).$$

f. Time dilation (time scaling). Consider $f \leftrightarrow F$, *where* $f \in L^1(\mathbb{R})$, *and take* $\lambda \in \mathbb{R}\backslash\{0\}$. *Define the* λ-dilation *of* f *by*

$$f_\lambda(t) = \lambda f(\lambda t).$$

Then $\hat{f}_\lambda(\gamma) = \frac{\lambda}{|\lambda|} \hat{f}(\frac{\gamma}{\lambda})$, *i.e.,*

$$f_\lambda(t) \longleftrightarrow \frac{\lambda}{|\lambda|} F(\frac{\gamma}{\lambda}).$$

The formal proof of this result is easy, but we should comment on the hypotheses of part *b*. It turns out that the integrability of f and F is sufficient for the validity of the pointwise inversion formula (1.1.1) which is used in the proof of part *b*. The verification of this sufficiency requires some work, and we shall deal with it in *Theorem 1.7.8.*

1.2.2 Example. DILATION AND MODULATION

a. *Dilation.* Let $\mathbf{1}_{[-T,T)}$ be the *rectangular pulse* on $\mathbb{R}$ defined as

$$\mathbf{1}_{[-T,T)}(t) = \begin{cases} 1, & -T \leq t < T, \\ 0, & \text{otherwise.} \end{cases}$$

If $T = 1$ we write $\mathbf{1} = \mathbf{1}_{[-1,1)}$. Now define the function f[redacted] $=$ $h\mathbf{1}_{[-T,T)}$[redacted], $h > 0$. To fix ideas, take $\lambda \in (0,1)$. Then the graphs of f and f_λ are —

Figure 1.1

b. *Modulation.* Consider a *carrier wave* $\cos 2\pi t\gamma_0$, where $\gamma_0 > 0$ is the *carrier frequency*. If $f \in L^1(\mathbb{R})$ then $f(t)\cos 2\pi t\gamma_0$ is the resulting *modulated signal.*

Figure 1.2

We shall compute the Fourier transform of $f(t)\cos 2\pi t\gamma_0$ in terms of $\widehat{f}(\gamma) = F(\gamma) = A(\gamma)e^{i\phi(\gamma)}$:

$$f(t)\cos 2\pi t\gamma_0 = f(t)\frac{e^{2\pi i t\gamma_0} + e^{-2\pi i t\gamma_0}}{2}$$

$$\longleftrightarrow \frac{1}{2}(F(\gamma-\gamma_0) + F(\gamma+\gamma_0)) = A_c(\gamma)e^{i\phi_c(\gamma)},$$

where $A_c(\gamma)$ is the amplitude of $\frac{1}{2}(F(\gamma-\gamma_0) + F(\gamma+\gamma_0))$ and $\phi_c(\gamma)$ is its phase angle.

Suppose $f \in L^1(\mathbb{R})$ has the property that $\widehat{f}$ vanishes off the interval $[-\Omega,\Omega]$. In this case, we say that f is *Ω-bandlimited*, and we use the notation $\Omega{-}BL$ to designate this property of a function. To fix ideas, consider the amplitude A as in *Figure 1.3.*

Figure 1.3

If $\gamma_0 \geq \Omega$, A_c has the following graph, indicating that A splits into two parts, each having half the amplitude of the original.

Figure 1.4

If $\gamma_0 < \Omega$, then the two lobes in *Figure 1.4* overlap, leading to *aliasing* problems, e.g., *Example 1.9.4* and *Remark 3.10.12.*

1.3 Examples

1.3.1 Example. THE SINC OR DIRICHLET FUNCTION

Let $f(t) = \mathbf{1}_{[-T,T)}(t)$. Clearly, we have $\widehat{\mathbf{1}}_{[-T,T)}(\gamma) = \frac{\sin 2\pi T\gamma}{\pi\gamma}$. Notationally, we write

$$d(\gamma) = \frac{\sin\gamma}{\pi\gamma}$$

and

$$\operatorname{sinc}\gamma = \frac{\sin\pi\gamma}{\pi\gamma},$$

so that $\widehat{\mathbf{1}}_{[-T,T)}(\gamma) = d_{2\pi T}(\gamma)$. If $\lambda > 0$ then $\widehat{\mathbf{1}}_{[-\frac{\lambda}{2\pi},\frac{\lambda}{2\pi}]} = d_\lambda$. We refer to d as the *Dirichlet function*, and shall see that $\int d(\gamma)\,d\gamma = 1$ (*Proposition 1.6.3*), noting that $d \notin L^1(\widehat{\mathbb{R}})$ (*Exercise 1.6*).

Figure 1.5

1.3.2 Example. THE POISSON FUNCTION

Let $f(t) = e^{-2\pi r|t|}$, $r > 0$. To compute $\widehat{f}$ we calculate

$$\int_{-T}^{T} e^{-2\pi r|t|}\, e^{-2\pi it\gamma}\,dt$$

$$= \int_{-T}^{0} e^{2\pi rt}e^{-2\pi it\gamma}\,dt + \int_{0}^{T} e^{-2\pi rt}e^{-2\pi it\gamma}\,dt$$

$$= \frac{1}{2\pi(r-i\gamma)}(1 - e^{-2\pi T(r-i\gamma)}) - \frac{1}{2\pi(r+i\gamma)}(e^{-2\pi T(r+i\gamma)} - 1)$$

$$= \frac{1}{2\pi}(\frac{1}{r-i\gamma} + \frac{1}{r+i\gamma}) + \epsilon(T).$$

Clearly, $\lim_{T\to\infty} \epsilon(T) = 0$ since $r > 0$; and, hence,

$$\hat{f}(\gamma) = \frac{1}{\pi}\frac{r}{r^2+\gamma^2}.$$

We write

$$p(\gamma) = \frac{1}{\pi(1+\gamma^2)}$$

so that $p_{\frac{1}{r}}(\gamma) = \hat{f}(\gamma)$. Thus, if $\lambda = 1/r > 0$ then

$$e^{-2\pi|t|/\lambda} \longleftrightarrow p_\lambda(\gamma).$$

We refer to p as the *Poisson function*, and observe that $\int p(\gamma)\,d\gamma = 1$, e.g., *Example 1.6.5*.

Note that the exponential decrease of f is transformed into the polynomial decrease of $\hat{f}$, cf., *Exercise 1.5* where the point of nondifferentiability in this example is replaced by a discontinuity.

1.3.3 Example. THE GAUSSIAN

Let $f(t) = e^{-\pi r t^2}$, $r > 0$. We could calculate $\hat{f} = F$ by means of contour integrals, but we choose a real, and by now classical approach [Fel66, page 476]. By definition of F, which is real and even, we have

$$F'(\gamma) = -2\pi i \int t e^{-\pi r t^2} e^{-2\pi i t\gamma}\,dt. \tag{1.3.1}$$

Noting that

$$\frac{d}{dt}(e^{-\pi r t^2}) = -2\pi r t\, e^{-\pi r t^2},$$

we rewrite (1.3.1) as

$$\begin{aligned} F'(\gamma) &= -2\pi i \int \frac{-1}{2\pi r}(e^{-\pi r t^2})' e^{-2\pi i t\gamma}\,dt \\ &= \frac{i}{r}\left[e^{-\pi r t^2} e^{-2\pi i t\gamma}\Big|_{-\infty}^{\infty} - \int e^{-\pi r t^2}(-2\pi i\gamma) e^{-2\pi i t\gamma}\,dt\right] \\ &= \frac{-2\pi\gamma}{r} F(\gamma). \end{aligned}$$

Thus, F is a solution of the differential equation,

$$F'(\gamma) = -\frac{2\pi\gamma}{r}F(\gamma); \tag{1.3.2}$$

and (1.3.2) is solved by elementary means with solution

$$F(\gamma) = Ce^{-\pi\gamma^2/r},$$

e.g., *Exercise 1.9.*

Taking $\gamma = 0$ and using the definition of the Fourier transform, we see that

$$C = \int e^{-\pi r t^2}\, dt.$$

In order to calculate C we first evaluate $a \equiv \int_0^\infty e^{-u^2}\, du$.

$$\begin{aligned} a^2 &= \int_0^\infty e^{-x^2}\, dx \int_0^\infty e^{-y^2}\, dy \\ &= \int_0^\infty \int_0^\infty e^{-(x^2+y^2)}\, dxdy = \int_0^{\pi/2} \int_0^\infty e^{-r^2}\, r dr d\theta \\ &= \frac{\pi}{4} \int_0^\infty e^{-u}\, du = \frac{\pi}{4}. \end{aligned}$$

Thus, $a = \frac{\sqrt{\pi}}{2}$ and so

$$\int e^{-u^2}\, du = \sqrt{\pi}.$$

Consequently,

$$C = \int e^{-\pi r t^2}\, dt = \frac{1}{\sqrt{\pi r}} \int e^{-u^2}\, du = \frac{1}{\sqrt{r}}.$$

Therefore, we have shown that

$$e^{-\pi r t^2} \longleftrightarrow \frac{1}{\sqrt{r}} e^{-\pi\gamma^2/r}.$$

We write

$$g(t) = \frac{1}{\sqrt{\pi}} e^{-t^2}$$

so that if $\lambda > 0$ then

$$g_\lambda(t) \longleftrightarrow e^{-(\pi\gamma/\lambda)^2}.$$

In particular,

$$\frac{1}{\sqrt{r}} g_{\sqrt{\pi r}} \longleftrightarrow g_{\sqrt{\pi / r}}$$

and hence $g_{\sqrt{\pi}} \longleftrightarrow g_{\sqrt{\pi}}$. We refer to g as the *Gauss function* or *Gaussian*, and note that $\int g(\gamma)\, d\gamma = 1$.

1.3.4 Example. THE FEJÉR FUNCTION

Let $f(t) = \max(1 - |t|, 0)$. On $[-1, 1]$, the graph of f consists of the equal legs of an isosceles triangle of height 1; f vanishes outside $[-1,1]$, e.g., *Figure 1.6*. The Fourier transform of f is

$$F(\gamma) = \int_0^1 (1-t) e^{-2\pi i t \gamma}\, dt + \int_{-1}^0 (1+t) e^{-2\pi i t \gamma}\, dt$$

$$= \left[\frac{1}{2\pi i \gamma} + \frac{1}{(2\pi i \gamma)^2}(e^{-2\pi i \gamma} - 1)\right] + \left[-\frac{1}{2\pi i \gamma} - \frac{1}{(2\pi i \gamma)^2}(1 - e^{2\pi i \gamma})\right]$$

$$= \frac{-2 + 2\cos 2\pi\gamma}{(2\pi i \gamma)^2} = \frac{2(1 - \cos 2\pi\gamma)}{(2\pi\gamma)^2}$$

$$= \frac{2(1 - [\cos \pi\gamma \cos \pi\gamma - \sin \pi\gamma \sin \pi\gamma])}{(2\pi\gamma)^2} = \frac{\sin^2 \pi\gamma}{(\pi\gamma)^2},$$

i.e.,

$$\hat{f}(\gamma) = \left(\frac{\sin \pi\gamma}{\pi\gamma}\right)^2.$$

Notationally, we set $\Delta = f$ and

$$w(\gamma) = \frac{1}{2\pi}\left(\frac{\sin \gamma/2}{\gamma/2}\right)^2,$$

so that $w_{2\pi}(\gamma) = \hat{\Delta}(\gamma)$, i.e.,

$$\Delta \longleftrightarrow w_{2\pi}. \tag{1.3.3}$$

Thus, if $\lambda > 0$ then

$$\max\left(1 - \frac{|2\pi t|}{\lambda}, 0\right) = \frac{\lambda}{2\pi} \Delta_{2\pi/\lambda}(t) \longleftrightarrow \omega_\lambda(\gamma). \tag{1.3.4}$$

We refer to w as the *Fejér function*, and note that

$$\int w(\gamma)\, d\gamma = 1$$

(*Proposition 1.6.3*). Fejér's surname at birth was Weiss.

Besides the integrability of w_λ, a key difference between d_λ and w_λ is the fact that $w_\lambda \geq 0$.

Figure 1.6

1.4 Analytic properties of Fourier transforms

1.4.1 Theorem. Analytic Properties of Fourier Transforms
Let $f \in L^1(\mathbb{R})$, $f \longleftrightarrow F$.

a. Boundedness. For each $\gamma \in \widehat{\mathbb{R}}$, $|F(\gamma)| \leq \|f\|_{L^1(\mathbb{R})}$.

b. Continuity. F is uniformly continuous on $\widehat{\mathbb{R}}$, i.e., for all $\epsilon > 0$, there is $\delta > 0$ such that for each γ and each λ for which $|\lambda| < \delta$, we have $|F(\gamma + \lambda) - F(\gamma)| < \epsilon$. In particular, F is continuous on $\widehat{\mathbb{R}}$.

c. *Riemann-Lebesgue Lemma.* $\lim_{|\gamma|\to\infty} F(\gamma) = 0.$

d. *Time differentiation. Suppose that* $f^{(m)}$, $m \geq 1$, *exists everywhere and that* $f^{(m)} \in L^1(\mathbb{R})$. *Assume*

$$f(\pm\infty) = \cdots = f^{(m-1)}(\pm\infty) = 0, \tag{1.4.1}$$

where $f(\pm\infty) = 0$ *indicates that* $\lim_{t\to+\infty} f(t) = 0$ *and* $\lim_{t\to-\infty} f(t) = 0$. *Then*

$$f^{(m)}(t) \longleftrightarrow (2\pi i\gamma)^m F(\gamma).$$

e. *Frequency differentiation. Suppose* $t^m f(t) \in L^1(\mathbb{R})$. *Then* $tf(t), \cdots, t^{m-1}f(t) \in L^1(\mathbb{R})$, $F^{(1)}, \cdots, F^{(m)}$ *exist everywhere, and*

$$\forall j = 0, 1, \cdots, m, \qquad (-2\pi it)^j f(t) \longleftrightarrow F^{(j)}(\gamma).$$

Proof. a. $|F(\gamma)| \leq \int |f(t)|\,|e^{-2\pi it\gamma}|\,dt \leq \|f\|_{L^1(\mathbb{R})}$.

b. We begin with the estimate,

$$\begin{aligned} &|F(\gamma+\lambda) - F(\gamma)| \\ &\leq \int |f(t)|\,|e^{-2\pi it\gamma}(e^{-2\pi it\lambda} - 1)|\,dt = \int |f(t)|\,|e^{-2\pi it\lambda} - 1|\,dt. \end{aligned} \tag{1.4.2}$$

Let $f_{[\lambda]}(t) = |f(t)||e^{-2\pi it\lambda} - 1|$ so that $\lim_{\lambda\to 0} f_{[\lambda]}(t) = 0$ for all t and $|f_{[\lambda]}(t)| \leq 2|f(t)|$. Thus, LDC (the Lebesgue Dominated Convergence Theorem, *Theorem A.9* from *Appendix A*) applies to the right side of (1.4.2), which is independent of $\gamma \in \widehat{\mathbb{R}}$. Consequently, we have

$$\forall \epsilon > 0, \quad \exists \lambda_0 > 0 \quad \text{such that} \quad \forall \lambda \in (0, \lambda_0) \quad \text{and} \quad \forall \gamma \in \widehat{\mathbb{R}},$$

$$|F(\gamma+\lambda) - F(\gamma)| < \epsilon.$$

This is the desired uniform continuity.

c. Suppose $f = \mathbf{1}_{[a,b]}$ and $\gamma \neq 0$. Then

$$|\widehat{f}(\gamma)| = \left| \int_a^b e^{-2\pi it\gamma}\,dt \right| = \frac{1}{2\pi|\gamma|}|e^{-2\pi ib\gamma} - e^{-2\pi ia\gamma}| \leq \frac{1}{\pi|\gamma|},$$

and this tends to 0 as $|\gamma|$ tends to infinity.

Therefore, $\lim_{|\gamma|\to\infty} |\widehat{f}(\gamma)| = 0$ if $f = \sum_{j=1}^{n} c_j \mathbf{1}_{[a_j,b_j]}$, where $b_j \leq a_{j+1}$.

For arbitrary $f \in L^1(\mathbb{R})$, $f \leftrightarrow F$, we take $\epsilon > 0$; and we shall find $\gamma_\epsilon > 0$ such that if $|\gamma| > \gamma_\epsilon$ then $|\widehat{f}(\gamma)| < \epsilon$. To this end we invoke *Theorem A.5* and choose

$$g = \sum_{j=1}^{n} c_j \mathbf{1}_{[a_j,b_j]}, \qquad g \longleftrightarrow G,$$

where $b_j \leq a_{j+1}$, for which $\|f - g\|_{L^1(\mathbb{R})} < \epsilon/2$. Consequently, we have

$$\forall \gamma \in \widehat{\mathbb{R}}, \quad |F(\gamma)| \leq |F(\gamma) - G(\gamma)| + |G(\gamma)|$$
$$\leq \|f - g\|_{L^1(\mathbb{R})} + |G(\gamma)| < \frac{\epsilon}{2} + |G(\gamma)|.$$

From the previous step we can take $\gamma_\epsilon > 0$ such that $|\gamma| > \gamma_\epsilon$ implies $|G(\gamma)| < \epsilon/2$. This completes the proof.

d. By integration by parts (*Theorem A.22*), we compute

$$\int_{-S}^{T} f^{(m)}(t)e^{-2\pi it\gamma}\, dt = f^{(m-1)}(t)e^{-2\pi it\gamma}\Big|_{-S}^{T}$$
$$+ 2\pi i\gamma \int_{-S}^{T} f^{(m-1)}(t)e^{-2\pi it\gamma}\, dt = f^{(m-1)}(t)e^{-2\pi it\gamma}\Big|_{-S}^{T}$$
$$+ 2\pi i\gamma\left(f^{(m-2)}(t)e^{-2\pi it\gamma}\Big|_{-S}^{T} + 2\pi i\gamma \int_{-S}^{T} f^{(m-2)}(t)e^{-2\pi it\gamma}\, dt\right)$$
$$= \cdots = \sum_{j=0}^{m-1} (2\pi i\gamma)^j$$
$$\times \left(f^{(m-(j+1))}(T)e^{-2\pi iT\gamma} - f^{(m-(j+1))}(-S)e^{2\pi iS\gamma}\right)$$
$$+ (2\pi i\gamma)^m \int_{-S}^{T} f(t)e^{-2\pi it\gamma}\, dt.$$

group better

Letting $S, T \to \infty$, the right side converges to $(2\pi i\gamma)^m F(\gamma)$ and the result is proved.

e.Without loss of generality let $m = 1$ and fix $\gamma \in \widehat{\mathbb{R}}$. Then

$$\frac{F(\gamma+\lambda) - F(\gamma)}{\lambda} = \int f(t)e^{-2\pi it\gamma}\left(\frac{e^{-2\pi it\lambda} - 1}{\lambda}\right) dt,$$

and we designate the integrand by $f(t,\lambda)$ (γ is fixed).

By the mean value theorem we have the estimate

$$\left|\frac{e^{-2\pi it\lambda}-1}{\lambda}\right| = \left|\frac{\cos 2\pi t\lambda - 1}{2\pi t\lambda}2\pi t - i\frac{\sin 2\pi t\lambda}{2\pi t\lambda}2\pi t\right|$$

$$\leq 2\pi|t|\left|\frac{\cos 2\pi t\lambda - 1}{2\pi t\lambda}\right| + 2\pi|t|$$

$$\leq 2\pi|t|\frac{|\sin\xi||2\pi t\lambda|}{|2\pi t\lambda|} + 2\pi|t| \leq 4\pi|t|.$$

Consequently,

$$|f(t,\lambda)| \leq 4\pi|tf(t)| \text{ a.e.,} \tag{1.4.3}$$

and we also know that

$$\lim_{\lambda\to 0} f(t,\lambda) = -2\pi itf(t)e^{-2\pi it\gamma} \text{ a.e.} \tag{1.4.4}$$

since

$$\lim_{\lambda\to 0}\frac{\cos 2\pi t\lambda - 1}{\lambda} = 2\pi t\lim_{\alpha\to 0}\frac{\cos\alpha - 1}{\alpha} = 0$$

and

$$\lim_{\lambda\to 0}\frac{-i\sin 2\pi t\lambda}{\lambda} = -2\pi it.$$

By (1.4.3) and (1.4.4) we can invoke LDC and assert that

$$\exists \lim_{\lambda\to 0}\tfrac{F(\gamma+\lambda)-F(\gamma)}{\lambda} = \int(-2\pi it)f(t)e^{-2\pi it\gamma}\,dt. \qquad \square$$

1.4.2 Remark. THE ROLE OF ABSOLUTE CONTINUITY

a. Suppose $f^{(m)}$ exists a.e. and $f^{(m)} \in L^1(\mathbb{R})$. If $a, b \in \mathbb{R}$ and $c \in \mathbb{C}$ then

$$F(t) = c + \int_a^t f^{(m)}(u)\,du$$

is absolutely continuous on $[a,b]$ and $F(a) = c$. FTCI (*Theorem A.20*) implies $F' = f^{(m)}$ a.e. on $[a,b]$. This does not imply that $f^{(m-1)} \in$

$AC_{\text{loc}}(\mathbb{R})$; and closely related to this phenomenon is the fact that it is *not* necessarily true that

$$(f')^\wedge(\gamma) = 2\pi i\gamma \hat{f}(\gamma) \quad \text{on } \hat{\mathbb{R}} \tag{1.4.5}$$

when $f, f' \in L^1(\mathbb{R})$. For example if f_C is the Cantor function for the usual 1/3-Cantor set C on $[0,1]$, e.g., [Ben76, page 22], then $f = \tau_{-1} f_C + (1 - f_C)\mathbf{1}_{[0,1]}$ defines a continuous compactly supported function of bounded variation on $\mathbb{R}$ for which $f' = 0$ a.e. In particular, $f, f' \in L^1(\mathbb{R})$ and (1.4.5) fails.

In this regard, note that if $f, f' \in L^1(\mathbb{R})$ and $f \in AC_{\text{loc}}(\mathbb{R})$ then

$$\int f'(t)\, dt = 0$$

[Ben76, Theorem 4.16].

b. The formula,

$$\int f^{(m)}(t) e^{-2\pi i t\gamma}\, dt = (2\pi i\gamma)^m \hat{f}(\gamma), \tag{1.4.6}$$

is true by the proof of *Theorem 1.4.1d* if we replace the hypothesis, $f^{(m)} \in L^1(\mathbb{R})$, by $f^{(m)} \in L^1_{\text{loc}}(\mathbb{R})$. In either case, the application of integration by parts in the proof, "going from m to $m-1$", is subtle since the everywhere differentiability of $f^{(m)}$ allows us to conclude that $f^{(m-1)} \in AC_{\text{loc}}(\mathbb{R})$ [Ben76, Theorem 4.15], and this smoothness allows us to integrate by parts.

c. Equation (1.4.6) is also valid, without the aformentioned subtlety, if the hypotheses, that $f^{(m)}$ is everywhere differentiable and $f^{(m)} \in L^1(\mathbb{R})$, are replaced by the hypothesis that $f^{(m)}$ be piecewise continuous, cf., the example of part *a* and the delicate issues that can arise in [Ben76, Section 4.6].

d. The hypothesis (1.4.1) of *Theorem 1.4.1d* is not required. For simplicity, let $m = 1$ and assume $f, f' \in L^1(\mathbb{R})$ and $f \in AC_{\text{loc}}(\mathbb{R})$, cf., part *b*. For fixed $a \in \mathbb{R}$ and $c \in \mathbb{C}$, set $F(t) = c + \int_a^t f'(u)\, du$. By FTCI, $F \in AC_{\text{loc}}(\mathbb{R})$ and $F' = f'$ a.e. Since $f \in AC_{\text{loc}}(\mathbb{R})$, we have $f = F + C$ on $[a, \infty)$ and so

$$\forall t \in [a, \infty), \quad f(t) = F(a) + C + \int_a^t f'(u)\, du.$$

Therefore, $f(a) = F(a) + C$ and

$$f(t) - f(a) = \int_a^t f'(u)\,du.$$

Thus, $f' \in L^1(\mathbb{R})$ implies $\lim_{t\to\pm\infty} f(t) = L_\pm$, and we have $L_+ = L_- = 0$ since $f \in L^1(\mathbb{R})$. This is (1.4.1). Also, this calculation shows that

$$\forall a \in \mathbb{R}, \quad f(a) = -\int_a^\infty f'(u)\,du.$$

e. In our proof of *Theorem 1.4.1d* we did not require that $f^{(j)} \in L^1(\mathbb{R})$ for $0 < j < m$. We only used the fact that each such $f^{(j)} \in AC_{\text{loc}}(\mathbb{R})$. It is true, however, that if $f, f^{(m)} \in L^1(\mathbb{R})$ then $f^{(j)} \in L^1(\mathbb{R})$ for $0 < j < m$, e.g., [BC49, pages 29-30].

Proposition 1.4.3 below is an extension for $m < 0$ of *Theorem 1.4.1d.*

1.4.3 Proposition.

Let $f \in L^1(\mathbb{R})$, $f \leftrightarrow F$. Define $g(t) = \int_{-\infty}^t f(u)\,du$ and assume $g \in L^1(\mathbb{R})$. (Note that $\int f(t)\,dt = 0$ since $g \in L^1(\mathbb{R})$, e.g., Exercise 1.10.) Then $F(\gamma) = 2\pi i\gamma\widehat{g}(\gamma)$ for $\gamma \in \widehat{\mathbb{R}}$, and so

$$\forall \gamma \in \widehat{\mathbb{R}}\backslash\{0\}, \quad \widehat{g}(\gamma) = \frac{1}{2\pi i\gamma}F(\gamma),$$

i.e.,

$$g(t) = \int_{-\infty}^t f(u)\,du \longleftrightarrow \frac{1}{2\pi i\gamma}F(\gamma),$$

where $\lim_{\gamma\to 0} F(\gamma)/(2\pi i\gamma) = \widehat{g}(0)$.

Proof. We calculate

$$\int_{-S}^T g'(t)e^{-2\pi it\gamma}\,dt$$

$$= g(t)(-2\pi i\gamma)e^{-2\pi it\gamma}\Big|_{-S}^T + \int_{-S}^T g(t)(2\pi i\gamma)e^{-2\pi it\gamma}\,dt.$$

Since $g(\pm\infty) = 0$ and $g'(t) = f(t)$ *a.e.* by FTCI, we can conclude that $2\pi i\gamma\widehat{g}(\gamma) = F(\gamma)$. □

1.4.4 Example. $C_0(\widehat{\mathbb{R}})\backslash A(\widehat{\mathbb{R}}) \neq \varnothing$

Theorem 1.4.1b,c allow us to conclude that $A(\widehat{\mathbb{R}}) \subseteq C_0(\widehat{\mathbb{R}})$, where $C_0(\widehat{\mathbb{R}})$ is the space of continuous functions F on $\widehat{\mathbb{R}}$ for which

$$\lim_{|\gamma|\to\infty} F(\gamma) = 0.$$

It is relatively easy to check that this inclusion is proper. For example, if F is defined as

$$F(\gamma) = \begin{cases} \frac{1}{\log\gamma}, & \text{if } \gamma > e, \\ \frac{\gamma}{e}, & \text{if } 0 \leq \gamma \leq e, \end{cases} \tag{1.4.7}$$

on $[0, \infty)$ and as $-F(-\gamma)$ on $(-\infty, 0]$ then $F \in C_0(\widehat{\mathbb{R}})$. The fact that $F \notin A(\widehat{R})$ depends on the divergence of $\int_e^\infty \frac{d\gamma}{\gamma\log\gamma}$. Instead of providing the details we refer ahead to *Example 3.3.4a* where the analogous calculation for Fourier series coefficients is not only verified, but motivated, cf., [Gol61, pages 8-9].

The function in (1.4.7) is not an isolated example. In fact, although $A(\widehat{\mathbb{R}})$ is only a set of first category in $C_0(\widehat{\mathbb{R}})$, we do have $\overline{A(\widehat{\mathbb{R}})} = C_0(\widehat{\mathbb{R}})$, *Exercise 1.40*. Even more, a Baire category argument can also be used to show the existence of $F \in C_c(\widehat{\mathbb{R}})$ for which $F \notin A(\widehat{\mathbb{R}})$. Explicit examples of such functions are more difficult to construct, but it is possible to do so, e.g., define

$$B(t) = \begin{cases} \frac{1}{n}\sin(2\pi 4^n t), & \text{if } \frac{1}{2^{n+1}} \leq |t| \leq \frac{1}{2^n}, \\ 0, & \text{if } t = 0 \text{ or } |t| > \frac{1}{2}, \end{cases}$$

[Her85]. ("B" is for "butterfly".)

It is natural to ask for an *intrinsic characterization* of $A(\widehat{\mathbb{R}})$, i.e., to seek a theorem of the form "$F \in C_0(\widehat{\mathbb{R}})$ is an element of $A(\widehat{\mathbb{R}})$ if and only if ...", where "..." is a statement about the behavior of F on $\widehat{\mathbb{R}}$. This is an open problem, e.g., [Kah70].

1.4.5 Remark. PERSPECTIVE ON THE OPERATIONAL CALCULUS

Theorem 1.4.1 is a major component of the operational calculus used in classical electrical engineering and in solving various differential equations. Typically, a calculus problem, e.g., a differential equation, is transformed into an algebra problem by *Theorem 1.4.1d*; the algebra problem is solved and the solution is transformed by an inversion formula into the solution of the original problem. A feature of this formalism is the notion of *convolution*.

1.5 Convolution

1.5.1 Definition. CONVOLUTION

Let $f, g \in L^1(\mathbb{R})$. The *convolution* of f and g, denoted by $f * g$, is

$$f * g(t) = \int f(t-u)g(u)\,du = \int f(u)g(t-u)\,du.$$

It is not difficult to prove that $f * g \in L^1(\mathbb{R})$ (*Exercise 1.31*). Later, we shall demonstrate the role of convolution in the method alluded to in *Remark 1.4.5*. The algebraic properties of convolution are the subject of *Exercise 1.32*.

We illustrate convolution as follows.

Figure 1.7

1.5.2 Proposition.

Let $f, g \in L^1(\mathbb{R})$, *with corresponding Fourier pairs* $f \leftrightarrow F$ *and* $g \leftrightarrow G$. *Then* $f * g \in L^1(\mathbb{R})$ *and* $(f * g)^\wedge = \widehat{f}\widehat{g}$, *i.e.,*

$$f * g \longleftrightarrow FG.$$

Proof. As indicated in *Definition 1.5.1*, the assertion that $f*g \in L^1(\mathbb{R})$ is the task of *Exercise 1.31*. Assuming this fact, we use the Fubini-Tonelli Theorem (*Theorem A.14*) to compute

$$\begin{aligned}
(f * g)^\wedge(\gamma) &= \iint f(t-u)g(u)e^{-2\pi i t\gamma}\, du dt \\
&= \iint f(t-u)g(u)e^{-2\pi i(t-u)\gamma}\, e^{-2\pi i u\gamma}\, du dt \\
&= \int \left(\int f(t-u)e^{-2\pi i(t-u)\gamma}\, dt \right) g(u)e^{-2\pi i u\gamma}\, du \\
&= \int \widehat{f}(\gamma)g(u)e^{-2\pi i u\gamma}\, du = \widehat{f}(\gamma)\widehat{g}(\gamma).
\end{aligned}$$

□

1.5.3 Remark. PERSPECTIVE ON THE OPERATIONAL CALCULUS

a. *Proposition 1.5.2* is another key ingredient in the operational calculus recipe mentioned in *Remark 1.4.5.* The complete story unfolds when we discuss distributions in *Chapter 2.*

b. A critical step in the proof of *Proposition 1.5.2* involved the translation invariance of the integral (the penultimate equality). This feature accounts for the effectiveness of time invariant systems in electrical engineering. Mathematically, it has to do with the fact that $\mathbb{R}$ is a (locally compact) group with an invariant measure.

1.6 Approximate identities and examples

The following notion is critical in approximating the unit impulse and for providing examples in applications including signal processing and spectral estimation.

1.6.1 Definition. APPROXIMATE IDENTITY

An *approximate identity* is a family $\{k_{(\lambda)} : \lambda > 0\} \subseteq L^1(\mathbb{R})$ of functions with the properties:

a. $\forall\lambda > 0, \int k_{(\lambda)}(t)\, dt = 1,$

b. $\exists K$ such that $\forall\lambda > 0,\ \|k_{(\lambda)}\|_1 \leq K,$

c. $\forall\eta > 0, \lim\limits_{\lambda\to\infty} \int_{|t|\geq\eta} |k_{(\lambda)}(t)|\, dt = 0.$

Caveat. The subscript "(λ)" in *Definition 1.6.1* does not necessarily denote a dilation. The following result, however, shows that dilations yield a large class of approximate identities.

1.6.2 Proposition.

Let $k \in L^1(\mathbb{R})$ have the property that $\int k(t)\, dt = 1$. The family $\{k_\lambda : k_\lambda(t) = \lambda k(\lambda t), \lambda > 0\} \subseteq L^1(\mathbb{R})$ of dilations is an approximate identity.

Proof. To verify the conditions of *Definition 1.6.1a*, we compute

$$\int k_\lambda(t)\, dt = \lambda \int k(\lambda t)\, dt = \int k(t)\, dt = 1.$$

For part b we compute

$$\int |k_\lambda(t)|\, dt = \lambda \int |k(\lambda t)|\, dt = \int |k(u)|\, du = K < \infty,$$

where K is finite since $k \in L^1(\mathbb{R})$.

For part c, take $\eta > 0$ and compute

$$\int_{|t|\geq\eta} |k_\lambda(t)|\, dt = \lambda \int_{|t|\geq\eta} |k(\lambda t)|\, dt = \int_{|u|\geq\lambda\eta} |k(u)|\, du;$$

this last term tends to 0 as λ tends to ∞ since $\eta > 0$ and because of the definition of the integral. □

1.6.3 Proposition.

$$\int_{-\infty}^{\infty} \frac{\sin t}{t}\, dt = \int_{-\infty}^{\infty} \frac{\sin^2 t}{t^2}\, dt = \pi.$$

(Contrary to our consistent notation, we have written the limits of integration $\pm\infty$ since the first integral is an improper Riemann integral.)

Proof. a. To prove that these integrals are equal, let $u = \sin^2 t$ and $dv = \frac{dt}{t^2}$ in the second integral so that

$$\begin{aligned}\int_0^\infty \frac{\sin^2 t}{t^2}\,dt &= -\frac{\sin^2 t}{t}\bigg|_0^\infty + \int_0^\infty \frac{2}{t}\sin t\cos t\,dt \\ &= \int_0^\infty \frac{\sin 2t}{t}\,dt = \int_0^\infty \frac{\sin t}{t}\,dt.\end{aligned}$$

b. We now show that

$$\int_0^\infty \frac{\sin t}{t}\,dt = \frac{\pi}{2}$$

to complete the result. There are several ways to accomplish this computation including a contour integral calculation. We choose the following method.

Let $F(\sigma)$ be the Laplace transform,

$$F(\sigma) = \int_0^\infty e^{-\sigma t}\frac{\sin t}{t}\,dt = \mathcal{L}(\frac{\sin t}{t})(\sigma).$$

We assert that $F(\sigma)$ is a continuous function on $[0,\infty)$ and that the formal calculation,

$$\forall\sigma > 0, \quad \exists F'(\sigma) = -\int_0^\infty e^{-\sigma t}\sin t\,dt,$$

is in fact true. (The verification of these claims involves uniform convergence.) The convergence of $F(0)$ is clear by an alternating series argument.

It is easy to see that $\mathcal{L}(\sin t)(\sigma) = \frac{1}{1+\sigma^2}, \sigma > 0$, either by direct calculation using integration by parts or by use of the general Laplace transform formula, $\mathcal{L}(g^{(2)})(\sigma) = \sigma^2\mathcal{L}(g)(\sigma) - \sigma g(0) - \sigma g'(0)$, for the special function $g(t) = \sin t$. Thus, using FTC, we compute

$$\begin{aligned}F(\sigma) - F(0) &= \int_0^\sigma F'(\eta)d\eta = -\int_0^\sigma \mathcal{L}(\sin t)(\eta)d\eta \\ &= -\int_0^\sigma \frac{d\eta}{1+\eta^2} = -\tan^{-1}\sigma, \quad \sigma > 0.\end{aligned}$$

We know $F(\infty) = 0$ by LDC, and so

$$F(0) = \lim_{\sigma \to \infty} \tan^{-1} \sigma = \frac{\pi}{2}. \qquad \square$$

1.6.4 Remark. THE DIRICHLET AND FEJÉR KERNELS

a. The family $\{d_\lambda\}$ of dilations of $d(t)$ is the *Dirichlet kernel*, and the family $\{w_\lambda\}$ of dilations of $w(t)$ is the *Fejér kernel*. The Fejér kernel is an approximate identity by *Propositions 1.6.2* and *1.6.3*. The Dirichlet kernel is not an approximate identity since $d_\lambda \notin L^1(\mathbb{R})$, whereas *Proposition 1.6.3* highlights a similarity between d_λ and w_λ, cf., *Exercises 1.17* and *1.46*. Although $\{d_\lambda\}$ is not an approximate identity, it does possess the property that its "mass" accumulates at the origin, while its Fourier transform tends to the function identically 1 on $\widehat{\mathbb{R}}$ as $\lambda \to \infty$.

b. By definition of convolution,

$$\frac{1}{T}\mathbf{1}_{[-\frac{T}{2},\frac{T}{2})} * \mathbf{1}_{[-\frac{T}{2},\frac{T}{2})}(t) = \max\left(1 - \frac{|t|}{T}, 0\right). \tag{1.6.1}$$

Thus, by *Example 1.3.1* and *Proposition 1.5.2*, we have another proof of the Fourier transform pairing (1.3.4), viz.,

$$\max\left(1 - \frac{|t|}{T}, 0\right) \longleftrightarrow w_{2\pi T}(\gamma),$$

since ~~the right side of (1.6.1) is~~ delete

$$\frac{1}{T} d_{\pi T}(\gamma)^2 = w_{2\pi T}(\gamma).$$

(1.6.1) asserts that the "convolution of rectangles" is a "triangle". Further steps are carried out in *Exercise 1.17*.

Yet another way to prove (1.3.4) is to introduce the following translation and dilation of the *Haar wavelet*, viz.,

$$f(t) = \begin{cases} \frac{1}{T}, & \text{if } t \in [-T, 0), \\ -\frac{1}{T}, & \text{if } t \in [0, T), \\ 0, & \text{otherwise.} \end{cases}$$

An elementary computation gives

$$g(t) = \int_{-\infty}^{t} f(u)\,du = \max\left(1 - \frac{|t|}{T}, 0\right),$$

and, in particular, $f, g \in L^1(\mathbb{R})$. From *Proposition 1.4.3* we obtain the Fourier transform pairing,

$$\max\left(1 - \frac{|t|}{T}, 0\right) \longleftrightarrow \frac{1}{2\pi i\gamma}\widehat{f}(\gamma). \tag{1.6.2}$$

Since $f = \frac{1}{T}\left(\tau_{-\frac{T}{2}}\mathbf{1}_{[-\frac{T}{2},\frac{T}{2})} - \tau_{\frac{T}{2}}\mathbf{1}_{[-\frac{T}{2},\frac{T}{2})}\right)$, we can easily compute

$$\frac{1}{2\pi i\gamma}\widehat{f}(\gamma) = w_{2\pi T}(\gamma),$$

and so (1.3.4) is obtained from (1.6.2).

1.6.5 Example. THE POISSON KERNEL

a. The family $\{p_\lambda\}$ of dilations of $p(t)$, e.g., *Example 1.3.2* is an approximate identity by *Proposition 1.6.2* and the fact that

$$\int \frac{dt}{\pi(1+t^2)} = 1. \tag{1.6.3}$$

$\{p_\lambda\}$ is the *Poisson kernel.*

b. We can do the elementary computation (1.6.3) by means of direct integration, cf., [Rud66, page 4] or in the following more complicated way. In *Example 1.3.2* we verified that $e^{-2\pi|t|} \leftrightarrow p(\gamma)$ so that, once we have the inversion theory of *Section 1.7*, we shall have the pairing $p(t) \leftrightarrow e^{-2\pi|\gamma|} \equiv P(\gamma)$. Thus, $\widehat{p}(0) = P(0) = 1$ by *Theorem 1.1.6* (*Theorem 1.7.6*), and this is (1.6.3).

c. The *Paley-Wiener Logarithmic Integral Theorem* is the following assertion. *Let ϕ be a non-negative function for which $\int \phi^2(\gamma)\,d\gamma < \infty$. There is a function f vanishing on $(-\infty, 0)$ for which $\int |f(t)|^2\,dt < \infty$ and $|\widehat{f}| = \varphi$ a.e. if and only if*

$$\int \frac{|\log\phi(\gamma)|}{1+\gamma^2}\,d\gamma < \infty,$$

cf., *Section 1.10* for the definition of $\widehat{f}$ in this case. A function f is referred to as a *causal signal* in the case it vanishes on $(-\infty, 0)$; and a function f is a *signal of finite energy* in the case $\int |f(t)|^2\, dt < \infty$.

$$p_\lambda(t) = \frac{\lambda}{\pi(1+\lambda^2 t^2)} \qquad\qquad \widehat{p}_\lambda(\gamma) = e^{-2\pi|\gamma|/\lambda}$$

Figure 1.8

The Paley-Wiener Logarithmic Integral Theorem is used to characterize causal signals, e.g., [OS75], [Pap77], cf., *Remark 3.7.10.* We have stated it now to highlight the appearance of the Poisson function, to mention finite energy signals which we shall study in *Section 1.10*, and to give an explicit result which alludes to the profound uniqueness and uncertainty principle properties of Fourier analysis, cf., *Example 1.10.6.*

1.6.6 Example. The Gauss Kernel

The family $\{g_\lambda\}$ of dilations of $g(t)$, e.g., *Example 1.3.3*, is an approximate identity by *Proposition 1.6.2* and the fact that $\int g(t)\, dt = (1/\sqrt{\pi}) \int e^{-t^2}\, dt = 1$. $\{g_\lambda\}$ is the *Gauss kernel.*

$$g_\lambda(t) = \frac{\lambda}{\sqrt{\pi}} e^{-(\lambda t)^2} \qquad \widehat{g}_\lambda(\gamma) = e^{-(\pi\gamma/\lambda)^2}$$

$(g_\lambda)^\wedge(\gamma) = \ldots$

Figure 1.9

1.6.7 Example. PROPERTIES OF THE POISSON FUNCTION

a. If $\{p_\lambda : \lambda > 0\}$ is the Poisson kernel then $\widehat{p}_\lambda$ is not differentiable at 0 even though $p_\lambda \in L^1(\mathbb{R})$ and decreases like $1/t^2$ as $|t| \to \infty$, cf., *Theorem 1.4.1e*. The verification of nondifferentiability is elementary from *Example 1.3.2*; in fact, expanding $\widehat{p}_\lambda$ in a Taylor series, we have

$$\lim_{\gamma\to 0\pm} \frac{\widehat{p}_\lambda(\gamma) - \widehat{p}_\lambda(0)}{\gamma - 0} = \mp\frac{2\pi}{\lambda}.$$

Further, $\widehat{p}_\lambda$ is even, convex, and decreasing to 0 on $(0, \infty)$.

b. Because of *Example 1.6.5a,c*, we next note that

$$I = \int \frac{\log|\gamma|}{1+\gamma^2}\, d\gamma = 0.$$

In fact, integrating over $(0, \infty)$ and letting $\lambda = 1/\gamma$ we see that $I/2 = -I/2$, cf., [AGR88] for a unifying principle to calculate such integrals by real methods.

c. The linear fractional transformation $w = \frac{z-i}{z+i}$ maps the upper half-plane onto the unit disk with $\mathbb{R}$ mapped onto the unit circle. As such, the Poisson function $p(t)$ assumes the role of the Jacobian:

$$\frac{1}{2\pi}\int_0^{2\pi} g(e^{i\theta})\,d\theta = \int f(t)p(t)\,dt,$$

where $g(e^{i\theta})$ is defined on the unit circle and

$$f(t) = g(e^{i\theta}) = g\left(\frac{t-i}{t+i}\right).$$

d. Note that if $a, b > 0$ then

$$p_{\frac{1}{a}} * p_{\frac{1}{b}} = p_{\frac{1}{a+b}}$$

because of *Example 1.3.2* and *Proposition 1.5.2.*

For a host of other examples we refer to [Harr78]

1.6.8 Example. Central Limit Theorem

a. We can compute

$$\forall t \in \mathbb{R}, \quad \lim_{n\to\infty}\left(\cos\frac{t}{\sqrt{n}}\right)^n = e^{-t^2/2}, \tag{1.6.4}$$

[Hint. We approximate the cosine by the first two terms of its Taylor series, and have the approximation,

$$\left(\cos\frac{t}{\sqrt{n}}\right)^n \approx \left(1 + \frac{1}{(\frac{-2n}{t^2})}\right)^{(\frac{-2n}{t^2})(\frac{-t^2}{2})}.$$

The right side converges to the right side of (1.6.4) by the definition of e. The "error terms" can be shown to tend to zero by a variety of methods.]

b. The *Central Limit Theorem* in probability theory is *equivalent* to the following result. *Let $f \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ be nonnegative and continuous, and assume that*

$$\int f(t)\,dt = 1, \quad \int t f(t)\,dt = 0, \quad \text{and}$$
(1.6.5)
$$\int t^2 f(t)\,dt = 1.$$

Define $f_1 = f, f_n = f * f_{n-1}$ *for each* $n \geq 2$, *and* $g_n(t) = n^{1/2} f(n^{1/2}t)$. *Then*

$$\forall a < b, \quad \lim_{n\to\infty} \int_a^b g_n(t)\,dt = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\,dt.$$

There is a dazzling treatment of this material in [Kör88, Chapter 70]. The hypotheses (1.6.5) are equivalent to the statement that f *is the probability density function of a random variable* X *having mean* 0 *and variance* 1. We shall explain this terminology in *Section 2.8.* The Central Limit Theorem deals with the asymptotic behavior of sample means as the sample size increases, and it quantifies the remarkable fact that the sum of a large number of independent random variables approximates a normal, i.e., Gaussian, distribution, e.g., [Kac59], [Lam66, Chapter 3], [Pri81]. It will not come as a surprise that (1.6.4) is related to the Central Limit Theorem.

The major *elementary* property of approximate identities is given in *Theorem 1.6.9a. Theorem 1.6.9b* is the special case for the Fejér kernel, and part c is the Fourier transform uniqueness theorem. We prove the uniqueness theorem as a corollary of part b.

1.6.9 Theorem. Approximation and Uniqueness

Let $f \in L^1(\mathbb{R})$.

a. If $\{k_{(\lambda)} : \lambda > 0\} \subseteq L^1(\mathbb{R})$ *is an approximate identity then*

$$\lim_{\lambda\to\infty} \|f - f * k_{(\lambda)}\|_{L^1(\mathbb{R})} = 0.$$

b. We have

$$\lim_{\lambda\to\infty} \int \Big|f(t) - \int_{-\lambda/2\pi}^{\lambda/2\pi} (1 - \frac{2\pi|\gamma|}{\lambda}) \hat{f}(\gamma) e^{2\pi i t\gamma}\,d\gamma\Big|\,dt = 0.$$

c. Uniqueness. *If $\widehat{f} = 0$ on $\widehat{\mathbb{R}}$ then f is the 0-function.*

Proof. a. We use the fact that $\int k_{(\lambda)}(t)\,dt = 1$ to compute

$$\begin{aligned}
&\|f - f * k_{(\lambda)}\|_{L^1(\mathbb{R})} \\
&= \int \left| \int k_{(\lambda)}(u) f(t)\,du - \int k_{(\lambda)}(u) f(t-u)\,du \right| dt \\
&\leq \int |k_{(\lambda)}(u)| \left(\int |f(t) - \tau_u f(t)|\,dt \right) du.
\end{aligned}$$

Let $\epsilon > 0$. By *Theorem A.5*, there is $\eta > 0$ with the property that

$$\forall |u| < \eta, \quad \|f - \tau_u f\|_{L^1(\mathbb{R})} < \epsilon / K, \tag{1.6.6}$$

where $\|k_{(\lambda)}\|_{L^1(\mathbb{R})} \leq K$. Therefore, we have the estimate

$$\begin{aligned}
\|f - f * k_{(\lambda)}\|_{L^1(\mathbb{R})} &\leq 2\|f\|_{L^1(\mathbb{R})} \int_{|u| \geq \eta} |k_{(\lambda)}(u)| du \\
&\quad + \frac{\epsilon}{K} \int_{|u| \leq \eta} |k_{(\lambda)}(u)| du \\
&\leq \epsilon + 2\|f\|_{L^1(\mathbb{R})} \int_{|u| \geq \eta} |k_{(\lambda)}(u)| du.
\end{aligned}$$

Consequently, by the definition of an approximate identity, we have

$$\overline{\lim_{\lambda \to \infty}} \|f - f * k_{(\lambda)}\|_{L^1(\mathbb{R})} \leq \epsilon;$$

and so we obtain part *a* since $\epsilon > 0$ can be chosen as small as we like.

b. To begin with, the calculation in *Example 1.3.4* shows that

$$w_\lambda(t) = \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\gamma|}{\lambda}\right) e^{2\pi i t \gamma}\,d\gamma.$$

Then, by the definition of convolution and an application of the Fubini-Tonelli Theorem, we compute

$$f * w_\lambda(t) = \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\gamma|}{\lambda}\right) \widehat{f}(\gamma) e^{2\pi i t \gamma}\,d\gamma.$$

Since $\{w_\lambda\}$ is an approximate identity, part *b* follows from part *a*.

Part *c* follows from part *b*. In fact, the hypothesis and part *b* imply $\|f\|_{L^1(\mathbb{R})} = 0$; and so f is the 0-function by *Theorem A.5*. □

1.6.10 Remark. INVERSION FORMULA FOR L^1-NORM

Theorem 1.6.9b has the flavor of the inversion results discussed in *Section 1.1.* For example we could compare *Theorem 1.6.9b* with *Theorem 1.1.6.* There are two differences:

i. *Theorem 1.1.6* is a pointwise result whereas we deal with L^1-convergence in *Theorem 1.6.9b*;

ii. the Dirichlet kernel is used in the statement of *Theorem 1.1.6*, whereas the Fejér kernel is used in *Theorem 1.6.9b.*

1.6.11 Proposition.

Let $f \in L^\infty(\mathbb{R})$ be continuous on $\mathbb{R}$. If $\{k_{(\lambda)} : \lambda > 0\} \subseteq L^1(\mathbb{R})$ is an approximate identity then

$$\forall t \in \mathbb{R}, \quad \lim_{\lambda \to \infty} f * k_{(\lambda)}(t) = f(t).$$

($L^\infty(\mathbb{R})$ is defined in Definition A.10.*)*

Proof. We first compute

$$\begin{aligned} |f(t) - f * k_{(\lambda)}(t)| &= \left| \int k_{(\lambda)}(u)(f(t) - f(t-u))\, du \right| \\ &\leq \int |k_{(\lambda)}(u)|\, |f(t) - f(t-u)|\, du, \end{aligned}$$

for a fixed $t \in \mathbb{R}$.

Let $\epsilon > 0$. Since f is continuous, there is $\eta > 0$ such that if $0 \leq |u| < \eta$ then $|f(t) - f(t-u)| < \epsilon/K$, where $\|k_{(\lambda)}\|_{L^1(\mathbb{R})} \leq K$. This yields the estimate,

$$|f(t) - f * k_{(\lambda)}(t)| \leq \epsilon + 2\|f\|_{L^\infty(\mathbb{R})} \int_{|u| \geq \eta} |k_{(\lambda)}(u)|\, du.$$

Consequently, by the definition of an approximate identity, we have

$$\overline{\lim_{\lambda \to \infty}}\, |f(t) - f * k_{(\lambda)}(t)| \leq \epsilon;$$

and so we obtain our result since ϵ can be chosen as small as we like. $\square$

1.7 Pointwise inversion of the Fourier transform

1.7.1 Remark. MOTIVATION FOR THE INVERSION THEOREM
The inversion formula in *Theorem 1.1.6* for f continuous is:

$$f(t) = \lim_{\Omega\to\infty} \int_{-\Omega}^{\Omega} F(\gamma)e^{2\pi i t\gamma}\, d\gamma, \quad f \leftrightarrow F. \tag{1.7.1}$$

To see that this formula is reasonable we begin with the formal calculation:

$$\begin{aligned} \int F(\gamma)e^{2\pi i t\gamma}\, d\gamma &= \iint f(u)e^{2\pi i(t-u)\gamma}\, du\, d\gamma \\ &= \int f(u)\left[\lim_{\Omega\to\infty}\int_{-\Omega}^{\Omega} e^{2\pi i(t-u)\gamma}\, d\gamma\right] du \\ = \lim_{\Omega\to\infty}\int f(u)&\frac{\sin 2\pi(t-u)\Omega}{\pi(t-u)}\, du = \lim_{\Omega\to\infty} f * d_{2\pi\Omega}(t). \end{aligned} \tag{1.7.2}$$

2Ω

$\tau_t d_{2\pi\Omega}$ L $t+\frac{1}{2\Omega}$

t

Figure 1.10

Observe that the area L of the major lobe of $\tau_t d_{2\pi\Omega}$ in *Figure 1.10* remains constant for all Ω:

$$\int_{t-\frac{1}{2\Omega}}^{t+\frac{1}{2\Omega}} \tau_t d_{2\pi\Omega}(u)\, du = \int_{-\frac{1}{2\Omega}}^{\frac{1}{2\Omega}} d_{2\pi\Omega}(u)\, du \tag{1.7.3}$$

$$= \int_{-\frac{1}{2\Omega}}^{\frac{1}{2\Omega}} \frac{\sin 2\pi u\Omega}{\pi u}\, du = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin u}{u}\, du.$$

The calculation (1.7.2) involves changing orders of operations and a principal value argument. These steps must be justified. The basic idea, however, is clear. We expect that $\lim_{\Omega\to\infty} f * d_{2\pi\Omega}(t) = f(t)$ (and this is (1.7.1)) since L remains constant as the major lobes converge to t and since the oscillations of the minor lobes on either side of t become increasingly rapid as $\Omega \to \infty$. The intuition is that the total contribution of the minor lobes will be negligible for large Ω since the Dirichlet kernel take positive and negative values. This intuition is not quite correct since

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin u}{u}\, du > 1$$

(why?) – see *Exercise 1.4*; but the cancellation is such that we can verify (1.7.1) under the conditions given in *Theorem 1.1.6*.

We shall have more to say about the phenomenon associated with the integral in (1.7.3) in *Section 1.9*. Another issue that arises in the calculation (1.7.2) is the fact that $\hat{f}$ need not be in $L^1(\hat{\mathbb{R}})$ for $f \in L^1(\mathbb{R})$.

1.7.2 Example. $f \in L^1(\mathbb{R})$ DOES NOT IMPLY $\hat{f} \in L^1(\hat{\mathbb{R}})$

Let $f(t) = H(t)e^{-2\pi rt}$ where $r > 0$ and H is the *Heaviside function* defined as $H = \mathbf{1}_{[0,\infty)}$. Then $\hat{f}(\gamma) = \frac{1}{2\pi(r+i\gamma)} \notin L^1(\hat{\mathbb{R}})$, e.g., *Exercise 1.5*. This should be compared with the fact that if $f(t) = e^{-2\pi r|t|}$ then $\hat{f} = p_{1/r} \in L^1(\hat{\mathbb{R}})$, cf., the end of *Example 1.3.2*.

Having made these introductory remarks, let us begin the proof of *Theorem 1.1.6*, which is Jordan's pointwise inversion formula for the Fourier transform. *Theorem 1.1.6* and *Theorem 1.7.6* are the same.

1.7.3 Lemma. SECOND MEAN VALUE THEOREM FOR INTEGRALS OR BONNET THEOREM

Let g be continuous on $[a,b]$ and let f be increasing on $[a,b]$. There is $\xi \in [a,b]$ such that

$$\int_a^b f(t)g(t)\,dt = f(a+)\int_a^{\xi} g(t)\,dt + f(b-)\int_{\xi}^b g(t)\,dt,$$

e.g., [Apo57, page 217] *and Exercise 1.48.*

1.7.4 Remark. JORDAN DECOMPOSITION

The classical form of the *Jordan Decomposition Theorem* for a function $f : [a,b] \longleftrightarrow \mathbb{C}$ asserts that $f \in BV[a,b]$, e.g., *Definition 1.1.5, if and only if* f can be expressed as the difference $f_1 - f_2$ *of two increasing functions on* $[a,b]$. The proof is not difficult. We set $f_1(a) = 0$ and define

$$\forall t \in (a,b], \quad f_1(t) = \sup\left\{\sum |f(t_j) - f(t_{j-1})|\right\},$$

where the supremum is taken over every finite set $a = t_0 < t_1 < \ldots < t_n = t$. It is easy to see that f_1 is increasing, and straightforward to check that $f_2 \equiv f_1 - f$ is increasing. This completes the proof.

There are standard measure theoretic generalizations of this result, e.g., [AB66], [Ben76, Section 5.1], [Rud66].

1.7.5 Lemma.

Let $g \in BV[0,\epsilon]$, $\epsilon > 0$. Then

$$\lim_{\Omega\to\infty}\int_0^{\epsilon} g(t)d_{2\pi\Omega}(t)\,dt = \frac{1}{2}g(0+). \tag{1.7.4}$$

Proof. a. By the Jordan Decomposition Theorem stated in *Remark 1.7.4*, we assume that g is increasing on $[0,\epsilon]$; and, in particular, from the definition of bounded variation, g is bounded on $[0,\epsilon]$.

b.i. Assume $g(0+) = 0$ and let $\eta > 0$. Since $\int d(t)\,dt = 1$ for the Dirichlet function d (*Proposition 1.6.3*), there is $C > 0$ such that

$$\forall a,b \in \mathbb{R}, \quad \left|\int_a^b \frac{\sin t}{\pi t}\,dt\right| \le C.$$

We shall verify that

$$\overline{\lim_{\Omega\to\infty}} \left| \int_0^\epsilon g(t) d_{2\pi\Omega}(t)\, dt \right| \leq \eta C; \tag{1.7.5}$$

and this will complete the proof of (1.7.4) for the $g(0+) = 0$ case since $\eta > 0$ is arbitrary.

b.ii. Since $g(0+) = 0$, there is $\nu = \nu(\eta) \in (0, \epsilon)$ such that $|g(t)| \leq \eta$ for all $t \in (0, \nu)$. Also, by *Lemma 1.7.3*, using the continuity of $d_{2\pi\Omega}$ and the boundedness and monotonicity of g, there is $\xi \in [0, \nu]$ for which

$$\begin{aligned} &\int_0^\nu g(t) d_{2\pi\Omega}(t)\, dt \\ &= g(0+) \int_0^\xi d_{2\pi\Omega}(t)\, dt + g(\nu-) \int_\xi^\nu d_{2\pi\Omega}(t)\, dt \\ &= g(\nu-) \int_{2\pi\Omega\xi}^{2\pi\Omega\nu} \frac{\sin u}{\pi u}\, du. \end{aligned}$$

Consequently,

$$\left| \int_0^\nu g(t) d_{2\pi\Omega}(t)\, dt \right| \leq C|g(\nu-)| \leq \eta C. \tag{1.7.6}$$

b.iii. Note that $(g(t)/t)\mathbf{1}_{[\nu,\epsilon]}(t) \in L^1(\mathbb{R})$ since we don't have to deal with the origin. Therefore, by the Riemann-Lebesgue Lemma,

$$\lim_{\Omega\to\infty} \int_\nu^\epsilon g(t) d_{2\pi\Omega}(t)\, dt = 0.$$

Using this fact, (1.7.6), and the inequality

$$\begin{aligned} &\left| \int_0^\epsilon g(t) d_{2\pi\Omega}(t)\, dt \right| \\ &\leq \left| \int_0^\nu g(t) d_{2\pi\Omega}(t)\, dt \right| + \left| \int_\nu^\epsilon g(t) d_{2\pi\Omega}(t)\, dt \right|, \end{aligned}$$

we obtain (1.7.5).

c. Finally, suppose $g(0+) \neq 0$. Let $h(t) = g(t) - g(0+)$ so that $h(0+) = 0$ and

$$\lim_{\Omega\to\infty} \int_0^\epsilon h(t) d_{2\pi\Omega}(t)\, dt = 0$$

by part *b*. Also, we know from *Proposition 1.6.3* that

$$\lim_{\Omega\to\infty}\int_0^\epsilon d_{2\pi\Omega}(t)\,dt = \lim_{\Omega\to\infty}\int_0^{2\pi\Omega\epsilon} d(t)\,dt = \frac{1}{2}.$$

Combining these facts, we compute

$$\lim_{\Omega\to\infty}\int_0^\epsilon g(t)d_{2\pi\Omega}(t)\,dt$$
$$= \lim_{\Omega\to\infty}\left(\int_0^\epsilon \big(h(t)+g(0+)\big)d_{2\pi\Omega}(t)\,dt\right) = \frac{g(0+)}{2},$$

since both limits exists when we expand the middle term. □

We can now complete the proof of Jordan's theorem.

1.7.6 Theorem. Jordan Theorem

Let $f \in L^1(\mathbb{R})$ and assume $f \in BV[t-\epsilon, t+\epsilon]$ for some $t \in \mathbb{R}$ and $\epsilon > 0$. Then

$$\frac{f(t+)+f(t-)}{2} = \lim_{\Omega\to\infty}\int_{-\Omega}^{\Omega} F(\gamma)e^{2\pi it\gamma}\,d\gamma,$$

where $f \leftrightarrow F$. If f is continuous at t then the left side can be replaced by $f(t)$.

Proof. For each $\Omega > 0$, define the "partial sums"

$$S_\Omega(t) = \int_{-\Omega}^{\Omega} e^{2\pi it\gamma}F(\gamma)\,d\gamma$$
$$= \int f(u)\Big(\int_{-\Omega}^{\Omega} e^{2\pi i(t-u)\gamma}\,d\gamma\Big)\,du = f * d_{2\pi\Omega}(t).$$

The calculation is justified by the Fubini-Tonelli Theorem since the double integral on $\mathbb{R}\times[-\Omega,\Omega]$ is absolutely convergent. We write $S_\Omega(t)$ as

$$S_\Omega(t) = \int f(t-u)d_{2\pi\Omega}(u)\,du$$
$$= \int_0^\infty (f(t+u)+f(t-u))d_{2\pi\Omega}(u)\,du.$$

Let $g(u) = f(t+u) + f(t-u)$, noting that t is fixed, and let $\epsilon > 0$. The result will be proved when we show

$$\lim_{\Omega\to\infty} \int_0^{\epsilon} g(u) d_{2\pi\Omega}(u)\, du = \frac{f(t+) + f(t-)}{2} \tag{1.7.7}$$

and

$$\lim_{\Omega\to\infty} \int_{\epsilon}^{\infty} g(u) d_{2\pi\Omega}(u)\, du = 0. \tag{1.7.8}$$

Equation (1.7.7) is an immediate consequence of *Lemma 1.7.5.* Equation (1.7.8) follows from the Riemann-Lebesgue Lemma and the fact that $\frac{g(t)}{t}\tau_{\epsilon}H(t) \in L^1(\mathbb{R})$, since $f \in L^1(\mathbb{R})$. □

1.7.7 Remark. THE JORDAN THEOREM AND PARTIAL SUMS

a. We could use other kernels besides the pairing, $d_{2\pi\Omega} \leftrightarrow \mathbf{1}_{[-\Omega,\Omega)}$, to obtain analogues of Jordan's Theorem. The advantage of Jordan's Theorem is that we really are dealing with the "partial sum" S_{Ω} and not some weighting of F, such as the factor,

$$\left(1 - \frac{2\pi|\gamma|}{\lambda}\right)\mathbf{1}_{[-\frac{\lambda}{2\pi},\frac{\lambda}{2\pi}]}(\gamma),$$

in *Theorem 1.6.9b.*

b. Jordan's Theorem is the analogue for Fourier transforms of Dirichlet's Theorem for Fourier series. Dirichlet proved his result much earlier in the 19th century, and his work contains the first proper definition of the notion of "function". We shall prove Dirichlet's Theorem in *Theorem 3.1.6.*

If $f \in L^1(\mathbb{R})$ and $\hat{f} \in L^1(\hat{\mathbb{R}})$, we can use *Theorem 1.6.9* to obtain the following pointwise inversion theorem, cf., *Exercise 1.37.*

1.7.8 Theorem. INVERSION FORMULA FOR $f \in L^1(\mathbb{R}) \cap A(\mathbb{R})$

Let $f \in L^1(\mathbb{R}) \cap A(\mathbb{R})$. Then

$$\forall\, t \in \mathbb{R}, \quad f(t) = \int \hat{f}(\gamma) e^{2\pi i t\gamma}\, d\gamma. \tag{1.7.9}$$

Proof. Since $\widehat{f} \in L^1(\widehat{\mathbb{R}})$ then g defined by the right side of (1.7.9) is uniformly continuous *Theorem 1.4.1b.*

Note that

$$\left\| \int \widehat{f}(\gamma)e^{2\pi it\gamma}\,d\gamma - \int_{\frac{-\lambda}{2\pi}}^{\frac{\lambda}{2\pi}} \left(1 - \frac{2\pi|\gamma|}{\lambda}\right)\widehat{f}(\gamma)e^{2\pi it\gamma}\,d\gamma \right\|_{L^\infty(\mathbb{R})}$$

$$\leq \left\| \int_{|\gamma|>\frac{\lambda}{2\pi}} \widehat{f}(\gamma)e^{2\pi it\gamma}\,d\gamma \right\|_{L^\infty(\mathbb{R})}$$

(1.7.10)

$$+\left\| \int_{\frac{-\lambda}{2\pi}}^{\frac{\lambda}{2\pi}} \frac{2\pi|\gamma|}{\lambda}\widehat{f}(\gamma)e^{2\pi it\gamma}\,d\gamma \right\|_{L^\infty(\mathbb{R})}$$

$$\leq \int_{|\gamma|>\frac{\lambda}{2\pi}} |\widehat{f}(\gamma)|\,d\gamma + \int_{\frac{-\lambda}{2\pi}}^{\frac{\lambda}{2\pi}} \frac{2\pi|\gamma|}{\lambda}|\widehat{f}(\gamma)|\,d\gamma.$$

We shall apply LDC to the second integral on the right side of (1.7.10). Let

$$F_\lambda(\gamma) = \frac{2\pi|\gamma|}{\lambda}|\widehat{f}(\gamma)|\mathbf{1}_{[-\frac{\lambda}{2\pi},\frac{\lambda}{2\pi}]}(\gamma),$$

so that $\lim_{\lambda\to\infty} F_\lambda = 0$ a.e. and $|F_\lambda| \leq 2\pi|\widehat{f}| \in L^1(\widehat{\mathbb{R}})$. Consequently, LDC applies, and hence

$$\lim_{\lambda\to\infty} \int_{-\lambda/2\pi}^{\lambda/2\pi} \frac{2\pi|\gamma|}{\lambda}|\widehat{f}(\gamma)|\,d\gamma = 0.$$

From the definition of $L^1(\widehat{\mathbb{R}})$, the first integral on the right side of (1.7.10) also tends to 0 as $\lambda \to \infty$. Therefore,

(1.7.11) $$\lim_{\lambda\to\infty} \left\| g(t) - \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\gamma|}{\lambda}\right)\widehat{f}(\gamma)e^{2\pi it\gamma} \right\|_{L^\infty(\mathbb{R})} = 0.$$

We now invoke *Theorem 1.6.9b*, viz.,

(1.7.12) $$\lim_{\lambda\to\infty} \left\| f(t) - \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\gamma|}{\lambda}\right)\widehat{f}(\gamma)e^{2\pi it\gamma} \right\|_{L^1(\mathbb{R})} = 0,$$

to obtain a pointwise a.e. result in the following way. Equation (1.7.12) implies that

(1.7.13)
$$\lim_{\lambda\to\infty}\int_{-\lambda/2\pi}^{\lambda/2\pi}\left(1-\frac{2\pi|\gamma|}{\lambda}\right)\widehat{f}(\gamma)e^{2\pi it\gamma}\,d\gamma = f(t) \text{ in measure.}$$

(By definition, $\lim_{\lambda\to\infty} f_\lambda = f$ *in measure* if

$$\forall\epsilon>0,\quad \lim_{\lambda\to\infty}|\{t:|f_\lambda(t)-f(t)|\geq\epsilon\}|=0,$$

e.g., [Ben76, page 89].) A basic result due to F. Riesz is that *convergence in measure implies convergence a.e. of a subsequence*, e.g., [Ben76, page 106] and *Example A.11*, cf., [RN55, page 100] for a motivating footnote from the master. Thus, (1.7.13) can be changed to convergence a.e. for some λ_n instead of λ. Combined with (1.7.11), for λ_n instead of λ, this adjustment of (1.7.13) yields the fact that $f = g$ a.e.; and the result follows since f and g are continuous. □

1.7.9 Remark. THE LEBESGUE SET

As remarked in the proof of *Theorem 1.7.8*, if $f \in L^1(\mathbb{R})$ then there is $\{\lambda_n\}\subseteq(0,\infty)$ such that

$$\text{(1.7.14)}\qquad \lim_{\lambda_n\to\infty}\int_{-\lambda_n/2\pi}^{\lambda_n/2\pi}\left(1-\frac{2\pi|\gamma|}{\lambda_n}\right)\widehat{f}(\gamma)e^{2\pi it\gamma}\,d\gamma = f(t) \text{ a.e.}$$

It turns out that λ_n can be replaced by λ in (1.7.14), and that the convergence a.e. can be enlarged to include all t in the *Lebesgue set* for f, e.g., [Gol61, pages 14-16]. The *Lebesgue set* L for $f \in L^1_{\text{loc}}(\mathbb{R})$ is the largest set of points $t\in\mathbb{R}$ for which

$$\lim_{h\to0}\frac{1}{h}\int_0^h|f(t+u)-f(t)|\,du=0.$$

Refining FTCI we see that $|\mathbb{R}\backslash L|=0$ and that L includes all points of continuity of f, e.g., [Ben76, Section 4.4].

1.7.10 Example. COMPUTATIONS WITH THE INVERSION FORMULA

a. Using the Jordan Theorem we have

(1.7.15)
$$\frac{1}{\pi}\int_0^\infty \frac{\cos 2\pi t\gamma + \gamma \sin 2\pi t\gamma}{1+\gamma^2}\,d\gamma = \begin{cases} 0, & \text{if } t < 0, \\ \frac{1}{2}, & \text{if } t = 0, \\ e^{-2\pi t}, & \text{if } t > 0. \end{cases}$$

To see this, let $f(t) = e^{-2\pi t} H(t)$, and note that $f \in L^1(\mathbb{R})$, and that $f \in BV(I)$, not only for bounded intervals I but also for $\mathbb{R}$. Thus, by *Theorem 1.7.6*,

(1.7.16)
$$\forall t \in \mathbb{R}, \quad \frac{f(t+) + f(t-)}{2} = \lim_{\Omega\to\infty} \frac{1}{2\pi}\int_{-\Omega}^{\Omega} \frac{e^{2\pi i t\gamma}}{1+i\gamma}\,d\gamma,$$

since $\widehat{f}(\gamma) = \frac{1}{2\pi(1+i\gamma)}$. Clearly, the left side of (1.7.16) is the right side of (1.7.15), so that it remains to verify that

(1.7.17)
$$\frac{1}{2\pi}\int_0^\infty \frac{\cos 2\pi t\gamma + \gamma \sin 2\pi t\gamma}{1+\gamma^2}\,d\gamma = \lim_{\Omega\to\infty}\int_{-\Omega}^{\Omega} \frac{e^{2\pi i t\gamma}}{2\pi(1+i\gamma)}\,d\gamma.$$

We have

$$\frac{e^{2\pi i t\gamma}}{1+i\gamma} = \frac{\cos 2\pi t\gamma + \gamma\sin 2\pi t\gamma + i(\sin 2\pi t\gamma - \gamma\cos 2\pi t\gamma)}{1+\gamma^2};$$

the imaginary part is odd so that the integral on the right side of (1.7.17) is

$$\frac{1}{2\pi}\int_{-\Omega}^{\Omega} \frac{\cos 2\pi t\gamma + \gamma\sin 2\pi t\gamma}{1+\gamma^2}\,d\gamma,$$

and this yields (1.7.17) since the integrand is even.

b. Similarly, we can use either *Theorem 1.7.6* or *Theorem 1.7.8* to verify that

$$\frac{2r}{\pi}\int_0^\infty \frac{\cos 2\pi t\gamma}{r^2+\gamma^2}\,d\gamma = e^{-2\pi r|t|},$$

e.g., *Exercise 1.23a*, cf., (1.7.15).

1.8 Partial differential equations

Harmonic analysis and partial differential equations (PDE) have a profound relationship and interaction, e.g., [Hör83]; and the subject of harmonic analysis owes its existence and formative years to PDE, e.g., [Fou1822]. As a result, the literature on harmonic analysis *and* PDE is extraordinarily extensive, e.g., [BMc66], [CB78], [Dav85], [Gus87], [Wei65] for some elementary books. For this reason, and because of our limitations (both space-time and neuronal) and goals, e.g., *Preface*, we have only selected the following small collection of PDEs. These give a flavor of the aformentioned interaction, but perhaps so little as to be misleading!

1.8.1 Example. A Diffusion Equation

Consider the heat equation with convection,

$$\frac{\partial u}{\partial t} = \frac{k}{4\pi^2}\frac{\partial^2 u}{\partial x^2} + \frac{c}{2\pi i}\frac{\partial u}{\partial x}, \tag{1.8.1}$$

with initial and boundary conditions,

$$u(x,0) = f(x) \tag{1.8.2}$$

and

$$\forall\, t \geq 0, \quad u(\pm\infty, t) = u_x(\pm\infty, t) = 0, \tag{1.8.3}$$

respectively. $u(x,t)$ represents the temperature at x when the time is t. The domain of u consists of $\mathbb{R}\times[0,\infty)$ and $f(x)$ is the given temperature on $\mathbb{R}$ when the process begins at time $t = 0$. The term $\frac{c}{2\pi i}\frac{\partial u}{\partial x}$ is the convection term, and the constants satisfy $k > 0$ and $c \in \mathbb{R}$. We shall compute $u(x,t)$.

Formally, we assume there is a solution $u(\cdot,t) \in L^1(\mathbb{R})$ for each $t \geq 0$ and we let $U(\cdot,\gamma)$ be its Fourier transform, i.e., for each $t \geq 0$, we have the pairing $u(x,t) \leftrightarrow U(\gamma,t)$. We also assume $f, \widehat{f} \in L^1$. Taking Fourier transforms of the functions in equation (1.8.1) we obtain

$$\begin{aligned} U_t(\gamma,t) &= \frac{k}{4\pi^2}\int u_{xx}(x,t)e^{-2\pi i x\gamma}\,dx + \frac{c}{2\pi i}\int u_x(x,t)e^{-2\pi i x\gamma}\,dx \\ &= \frac{k}{4\pi^2}\left[u_x(x,t)e^{-2\pi i x\gamma}\Big|_{-\infty}^{\infty} + 2\pi i\gamma\int u_x(x,t)e^{-2\pi i x\gamma}\,dx\right] \\ &\quad + \frac{c}{2\pi i}\left[u(x,t)e^{-2\pi i x\gamma}\Big|_{-\infty}^{\infty} + 2\pi i\gamma\int u(x,t)e^{-2\pi i x\gamma}\,dx\right] \\ &= \frac{2\pi i\gamma k}{4\pi^2}\left[u(x,t)e^{-2\pi i x\gamma}\Big|_{-\infty}^{\infty} + 2\pi i\gamma\int u(x,t)e^{-2\pi i x\gamma}\,dx\right] \\ &\qquad + c\gamma U(\gamma,t) = (-k\gamma^2 + c\gamma)U(\gamma,t), \end{aligned}$$

where we have used (1.8.3) and assumed $u_x(\cdot,t), u_{xx}(\cdot,t) \in L^1(\mathbb{R})$ for each $t > 0$. Consequently, for each fixed $\gamma \in \widehat{\mathbb{R}}$ we have the ordinary differential equation

$$U_t(\gamma,t) = (-k\gamma^2 + c\gamma)U(\gamma,t), \quad t > 0;$$

and, as in *Example 1.3.3*, we can solve it by elementary methods to obtain

$$U(\gamma,t) = C(\gamma)\exp([-k\gamma^2 + c\gamma]\,t). \tag{1.8.4}$$

Hence, $U(\gamma,0) = C(\gamma)$ so that, since $U(\gamma,0) = \widehat{f}(\gamma)$, we can write

$$U(\gamma,t) = \widehat{f}(\gamma)\exp([-k\gamma^2 + c\gamma]\,t) \tag{1.8.5}$$

for all $(\gamma,t) \in \widehat{\mathbb{R}} \times (0,\infty)$. (At this stage we do not choose to be careful about letting $t = 0$ in (1.8.4).) By completing the square, (1.8.5) becomes

$$U(\gamma,t) = \exp\Big(\frac{tc^2}{4k}\Big)\widehat{f}(\gamma)\exp\Big(-tk(\gamma - \frac{c}{2k})^2\Big) \tag{1.8.6}$$

for all $(\gamma,t) \in \widehat{\mathbb{R}} \times (0,\infty)$, and, in particular, $\widehat{f}(\gamma)\exp(-tk(\gamma - \frac{c}{2k})^2) \in L^1(\widehat{\mathbb{R}})$ for each fixed $t > 0$. Thus, we can apply the inversion theorem,

Theorem 1.7.8, to obtain our solution

$$(1.8.7)\qquad u(x,t) = \exp\left(\frac{tc^2}{4k}\right)\left[f(y) * \left(e^{\pi iyc/k}\, g_{\pi/(tk)^{1/2}}(y)\right)\right](x)$$

by taking the inverse Fourier transform of (1.8.6) for each $t > 0$.

The calculation of (1.8.7) depends on the convolution formula, *Proposition 1.5.2*, and the fact that

$$\begin{aligned}\int e^{-tk(\gamma-\frac{c}{2k})^2}\, e^{2\pi iy\gamma}\, d\gamma &= e^{\pi iyc/k}\int e^{-(tk)\lambda^2} e^{2\pi iy\lambda}\, d\lambda\\ &= e^{\pi iyc/k}\, g_{\pi/(tk)^{1/2}}(y).\end{aligned}$$

We now have to check that $u(x,t)$ in (1.8.7) is really a solution of the system (1.8.1)–(1.8.3). We leave this as *Exercise 1.39.* Technically, we have no right to begin with a function u, as we did, unless we had available an existence theorem, which, in fact, does exist (sic).

1.8.2 Example. A DIRICHLET PROBLEM

a. Consider Laplace's equation

$$(1.8.8)\qquad \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

on the upper half-plane $\mathbb{R} \times [0,\infty)$ with boundary condition,

$$(1.8.9)\qquad u(x,0) = f(x),$$

where $f(x)$ is a given function. The *Dirichlet problem* is to determine whether or not the system (1.8.8) and (1.8.9) has a unique solution in the upper half-plane.

To focus on the problem mathematically, we assume equation (1.8.8) is valid on $\mathbb{R} \times (0,\infty)$. Other natural assumptions will arise as we proceed with the calculation.

Physically, the system (1.8.8) and (1.8.9) models a steady state temperature distribution problem. "Steady state" indicates that the average temperature doesn't change with time, i.e., the rate at which heat flows into the upper half-plane is 0. This should be compared with *Example 1.8.1* which is not a steady-state problem as reflected by the presence of the term "$\frac{\partial u}{\partial t}$" on the left hand side of (1.8.1). (Lest there be any confusion, *Example 1.8.1* is a one-dimensional temperature distribution problem, and this example is 2-dimensional.) Condition (1.8.9) indicates a known temperature distribution along the boundary, and the Dirichlet problem is to determine if a steady state system with a known temperature distribution on the boundary, e.g., predictable radiators along the walls of a room, characterizes the temperature in the interior.

Formally, we assume there is a solution $u(x,y)$ and that $u(\cdot,y) \in L^1(\mathbb{R})$ for each $y \geq 0$. If $U(\cdot,y)$ is the Fourier transform of $u(\cdot,y)$ then (1.8.8) yields

$$\forall \gamma \in \widehat{\mathbb{R}}, \quad -4\pi^2 \gamma^2 U(\gamma,y) + \frac{d^2}{dy^2} U(\gamma,y) = 0. \tag{1.8.10}$$

For each fixed $\gamma \in \widehat{\mathbb{R}}$ we view (1.8.10) as an ordinary differential equation in y. The corresponding characteristic equation is $r^2 - 4\pi^2\gamma^2 = 0$ so that $r = \pm 2\pi|\gamma|$. Hence, we have

$$U(\gamma,y) = a(\gamma)e^{2\pi y|\gamma|} + b(\gamma)e^{-2\pi y|\gamma|}.$$

Now, we refine the assumption, $u(\cdot,y) \in L^1(\mathbb{R})$, to include the estimate,

$$\exists M \quad \text{such that} \quad \forall y \geq 0, \quad \int |u(x,y)|\, dx \leq M.$$

As such we see that $a(\gamma) = 0$ for $\gamma \neq 0$ since

$$|U(\gamma,y)| \leq \int |u(x,y)|\, dx \leq M.$$

Thus, we obtain

$$U(\gamma,y) = b(\gamma)e^{-2\pi y|\gamma|}.$$

Formally, we have

$$b(\gamma) = \widehat{f}(\gamma) \tag{1.8.11}$$

and

$$u(x,y) = \int \frac{y f(u)\, du}{\pi(y^2 + (x-u)^2)}. \tag{1.8.12}$$

In fact, (1.8.11) is clear by (1.8.9) and the equation for U; and equation (1.8.12) follows by the calculation

$$u(x,y) = [f * (e^{-2\pi y|\gamma|})^{\vee}](x) = f * p_{1/y}(x),$$

which is valid by *Theorem 1.7.8* when $f, \hat{f} \in L^1$.

The right hand side of (1.8.12) is the *Poisson integral formula for the upper half-plane.* It is easy to check that $u(x,y)$ so defined is a solution of Laplace's equation on $\mathbb{R} \times (0,\infty)$. Further, since $f \in L^1(\mathbb{R})$ and $\{p_\lambda\}$ is an approximate identity, we know that

$$\lim_{y \to 0} u(x,y) = f(x)$$

in L^1-norm. In this sense we have solved the Dirichlet problem as far as obtaining a solution.

The uniqueness of this solution is intuitively clear for the following reason. Suppose $f = 0$ on $\mathbb{R}$, the boundary of the upper half-plane. By the definition of a steady state system, the temperature $u(x,y)$ would also have to be 0 since the temperature of the upper-half plane is not influenced by any other heat flow. Of course, there is a highly developed mathematical theory of uniqueness for solution of partial differential equations, e.g., [Hör83].

1.8.3 Example. A DIFFUSION EQUATION AND IMAGE PROCESSING

Combining the two-dimensionality of *Example 1.8.2* with the time dependence of *Example 1.8.1*, let us consider the diffusion equation

$$\Delta u = \frac{\partial u}{\partial s} \tag{1.8.13}$$

on $\mathbb{R}^2 \times [0,\infty)$ with initial condition,

$$u(x,y,0) = f(x,y), \tag{1.8.14}$$

where Δ is defined by (1.8.8), f is a given function, and $(x,y,s) \in \mathbb{R}^2 \times [0,\infty)$. Since such problems are a well established part of partial differential equations and classical physics, we choose to look at

(1.8.13) and (1.8.14) in terms of a more recent interpretation in *image processing.*

Let

$$u_g(x,y,s) = g_{\sqrt{\frac{1}{4s}}}(x)g_{\sqrt{\frac{1}{4s}}}(y). \tag{1.8.15}$$

We are using "s" instead of "t" in the dilations of the Gaussian on $\mathbb{R}$ to denote scale instead of time. Note that if we define the function k on $\mathbb{R}^2$ as

$$k(x,y) = g(x)g(y)$$

then

$$k_\lambda(x,y) = \lambda^2\, k(\lambda x, \lambda y), \quad \lambda > 0$$

is the $L^1(\mathbb{R}^2)$-dilation of k by λ; and $\{k_\lambda\}$ is an approximate identity for $L^1(\mathbb{R}^2)$ as $\lambda \to \infty$. By (1.8.15), we see that u_g is the dilation

$$u_g(x,y,s) = k_{\sqrt{\frac{1}{4s}}}(x,y).$$

Further, for this u we have

$$\Delta u_g(x,y,s) = \frac{\partial u_g}{\partial s}, \tag{1.8.16}$$

that is

$$\Delta k_{\sqrt{\frac{1}{4s}}}(x,y) = \frac{\partial}{\partial s} k_{\sqrt{\frac{1}{4s}}}(x,y).$$

It is easy to see that $u = u_g * f$ is a solution of the system (1.8.13) and (1.8.14) when f satisfies natural hypotheses. If f is an *image* to be processed then $u(\cdot,\cdot,s)$ is a blurred image of f, more blurred for larger s. An idea in image processing is to reconstruct f from blurred versions of f, which may not require too much data or "expense", along with available detailed data. This is associated with the notion of multiresolution (in wavelet theory) due to Mallat [Mall89] and Meyer [Mey90], and the discrete version of this scheme due to Burt and Adelson [BA83].

The convolution $u_g * f$ is, in engineering terms, *filtering the image f by the Gaussian u_g*. Now, it also makes sense to filter the image f by Δu_g for the following reason [KJ55]. An elementary calculation shows that $\iint \Delta u_g(x,y,s)\,dxdy = 0$. For small s, we can think of Δu_g as being concentrated in a small region; and so if the image f is constant in a region R of comparable size (or larger), then $(\Delta u_g) * f(x,y) = 0$ for (x,y) near the center of R. Similarly, if f is not constant on R then this convolution procedure can have the effect of detecting *edges* in the image.

Thus, since filtering by Δu_g is potentially important, we see the value of (1.8.13) in image analysis for the following reason. Because of (1.8.16) we can estimate Δu_g by considering a scaled difference of Gaussians, thereby reducing computational complexity, e.g., [HM89].

1.9 Gibbs phenomenon

1.9.1 Remark. Description of Gibbs Phenomenon

Let f be a function on $\mathbb{R}$ and suppose f is continuous on

$$I = [t_0 - T, t_0 + T],\ T > 0,$$

except for a jump discontinuity at t_0, i.e., $f(t_0+)$ and $f(t_0-)$ exist and $f(t_0+) - f(t_0-) \neq 0$.

In the case $f \in L^1(\mathbb{R}) \cap BV(I)$, we showed in Jordan's Theorem *(Theorem 1.7.6)* that

$$\lim_{\Omega\to\infty} f * d_{2\pi\Omega}(t_0) = \frac{f(t_0+) + f(t_0-)}{2}. \tag{1.9.1}$$

We shall now investigate this limit more closely and shall detect a remarkable behavior of the "partial sums" $S_\Omega = f * d_{2\pi\Omega}$. This behavior is called *Gibbs phenomenon*.

To fix ideas let $f = H$, the Heaviside function. Even though $H \notin L^1(\mathbb{R})$, the method of Jordan's Theorem is valid; and the partial sums $S_\Omega \equiv H * d_{2\pi\Omega}$ exist and satisfy

$$\forall t \in \mathbb{R},\quad \lim_{\Omega\to\infty} S_\Omega(t) = \frac{H(t+) + H(t-)}{2}.$$

In fact,

$$S_\Omega = H * d_{2\pi\Omega}(t) = \int_0^\infty \frac{\sin 2\pi\Omega(t-u)}{\pi(t-u)}\,du$$

(1.9.2)

$$= \int_{-\infty}^{2\pi\Omega t} \frac{\sin u}{\pi u}\,du = \frac{1}{2} + \int_0^{2\pi\Omega t} \frac{\sin u}{\pi u}\,du,$$

noting that t can be negative. Clearly, for a fixed Ω, the last integral achieves its maximum (on $[0,\infty)$, say) at $t = 1/(2\Omega)$ since for larger t the integrand alternates between "decreasing" negative and positive values. Thus, for each Ω, $H * d_{2\pi\Omega}$ is maximized at $t = 1/(2\Omega)$, and, similarly, is minimized at $t = -1/(2\Omega)$. The values of $H * d_{2\pi\Omega}$ at these points are the two *constants*,

$$G = H * d_{2\pi\Omega}(\frac{1}{2\Omega}) = \frac{1}{2} + \int_0^\pi \frac{\sin u}{\pi u}\,du > 1$$

and

$$H * d_{2\pi\Omega}(\frac{-1}{2\Omega}) = \frac{1}{2} - \int_0^\pi \frac{\sin u}{\pi u}\,du < 0.$$

These two inequalities are clear from our knowledge of the Dirichlet function d.

Figure 1.11

The fact that the convergence of $\{H * d_{2\pi\Omega}\}$ to H on $\mathbb{R}\backslash\{0\}$ involves the intrinsic "overshoot" G (and the corresponding behavior on the negative axis) is the *Gibbs phenomenon*, e.g., *Figure 1.11.* This pointwise convergence is uniform on closed bounded subintervals $K \subseteq \mathbb{R}\backslash\{0\}$; but the behavior of the "partial sums" $H * d_{2\pi\Omega}$ near the jump discontinuity always exhibits a fixed rise $G > H(t) = 1$ at $t = \frac{1}{(2\Omega)}$ as Ω increases to infinity.

Note that

$$\forall \Omega > 0, \quad H * d_{2\pi\Omega}(0) = \frac{1}{2}.$$

1.9.2 Remark. Historical Note

All of the early work on Gibbs phenomenon was in the context of Fourier series, e.g., [Car30], [HH79]. The term "Gibbs phenomenon" is due to Bôcher (1906), who also provided a proof of Gibbs' original assertion, cf., [HH79, pages 155-156] for a later less-than-civilized development.

Apparently, Henry Wilbraham (1848) was the first to understand the presence of the overshoot G in the Fourier analysis of functions having jump discontinuities. Wilbraham dealt with the function

$$\text{(1.9.3)} \qquad \frac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k\frac{\cos(2k+1)t}{2k+1} = \begin{cases} 1, & \text{if } |t| < \frac{1}{2}\pi, \\ -1, & \text{if } \frac{1}{2}\pi < |t| < \pi, \end{cases}$$

defined 2π-periodically on $\mathbb{R}$, e.g., *Definition 3.1.1.* Without knowing of Wilbraham's work, Gibbs made his fundamental contribution to the topic on April 27, 1899, as part of a lively interchange in *Nature (Volumes 58-60, 1898-1899)* initiated by Michelson, and also involving Love and Poincaré. They dealt with the function

$$\text{(1.9.4)} \qquad \forall\, t \in (-\pi, \pi), \quad 2\sum_{k=1}^{\infty}(-1)^{k+1}\frac{\sin kt}{k} = t,$$

defined 2π-periodically on $\mathbb{R}$.

For perspective, it is interesting to note that in 1898 Michelson and Stratton designed their *harmonic analyzer*, complete with graphs of partial sums of the series in (1.9.3) and (1.9.4). This work probably inspired Michelson's letter to *Nature*. Michelson and Stratton had collaborated on X-ray research; and Stratton was the founder of the National Bureau of Standards. Their harmonic analyzer machine was used for the decomposition of sound and electrical waves into simpler components, cf., our remarks on *harmonic analysis* in the *Preface*. Harmonic analyzers *and* synthesizers were first designed by Lord Kelvin, cf., *Exercise 1.47*, for tidal analysis.

The study of tides received its scientific basis with Newton's law of gravitation for the sun, moon, and earth – even the ancient mariner knew that tidal phenomena were related to astronomical factors. Then Laplace was able to separate various cyclic influences of the sun and moon (on tides) by defining a model of the sun, moon, and earth having a number of tide-affecting satellites; and the Newtonian solution of this model associated an elementary tidal constituent with each satellite in such a way that the tide was viewed as a combination of these constituents. Lord Kelvin systematized this method, and the analyzing and synthesizing machines he designed were meant to determine the constituents and to reconstruct the tide (from the constituents), respecitvely, e.g., [God72], [Mac66]. These mechanical analyzers and synthesizers may be on view at your local science museum!

Instead of the standard partial sum, as in (1.9.2), it is possible to consider functions f having a jump discontinuity in the case that an approximate identity $\{k_{(\lambda)}\}$ replaces the Dirichlet kernel $\{d_{2\pi\Omega}\}$, e.g., *Theorem 1.6.9.* In this situation, if each $k_{(\lambda)}$ is nonnegative and even then $H * k_{(\lambda)}(0) = \frac{1}{2}$ and

$$\forall t \in \mathbb{R}, \quad 0 \leq H * k_{(\lambda)}(t) \leq 1,$$

i.e., *the Gibbs phenomenon is eliminated for nonnegative approximate identities.*

Using the Gaussian approximate identity, Weyl (1910) studied what he called *heat conduction partial sums*,

$$f * g_{\pi/(tk)^{1/2}}.$$

From our discussion in *Example 1.8.1*, these functions are solutions of the *heat equation*, i.e., the convection equation (1.8.1) for $c = 0$, cf., *Exercise 1.47.* Weyl was interested in a two temperature boundary value function f, e.g., f is either $\alpha = 100°C$ or $\beta = 0°C$ [Wey50b],and the corresponding eigenfunction expansion, cf., *Exercise 1.27.* He made note of the Gibbs phenomenon for such expansions; and had to deal with rational approximation of irrational numbers (at specified rates) to complete his analysis for arbitrary α and β. Such approximations are part of Diophantine Approximation, a branch of number theory. Later, because of work by P. Bohl and the problem of mean motion in Lagrange's linear theory of perturbation for the planetary system, Weyl (1916) generalized the aforementioned approximation procedure; and gave his definition and characterization of *equidistribution mod 1*, e.g., *Exercise 3.40*, [HW65], [KK64], [Sal63]. This characterization, the *Weyl Equidistribution Theorem*, is not only an important result in number theory, but has a host of applications in a variety of fields including Wiener's Generalized Harmonic Analysis, e.g., [Bas84], [KK64, pages 23-28, by J.-P. Bertrandias], cf., *Section 2.9.* Our point in highlighting Weyl's work is to trace the intellectual excursion relating Gibbs phenomenon to heat conduction problems to number theory.

1.9.3 Example. COMPUTATION OF G

We can estimate G by expanding $\sin u$ in a Taylor series, noting that $\pi \in (3.141, 3.142)$ and estimating the integral $\int_0^\pi \frac{\sin u}{\pi u}\, du$, e.g., *Exercise 1.29.* In fact,

$$G = \frac{1}{2} + \frac{1}{\pi}\int_0^\pi \frac{\sin u}{u}\, du$$

$$= \frac{1}{2} + \frac{1}{\pi}\int_0^\pi \frac{1}{u}\sum_{n=0}^{\infty}(-1)^n \frac{u^{2n+1}}{(2n+1)!}\, du$$

$$= \frac{1}{2} + \frac{1}{\pi}\sum_{n=0}^{\infty}(-1)^n \frac{\pi^{2n+1}}{(2n+1)(2n+1)!}$$

$$= \frac{1}{2} + \left[1 - \frac{\pi^2}{3\cdot 3!} + \frac{\pi^4}{5\cdot 5!} - \frac{\pi^6}{7\cdot 7!} + \frac{\pi^8}{9\cdot 9!} - \cdots\right].$$

1.9.4 Example. PERSPECTIVE ON BANDLIMITED APPROXIMANTS

Let f be a function and let $\Omega > 0$. A basic approximation problem, associated with the notion of *aliasing* in engineering, is to approximate f in a realistic way by an $\Omega - BL$ function, this latter notion having been defined in *Example 1.2.2.* Of course, the *criterion* for approximation must be specified.

Suppose the criterion is pointwise convergence. Then our discussion of the Gibbs phenomenon shows that, because of the overshoot and undershoot, the natural $\Omega - BL$ approximant $f * d_{2\pi\Omega}$ of f is poor in a neighborhood of any jump discontinuity of f, even when the band $[-\Omega, \Omega]$ is large. As we saw in *Remark 1.9.2* the overshoot and undershoot are obviated when dealing with the $\Omega - BL$ approximant $f * w_{2\pi\Omega}$, cf., *Remark 1.10.8*, *Proposition 1.10.9*, and *Remark 1.10.10.*

Another pointwise approach involves dealing with the continuous function $f * \left(\frac{1}{2T}\mathbf{1}_{[-T,T)}\right)$, and then defining the $\Omega - BL$ approximant

$$f_{T,\Omega} = f * \left(\frac{1}{2T}\mathbf{1}_{[-T,T)}\right) * d_{2\pi\Omega}.$$

Notice that $f * \left(\frac{1}{2T}\mathbf{1}_{[-T,T)}\right)(t)$ is the average value,

$$\frac{1}{2T}\int_{t-T}^{t+T} f(u)\, du,$$

of f on $[t-T, t+T]$. In this case, $f_{T,\Omega}$ can be a good pointwise (or even uniform) $\Omega - BL$ approximant of f if Ω is large and T is small. For example, if $f \in L^1(\mathbb{R}) \cap A(\mathbb{R})$ then

$$\forall t \in \mathbb{R}, \quad |f(t) - f_{T,\Omega}(t)|$$
$$\leq \int_{|\gamma|\geq\Omega} |\widehat{f}(\gamma)|\, d\gamma + \int_{-\Omega}^{\Omega} |\widehat{f}(\gamma)||1 - \frac{\sin 2\pi T\gamma}{2\pi T\gamma}|\, d\gamma.$$

An important case of our approximation problem deals with an approximation criterion motivated by physical considerations such as variance and energy. The mathematical setting is $L^2(\mathbb{R})$, the space of square integrable functions.

1.10 The $L^2(\mathbb{R})$ theory

1.10.1 Definition. $L^2(\mathbb{R})$

$$L^2(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{C} : \|f\|_{L^2(\mathbb{R})} = \left(\int |f(t)|^2 \, dt \right)^{1/2} < \infty \right\}.$$

$L^2(\mathbb{R})$ is the space of *square-integrable functions* f or *signals* f *having finite energy* $\|f\|^2_{L^2(\mathbb{R})}$.

The major result about $L^2(\mathbb{R})$ is the following theorem [Pla10], [Pla15].

1.10.2 Theorem. Plancherel Theorem

There is a unique linear bijection $\mathcal{F} : L^2(\mathbb{R}) \longrightarrow L^2(\widehat{\mathbb{R}})$ *with the properties*

a. $\forall f \in L^2(\mathbb{R}), \quad \|f\|_{L^2(\mathbb{R})} = \|\mathcal{F}f\|_{L^2(\widehat{\mathbb{R}})}$,

b. $\forall f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ *and* $\forall \gamma \in \widehat{\mathbb{R}}, \quad \widehat{f}(\gamma) = (\mathcal{F}f)(\gamma)$,

c. $\forall f \in L^2(\mathbb{R}), \exists \{f_n : n = 1, \cdots\} \subseteq L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ *for which*

$$\lim_{n\to\infty} \|f_n - f\|_{L^2(\mathbb{R})} = 0 \text{ and } \lim_{n\to\infty} \|\widehat{f_n} - \mathcal{F}f\|_{L^2(\widehat{\mathbb{R}})} = 0.$$

Proof. Our outline of proof, which is one of the usual schemes to prove the Plancherel Theorem, has four steps: verification that $\|f\|_{L^2(\mathbb{R})} = \|\widehat{f}\|_{L^2(\widehat{\mathbb{R}})}$ for $f \in X \subseteq L^2(\mathbb{R})$ (part *i*), closure results on $\mathbb{R}$ and $\widehat{\mathbb{R}}$ (parts *ii* and *iii*, respectively), and a routine functional analysis argument to obtain the result from parts *i, ii, iii* (part *iv*). There are trade-offs in difficulty between parts *i, ii,* and *iii* depending on which space X one chooses to use.

i. Let $X = C_c(\mathbb{R})$, the space of continuous functions (on $\mathbb{R}$) having compact support, and consider the involution $\tilde{f}(t) \equiv \bar{f}(-t)$ of $f \in C_c(\mathbb{R})$. We shall prove that $\widehat{f} \in L^2(\widehat{\mathbb{R}})$. Clearly, $\widehat{f} \in A(\widehat{\mathbb{R}})$ since $C_c(\mathbb{R}) \subseteq L^1(\mathbb{R})$. Define $g = f * \tilde{f}$ so that g is continuous, $g \in L^1(\mathbb{R})$, e.g., *Definition 1.5.1* or by a direct calculation, and

$$g(0) = \|f\|^2_{L^2(\mathbb{R})}. \tag{1.10.1}$$

Also,

$$\forall \gamma \in \widehat{\mathbb{R}}, \quad \widehat{g}(\gamma) = |\widehat{f}(\gamma)|^2 \tag{1.10.2}$$

by the Fubini-Tonelli Theorem and the translation invariance of Lebesgue measure (on the group $\mathbb{R}$). By *Proposition 1.6.11* and (1.10.2), since $g \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ is continuous, we have

$$\lim_{\lambda \to \infty} \int_{-\lambda/2\pi}^{\lambda/2\pi} \left(1 - \frac{2\pi|\gamma|}{\lambda}\right) |\widehat{f}(\gamma)|^2 \, d\gamma = g(0). \tag{1.10.3}$$

The Beppo Levi Theorem and (1.10.3) allow us to assert that $\widehat{f} \in L^2(\widehat{\mathbb{R}})$ and that

$$\|\widehat{f}\|^2_{L^2(\widehat{\mathbb{R}})} = g(0) = \|f\|_{L^2(\widehat{\mathbb{R}})},$$

where the second equality follows from the definition of g.

ii. We shall now note that $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, and prove that $C_c(\mathbb{R})$ is dense in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ taken with the norm $\| \cdots \| = \| \cdots \|_{L^1(\mathbb{R})} + \| \cdots \|_{L^2(\mathbb{R})}$. These facts imply $\overline{C_c(\mathbb{R})} = L^2(\mathbb{R})$.

Let $f \in L^2(\mathbb{R})$ and define $f_T = f \mathbf{1}_{[-T,T]}$. For this proof, f_T does not designate dilation. $f_T \in L^2(\mathbb{R})$ since $|f_T| \leq |f|$, and $f_T \in L^1(\mathbb{R})$ by Hölder's Inequality. Clearly,

$$\|f - f_T\|_{L^2(\mathbb{R})} = \left(\int_{|t|>T} |f(t)|^2 \, dt \right)^{1/2},$$

and $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$.

Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then $\lim_{T \to \infty} \|f - f_T\| = 0$ by the argument of the previous paragraph. Next, set $f_{S,T} = f_T * \Delta_S$ (Δ_S is the dilation of the triangle function Δ defined in *Example 1.3.4*), so that $f_{S,T} \in C_c(\mathbb{R})$. Letting $\epsilon > 0$, we shall find $S, T > 0$ such that $\|f - f_{S,T}\| < \epsilon$.

First, there is $T = T_\epsilon$ for which $\|f - f_T\| < \epsilon/2$. We keep this T fixed, and have the estimate

$$\|f_T - f_{S,T}\|_{L^1(\mathbb{R})} \leq \iint \Delta_S(u) |f_T(t) - f_T(t-u)| \, du dt$$
$$= \int_{|u| \leq 1/S} \Delta_S(u) \left(\int |f_T(t) - f_T(t-u)| \, dt \right) du.$$

Next, choose S_1 such that

$$\forall S \geq S_1 \text{ and } \forall |u| \leq 1/S, \quad \|f_T - \tau_u f_T\|_{L^1(\mathbb{R})} < \epsilon/4.$$

Thus, for such S,

$$\|f_T - f_{S,T}\|_{L^1(\mathbb{R})} < \epsilon/4. \tag{1.10.4}$$

Finally, we use Minkowski's Inequality (*Theorem A.16*) to obtain the estimate

$$\begin{aligned}\|f_T - f_{S,T}\|_{L^2(\mathbb{R})} &\leq \int \left(\int \left| \Big(f_T(t) - f_T(t-u) \Big) \Delta_S(u) \right|^2 dt \right)^{1/2} du \\ &= \int \Delta_S(u) \|f_T - \tau_u f_T\|_{L^2(\mathbb{R})}\, du.\end{aligned}$$

We can choose $S \geq S_1$ for which

$$\forall |u| \leq 1/S, \quad \|f_T - \tau_u f_T\|_{L^2(\mathbb{R})} < \epsilon/4;$$

and, therefore,

$$\|f_T - f_{S,T}\|_{L^2(\mathbb{R})} < \epsilon/4. \tag{1.10.5}$$

Combining (1.10.4) and (1.10.5) with the above estimate $\|f - f_T\| < \epsilon/2$, we have the desired inequality, viz., $\|f - f_{S,T}\| < \epsilon$.

iii. We shall now prove that $C_c(\mathbb{R})^\wedge \subseteq A(\widehat{\mathbb{R}}) \cap L^2(\widehat{\mathbb{R}})$ is a dense subspace of $L^2(\widehat{\mathbb{R}})$.

Let $G \in L^2(\widehat{\mathbb{R}})$ and suppose

$$\forall\, f \in C_c(\mathbb{R}), \quad \int \widehat{f}(\gamma)\overline{G(\gamma)}\, d\gamma = 0. \tag{1.10.6}$$

If $f \in C_c(\mathbb{R})$ then $\tau_u f \in C_c(\mathbb{R})$ and so (1.10.6) implies

$$\forall f \in C_c(\mathbb{R}) \text{ and } \forall u \in \mathbb{R}, \quad \int \widehat{f}(\gamma)\overline{G(\gamma)} e^{-2\pi i u \gamma}\, d\gamma = 0. \tag{1.10.7}$$

By Hölder's Inequality, $\widehat{f}\overline{G} \in L^1(\widehat{\mathbb{R}})$, and so (1.10.7) allows us to invoke the uniqueness theorem, *Theorem 1.6.9c*, to conclude that $\widehat{f}\overline{G} = 0$ a.e. for each $f \in C_c(\mathbb{R})$.

Note that

$$\forall f \in C_c(\mathbb{R}) \text{ and } \forall \gamma \in \widehat{\mathbb{R}}, \quad e^{2\pi i t\gamma} f(t) \in C_c(\mathbb{R}).$$

Thus, $C_c(\mathbb{R})^\wedge$ is translation invariant, i.e.,

$$\forall f \in C_c(\mathbb{R}) \text{ and } \forall \gamma \in \widehat{\mathbb{R}}, \quad \tau_\gamma \widehat{f} \in C_c(\mathbb{R})^\wedge.$$

From this we can conclude that *for each* $\gamma_0 \in \widehat{\mathbb{R}}$, *there is* $f = f_{\gamma_0} \in C_c(\mathbb{R})$ *for which* $|\widehat{f}| > 0$ *on an interval* I *centered about* γ_0. To verify this claim, suppose there is γ_0 such that for each $f \in C_c(\mathbb{R})$ and for each interval I centered at γ_0, $\widehat{f}$ has a zero in I. Consequently, $\widehat{f}(\gamma_0) = 0$ for each $f \in C_c(\mathbb{R})$. By the translation invariance of $C_c(\mathbb{R})^\wedge$, $\tau_\gamma \widehat{f} \in C_c(\mathbb{R})^\wedge$ for each $\gamma \in \widehat{\mathbb{R}}$, and so

$$\forall f \in C_c(\mathbb{R}) \text{ and } \forall \gamma \in \widehat{\mathbb{R}}, \quad (\tau_\gamma \widehat{f})(\gamma_0) = 0.$$

i.e., $\widehat{f} = 0$ on $\widehat{\mathbb{R}}$ for each $f \in C_c(\mathbb{R})$. This contradicts the uniqueness theorem, *Theorem 1.6.9c* (for just one $f \in C_c(\mathbb{R}) \backslash \{0\}$), and the claim is proved.

Therefore, if we assume (1.10.6) we can conclude that $G = 0$ a.e. Consequently, by the Hahn-Banach Theorem (*Theorem B.12*) and the fact that $L^2(\widehat{\mathbb{R}})' = L^2(\widehat{\mathbb{R}})$ (*Theorem B.14*), we have that $C_c(\mathbb{R})^\wedge$ is dense in $L^2(\widehat{\mathbb{R}})$.

iv. We have shown that $\mathcal{F}$ is a continuous linear injection $C_c(\mathbb{R}) \to L^2(\widehat{\mathbb{R}})$ (part *i*), when $C_c(\mathbb{R})$ is endowed with the L^2-norm, and so $\mathcal{F}$ has a unique linear injective extension to $L^2(\mathbb{R})$ (*Theorem B.7*) by (part *ii*). Also, $\mathcal{F}(C_c(\mathbb{R}))$ is closed and dense in $L^2(\widehat{\mathbb{R}})$ by parts *i* and *iii*. Thus, $\mathcal{F}$ is also surjective. The remaining claims of the theorem are now immediate. □

Notationally, because of the Plancherel Theorem, we refer to $\mathcal{F}f$ as the *Fourier transform* of $f \in L^2(\mathbb{R})$, and we write the pairing between $f \in L^2(\mathbb{R})$ and $\mathcal{F}f$ in one of the following ways:

$$\mathcal{F}f = \widehat{f} = F, \quad f \longleftrightarrow F, \quad f = F^\vee. \tag{1.10.8}$$

1.10.3 Theorem. PARSEVAL FORMULA

Consider $f \leftrightarrow F$ and $g \leftrightarrow G$, where $f, g \in L^2(\mathbb{R})$. Then we have the formulas

$$\|f\|_{L^2(\mathbb{R})} = \|F\|_{L^2(\widehat{\mathbb{R}})}, \tag{1.10.9}$$

$$\int f(t)\overline{g(t)}\,dt = \int F(\gamma)\overline{G(\gamma)}\,d\gamma, \tag{1.10.10}$$

$$\int f(t)g(t)\,dt = \int F(\gamma)G(-\gamma)\,d\gamma, \tag{1.10.11}$$

and

$$\forall \gamma \in \widehat{\mathbb{R}}, \quad \int f(t)g(t)e^{-2\pi i t\gamma}\,dt = \int F(\lambda)G(\gamma-\lambda)\,d\lambda. \tag{1.10.12}$$

Proof. (1.10.9) is part of *Theorem 1.10.2*. (1.10.10) is a consequence of (1.10.9) and the fact that $4f\bar{g} = |f+g|^2 - |f-g|^2 + i|f+ig|^2 - i|f-ig|^2$.

(1.10.11) can be proved similarly or by the following formal calculation, which can be made valid, e.g., *Exercise 1.44*. This calculation actually gives (1.10.12), from which (1.10.11) follows for $\gamma = 0$. Note that $fg \in L^1(\mathbb{R})$ and $F * G \in A(\widehat{\mathbb{R}})$. We compute

$$\begin{aligned}(F*G)^\vee(t) &= \int F*G(\gamma)e^{2\pi i t\gamma}\,d\gamma \\ &= \iint F(\lambda)G(\gamma-\lambda)e^{2\pi i t\gamma}\,d\gamma d\lambda \\ &= g(t)\int F(\lambda)e^{2\pi i t\lambda}\,d\lambda = f(t)g(t),\end{aligned}$$

and, hence,

$$\int f(t)g(t)e^{-2\pi i t\gamma}\,dt = \int F(\lambda)G(\gamma-\lambda)\,d\lambda. \qquad \square$$

We shall refer to *Theorem 1.10.2* and *Proposition 1.10.3* as the *Parseval-Plancherel Theorem*, and we shall refer to equations (1.10.10) or (1.10.11) as *Parseval's formula*. Parseval was a French engineer who gave a formal verification of the Fourier series version of (1.10.9) in 1799; the publication is dated 1805.

The following version of (1.10.10) is required in *Example 2.4.8*.

1.10.4 Proposition.

Consider $f \leftrightarrow F$ and $g \leftrightarrow G$, where $f \in L^1(\mathbb{R})$ and $g \in A(\mathbb{R})$, i.e., $G \in L^1(\widehat{\mathbb{R}})$ and $g(t) = \int G(\gamma)e^{2\pi it\gamma}\,dt$. Then

$$\int f(t)\overline{g(t)}\,dt = \int F(\gamma)\overline{G(\gamma)}\,d\gamma$$

and

$$\int f(t)g(t)\,dt = \int F(\gamma)G(-\gamma)\,d\gamma.$$

Proof. Consider the first formula. The right side is

$$\iint f(t)e^{-2\pi it\gamma}\,\overline{G(\gamma)}\,dtd\gamma$$

$$= \int f(t)\left(\overline{\int G(\gamma)e^{2\pi it\gamma}\,d\gamma}\right)dt = \int f(t)\overline{g(t)}\,dt,$$

where the first equality is a consequence of the Fubini-Tonelli Theorem. □

1.10.5 Example. Powers of the Dirichlet Kernel

Consider the pairing,

$$\frac{\sin t}{t} \longleftrightarrow \pi\mathbf{1}_{[-\frac{1}{2\pi},\frac{1}{2\pi})},$$

noting that $(\sin t)/t \in L^2(\mathbb{R})\backslash(L^2(\mathbb{R}) \cap L^1(\mathbb{R}))$. We can compute

$$\int \frac{\sin^n t}{t^n}\,dt, \quad n \geq 2,$$

by means of the Parseval-Plancherel Theorem.

For example,

$$\int \frac{\sin^3 t}{t^3}\,dt = \frac{3\pi}{4}$$

because of the pairing,

$$\frac{\sin^2 t}{t^2} \longleftrightarrow \pi\max(1 - |\pi\gamma|, 0),$$

and by the Parseval-Plancherel Theorem.

Similarly,

$$\int \frac{\sin^4 t}{t^4}\,dt = \frac{2}{3}\pi,$$

e.g., *Exercise 1.17.*

1.10.6 Example. A FOURIER UNIQUENESS PROPERTY

Do there exist $f \in (L^1(\mathbb{R}) \cap L^2(\mathbb{R}))\backslash\{0\}$ and $a \in \mathbb{C}$ for which

$$\forall t \in \mathbb{R}, \quad af(t)\mathbf{1}_{[-T,T)}(t) = f * d_{2\pi\Omega}(t)? \tag{1.10.13}$$

To answer this question we distinguish two cases, $a = 0$ and $a \neq 0$. Also, we say that a function k is *supported* by $[A, B] \subseteq \mathbb{R}$ if $k = 0$ on $[A, B]^c \equiv \mathbb{R}\backslash[A, B]$; in this case we write $\operatorname{supp} k \subseteq [A, B]$, cf., *Definition 2.2.1a.*

a. If $a = 0$ then $\hat{f}\mathbf{1}_{[-\Omega,\Omega]} = 0$; and so any $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, for which $\hat{f} = 0$ on $(-\Omega, \Omega)$, is a solution of (1.10.13).

b. If $a \neq 0$ and a solution f is constrained to be supported by $[-T, T]$, then (1.10.13) implies $a\hat{f} = \hat{f}\mathbf{1}_{[-\Omega,\Omega]}$. Thus, $a = 1$ and $\operatorname{supp} \hat{f} \subseteq [-\Omega, \Omega]$. This contradicts the analyticity of f unless $f = 0$; in fact, if $\operatorname{supp} f \subseteq [-T, T]$, then $\hat{f}$ is entire, and so $\operatorname{supp} \hat{f} \subseteq [-\Omega, \Omega]$ can only occur if $\hat{f} = 0$ on $\hat{\mathbb{R}}$.

c. Generally, let $a \neq 0$, and let g denote $af\mathbf{1}_{[-T,T)}$ and $f * d_{2\pi\Omega}$. In particular, $\operatorname{supp} g \subseteq [-T, T]$ and $\operatorname{supp} \hat{g} \subseteq [-\Omega, \Omega]$, so that $g = 0$ as in part *b*. Consequently, $f = 0$ on $(-T, T)$ and $\hat{f} = 0$ on $(-\Omega, \Omega)$. It is a remarkable fact that there are functions $f \in (L^1(\mathbb{R}) \cap L^2(\mathbb{R}))\backslash\{0\}$ with this property, e.g., [Len72], [ABe77], [Bene84, Theorem 6]. This result is related to topics about the uncertainty principle and uniqueness, e.g., [Pric85, pages 149–170].

1.10.7 Example. AN IDEMPOTENT PROBLEM IN $L^1(\mathbb{R})$ AND $L^2(\mathbb{R})$

Consider the equation

$$f = f * f. \tag{1.10.14}$$

a. If we ask whether (1.10.14) has a solution $f \in L^1(\mathbb{R})\backslash\{0\}$, the answer is "no" for the following reason, cf., *Proposition 2.1.1.* If there were such an f then $\widehat{f} = (\widehat{f})^2$ so that $\widehat{f}$ takes values 0 or 1. If $\widehat{f} = 0$ on $\mathbb{R}$ then $f = 0$ by the uniqueness theorem. If $\widehat{f} = 1$ on $\widehat{\mathbb{R}}$ then $f \notin L^1(\mathbb{R})$ since $L^1(\mathbb{R})^\wedge \subseteq C_0(\widehat{\mathbb{R}})$. If $\widehat{f}$ takes both 0 and 1 values we contradict the continuity of $\widehat{f}$.

b. If we ask whether (1.10.14) has a solution $f \in L^2(\mathbb{R})\backslash\{0\}$, the answer is "yes". In fact, let $\widehat{f} = \mathbf{1}_E$ where $|E| < \infty$. (We are using the Parseval-Plancherel Theorem here to assert the existence of $f \in L^2(\mathbb{R})$ for which $\widehat{f} = \mathbf{1}_E$.)

1.10.8 Remark. ROOT MEAN SQUARE (RMS) PROBLEM

In *Example 1.9.4* we defined an approximation problem, and alluded to a case which we now give.

For $\Omega > 0$, let $PW_\Omega = \{g \in L^2(\mathbb{R}) : \operatorname{supp} \widehat{g} \subseteq [-\Omega, \Omega]\}$; this means that $g \in L^2(\mathbb{R})$ and $\widehat{g} = 0$ on $[-\Omega, \Omega]^c \equiv \mathbb{R}\backslash[-\Omega, \Omega]$, cf., *Definition 2.2.1a*. PW_Ω is the *Paley-Wiener space*. Let $f \in L^2(\mathbb{R})$. The approximation problem in this setting is to find $g_f \in PW_\Omega$ such that

$$(1.10.15)\quad \forall g \in PW_\Omega, \quad \|f - g_f\|_{L^2(\mathbb{R})} \leq \|f - g\|_{L^2(\mathbb{R})} = e_g^{1/2}.$$

It is not clear that such a "minimizer" g_f exists. The quantity e_g is the *RMS error* corresponding to f and g.

1.10.9 Proposition.

For $f \in L^2(\mathbb{R})$ and $\Omega > 0$, let $g_f = f * d_{2\pi\Omega}$. Then $g_f \in PW_\Omega$, (1.10.15) is valid, and the RMS error corresponding to f and g_f is $\int_{|\gamma|\geq\Omega} |\widehat{f}(\gamma)|^2\, d\gamma$.

Proof. We take $g \in PW_\Omega$ and compute

$$\|f - g\|^2_{L^2(\mathbb{R})} = \|\widehat{f} - \widehat{g}\|^2_{L^2(\widehat{\mathbb{R}})}$$

$$(1.10.16)\qquad = \int_{|\gamma|\geq\Omega} |\widehat{f}(\gamma)|^2\, d\gamma + \int_{-\Omega}^{\Omega} \left(\widehat{f}(\gamma) - \widehat{g}(\gamma)\right)\overline{\left(\widehat{f}(\gamma) - \widehat{g}(\gamma)\right)}\, d\gamma.$$

This last term is zero when $g = g_f$; and since the integrand, $(\widehat{f} - \widehat{g})\overline{(\widehat{f} - \widehat{g})}$, is non-negative we have verified (1.10.15). Also, because of (1.10.16), the RMS error corresponding to f and g is $\int_{|\gamma|\geq\Omega} |\widehat{f}(\gamma)|^2\, d\gamma$. □

1.10.10 Remark. PERSPECTIVE ON BANDLIMITED APPROXIMANTS

We saw in the case of a jump discontinuity that $f * w_{2\pi\Omega}$ was preferable to $f * d_{2\pi\Omega}$ as an approximant to f under pointwise convergence. On the other hand, *Proposition 1.10.9* shows that $f * d_{2\pi\Omega}$ is the best approximant to f in the sense of minimizing the RMS error.

Another aspect of the approximation problem discussed in *Example 1.9.4* is the following result.

1.10.11 Theorem. TIMELIMITED AND BANDLIMITED APPROXIMATION

Let $T, \Omega > 0$ *and let*

$$L^2_T(\mathbb{R}) = \{f \in L^2(\mathbb{R}) : \operatorname{supp} f \subseteq [-T, T]\},$$

i.e., $f \in L^2(\mathbb{R})$ *and* $f = 0$ *on* $[-T, T]^c \equiv \mathbb{R}\backslash[-T, T]$. *Then*

$$Y = \{F : [-\Omega, \Omega] \longrightarrow \mathbb{C} : \exists k \in L^2_T(\mathbb{R}) \text{ such that } \widehat{k} = F \text{ a.e. on } [-\Omega, \Omega]\}$$

is dense in $L^2[-\Omega, \Omega]$.

Proof. Suppose the result is not true. Then by the Hahn-Banach Theorem (*Theorem B.12*) there is $G \in L^2[-\Omega, \Omega]'\backslash\{0\} = L^2[-\Omega, \Omega]\backslash\{0\}$ such that for all $K \in Y$, where $k \leftrightarrow K$, we have

$$0 = \int_{-\Omega}^{\Omega} G(\gamma)\overline{K(\gamma)}\, d\gamma$$
$$= \int \left(G\mathbf{1}_{[-\Omega,\Omega)}\right)(\gamma)\overline{K(\gamma)}\, d\gamma = \int_{-T}^{T} \left(G\mathbf{1}_{[-\Omega,\Omega)}\right)^{\vee}(t)\overline{k(t)}\, dt.$$

Since this is true for all $k \in L^2_T(\mathbb{R})$ we have $\left(G\mathbf{1}_{[-\Omega,\Omega)}\right)^{\vee} = 0$ on $[-T, T]$, and, by definition, $G\mathbf{1}_{[-\Omega,\Omega)} = 0$ off $[-\Omega, \Omega]$, i.e., $\operatorname{supp} G\mathbf{1}_{[-\Omega,\Omega)}$

is compact. This last property implies $(G\mathbf{1}_{[-\Omega,\Omega)})^\vee$ is entire and so $\left(G\mathbf{1}_{[-\Omega,\Omega)}\right)^\vee = 0$ on $\mathbb{R}$. By the L^1-uniqueness theorem and the fact that $G\mathbf{1}_{[-\Omega,\Omega)} \in L^1(\widehat{\mathbb{R}})$ we have $G\mathbf{1}_{[-\Omega,\Omega)} = 0$ on $\widehat{\mathbb{R}}$ and so $G = 0$ on $[-\Omega, \Omega]$, a contradiction. Thus, we have the desired density. □

One proof of *Theorem 1.10.2*, due to Norbert Wiener, e.g., [Wie33], involves defining $\widehat{f}$, for given $f \in L^2(\mathbb{R})$, as an eigenfunction expansion where the eigenfunctions are the so-called Hermite functions. For this reason we give the following example.

1.10.12 Example. Eigenvalues of the Fourier Transform

a. Suppose we are given the pairing $f \leftrightarrow F$ and that we can compute $\widehat{f}^\wedge$, $\widehat{f}^{\wedge\wedge}$, etc. Formally, we have

$$\widehat{f}^\wedge(\gamma) = f(-\gamma),$$

and

$$\widehat{f}^{\wedge\wedge}(\gamma) = f(-t)^\wedge(\gamma) = \widehat{f}(-\gamma);$$

and, hence,

$$\widehat{f}^{\wedge\wedge\wedge} = f.$$

b. Next, consider the operator,

$$\begin{array}{rcl} \mathcal{F} : L^2(\mathbb{R}) & \longrightarrow & L^2(\widehat{\mathbb{R}}), \\ f & \longmapsto & \widehat{f}, \end{array}$$

defined on a sufficiently well-behaved subset of $L^2(\mathbb{R})$, e.g., the space $\mathcal{S}(\mathbb{R})$ defined in *Definition 2.4.3*. Consider the eigenvalue problem,

$$\mathcal{F}f = \lambda f.$$

We have $\widehat{f} = \lambda f$, $\widehat{f}^\wedge = \lambda \widehat{f} = \lambda^2 f$, $\widehat{f}^{\wedge\wedge} = \lambda^2 \widehat{f} = \lambda^3 f$; and so, from part *a*, we obtain $f = \widehat{f}^{\wedge\wedge\wedge} = \lambda^3 \widehat{f} = \lambda^4 f$. Consequently, the eigenvalues of $\mathcal{F}$ are $\lambda = 1, i, -1, -i$; and the Hermite functions arise in this setting as the eigenfunctions of $\mathcal{F}$, e.g., *Exercise 1.26*, *Exercise 1.27*, and *Remark 2.4.11*.

c. If g is Gaussian, we saw in *Example 1.3.3* that $\widehat{g}_{\sqrt{\pi}} = g_{\sqrt{\pi}}$. By part a, we see that there are many functions k for which $\widehat{k} = k$. In fact, for any $f \in L^1(\mathbb{R}) \cap A(\mathbb{R})$, let

$$k = f + \widehat{f} + \widehat{f}^{\wedge} + \widehat{f}^{\wedge\wedge},$$

cf., *Excercise 1.28.*

1.10.13 Example. COMPUTATION OF INTEGRALS

Using the Parseval-Plancherel Theorem we can compute integrals of the form

$$\int \frac{dt}{(t^2+a^2)(t^2+b^2)}, \quad a, b > 0.$$

[Hint: $p_{1/c} \longleftrightarrow e^{-2\pi c|\gamma|}$.]

1.10.14 Example. COMPUTATION OF L^2-FOURIER TRANSFORMS

a. Let $f \in L^2(\mathbb{R})$ and let $f_n = f\mathbf{1}_{[-n,n]}$. Clearly, $f_n \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ (by Hölder's Inequality) and $\lim_{n\to\infty} \|f - f_n\|_{L^2(\mathbb{R})} = 0$. Thus, by the Parseval-Plancherel Theorem,

$$\lim_{n\to\infty} \|\widehat{f}(\gamma) - \int_{-n}^{n} f(t)e^{-2\pi it\gamma}\,dt\|_{L^2(\mathbb{R})} = 0. \tag{1.10.17}$$

b. We saw that the Dirichlet function

$$d_{2\pi} \in L^2(\mathbb{R}) \backslash (L^1(\mathbb{R}) \cap L^2(\mathbb{R})).$$

By (1.10.17),

$$\lim_{n\to\infty} \left\| \widehat{d}_{2\pi}(\gamma) - \int_{-n}^{n} \frac{\sin 2\pi t}{\pi t} \cos 2\pi t\gamma\,dt \right\|_{L^2(\mathbb{R})} = 0; \tag{1.10.18}$$

and it is easy to see, by *Theorem 1.7.6*, that

$$\forall \gamma \in \widehat{\mathbb{R}}, \quad \lim_{n\to\infty} \int_{-n}^{n} \frac{\sin 2\pi t}{\pi t} \cos 2\pi t\gamma\,dt = \begin{cases} 1, & \text{if } |t| < 1, \\ \frac{1}{2}, & \text{if } |t| = 1, \\ 0, & \text{if } |t| > 1, \end{cases} \tag{1.10.19}$$

e.g., *Exercise 1.23b*. Combining (1.10.18), (1.10.19), and F. Riesz's result used at the end of *Theorem 1.7.8*, we see that the L^2–Fourier transform of $d_{2\pi}$ is $\widehat{d}_{2\pi} = \mathbf{1}_{[-1,1]}$ a.e., which, of course, is the way it had to turn out!

Chapter 1. Exercises

Exercises 1.1–1.30 are appropriate for *Course I*. Recall that H is the Heaviside function.

1.1. Suppose that $pv \int g(t)\, dt$ exists. Prove that if g is even or g is non-negative then $\int g(t)\, dt$ exists and

$$\int g(t)\, dt = pv \int g(t)\, dt.$$

1.2. Consider the formal pairing $f \leftrightarrow F$. Verify (formally) the following.

a. If f is imaginary then $\overline{F(\gamma)} = -F(-\gamma)$ and

$$f(t) = 2i \operatorname{Im} \int_0^\infty F(\gamma) e^{2\pi i t\gamma}\, d\gamma.$$

b. f is imaginary and odd if and only if F is real and odd.

1.3. Let $f \in L^1(\mathbb{R})$ be real-valued, $f \leftrightarrow F$.

a. Verify whether or not $|F|^2$ is even, odd, or neither.

b. Verify whether or not $|F|^3$ is even, odd, or neither.

1.4. Prove that

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \frac{\sin u}{u}\, du > 1.$$

1.5. Prove that $\widehat{f} \notin L^1(\widehat{\mathbb{R}})$ for $f(t) = H(t)e^{-2\pi r t}$, where $r > 0$.

1.6. Verify whether or not the following functions f defined on $\mathbb{R}$ are elements of $L^1(\mathbb{R})$. For those that are not, verify whether or not they are elements of $L^1_{\text{loc}}(\mathbb{R})$.

a. $f(t) = H(t)$. You may assign $H(0)$ any value you like; it will not affect the validity of your answer.

b. $f(t) = \frac{1}{1+t^2}$.

c. $f(t) = \begin{cases} \frac{1}{t^2}, & \text{if } t \neq 0, \\ 0, & \text{if } t = 0. \end{cases}$

d. $f(t) = \begin{cases} \frac{1}{t}, & \text{if } t \neq 0, \\ 0, & \text{if } t = 0. \end{cases}$

e. $f(t) = \cos 2\pi t\gamma_0 + i \sin 2\pi t\gamma_0 = e^{2\pi i t\gamma_0}$.

f. $f(t) = \operatorname{sgn} t$, where $\operatorname{sgn} t \equiv H(t) - H(-t)$.

g. $f(t) = d(t)$, the Dirichlet function.

1.7. Let $f\ \mathbb{R} \longrightarrow \mathbb{C}$ be a function. The *even part* of f is the function $f_e(t) = \frac{1}{2}(f(t) + f(-t))$, and the *odd part* of f is the function $f_o(t) = \frac{1}{2}(f(t) - f(-t))$.

a. Compute $f_e + f_o$.

b. Compute the even and odd parts of the functions in *Exercise 1.6.*

c. For functions f and g, verify that

$$(fg)_e = f_e g_e + f_o g_o.$$

What is the corresponding decomposition for $(fg)_o$?

1.8. Compute the Fourier transforms of the following functions.

a. $f = \tau_u \mathbf{1}_{[-T,T)}$.

b. $f(t) = \frac{-\cos t}{\pi(t-\frac{\pi}{2})}$. [Hint. Evaluate $\tau_{\pi/2}\, d$.]

c. $f = \tau_u(\mathbf{1}_{[-T,T)} H)$.

d. $f = \mathbf{1}_{[-T,T)}(\tau_u H)$.

e. $f = \tau_u g_\lambda$, where g is the Gaussian.

1.9. Solve the differential equation $F'(\gamma) = \frac{-2\pi\gamma}{r} F(\gamma)$, $r > 0$, in *Example 1.3.3.*

1.10. Prove that $\int f(t)\, dt = 0$ in *Proposition 1.4.3.*

[Hint. $\underline{\lim}_{t\to\infty} |g(t)| > 0$ implies $g \notin L^1(\mathbb{R})$.]

1.11. Let $f \in L^1(\mathbb{R})$ and let $g(t) = f(at+b)$ for fixed $a, b \in \mathbb{R}$, $a \neq 0$. Compute $\widehat{g}$ in terms of $\widehat{f}$.

1.12. Compute the Fourier transforms of the following functions.

a. $f(t) = \frac{t^n}{n!} e^{-2\pi r t} H(t), \quad r > 0$ and $n \in \mathbb{N}$.

b. $f(t) = e^{-2\pi r t} H(t) \sin 2\pi t \gamma_0, \quad r > 0$.

1.13. Let $m, n \geq 0$ and let $f \in L^1(\mathbb{R})$. Suppose that $t^n f(t) \in L^1(\mathbb{R})$ and that the m^{th} derivative $(t^n f(t))^{(m)}$ exists everywhere and is an element of $L^1(\mathbb{R})$. Verify the Fourier transform pairing,

$$(t^n f(t))^{(m)} \longleftrightarrow (-1)^n (2\pi i)^{m-n} \gamma^m (\widehat{f})^{(n)}(\gamma).$$

Use this result to show that

$$\forall \gamma \in \widehat{\mathbb{R}} \backslash \{0\}, \quad |\widehat{f}(\gamma)| \leq \frac{1}{(2\pi|\gamma|)^m} \|f^{(m)}\|_{L^1(\mathbb{R})}.$$

1.14. Using the methods of this chapter, compute

$$\int \frac{dt}{t^4 + 7t^2 + 6}.$$

1.15. We have defined the Fourier transform in *Definition 1.1.2* and *Remark 1.1.3* using the kernels $e^{-2\pi it\gamma}$ and $e^{2\pi it\gamma}$, when the latter is required to ensure the validity of the inversion formula. These are the kernels of choice in major theoretical treatises, such as Stein and Weiss' masterpiece on Euclidean harmonic analysis [SW71], as well as in fundamental tools and algorithms such as the Discrete Fourier Transform (DFT) and the Fast Fourier Transform (FFT), e.g., [Walk91] and MATLAB. We find the above pairing to be computationally convenient. Other pairings in the literature include $e^{-it\gamma}$ and $\frac{1}{2\pi}e^{it\gamma}$, and $\frac{1}{\sqrt{2\pi}}e^{-it\gamma}$ and $\frac{1}{\sqrt{2\pi}}e^{it\gamma}$. For this exercise, do not use any of these latter kernels, or their variations in any of the other exercises!

1.16. Let $f(t) = e^t H(t)$, $g(t) = e^{-t}H(t)$, and $k(t) = e^t H(-t)$.

a. Compute $f * \cdots * f$, where there are n factors.

b. Compute $g * \cdots * g$, where there are n factors.

c. Compute $k * \cdots * k$, where there are n factors.

d. Compute $f * g$ and $g * k$. Comment on $f * k$.

e. Compute $H * f$, $H * g$, and $H * k$.

1.17. a. Compute $d_{2\pi\Omega} * \cdots * d_{2\pi\Omega}$, where there are n factors and where d is the Dirichlet function.

b. Compute and graph $\left(\mathbf{1}_{[-\Omega,\Omega)}\right)^{*n} \equiv \mathbf{1}_{[-\Omega,\Omega)} * \cdots * \mathbf{1}_{[-\Omega,\Omega)}$, where there are n factors, for the cases $n = 1, \ldots, 6$.

The n-fold convolution of part b is an example of a *spline* supported by $[-n\Omega, n\Omega]$, e.g., [dHR93], [Scho73], cf., *Exercise 2.48*. The general formula,

$$\left(\mathbf{1}_{[-\Omega,\Omega)}\right)^{*n}(t) = \sum_{k=0}^{n} \binom{n}{k} \frac{(-1)^{n-k}}{(n-1)!}(t-(n-2k)\Omega)^{n-1}\mathbf{1}_{[(n-2k)\Omega,\infty)}(t),$$

is not difficult to verify.

1.18. Find the values of $n \in \mathbb{N}$ for which $d^{(n)}_{2\pi\Omega} \in L^2(\mathbb{R})$. [Hint. Use *Theorem 1.4.1e* and the Parseval-Plancherel Theorem.]

1.19. Let $g(t) = \frac{1}{\sqrt{\pi}}e^{-t^2}$, the Gaussian.

a. Compute $g_a * g_b$ for $a, b > 0$.

b. Determine the points of inflection and regions of convexity and concavity of $\{g_\lambda\}$.

1.20. Let $f_{a,\lambda} = \tau_a d_\lambda + \tau_{-a} d_\lambda$, where d is the Dirichlet function.

a. Compute $f_{a,\lambda} * f_{a,\lambda}$.

b. Compute $\lim_{a \to a_0} f_{a,\lambda} * f_{a,\lambda}$, for $a_0 = 0, \infty$.

1.21. a. Compute the Fourier transforms of $tg_\lambda(t)$ and $tw_\lambda(t)$, where g is the Gaussian and w is the Fejér function. [Hint. The fact that $tw_\lambda(t) \notin L^1(\mathbb{R})$ is not a problem. Begin by guessing at the answer using *Theorem 1.4.1*e, and then justify your guess.]

b. Graph the functions of part a and their Fourier transforms.

1.22. Let $g(t) = \frac{1}{\sqrt{\pi}}e^{-t^2}$, the Gaussian. Compute the Fourier transforms of the functions,

$$\frac{d^{10}}{dt^{10}}(g_a * g_b)$$

and

$$\left(\frac{d^5}{dt^5}g_a\right) * \left(\frac{d^5}{dt^5}g_b\right),$$

where $a, b > 0$.

1.23. Use the inversion theorems to prove the following:

a. $$\frac{2r}{\pi}\int_0^\infty \frac{\cos 2\pi t\gamma}{r^2+\gamma^2}\,d\gamma = e^{-2\pi r|t|}, \quad r > 0.$$

b. $$\frac{2}{\pi}\int_0^\infty \frac{\sin 2\pi\gamma \cos 2\pi t\gamma}{\gamma}\,d\gamma = \begin{cases} 1, & \text{and} \quad |t| < 1, \\ \frac{1}{2}, & \text{and} \quad |t| = 1, \\ 0, & \text{and} \quad |t| > 1. \end{cases}$$

1.24. Let $f_{(\Omega)}(t) = \mathbf{1}_{[-1,1)}(t) \cos 2\pi t\Omega$. Compute

$$\lim_{\Omega \to 0} \int |\hat{f}_{(\Omega)}(\gamma) - \hat{\mathbf{1}}_{[-1,1)}(\gamma)|^2 \, d\gamma.$$

1.25. Compute the Fourier transform of

$$f(t) = g_a(t) \sin 2\pi bt, \quad a > 0,$$

in terms of a hyperbolic trigonometric function. g is the Gaussian.

1.26. The first six *Hermite polynomials* are

$$H_0(t) = 1, \quad H_1(t) = t, \quad H_2(t) = t^2 - 1,$$
$$H_3(t) = t^3 - 3t, \quad H_4(t) = t^4 - 6t^2 + 3,$$
$$H_5(t) = t^5 - 10t^3 - 15t.$$

In general,

$$\forall n \geq 0, \quad H_n(t) = (-1)^n e^{\frac{t^2}{2}} \frac{d^n}{dt^n} e^{-\frac{t^2}{2}},$$

e.g., [Wie33], [CH53], [Jac41].

a. Verify, in fact, that H_n is a polynomial of n^{th} degree with the coefficient of t^n equal to 1. [Hint. $H_{n+1}(t) = tH_n(t) - H_n'(t)$.]

b. Prove the orthogonality relations, viz., if $0 \leq m < n$ then

$$\int H_m(t) \, H_n(t) e^{-t^2/2} \, dt = 0.$$

1.27. If the convection term in (1.8.1) is replaced by $V(x)u$, where V is a potential energy, then (with an adjustment of constants) we obtain the one dimensional *Schrödinger equation*,

$$\frac{\partial^2 u}{\partial x^2} - V(x)u - c\frac{\partial u}{\partial t} = 0,$$

e.g., [Wey50a, pages 54–60]. Assuming solutions of the form $u(x,t) = X(x)T(t)$, the *separation of variables method* leads to

$$\frac{1}{X(x)}\frac{d^2X(x)}{dx^2} - V(x) = \frac{c}{T(t)}\frac{dT(t)}{dt};$$

and, if $-\lambda$ is a constant common value of both sides of this equation, we are led to the problem of determining *eigenvalues* λ and *eigenfunctions* X_λ for which the *oscillator equation*,

$$\frac{d^2X_\lambda(x)}{dx^2} + (\lambda - V(x))X_\lambda(x) = 0,$$

is valid, cf., [Str88, pages 243–253] to see the relationship between the solution of differential equations and eigenvalue problems. Solve this particular problem for the potential energy $V(x) = x^2/4$. [Hint. Let $X_n(x) = e^{-x^2/4}H_n(x)$ (*Exercise 1.26*), and obtain $\lambda_n = (n + \frac{1}{2})$, $n \geq 0$.]

1.28. Let $f(t) = e^{-\pi(t/r)^2} + re^{-\pi(rt)^2}$. For what numbers $r \in \mathbb{R}\backslash\{0\}$ do we have $f = \hat{f}$? Note the case $r = -1$.

1.29. Compute the Gibbs "overshoot" G (from *Remark 1.9.1* and *Example 1.9.3*) accurately to 6 decimal places.

1.30. Compute $(H(u)\sin u) * (H(u)\sin u)(t)$ for $t = \pm\pi/2$.

1.31. Let $f, g \in L^1(\mathbb{R})$.

a. Verify that

$$(E1.1) \qquad \int |f(t-u)||g(u)|\,du < \infty \quad t\text{–a.e. in } \mathbb{R}.$$

b. For the set X of $t \in \mathbb{R}$ which satisfy (E1.1) in part *a*, define

$$f * g(t) = \int f(t-u)g(u)\,du.$$

Prove that $f * g \in L^1(\mathbb{R})$, and that

$$\|f * g\|_{L^1(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})}\|g\|_{L^1(\mathbb{R})}.$$

1.32. The set $\mathbb{R}$ of real numbers satisfies the commutative and associative laws under both addition and multiplication, and it satisfies the distributive law under addition and multiplication. Prove that $L^1(\mathbb{R})$ satisfies the same properties when multiplication is replaced by convolution, i.e., show that

$$f+g=g+f, \quad f*g=g*f,$$
$$(f+g)+h=f+(g+h), \quad f*(g*h)=(f*g)*h,$$
$$f*(g+h)=f*g+f*h$$

for $f, g, h \in L^1(\mathbb{R})$.

1.33. Let $f, g \in L^1(\mathbb{R})$ and define the L^1-*cross-correlation*

$$f \circledast g(t) = \int f(t+u)\overline{g(u)}\, du$$

of f and g. Which properties of *Exercise 1.32* are valid? Note that $f \circledast g = f * \tilde{g}$, where $\tilde{g}$ is the *involution* defined as $\tilde{g}(t) = \overline{g(-t)}$. Clearly, $\tilde{f}^{\sim} = f$ and $f * \tilde{g} = (g * \tilde{f})^{\sim}$.

1.34. Using the methods of this chapter, verify that

$$4 \int \frac{\sin \pi\gamma}{\gamma} e^{-2\pi|\gamma| + \pi i \gamma}\, d\gamma = \pi.$$

1.35.
a. Let $f \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Prove that $f \in L^2(\mathbb{R})$.

b. Construct $f \in L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$, resp., $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \cap L^\infty_{\text{loc}}(\mathbb{R})$, such that $f \notin L^1(\mathbb{R})$, resp., $f \notin L^\infty(\mathbb{R})$.

c. Clearly, if $f \equiv 1$ then $f \in L^\infty(\mathbb{R})$ and $f \notin L^1(\mathbb{R}) \cup L^2(\mathbb{R})$. Construct $f \in L^1(\mathbb{R})$, resp., $f \in L^2(\mathbb{R})$, such that $f \notin L^2(\mathbb{R}) \cup L^\infty(\mathbb{R})$, resp., $f \notin L^1(\mathbb{R}) \cup L^\infty(\mathbb{R})$.

d. Are the constructions of part *c* possible if $L^1(\mathbb{R})$, resp., $L^2(\mathbb{R})$, is replaced by $L^1(\mathbb{R}) \cap L^\infty_{\text{loc}}(\mathbb{R})$, resp., $L^2(\mathbb{R}) \cap L^\infty_{\text{loc}}(\mathbb{R})$?

1.36. Verify *Viète's formula*

$$(E1.2) \qquad \prod_{k=1}^{\infty} \cos\frac{\gamma}{2^k} = \frac{\sin\gamma}{\gamma}.$$

In particular,

$$d_{2\pi T}(\gamma) = 2T \prod_{k=0}^{\infty} \cos\left(\frac{\pi T\gamma}{2^k}\right).$$

Viète proved the case $\gamma = \pi/2$:

$$\frac{2}{\pi} = \frac{\sqrt{2}}{2}\frac{\sqrt{2+\sqrt{2}}}{2}\frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2}\cdots.$$

François Viète was a lawyer from Poitou, who became one of the foremost mathematicians of the 16th century because of his contributions to algebra, e.g., [Kli72]. He was also the major "codebreaker" for Henry IV of France [Kahn67, pages 116–117] amidst the Holy League, the Huguenots, the Hapsburgs, the demise of the House of Valois, and the death of the Dark Eminence, Catherine de Medici.

1.37. Prove the following results, where $f \in L^1(\mathbb{R})$ and $\{k_{(\lambda)}\}$ is an approximate identity.

a. There is $\{\lambda_n\} \subseteq (0,\infty)$ such that

$$\lim_{n\to\infty} f * k_{(\lambda_n)} = f \ \text{ a.e.}$$

b. For each $\gamma \in \widehat{\mathbb{R}}$, $\lim_{\lambda\to\infty} \widehat{k}_{(\lambda)}(\gamma) = 1.$

c. Assume $\widehat{f} \in L^1(\widehat{\mathbb{R}})$, $\widehat{k}_{(\lambda)} \in L^1(\widehat{\mathbb{R}})$, and

$$(E1.3) \qquad \forall t \in \mathbb{R}, \quad k_{(\lambda)}(t) = \int \widehat{k}_{(\lambda)}(\gamma) e^{2\pi i t\gamma}\, d\gamma.$$

Then

$$\lim_{\lambda\to\infty} \left\| \int \widehat{f}(\gamma) e^{2\pi i t\gamma}\, d\gamma - f * k_{(\lambda)}(t) \right\|_{L^\infty(\mathbb{R})} = 0.$$

[Hint. Add and subtract $\int \widehat{f}(\gamma)\widehat{k}_{(\lambda)}(\gamma)e^{2\pi it\gamma}\,d\gamma$, use part *b* and LDC for one part and the Fubini-Tonelli Theorem for the other part.]

The Gauss kernel $\{g_\lambda\}$ is an example of an approximate identity where we have already verified that (E1.3) is satisfied. There are also direct calculations for the Poisson and Fejér kernels.

The point of this exercise is to illustrate the idea behind the proof of *Theorem 1.7.8.* In fact, we obtain that proof by combining parts *a* and *c*.

1.38. Let $f \in L^1(\mathbb{R})$ and assume that $|\int f(t)dt| = 1$. Prove that

$$\{f_\lambda * \widetilde{f_\lambda}\}$$

is an approximate identity, where f_λ, $\lambda > 0$, is the L^1-dilation of f.

1.39. Let $f \in L^1(\mathbb{R})$, and let $u(x,t)$ be defined by (1.8.7). Prove that u is a solution of the system (1.8.1)–(1.8.3). For the case of (1.8.2), prove only that

$$\lim_{t\to 0}\left\|u(x,t) - f(x)\right\|_{L^1(\mathbb{R})} = 0.$$

1.40. a. Prove that $A(\widehat{\mathbb{R}})$ is dense in $C_0(\widehat{\mathbb{R}})$.

b. Prove that $A(\widehat{\mathbb{R}})$ is a set of first category in $C_0(\widehat{\mathbb{R}})$.

1.41. Let $f \in L^1(\mathbb{R})$, $g \in L^2(\mathbb{R})$. Prove that $f * g \in L^2(\mathbb{R})$ and that

$$(E1.4) \qquad \|f * g\|_{L^2(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})}\|g\|_{L^2(\mathbb{R})}.$$

[Hint. Use Minkowski's inequality.] Verify that $(f * g)^\wedge = \widehat{f}\widehat{g}$.

(E1.4) should be compared with *Exercise 1.31b.* (E1.4) is a special case of *W. H. Young's inequality*,

$$\|f * g\|_{L^q(\mathbb{R})} \leq \|f\|_{L^p(\mathbb{R})}\|g\|_{L^r(\mathbb{R})},$$

where $\frac{1}{q} = \frac{1}{p} + \frac{1}{r} - 1$, e.g., [SW71, pages 178–183].

1.42. Let $f \in C_c(\mathbb{R})$. Verify whether or not $f * \tilde{f}$ is a function of bounded variation. If $f * \tilde{f}$ is a function of bounded variation then we can use the Jordan inversion theorem in place of *Proposition 1.6.11* in *Theorem 1.10.2*. [Hint. Using the Cantor function $f = f_C$, e.g., [Ben76, page 22] and *Remark 1.4.2*, we can show that $f * \tilde{f}$ need not be absolutely continuous. It is more difficult to prove that $f * \tilde{f}$ need not have bounded variation, cf., the remarks on continuous nowhere differentiable functions in *Section 3.2.4*.]

1.43. The *de la Vallée-Poussin function* v is defined as $v = 2w_2 - w_1$, where w is the Fejér function. Note that $\int v(t)dt = 1$.

  a. Verify that the dilations v_λ are equal to $v_{(\lambda)}$, where $v_{(\lambda)} \equiv 2w_{2\lambda} - w_\lambda$. Clearly, the *de la Vallée-Poussin kernel* $\{v_\lambda\}$ is an approximate identity.
  b. Graph $\widehat{v}_\lambda$.
  c. Clearly, $1 \leq \|v_\lambda\|_{L^1(\mathbb{R})} = \|v\|_{L^1(\mathbb{R})}$. Estimate $\|v\|_{L^1(\mathbb{R})}$.

1.44. Prove (1.10.11) and (1.10.12).

1.45. Verify that

$$\lim_{\Omega\to\infty} \int \left(\frac{\sin t}{t}\right)^j d_{2\pi\Omega}(t)\, dt = 1$$

for $j \in \mathbb{N}$. [Hint. Consider the integral over $(0,\infty)$, verify that

$$\frac{1}{t}\left[\left(\frac{\sin t}{t}\right)^j - 1\right] \in L^1(0,1),$$

and use the Riemann-Lebesgue Lemma.]

1.46. Compute

$$I = \int \left(\frac{\sin t}{t}\right)^j \frac{\sin rt}{t} dt$$

for $j \in \mathbb{N}$ and real $r \geq j$. This generalizes the $n = 2$ case of *Example 1.10.5*, cf., *Exercises 1.17, 1.34*, and *1.45*. [Hint. By Parseval's formula,

$$I = \pi^{j+1} \int_{\frac{-r}{2\pi}}^{\frac{r}{2\pi}} \mathbf{1}_{[-\frac{1}{2\pi},\frac{1}{2\pi}]} * \cdots * \mathbf{1}_{[-\frac{1}{2\pi},\frac{1}{2\pi}]}(\gamma)\, d\gamma.$$

The support of the integrand is contained in $[-\frac{j}{2\pi}, \frac{j}{2\pi}]$, e.g., *Exercise 2.11.* Thus,

$$I = \pi^{j+1}\hat{\mathbf{1}}_{[-\frac{1}{2\pi},\frac{1}{2\pi}]}(0)^j.]$$

1.47. If the potential energy V of the Schrödinger equation in *Exercise 1.27* is a constant b, we obtain the fundamental equation of *linear cable theory*. Equations such as this were used by Lord Kelvin (1856) to analyze electric current flow in submarine cables, cf., [Kör88, pages 332–337] where the heat equation, i.e., $b = 0$, is analyzed vis a vis the success of the transatlantic cable. Solve the *cable equation*

$$\frac{\partial^2 u}{\partial x^2} - bu - c\frac{\partial u}{\partial t} = f$$

for a given forcing function $f(x,t)$. Nowadays, the cable equation plays a critical role in electric models of neuron excitation, e.g., [JNT75].

1.48. Prove *Lemma 1.7.3.* [Hint. The proof is based on the following argument in the case that f is not only increasing but also absolutely continuous. Without loss of generality, let g be real-valued and let $G(t) = \int_a^t g(u)\,du$. Then FTC and integration by parts give

$$\int_a^b f(t)g(t)\,dt = G(b)f(b) - \int_a^b G(t)f'(t)\,dt.$$

Since $f' \geq 0$ a.e. and G is continuous on $[a,b]$ we have

$$m\int_a^b f'(t)\,dt \leq \int_a^b G(t)f'(t)\,dt \leq M\int_a^b f'(t)\,dt,$$

where $m = \inf\{G(t)\}$ and $M = \sup\{G(t)\}$. Thus, by the intermediate value theorem there is $\xi \in [a,b]$ such that

$$\int_a^b f(t)g(t)\,dt = G(b)f(b) - G(\xi)\int_a^b f'(t)\,dt.$$

Another application of FTC gives the result.]

1.49. Let $g \in L^1(\mathbb{R})$, and, for $s > 0$ and $(t,\gamma) \in \mathbb{R} \times \widehat{\mathbb{R}}$, set

$$g_{s,t,\gamma}(u) = g_s(u-t)e^{2\pi i u\gamma}.$$

Compute $\widehat{g}_{s,t,\gamma}$, cf., *Theorem 1.2.1d, e, f* and *Exercise 1.11.*

The algebraic (and geometric) operations of dilation, modulation, and translation are the fundamental "invariants" under the action of the Fourier transform on $\mathbb{R}$ (*Theorem 1.2.1*). As such it is natural to attempt to "synthesize" or reconstruct functions f in terms of scale-time-frequency harmonics $\{g_{s,t,\gamma} : s > 0 \text{ and } (t,\gamma) \in \mathbb{R} \times \widehat{\mathbb{R}}\}$ for a given "analyzing" function g. For a fixed frequency $\gamma \in \widehat{\mathbb{R}}$, this program of signal reconstruction is a fundamental part of wavelet theory [Dau92], [Mey90]; and for a fixed scale $s > 0$, this program becomes the Weyl-Heisenberg or Gabor theory [Gabo46], [vN55], [BW94]. The signal processing program for the complete set $\{g_{s,t,\gamma}\}$ has important applications, e.g., [MZ93], and theoretical developments in terms of the metaplectic group, e.g., [HL95]. There have been many contributors to the whole program, and we refer to [HL95] for an overview. The metaplectic group contains the Heisenberg group as a subgroup, as well as an isomorphic copy of the affine group. These two groups correspond to the Gabor and wavelet theories, respectively.

1.50. Let $0 < \alpha < \beta$ and let K be the trapezoid

$$K(\gamma) = \begin{cases} 1, & \text{if } |\gamma| \leq \alpha, \\ 0, & \text{if } |\gamma| \geq \beta, \\ \text{linear}, & \text{if } \alpha \leq |\gamma| \leq \beta. \end{cases}$$

a. Compute $k \equiv K^\vee$, cf., *Exercise 1.43.*

b. Show that $\|K\|_{A(\widehat{\mathbb{R}})} \leq \frac{\beta+\alpha}{\beta-\alpha}$, and hence $\|K\|_{A(\widehat{\mathbb{R}})} \leq 3$ when $\beta = 2\alpha$, cf., (3.5.4) in *Example 3.5.3.*

c. Besides the estimate in part b, it is obvious that $\|K\|_{L^\infty(\widehat{\mathbb{R}})} = 1$. Compute $\|K\|_{L^2(\widehat{\mathbb{R}})}$.

1.51. Let $f \in PW_\Omega$. Prove that

$$\forall t \in \mathbb{R}, \quad f(t) = \int \widehat{f}(\gamma) e^{2\pi i t \gamma}\, d\gamma,$$

cf., *Theorems 1.1.6* and *1.1.7*.

Chapter 2

Measures and Distribution Theory

2.1 Approximate identities and δ

In $\mathbb{R}$, we know that $7 \times 1 = 7$, $\pi \times 1 = \pi$, $0 \times 1 = 0, \cdots$. 1 is the *multiplicative unit* for $\mathbb{R}$. In $L^1(\mathbb{R})$, convolution is the multiplication (*Section 1.5*), and the norm $\| \cdots \|_{L^1(\mathbb{R})}$ is the absolute value (measure of distance).

2.1.1 Proposition.
$L^1(\mathbb{R})$ does not have a unit under convolution.

Proof. Suppose $u \in L^1(\mathbb{R})$ were a unit. Choose $f \in L^1(\mathbb{R})$ for which $\widehat{f}$ never vanishes, e.g., $f(t) = e^{-2\pi|t|}$. Then $\|\widehat{f} - \widehat{f}\widehat{u}\|_{L^\infty(\widehat{\mathbb{R}})} \leq \|f - f * u\|_{L^1(\mathbb{R})} = 0$, and so $\widehat{u}$ is identically 1 on $\widehat{\mathbb{R}}$. This contradicts the Riemann-Lebesgue Lemma, and so $u \notin L^1(\mathbb{R})$. □

On the other hand, we did show in *Theorem 1.6.9* that there are families $\{k_{(\lambda)}\} \subseteq L^1(\mathbb{R})$ of functions, appropriately called approximate identities, with the property that

$$\text{(2.1.1)} \qquad \forall f \in L^1(\mathbb{R}), \quad \lim_{\lambda \to \infty} \|f - f * k_{(\lambda)}\|_{L^1(\mathbb{R})} = 0.$$

Similarly, recall from *Proposition 1.6.11* that if $f \in L^{\infty}(\mathbb{R})$ is continuous on $\mathbb{R}$ and $\{k_{(\lambda)}\}$ is an approximate identity, then

$$\forall t \in \mathbb{R}, \quad \lim_{\lambda \to \infty} f * k_{(\lambda)}(t) = f(t). \tag{2.1.2}$$

2.1.2 Remark. MOTIVATION FOR δ

We can think of "$f*k_{(\lambda)} \longrightarrow f$" in (2.1.1) or (2.1.2) as a set $\{f*k_{(\lambda)}\}$ of functions approximating the function f; or we can think of the family $\{k_{(\lambda)}\}$ as approximating something, call it δ, which plays the role of an identity under convolution for elements in $L^1(\mathbb{R})$, i.e., "$f * \delta = f$" even though $\delta \notin L^1(\mathbb{R})$.

In order to quantify this latter interpretation, assume each $k_{(\lambda)}$ is an even function. Then fix the point $t = 0$ in (2.1.2), and think of the integral $f * k_{(\lambda)}(0)$ as a function,

bigger "∫"?

$$\begin{array}{rcl} k_{(\lambda)} : C_b(\mathbb{R}) & \longrightarrow & \mathbb{C} \\ f & \longmapsto & f * k_{(\lambda)}(0) = \int k_{(\lambda)}(t) f(t)\, dt, \end{array} \tag{2.1.3}$$

whose domain $C_b(\mathbb{R})$ is the set of bounded, continuous functions on $\mathbb{R}$.

2.1.3 Definition. δ

a. δ is the function

$$\begin{array}{rcl} \delta : C_b(\mathbb{R}) & \longrightarrow & \mathbb{C} \\ f & \longmapsto & f(0), \end{array} \tag{2.1.4}$$

i.e., for each $f \in C_b(\mathbb{R})$, $\delta(f) \equiv f(0)$. A function (such as (2.1.4)), whose domain is a set of functions and whose range is a set of numbers, is called a *functional*, cf., *Definition 2.2.3c*.

δ is often called the *Dirac δ-function* in spite of the fact that it was neither discoverd by Dirac nor is it an ordinary function on $\mathbb{R}$. It is also called the *unit impulse*. We shall refer to δ as the *Dirac measure*. The concept of a *measure* generalizes that of a locally integrable function, i.e., an element of $L^1_{\text{loc}}(\mathbb{R})$, and is a special type of distribution. We shall see that δ is a measure (*Definition 2.3.6c*).

b. We sometimes write $\delta = \delta_0$ since the definition of δ in (2.1.4) specifies values of functions f (in the domain of δ) at 0. Similarly, we define the *Dirac measure* δ_r *at* $r \in \mathbb{R}$ as the function, with domain $C_b(\mathbb{R})$ (*not* $\mathbb{R}$),

$$\begin{array}{rcl} \delta_r : C_b(\mathbb{R}) & \longrightarrow & \mathbb{C} \\ f & \longmapsto & f(r), \end{array}$$

i.e., for each $f \in C_b(\mathbb{R})$, $\delta_r(f) \equiv f(r)$. We also write $\tau_r\delta$ in place of δ_r for reasons that will become apparent in *Definition 2.5.3*.

c. Recall our introduction of δ in *Remark 1.1.4* and the "formula" (δ) there. The point of that formula was that if δ were an ordinary function then it would be 0 everywhere except at the origin, where it would be so large that "$\int \delta(u)\,du = 1$". This is of course nonsense because of the definition of the integral. It is not nonsense from the point of view and needs of engineers and physicists (such as Dirac!); and the definition (2.1.4), motivated by the theorems expressed in (2.1.1) and (2.1.2), is a mathematically sound way to legitimize the ingenuity of these scientists in their formulas such as (δ).

d. To be consistent, (2.1.3) also tells us that any element $f \in L^1_{\text{loc}}(\mathbb{R})$ can be thought of as a function whose domain $C_c(\mathbb{R})$ is the vector space of all compactly supported, continuous functions on $\mathbb{R}$. In fact, in this case, we can formulate the notation $g(f)$ to mean

$$\forall f \in C_c(\mathbb{R}), \quad g(f) = \int g(t)f(t)\,dt, \tag{2.1.5}$$

cf., *Section 2.2.* The domain $C_c(\mathbb{R})$ is used instead of $C_b(\mathbb{R})$ so that the integral on the right side of (2.1.5) is well-defined for $f \in L^1_{\text{loc}}(\mathbb{R})$.

e. This approach of defining objects such as δ, which beg to exist but do not exist as ordinary functions, as functions (functionals) whose domain is a space of functions, is rooted in ideas associated with Parseval's formula and weak solutions in physics. These generalized objects are called *distributions* or *generalized functions*; and in this chapter we shall develop their elementary properties, especially those used in applications.

The following result is expected in light of *Remark 2.1.2*; its proof is part of *Exercise 2.38*.

2.1.4 Proposition.
Let $\{k_{(\lambda)}\} \subseteq L^1(\mathbb{R})$ *be an approximate identity. Then*

$$\forall f \in C_b(\mathbb{R}), \quad \lim_{\lambda \to \infty} \int k_{(\lambda)}(t) f(t)\, dt = \delta(f).$$

2.1.5 Remark. $\delta(t)$ AND COMPOSITION
In spite of the fact that the domain of δ is a space of functions, we shall sometimes write $\delta(t)$ instead of δ, and in this case we replace the notation $\delta(f)$ (after (2.1.4)) by the seemingly more cumbersome notation

$$\delta(t)(f(t)).$$

The reasons for doing this are that δ is intuitively perceived as an ordinary function by the (technically false) description of it in *Definition 2.1.3c*, and δ is also approximated by ordinary functions as in *Proposition 2.1.4*.

The theoretical rationale in the previous paragraph would be an effete exercise if it were not for the importance of *composition* in mathematics, e.g., [AZ90], in neural nets, e.g., [Hay94], for compression problems in signal processing, etc. We shall illustrate a role of composition for data compression in *Example 2.3.10*, but for now are interested in giving an intelligent meaning to the compelling notation "$\delta \circ g$", noting that if $g(t) = t$ then we are really dealing with δ itself, cf., [AMS73], [Jon82].

2.1.6 Example. $\delta(at + b)$
We give a reasonable meaning to "$\delta(at + b)$", where $a \neq 0$; and in the process we show that

$$\delta(at + b) = \frac{1}{|a|}\delta(t + \frac{b}{a}) = \frac{1}{|a|}\, \tau_{-b/a}\delta, \tag{2.1.6}$$

cf., *Exercises 2.10* and *2.21*. Let $\{k_{(\lambda)}\}$ be an approximate identity. Then, by (2.1.5) and *Proposition 2.1.4*, it is reasonable to define $\delta(at+b)$ by the limit,

$$\forall f \in C_b(\mathbb{R}), \quad \lim_{\lambda \to \infty} \int k_{(\lambda)}(at + b) f(t)\, dt = \delta(at + b)(f(t)).$$

The integral on the left side is

$$\frac{1}{|a|}\int_{-\infty}^{\infty} k_{(\lambda)}(u) f\left(\frac{u-b}{a}\right) du,$$

and this converges to $1/|a|\, f(-b/a)$. (2.1.6) follows by definition of $\delta_{-b/a}$.

2.2 Definition of distributions

The purpose of the theory of distributions is to provide a unified setting and *calculus* for many of the objects arising in analysis. These objects include the customary functions, viz., the elements of the space $L^1_{\text{loc}}(\mathbb{R})$ of locally integrable functions. They also include impulses (Dirac measures), dipoles, and other notions from the sciences, whose role and mathematical identity could not be assimilated by the 17th century calculus, or the spectacular 18th and 19th century developments of this calculus, and the 19th and early 20th century theory of real analysis.

A key feature of the theory of distributions is that all of these objects (distributions) can be differentiated in a natural way inspired by the integration by parts formula, e.g., *Section 2.3* and *Theorem A.22.* Further, some of the important results from real analysis allowing for switching of operations, such as summation and differentiation, are true without hypotheses in the case of distributions, e.g., *Exercise 2.46*, cf., *Example 2.2.2b.*

2.2.1 Definition. $C_c^\infty(\mathbb{R})$

a. $C^\infty(\mathbb{R})$ denotes the space of infinitely differentiable complex-valued functions on $\mathbb{R}$.

If $f : \mathbb{R} \longrightarrow \mathbb{C}$ is any function on $\mathbb{R}$ then the *support* of f, denoted by $\operatorname{supp} f$, is the smallest closed set outside of which f vanishes, e.g., *Example 1.2.2b.*

Let

$$C_c^\infty(\mathbb{R}) = \{f : f \in C^\infty(\mathbb{R}) \text{ and } \operatorname{supp} f \text{ is compact}\}.$$

(*Compact sets* of $\mathbb{R}$ are precisely the closed and bounded sets of $\mathbb{R}$, cf., *Definition B.1.* A set $B \subseteq \mathbb{R}$ is *bounded* if there is $T > 0$ such that $B \subseteq [-T, T]$. $C \subseteq \mathbb{R}$ is *closed* if its complement $\mathbb{R} \backslash C$ is *open*; and $U \subseteq \mathbb{R}$ is open if U is the countable union of disjoint (open) intervals (a,b). The structure of open sets in $\mathbb{R}^2$ is more complicated, e.g., [Ben76, pages 15–16]. Clearly, closed intervals $[-T, T]$ are compact.)

b. Just as $C_b(\mathbb{R})$ allowed us to define the object δ, the even smaller space $C_c^\infty(\mathbb{R})$ will allow us to define more unusual distributions T than δ. T exists as an element of a large space of distributions (which we shall soon define), but can generally only be realized or evaluated by operating on a "test function" $f \in C_c^\infty(\mathbb{R})$.

2.2.2 Example. $C_c^\infty(\mathbb{R})$

$C_c^\infty(\mathbb{R})$ is not the trivial set $\{0\}$. This statement is not so absurd since there are no analytic functions in $C_c^\infty(\mathbb{R})$. Why? (*Exercise 2.5*). In fact, $C_c^\infty(\mathbb{R})$ is an infinite dimensional space, so let us write down at least one element.

a. Let

$$\phi(t) = \begin{cases} e^{1/t}, & \text{if } t < 0, \\ 0, & \text{if } t \geq 0, \end{cases}$$

and define $f(t) = c\phi(|t|^2 - 1)$. Clearly, $f(t) = 0$ if $|t| \geq 1$ so that $\operatorname{supp} f \subseteq [-1, 1]$. The constant c is chosen so that $\int f(t)\, dt = 1$, and it is a straightforward calculation to show that f is infinitely differentiable, e.g., *Exercise 2.5.*

Of course, the set $\{f_\lambda\}$ of dilations is an approximate identity since $\int f(t)\, dt = 1$.

b. To generate other examples of elements in $C_c^\infty(\mathbb{R})$, take any $g \in L^1(\mathbb{R})$ having compact support. Then $f * g \in C_c^\infty(\mathbb{R})$. To see this we must verify that $\operatorname{supp} f * g$ is compact and that $f * g$ is infinitely differentiable. The first fact is routine to prove and the second involves checking conditions to switch the operations of differentiation and integration, e.g., [Apo57] for the conditions and [Fey86, pages 72 and 93] for motivation.

c. We shall have occasion to need elements of $C_c^\infty(\mathbb{R})$ with special properties such as vanishing moments, etc. In this regard let $N \in \mathbb{N}$.

We shall construct real valued even functions $k \in C_c^\infty(\mathbb{R})$ such that

$$\forall\, 0 \leq n \leq N, \quad \int t^n k(t)\, dt = 0 \tag{2.2.1}$$

and

$$\forall \gamma \in \widehat{\mathbb{R}}\backslash\{0\}, \quad \int_0^\infty \widehat{k}(\lambda\gamma)^2 \frac{d\lambda}{\lambda} = 1. \tag{2.2.2}$$

Let $h = f^{(2j)}$ for some $j > N/2$, where $f \in C_c^\infty(\mathbb{R})$ is real and even. Clearly, $h \in C_c^\infty(\mathbb{R})$, h is real-valued, and $\operatorname{supp} h$ is compact. By definition of the derivative, one computes that h is even. Noting that $n \leq N < 2j$, we compute

$$\begin{aligned}
\int t^n f^{(2j)}(t)\, dt &= -n \int t^{n-1} f^{(2j-1)}(t)\, dt \\
&= (-1)^2 n(n-1) \int t^{n-2} f^{(2j-2)}(t)\, dt \\
&= (-1)^n n! \int f^{(2j-n)}(t)\, dt = 0
\end{aligned}$$

by integration by parts, using the facts that $2j - n > 0$ and that $f \in C_c^\infty(\mathbb{R})$. This is (2.2.1) for $k = h$.

Since $h \in C_c^\infty(\mathbb{R})$ then $\widehat{h}(\lambda\gamma)$ is rapidly decreasing as $\lambda \longrightarrow \infty$ for any fixed $\gamma \in \widehat{\mathbb{R}}\backslash\{0\}$, e.g., *Exercise 1.13*. Thus, $\int_1^\infty \widehat{h}(\lambda\gamma)^2 \frac{d\lambda}{\lambda}$ is an absolutely convergent integral. By (2.2.1), $\int h(t)\, dt = 0$ and so $\widehat{h}(0) = 0$. Clearly, $\widehat{h} \in C^\infty(\mathbb{R})$, and so, for a fixed $\gamma \in \widehat{\mathbb{R}}\backslash\{0\}$, $\widehat{h}(\lambda\gamma)^2/\lambda$ is bounded in a neighborhood of 0 since $\widehat{h}(0) = 0$ and by the definition of derivative. Thus, $\int_0^1 \widehat{h}(\lambda\gamma)^2 \frac{d\lambda}{\lambda}$ is an absolutely convergent integral. We compute

$$\forall \gamma \in \widehat{\mathbb{R}}\backslash\{0\}, \quad \int_0^\infty \widehat{h}(\lambda\gamma)^2 \frac{d\lambda}{\lambda} = \int_0^\infty \widehat{h}(\lambda)^2 \frac{d\lambda}{\lambda} = c^2, \tag{2.2.3}$$

using the fact that $\widehat{h}$ is even. Note that $c \neq 0$ since the integrands of (2.2.3) are positive. (2.2.2) is therefore obtained by setting $k = c^{-1}h$.

2.2.3 Definition. DISTRIBUTIONS

a. $C_c^\infty(\mathbb{R})$ is a vector space. A linear function,

$$\begin{aligned} T : C_c^\infty(\mathbb{R}) &\longrightarrow \mathbb{C} \\ f &\longmapsto T(f), \end{aligned}$$

is a *distribution* or *generalized function* if $\lim_{n\to\infty} T(f_n) = 0$ for every sequence $\{f_n\} \subseteq C_c^\infty(\mathbb{R})$ satisfying the properties:

i. $\exists K \subseteq \mathbb{R}$, compact, such that $\forall n$, $\operatorname{supp} f_n \subseteq K$,

ii. $\forall k \geq 0$, $\lim_{n\to\infty} \|f_n^{(k)}\|_{L^\infty(\mathbb{R})} = 0$.

b. A distribution T is *positive*, written $T \geq 0$, if $T(f) \geq 0$ for all nonnegative functions $f \in C_c^\infty(\mathbb{R})$.

c. The space of all distributions on $\mathbb{R}$ is denoted by $D'(\mathbb{R})$. We incorporate the prime "$'$" in this notation to continue established notation and to emphasize the fact that $D'(\mathbb{R})$ is the dual space of $C_c^\infty(\mathbb{R})$, i.e., the space of all continuous linear functionals on $C_c^\infty(\mathbb{R})$.

We use the word "functional" just as we use the word "function"; the only reason we make any distinction is because of our new situation dealing with a space of functions as domain and $\mathbb{C}$ as range, e.g., (2.1.4) and (2.1.5). We denote the operation of T on f by $T(f)$. The *linearity* of T means that

$$T(c_1 f_1 + c_2 f_2) = c_1 T(f_1) + c_2 T(f_2)$$

for all $c_1, c_2 \in \mathbb{C}$ and $f_1, f_2 \in C_c^\infty(\mathbb{R})$.

d. $D'(\mathbb{R})$ is a vector space. In fact, if $T_1, T_2 \in D'(\mathbb{R})$ and $c_1, c_2 \in \mathbb{C}$ then $c_1T_1 + c_2T_2$ is a well-defined distribution, defined by the rule,

$$\forall f \in C_c^\infty(\mathbb{R}), \quad (c_1T_1 + c_2T_2)(f) = c_1T_1(f) + c_2T_2(f).$$

e. Occasionally we shall have to keep track of the underlying variable $t \in \mathbb{R}$ in dealing with the notation $T(f)$. In such cases, as in *Remark 2.1.5* and *Example 2.1.6*, we shall denote $T(f)$ by

$$T(t)(f(t)).$$

2.2.4 Remark. HISTORICAL AND BIBLIOGRAPHICAL NOTE

Laurent Schwartz received the Fields medal in 1950 for developing the theory of distributions. His classic book is *Théorie des distributions* [Sch66], [Sch61]. The first edition was published in two volumes in 1950 and 1951. These volumes are a compendium of diverse past accomplishments, a unification of technologies, an original formulation of ideas both new and old, and a research manual leading to new mathematics and applications.

Two other monumental contributions are Gel'fand and Shilov's *Generalized Functions*, in five volumes, and Hörmander's three volumes, *The Analysis of Linear Partial Differential Equations*, e.g., [Hör83].

The origins of distribution theory are based in the operational calculus from engineering, e.g., *Section 2.6*, and the concepts of "turbulent" and "weak" solutions of partial differential equations from physics. Schwartz's Introduction and Gel'fand and Shilov's bibliographic notes give a nice overview.

A great number of books has been written on the theory of distributions, running the gamut from pure topological vector space presentations [Hor66] to applications in optics and supersonic wing theory [deJ64]. We hesitate listing excellent books that we know, since our omissions would surely include comparably excellent ones.

2.2.5 Definition. EQUALITY OF DISTRIBUTIONS

a. Let $T_1, T_2 \in D'(\mathbb{R})$. T_1 *equals* T_2, i.e., T_1 is the same distribution as T_2, if

$$\forall f \in C_c^\infty(\mathbb{R}), \quad T_1(f) = T_2(f).$$

Notationally, in this case, we write $T_1 = T_2$. In particular, $T = 0$ if $T(f) = 0$ for all $f \in C_c^\infty(\mathbb{R})$.

b. This notion of equality can be explained mathematically in functional analytic terms, e.g., *Definition B.6*. Intuitively, however, the idea is clear: $T(f) = 0$ for all f in the domain $C_c^\infty(\mathbb{R})$ of T implies T is the 0-distribution just as $g(t) = 0$ for all $t \in \mathbb{R}$ implies g is the 0-function.

Another compelling reason to accept this notion of equality is that if $g \in L^1_{\text{loc}}(\mathbb{R})$ and $g(f) = 0$ for all $f \in C_c^\infty(\mathbb{R})$ then g is the 0-function, e.g., *Exercise 2.8.*

2.2.6 Example. Locally Integrable Distributions

a. Let $g \in L^1_{\text{loc}}(\mathbb{R})$ and define the functional,

$$\begin{aligned} T_g : C_c^\infty(\mathbb{R}) &\longrightarrow \mathbb{C}, \\ f &\longmapsto T_g(f), \end{aligned}$$

where T_g is defined as

$$\forall f \in C_c^\infty(\mathbb{R}), \quad T_g(f) = \int g(t)f(t)\,dt, \tag{2.2.4}$$

cf., (2.1.5) where we wrote $g(f)$ instead of $T_g(f)$. It is easy to check that $T_g \in D'(\mathbb{R})$.

b. Not only does T_g define an element of $D'(\mathbb{R})$, but the linear mapping

$$\begin{aligned} L : L^1_{\text{loc}}(\mathbb{R}) &\longrightarrow D'(\mathbb{R}) \\ g &\longmapsto L(g) = T_g \end{aligned} \tag{2.2.5}$$

allows us to identify $L^1_{\text{loc}}(\mathbb{R})$ with a subset of $D'(\mathbb{R})$.

To see this, we must show that the mapping L is injective, i.e., a one-to-one function. This means we must prove that if $T_{g_1} = T_{g_2}$ in $D'(\mathbb{R})$ then $g_1 = g_2$ a.e., or, equivalently (by the linearity), if $g(= g_1 - g_2)$ is not the 0-function in $L^1_{\text{loc}}(\mathbb{R})$ then there is $f \in C_c^\infty(\mathbb{R})$ for which $\int g(t)f(t)\,dt \neq 0$. As pointed out in *Definition 2.2.5b*, this fact is a consequence of *Exercise 2.8.*

Since the mapping is injective, there is no ambiguity in identifying $g \in L^1_{\text{loc}}(\mathbb{R})$ with $L(g) = T_g \in D'(\mathbb{R})$.

c. Examples of such distributions T_g arise from $g = H$ and $g(t) = 1/|t|^{1/2}$. On the other hand, $g(t) = 1/t \notin L^1_{\text{loc}}(\mathbb{R})$, cf., *Example 2.3.8c.*

d. The domain of the Dirac measure defined in (2.1.4) can be restricted to $C_c^\infty(\mathbb{R})$, in which case we have

$$\begin{aligned} \delta : C_c^\infty(\mathbb{R}) &\longrightarrow \mathbb{C} \\ f &\longmapsto \delta(f) = f(0). \end{aligned} \tag{2.2.6}$$

It is easy to check that δ defined by (2.2.6) is a distribution. We hasten to point out that the Dirac measure is not an element of $L^1_{\text{loc}}(\mathbb{R})$, i.e., δ is not in the range of the function L defined in (2.2.5), e.g., *Exercise 2.39.*

2.2.7 Definition. Support

a. We have defined the 0-distribution. We now "localize this definition in the following way. $T \in D'(\mathbb{R})$ is *zero on the open set* $U \subseteq \mathbb{R}$, written $T = 0$ on U, if

$$\forall f \in C_c^\infty(\mathbb{R}), \quad \text{such that} \quad \operatorname{supp} f \subseteq U, \quad T(f) = 0.$$

The *support* of T, denoted by $\operatorname{supp} T$, is the smallest closed set $C = \mathbb{R}\backslash U$ outside of which T is 0.

b. The definition in part *a* is usually easier to check than you might imagine. In particular, we have

$$\forall r \in \mathbb{R}, \quad \operatorname{supp} \delta_r = \{r\} \tag{2.2.7}$$

and

$$\operatorname{supp} T_g = \operatorname{supp} g, \tag{2.2.8}$$

where $g \in L^1_{\text{loc}}(\mathbb{R})$ and where the right side of (2.2.8) was defined in *Definition 2.2.1*, e.g., *Exercise 2.25*, cf., *Exercise 2.33.*

2.3 Differentiation of distributions

2.3.1 Definition. Distributional Differentiation

a. The duality between the small space $C_c^\infty(\mathbb{R})$ and the large space $D'(\mathbb{R})$, allowing us to define so many objects T in $D'(\mathbb{R})$, can be coupled with the integration by parts formula to provide a definition of "T'", the distributional derivative of T, in part *b* below.

Let $C^1(\mathbb{R})$ be the space of continuously differentiable functions on $\mathbb{R}$. If $f \in C_c^\infty(\mathbb{R})$ and g is sufficiently smooth, e.g., if $g \in C^1(\mathbb{R})$ or even if g is only an element of $AC_{\text{loc}}(\mathbb{R})$, then

$$\int g'(t) f(t)\, dt = -\int g(t) f'(t)\, dt. \tag{2.3.1}$$

The integration by parts formula in (2.3.1) is the distributional "duality formula"

$$\forall f \in C_c^\infty(\mathbb{R}), \quad g'(f) = -g(f'). \tag{2.3.2}$$

Since the right side of (2.3.2) is well-defined when g is replaced by any $T \in D'(\mathbb{R})$, we are motivated to make the following definition.

b. The *distributional derivative* T' of $T \in D'(\mathbb{R})$ is defined by the formula

$$\forall f \in C_c^\infty(\mathbb{R}), \quad T'(f) = -T(f'). \tag{2.3.3}$$

c. To establish the viability of (2.3.3) as an effective definition of the notion of derivative we must prove the following:

$$T' \in D'(\mathbb{R}), \tag{2.3.4}$$

and

$$\forall g \in AC_{\text{loc}}(\mathbb{R}), \quad T'_g = T_{Dg}, \tag{2.3.5}$$

where Dg denotes the ordinary pointwise derivative. The verification of (2.3.4) and (2.3.5) is routine, e.g., *Exercise 2.7.* In concert with (2.3.3) we write $T'_g = g'$.

d. For each $T \in D'(\mathbb{R})$ we define $T^{(n)}$, the *nth distributional derivative* of T, as the distribution defined by the formula,

$$\forall f \in C_c^\infty(\mathbb{R}), \quad T^{(n)}(f) = (-1)^n T(f^{(n)}). \tag{2.3.6}$$

2.3.2 Remark. DEFINITION BY DUALITY

a. The use of "duality formulas" such as (2.3.2) to define an analytic operation (such as differentiation) on an arbitrary distribution T is a critical aspect of the theory. The idea is to define an operation on T in terms of the same (or similar) operation on $C_c^\infty(\mathbb{R})$, where it makes perfectly good sense; the vehicle for effecting this definition is a "duality formula". As we shall soon see, Parseval's formula is the "duality formula" which allows us to define the Fourier transform of distributions.

b. In general, we cannot extend (2.3.5) to arbitrary elements $g \in L^1_{\text{loc}}(\mathbb{R})$ even though T_g is a well-defined generalized function. In fact, if g is infinitely differentiable on $\mathbb{R}\backslash\{t_0\}$ in the ordinary pointwise sense and Dg is the ordinary pointwise derivative of g defined everywhere except at t_0, then, in general, T'_g and T_{Dg} are distributions but $T'_g \neq T_{Dg}$, e.g., *Example 2.3.3.*

2.3.3 Example. $H' = \delta$

Let H be the Heaviside function.

a. The ordinary pointwise derivative DH exists and takes the value 0 on $\mathbb{R}\backslash\{0\}$. Thus, $DH \in L^1_{\text{loc}}(\mathbb{R})$ and DH is the 0 distribution.

b. The distributional derivative H' is evaluated as follows. Choose $f \in C_c^\infty(\mathbb{R})$ and compute

$$H'(f) = -H(f') = -\int_0^\infty f'(t)\, dt = f(0) = \delta(f). \tag{2.3.7}$$

Since $H'(f) = \delta(f)$ for all $f \in C_c^\infty(\mathbb{R})$ we can conclude that H' and δ are the same distribution, i.e.,

$$H' = \delta,$$

where H' is the distributional derivative of H.

2.3.4 Remark. NOTATION FOR DIFFERENTIATION

Let $g : \mathbb{R} \longrightarrow \mathbb{C}$ be a function, and suppose the ordinary pointwise derivative of g exists at t. As indicated above, we denote this value by $Dg(t)$, and if $Dg(t)$ exists a.e., we denote the resulting function by Dg.

On the other hand, if g defines a distribution T_g then g' denotes the distributional derivative T_g'.

In general, the distinction will be clear; and when no confusion arises we shall be less than compulsive about the "D" notation.

2.3.5 Proposition.

Let $g \in C^1(\mathbb{R}\backslash\{0\})$, i.e., g is continuously differentiable on $\mathbb{R}\backslash\{0\}$. Assume that g has a jump discontinuity at the origin with jump $\sigma_0 = g(0+) - g(0-)$. Then we have the distributional equation,

$$g' = Dg + \sigma_0\delta,$$

i.e.,

$$\forall f \in C_c^\infty(\mathbb{R}), \quad g'(f) = \int (Dg)(t) f(t)\, dt + \sigma_0 f(0),$$

where g corresponds to $T_g \in D'(\mathbb{R})$, Dg corresponds to T_{Dg}, and $g' = T_g'$.

Proof. For $f \in C_c^\infty(\mathbb{R})$ we compute

$$\begin{aligned} g'(f) = -g(f') &= -\int g(t)f'(t)\,dt \\ &= -\int_{-\infty}^{0} g(t)f'(t)\,dt - \int_{0}^{\infty} g(t)f'(t)\,dt \\ &= -\left[g(t)f(t)\Big|_{-\infty}^{0} - \int_{-\infty}^{0} (Dg)(t)f(t)\,dt \right] \\ &\quad - \left[g(t)f(t)\Big|_{0}^{\infty} - \int_{0}^{\infty} (Dg)(t)f(t)\,dt \right] \\ &= \sigma_0 f(0) + \int (Dg)(t)f(t)\,dt. \end{aligned}$$

□

We can not conclude from *Proposition 2.3.5* that $Dg \in L^1_{\text{loc}}(\mathbb{R})$. For example, let

$$g(t) = \begin{cases} t\exp(ie^{-1/t^2}), & \text{if } t > 0, \\ 0, & \text{if } t < 0. \end{cases}$$

2.3.6 Definition. RADON MEASURES

a. The space $M(\mathbb{R}) \subseteq D'(\mathbb{R})$ of *Radon measures* is defined as

$$M(\mathbb{R}) = \{F' \in D'(\mathbb{R}) : F \in BV_{\text{loc}}(\mathbb{R})\}, \tag{2.3.8}$$

i.e., $T \in M(\mathbb{R})$ if there is $F \in BV_{\text{loc}}(\mathbb{R})$ for which $T = F'$, the first distributional derivative of F. The space $M_b(\mathbb{R}) \subseteq M(\mathbb{R})$ of *bounded Radon measures* is defined as

$$M_b(\mathbb{R}) = \{F' \in D'(\mathbb{R}) : F \in BV(\mathbb{R})\}. \tag{2.3.9}$$

(2.3.8) and (2.3.9) are well-defined since $BV_{\text{loc}}(\mathbb{R}) \subseteq L^1_{\text{loc}}(\mathbb{R})$, e.g., *Exercise 2.23.*

The elements of $M(\mathbb{R})$, resp., $M_b(\mathbb{R})$, are often referred to as *measures*, resp., *bounded measures.* It is also often the case that measures are denoted by Greek letters such as μ and ν. Further, if $\mu \in M(\mathbb{R})$ then, notationally, we write

$$\forall f \in C_c^\infty(\mathbb{R}), \quad \mu(f) = \int f(t)\,d\mu(t). \tag{2.3.10}$$

For a given measure μ the domain of function f for which $\mu(f)$ can be defined is much larger than $C_c^\infty(\mathbb{R})$.

b. A measure, resp., bounded measure, μ is *positive* if $\mu(f) \geq 0$ for all nonnegative functions $f \in C_c^\infty(\mathbb{R})$. The space of positive measures, resp., bounded positive measures, is denoted by $M_+(\mathbb{R})$, resp., $M_{b+}(\mathbb{R})$. It turns out that $\mu \in M(\mathbb{R})$ is positive if and only if each of the corresponding functions F (for which $F' = \mu$) is increasing on $\mathbb{R}$, e.g., *Section 2.7*.

c. Since the Heaviside function H is an element of $BV(\mathbb{R})$, we see that $\delta = H' \in M_b(\mathbb{R})$. *Lebesgue measure* $\mu \in M(\mathbb{R})\backslash M_b(\mathbb{R})$ is $\mu = F'$, where $F \in BV_{\text{loc}}(\mathbb{R})$ is defined as $F(t) = t$, i.e., Lebesgue measure can be identified with the distribution which is the constant 1, cf., *Definition A.4*.

Further, note that $L^1(\mathbb{R}) \subseteq M_b(\mathbb{R})$ and $L^1_{\text{loc}}(\mathbb{R}) \subseteq M(\mathbb{R})$, e.g., *Exercise 2.39*. The *norm* of $\mu \in M_b(\mathbb{R})$ is defined as

$$\|\mu\|_1 = \sup\{|\mu(f)| : f \in C_0(\mathbb{R}) \text{ and } \|f\|_{L^\infty(\mathbb{R})} \leq 1\}.$$

If $g \in L^1(\mathbb{R})$ then $\|g\|_{L^1(\mathbb{R})} = \|g\|_1$.

d. The definition of "measure" in part *a* is equivalent to that from real analysis. The underlying idea establishing this equivalence is the Riesz Representation Theorem, e.g., *Section 2.7*, cf., [Ben76, Appendix III], [Bou65, Chapter III.1], [Sch66, Chapitre II.4, Théorème II].

2.3.7 Remark. MULTIPOLES AND $\delta^{(n)}$

The first distributional derivative of δ, viz., δ', is not an element of $M(\mathbb{R})$ since $\delta \notin BV_{\text{loc}}(\mathbb{R})$. δ' is the *dipole* at the origin, and the distributions $\delta^{(n)}, n \geq 1$, are the *multipoles* which arise in several imporant applications. For example, in fluid mechanics, a dipole is the limiting case in fluid flow of a source and a sink of equal strength approaching each other under the constraint that the product of the distance between them and their strength is constant.

To quantify the fact that multipoles arise in applications, let us consider the case of electromagnetism and the potential due to point charges. The laws or equations of electromagnetism can be developed and formulated beginning with the notions of *length, mass, time*, and the electrical quantity of *charge*. Coulomb's law asserts that the force F_{jk} between two charges q_j, q_k at points $\mathbf{v}_j, \mathbf{v}_k \in \mathbb{R}^3$ is

$$F_{jk} = c\frac{q_j q_k}{|\mathbf{v}_j - \mathbf{v}_k|^2}\mathbf{u}_{j,k},$$

where c is the permittivity constant and $\mathbf{u}_{j,k} \in \mathbb{R}^3$ is the unit vector directed from $\mathbf{v}_j$ to $\mathbf{v}_k$. If there are charges $q_1, \cdots, q_n$ at $\mathbf{v}_1, \cdots, \mathbf{v}_n \in \mathbb{R}^3$ then the electric field $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ due to these charges can be formulated in terms of Coulomb's law and linear superposition. For conservative fields there is a potential function $V : \mathbb{R}^3 \longrightarrow \mathbb{R}$ whose gradient is F. In the case the charges are close to the origin and $r = |\mathbf{x}|$ is much larger than any $|\mathbf{v}_j|$ then $V(\mathbf{x})$ is of the form

$$V(\mathbf{x}) = c\left\{\frac{1}{r}\sum_{j=1}^{n} q_j + \frac{1}{r^2}\left[\frac{x_1}{r}\sum_{j=1}^{n} q_j v_{j1} + \frac{x_2}{r}\sum_{j=1}^{n} q_j v_{j2} + \frac{x_3}{r}\sum_{j=1}^{n} q_j v_{j3}\right]\right.$$

$$(2.3.11) \qquad \left. + \frac{1}{r^3}[\cdots] + \cdots\right\},$$

where $\mathbf{x} = (x_1, x_2, x_3), \mathbf{v}_j = (v_{j1}, v_{j2}, v_{j3}) \in \mathbb{R}^3$, e.g., [Con58, Chapter 8]. The first sum, Σq_j, is the total charge or *monopole moment* of the charge distribution. The next three sums of (2.3.11) are the components of the *dipole moment* of the charge distribution, and we have omitted writing the quadrupole moment, octupole moment, etc.

Now consider the particular example of two charges on the x-axis, viz., q at $\epsilon/2$ and $-q$ at $-\epsilon/2$. The monopole moment vanishes, and the dipole moment is $q\epsilon$. If $q = q(\epsilon) = -1/\epsilon$, then it is natural to formulate the notion of the *dipole of moment* 1 *at the origin* to be

$$-\lim_{\epsilon \to 0} \frac{1}{\epsilon}\left(\delta_{\epsilon/2} - \delta_{-\epsilon/2}\right),$$

which, in turn, is δ'. In fact,

$$-\lim_{\epsilon \to 0} \frac{1}{\epsilon}\left(\delta_{\epsilon/2} - \delta_{-\epsilon/2}\right)(f) = -\lim_{\epsilon \to 0} \frac{f(\frac{\epsilon}{2}) - f(-\frac{\epsilon}{2})}{\epsilon}$$

$$= -f'(0) = -\delta(f') = \delta'(f).$$

2.3.8 Example. THE PRINCIPAL VALUE DISTRIBUTION

a. We define the functional T by the formula

$$\forall f \in C_c^\infty(\mathbb{R}), \quad T(f) = \lim_{\epsilon \to 0} \int_{|t| \geq \epsilon} \frac{f(t)}{t}\, dt.$$

We shall verify that T is well-defined (part *b*), observe that $T \in D'(\mathbb{R})$ (*Exercise 2.43*), and show that $T \notin M(\mathbb{R})$. T is the *principal value distribution*, and is denoted by

$$pv(\frac{1}{t}).$$

b. To show that T is well-defined, note that

$$\int_{|t|\geq\epsilon} \frac{f(t)}{t}\, dt = \int_{R\geq|t|\geq\epsilon} \frac{f(t)-f(0)}{t-0}\, dt + f(0)\int_{R\geq|t|\geq\epsilon} \frac{dt}{t}$$

$$= \int_{R\geq|t|\geq\epsilon} \frac{f(t)-f(0)}{t-0}\, dt$$

for $f \in C_c^\infty(\mathbb{R})$ and some $R > 0$, depending on f. If

$$h_\epsilon(t) = \frac{f(t)-f(0)}{t} \mathbf{1}_{[-R,-\epsilon]\cup[\epsilon,R]}(t),$$

then $|h_\epsilon(t)| \leq \|f'\|_{L^\infty(\mathbb{R})}$ by the mean value theorem, the constant function $g \equiv \|f'\|_{L^\infty(\mathbb{R})}$ belongs to $L^1[-R,R]$, and, for each $t \neq 0$,

$$\lim_{\epsilon\to 0} h_\epsilon(t) = \frac{f(t)-f(0)}{t}\mathbf{1}_{[-R,R]}(t).$$

LDC applies, and, thus, $T(f)$ is well-defined.

c. Let $g(t) = \log|t|$, $t \in \mathbb{R}\backslash\{0\}$. Clearly, $g \in L^1_{\text{loc}}(\mathbb{R})$. In fact, if $a \geq 1$, then we have

$$\int_0^a |\log|t||\, dt = -\int_0^1 \log t\, dt + \int_1^a \log t\, dt \tag{2.3.12}$$

and $D(t\log t - t) = \log t$, so that the right side of (2.3.12) is finite. Obvious adaptations of this calculation yield the local integrability of g. Thus, $g \in D'(\mathbb{R})$ and we compute g'. For any $f \in C_c^\infty(\mathbb{R})$ we have

$$
\begin{aligned}
g'(f) &= -g(f') = -\int \log|t|\, f'(t)\, dt \\
&= -\int_{-\infty}^{0} \log|t|\, f'(t)\, dt - \int_0^\infty \log|t|\, f'(t)\, dt.
\end{aligned}
\tag{2.3.13}
$$

The second term on the right side of (2.3.13) is

$$
\begin{aligned}
-\int_0^\infty (\log t) f'(t)\, dt &= -\lim_{\epsilon\to 0}\int_\epsilon^\infty (\log t) f'(t)\, dt \\
&= -\lim_{\epsilon\to 0}\left[(\log t) f(t)\Big|_\epsilon^\infty - \int_\epsilon^\infty \frac{1}{t} f(t)\, dt\right] \\
&= \lim_{\epsilon\to 0}\left[f(\epsilon)\log\epsilon + \int_\epsilon^\infty \frac{1}{t} f(t)\, dt\right].
\end{aligned}
\tag{2.3.14}
$$

Note that $\log\epsilon\,(f(\epsilon) - f(0)) \longrightarrow 0$ as $\epsilon \longrightarrow 0$ since

$$\log\epsilon\,(f(\epsilon) - f(0)) = \epsilon(\log\epsilon)\frac{f(\epsilon) - f(0)}{\epsilon} \longrightarrow 0 \cdot f'(0) = 0, \quad \epsilon \longrightarrow 0.$$

Thus, replacing $f(\epsilon)\log\epsilon$ by

$$\log\epsilon\,(f(\epsilon) - f(0)) + f(0)\log\epsilon$$

in (2.3.14), we obtain

$$-\int_0^\infty (\log t) f'(t)\, dt = \lim_{\epsilon\to 0}\left[f(0)\log\epsilon + \int_\epsilon^\infty \frac{f(t)}{t}\, dt\right]_+,$$

where we have used the fact that $\log t \in L^1_{\text{loc}}(0,\infty)$ (a primitive of $\log t$ is $t\log t - t$, and therefore $\int_0^a \log t\, dt = a\log a - a$) and properties of sums of limits. $[\cdots]_+$ designates the fact that we are integrating over $(0,\infty)$ on the left side.

Similarly, for the first term on the right side of (2.3.13), we compute

$$-\int_{-\infty}^{0} \log|t|\, f'(t)\, dt = \lim_{\epsilon\to 0}\left[-f(0)\log\epsilon + \int_{-\infty}^{-\epsilon} \frac{f(t)}{t}\, dt\right]_{-},$$

where $[\cdots]_-$ designates that we are integrating over $(-\infty, 0)$ on the left side.

Since $\lim_{\epsilon\to 0}[\cdots]_+$ and $\lim_{\epsilon\to 0}[\cdots]_-$ exist we see that

$$\begin{aligned}\lim_{\epsilon\to 0}[\cdots]_+ + \lim_{\epsilon\to 0}[\cdots]_- &= \lim_{\epsilon\to 0}([\cdots]_+ + [\cdots]_-)\\ &= \lim_{\epsilon\to 0}\int_{|t|>\epsilon} \frac{f(t)}{t}\, dt,\end{aligned} \tag{2.3.15}$$

cf., part *b*. On the other hand, equation (2.3.13) shows that the left side of (2.3.15) is $g'(f)$, where $g(t) \equiv \log|t|$, $t \in \mathbb{R}\backslash\{0\}$. Therefore, since the right side of (2.3.15) defines the principal value distribution, we obtain

$$(\log|t|)' = pv(\frac{1}{t}). \tag{2.3.16}$$

Clearly, $\log|t| \notin BV_{\text{loc}}(\mathbb{R})$ and so $pv(\frac{1}{t}) \notin M(\mathbb{R})$, e.g., *Exercise 2.30*.

2.3.9 Example. FIRST MOMENT OF A MEASURE

a. We shall verify that

$$\int t\, d\mu(t) = 1 - \int_0^1 F(t)\, dt, \tag{2.3.17}$$

where $\mu = F' \in M_{b+}(\mathbb{R})$ has the property that $F = 0$ on $(-\infty, 0]$, $F = 1$ on $[1, \infty)$, and F is increasing on $[0, 1]$. In fact, if $f \in C_c^\infty(\mathbb{R})$, then

$$\begin{aligned}(tF'(t))(f(t)) &= F'(t)(tf(t)) = -F(t)(tf'(t) + f(t))\\ &= -\int_0^1 tf'(t)F(t)\, dt - \int_1^\infty tf'(t)\, dt - \int_0^1 f(t)F(t)\, dt\\ -\int_1^\infty f(t)\, dt &= -\int_0^1 tf'(t)F(t)\, dt + f(1) - \int_0^1 f(t)F(t)\, dt.\\ &(\delta_1 - \mathbf{1}_{[0,1]}F)(f) - \int_0^1 tf'(t)F(t)\, dt.\end{aligned} \tag{2.3.18}$$

Consequently, if $f = 1$ on $[0,1]$ we obtain (2.3.17) since the left side of (2.3.18) is $\int t\, d\mu(t)$ in this case.

b. Let F in part a be defined on $[0,1]$ as the Cantor functions f_C for the $\frac{1}{3}$-Cantor set C, cf., *Remark 1.4.2a*. In this case, $DF = 0$ a.e. and $F' = \mu_C \in M_{b+}(\mathbb{R})$. μ_C is the *Cantor measure*, and has the property that $\int d\mu_C(t) = 1$ and $\operatorname{supp} \mu_C = C$, e.g., [Ben75, pages 28-29], [KS63, Chapitre I], [Zyg59, Volume I, pages 194-196]. $F \in BV(\mathbb{R}) \cap C(\mathbb{R})$, but $F \notin AC_{\text{loc}}(\mathbb{R})$.

Using (2.3.17) we have the fact that

$$\int t\, d\mu_C(t) = \frac{1}{2},$$

e.g., *Exercise 2.44*.

2.3.10 Example. DATA COMPRESSION AND LATERAL NETWORKS

Suppose f is a signal which can not be understood by direct examination or which must be reconstructed after less than ideal transmission. f could be a speech signal, radar data, an image, MRI or EEG data, etc. Let $F(t, \omega)$ be data collected from the signal. F could be a spectrogram, a scalogram, a radar ambiguity function, etc., e.g., [BF94], [Mey91], [Rih85]. The goal is to understand or reconstruct f from F or a modified version of F. A possible modification of F, because of excessive volume, is the compressed data

$$H \circ F(t, \omega)$$

consisting of 0s and 1s. Here, $H = H(t)$. In some systems, the variables (t, ω), which could be time-frequency, or time-scale, etc., are interrelated because of physical constraints. For example, in the auditory system and for ω being a scale variable, the modified data can be of the form

$$\partial_\omega (H \circ F)(t, \omega) = [\delta \circ F(t, \omega)] \partial_\omega F(t, \omega),$$

e.g., [BT93]; in particular, compositions of the form $\delta \circ F$ arise in a natural way.

2.4 The Fourier transform of distributions

In Equations (2.3.1) and (2.3.2) we used the integration by parts formula to motivate the definition of the distributional derivative. In this section we shall use the Parseval formula to motivate the definition of the distributional Fourier transform. In this role, such classical formulas become *creative formulas* in the sense of the following remark.

2.4.1 Remark. CREATIVE FORMULAS

What do we mean by a creative formula? The Pythagorean Theorem is an example of such a formula. It leads to the proper definition of distance between points in Euclidean space. It is fundamental in defining lengths of curves, areas of surfaces, etc. It extends to defining the notion of distance by means of the L^2-norm in the space of square integrable functions. It provides a fundamental guideline in the geometry of Hilbert space. It is a backdrop from which various non-Euclidean geometries are assessed. It inspires new developments in the sciences, with concepts such as *energy* or fields of study such as *quantum mechanics*, e.g., [vN55].

2.4.2 Definition. FOURIER TRANSFORM OF DISTRIBUTIONS

a. Recall the Parseval formula (1.10.10),

$$\forall f, g \in L^2(\mathbb{R}), \quad \int \widehat{g}(\gamma)\overline{\widehat{f}(\gamma)}\, d\gamma = \int g(t)\overline{f(t)}\, dt. \tag{2.4.1}$$

In the notation of (2.1.5), and thinking of f as a test function, we rewrite (2.4.1) as

$$\widehat{g}(\overline{\widehat{f}}) = g(\overline{f}). \tag{2.4.2}$$

In the notation of (2.2.4), we rewrite (2.4.2) as

$$T_{\widehat{g}}(\overline{\widehat{f}}) = T_g(\overline{f}). \tag{2.4.3}$$

Since $L^2(\mathbb{R})$ is a Hilbert space, e.g., [GG81], it is customary to introduce the notation of inner products and write

$$\langle g, f \rangle = \int g(t)\overline{f(t)}\, dt.$$

Thus, (2.4.1) is

$$\langle \widehat{g}, \widehat{f} \rangle = \langle g, f \rangle,$$

and (2.4.3) is

$$\langle T_{\widehat{g}}, \widehat{f} \rangle = \langle T_g, f \rangle.$$

b. Motivated by part *a*, the *Fourier transform* $\widehat{T}$ of a distribution T is *formally defined* by the equation

$$\widehat{T}(\overline{\widehat{f}}) = T(\overline{f}) \tag{2.4.4}$$

or, equivalently,

$$\langle \widehat{T}, \widehat{f} \rangle = \langle T, f \rangle, \tag{2.4.5}$$

for all functions f in an appropriate space of test functions, cf., *Definition 2.2.1b.*

c. We have been purposely vague about specifying the distributions and test functions in part *b*, since the meaning of (2.4.4) is difficult to formulate for *all* distributions. In fact, if $T \in D'(\mathbb{R})$ and $f \in C_c^\infty(\mathbb{R})$, then the right side of (2.4.4) is well defined, whereas $\widehat{f} \notin C_c^\infty(\widehat{\mathbb{R}})$, e.g., *Exercise 2.5.* Thus, $\widehat{T}$ on the left side is not necessarily defined on $C_c^\infty(\widehat{\mathbb{R}})$, and would not necessarily be a distribution. Even if we worked with $\widehat{f} \in C_c^\infty(\widehat{\mathbb{R}})$ and $\widehat{T} \in D'(\widehat{\mathbb{R}})$, then T would not be well-defined on $C_c^\infty(\mathbb{R})$ for the same reason.

This quandary is resolved by introducing the *Schwartz space* of test functions and the space of *tempered distributions* in the following material.

2.4.3 Definition. THE SCHWARTZ SPACE

a. An infinitely differentiable function $f : \mathbb{R} \to \mathbb{C}$ is an element of the *Schwartz space* $\mathcal{S}(\mathbb{R})$ if

(2.4.6)
$$\forall n = 0, 1, \ldots, \quad \|f\|_{(n)} = \sup_{0 \le j \le n} \sup_{t \in \mathbb{R}} (1 + |t|^2)^n |f^{(j)}(t)| < \infty.$$

b. Note that the Gaussian $g(t) = \frac{1}{\sqrt{\pi}} e^{-t^2} \in \mathcal{S}(\mathbb{R})$, and that

$$C_c^\infty(\mathbb{R}) \subseteq \mathcal{S}(\mathbb{R}) \subseteq L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \cap A(\mathbb{R}).$$

c. Using (2.4.6) we define the function $\rho : \mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R}) \to \mathbb{R}^+$ as

$$\forall f, g \in \mathcal{S}(\mathbb{R}), \quad \rho(f,g) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{\|f-g\|_{(n)}}{1+\|f-g\|_{(n)}}. \tag{2.4.7}$$

The following theorem is a basic result about the Schwartz space, and its verification is left as *Exercise 2.45.*

2.4.4 Theorem. $\mathcal{S}(\mathbb{R})^\wedge = \mathcal{S}(\widehat{\mathbb{R}})$

a. The mapping $\mathcal{S}(\mathbb{R}) \longrightarrow \mathcal{S}(\widehat{\mathbb{R}})$, $f \longmapsto \widehat{f}$, *is a bijection.*

b. The mapping ρ defined by (2.4.7) is a metric on $\mathcal{S}(\mathbb{R})$, and, as such, the metric space $\mathcal{S}(\mathbb{R})$ is complete.

c. The Fourier transform mapping of part a, where $\mathcal{S}(\mathbb{R})$ is given the metrizable topology of part b is bicontinuous. Thus, the Fourier transform,

$$\begin{array}{rcl} \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathcal{S}(\widehat{\mathbb{R}}) \\ f & \longmapsto & \widehat{f}, \end{array}$$

is a metric space isomorphism, cf., Theorem 1.10.2 *and* Exercise 2.47.

2.4.5 Definition. TEMPERED DISTRIBUTIONS

a. A linear functional,

$$\begin{array}{rcl} T : \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathbb{C} \\ f & \longmapsto & T(f), \end{array}$$

is a *tempered distribution* if $\lim_{n\to\infty} T(f_n) = 0$ for every sequence $\{f_n\} \subseteq \mathcal{S}(\mathbb{R})$ satisfying the properties

$$\forall k, m \geq 0, \quad \lim_{n\to\infty} \|t^m f_n^{(k)}(t)\|_{L^\infty(\mathbb{R})} = 0. \tag{2.4.8}$$

The space of all tempered distributions on $\mathbb{R}$ is denoted by $\mathcal{S}'(\mathbb{R})$.

b. It is not difficult to prove that "$f_n \to 0$" in the sense of (2.4.8) if and only if $\lim_{n\to\infty} \rho(f_n, 0) = 0$, where ρ is defined in (2.4.7), e.g., *Exercise 2.45.*

c. In light of our terminology, it is natural to ask if $\mathcal{S}'(\mathbb{R}) \subseteq D'(\mathbb{R})$. The answer is "yes". The proof of this fact depends on functional analysis duality results (*Remark B.16*) and the following easily verifiable facts: the mapping

$$\begin{array}{rcl} C_c^\infty(\mathbb{R}) & \longrightarrow & \mathcal{S}(\mathbb{R}) \\ f & \longmapsto & f, \end{array} \tag{2.4.9}$$

is continuous, and

$$\overline{C_c^\infty(\mathbb{R})} = \mathcal{S}(\mathbb{R}), \tag{2.4.10}$$

cf., *Example 2.4.6h*. (2.4.9) means that if "$f_n \to 0$" in the sense of *Definition 2.2.3a* then "$f_n \to 0$" in the sense of (2.4.8). The density (2.4.10) means that if $f \in \mathcal{S}(\mathbb{R})$ then there is $\{f_n\} \subseteq C_c^\infty(\mathbb{R})$ for which "$f - f_n \to 0$" in the sense of (2.4.8) (in fact, let $f_n = fg_n$ where $g_n \in C_c^\infty(\mathbb{R})$ is even, $g_n = 1$ on $[-n, n]$, $\operatorname{supp} g_n \subseteq [-(n+1), n+1]$, and g_n has the same shape on $[n, n+1]$ as g_m has on $[m, m+1]$).

The following section contains some advanced material and relevant references, but no details.

2.4.6 Example. DISTRIBUTIONS: POTPOURRI AND TITILLATION

a. Let $T \in D'(\mathbb{R})$. Then $T \in \mathcal{S}'(\mathbb{R})$ if and only if there is $g \in C_b(\mathbb{R})$, and $k, m \geq 0$, such that

$$T = f^{(k)} \text{ (distributionally)},$$

where $f(t) = (1 + t^2)^{m/2} g(t)$ [Sch66, Chapitre VII, §4, Théorème VI]. For example, if $f \in C_b(\mathbb{R})$, then $f^{(k)} \in \mathcal{S}'(\mathbb{R})$. In particular, taking $f(t) = tH(t)$, we see that $\delta^{(k)} \in \mathcal{S}'(\mathbb{R})$ for each $k \geq 0$.

b. If $T \in D'(\mathbb{R})$ and $\operatorname{supp} T \subseteq \mathbb{R}$ is compact, then $T \in \mathcal{S}'(\mathbb{R})$ [Sch66]. The space of distributions having compact support is denoted by $\mathcal{E}'(\mathbb{R})$, and so

$$\mathcal{E}'(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R}).$$

c. If $g(t) = e^{|t|}$ then $g \notin \mathcal{S}'(\mathbb{R})$.

d. For each $p \in [1, \infty]$,

$$L^p(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R}).$$

e. $\mathcal{S}'(\mathbb{R})$ is the smallest subspace of $D'(\mathbb{R})$ which contains $L^1(\mathbb{R})$ and which is invariant under differentiation and multiplication by polynomials, cf., *Definition 2.5.7.*

f. *Spectral synthesis.* If the norm of $f \in A(\mathbb{R})$ is defined as $\|f^\vee\|_{L^1(\mathbb{R})}$ then $A(\mathbb{R})$ is a Banach space (in fact, it is a Banach algebra under pointwise multiplication of functions). The dual space $A'(\mathbb{R})$ of $A(\mathbb{R})$, i.e., the space of continuous linear functionals $T : A(\mathbb{R}) \to \mathbb{C}$, is the space of *pseudo-measures*. We have the inclusion

$$M_b(\mathbb{R}) \subseteq A'(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R}).$$

An important area of harmonic analysis, rooted in physical considerations and with many unsolved problems, is *spectral synthesis*, e.g., [Ben75], [Beu89], [KS63], [Kah70], [Kat76]. Its creation, canonicity, and depth were revealed to normal man by Beurling. In deceptive form, an essential *problem of spectral synthesis* is to determine the pseudo-measures $T \in A'(\mathbb{R})$ for which there is a sequence $\{\mu_n\} \subseteq M_b(\mathbb{R})$ satisfying the properties that $\operatorname{supp} \mu_n \subseteq \operatorname{supp} T$ for each n, and

$$\forall f \in A(\mathbb{R}), \quad \lim_{n \to \infty} \mu_n(f) = T(f).$$

g. *Riemann Hypothesis.* The most elebrated problem in analytic number theory is to settle the validity or not of the *Riemann Hypothesis.*

The *Riemann zeta function* $\zeta(s)$ is defined as

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}; \quad \operatorname{Re} s > 1,$$

and it has an analytic continuation, whereby it is analytic on $\mathbb{C}\backslash\{1\}$ and has a simple pole at $s = 1$. The *Riemann Hypothesis* is the statement that the complex zeros of $\zeta(s)$ all have real part equal to $\frac{1}{2}$, e.g., [Edwa74], [Tit51].

The *Weil distribution* $W \in D'(\mathbb{R})$ is defined as

$$\forall f \in C_c^\infty(\mathbb{R}), \quad W(f) = \Sigma_\rho \Phi_f(\rho),$$

where $\zeta(\rho) = 0$ for $0 \leq \operatorname{Re}\rho \leq 1$, and where

$$\Phi_f(s) = \int f(t)e^{(s-\frac{1}{2})t}\,dt;$$

this integral is the bilateral Laplace transform of $f(t)e^{-\frac{1}{2}t}$. Tempered distributions arise in the following result. *The Riemann Hypothesis is valid if and only if* $W \in \mathcal{S}'(\mathbb{R})$ [Ben80], [Joy86, page 6].

The Riemann Hypothesis can be considered a strong form of the Prime Number Theorem; and Wiener's Tauberian Theorem, a fundamental result in harmonic analysis, is an indispensable tool in this type of analysis, e.g., [Ben75, *Section 2.3*].

h. If $D'(\mathbb{R})$ is taken with the canonical dual space topology from $C_c^\infty(\mathbb{R})$, then the natural injection $C_c^\infty(\mathbb{R}) \longrightarrow D'(\mathbb{R})$, $f \longmapsto T_f$, is continuous. Further, if the topological vector space X has the properties that $C_c^\infty(\mathbb{R}) \subseteq X \subseteq D'(R)$, the inclusions are continuous, and $C_c^\infty(\mathbb{R})$ is dense in X, then the dual X' is a subspace of $D'(\mathbb{R})$.

2.4.7 Definition. THE FOURIER TRANSFORM OF TEMPERED DISTRIBUTIONS

a. The *Fourier transform* of $T \in \mathcal{S}'(\mathbb{R})$ is $\hat{T}$, defined by

$$\forall f \in \mathcal{S}(\mathbb{R}), \quad \hat{T}(\overline{\hat{f}}) = T(\overline{f}). \tag{2.4.11}$$

b. Equation (2.4.11) is a quantified version of (2.4.4), and there is the equivalent quantified analogue of (2.4.5). Because of *Theorem 2.4.4*, we know that $\hat{f} \in \mathcal{S}(\mathbb{R})$, and, thus, it is straightforward to show that $\hat{T} \in \mathcal{S}'(\mathbb{R})$. In fact, the mapping,

$$\begin{array}{ccc} \mathcal{S}'(\mathbb{R}) & \longrightarrow & \mathcal{S}'(\hat{\mathbb{R}}), \\ T & \longrightarrow & \hat{T} \end{array} \tag{2.4.12}$$

is a linear bijection. Further, there is a natural convergence criterion (topology) on $\mathcal{S}'$ so that (2.4.12) is bicontinuous, e.g., [Hor66], [Sch66], and, hence, (2.4.12) is a topological vector space isomorphism.

We now want to compute Fourier transforms with this new definition; and, in particular, in light of *Example 2.4.6d*, we want to make sure that the new definition reduces to the classical definition of *Chapter 1.*

2.4.8 Example. THE FOURIER TRANSFORM OF $L^1(\mathbb{R})$ AND $\delta^{(n)}$

a. Let $g \in L^1(\mathbb{R})$. We shall verify that $\widehat{T}_g = T_{\widehat{g}}$ so that the distributional Fourier transform is a generalization of the usual, i.e., $L^1(\mathbb{R})$, definition. For each $f \in S(\mathbb{R})$ we compute

$$\begin{aligned} \langle \widehat{T}_g, \widehat{f} \rangle = \langle T_g, f \rangle &= \int g(t)\overline{f(t)}\,dt \\ &= \int \widehat{g}(\gamma)\overline{\widehat{f}(\gamma)}\,d\gamma = \langle T_{\widehat{g}}, \widehat{f} \rangle, \end{aligned} \tag{2.4.13}$$

where we have used the Parseval formula for $g \in L^1(\mathbb{R})$ and $\widehat{f} \in L^1(\widehat{\mathbb{R}})$, e.g., *Proposition 1.10.4*. By our discussion of equality for distributions we can conclude that $\widehat{T}_g = T_{\widehat{g}}$ since (2.4.13) is true for each $\widehat{f} \in \mathcal{S}(\widehat{\mathbb{R}})$.

b. Let $T = \delta^{(n)}$. From *Example 2.4.6a* or *b*, we know that $T \in \mathcal{S}'(\mathbb{R})$. To evaluate $\widehat{T}$ we compute

$$\langle \widehat{\delta^{(n)}}, \widehat{f} \rangle = \langle \delta^{(n)}, f \rangle = (-1)^n \langle \delta, f^{(n)} \rangle$$

$$= (-1)^n \langle \delta(t), \int (2\pi i\gamma)^n \widehat{f}(\gamma) e^{2\pi i t\gamma}\,d\gamma \rangle$$

$$= (-1)^n \langle (\bar{i})^n (2\pi\gamma)^n, \widehat{f}(\gamma) \rangle = \langle (2\pi i\gamma)^n, \widehat{f}(\gamma) \rangle$$

for each $\widehat{f} \in S(\widehat{\mathbb{R}})$. Consequently, we have

$$\left(\delta^{(n)}\right)^{\wedge}(\gamma) = (2\pi i\gamma)^n \tag{2.4.14}$$

for each $n \in \mathbb{N} \cup \{0\}$. In particular, $\widehat{\delta} = 1$, which is compatible with our discussion of approximate identities.

An important point about (2.4.14) is that $(\delta^{(n)})^{\wedge}$ has polynomial growth at $\pm\infty$ as opposed to the behavior of $\widehat{f}$ at $\pm\infty$ for $f \in L^1(\mathbb{R})$ (or $L^2(\mathbb{R})$).

2.4.9 Example. THE FOURIER TRANSFORM OF THE HEAVISIDE FUNCTION

The Heaviside function $H \in L^1_{\text{loc}}(\mathbb{R}) \cap S'(\mathbb{R}), \delta \in M_b(\mathbb{R}) \subseteq S'(\mathbb{R})$, and $pv(\frac{1}{t}) \in S'(\mathbb{R})$, whereas $pv(\frac{1}{t}) \notin M(\mathbb{R})$. The Fourier transforms of any one of these distributions have a nice formulation in terms of the others or their Fourier transforms. We shall verify

$$(2.4.15)\qquad \begin{aligned} H(t) &\longleftrightarrow \tfrac{1}{2\pi i}pv(\tfrac{1}{\gamma}) + \tfrac{1}{2}\delta(\gamma) \\ \tfrac{1}{2}\delta(t) - \tfrac{1}{2\pi i}pv(\tfrac{1}{t}) &\longleftrightarrow H(\gamma). \end{aligned}$$

For $f \in S(\mathbb{R})$ and $\hat{f} = F$, we compute

$$\langle (pv(\frac{1}{t}))^\wedge(\gamma), F(\gamma)\rangle = \langle pv(\frac{1}{t}), f(t)\rangle = \lim_{\epsilon\to 0}\int_{|t|\geq\epsilon} \frac{\overline{f(t)}}{t}\,dt$$

$$(2.4.16)\qquad = \lim_{\epsilon\to 0}\iint_{|t|\geq\epsilon} \overline{F(\gamma)}\frac{e^{-2\pi it\gamma}}{t}\,d\gamma dt$$

$$= \lim_{\epsilon\to 0}\int \overline{F(\gamma)}\int_{|t|\geq\epsilon}\frac{e^{-2\pi it\gamma}}{t}\,dtd\gamma.$$

Now note that if $\gamma \neq 0$ then

$$\lim_{\epsilon\to 0,\, T\to\infty}\int_{1/\epsilon}^{T}\frac{\cos 2\pi t\gamma}{t}\,dt = 0,$$

where $1/\epsilon < T$. In fact, by the second mean value theorem for integrals (*Lemma 1.7.3*), we have

$$\int_{1/\epsilon}^{T}\frac{\cos 2\pi t\gamma}{t}\,dt = \epsilon\int_{1/\epsilon}^{\xi}\cos 2\pi t\gamma\,dt + \frac{1}{T}\int_{\xi}^{T}\cos 2\pi t\gamma\,dt$$

$$= \frac{\epsilon\sin 2\pi t\gamma}{2\pi\gamma}\Big|_{t=1/\epsilon}^{\xi} + \frac{\sin 2\pi t\gamma}{2\pi T\gamma}\Big|_{t=\xi}^{T};$$

and for fixed $\gamma \neq 0$ this last term tends to 0 as $\epsilon \to 0$ and $T \to \infty$. A similar calculation for the domain $[-S, -\frac{1}{\epsilon}]$ (where $S > \frac{1}{\epsilon}$) combined with an application of LDC allow us to write the right side of (2.4.16) as

$$(2.4.17)\qquad \int \overline{F(\gamma)}\left[\lim_{\epsilon\to 0}\int_{\epsilon\leq|t|\leq 1/\epsilon}\frac{e^{-2\pi it\gamma}}{t}\,dt\right]d\gamma.$$

(The use of LDC requires certain hypotheses, e.g., *Exercise 2.19.*)

Next, we compute

$$\lim_{\epsilon\to 0}\int_{\epsilon\le|t|\le 1/\epsilon}\frac{e^{-2\pi it\gamma}}{t}\,dt = -\lim_{\epsilon\to 0} i\int_{\epsilon\le|t|\le 1/\epsilon}\frac{\sin 2\pi t\gamma}{t}\,dt$$

(2.4.18)

$$= \begin{cases} -\pi i, & \text{if } \gamma > 0, \\ \pi i, & \text{if } \gamma < 0. \end{cases}$$

We can write the right side of (2.4.18) as

$$\pi i - 2\pi i H(\gamma) = \pi i\widehat{\delta}(\gamma) - 2\pi i H(\gamma)$$

for $\gamma \neq 0$, and thus we have proved that

$$\forall f \in S(\mathbb{R}), \quad \langle (pv(\frac{1}{t}))^\wedge(\gamma), F(\gamma)\rangle = \langle \pi i\widehat{\delta}(\gamma) - 2\pi i H(\gamma), F(\gamma)\rangle,$$

where $\widehat{f} = F$. Therefore, $(pv(\frac{1}{t}))^\wedge(\gamma) = \pi i\widehat{\delta}(\gamma) - 2\pi i H(\gamma)$ and (2.4.15) is obtained.

2.4.10 Remark. PERSPECTIVE ON $\widehat{H}$

Our calculation of (2.4.15) is relatively honest but too long-winded because of the distributional setup. The formula was certainly known and used long before distribution theory, and several formal, short, and essentially correct calculations give the result, e.g., (2.4.18) contains the essential details.

2.4.11 Remark. EIGENFUNCTIONS

a. We defined the Hermite polynomials $H_n, n \geq 0$, in *Exercise 1.26*, and computed the eigenvalues of the Fourier transform mapping $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\widehat{\mathbb{R}})$ in *Example 1.10.12*. The *Hermite functions* $h_n(t) \equiv e^{-\pi t^2} H_n(2\sqrt{\pi}\, t)$ are the eigenfunctions of $\mathcal{F}$, and, in fact, $\widehat{h}_n = (-i)^n h_n$. Further, $\{h_n/\|h_n\|_{L^2(\mathbb{R})} : n \geq 0\}$ is an orthonormal basis of $L^2(\mathbb{R})$. This material was developed by Norbert Wiener [Wie33, pages 51-71] to give his proof of the Parseval-Plancherel Theorem, cf., [Wie33, page 70] for an interesting historical note and [Wie81, article 29d], which is related to earlier work of Hermann Weyl and which establishes the Fourier transform of fractional order in terms of Hermite polynomials.

b. Since the Hermite functions are contained in $\mathcal{S}(\mathbb{R})$, they are also eigenfunctions of the mapping $\mathcal{F} : \mathcal{S}(\mathbb{R}) \longrightarrow \mathcal{S}(\widehat{\mathbb{R}})$. In this context it is natural to determine the eigenfunctions of the Fourier transform mapping $\mathcal{F} : \mathcal{S}'(\mathbb{R}) \longrightarrow \mathcal{S}'(\widehat{\mathbb{R}})$. In this regard, we do note now that $\Sigma\delta_n \in M(\mathbb{R}) \cap \mathcal{S}'(\mathbb{R})$ and

$$\left(\sum \delta_n\right)^\wedge = \sum \delta_n. \tag{2.4.19}$$

Equation (2.4.19) is a form of the *Poisson Summation Formula*, e.g., *Theorem 3.10.8.*

2.5 Convolution of distributions

In *Section 1.5* we defined the convolution $g * h$ for $g, h \in L^1(\mathbb{R})$. In this section, we shall see to what extent we can define the *convolution* $S*T$ for $S, T \in D'(\mathbb{R})$. In *Proposition 1.5.2* we established *the exchange formula* $(g * h)^\wedge = \widehat{g}\widehat{h}$ for $g, h \in L^1(\mathbb{R})$. In this section we shall see to what extent this formula is valid for distributions; and, in view of the right side of the exchange formula, we shall see to what extent the *multiplication of distributions* is well-defined.

Convolution and the exchange formula are major components of the operational calculus (*Section 2.6*). Further, the formulation of convolution and multiplication of distributions is a formidable mathematical task which we shall really only address at the motivational and computational levels.

2.5.1 Definition. CONVOLUTION

a. If $g, h \in L^1(\mathbb{R})$, then for each $f \in C_c^\infty(\mathbb{R})$ we have

$$\begin{aligned} (g * h)(f) &= \int g * h(t) f(t)\, dt = \iint g(t-u) h(u) f(t)\, du dt \\ &= \int h(u) \left(\int g(t-u) f(t)\, dt \right) du \\ &= \int h(u) \int g(v) f(u+v)\, dv du. \end{aligned} \tag{2.5.1}$$

b. Because of the calculation in part *a*, we define the *convolution* $S * T$ of certain distributions S and T by

$$\forall f \in C_c^\infty(\mathbb{R}), \quad (S * T)(f) = T(u)\big(S(v)(f(u+v))\big), \tag{2.5.2}$$

where, although T is not necessarily an ordinary function on $\mathbb{R}$, we write $T(u)$ to indicate its dependence on the u variable.

c. We have been deliberately noncommital in part *b* about which distribution S and T we can convolve; and whether, in case $S*T$ exists, it is an element of $D'(\mathbb{R})$, cf., *Theorem 2.5.4*. In fact,

$$\forall f \in C_c^\infty(\mathbb{R})\backslash\{0\}, \quad f(u+v) \notin C_c^\infty(\mathbb{R} \times \mathbb{R}); \tag{2.5.3}$$

and $S(v)(f(u+v))$ is not necessarily an element of $C_c^\infty(\mathbb{R})$ (as a function of u), so that the right side of (2.5.2) may not make sense. To visualize (2.5.3) it is instructive to let $\operatorname{supp} f = [a, b]$, $a > 0$, and then note that the support X of $f(u+v)$ as a function of two variables is the diagonal strip of $\mathbb{R} \times \mathbb{R}$ whose intersection with both the u and v axes is $[a, b]$, i.e.,

$$X = \{(u, v) \in \mathbb{R} \times \mathbb{R} : a \leq u + v \leq b\}. \tag{2.5.4}$$

2.5.2 Example. PROPERTIES OF CONVOLUTION

a. The calculation in part *a* of *Definition 2.5.1* shows that (2.5.2) generalizes the definition of convolution in $L^1(\mathbb{R})$. It can also be shown that $\mu * \nu \in M_b(\mathbb{R})$ for $\mu, \nu \in M_b(\mathbb{R})$. This can be proved measure theoretically, e.g., [Rud66], or by our definition of $M_b(\mathbb{R})$ in *Definition 2.3.6* combined with the *Riesz Representation Theorem* in *Section 2.7*, e.g., *Exercise 2.49*.

b. If $S, T \in D'(\mathbb{R})$ and $S * T$, defined by (2.5.2), is a well-defined element of $D'(\mathbb{R})$, then $T * S$ is also a well-defined element of $D'(\mathbb{R})$, and $S * T = T * S$. Thus, convolution is a *commutative* operation.

c. Convolution is *not* necessarily *associative*. In fact,

$$(1 * \delta') * H = 0 * H = 0$$

and

$$1 * (\delta' * H) = 1 * \delta = 1.$$

d. Let $S = T = 1$. Then $S, T \in \mathcal{S}'(\mathbb{R}) \cap M(\mathbb{R})$ and

$$1(v)(f(u+v)) = \int f(t)\,dt.$$

Therefore, $S * T(f) = \infty$ for each nonnegative $f \in C_c^\infty(\mathbb{R})\backslash\{0\}$, and, in particular, $S * T$ does not exist.

e. If $T \in D'(\mathbb{R})$ and $g \in C_c^\infty(\mathbb{R})$ then $T * g \in C^\infty(\mathbb{R})$.

2.5.3 Definition. TRANSLATION

a. If $g \in L^1_{\text{loc}}(\mathbb{R})$, then for each fixed $t \in \mathbb{R}$ and each $f \in C_c^\infty(\mathbb{R})$, we have

$$\begin{aligned}(T_{\tau_t g})(f) &= \int g(v-t)f(v)\,dv = \int g(u)f(u+t)\,du \\ &= T_g(u)(f(u+t)) = T_g(u)\Big(\delta_t(v)(f(u+v))\Big) \\ &= (\delta_t * T_g)(f).\end{aligned}$$

b. Because of the calculation in part *a*, we define the *translation* $\tau_t T$ of $T \in D'(\mathbb{R})$, by $t \in \mathbb{R}$ to be

$$\tau_t T = \delta_t * T. \tag{2.5.5}$$

c. It should be noted that the right side of (2.5.5) is a well-defined element of $D'(\mathbb{R})$, although there is something to prove, e.g., *Exercise 2.51*.

We have seen that $\delta_t * T \in D'(\mathbb{R})$ for arbitrary distributions T, whereas $S * T$ is not necessarily a distribution for $S, T \in \mathcal{S}'(\mathbb{R})$ (*Example 2.5.2d*). The following result records some of the elementary theory for the existence of $S * T$, e.g., [Hor66, pages 365-401, esp., pages 382-388], [Sch66, Chapitre VI].

2.5.4 Theorem. EXISTENCE OF CONVOLUTION

a. Let $S, T \in D'(\mathbb{R})$ satisfy the property that for each compact set $C \subseteq \mathbb{R}$,

$$((\operatorname{supp} S) \times (\operatorname{supp} T)) \cap C^\Delta$$

is compact in $\mathbb{R} \times \mathbb{R}$, *where* $C^{\Delta} = \{(u, v) \in \mathbb{R} \times \mathbb{R} : u + v \in C\}$, *cf.*, *(2.5.3) and (2.5.4). Then* $S * T \in D'(\mathbb{R})$.

b. Let $S \in \mathcal{E}'(\mathbb{R})$ *and* $T \in D'(\mathbb{R})$. *Then* $S * T \in D'(\mathbb{R})$.

2.5.5 Proposition.

Let $T \in D'(\mathbb{R})$ *and let* $n \geq 0$. *Then*

$$T * \delta^{(n)} = T^{(n)};$$

in particular, $T * \delta = T$.

Proof. For each $f \in C_c^{\infty}(\mathbb{R})$, we compute

$$\begin{aligned} T * \delta^{(n)}(f) &= T(u)\left(\delta^{(n)}(v)(f(u+v))\right) \\ &= (-1)^n T(u)\left(\delta(v)(f^{(n)}(u+v))\right) \\ &= (-1)^n T(u)(f^{(n)}(u)) = T^{(n)}(f). \end{aligned}$$

□

2.5.6 Definition. Exchange Formula

a. If $g, h \in L^1(\mathbb{R})$, then, as mentioned at the beginning of this section, we have the *exchange formula*,

$$(g * h)^{\wedge} = \hat{g}\hat{h}. \tag{2.5.6}$$

Besides the direct proof in *Section 1.5*, we could also prove it "distributionally" as follows. For each $f \in C_c^{\infty}(\mathbb{R})$,

$$(T_g * T_h)^{\wedge}\left(\overline{f}\right) = T_g * T_h\left(\overline{\hat{f}}\right) = T_g(u)\left(T_h(v)\left(\overline{\hat{f}}(u+v)\right)\right)$$

$$= T_g(u)\left(T_h(v)\left(\int \overline{\hat{f}}(\gamma) e^{-2\pi i (u+v)\gamma}\, d\gamma\right)\right) \tag{2.5.7}$$

$$= \left[T_g(u)\left(e^{-2\pi i u\gamma}\, T_h(v)(e^{-2\pi i v\gamma})\right)\right]\left(\overline{\hat{f}}(\gamma)\right) = \left(\hat{T}_g \hat{T}_h\right)\left(\overline{f}\right).$$

The calculation (2.5.7) is valid for each $f \in C_c^{\infty}(\mathbb{R})$, and so we can conclude that $(T_g * T_h)^{\wedge} = \hat{T}_g \hat{T}_h$, which is precisely (2.5.6).

Although the notation in (2.5.7) is labyrinthine, it does inspire part *b* and possible proof of the formula therein, when T_g is replaced by S and T_h is replaced by T.

b. The *exchange formula* for certain distributions S and T is

$$(S * T)^{\wedge} = \widehat{S}\widehat{T}. \tag{2.5.8}$$

c. In the best of all worlds, if $S, T \in \mathcal{S}'(\mathbb{R})$ then we could conclude that $S * T$ exists and is an element of $\mathcal{S}'(\mathbb{R})$, that the multiplication $\widehat{S}\widehat{T}$ is well-defined and $\widehat{S}\widehat{T} \in \mathcal{S}'(\mathbb{R})$, and finally that (2.5.8) is valid, cf., [Hor66, page 424], [Sch66, Chapitre VII.8] for the first distributional version of the exchange formula. Also, we already saw that $S * T$ need not exist for $S, T \in \mathcal{S}'(\mathbb{R})$. For the multiplication of distributions we refer to *Definition 2.5.7*, and for a reasonable *theorem* yielding the validity of the exchange formula we refer to *Theorem 2.5.9.*

2.5.7 Definition. MULTIPLICATION OF DISTRIBUTIONS

a. Let $g \in C^{\infty}(\mathbb{R})$, resp., $\mathcal{S}(\mathbb{R})$ or $C^{\infty}(\mathbb{R})$, and let $T \in D'(\mathbb{R})$, resp., $\mathcal{S}'(\mathbb{R})$ or $\mathcal{E}'(\mathbb{R})$. The *product* $gT \in D'(\mathbb{R})$, resp., $\mathcal{S}'(\mathbb{R})$ or $\mathcal{E}'(\mathbb{R})$, is defined by

$$\forall f \in C_c^{\infty}(\mathbb{R}), \text{ resp., } \mathcal{S}(\mathbb{R}) \text{ or } C^{\infty}(\mathbb{R}), \ (gT)(f) = T(gf). \tag{2.5.9}$$

It is easy to see that gT, defined by the right side of (2.5.9) is an element of $D'(\mathbb{R})$, resp., $\mathcal{S}'(\mathbb{R})$ or $\mathcal{E}'(\mathbb{R})$.

Similarly, $gT \in \mathcal{E}'(\mathbb{R})$ for $g \in C_c^{\infty}(\mathbb{R})$ and $T \in D'(\mathbb{R})$.

b. Let $S, T \in D'(\mathbb{R})$, and let $\{f_n\} \subseteq C_c^{\infty}(\mathbb{R})$ be an approximate identity. If the limit

$$\lim_{n \to \infty} [(S * f_n)(T * f_n)] (f)$$

exists for each $f \in C_c^{\infty}(\mathbb{R})$ and is independent of $\{f_n\}$, then S, T are *multiplicable* with *product* $ST \in D'(\mathbb{R})$ defined by

$$\forall f \in C_c^{\infty}(\mathbb{R}), \quad (ST)(f) = \lim_{n \to \infty} [(S * f_n)(T * f_n)] (f) \tag{2.5.10}$$

for any fixed $\{f_n\}$.

If $S = g \in C^\infty(\mathbb{R})$ then (2.5.10) reduces to (2.5.9).

c. Defining the multiplication of distributions is not an annoying technical problem, but rather a large theoretical program motivated by quantum electrodynamics, nonlinear shock waves, and the harmonic analysis associated with Sobolev spaces and the Littlewood-Paley theory, e.g, [Col85], [Mey81], [Obe92]. In quantum field theory, the products of distributions which arose led to Feynman integrals and renormalization theory, cf., [Bre65], [deJ64] for classical background and [Glei92] for a thrilling layman's approach to this material.

The definition of multiplication in (2.5.10) is due to Mikusiński (1960) and is equivalent to one of Hirata and Ogata (1958), cf., [SI64].

2.5.8 Definition. $\mathcal{S}'$-CONVOLUTION

a. Let $S, T \in \mathcal{S}'(\mathbb{R})$. The *$\mathcal{S}'$-convolution* $S * T$ exists if

$$\forall f, g \in \mathcal{S}(\mathbb{R}), \quad (S * f)(\tilde{T} * g) \in L^1(\mathbb{R}),$$

where $\widetilde{T}(f) = T(t)(f(-t))$. In this case, $S * T$ is the unique element of $\mathcal{S}'(\mathbb{R})$ for which

$$\forall f, g \in \mathcal{S}(\mathbb{R}), \quad ((S * T) * f)(g) = \int (S * f)(t)(\tilde{T} * g)(t)\, dt.$$

b. This definition is due to Hirata and Ogata (1958). That the analogous definition for $D'(\mathbb{R})$ reduces to (2.5.2) was proved by results due to L. Schwartz (1954), Shiraishi (1959), and Horváth (1974), cf., [DV78].

The first part of the following theorem was given by Hirata and Ogata (1958), with a lovely proof in [SI64]; and there is a generalization of the result due to Oberguggenberger (1986), e.g.,[Obe92]. The second part of the theorem can be proved directly [Ben75] or as a corollary of the first part.

2.5.9 Theorem. EXCHANGE FORMULA

a. Assume the $\mathcal{S}'$-convolution of $S, T \in \mathcal{S}'$ exists. Then $\widehat{S}, \widehat{T}$ are multiplicable and

$$(S * T)^\wedge = \widehat{S}\widehat{T}.$$

b. Let $g \in L^1(\mathbb{R})$, and $h \in L^\infty(\mathbb{R})$. Then $g * h \in L^\infty(\mathbb{R})$ and

$$(g * h)^\wedge = \widehat{g}\widehat{h}.$$

Further, $\widehat{g} \in A(\widehat{\mathbb{R}})$, $\widehat{h} \in A'(\widehat{\mathbb{R}})$ (the space of pseudo-measures defined in Example 2.4.6f*), and $\widehat{g}, \widehat{h}$ are not only multiplicable in the sense of* Definition 2.5.7 *but in the sense of (2.5.9) where the domain space of functions is $A(\widehat{\mathbb{R}})$.*

2.5.10 Example. CONVOLUTION AND MULTIPLICATION

a. Recall that $L^1_{\text{loc}}(\mathbb{R})$ is not closed under multiplication pointwise a.e. For example, let $g(t) = 1/|t|^{1/2}$, $t \neq 0$, so that $g \in L^1_{\text{loc}}(\mathbb{R})$ and $g^2 \notin L^1_{\text{loc}}(\mathbb{R})$.

b. Does the product δ^2 exist as a distribution? It would be a success of the theory if $\delta^2 \in D'(\mathbb{R})$, since formulas such as

$$\delta^2 - \frac{1}{(\pi t)^2} = -\left(\frac{\pi}{t}\right)^2$$

arise in the surreal quantum world.

First, the product δ^2 doesn't exist in the sense of (2.5.10), as dealing with the approximate identity $\{\frac{n}{2}\mathbf{1}_{[-\frac{1}{n},\frac{1}{n}]}\}$ shows.

Even with this setback, in light of the "naturalness" of the exchange formula, which afterall could be one of our "creative formulas", we are tempted to define δ^2 as $\delta^\vee * \delta^\vee$. However,

$$\delta^\vee * \delta^\vee(\bar{f}) = \delta^\vee(u)\Big(\delta^\vee(v)(\bar{f}(u+v))\Big) = 1\Big(1(\bar{f}(u+v))\Big)$$

$$= 1\left(\int \bar{f}(u+v)\,dv\right) = \iint \bar{f}(u+v)\,dvdu = \int\left(\int \bar{f}(v)\,dv\right)du,$$

which is divergent for test functions f for which $\int f(v)\,dv \neq 0$.

Fortunately, δ^2 *does* fit into Colombeau's theory [Col85]!

c. By the exchange formula it is easy to see that if $T(t) = \frac{1}{\pi}pv(1/t)$ then

$$T * T = \delta. \tag{2.5.11}$$

Note that $\operatorname{supp} T = \mathbb{R}$, and that, even though convolution is intuitively a smoothing operation, $\operatorname{supp} T * T = \{0\}$! It is instructive to graph $1/t$ and its translates to see how the cancellation in (2.5.11) can occur.

d. In *Section 1.5* we saw that $L^1(\mathbb{R})$ is an algebra with convolution as the multiplicative operation; and in *Proposition 2.1.1* we noted that $L^1(\mathbb{R})$ does not have a unit under convolution. On the other hand, we saw that $T*\delta = T$ for all $T \in D'(\mathbb{R})$, and, hence, $\mathcal{E}'(\mathbb{R})$ is a convolution algebra with unit δ.

e. δ, H are multiplicable in the sense of (2.5.10), and

$$\delta H = \frac{1}{2}\delta. \tag{2.5.12}$$

To verify (2.5.12), let $\{f_n\} \subseteq L^1(\mathbb{R})$ be an approximate identity, and note that

$$(\delta * f_n)(H * f_n) = \left(\frac{(H * f_n)^2}{2}\right)'.$$

Clearly, $\lim_{n\to\infty}[(H * f_n)^2/2](f) = (H/2)(f)$ for all $f \in C_c^\infty(\mathbb{R})$, and so the result is obtained.

The exchange formula and formula (2.4.15) for $\widehat{H}$ also yield (2.5.12), when we take note of the fact that

$$\frac{1}{2\pi i}pv(\frac{1}{\gamma})\Big(1(\lambda)(f(\lambda+\gamma))\Big) = \frac{1}{2\pi i}pv(\frac{1}{\gamma})\Big(\int f(\lambda)\,d\lambda\Big) = 0.$$

f. It turns out that the $\mu, \mathbf{1}_{[a,b]}$ are multiplicable for $\mu \in M(\mathbb{R})$, but that the product $T\mathbf{1}_{[a,b]}$ is more elusive and is related to spectral synthesis in the case T is a pseudo-measure, e.g., [Ben75]. A related and equally challenging issue is to define the notion of the point value of a distribution, e.g., [Loj57].

We shall conclude *Section 2.5* with a discussion of the Hilbert transform, which is a special but far reaching convolution.

2.5.11 Definition. HILBERT TRANSFORM

a. The *Hilbert transform* $\mathcal{H}T$ of a distribution T is the convolution

$$\mathcal{H}T(t) = \left(\frac{1}{\pi}pv\left(\frac{1}{u}\right) * T(u)\right)(t),$$

where we have used the notation of point functions to deal with "$pv(\frac{1}{u})$". Thus, *formally*, the Hilbert transform $\mathcal{H}g$ of a function g is

$$\mathcal{H}g(t) = \lim_{\epsilon \to 0} \frac{1}{\pi} \int_{|t-u| \geq \epsilon} \frac{g(u)}{t-u} \, du.$$

We are purposely vague about which distributions (or functions) yield a well-defined Hilbert transform; and, in fact, some of the fundamental theory of Hilbert transforms is associated with the domain X and range Y of the Hilbert transform operator $\mathcal{H} : X \longrightarrow Y$.

b. The Hilbert transform opens the door to a large and profound area of harmonic analysis associated with the theory, relevance, and importance of *singular integral operators*, e.g., [Ste70]. A magnificent exposition of the basic theory of Hilbert transforms is due to Neri [Ner71]. The following result is fundamental.

2.5.12 Theorem. $\mathcal{H} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$

a. $\mathcal{H} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$ *is a well-defined linear bijection with the property that*

$$\forall f \in L^2(\mathbb{R}), \quad \|\mathcal{H}f\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}.$$

b. Let $\sigma(\mathcal{H})(\gamma) \equiv -i \operatorname{sgn} \gamma$. *Then*

$$\mathcal{H} = \mathcal{F}^{-1} \, \sigma(\mathcal{H}) \, \mathcal{F},$$

where $\mathcal{F} : L^2(\mathbb{R}) \longrightarrow L^2(\widehat{\mathbb{R}})$ *is the Fourier transform mapping, e.g.,* Theorem 1.10.2.

c. $\mathcal{H} \circ \mathcal{H} = -I$, *where* I *is the identity mapping on* $L^2(\mathbb{R})$.

Part b and c of *Theorem 2.5.12* are easy to prove. Using (2.4.15) we see that

$$\frac{1}{\pi} pv(\frac{1}{t}) \longleftrightarrow \sigma(\mathcal{H})(\gamma),$$

where $\sigma(\mathcal{H})$ is called the *symbol* of $\mathcal{H}$. By the exchange formula, we have

$$(\mathcal{H}f)^{\wedge} = \sigma(\mathcal{H})\widehat{f}, \tag{2.5.13}$$

and part *b* follows. In the context of (2.5.13), $\sigma(\mathcal{H})$ is called a *multiplier*, e.g., [Ste70], [Lar71]. Part *c* is a consequence of part *b* and the calculation

$$(\mathcal{H}(\mathcal{H}f))^\wedge = \sigma(\mathcal{H})(\mathcal{H}f)^\wedge = \sigma(\mathcal{H})^2 \hat{f} = -\hat{f}.$$

2.5.13 Example. A DISTRIBUTIONAL DOMAIN FOR $\mathcal{H}$

a. It can be shown by direct calculation that $\mathcal{H}f \in L^\infty(\mathbb{R})$ for all $f \in \mathcal{S}(\mathbb{R})$, cf., *Exercise 2.53* for an ingenious calculation due to Logan yielding a stronger result [Log83].

Then, by *Definition 2.5.8*, $\mathcal{H}T \in \mathcal{S}'(\mathbb{R})$ exists for those $T \in \mathcal{S}'(\mathbb{R})$ for which $T * g \in L^1(\mathbb{R})$ whenever $g \in \mathcal{S}(\mathbb{R})$, cf., *Exercise 2.54*.

b. Besides our method in part *a*, the Hilbert transform of distributions can be defined by other methods, which lead to general real variable formulations, e.g., [Cart91], [Jon82], as well as complex variable formulations, e.g., [Bre65] and research by Lauwerier, Martineau, Orton, and Tillmann.

Also note that if $T \in \mathcal{S}'(\mathbb{R})$ and $\hat{T}$ is 0 on $(-a, a)$, then *Theorem 2.5.12b* allows us to define $\mathcal{H}T$ as $\mathcal{F}^{-1}\sigma(H)\mathcal{F}T$.

2.5.14 Remark. PERSPECTIVE ON $\mathcal{H}$

Suppose $f \in L^2(\mathbb{R})$. Then it is easy to check that

$$\forall t \in \mathbb{R}, \quad \mathcal{H}(\tau_t f) = \tau_t(\mathcal{H}f) \tag{2.5.14}$$

and

$$\forall \lambda > 0, \quad \mathcal{H}f_\lambda = (\mathcal{H}f)_\lambda, \tag{2.5.15}$$

i.e., the Hilbert transform $\mathcal{H} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$, which is continuous by *Theorem 2.5.12*, commutes with translations and positive dilations. The algebraic facts (2.5.14) and (2.5.15) should be juxtaposed with the fact that they only continuous operators $L : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$ which commute with translations and both positive and negative dilations are constant multiples of the identity, e.g., [Ste70, pages 55-56].

The fact, that $\mathcal{H}$ commutes with translations and has a multiplier associated with it, is a feature in common with a large class of operators, some of which play an important role in the next section.

2.6 Operational calculus

As Norbert Wiener points out in his paper, "The operational calculus" [Wie81, 26c], "the operational calculus owes its inception to Leibniz, who was struck by the resemblance between the formula for the n-fold differentiation of a product and the nth power of a sum", e.g., *Exercise 2.28.* To fix ideas, we shall think of the *operational calculus* as a symbolic method of solving an equation. For example, if we are given the differential equation $Lf = g$ for a given forcing function g and differential operator L, the goal is to design an operator L^{-1} so that $L^{-1}g \equiv f$ is a solution of the equation. Major contributions in this area are due to Lagrange, Boole, and Pincherle; but the most spectacular, nonrigorous, and successful "formal theory" is Oliver Heaviside's operational calculus [Hea1894], cf., Heaviside's biographies [Nah88], [Sea87], the former professional, the latter personal, and both exquisite.

Before developing his operational calculus, Heaviside had made a profound contribution to the Atlantic submarine cable problem, cf., *Exercise 1.47.* The operators in his operational calculus applied to voltages and currents, gave a meaning to fractional differentiation, led to asymptotic series which could be successfully applied in computations, and infuriated the "wooden-headed" (Heaviside's appellation) mathematicians of his day. Lest we mathematicians become a splinter (sic) group, Laurent Schwartz came along with his theory of distributions to legitimize Heaviside's ideas – which, of course, may have infuriated Heaviside! A concept in Wiener's paper on the operational calculus [Wie81, 26c, Section 8] played a role in Schwartz's motivation to define distributions; although, at the time of his research, Schwartz was unaware of Wiener's contribution and attributed it to related work done by others after Wiener, e.g., [Wie81, Volume II, pages 426-427], [Sch66, page 4].

Other 20th century work on the operational calculus includes contributions by Volterra (1920), Carson (1925), van der Pol and Bremmer (1955), Doetsch (1958), and Mikusiński (1959), as well as books such as [Erd62], [Sch61, pages 123-140 and pages 230-235], [Yos84], and [Zem87].

In the following result the distributional proof is correct up to the last step. The subtleties are discussed in *Remark 2.6.2.*

2.6.1 Theorem. DIFFERENTIAL EQUATION AND OPERATIONAL SOLUTION

Let $\{a_n : n = 0, \ldots, N\} \subseteq \mathbb{C}$, $a_N \neq 0$, *and let* $S \in \mathcal{S}'(\mathbb{R})$. *A solution* $X \in \mathcal{S}'(\mathbb{R})$ *of the differential equation*

$$LX = \sum_{n=0}^{N} a_n X^{(n)} = S \tag{2.6.1}$$

is

$$X = \left(\frac{\widehat{S}}{P}\right)^{\vee}, \tag{2.6.2}$$

where

$$P(\gamma) = \sum_{n=0}^{N} (2\pi i)^n a_n \gamma^n.$$

Proof. By *Proposition 2.5.5*, equation (2.6.1) can be written as

$$X * \left(\sum_{n=0}^{N} a_n \delta^{(n)}\right) = S.$$

Thus, by the exchange formula, $\widehat{X}P = \widehat{S}$. Dividing by P and taking the inverse transform yield (2.6.2). □

2.6.2 Remark. DIVISION OF DISTRIBUTIONS

a. If the polynomial P of *Theorem 2.6.1* does not have real zeros then the reasoning in the last step of the proof (of *Theorem 2.6.1*) is correct, and little further detail is needed. If P has real zeros then a more elaborate argument is required which depends on the relatively elementary fact that *if* $T \in D'(\mathbb{R})$ *and* P *is a polynomial on* $\mathbb{R}$ *then there is a distribution* $X \in D'(\mathbb{R})$ *for which* $PX = T$, e.g., [Sch66, Chapitre V.4, pages 123-126].

b. The *division problem* for $\mathbb{R}^d$, $d \geq 2$, i.e., the solution of the distributional equation $PX = T$ on $\mathbb{R}^d$ for given distributions T and polynomials P, is difficult. It was solved by Hörmander and Łojasiewicz independently in the late 1950s. There is a wonderful exposition of the topic including its relationship with partial differential equations in [Sch63].

c. Let L be a partial differential operator on $\mathbb{R}^d$. A distribution E for which $LE = \delta$ is a *fundamental solution* of L. A great success of distribution theory is the theorem that *if L has constant coefficients then it has a fundamental solution.* This result was proved by Ehrenpreis and Malgrange independently in 1953, e.g., [Hör83]. There is now a elementary proof of this theorem in [Ros91].

Suppose L has constant coefficients and $S \in \mathcal{E}'(\mathbb{R}^d)$ is a "forcing function". Then $X = E * S$ is a solution of the equation $LX = S$ since, formally,

$$L(E * S) = L\delta * (E * S) = \delta * S = S, \tag{2.6.3}$$

cf., the proof of *Theorem 2.6.3b*, and every other solution is of the form $E * S + T$, where T is a solution of the corresponding homogeneous equation, i.e., $LT = 0$.

We shall not get involved in aspects of Heaviside's calculus, such as his *expansion theorem*, which caused "gorges to rise" in yesteryear. However, we shall prove the following result, which quantifies the point of view of the first paragraph of this section. In light of the proof of *Theorem 2.6.1*, let us first clarify the notation for a differential operator L. If LX has the designation

$$LX = \sum_{n=0}^{N} a_n X^{(n)},$$

then

$$L\delta = \sum_{n=0}^{N} a_n \delta^{(n)} \in D'(\mathbb{R}),$$

and so

$$LX = (L\delta) * X.$$

2.6.3 Theorem. DIFFERENTIAL EQUATION AND FUNDAMENTAL SOLUTION

Let $\{a_n : n = 0, \ldots, N\} \subseteq \mathbb{C}$, $a_N = 1$, and let $h \in C^\infty(\mathbb{R})$ be the solution of the initial value problem,

$$(2.6.4) \qquad \begin{aligned} (L\delta) * h = Lh = 0 \quad \text{on} \quad \mathbb{R}, \\ h(0) = \cdots = h^{(N-2)}(0) = 0, \quad h^{(N-1)}(0) = 1. \end{aligned}$$

(The method to find the solution of (2.6.4) is elementary and well-known in ordinary differential equations, e.g., [Swe96].)

a. $(L\delta) * (Hh) = \delta$, where H is the Heaviside function. Thus, $E \equiv Hh$ is a fundamental solution of L.

b. Let $S \in \mathcal{E}'(\mathbb{R})$. Then $X = E * S$ is a solution of the differential equation $LX = S$.

Proof. a. We shall evaluate $(Hh)^{(n)}, n = 1, \ldots, N$. Note that $(Hh)^{(1)} = Hh^{(1)} + \delta h$, e.g., *Exercise 2.28*, and $\delta h = h(0)$. The second fact results from the calculation $(\delta h)(f) = \delta(hf) = h(0)f(0) = (h(0)\delta)(f)$. By hypothesis, then,

$$(Hh)^{(1)} = Hh^{(1)}.$$

Continuing to use the product rule and (2.6.4), in conjunction with the calculation from the previous step, we obtain

$$\begin{array}{lllll} (Hh)^{(2)} & = & Hh^{(2)} + \delta h^{(1)} & = & Hh^{(2)}, \\ \cdots & & & & \\ (Hh)^{(N-1)} & = & Hh^{(N-1)} + \delta h^{(N-2)} & = & Hh^{(N-1)}, \\ (Hh)^{(N)} & = & Hh^{(N)} + \delta h^{(N-1)} & = & Hh^{(N)} + \delta. \end{array}$$

Thus,

$$(L\delta) * (Hh) = \sum_{n=0}^{N} a_n (Hh)^n = H(Lh) + \delta = \delta.$$

b. We calculate

$$L(E * S) = (L\delta) * (E * S) = (L\delta * E) * S = \delta * S = S.$$

Note that we have assumed associativity of convolution in the second equality, cf., *Example 2.5.2c* where we showed such associativity is not generally valid. We are justified in this case since the distributions $L\delta, E = Hh$, and S are each supported in an interval of the $[c, \infty)$, e.g., [Sch66, page 172]. □

2.6.4 Example. $L\delta = \sum_{n=0}^{N} a_n \delta^{(n)}$

It is easy to see that $\operatorname{supp} L\delta = \{0\}$. Conversely, *if* $T \in D'(\mathbb{R})$ *and* $\operatorname{supp} T = \{0\}$ *then there is* $\{a_0, \ldots, a_N\} \subseteq \mathbb{C}$ *such that* $T = L\delta$.

This important fact is a corollary of a theorem, e.g., [Hör83, Volume I, pages 46-47], which depends on a "philosophy" (a big word, but "point of view" is not quite accurate in this case) underlying significant parts of spectral synthesis [Ben75], [KS63] and potential theory [Hed80]. To describe this "philosophy", first note that *if* $g \in L^1(\mathbb{R})$ *and* $f \in C_0(\mathbb{R})$ *vanishes on* $\operatorname{supp} g$ *then* $\int g(t)f(t)\,dt = 0$. Similarly, *if* $\mu \in M_b(\mathbb{R})$ *and* $f \in C_0(\mathbb{R})$ *vanishes on* $\operatorname{supp} \mu$ *then* $\mu(f) = \int f(t)d\mu(t) = 0$. However, *if* $T \in A'(\mathbb{R})$ *and* $f \in A(\mathbb{R})$ *vanishes on* $\operatorname{supp} T$ *then it is not necesssarily true that* $T(f) = 0$. This last fact is a deoderized version of Malliavin's profound example of non-synthesis. Suppose that $f \in X$ and $T \in X'$, where $X \supseteq C_c^\infty(\mathbb{R})$ and $X' \subseteq D'(\mathbb{R})$ satisfy the natural condition of *Example 2.4.6h*. The "philosophy", alluded to above, is that if $f = 0$ on $\operatorname{supp} T$, and if the set $\operatorname{supp} T$ is regular enough vis a vis the smoothness of f near the boundary of $\operatorname{supp} T$, then $T(f) = 0$.

2.6.5 Definition. LINEAR TRANSLATION INVARIANT SYSTEMS

a. Let X be a linear subspace of $D'(\mathbb{R})$ with the properties that $\delta \in X$, and that if $T \in X$, then

$$\forall t \in \mathbb{R}, \quad \tau_t T \in X.$$

A *linear translation invariant (LTI) system* is a linear operator $L : X \to X$ for which

$$\forall T \in X \text{ and } \forall t \in \mathbb{R}, \quad L(\tau_t T) = \tau_t L(T). \tag{2.6.5}$$

Property (2.6.5) is *translation invariance*, and for many applications this reflects *time invariance*. The property,

$$\forall S, T \in X \quad \text{and} \quad \forall a, b \in \mathbb{C}, \quad L(aS + bT) = aL(S) + bL(T),$$

is the *linearity* of L.

The *impulse response* of the system L is

$$L\delta = h \in X.$$

The *filter* corresponding to L is $\widehat{h}$; $\widehat{h}$ is also referred to as the *frequency response* or *transfer function* corresponding to L.

b. Norbert Wiener argues persuasively for studying translation invariant operators on physical grounds [Wie33, Introduction] , i.e., although time shifts may distort some astronomical observations, laboratory experiments should generally be time invariant. Wavelets and nonstationary methods can be used in dealing with *time varying* events.

The notion of an LTI system L is an important and basic engineering concept, e.g., [OS75], [OW83], [Pap77], which also has a long history in mathematics, e.g., [Ben75, pages 216-217]. Suppose X is closed under convolution, e.g., $X = M_b(\mathbb{R})$. Then, if L satisfies certain natural conditions, e.g., part c below, the impulse response h and filter $\widehat{h}$ play a central role in quantifying L, viz.,

$$\forall f \in X, \quad Lf = h * f. \tag{2.6.6}$$

c. We have already defined the notion of a causal signal. We now say that an LTI system L is *causal* if, whenever $T \in \mathbb{R}$ and $f \in X$ vanishes on $(-\infty, T)$ then Lf vanishes on $(-\infty, T)$. This means that there can not be an output signal from the system L before there is an input signal – a reasonable point a view.

$\longrightarrow Y$

Definition 2.6.5 can obviously be extended to operators $L : X \longrightarrow Y$. The following result is due to [AN79], and a precursor depending on the continuity of L is due to [Sch66, pages 197-198]. *Let* $L : C_c^\infty(\mathbb{R}) \to D'(\mathbb{R})$ *be a causal* LTI *system. Then there is a unique distribution* $h \in D'(\mathbb{R})$ *such that* $Lf = h * f$ *for all* $f \in C_c^\infty(\mathbb{R})$. A feature of this theorem is that continuity of L is not required à priori to obtain (2.6.6).

2.6.6 Example. TRANSLATION, CONVOLUTION, AND $Lf = h * f$

A fundamental theorem in the *theory of multipliers* is that *the translation invariant continuous linear operators* $L : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ *are precisely of the form,*

$$\forall f \in L^1(\mathbb{R}), \quad Lf = \mu * f, \tag{2.6.7}$$

where $\mu \in M_b(\mathbb{R})$, e.g., [Lar71]. It is a routine calculation to prove that if $\mu \in M_b(\mathbb{R})$ and L is defined by (2.6.7) then $L : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ is a translation invariant continuous linear operator, e.g., *Exercise 2.58a.* The converse is more difficult, and the initial steps of its proof are the content of *Exercise 2.58b,c.*

2.6.7 Example. BANDPASS AND LOWPASS FILTERS - THE HILBERT TRANSFORM

In *Example 1.2.2b* we discussed modulated signals and their carrier frequencies. This example continues that discussion in terms of lowpass and bandpass filters, and the fact that they are related by the Hilbert transform. To fix ideas, we say that a filter $\widehat{h}$ is *bandpass* if $\operatorname{supp}\widehat{h} = [-\beta, -\alpha] \bigcup [\alpha, \beta]$, where $\beta > \alpha > 0$, and it is *lowpass* if $\operatorname{supp}\widehat{h} = [-\alpha, \alpha]$.

a. Let L be an LTI system with a real impulse response $h \in PW_\Omega$; in particular, $\widehat{h}$ is a lowpass filter. Then the modulated signal $h_b(t) \equiv h(t)\cos 2\pi t\gamma_0$, with carrier frequency $\gamma_0 > \Omega$, is the impulse response for the bandpass filter $\widehat{h}_b = A_b e^{i\varphi_b}$ with *bands* $[\gamma_0 - \Omega, \gamma_0 + \Omega], [-\gamma_0 - \Omega, -\gamma_0 + \Omega]$; and A_b and φ_b satisfy the properties

$$\forall \gamma \in [0, \Omega], A_b(\gamma + \gamma_0) = A_b(-\gamma + \gamma_0) \text{ and}$$

$$\varphi_b(\gamma + \gamma_0) = -\varphi_b(-\gamma + \gamma_0), \tag{2.6.8}$$

i.e., in the band $[\gamma_0 - \Omega, \gamma_0 + \Omega]$, A_b satisfies mirror symmetry and φ_b satisfies "point" symmetry about γ_0, e.g., *Exercise 2.59c.*

b. Conversely, let L_b be an LTI system with a real impulse response h_b and bandpass filter $\widehat{h}_b = A_b e^{i\varphi_b}$. Assume the bands of $\widehat{h}_b$ are $[\gamma_0 - \Omega, \gamma_0 + \Omega], [-\gamma_0 - \Omega, -\gamma_0 + \Omega]$, where $\gamma_0 > \Omega$, and that the symmetry condition (2.6.8) is satisfied. We shall verify that h_b is a modulated signal $h_b(t) = h(t)\cos 2\pi t\gamma_0$, with carrier frequency γ_0, where h is the impulse response for a lowpass filter $\widehat{h}$ for which $\operatorname{supp}\widehat{h} = [-\Omega, \Omega]$.

Let $\widehat{h} = Ae^{i\varphi}$ be defined as

$$\widehat{h}(\gamma) = 2(\widehat{h}_b\, H)(\gamma + \gamma_0), \tag{2.6.9}$$

where H is the Heaviside function. By definition, it is easy to see that $\operatorname{supp} \widehat{h} = [-\Omega, \Omega]$, e.g., *Figure 2.1. Exercise 2.57a*, where it is required to verify

$$f + i\mathcal{H}f \longleftrightarrow 2\widehat{f}H,$$

and (2.6.9) allow us to write

$$h(t) = (h_b + i\mathcal{H}h_b)(t)e^{-2\pi i t\gamma_0}. \tag{2.6.10}$$

It turns out that (2.6.8) allows us to conclude that h is real valued. In fact, (2.6.8) implies $\widehat{h}_b(\gamma + \gamma_0) = \overline{\widehat{h}_b(-\gamma + \gamma_0)}$ for $\gamma \in [-\Omega, \Omega]$; and hence, by (2.6.8), $\widehat{h}(\gamma) = \overline{\widehat{h}(-\gamma)}$, on $\widehat{\mathbb{R}}$, i.e., h is real valued.

Thus, by (2.6.10), we have $h_b(t) = h(t) \cos 2\pi t\gamma_0$ and our calculation is complete.

Figure 2.1

c. Parts a and b combine for a result of both theoretical and practical value. Theoretically, we establish isomorphisms between PW_Ω and spaces of $L^2(\mathbb{R})$ functions whose Fourier transforms are supported by more unusual sets than $[-\Omega, \Omega]$. Practically, ideas based on the above calculations are used in bandpass sampling, in the design of equivalent systems where one may be more efficient than another in some desired way, or in narrow band communications theory, e.g., [OS75, Chapters 7 and 10].

2.7 Measure theory

We shall address two profound ideas from modern analysis: the Riesz Representation Theorem (RRT) and the Herglotz-Bochner Theorem. In the process we hope to tie-together our discussions of Radon measures in *Definition 2.3.6* and of real analysis in *Appendix A*. Integration and measure, truly a "keystone combination" for analysis, began with ancient efforts to *define* areas of nonrectilinear regions, e.g., [Ben77] on Archimedes and integration, reached a new level of precision and generality in the mid 19th century to cope with emerging requirements of trigonometric series, e.g., [Rie1873] and *Section 3.2*, generalized around 1900 to measure bizarre sets and to grapple with subjects such as statistical mechanics, e.g., [Wie81, Volume II, page 801] on sets of measure 0, and assumed a sophisticated 20th century functional analytic identity by means of the RRT, e.g., [Ries49] for a readable history for the first half of this century by the master, cf., [Gra84].

2.7.1 Definition. DUAL SPACES ASSOCIATED WITH RRT

a. A linear functional $T : C_c(\mathbb{R}) \longrightarrow \mathbb{C}$ is an element of the *dual space* $C_c(\mathbb{R})'$ (of the vector space $C_c(\mathbb{R})$) if $\lim_{n\to\infty} T(f_n) = 0$ for every sequence $\{f_n\} \subseteq C_c(\mathbb{R})$ satisfying the properties:

i. $\exists K \subseteq \mathbb{R}$, compact, such that $\forall n$, $\operatorname{supp} f_n \subseteq K$,

ii. $\lim_{n\to\infty} \|f_n\|_{L^\infty(\mathbb{R})} = 0$.

These two properties define a topology on $C_c(\mathbb{R})$, and $C_c(\mathbb{R})'$ is the space of *continuous linear functionals* on $C_c(\mathbb{R})$.

b. A linear functional $T : C_0(\mathbb{R}) \longrightarrow \mathbb{C}$ is an element of the *dual space* $C_0(\mathbb{R})'$ (of the vector space $C_0(\mathbb{R})$) if $\lim T(f_n) = 0$ for every sequence $\{f_n\} \subseteq C_0(\mathbb{R})$ for which $\lim_{n\to\infty} \|f_n\|_{L^\infty(\mathbb{R})} = 0$. The sup norm, $\|\ldots\|_{L^\infty(\mathbb{R})}$, defines a topology on $C_0(\mathbb{R})$, by which it becomes a

Banach space; and $C_0(\mathbb{R})'$ is the space of *continuous linear functionals* on $C_0(\mathbb{R})$, cf., *Example B.17*.

2.7.2 Remark. PRELIMINARIES FOR RRT

a. If $F \in BV_{\text{loc}}(\mathbb{R})$ there are two Radon measures, T_F and S_F, defined in terms of F, which are fundamental for RRT. Recalling that $BV_{\text{loc}}(\mathbb{R}) \subseteq L^1_{\text{loc}}(\mathbb{R})$ (*Exercise 2.23*), T_F is defined by (2.2.4) and (2.2.5) as

$$\forall f \in C_c^\infty(\mathbb{R}), \quad T_F(f) = \int F(t)f(t)\,dt. \tag{2.7.1}$$

S_F is the *Riemann-Stieltjes Radon measure* defined in terms of the Riemann-Stieltjes integral [Apo57] as

$$\forall f \in C_c^\infty(\mathbb{R}), \quad S_F(f) = \int f(t)dF(t). \tag{2.7.2}$$

It is easy to check that T_F and S_F are, in fact, Radon measures.

b. (2.7.2) defines a mapping,

$$\begin{aligned} L_S : BV_{\text{loc}}(\mathbb{R}) &\longrightarrow M(\mathbb{R}), \\ F &\longmapsto S_F, \end{aligned} \tag{2.7.3}$$

which is *not* injective. In fact, if $F \in BV_{\text{loc}}(\mathbb{R})$ and $C \in \mathbb{C}$ then $F + C \in BV_{\text{loc}}(\mathbb{R})$ and $S_F = S_{F+C}$. This should be compared to the mapping L of (2.2.5), restricted to $BV_{\text{loc}}(\mathbb{R})$, which is injective.

c. Let $F \in BV_{\text{loc}}(\mathbb{R})$. Clearly, the integration by parts formula for Riemann-Stieltjes integrals [Apo57],

$$\forall f \in C_c^\infty(\mathbb{R}), \quad \int f(t)dF(t) = -\int f'(t)F(t)\,dt,$$

can be rewritten as $S_F = T'_F(= F')$.

The space $M(\mathbb{R})$, resp., $M_b(\mathbb{R})$, of Radon measures, resp., bounded Radon measures, was defined in *Definition 2.3.6* in terms of distributional derivatives.

2.7.3 Theorem. RIESZ REPRESENTATION THEOREM

a. $C_c(\mathbb{R})' = M(\mathbb{R})$, *the space of Radon measures. In fact, for every* $\mu \in M(\mathbb{R})$ *there is* $F \in BV_{\text{loc}}(\mathbb{R})$ *such that* $\mu = S_F = F'$.

b. $C_0(\mathbb{R})' = M_b(\mathbb{R})$, *the space of bounded Radon measures. In fact, for every* $\mu \in M_b(\mathbb{R})$ *there is* $F \in BV(\mathbb{R})$ *such that* $\mu = S_F = F'$.

2.7.4 Remark. RRT: THEOREM TO DEFINITION

a. RRT asserts that the mapping (2.7.3) is a surjection, and that its restriction to $BV(\mathbb{R})$ is a surjection onto $M_b(\mathbb{R})$. An important part of the proof of RRT utilizes the theory of integration from *Appendix A* to extend $\mu \in C_0(\mathbb{R})'$ to a functional on the vector space generated by linear combinations of characteristic functions of intervals.

Frigyes (Frederick) Riesz proved RRT in 1909. Although Riesz did not deal with distributions, it is most efficient to write RRT as we did in *Theorem 2.7.3*, e.g., [Sch66, pages 53-54], cf., [Ries14], [RN55], [Ben76, pages 255-257] for classical and readable proofs.

RRT has evolved from a *theorem*, associating certain linear functionals with elements of $BV_{\text{loc}}(\mathbb{R})$ by means of the Riemann-Stieltjes integral, to the *definition of integral* in terms of such functionals, e.g., [Bou65]. Notation reflects this metamorphisis, e.g., (2.3.10). If $\mu \in M(\mathbb{R})$ then μ is a distribution $\mu = F'$ for some $F \in BV_{\text{loc}}(\mathbb{R})$, $\mu \in C_c(\mathbb{R})'$, where $\mu(f)$ is denoted by $\int f(t)d\mu(t)$ but is really the Riemann-Stieltjes integral $\int f(t)dF(t)$, *and* there is a space $L^1_\mu(\mathbb{R})$ of functions *integrable with respect to* μ. This last concept is developed in integration theory by extending the functional $\mu : C_c(\mathbb{R}) \to \mathbb{C}$ to a large space, viz., $L^1_\mu(\mathbb{R})$, e.g., [Ben76], [Bou65], [Mal82], [Rud66], cf., part *b*. In this setting $L^1(\mathbb{R})$ is the space of functions integrable with respect to Lebesgue measure.

b. In integration theory we define the *Borel algebra* $\mathcal{B}(\mathbb{R})$ to be the smallest σ-algebra of subsets of $\mathbb{R}$ that contains all the open sets. A function $\mu : \mathcal{B}(\mathbb{R}) \to \mathbb{R}^+ \cup \{\infty\}$ is a *locally finite Borel measure* if the following conditions are satisfied:

i. $\mu(\cup B_n) = \Sigma\mu(B_n)$ for every disjoint sequence $\{B_n\} \subseteq \mathcal{B}(\mathbb{R})$,

ii. $\mu(K) < \infty$ for every compact set $K \subseteq \mathbb{R}$.

Let $\mathcal{M}_+(\mathbb{R})$ be the space of locally finite Borel measures.

An integral (sic) part of the technique and philosophy associated with RRT is the following result. Define the mapping

$$\begin{array}{rcl} I : \mathcal{M}_+(\mathbb{R}) & \longrightarrow & M_+(\mathbb{R}) \\ \mu & \longmapsto & I_\mu, \end{array}$$

where, for all $f \in C_c(\mathbb{R})$, $I_\mu(f)$ is the integral of f with respect to μ, as defined in integration theory. F. Riesz and Radon proved the existence and uniqueness theorem that I *is a bijection*, e.g., [Mal82, pages 61-76].

2.7.5 Definition. CONTINUOUS AND DISCRETE MEASURES

a. A Radon measure $\mu \in M(\mathbb{R})$ is *discrete* if μ is of the form $\mu = \sum_{x\in S} a_x\delta_x$, where each $a_x \in \mathbb{C}\backslash\{0\}$ and $\sum_{x\in S} |a_x| < \infty$. Thus, each such index set S is countable, cf., *Exercise 2.33.* $M_d(\mathbb{R})$ denotes the space of discrete measures. Clearly, $M_d(\mathbb{R}) \subseteq M_b(\mathbb{R})$.

b. A Radon measure $\mu \in M(\mathbb{R})$ is *continuous* if $F \in BV_{\text{loc}}(\mathbb{R})$ can be chosen as a continuous function on $\mathbb{R}$ in the representation $F' = \mu$.

The Cantor measure μ_C defined in *Example 2.3.9b* is a continuous measure, as are the elements of $L^1(\mathbb{R})$.

c. The *problem of spectral estimation* can be viewed in terms of determining the discrete part of Radon measures, e.g., *Definition 2.8.6.*

The following result is one of the basic decompositions of real analysis, e.g., [Ben76], [Bou65], [Rud66].

2.7.6 Theorem. DECOMPOSITION OF MEASURES

Let $F \in BV(\mathbb{R})$, $\mu \in M_b(\mathbb{R})$, and assume $F' = \mu$.

a. $F = F_{ac} + F_{sc} + \Sigma a_x\tau_x H$, where H is the Heaviside function, $F_{ac} \in BV(\mathbb{R}) \cap AC_{\text{loc}}(\mathbb{R})$, $\Sigma|a_x| < \infty$, F_{sc} is continuous on $\mathbb{R}$, and $DF_{sc} = 0$ a.e., where D is ordinary pointwise differentiation.

b. $\mu = g + \mu_{sc} + \Sigma a_x\delta_x$, where $g \in L^1(\mathbb{R})$, $\Sigma|a_x| < \infty$, and $\mu_{sc} \in M_b(\mathbb{R})$ is designated the continuous singular *part of μ.*

c. $g = F'_{ac}$, $\mu_{sc} = F'_{sc}$, and $\Sigma a_x\delta_x = (\Sigma a_x\tau_x H)'$.

d. For all $f \in C_c(\mathbb{R})$,

$$\mu(f) = \int g(t)f(t)\,dt + \int f(t)dF_{sc}(t) + \Sigma a_x f(x),$$

where the integrals on the right side are Lebesgue and Riemann-Stieltjes integrals, respectively.

e. $\|\mu\|_1 = \|g\|_{L^1(\mathbb{R})} + \|\mu_{sc}\|_1 + \Sigma|a_x|$.

2.7.7 Theorem. POSITIVE DISTRIBUTIONS

If $T \in D'(\mathbb{R})$ is positive then $T \in M_+(\mathbb{R})$.

Proof. Let $\{f_n\} \subseteq C_c^\infty(\mathbb{R})$ satisfy the properties that there is a compact set $K \subseteq \mathbb{R}$ such that supp $f_n \subseteq K$ for each n, and that $\lim_{n\to\infty} \|f_n\|_{L^\infty(\mathbb{R})} = 0$. We shall prove that $\lim_{n\to\infty} T(f_n) = 0$ from which we can conclude that $T \in M(\mathbb{R})$ by RRT.

Let nonnegative $f \in C_c^\infty(\mathbb{R})$ equal 1 on K. By the uniform convergence of $\{f_n\}$ there is $\{\epsilon_n\} \subseteq (0,\infty)$ decreasing to 0 such that

$$\forall t \text{ and } \forall n, \quad |f_n(t)| \le \epsilon_n f(t).$$

A straightforward calculation yields the fact that the real and imaginary parts of each f_n are in $C_c^\infty(\mathbb{R})$, and so we assume that each f_n is real-valued. Thus,

$$\forall t \text{ and } \forall n, \quad -\epsilon_n f \le f_n \le \epsilon_n f,$$

so that, by the positivity assumption, $\lim_{n\to\infty} T(f_n) = 0$, and, hence, $T \in M(\mathbb{R})$. □

2.7.8 Definition. POSITIVE DEFINITE FUNCTIONS

a. A function $P : \widehat{\mathbb{R}} \to \mathbb{C}$ is *positive definite* if

$$\forall n, \forall c_1, \ldots, c_n \in \mathbb{C}, \text{ and } \forall \gamma_1, \ldots, \gamma_n \in \widehat{\mathbb{R}},$$

(2.7.4)
$$\sum_{j,k} c_j \bar{c}_k P(\gamma_j - \gamma_k) \ge 0.$$

In this case we write $P >> 0$.

b. If $P >> 0$ then

i. $\forall \gamma \in \widehat{\mathbb{R}},\ P(\gamma) = \overline{P(-\gamma)}$,

ii. $\forall \gamma \in \widehat{\mathbb{R}},\ |P(\gamma)| \le P(0)$, and hence $P(0) \ge 0$ and $P \in L^\infty(\widehat{\mathbb{R}})$,

iii. $\forall \lambda, \gamma \in \widehat{\mathbb{R}}$,

$$|P(\gamma) - P(\lambda)|^2 \leq 2P(0)\,\mathrm{Re}(P(0) - P(\gamma - \lambda)),$$

iv. P continuous at 0 implies P is continuous on $\widehat{\mathbb{R}}$ (by part *iii*), e.g., *Exercise 2.52* for verification of these facts and the relationship to positive definite matrices.

c. Positive definite functions arise in *moment problems*, e.g., [RN55], [Wid41]; and, since (2.7.4) is a pure quadratic form, such functions also arise in minimization problems (in economics, for example) and minimum principles (in physics, for example), e.g., [CH53], [Str88].

2.7.9 Example. Positive Definite Functions

a. If $F \in L^2(\mathbb{R})$ then the *L^2-autocorrelation* of F, viz., $P \equiv F * \widetilde{F}$, is positive definite. In fact,

$$\sum_{j,k} c_j\bar{c}_k P(\gamma_j - \gamma_k) = \int \sum_{j,k} c_j\bar{c}_k F(\gamma_j - \lambda)\overline{F}(\gamma_k - \lambda)d\lambda$$
$$= \int |\Sigma c_j F(\gamma_j - \lambda)|^2 d\lambda \geq 0.$$

Clearly, $P \in A(\widehat{\mathbb{R}})$ since $(F * \widetilde{F})^\vee = |F^\vee|^2 \in L^1(\mathbb{R})$.

b. If $\mu \in M_{b+}(\mathbb{R})$ then $P \equiv \widehat{\mu} >> 0$ and P is continuous.

For the continuity, if $\mu \in M_b(\mathbb{R})$, then it is easy to check that $\mu \in \mathcal{S}'(\mathbb{R})$ and

$$\forall \gamma \in \widehat{\mathbb{R}}, \quad \widehat{\mu}(\gamma) = \int e^{-2\pi i t\gamma} d\mu(t),$$

e.g., *Remark B.16*. This formulation in terms of an integral allows us to conclude that $\widehat{\mu}$ is continuous on $\widehat{\mathbb{R}}$.

To prove that $P >> 0$, it is only necessary to make the computation,

$$\sum_{j,k} c_j c_k \widehat{\mu}(\gamma_j - \gamma_k) = \int |\sum c_j e^{-2\pi i t\gamma_j}|^2 d\mu(t) \geq 0.$$

c. The set of positive definite functions is not a vector space. However, if $P, Q >> 0$ then $P + Q >> 0$ and $PQ >> 0$. The latter fact follows from a theorem of I. Schur, and is an easy calculation such as those of parts *a, b* in the case $Q \in M_{b+}(\mathbb{R})^\wedge$.

In *Example 2.7.9b* we showed that if $\mu \in M_{b+}(\mathbb{R})$ then $\widehat{\mu}$ is a continuous positive definite function on $\widehat{\mathbb{R}}$. The converse of this fact is the *Herglotz-Bochner Theorem* (*Theorem 2.7.10*). Herglotz (1911) proved it for positive definite functions on $\mathbb{Z}$. The proofs for positive definite functions on $\widehat{\mathbb{R}}$ and locally compact abelian groups are due to Bochner (1933) and Weil (1938), respectively. There are several conceptually different proofs, e.g., [Don69], [Kat76], [Rud62], cf., [Sch66].

2.7.10 Theorem. HERGLOTZ-BOCHNER THEOREM

Let $P : \widehat{\mathbb{R}} \to \mathbb{C}$ be a continuous positive definite function on $\widehat{\mathbb{R}}$. Then there is a unique positive bounded Radon measure $\mu \in M_b(\mathbb{R})$ for which $\widehat{\mu} = P$ on $\widehat{\mathbb{R}}$.

Proof. a. Let $F \in \mathcal{S}(\widehat{\mathbb{R}})$. Since $P \in C_b(\widehat{\mathbb{R}})$, the integral

$$\iint P(\gamma - \lambda)F(\gamma)\overline{F(\lambda)}\, d\gamma d\lambda \tag{2.7.5}$$

is the limit of the Riemann sums

$$\sum_j \sum_k P(\gamma_j - \lambda_k)F(\gamma_j)\overline{F(\lambda_k)}\Delta\gamma_j\Delta\lambda_k. \tag{2.7.6}$$

$P >> 0$ and (2.7.6) allow us to conclude that the integral of (2.7.5) is nonnegative. Rewriting (2.7.5), we have

$$0 \leq \iint P(\lambda)F(\gamma)\overline{F(\gamma - \lambda)}\, d\gamma d\lambda = \int P(\lambda)F * \tilde{F}(\lambda)\, d\lambda.$$

Since $C_b(\widehat{\mathbb{R}}) \subseteq \mathcal{S}'(\widehat{\mathbb{R}})$ we can compute

$$P(F * \tilde{F}) = P^\vee(|\check{F}|^2);$$

and so $P^\vee \in \mathcal{S}'(\mathbb{R})$ is nonnegative on all $k \in \mathcal{S}(\mathbb{R})$ of the form $k = |h|^2$, $h \in \mathcal{S}(\mathbb{R})$.

Now let $f \in C_c^\infty(\mathbb{R})$ be nonnegative. For each $\epsilon > 0$, $f + \epsilon^2 g \in \mathcal{S}(\mathbb{R})$ is positive on $\mathbb{R}$, where g is the normalized Gaussian; and $h \equiv (f + \epsilon^2 g)^{1/2} \in \mathcal{S}(\mathbb{R})$. Thus, $P^\vee(f + \epsilon^2 g) = P^\vee(h^2) = P^\vee(|h|^2) \geq 0$. By linearity and the fact that $\epsilon > 0$ is arbitrary, we see that $P^\vee(f) \geq 0$; and so $\mu \equiv P^\vee \in M(\mathbb{R})$ by *Theorem 2.7.7.*

Finally, we must show that $\mu \in M_b(\mathbb{R})$. First note that

$$\forall \alpha > 0, \quad P * g_\alpha(0) \leq P(0),$$

where g_α is the dilation of g. This follows since $|P(\gamma)| \leq P(0)$ and

$$\forall \gamma \in \widehat{\mathbb{R}},$$

$$|P * g_\alpha(\gamma)| \leq \int |P(\gamma - \lambda)| g_\alpha(\lambda)\, d\lambda \leq P(0) \int g_\alpha(\lambda)\, d\lambda = P(0).$$

Next note that the approximate identity $\{g_\alpha\}$ satisfies the following properties: $g_\alpha \geq 0$, $(g_\alpha)^\vee \geq 0$, and g_α is even. Consequently, we have the computation

$$P(0) \geq |P * g_\alpha(0)| = |\int P(\gamma) g_\alpha(\gamma) d\gamma|$$

$$= |\alpha P^\vee((g_{1/\alpha})^\vee)| = \int g^\vee(\alpha t) d\mu(t) \geq 0.$$

We can apply the Beppo Levi Theorem (*Theorem A.8*) since $0 \leq g^\vee(\alpha t) \leq g^\vee(\beta t)$ for $\alpha \leq \beta$ and since $\lim_{\alpha \to \infty} g^\vee(\alpha t) = 1$ by *Proposition 1.6.11*. Thus

$$0 \leq \int d\mu(t) \leq P(0)$$

for the positive measure μ, and this allows us to conclude that $\mu \in M_b(\mathbb{R})$. □

2.7.11 Definition. FOURIER-STIELTJES TRANSFORMS

a. The space of Fourier transforms of bounded Radon measures is denoted by $B(\widehat{\mathbb{R}})$, i.e.,

$$B(\widehat{\mathbb{R}}) = \{F : \widehat{\mathbb{R}} \to \mathbb{C} : \exists \mu \in M_b(\mathbb{R}) \text{ such that } \widehat{\mu} = F\}.$$

An element of $B(\widehat{\mathbb{R}})$ is a *Fourier-Stieltjes transform*, e.g., [Rud62] for an elegant, incisive, authoritative exposition.

b. It is elementary to check that the elements of $B(\widehat{\mathbb{R}})$ are uniformly continuous members of $C_b(\widehat{\mathbb{R}})$. It is natural to ask for an intrinsic characterization of $B(\widehat{\mathbb{R}})$ as we did for $A(\widehat{\mathbb{R}})$ in *Example 1.4.4*, i.e., to seek a theorem of the form, "a uniformly continuous element $F \in C_b(\widehat{\mathbb{R}})$ is an element of $B(\widehat{\mathbb{R}})$ if and only if ...", where "..." is a statement about the behavior of F on $\widehat{\mathbb{R}}$. In spite of *Theorem 2.7.10*, and some wonderful contributions by A. C. Berry (1931) for $A(\widehat{\mathbb{R}})$, Bochner (1934), Schoenberg (1934), Kreĭn (1940), Yosida (1944), R. S. Phillips (1950), and Doss (1971), the problem remains unsolved, cf., *Exercise 2.60.*

Even though $B(\widehat{\mathbb{R}})$ is much larger than $A(\widehat{\mathbb{R}})$ there are still elements from $C_0(\widehat{\mathbb{R}})$, and even $C_c(\widehat{\mathbb{R}})$, which are not in $B(\widehat{\mathbb{R}})$. In fact, membership in $B(\widehat{\mathbb{R}})$ is a predominantly local property, and in that sense is closely related to membership in $A(\widehat{\mathbb{R}})$ except at infinity, e.g., [Ben75, *Definition 2.4.2*] for an explanation of this opaque remark as well as further references, cf., *Definition 3.5.6.*

c. The hypotheses for the inversion formula in *Theorem 1.7.8* can be weakened. In fact, if $f \in L^1(\mathbb{R}) \cap B(R)$ then $f \in A(\mathbb{R})$. To see this, let $\mu \in A(\widehat{\mathbb{R}}) \cap M_b(\widehat{\mathbb{R}})$ have the property that $\mu^\vee = f$. Since $\mu \in A(\widehat{\mathbb{R}})$ we have $\mu \in L^1_{\text{loc}}(\widehat{\mathbb{R}})$, and so μ has no discrete or continuous singular part. Thus, $\mu \in L^1(\widehat{\mathbb{R}})$ by *Theorem 2.7.6.*

2.8 Definitions from probability theory

After our statement of the *Central Limit Theorem* in *Example 1.6.8*, we noted that the hypotheses involved the notion of a *probability density function (of a random variable) having mean* 0 *and variance* 1. We shall now define these and some other probabilistic terms which arise in Fourier analysis; but we do not make any pretense about explaining probabilistic ideas, cf., [Lam66], [Lam77], [Pri81] for such explanations.

2.8.1 Definition. PROBABILISTIC SETUP

a. A *probability measure* on $\mathbb{R}$ is an element $p \in M_{b+}(\mathbb{R})$ for which $\|p\|_1 = 1$. The pair $(\mathbb{R}, p)$ is a *probability space*. The *distribution function* $F \in BV(\mathbb{R})$ associated with p is the increasing function

$$F : \mathbb{R} \longrightarrow [0, 1]$$

having the properties

i. $\forall \alpha \in \mathbb{R}, \quad \lim_{\beta \to \alpha+} F(\beta) = F(\alpha).$

ii. $F' = p$,

e.g., let p be the Dirac measure δ. Caveat: The (Schwartz) distributional derivative in part *ii* has *nothing* to do with the probabilistic designation, *distribution function*, that is given to F.

b. *A random variable*,

$$X : (\mathbb{R}, p) \longrightarrow \mathbb{R},$$

is a measurable function $X : \mathbb{R} \longrightarrow \mathbb{R}$. Measurable functions are defined as limits of sequences of continuous functions in *Definition A.10a*; there is also a primordial measure theoretic definition. The *cumulative distribution function* F_X defined on the range of X is defined as

$$\forall t \in \mathbb{R}, \quad F_X(t) = \int \mathbf{1}_{S(t)}(\alpha)\, dp(\alpha),$$

where integration is over the probability space $(\mathbb{R}, p)$ and

$$S(t) = \{\alpha \in \mathbb{R} : X(\alpha) \leq t\}.$$

$F_X \in BV(\mathbb{R})$ is an increasing function

$$F_X : \mathbb{R} \longrightarrow [0, 1].$$

We have the following figure.

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{F} & [0,1] \\ X \downarrow & & \\ \mathbb{R} & \xrightarrow{F_X} & [0,1] \end{array}$$

Figure 2.2

F'_X is a probability measure on the range of X. X has a *probability density function* (pdf) f_X if

$$F'_X = f_X \in L^1(\mathbb{R}),$$

i.e., f is the *pdf* of a random variable X if $f \geq 0$, $\int f(t)\, dt = 1$, and F_X can be written as

$$F_X(t) = \int_{-\infty}^{t} f(u)\, du.$$

2.8.2 Definition. MEAN AND VARIANCE

a. The *mean* or *expectation* of a random variable $X \in L^1_p(\mathbb{R})$ defined on a probability space $(\mathbb{R}, p)$ is

$$m_X = E\{X\} = \int X(\alpha)\, dp(\alpha),$$

where the value of the integral, which defines the notation m_X and $E\{X\}$, is $p(X)$ in our distributional notation, e.g., *Remark 2.7.4*. $X \in L^1_p(\mathbb{R})$ means that $p(|X|) \equiv \int |X(\alpha)|\, dp(\alpha) < \infty$.

b. We can prove by straightforward calculations that

$$E\{X\} = \int X(\alpha)\, dF(\alpha) = \int t dF_X(t), \tag{2.8.1}$$

where the integrals are Riemann-Stieltjes integrals over $\mathbb{R}$, the domain $\mathbb{R}$ of the first integral being the probability space $(\mathbb{R}, p)$, and the domain $\mathbb{R}$ of the second integral being the range space of X. In the case X has a pdf f_X then (2.8.1) allows us to write

$$E\{X\} = \int t f_X(t)\, dt.$$

c. Let $X \in L^2_p(\mathbb{R})$, i.e.,

$$\|X\|_{2,p} = \left(\int |X(\alpha)|^2\, dp(\alpha)\right)^{1/2} < \infty.$$

The *variance* σ_X^2 of the random variable X is

$$\sigma_X^2 = E\{(X - m_X)^2\}.$$

Therefore,

$$\sigma_X^2 = E\{X^2\} - (E\{X\})^2 = E\{X^2\} - m_X^2.$$

σ_X is the *standard deviation* of X. In the case X has a pdf f_X then

$$\sigma_X^2 = \int (t - m_X)^2 f_X(t)\, dt.$$

In fact, a formal calculation shows that

$$\int g(X)(\alpha)\, dp(\alpha) = \int g(t) f_X(t)\, dt.$$

2.8.3 Example. VARIANCE AND DISPERSION

The notion of variance can be viewed as providing more relevant information than that provided by knowing a mean value. For example, the complacency, of knowing that the mean or average family income in a nation can provide a good standard of living, would be offset by a revolution for justice if such a nation had some destitute and homeless in its population. Thus, it is important to know how a nation's wealth is *dispersed* among the population. Variance is a measure of dispersion.

To fix ideas, let P be a finite set, for example, a class of students; and let $X : P \longrightarrow \mathbb{R}$ be a random variable, for example, a test score, i.e., $X(\alpha)$ is the score that student $\alpha \in P$ received. (Technically, we should first put a measure p on P.) The *mean* (score) of X is

$$m_X = \frac{1}{\operatorname{card} P} \sum_{\alpha \in P} X(\alpha),$$

which is a discrete way of writing $m_X = \int_P X(\alpha)\, dp(\alpha)$, where p is a "probability measure" defined by

$$\forall A \subseteq P, \quad p(A) = \frac{\operatorname{card} A}{\operatorname{card} P}.$$

The *average squared distance of* X *from the average* is the variance, viz.,

$$\sigma_X^2 = \frac{1}{\operatorname{card} P} \sum_{\alpha \in P} \left(X(\alpha) - m_X\right)^2 .$$

Consequently, the standard deviation is

$$\sigma_X = \text{“}\|X - m_X\|_{L^2_p(P)}\text{”};$$

and it makes sense to discuss *dispersion* in terms of how many standard deviations a given value (or score) $X(\alpha)$ is from the mean.

2.8.4 Definition. STOCHASTIC PROCESSES

a. Let $(\mathbb{R}, p)$ be a probability space with elements $\alpha \in \mathbb{R}$. The mapping

$$\begin{aligned} X : \mathbb{R} \times \mathbb{R} &\longrightarrow \mathbb{C} \\ (t, \alpha) &\longmapsto X(t, \alpha) \end{aligned}$$

is a *weakly stationary stochastic process* (WSSP), or equivalently, a *wide-sense stationary stochastic* process, if the following properties are valid:

i. $\forall t \in \mathbb{R}, \quad X(t, \cdot) \in L^2_p(\mathbb{R})$;

ii. $\exists m \in \mathbb{C}$ such that $\forall t, \quad E\{X(t)\} = m$,

where

$$E\{X(t)\} = \int X(t, \alpha)\, dp(\alpha);$$

iii. $\forall t, y, h \in \mathbb{R}, \quad E\{X(t+u)\overline{X(t+h)}\} = E\{X(u)\overline{X(h)}\}$;

iv. $\lim_{t \to 0} E\{|X(t) - X(0)|^2\} = 0$.

This definition makes sense for mappings $X : \mathbb{R} \times P \to \mathbb{C}$, where (P, p) is a quite general probability space (which we haven't defined) or something as specific as the set P of *Example 2.8.3.*

b. Part *iii* of the above definition implies $E\{X(t+u)\overline{X(t)}\} = E\{X(u)\overline{X(0)}\}$ for all $t, u \in \mathbb{R}$; and the function R, defined by

$$\forall t \in \mathbb{R}, \quad R(t) = E\{X(t+u)\overline{X(u)}\}$$

is the *stochastic autocorrelation* of the WSSP X. The *autocovariance* of X is the function C, defined by

$$\forall t \in \mathbb{R}, \quad C(t) = E\{(X(t+u) - m)(\overline{X(u) - m})\} = R(t) - |m|^2.$$

Thus, the variance of the WSSP X is

$$\sigma_X^2 = C(0) = R(0) - |m|^2.$$

2.8.5 Proposition.

The stochastic autocorrelation R of a WSSP X is a continuous positive definite function.

This result is proved by an elementary calculation, and the continuity of R results from part *iv* of the definition of a WSSP. If X does not satisfy part *iv* then R is still positive definite, and it has the decomposition $R = R_C + R_0$, where $R_C, R_0 >> 0, R_C$ is continuous, and $R_0 = 0$ a.e. This decomposition is due mostly to F. Riesz (Acta Sci. Math., 1933).

2.8.6 Definition. Power Spectrum and Spectral Estimation

a. Let X be a WSSP with stochastic autocorrelation R. Because of *Proposition 2.8.5* and the Herglotz-Bochner Theorem, there is $\mu \in M_{b+}(\widehat{\mathbb{R}})$ for which $\widehat{\mu} = R$ on $\mathbb{R}$. By notational tradition we set $S = \mu$, and, by definition, S is the *power spectrum* of X.

b. The *spectral estimation problem* is the clarify and quantify the statement: find periodicities in a signal X recorded over a fixed time interval. In more picturesque language, we want to filter the noise from the incoming signal X in order to determine the intelligent message (periodicities) therein, e.g., [Chi78], [IEEE82].

Such signals can sometimes be modelled as WSSPs [Bar78], [BTu59], [Bri81], [Pri81], and, then, the spectral estimation problem is one of power spectrum computation [Ben83, Part IV] or approximation.

2.8.7 Definition. Periodogram

Let X be a WSSP, for which $X(\cdot, \alpha) \in L^{\infty}(\mathbb{R})$ for each $\alpha \in \mathbb{R}$, and let $b \in L^1(\mathbb{R})$. The function

$$\begin{array}{rcl} S_b : \widehat{\mathbb{R}} \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (\gamma, \alpha) & \longmapsto & S_b(\gamma, \alpha) \end{array}$$

defined as

$$S_b(\gamma, \alpha) = \left| \int b(t) X(t, \alpha) e^{-2\pi i t \gamma} \, dt \right|^2$$

is the *periodogram* associated with the process X and the *data window* b.

Schuster initiated periodogram analysis, and his work was one of the major influences on Wiener's Generalized Harmonic Analysis (1930), e.g., [Wie81, Volume II, pages 183-324] and *Section 2.9*. The following calculation shows the role of periodograms in spectral estimation.

2.8.8 Proposition.

Let X be a real-valued WSSP, for which $X(\cdot, \alpha) \in L^{\infty}(\mathbb{R})$ for each $\alpha \in \mathbb{R}$, and let S_b be the periodogram associated with X for the real and even data window $b \in L^1(\mathbb{R})$. Assume $B^2 \equiv (\hat{b})^2 \in L^1(\hat{\mathbb{R}})$ and $\int B^2(\gamma)\,d\gamma = 1$. If S is the power spectrum of X then

$$E\{S_b(\gamma)\} = S * B^2(\gamma), \tag{2.8.2}$$

and

$$\lim_{\lambda\to\infty} E\{S_{\frac{1}{\sqrt{\lambda}}b(\frac{t}{\lambda})}(\gamma)\} = \lim_{\lambda\to\infty} S * (B^2)_\lambda = S \tag{2.8.3}$$

in the sense that

$$\forall F \in C_b(\hat{\mathbb{R}}), \quad \lim_{\lambda\to\infty} (S * (B^2)_\lambda)(F) = S(F), \tag{2.8.4}$$

where the subscript "λ" of B^2 designates dilation.

Proof. Our hypotheses allow us to verify (2.8.2) as follows:

$$\begin{aligned} E\{S_b(\gamma)\} &= \iint b(t)\overline{b(u)}e^{-2\pi i(t-u)\gamma}\hat{S}(t-u)\,dtdu \\ &= \int |B(\gamma+\omega)|^2\,dS(\omega) = S * B^2(\omega). \end{aligned}$$

The last step is a consequence of the following: since b is real and even, $|B|^2 = B^2$; and S is even since X real allows us to conclude that R is real and even.

(2.8.3) is a consequence of the facts that $\{(B^2)_\lambda\}$ is an approximate identity, and that $\frac{1}{\sqrt{\lambda}}b(\frac{t}{\lambda}) \longleftrightarrow \sqrt{\lambda}B(\lambda\gamma)$. In fact,

$$G(\alpha) = \int F(\alpha+\beta)\,dS(\beta) \in C_b(\hat{\mathbb{R}}),$$

so that by the definition of convolution and the evenness of B^2 we can apply (2.1.2), which in turn yields (2.8.3), cf., *Exercise 2.38.* □

2.8.9 Remark. ASYMPTOTICALLY UNBIASED ESTIMATOR

For any real and even data window $b \in L^1(\mathbb{R})$, *Proposition 2.8.8* allows us to refer to the periodograms $\{S_{\frac{1}{\sqrt{\lambda}}b(\frac{t}{\lambda})}\}$ as an *asymptotically unbiased estimator* of the power spectrum S.

The relatively weak type of convergence in (2.8.3) and (2.8.4) allows for a great deal of mischief on the part of the raw periodogram S_b if one thinks of it as an approximant to S. There are results due to Beurling, Herz, and Pollard, which are similar to (2.8.4), but for which the convergence is much "stronger". An example of such a theorem for closed intervals I, disjoint from $\operatorname{supp} S$, is

$$\lim_{T\to\infty} \int_{-T}^{T} R(t)e^{2\pi it\gamma}\, e^{|t|/T}\, dt = 0 \quad \text{uniformly on} \quad I,$$

cf., [Ben75, Section 2.1]. These results provide quantitative estimates for the support of S in terms of the support of the approximants.

2.8.10 Example. Michelson Interferometer and Spectral Estimation

a. In *Example 2.7.9* we defined a "deterministic" version of autocorrelation P for $f \in L^2(\mathbb{R})$, viz., the *L^2-autocorrelation*

$$P(t) = f * \tilde{f}(t) = \int f(t+u)\overline{f(u)}\, du;$$

and we noted that P is positive definite, just as the stochastic autocorrelation R of a WSSP is positive definite, cf., *Exercise 1.33.* We mention this to illustrate the following specific, but in some sense typical, use of autocorrelation.

b. It is often difficult to measure a signal directly, whereas one can experimentally measure its power. To be more precise, the spectral analysis of a beam of light $f(t)$, a real valued signal, can be made by a Michelson interferometer in the following way. The power or intensity of the beam is the energy flow per unit time (assuming area normalization) and is measured by a power-sensitive photometer. The interferometer allows the beam to take different paths of different length to the photometer. As such the intensity of $f(u) + f(u+t)$ can be measured for various lags t. Thus, the left side of the equation,

$$\int |f(u) + f(u+t)|^2\, du - 2\int |f(u)|^2\, du = 2\int f(u+t)f(u)\, du,$$

can be measured, noting that $\int |f(u+t)|^2\, du = \int |f(u)|^2\, du$. Consequently, the L^2-autocorrelation P is computable even though f may not be.

c. Summarizing, suppose we have a complicated signal f, considered deterministically on $\mathbb{R}$ or modelled by a WSSP X. Suppose, further, that f or X can not be analyzed directly, but that there are power measuring devices that allow us to quantify the autocorrelations $P >> 0$ or $R >> 0$, as we did in part b. Then the computation of the power spectra $P^\vee$ or $R^\vee$ allows us to determine significant frequencies of $P^\vee$ or $R^\vee$; and these frequencies are also significant in the behavior of f or X since, in the case of $P^\vee$, $P^\vee = |\hat{f}|^2$. This theme is further developed in *Section 2.9.*

2.8.11 Example. THE WAVE FUNCTION AND THE UNCERTAINTY PRINCIPLE

The wave theory in quantum mechanics arose since electron beams diffracted through crystals produced an effect analogous to Newton's spectral theory of white light diffracted through a prism. (Wiener's spectral theory explaining the polychromatic nature of sunlight, i.e., white light, is the Generalized Harmonic Analysis of *Section 2.9* [Wie81, Volume II, pages 183-324].)

For a fixed time t_0, the wave function $\Psi(\gamma)$, normalized so that $\|\Psi\|_{L^2(\widehat{\mathbb{R}})} = 1$, is a solution of Schrödinger's equation for a freely moving particle X (Schrödinger, 1926); and an important aspect of its physical significance is that the "probability" that X is in a given subset $A \subseteq \widehat{\mathbb{R}}$ is $\int_A |\Psi(\gamma)|^2\, d\gamma$, e.g., [Schi68]. This assertion defines a probability measure $p \in M_{b+}(\mathbb{R})$, and allows us to think of X as a random variable $X : (\mathbb{R}, p) \longrightarrow \widehat{\mathbb{R}}$, i.e., considered as a "measure of subsets" of $(\mathbb{R}, p)$ as in *Remark 2.7.4b*, p is defined as

$$p\{\alpha : X(\alpha) \in A\} = \int_A |\Psi(\gamma)|^2\, d\gamma.$$

The associated pdf is $f_X \equiv |\Psi|^2$. For the case of 0-mean, the *classical* (or *Heisenberg*) *uncertainty principle inequality* associated with the wave function Ψ and $\psi \equiv \Psi^\vee$ is

$$(2.8.5) \qquad 1 = \|\Psi\|^2_{L^2(\widehat{\mathbb{R}})} \leq 4\pi \|t\psi(t)\|_{L^2(\mathbb{R})} \|\gamma\Psi(\gamma)\|_{L^2(\widehat{\mathbb{R}})},$$

cf., [BF94, Chapter 7] as well as the "intuitive calculation" of the Heisenberg inequality in [Ben75, pages 77-79].

2.9 Wiener's Generalized Harmonic Analysis (GHA)

In 1930, Norbert Wiener [Wie81, Volume II, pages 183-324] proved an analogue of the Parseval-Plancherel formula, $\|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\hat{\mathbb{R}})}$, for functions which are *not* elements of $L^2(\mathbb{R})$. We refer to his formula as the *Wiener-Plancherel* formula, e.g., (2.9.2). It became a beacon in his perception and formulation of the statistical theory of communication, e.g., [Wie49], [Lee60]. Wiener even chose to have the formula appear on the cover of his autobiography, *I Am a Mathematician.* (This is a 20th century analogue of Archimedes' tombstone, which had a carving of a sphere inscribed in a cylinder to commemorate his "1:2:3" theorem, e.g., [Ben77] for details concerning the mathematical results, Cicero's role, and a recent update.)

Besides the motivation for GHA mentioned in *Example 2.8.11*, Wiener discussed the background for GHA in [Wie81, Volume II, pages 183-324]; and this background has been explained scientifically and historically in a virtuoso display of scholarship by Masani, e.g., Masani's remarkable commentaries in [Wie81, Volume II, pages 333-379], as well as [Mas90]. Two precursors, whose work Wiener studied and who should be mentioned vis a vis GHA, were Sir Arthur Schuster, cf., *Definition 2.8.7*, and Sir Geoffrey I. Taylor. Schuster pointed out analogies between the harmonic analysis of light and the statistical analysis of hidden periods associated with meteorological and astronomical data. Taylor conducted experiments in fluid mechanics dealing with the *onset to turbulence*, and formulated a special case of correlation. A third scientist, whose work (1914) vis a vis GHA was not known to Wiener, was Albert Einstein. Einstein writes: "Suppose the quantity y (for example, the number of sun spots) is determined empirically as a function of time, for a very large interval, T. How can one represent the statistical behavior of y?" In his heuristic answer to this question he came close to the notions of autocorrelation and power spectrum, e.g., *Section 2.8* and *Definition 2.9.5*, cf., [Mas90, pages 112-113], Einstein's paper (in Archive des Sciences Physiques et Naturelles, 37 (1914), 254-255), and commentaries by Masani and Yaglom.

The Fourier analysis of $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$ (*Chapter 1*) or the theory of Fourier series (*Chapter 3*) were inadequate tools to analyze the issues confronting Schuster, Taylor, and Einstein. On the other hand, GHA became a successful device to gain some insight into their problems, as well as other problems where the data and/or noises can not be modelled by the Fourier transform decay, finite energy, or periodicity inherent in the above classical theories, e.g., [Ars66, Chapter II], [Bas84], [Ric54].

The material in *Sections 2.9.1-2.9.10* outlines GHA and is due to Wiener [Wie81, Volume II, pages 183-324 and pages 519-619], [Wie33], cf., [Ben75, Chapter 2], [Ber87]. The higher dimensional theory, with its geometrical ramifications, is found in [BBE89], [Ben91a], cf., [AKM80].

2.9.1 Definition. BOUNDED QUADRATIC MEANS

The space $BQM(\mathbb{R})$ of functions having *bounded quadratic means* is the set of all functions $f \in L^2_{\text{loc}}(\mathbb{R})$ for which

$$\sup_{T>0} \frac{1}{2T} \int_{-T}^{T} |f(t)|^2 \, dt < \infty.$$

The *Wiener space* $W(\mathbb{R})$ is the set of all functions $f \in L^2_{\text{loc}}(\mathbb{R})$ for which

$$\int \frac{|f(t)|^2}{1+t^2} \, dt < \infty.$$

2.9.2 Theorem. INCLUSIONS FOR GHA

$$L^\infty(\mathbb{R}) \subseteq BQM(\mathbb{R}) \subseteq W(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R}),$$

and the inclusions are proper.

2.9.3 Definition. THE WIENER s-FUNCTION

The *Wiener s-function* associated with $f \in BQM(\mathbb{R})$ is defined as the sum $s = s_1 + s_2$ where

$$s_1(\gamma) = \int_{-1}^{1} f(t) \frac{e^{-2\pi i t\gamma} - 1}{-2\pi i t} \, dt$$

and

$$s_2(\gamma) = \int_{|t|\geq 1} f(t)\frac{e^{-2\pi it\gamma}}{-2\pi it}\,dt.$$

Since $f \in L^1[-1,1]$, we have $s_1 \in C(\mathbb{R})$ and $|s_1(\gamma)| \leq 2|\gamma|\|f\|_{L^1[-1,1]}$. Since $f \in BQM(\mathbb{R})$, *Theorem 2.9.2* and the Parseval-Plancherel Theorem allow us to conclude that $s_2 \in L^2(\mathbb{R})$. In particular, $s \in L^2_{\text{loc}}(\widehat{\mathbb{R}}) \cap \mathcal{S}'(\widehat{\mathbb{R}})$.

2.9.4 Theorem. THE DERIVATIVE OF THE WIENER s-FUNCTION

Let $f \in BQM(\mathbb{R})$. Then $f \in \mathcal{S}'(\mathbb{R})$ and

$$s' = \widehat{f},$$

where $s \in L^2_{\text{loc}}(\widehat{\mathbb{R}}) \cap \mathcal{S}'(\widehat{\mathbb{R}})$ is the Wiener s-function associated with f (Exercise 2.61).

2.9.5 Definition. DETERMINISTIC AUTOCORRELATION

The *deterministic autocorrelation* R of a function $f : \mathbb{R} \to \mathbb{C}$ is formally defined as

$$R(t) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(u+t)\overline{f(u)}\,du.$$

To fix ideas, suppose R exists for each $t \in \mathbb{R}$. It is easy to prove that $R >> 0$, and so $R = \widehat{S}$ for some $S \in M_{b+}(\widehat{\mathbb{R}})$. We have used the same notation, viz., R, to denote both deterministic and stochastic autocorrelation since they are often same, e.g., *Theorem 2.9.11.* As in the stochastic case, S is called the *power spectrum* of f.

The *Wiener-Plancherel formula* is equation (2.9.2) in the following result.

2.9.6 Theorem. WIENER-PLANCHEREL FORMULA

Let $f \in BQM(\mathbb{R})$, and suppose its deterministic autocorrelation $R = \widehat{S}$ exists for each $t \in \mathbb{R}$.

a. Then

$$\forall t \in \mathbb{R}, \quad R(t) = \lim_{\epsilon\to 0} \frac{2}{\epsilon}\int |\Delta_\epsilon s(\gamma)|^2 e^{-2\pi it\gamma}\,d\gamma, \tag{2.9.1}$$

where $\Delta_\epsilon s(\gamma) \equiv \frac{1}{2}(s(\gamma+\epsilon) - s(\gamma-\epsilon))$.

b. In particular,

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |f(t)|^2\,dt = \lim_{\epsilon\to 0} \frac{2}{\epsilon}\int |\Delta_\epsilon s(\gamma)|^2\,d\gamma. \tag{2.9.2}$$

2.9.7 Example. RELATED FORMULAS AND SPECTRAL ESTIMATION

a. Because of (2.9.1) and assuming the setup of *Theorem 2.9.6*, the following formulas are true under the proper hypotheses, e.g., [Ben75, page 90], [Ben91b, page 847]:

$$\lim_{\epsilon\to 0} \frac{2}{\epsilon}|\Delta_\epsilon s(\gamma)|^2 = S, \tag{2.9.3}$$

and

$$\begin{aligned} \int |\hat{k}(\gamma)|^2\,dS(\gamma) &= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} |k * f(t)|^2\,dt \\ &= \lim_{\epsilon\to 0}\frac{2}{\epsilon}\int |\hat{k}(\gamma)\Delta_\epsilon s(\gamma)|^2\,d\gamma. \end{aligned} \tag{2.9.4}$$

b. Formally, (2.9.4) is (2.9.2) for the case $k = \delta$. For $k \in C_c(\mathbb{R})$ the first equality of (2.9.4) is not difficult, e.g., [Ben91b, pages 847-848]. The second equality, or, equivalently, *Theorem 2.9.6*, requires *Wiener's Tauberian Theorem*, e.g., *Theorem 2.9.12*.

c. The following diagram illustrates the action and "levels" of the functions and measure in *Theorem 2.9.6* for a given signal f.

$$\begin{array}{ccccc} f & \longleftrightarrow & \hat{f} = s' & & s \\ \downarrow & & & & \downarrow \\ R = \hat{S} & \longleftrightarrow & S & & \{\frac{2}{\epsilon}|\Delta_\epsilon s|^2|\} \end{array}$$

Figure 2.3

d. Since S is the "power" spectrum, (2.9.2) and (2.9.3) allow us to assert that

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} |f(t)|^2\,dt$$

is a measure of the total power of f, cf., Wiener's comparison of energy and power in [Wie49, pages 39-40 and 42]. In light of the spectral estimation problem of *Definition 2.8.6*, the middle term of (2.9.4) is a measure of the power in a frequency band $[\alpha, \beta]$ if $\widehat{k} = \mathbf{1}_{[\alpha,\beta]}$ in the first term of (2.9.4), cf., [Ben91b, Theorem 5.2].

2.9.8 Remark. WIENER-PLANCHEREL FORMULA

The Parseval-Plancherel formula, $\|f\|_{L^2(\mathbb{R})} = \|\widehat{f}\|_{L^2(\widehat{\mathbb{R}})}$, allowed us to define the Fourier transform of a square integrable function (*Theorem 1.10.2*), and, at certain levels of abstraction, it is considered to characterize what is meant by an harmonic analysis of f. On the other hand, for most applications in $\mathbb{R}$, the formula assumes the workaday role of an effective tool used to obtain quantitative results. It is this latter role that was envisaged for the Wiener-Plancherel formula in dealing with the non-square-integrable case. After all, distribution theory gives the proper definition of the Fourier transform of tempered distributions. The real issue is to obtain quantitative results for problems where an harmonic analysis of a non-square-integrable function is desired. As mentioned above, a host of such problems comes under the heading of an harmonic (spectral) analysis of signals containing non-square-integrable noise and/or random components, whether it be speech recognition, image processing, geophysical modeling, or turbulence in fluid mechanics. Such problems can be attacked by Beurling's profound theory of spectral synthesis, e.g., *Example 2.4.6f*, as well as by the extensive multifaceted theory of time series, e.g., *Section 2.8.* Beurling's spectral synthesis does not deal with energy and power considerations, i.e., quadratic criteria, and time series relies on a stochastic point of view. The Wiener-Plancherel formula deals with these problems deterministically, and, hence, with potential for real implementation, e.g., *Example 2.9.7d.*

2.9.9 Example. ELEMENTARY POWER SPECTRA

a. The value of the autocorrelation R is that it can be measured in many cases where the underlying signal f can not be quantified, e.g., *Example 2.8.10.* Also, the discrete part of the power spectrum S characterizes periodicities in f, e.g., [Wie48, Chapter X]. This first is illustrated by taking $f(t) = \sum_{k=1}^{n} r_k e^{-2\pi i t \lambda_k}$, $r_k \in \mathbb{C}$, $\lambda_k \in \widehat{\mathbb{R}}$. The L^2-autocorrelation is not defined, but the deterministic autocorrelation is $R(t) = \sum_{k=1}^{n} |r_k|^2 e^{-2\pi i t \lambda_k}$ (by direct calculation); and hence the power spectrum is

$$S = \sum_{k=1}^{n} |r_k|^2 \, \delta_{\lambda_k}.$$

b. If $f : \mathbb{R} \to \mathbb{C}$ has the property that $\lim_{|t| \to \pm\infty} f(t) = 0$, then $S = 0$. It is elementary to construct examples f for which $S = 0$ whereas $\overline{\lim}_{|t| \to \pm\infty} |f(t)| > 0$, cf., [Wie33, pages 151-154], [Bas84, pages 99-100], [Ben75, pages 84 and 87], [Ben83, Section IV].

2.9.10 Definition. CORRELATION ERGODICITY

Let X be a WSSP with stochastic autocorrelation R. X is a *correlation ergodic process* if

$$\text{(2.9.5)} \qquad \forall t \in \mathbb{R}, \quad \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} X(t+u, \alpha) \overline{X(u, \alpha)} \, du = R(t)$$

in measure. (Convergence in measure is defined in *Example A.11c.*)

Because of (2.9.5), the following result establishes the relationship between the notions of deterministic and stochastic autocorrelation, cf., [Pap77, pages 354-360].

2.9.11 Theorem. CRITERION FOR CORRELATION ERGODICITY

Let X be a WSSP with stochastic autocorrelation R. X is a correlation ergodic process if

$$\forall t \in \mathbb{R}, \quad \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} C(t,v)\left(1 - \frac{|v|}{2T}\right) dv = 0,$$

where $C(t,v) = E\{X(t+u+v)\overline{X(u+v)}X(t+u)\overline{X(u)}\} - |R(t)|^2$.

As mentioned in *Example 2.9.7*, the following result is required to prove the Wiener-Plancherel formula.

2.9.12 Theorem. WIENER TAUBERIAN THEOREM

Let $g \in L^1(\mathbb{R})$ have a nonvanishing Fourier transform and let $\varphi \in L^\infty(\mathbb{R})$. If

$$\lim_{t\to\infty} g * \varphi(t) = r \int g(u)\, du \tag{2.9.6}$$

then

$$\forall f \in L^1(\mathbb{R}), \quad \lim_{t\to\infty} f * \varphi(t) = r \int f(u)\, du. \tag{2.9.7}$$

2.9.13 Remark. WIENER TAUBERIAN THEOREM

a. *Theorem 2.9.12* has the format of *classical* Tauberian theorems: a boundedness (or some other) condition and "summability" by a certain method yield "summability" by other methods. In *Theorem 2.9.12*, the boundedness or "Tauberian" condition is the hypothesis that $\varphi \in L^\infty(\mathbb{R})$. The given summability is (2.9.6), where g represents a so-called "summability method". The conclusion (2.9.7) of the theorem is summability for a whole class of summability methods, viz., for all $f \in L^1(\mathbb{R})$. A classical and masterful treatment of summability methods is due to Hardy [Har49].

If g is the Gaussian defined in *Example 1.3.3*, then $\widehat{g}$ never vanishes. Thus, in this case, if $\varphi \in L^\infty(\mathbb{R})$ has the property that

$$\lim_{t\to\infty} g * \varphi(t) = r$$

then

$$\forall \lambda, \quad \lim_{t\to\infty} w_\lambda * \varphi(t) = r,$$

where $\{w_\lambda\}$ is the Fejér kernel.

The particular functions used by Wiener to prove his Wiener Tauberian formulas are found in [Wie33], [Ben75, pages 91-92].

b. *Modern* Tauberian theorems have a more algebraic and/or functional analytic flavor to them. For example, the Wiener Tauberian Theorem is a special case of the fact that *if $\widehat{g} \in A(\mathbb{R})$, $T \in A'(\mathbb{R})$, and $T\widehat{g} = 0$, then $\widehat{g} = 0$ on* $\operatorname{supp} T$. In fact, the generalizations of *Theorem 2.9.12* are much more far reaching than this. [Ben75] gives an extensive treatment of both classical and modern Tauberian theory, as well as the history of the subject, and applications to spectral synthesis and analytic number theory.

Because of the importance of translation invariant systems and the theory of multipliers, e.g., *Definition 2.6.5* and *Example 2.6.6*, we define the *closed translation invariant subspace* V_g *generated by* $g \in X$, where X is $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$, to be the closure in X of the linear span of $\{\tau_t g : t \in \mathbb{R}\}$. We write

$$V_g = \overline{\text{span}}\{\tau_t g : t \in \mathbb{R}\}. \tag{2.9.8}$$

2.9.14 Theorem. ZERO SETS AND DENSE SUBSPACES

a. If $g \in L^1(\mathbb{R})$ *and* $\widehat{g}$ *never vanishes then* $V_g = L^1(\mathbb{R})$.

b. If $g \in L^2(\mathbb{R})$ *and* $|\widehat{g}| > 0$ *a.e. then* $V_g = L^2(\mathbb{R})$.

Proof. Part *a* is the Wiener Tauberian Theorem, and we refer to [Wie33], [Ben75, pages 25-26, 49-50, 94-95, and Section 2.3] for proofs.

The proof of part *b* is much simpler than that of part *a*, and so we shall give it here. Suppose $V_g \neq L^2(\mathbb{R})$. Then there is $h \in L^2(\mathbb{R})\backslash\{0\}$ such that

$$\forall t \in \mathbb{R}, \quad \int (\tau_t g)(u)\overline{h(u)}\, du = 0. \tag{2.9.9}$$

Equation (2.9.9) is a consequence of the Hahn-Banach Theorem and the fact that $L^2(\mathbb{R})' = L^2(\mathbb{R})$, e.g., *Theorem B.14*. By the Parseval-Plancherel Theorem,

$$\forall t \in \mathbb{R}, \quad \int \widehat{g}(\gamma)\overline{\widehat{h}(\gamma)}e^{-2\pi i t\gamma}\, d\gamma = 0.$$

$\widehat{g}\overline{\widehat{h}} \in L^1(\widehat{\mathbb{R}})$ by Hölder's inequality, and so, by the L^1-uniqueness theorem (*Theorem 1.6.9c*), $\widehat{g}\overline{\widehat{h}} = 0$ a.e. Since $|\widehat{g}| > 0$ a.e. we conclude that $\widehat{h} = 0$ a.e., and this contradicts the hypothesis on h. Thus, $V_g = L^2(\mathbb{R})$. □

Subspaces such as V_g in (2.9.8) play an important role in Gabor and wavelet decompositions in the case that the set of translates $\tau_t g$ is reduced to $\{\tau_r g : r \in D\}$ where D is a discrete subset of $\mathbb{R}$, e.g., [Mey90], [Dau92], [BF94].

2.9.15 Remark. GHA AND THE WIENER-KHINCHIN THEOREM

a. In GHA, a function f is analyzed for its frequency information by computing its autocorrelation R and its power spectrum $S = R^{\vee} \in M_{b+}(\widehat{\mathbb{R}})$. Mathematically, this is a mapping between a class of functions f and a class of measures $S \in M_{b+}(\widehat{\mathbb{R}})$. A natural question to ask is the following: for any $\mu \in M_{b+}(\widehat{\mathbb{R}})$, does there exist f whose autocorrelation R exists, and for which $R^{\vee} = \mu$?

b. The question of part *a* is answered affirmatively in the WSSP case by the Wiener-Khinchin Theorem: *a necessary and sufficient condition for R to be the stochastic autocorrelation of some* WSSP *X is that there exist $S \in M_{b+}(\widehat{\mathbb{R}})$ for which $\widehat{S} = R$.* In one direction, if R is the stochastic autocorrelation of a WSSP X then $S \equiv R^{\vee} \in M_{b+}(\widehat{\mathbb{R}})$ by the Herglotz-Bochner Theorem. The question in part *a* deals with the opposite direction, and the positive answer is not difficult to prove, e.g., [Pri81, pages 221-222], [DM76, pages 62-63 and 72-73]. Khinchin's proof dates from 1934, and there were further probabilistic contributions by Wold (1938), Cramér (1940), and Kolmogorov [Kol41], cf., [Ben92a].

c. The deterministic and constructive affirmative answer to the question in part *a* is the *Wiener-Wintner Theorem* (1939) [Wie81]. Bass and Bertrandis made significant contributions to this result, e.g., [Bas84]; and the multidimensional version is found in [Ben91b], [Ker90].

2.9.16 Theorem. WIENER-WINTNER THEOREM

Let $\mu \in M_{b+}(\widehat{\mathbb{R}})$. There is a constructible function $f \in L^{\infty}_{\mathrm{loc}}(\mathbb{R})$ such that its deterministic autocorrelation R exists for all $t \in \mathbb{R}$, and $R^{\vee} = \mu$.

2.10 exp$\{it^2\}$

The function

$$s(t) = \frac{1}{\sqrt{\pi i}} e^{it^2}$$

has properties which serve as a paradigm for the *method of stationary phase* (*Example 2.10.4*), as well as being an underlying kernel for the *oscillatory integrals* which arise in areas as diverse as analytic number theory, e.g., [Tit51, pages 61–70 and 83–84], and partial differential equations, e.g., [Hör83, Sections 7.7 and 7.8], cf., [Ste93]. In optics, s is the convolution kernel in *Fresnel's approximation* to the Huygens-Fresnel principle. (Fresnel's approximation allows for realistic diffraction-pattern calculations, e.g., [Good68].) In signal processing, s is the linear *chirp signal* whose frequency changes linearly with time, e.g., *Figure 2.4*. This is an example from the class of *frequency modulated (FM) signals* $e^{2\pi i\varphi(t)}$ that arise in subjects such as radar and sonar, e.g., [Rih85], [BMW91], and that are characterized by the property that φ' is not a constant.

no space

With regard to *Fourier optics*, the integrals $x(t) = \int_0^t \cos^2 u\, du$ and $y(t) = \int_0^t \sin^2 u\, du$ in *Theorem 2.10.1* are called *Fresnel integrals*, cf., *Example 2.10.7*. Also, in light of (2.10.1), note that $s \notin L^1(\mathbb{R}) \cup L^2(\mathbb{R})$.

less space

2.10.1 Theorem. FRESNEL INTEGRALS

a. Let s_λ be the dilation of s for $\lambda > 0$. Then

(2.10.1)
$$\int s_\lambda(t)\, dt = 1$$

in the sense that the improper Riemann integral $\int_0^\infty s_\lambda(t)\, dt$ equals $1/2$.

place earlier in §2.10

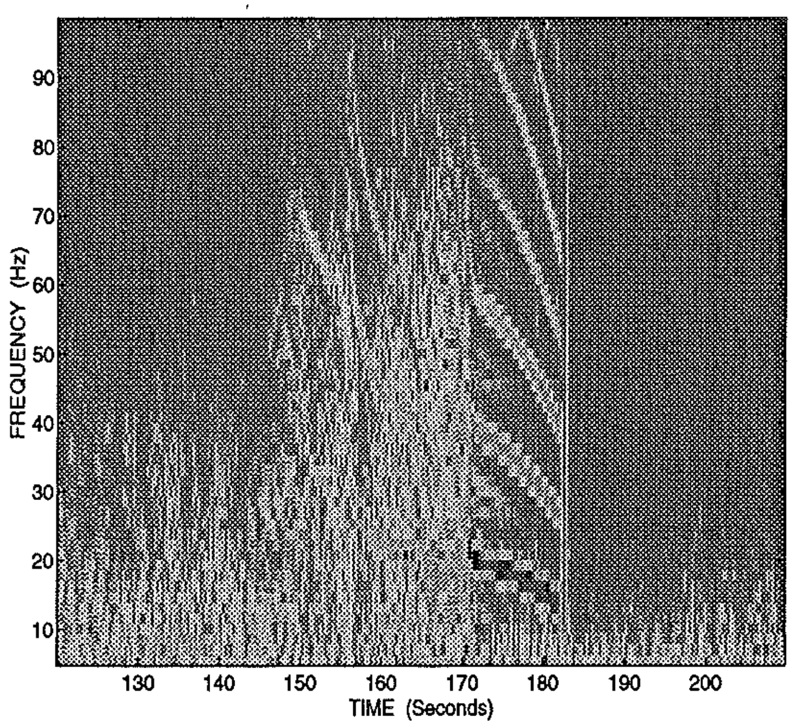

Figure 2.4

Spectrogram for electrocorticogram data [BeC95]. Chirp signal behavior is exhibited in the time interval 170 seconds to 182 seconds.

b.

$$\int_0^\infty \cos t^2\,dt = \int_0^\infty \sin t^2\,dt = \frac{1}{2}\sqrt{\frac{\pi}{2}}. \tag{2.10.2}$$

Proof. To obtain (2.10.2) from (2.10.1) we compute

$$\sqrt{i\pi} = \int_{-\infty}^{\infty} \cos t^2\,dt + i\int_{-\infty}^{\infty} \sin t^2\,dt$$

$$= 2\int_0^{\infty} \cos t^2\,dt + 2i\int_0^{\infty} \sin t^2\,dt.$$

Thus, since $i^{1/2} = (e^{i\frac{\pi}{2}})^{\frac{1}{2}} = e^{i\frac{\pi}{4}} = \frac{1}{\sqrt{2}}(1+i)$, we have (2.10.2).

Now, let us verify (2.10.1). Let $s(z) = \frac{1}{\sqrt{i\pi}}e^{iz^2}$ and consider the positively oriented wedge $C = [0, R] \cup C_R \cup L_R$ as illustrated in *Figure 2.5*:

Figure 2.5

By Cauchy's theorem, e.g., [Rud66], we compute

$$\sqrt{i\pi}\int_C s(z)\,dz = \int_0^R e^{ix^2}\,dx + \int_{C_R} e^{iz^2}\,dz + \int_{L_R} e^{iz^2}\,dz = 0. \tag{2.10.3}$$

The parametric representation of C_R is $z : [0, \frac{\pi}{4}] \to \mathbb{C}$, $\theta \mapsto \mathrm{Re}^{i\theta}$, and the parametric representation of $-L_R$ is $z : [0, R] \to \mathbb{C}$, $r \mapsto re^{i\pi/4}$. We estimate and compute the integrals in (2.10.3) as follows:

$$\begin{aligned} \left|\int_{C_R} e^{iz^2}\,dz\right| &= \left|iR\int_0^{\frac{\pi}{4}} e^{iR^2(\cos\theta + i\sin\theta)^2} e^{i\theta}\,d\theta\right| \\ &\le R\int_0^{\frac{\pi}{4}} e^{-(2R^2\cos\theta)\sin\theta}\,d\theta \\ &\le R\int_0^{\frac{\pi}{4}} e^{-(2R^2\cos\frac{\pi}{4})\sin\theta}\,d\theta \\ &\le R\int_0^{\frac{\pi}{2}} e^{-(2R^2\cos\frac{\pi}{4})\sin\theta}\,d\theta < \frac{\pi}{2}\left(\frac{1}{2R^2\cos\frac{\pi}{4}}\right) \end{aligned} \tag{2.10.4}$$

and

$$\begin{aligned}\int_{L_R} e^{iz^2}\,dz &= -\int_0^R e^{ir^2 e^{i\pi/2}} e^{i\pi/4}\,dr \\ &= -e^{i\pi/4}\left(\int_0^R e^{-r^2}\,dr\right).\end{aligned} \tag{2.10.5}$$

The last inequality in (2.10.4) follows from Jordan's inequality,

$$\int_0^{\frac{\pi}{2}} e^{-r\sin\theta}\,d\theta < \frac{\pi}{2r},$$

e.g., *Exercise 2.62a*. The right side of (2.10.4) tends to 0 as $R \to \infty$ and the right side of (2.10.5) tend to $-\frac{\sqrt{\pi}}{2}\left(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right)$ as $R \to \infty$. Equation (2.10.1) then follows from (2.10.3). □

2.10.2 Example. $\lambda\cos(\lambda t)^2$

Figure 2.6 is the graph of the dilation $\lambda\cos(\lambda t)^2$. Thus, for large values of λ, $\lambda\cos(\lambda t)^2$ is rapidly oscillating, especially near the origin. In light of the fact that $\int s_\lambda(t)\,dt = 1$, it should be noted that, except for the large amplitude at the origin, *Figure 2.6* does not resemble the approximate identities of *Chapter 1*, cf., *Theorem 2.10.3b.*

Figure 2.6

Note that the difference between consecutive "crests" of $\lambda\cos(\lambda t)^2$ has order of magnitude $\frac{1}{\lambda\sqrt{n}}$, and so tends to 0 as $n \to \infty$.

In fact, we have

$$\frac{1}{\lambda}\left(\sqrt{2\pi n}-\sqrt{2\pi(n-1)}\right)=\frac{\sqrt{2\pi}}{\lambda}\left(\frac{\sqrt{n}-\sqrt{n-1}}{\sqrt{n}+\sqrt{n-1}}\right)\left(\sqrt{n}+\sqrt{n-1}\right)$$
$$\frac{\sqrt{2\pi}}{\lambda}\frac{1}{\sqrt{n}+\sqrt{n-1}}\longrightarrow 0,\ n\longrightarrow\infty.$$

At this point we know that $\{s_\lambda\} \subseteq L^\infty(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R})$, $s_\lambda \notin L^1(\mathbb{R}) \cup L^2(\mathbb{R})$, and $\int s_\lambda(t)\,dt = 1$. Thus, theoretically, $\widehat{s}_\lambda \in A'(\widehat{\mathbb{R}}) \subseteq \mathcal{S}'(\widehat{\mathbb{R}})$ exists, and, tantalizingly, we'd like to know to what extent $\{s_\lambda\}$ is an "approximate identity", even though it isn't an approximate identity as defined in *Chapter 1.* The following result computes $\widehat{s}_\lambda$ and answers the "approximate identity" query.

2.10.3 Theorem. $\widehat{s}_\lambda$ AND $s_\lambda \to \delta$

a. The distributional Fourier transform $\widehat{s}_\lambda$ of s_λ is

$$\forall\gamma\in\widehat{\mathbb{R}},\quad \widehat{s}_\lambda(\gamma)=e^{-i(\pi\gamma/\lambda)^2}, \tag{2.10.6}$$

and so $s_\lambda \in L^\infty(\mathbb{R}) \cap A'(\mathbb{R})$.

b. $\lim_{\lambda\to\infty} s_\lambda = \delta$ in the sense that

$$\forall f\in L^1(\mathbb{R})\cap A(\mathbb{R}),\quad \lim_{\lambda\to\infty}\int s_\lambda(t)f(t)\,dt=\delta(f). \tag{2.10.7}$$

Proof. i. Formally, by completing the square and invoking *Theorem 2.10.1a* we compute

$$\widehat{s}_\lambda(\gamma)=\frac{\lambda}{\sqrt{\pi i}}\int \exp i\left\{\left(\lambda t-\frac{\pi\gamma}{\lambda}\right)^2-\left(\frac{\pi\gamma}{\lambda}\right)^2\right\}dt=e^{-i(\pi\gamma/\lambda)^2}.$$

Similarly,

$$\int s_\lambda(t)\overline{f(t)}\,dt=\int e^{-i(\pi\gamma/\lambda)^2}\overline{\widehat{f}(\gamma)}\,d\gamma. \tag{2.10.8}$$

ii. The formal calculation to obtain (2.10.8) can be justified for $f \in L^1(\mathbb{R}) \cap A(\mathbb{R})$. In this case, the integrals of (2.10.8) are Lebesgue integrals, and the validity of the equality is a consequence of the inversion formula (*Theorem 1.7.8*), LDC, and the fact that

$$\exists M \text{ such that } \forall a, b, \quad \left| \int_a^b e^{it^2}\, dt \right| \leq M,$$

e.g., *Exercises 2.64* and *2.68*, cf., *Exercise 2.63*. Equation (2.10.6) follows from (2.10.8) since $\mathcal{S}(\mathbb{R}) \subseteq L^1(\mathbb{R}) \cap A(\mathbb{R})$.

iii. Noting that

$$\forall \gamma, \quad \lim_{\lambda \to \infty} e^{-i(\pi\gamma/\lambda)} = 1,$$

we can use LDC and the inversion formula again to obtain (2.10.7). □

2.10.4 Example. STATIONARY PHASE

a. For a given compactly supported function f and real-valued phase $\varphi \in C^2(\mathbb{R})$, there is the associated *oscillatory integral*

$$F_\varphi(\gamma) = \int f(t) e^{2\pi i \varphi(t)\gamma}\, dt. \tag{2.10.9}$$

Obviously, the Fourier transform is an oscillatory integral. It is often important to investigate the behavior of F_φ for large values of γ. The *method of stationary phase* asserts that this behavior is determined by the so-called *stationary points* t for which $\varphi'(t) = 0$; and the method provides a means for quantifying this behavior, e.g., part b. Of course, in the case of the Fourier transform of $f \in L^1(\mathbb{R})$, where $\varphi' = 1$, we have $\lim_{\gamma \to \infty} F_\varphi(\gamma) = 0$, cf., *Exercise 2.63* for more general phases φ without stationary points.

The investigation of integrals F_φ in this spirit goes back to Airy (1838) and Stokes (1850), and the method of stationary phase is just another brilliant chapter in Riemann's thesis on trigonometric series [Rie1873, Section XIII]. It should be mentioned that *Laplace's asymptotic method* (1820), e.g., part c, preceded stationary phase; and, although Laplace's method doesn't deal directly with oscillatory integrals, it has striking resemblances in both technique and result with stationary phase, e.g., [Olv74], [Wid41].

b. Lord Kelvin (1887) made the following sort of observation about F_φ, most probably without knowledge of Riemann's work. If $f \in C_c(\mathbb{R})$ and γ is large, then the integral in (2.10.9) is very small because of the cancellation resulting from oscillation, *except* possibly near stationary points of φ since φ changes slowly near such points.

To quantify Lord Kelvin's point of view, assume $\operatorname{supp} f = [a,b]$ for $f \in C_c(\mathbb{R})$, $\varphi \in C^2(\mathbb{R})$, $\varphi'(t_0) = 0$ for some $t_0 \in (a,b)$, $\varphi'(t) \neq 0$ for all $t \in [a,b] \setminus \{t_0\}$, and $\varphi^{(2)}(t_0) > 0$. Then, because of the oscillation, Lord Kelvin's observation asserts that

$$F_\varphi(\gamma) \approx \int_I f(t_0) \exp\left\{2\pi i \gamma \left[\varphi(t_0) + \tfrac{1}{2}(t-t_0)^2 \varphi^{(2)}(t_0)\right]\right\} dt \tag{2.10.10}$$

for large γ, where I is a small interval about t_0. By the oscillation again, (2.10.10) can be replaced by

$$\begin{aligned} F_\varphi(\gamma) &\approx f(t_0) e^{2\pi i \varphi(t_0)\gamma} \int e^{\pi i t^2 \varphi^{(2)}(t_0)\gamma}\, dt \\ &= e^{\pi i/4} f(t_0) e^{2\pi i \varphi(t_0)\gamma} \frac{1}{(\gamma \varphi^{(2)}(t_0))^{1/2}}, \end{aligned} \tag{2.10.11}$$

where the right side is a consequence of *Theorem 2.10.1.* We shall say that F is *asymptotic* to G as $\gamma \to \infty$ to mean that $\lim_{\gamma\to\infty} F(\gamma)/G(\gamma) = 1$. The notation for this asymptotic behavior is $F(\gamma) \sim G(\gamma)$, $\gamma \to \infty$. As such, and noting that the heuristic argument leading to (2.10.11) can be made rigorous, the method of stationary phase allows us to assert that

$$F_\varphi(\gamma) \sim e^{\pi i/4} f(t_0) e^{2\pi i \varphi(t_0)\gamma} \frac{1}{(\gamma \varphi^{(2)}(t_0))^{1/2}}, \quad \gamma \longrightarrow \infty, \tag{2.10.12}$$

given our hypotheses on f and φ, cf., [Hör83], [Olv74], [Ste93] for more advanced results.

c. For comparison with (2.10.12), the asymptotic relation (2.10.13) is an elementary form of Laplace's asymptotic method. Let

$$G_\varphi(\gamma) = \int f(t) e^{2\pi \varphi(t)\gamma}\, dt.$$

Assume $\operatorname{supp} f = [t_0, b]$, $f \in C[t_0, b]$, $f(t_0) \neq 0$, $\varphi \in C^2[t_0, b]$, $\varphi'(t_0) = 0$, $\varphi^{(2)}(t_0) < 0$, *and* φ *non-increasing on* $[t_0, b]$. *Then*

$$G_\varphi(\gamma) \sim f(t_0) e^{2\pi\varphi(t_0)\gamma} \frac{1}{(-4\gamma\varphi^{(2)}(t_0))^{1/2}}, \ \gamma \longrightarrow \infty, \tag{2.10.13}$$

cf., [Wid41, Chapter 7] for this and more advanced results.

We shall define the Stieltjes transform in *Exercise 2.56*, as well as noting its close relationship to the Hilbert and iterated Laplace transforms. It turns out that a version of Laplace's asymptotic method is central to establish an inversion theory for the Stieltjes transforms of distributions, e.g., our theory of Stieltjes transforms in *Analytic representation of generalized functions*, Math. Zeitschr., 97(1967), 303–319.

2.10.5 Example. CHIRP TRANSFORM ALGORITHM

The *chirp transform algorithm* is the formula

$$\forall f \in L^1(\mathbb{R}), \quad \widehat{f} = \left[(f\overline{s}_{\sqrt{\pi}}) * s_{\sqrt{\pi}}\right] (\overline{i^{1/2} s_{\sqrt{\pi}} s_{\sqrt{\pi}}}), \tag{2.10.14}$$

where $s_{\sqrt{\pi}}$ is the $\sqrt{\pi}$-dilation of s. The verification of (2.10.14) is elementary:

$$(f\overline{s}_{\sqrt{\pi}}) * s_{\sqrt{\pi}}(\gamma) = \int f(u)(\overline{i^{-1/2}}) e^{-i\pi u^2} (i^{-1/2}) e^{i\pi(\gamma - u)^2}\, du = e^{i\pi\gamma^2} \widehat{f}(\gamma).$$

Equation (2.10.14) can be "implemented" for a given functin f by the following sequence of operations: multiply f by $\overline{s}_{\sqrt{\pi}}$, convolve $f\overline{s}_{\sqrt{\pi}}$ with $s_{\sqrt{\pi}}$, multiply $(f\overline{s}_{\sqrt{\pi}}) * s_{\sqrt{\pi}}$ by $(\overline{i^{1/2} s_{\sqrt{\pi}}})$. In signal processing, there is a simple block diagram for these operations, cf., [OW83, pages 511–512] or the terminology in *Definition 2.6.5*. We have chosen the word "algorithm" to describe the "fact" (2.10.14), since both f and $\widehat{f}$ have the same domain $\mathbb{R}$; in fact, our point of view is to consider f and $\widehat{f}$ as being defined on different spaces, whether they be time and frequency axes or dual groups.

2.10.6 Example. INFINITE FREQUENCIES

a. Let s_λ, $\lambda > 0$, be the λ-dilation of s. Then the deterministic autocorrelation R_λ of s_λ is

$$R_\lambda(t) = \begin{cases} 0, & \text{if } t \neq 0, \\ \frac{\lambda^2}{\pi}, & \text{if } t = 0. \end{cases} \tag{2.10.15}$$

To verify (2.10.15), we need only substitute into the definition of R_λ, and compute

$$R_\lambda(t) = \lim_{T\to\infty} \frac{\lambda^2}{2\pi T} e^{i(\lambda t)^2} \int_{-T}^{T} e^{2i\lambda^2 ut}\, du.$$

$R_\lambda >> 0$ is discontinuous, and the power spectrum S_λ of s_λ is the 0-measure, cf., *Exercise 2.66.*

b. Using the terminology of *Example 2.9.7d*, we see that the total power of s_λ is $R_\lambda(0)$, and that $R_\lambda(0) > \lim_{t\to 0} R_\lambda(t) = 0$. Since $S_\lambda = 0$, there is a portion of the power not represented by any finite frequencies; and Wiener reasoned that s_λ draws part of its power from so-called *infinite frequencies*, e.g., [Wie49, page 40].

c. For perspective with regard to the function s, if $f(t) = e^{2\pi i|t|}$ then its deterministic autocorrelation is $R(t) = \cos 2\pi t$ and its power spectrum is $S = \frac{1}{2}(\delta_1 + \delta_{-1})$. In fact, if $t > 0$, then $|t+u| - |u| = t$ for $u \geq 0$ and so

$$\lim_{T\to\infty} \frac{1}{2T} \int_0^T e^{2\pi i(|t+u|-|u|)}\, du = \frac{1}{2} e^{2\pi i t}.$$

Similarly, for $t > 0$, we consider the intervals $[-T, -t)$ and $[-t, 0]$ separately, and note that $|u+t| - |u| = -t$ for $u \in [-T, -t)$; thus

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^0 e^{2\pi i(|t+u|-|u|)}\, du = \frac{1}{2}\, e^{-2\pi i t}.$$

As another example, let

$$f(t) = e^{2\pi i|t|^{1/2}}.$$

Then, expanding

$$\exp 2\pi i \left(\frac{|t+u| - |u|}{|t+u|^{1/2} + |u|^{1/2}} \right),$$

we see that the deterministic autocorrelation of f is $R = 1$, and its power spectrum $S = \delta$.

2.10.7 Example. CURVATURE OF THE CORNU SPIRAL

The *Cornu spiral* C is the curve in the complex plane defined by the Fresnel integrals defined at the beginning of *Section 2.10*, i.e., $C = \{(x(t), y(t)) : t \geq 0\}$, where

$$x(t) = \int_0^t \cos u^2 \, du \text{ and } y(t) = \int_0^t \sin u^2 \, du.$$

The length of C from the origin to $(x, y) \in C$ is easily computed. To compute the curvature $\kappa(t)$ of C at t, we define the vector

$$\mathbf{r}(t) = (x(t), y(t), 0),$$

as well as its velocity $\mathbf{r}^{(1)}(t) = (\cos t^2, \sin t^2, 0)$ and acceleration $\mathbf{r}^{(2)}(t) = (-2t \sin t^2, 2t \cos t^2, 0)$. Then

$$\mathbf{r}^{(1)}(t) \times \mathbf{r}^{(2)}(t) = (0, 0, 2t\cos^2 t^2 + 2t\sin^2 t^2) = (0, 0, 2t),$$

and, hence,

$$\kappa(t) = \frac{\|\mathbf{r}^{(1)}(t) \times \mathbf{r}^{(2)}(t)\|}{\|\mathbf{r}^{(1)}(t)\|^3} = 2t,$$

where $\|\ldots\|$ is the Euclidean norm in $\mathbb{R}^3$.

If the Cornu spiral is replaced by the curve

$$\int_0^t e^{i\varphi(u)} \, du,$$

then the curvature at t of the resulting curve is $\varphi'(t)$.

Chapter 2. Exercises

Exercises 2.1–2.30 are appropriate for *Course I*.

2.1. Compute $\mathbf{1}'_{[-T,T)}$ (the distributional derivative).

2.2. Compute the following distributional derivatives, where H is the Heaviside function.

a. $\left(H(t)\cos t\right)'$.

b. $\left(H(t)\sin t\right)'$.

c. $\left(\mathbf{1}_{[-\pi/2,\pi/2)}(t)\cos t\right)^{(2)}$.

d. g' where $g(t) = \begin{cases} \sin t, & t < 0, \\ 3e^{-t}, & t > 0. \end{cases}$

e. g' where $g(t) = \begin{cases} 0, & \text{if } t \leq 0, \\ t, & \text{if } t > 0. \end{cases}$

2.3. Compute the nth distributional derivative of $g(t) = |t|$, $n = 1, 2, 3$.

2.4. Compute the nth distributional derivative of $g(t) = |\cos t|$, $n = 1, 2, 3$.

2.5. a. Prove that the function f defined in *Example 2.2.2* is infinitely differentiable.

b. Prove that there are no analytic functions in $C_c^\infty(\mathbb{R})\backslash\{0\}$.

2.6. Prove that if $f \in C_c^\infty(\mathbb{R})$ and $g \in L^1(\mathbb{R})$ has compact support then $f * g \in C_c^\infty(\mathbb{R})$.

2.7. Prove (2.3.4) and (2.3.5).

2.8. From *Example 2.2.6b* show that if $g \in L^1_{\text{loc}}(\mathbb{R}) \setminus \{0\}$ then there is $f \in C_c^\infty(\mathbb{R})$ such that

$$\int f(t)g(t)\, dt \neq 0.$$

2.9. Prove that

$$\forall n \geq 0, \quad t^n \delta^{(n)}(t) = (-1)^n n! \delta(t).$$

2.10. a. Show that $\delta(t) = \delta(-t)$.

b. Consider the approximate identity $\{k_\lambda\}$, where $k = \mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]}$. In the spirit of *Example 2.1.6*, evaluate $\delta(t^2)\big(f(t)\big)$, for $f \in C_b(\mathbb{R})$ and $f(0) \neq 0$, by computing

$$\lim_{\lambda\to\infty} \int k_\lambda(t^2) f(t)\, dt.$$

c. Evaluate the limit in part *b* for $f \in C_c^\infty(\mathbb{R})$ for which $f(0) = f'(0) = 0$.

2.11. a. Let $f \in C_c^\infty(\mathbb{R})$ and assume $g \in L^1(\mathbb{R})$ has compact support. Prove that $\operatorname{supp}(f * g) \subseteq \operatorname{supp} f + \operatorname{supp} g$, cf., *Exercise 2.6.*

b. Prove the generalization of part *a* when g is replaced by $T \in D'(\mathbb{R})$.

2.12. Prove that

$$\sum \delta(t-n) + \sum \delta\left(t - n - \frac{1}{2}\right) = 2\sum \delta(2t-n)$$

$$= \left(\sum \delta(t-n)\right) * \left(4\left(M_{\frac{1}{2}}\delta\right)\left(2t - \frac{1}{2}\right)\right),$$

where $M_{\frac{1}{2}}\delta$ is the mean

$$\left(M_{\frac{1}{2}}\delta\right)(t) = \frac{1}{2}\left(\delta\left(t + \frac{1}{2}\right) + \delta\left(t - \frac{1}{2}\right)\right).$$

The middle term arises in wavelet theory.

2.13. a. Compute the "L^2-autocorrelation" of δ, i.e., compute

$$P_\delta(t) = \delta(t) * \overline{\delta(-t)},$$

cf., *Exercises 1.33* and *2.10.* $\overline{\delta}$ and $\overline{\delta(-t)}$ can be computed using the method of *Example 2.1.6.*

b. Compute the L^2-autocorrelation P_Δ of the triangle function Δ defined in *Example 1.3.4.*

c. Compute the L^2-autocorrelation $P_{\Delta_{2\pi/\lambda}}$ of the dilation $\Delta_{2\pi/\lambda}$.

d. In what sense does $\lim_{\lambda \to 0} P_{\Delta_{2\pi/\lambda}} = P_\delta$?

2.14. Let

$$f - \frac{1}{16\pi^4} f^{(4)} = 4\pi^2 \delta + \delta^{(2)}$$

be a differential equation on $\mathbb{R}$. Solve for f using the method of *Theorem 2.6.1.*

2.15. a. Compute $\delta^{(n)} * H$.

b. Compute the Fourier transform of $\delta^{(n)} * H$.

2.16. a. Compute the Fourier transform of $\delta^{(n)} * g$, where $g(t) = tH(t)$ and $n = 1, 2, 3$.

b. Compute the Fourier transform of $\delta^{(n)} * g$, where $g(t) = t^2 H(t)$ and $n = 1, 2$.

2.17. Verify whether or not there are integrable or square integrable solutions of the following differential equations defined on $\mathbb{R}$.

a. $f^{(5)} + f^{(1)} = \delta^{(1)}$.

b. $f^{(5)} - i\, f^{(1)} = \delta^{(4)}$.

2.18. a. Compute $t^2 \delta^{(1)}(t)$, cf., *Exercise 2.9.*

b. Compute $t\delta^{(2)}(t)$.

c. Compute $t^{12}\delta^{(5)}(t)$.

d. Compute $t^5 \delta^{(12)}(t)$.

e. Compute $t^m \delta^{(n)}(t)$ for $m, n \geq 0$.

2.19. Verify that (2.4.17) is a consequence of (2.4.16). The calculation is elementary but involved.

2.20. Let $f_n = \frac{1}{n} f$, where $f \in C_c^\infty(\mathbb{R})$.

a. Prove that $f_n \to 0$ in the sense that *i* and *ii* of *Definition 2.2.3* are satisfied.

b. Let T be any linear functional on $C_c^\infty(\mathbb{R})$. Prove that

$$\lim T(f_n) = 0.$$

c. Show that there is a linear functional on $C_c^\infty(\mathbb{R})$ which is not a distribution.[Hint. Let $g \in C_c^\infty(\mathbb{R})\backslash\{0\}$ and let $g_n = \frac{1}{n}\tau_{1/n}g$. A Fourier transform argument shows that $\{g_n\}$ is linearly independent. There is a basis (in the algebraic sense) of $C_c^\infty(\mathbb{R})$ which contains $\{g_n\}$, e.g., [Tay58, pages 44–45] on Hamel bases. Define $T(g_n) = 1$ for each n and extend T linearly to $C_c^\infty(\mathbb{R})$.]

2.21. Verify that

$$\delta(t^3 + 3t) = \frac{1}{3}\delta(t)$$

and

$$\delta'(t^3 + 3t) = \frac{1}{9}\delta'(t).$$

2.22. a. We evaluated $\int \frac{\sin t}{t}\, dt$ in *Proposition 1.6.3.* Now show that

$$\forall T > 0, \quad \left| \int_{-T}^{T} \frac{\sin t}{t}\, dt - \pi \right| < \frac{\pi}{T}.$$

[Hint. Use Jordan's inequality, which is stated in *Theorem 2.10.1* and *Exercise 2.62a.*]

b. Refine part a by proving that

$$\forall T > 0, \quad \int_{-T}^{T} \frac{\sin t}{t}\, dt - \pi = -\frac{2\cos T}{T} - \frac{2\sin T}{T^2} + \varepsilon_T,$$

where $|\varepsilon_T| \leq 2\pi/T^2$. In particular, if $T_n = \frac{\pi}{2} + \pi n$, then

$$\left| \int_{-T_n}^{T_n} \frac{\sin t}{t}\, dt - \pi \right| \leq \frac{2(1+\pi)}{T_n^2}.$$

2.23. Prove that $BV_{\text{loc}}(\mathbb{R}) \subseteq L^1_{\text{loc}}(\mathbb{R})$. Technically, if $F \in BV_{\text{loc}}((\mathbb{R})$, then it is an ordinary point function, whereas the elements of $L^1_{\text{loc}}(\mathbb{R})$ are sets of ordinary functions which are equal a.e. Show that this situation does not cause any problems in this exercise.

2.24. Verify that

$$L^1(\mathbb{R}) = \left\{F' : F \in BV(\mathbb{R}) \cap AC_{\text{loc}}(\mathbb{R})\right\},$$

where F' designates distributional differentiation.

2.25. Prove (2.2.7) and (2.2.8), cf., *Exercise 2.33.*

2.26. Let $f \longleftrightarrow F$, where $f,\ f' \in L^1(\mathbb{R})$. Verify that

$$(f'(u) * pv(\frac{1}{u}))(t) \longleftrightarrow 2\pi|\gamma|F(\gamma).$$

2.27. a. Compute $\left(1 * \delta^{(n)}\right)^{\wedge}$.

b. Compute $\left(\delta^{(m)} * \delta^{(n)}\right)^{\wedge}$.

2.28. Let $T \in D'(\mathbb{R})$ and let $g \in C^\infty(\mathbb{R})$. Prove that

$$(Tg)^{(n)} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} T^{(k)} g^{(n-k)}.$$

2.29. Consider the distribution

$$(E2.1) \qquad Apv\left(\frac{1}{t}\right) + BH(t) + C\delta(t) + D\delta'(t) + E\log|t|.$$

For each of the following distributions T, list the coefficients in (E2.1) which must be 0. For example, if T is H', then the answer is A, B, D, and E.

a. $\widehat{H}$ b. $(\log|t|)'$ c. $\hat{1}$ d. $[tH(t)]^{(3)}$.

2.30. Let $T \in D'(\mathbb{R})$ and assume $T' = 0$. Prove that T is a constant. [Hint. Let $f_0 \in C_c^\infty(\mathbb{R})$ have the property that $\int f_0(t)\,dt = 1$. First prove that each $f \in C_c^\infty(\mathbb{R})$ has a unique decomposition $f = c_f f_0 + g$ where $c_f = \int f(t)\,dt$ and $g = h'$ for some $h \in C_c^\infty(\mathbb{R})$, e.g., [Sch66, pages 51–52]. Then compute $T(f) = c_f T(f_0) - T'(h) = T(f_0)1(f)$.]

2.31. Using MATLAB, graph the Cornu spiral C defined in *Example 2.10.7.*

2.32. Define

$$\forall \sigma \in (\frac{1}{2}, 1), \quad f_\sigma(t) = \frac{e^{\sigma t}}{1 + \exp e^t}.$$

a. Compute $\widehat{f_\sigma}$ in terms of the Riemann zeta function ζ, cf., [Ben75, pages 137–138]. ζ and the Riemann Hypothesis were defined in *Example 2.4.6g.*

b. Using the notation V_g defined in (2.9.8), prove that the Riemann Hypothesis is true if and only if $V_{f_\sigma} = L^1(\mathbb{R})$ for each $\sigma \in (\frac{1}{2}, 1)$. This elementary observation is due to Salem (1953) [Sal67].

2.33. a. Let $\mu = \sum_{n=1}^\infty \frac{1}{n^2}\, \delta_{1/n}$. Prove that $\mu \in M_d(\mathbb{R})$ and that $\operatorname{supp} \mu = \{0, 1/n : n \in \mathbb{N}\}$.

b. Let $\{r_n\}$ be the subset of the $1/3$-Cantor set $C \subseteq [0,1]$ where each r_n is of the form $k/3^m$, and let $\mu = \sum_{n=1}^\infty \frac{1}{n^2}\, \delta_{r_n}$. Prove that $\mu \in M_d(\mathbb{R})$ and that *supp* $\mu = C$. Recall that C is a closed, uncountable set without isolated points.

c. Let $\{r_n\} = \mathbb{Q}$, the set of rational numbers, and let $\mu = \sum_{n=1}^\infty \frac{1}{n^2}\, \delta_{r_n}$. Prove that $\mu \in M_d(\mathbb{R})$ and that *supp* $\mu = \mathbb{R}$.

2.34. Compute the Fourier transforms of $f(t) = \sin(\pi t) + \cos(\pi t)$ and $f(t) = \sin(\pi t)^2 + \cos(\pi t)^2$.

2.35. Let $n \in \mathbb{N} \cup \{0\}$ and define $T(t) = a\delta^{(n)}(t) + bt^n$ for $a, b \in \mathbb{C} \backslash \{0\}$. For which values of a, b, and n do we have $\widehat{T} = T$, cf., *Example 1.10.12c* and *Exercise 1.28*?

2.36. Let $\mu_n = \frac{1}{2}(\delta_{1/n!} + \delta_{-1/n!})$, $n \geq 2$, and define $\mu = \mu_2 * \mu_3 * \ldots$. Prove that $\|\mu\|_1 = 1$ and $\operatorname{supp} \mu \subseteq [-1, 1]$. Compute $\widehat{\mu}$, cf., Viète's formula in *Exercise 1.36*, where $\prod_{k=1}^{\infty} \cos \frac{\gamma}{2^k} \in L^2(\widehat{\mathbb{R}})$.

The *Wiener-Pitt Theorem* assserts that *if* $\mu = g + \mu_{sc} + \sum a_x \delta_x$, *with notation as in Theorem 2.7.6, satisfies the properties that* $|\widehat{\mu}| > 0$ *on* $\widehat{\mathbb{R}}$ *and*

$$\|\mu_{sc}\|_1 < \inf \left\{ \left| \sum a_x e^{-2\pi i x \gamma} \right| : \gamma \in \widehat{\mathbb{R}} \right\}$$

then there is $\nu \in M_b(\mathbb{R})$ *for which* $\widehat{\nu} = 1/\widehat{\mu}$, e.g., [Ben75, pages 147–149] where the proof uses Kronecker's Theorem (*Exercises 3.40* and *3.41*). John Williamson (International Congress of Mathematics, 1958) used the example of this exercise in analyzing the Wiener-Pitt phenomenon in another setting.

2.37. Consider the convolution equation

$$g - k * g = f \quad \text{on} \quad \mathbb{R},$$

where $k, f \in L^1(\mathbb{R})$ are a given kernel and forcing function, respectively. Assume $|1 - \widehat{k}| > 0$ on $\widehat{\mathbb{R}}$. Prove that there is a solution $g \in L^1(\mathbb{R})$. [Hint. Take the Fourier transform of the equation, invoke the Wiener-Pitt Theorem (stated in *Exercise 2.36*), and apply *Exercise 2.49b*.] There are other proofs.

The *Wiener-Hopf equation*,

$$g(t) - \int_0^{\infty} k(t-u) g(u)\, du = f(t), \quad t \geq 0,$$

is more difficult, but has also been solved, e.g., [Wie49, Appendix C] for applications and solution, cf., Constructive Methods of Wiener-Hopf Factorization, I. Gohberg and M. A. Kaashoek, editors, Birkhauser Verlag, Basel, 1986.

2.38. Let $\{k_{(\lambda)}\}$ be an approximate identity. Prove that

$$\lim_{\lambda \to \infty} k_{(\lambda)} = \delta$$

in the sense that

$$(E2.2) \qquad \forall f \in C_b(\mathbb{R}), \quad \lim_{\lambda\to\infty} \int k_{(\lambda)}(t)f(t)\,dt = f(0).$$

[Hint. Add and subtract $f(0)\int k_{(\lambda)}(u)du$.] If we replace the condition, $\int k_{(\lambda)}(u)du = 1$, of an approximate identity by the condition $\widehat{k}_{(\lambda)}(0) = \int k_{(\lambda)}(u)du = K$ for each λ (while retaining the other properties of an approximate identity), then the conclusion (E2.2) is replaced by

$$\forall f \in C_b(\mathbb{R}), \quad \lim_{\lambda\to\infty} \int k_{(\lambda)}(t)f(t)\,dt = Kf(0).$$

2.39. a. Prove that $\delta \notin L^1_{\text{loc}}(\mathbb{R})$ and that $\delta' \notin M(\mathbb{R})$, cf., *Remark 2.3.7.*

b. Prove that $L^1(\mathbb{R}) \subseteq M_b(\mathbb{R})$ and that $L^1_{\text{loc}}(\mathbb{R}) \subseteq M(\mathbb{R})$, e.g., *Theorem 2.7.6.*

2.40. Let $g \in L^1(\mathbb{R})$, and, for $s > 0$ and $(t,\gamma) \in \mathbb{R}\times\widehat{\mathbb{R}}$, set $g_{s,t,\gamma}(u) = g_s(u-t)e^{2\pi i u\gamma}$ as in *Exercise 1.49.* Prove that

$$\lim_{s\to\infty} g_{s,t,\gamma} = \widehat{g}(0)e^{2\pi i t\gamma}\delta_t$$

in the sense that

$$\forall f \in C_b(\mathbb{R}), \quad \lim_{s\to\infty} \int g_{s,t,\gamma}(u)f(u)\,du = (\widehat{g}(0)e^{2\pi i t\gamma}\tau_t\delta)(f)$$

for each fixed $(t,\gamma) \in \mathbb{R}\times\widehat{\mathbb{R}}$, cf., *Exercise 2.38.*

2.41. We defined $\mathcal{E}'(\mathbb{R})$ in *Example 2.4.6b.* With convolution as the multiplicative operation, $\mathcal{E}'(\mathbb{R})$ is a commutative, associative algebra with unit δ, cf., *Example 2.5.2c* and *Example 2.5.10.* Note that if $h_s(t) \equiv e^{-st}$ for any fixed $s \in \mathbb{C}$, then $h_s : \mathcal{E}'(\mathbb{R}) \longrightarrow \mathbb{C}$ is a homomorphism. Prove that if $h \in C^\infty(\mathbb{R})$ is a homomorphism $\mathcal{E}'(\mathbb{R}) \longrightarrow \mathbb{C}$ then $h(t) = e^{-st}$ for some $s \in \mathbb{C}$.

The algebra $\mathcal{E}'(\mathbb{R})$ and its space of (continuous) homomorphisms $\{h_s : s \in \mathbb{C}\}$ leads to the definition of the bilateral

Laplace transform $T(t)(e^{-st})$ of $T \in \mathcal{E}'(\mathbb{R})$, cf., *Example 2.4.6g*. In fact, as a general point of view, any *algebra* $\mathcal{A}$ and a set $\mathcal{M}$ (for *maximal ideal space*) of its *homomorphisms* $\mathcal{A} \longrightarrow \mathbb{C}$ gives rise to a *transform* $\mathcal{T}$ for which there is an *exchange formula* $\mathcal{T}(A * B) = \mathcal{T}(A)\mathcal{T}(B)$, e.g., *Theorem 2.5.9*. For example, if $\mathcal{A} = L^1(\mathbb{R})$ and $\mathcal{M} = \{e^{-2\pi i t\gamma} : \gamma \in \widehat{R}\}$ then we have Fourier analysis.

2.42. Prove that the Heaviside function H is an unbounded continuous measure.

2.43. Prove that the principal value distribution defined in *Example 2.3.8* is, in fact, a distribution.

2.44. Let $\mu_C \in M_{b+}(\mathbb{R})$ be the Cantor measure corresponding to the 1/3-Cantor set $C \subseteq [0,1]$. Prove that

$$\int t \; d\mu_C(t) = \frac{1}{2},$$

noting that $\int d\mu_C(t) = 1$, cf., *Example 2.3.9b*.

2.45. a. Prove that the mapping $\mathcal{S}(\mathbb{R}) \to \mathcal{S}(\widehat{\mathbb{R}})$, $f \longmapsto \widehat{f}$, is a linear bijection. [Hint. The linear injection follows by a straightforward calculation and the uniqueness theorem. For the surjectivity, let $F \in \mathcal{S}(\widehat{\mathbb{R}})$ and denote $F^\vee$ by g. $g \in \mathcal{S}(\mathbb{R})$ by the first part of the calculation, and the goal is to show that $\widehat{g} = F$. This follows by the inversion theorem for "$\vee$" instead of "$\wedge$", i.e., $F(\gamma) = \int F^\vee(t)e^{-2\pi i t\gamma}dt$ which is the same as $\widehat{g} = F$.]

b. Prove that ρ defined by (2.4.7) is a metric on $\mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R})$, and that, as such, $\mathcal{S}(\mathbb{R})$ is a complete metric space.

c. Prove that the mapping of part a, where $\mathcal{S}$ is given the metrizable topology of part b, is bicontinuous.

d. Prove that "$f_n \to 0$" in the sense of (2.4.8) if and only if $\lim_{n\to\infty} \rho(f_n, 0) = 0$.

2.46. Let $\{T_j\} \subseteq D'(\mathbb{R})$. Define $\lim_{j\to\infty} T_j = T$ for some $T \in D'(\mathbb{R})$ to mean that

$$\forall f \in C_c^\infty(\mathbb{R}), \quad \lim_{j\to\infty} T_j(f) = T(f).$$

Similarly, define $\sum T_j = T$ for some $T \in D'(\mathbb{R})$ to mean that

$$\forall f \in C_c^\infty(\mathbb{R}), \quad \sum T_j(f) = T(f).$$

a. Prove that if $\sum T_j(f)$ exists for each $f \in C_c^\infty(\mathbb{R})$ then there is $T \in D'(\mathbb{R})$ for which $\sum T_j = T$.

b. Prove that if $\sum T_j = T$ for some $T \in D'(\mathbb{R})$, then

$$\forall n \in \mathbb{N}, \quad \left(\sum T_j\right)^{(n)} = \sum T_j^{(n)}.$$

2.47. Note that $\mathcal{S}(\mathbb{R}) \subseteq X(\mathbb{R}) \equiv L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \cap A(\mathbb{R})$. If the norm of $f \in X(\mathbb{R})$ is defined as $\|f\|_{X(\mathbb{R})} = \|f\|_{L^1(\mathbb{R})} + \|f\|_{L^2(\mathbb{R})} + \|f\|_{A(\mathbb{R})}$, where $\|f\|_{A(\mathbb{R})} \equiv \|\hat{f}\|_{L^1(\hat{\mathbb{R}})}$, then $\overline{\mathcal{S}(\mathbb{R})} = X(\mathbb{R})$ and $X(\mathbb{R})^\wedge = X(\hat{\mathbb{R}})$. Thus, since $X'(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R})$, the distributional Fourier transform is a Banach space isomorphism of $X'(\mathbb{R})$ onto $X'(\hat{\mathbb{R}})$. Describe the elements of $X'(\mathbb{R})$.

2.48. Consider the n-fold convolution $g = 1_{[-\Omega,\Omega]} * \cdots * 1_{[-\Omega,\Omega]}$, introduced in terms of splines in *Exercise 1.17.* Compute $g^{(n)}$.

Motivated by the results of this calculation, we introduce the following "usual" definition. A *spline of order* r *on* $\mathbb{R}$ *with knots at the integers* is a function $g \in L^2(\mathbb{R})$, whose restriction to each interval $[n, n+1)$ is a polynomial of degree at most $r-1$, and which is in $C^{r-2}(\mathbb{R})$, i.e., $g \in L^2(\mathbb{R}) \cap C^{r-2}(\mathbb{R})$ and $g^{(r)} = \sum a_n \delta_n$.

2.49. $M_b(\mathbb{R})$ is a Banach algebra with unit δ under convolution, and $L^1(\mathbb{R}) \subseteq M_b(\mathbb{R})$ is a closed ideal. This "algebraic" fact is a consequence of the following exercises.

a. Prove that, for all μ, $\nu \in M_b(\mathbb{R})$, we have $\mu * \nu \in M_b(\mathbb{R})$, and, in fact,

$$\|\mu * \nu\|_1 \leq \|\mu\|_1 \|\nu\|_1.$$

b. We know $L^1(\mathbb{R}) \subseteq M_b(\mathbb{R})$, e.g., *Exercise 2.39b*. Prove that $\|\ldots\|_1$ reduces to $\|\ldots\|_{L^1(\mathbb{R})}$ on $L^1(\mathbb{R})$, and that, for all $\mu \in M_b(\mathbb{R})$, $f \in L^1(\mathbb{R})$, we have $f * \mu \in L^1(\mathbb{R})$, cf., *Exercise 2.54*.

The characterization of the ideal structure of $L^1(\mathbb{R})$ is equivalent to solving the problems of spectral synthesis (for $L^1(\mathbb{R})$) mentioned in *Example 2.4.6f*, e.g., [Ben75]. The characterization of the ideal structure of $M_b(\mathbb{R})$, even just its maximal ideals, is also a deep topic, e.g., [DR71], [Kat76, Chapter 8].

2.50. Let $T \in D'(\mathbb{R})$ and let $h \in C^\infty(\mathbb{R})$ be a strictly monotonic surjection $\mathbb{R} \to \mathbb{R}$. In light of *Example 2.1.6*, *Exercise 2.10*, and *Exercise 2.21*, prove that the *composition* $T \circ h^{-1}$ is an element of $D'(\mathbb{R})$ if it is defined as $(T \circ h^{-1})(f) \equiv T((f \circ h)|h'|)$ for $f \in C_c^\infty(\mathbb{R})$.

2.51. Let $T \in D'(\mathbb{R})$ and let $f \in C_c^\infty(\mathbb{R})$. Prove that $T(\tau_t f) \in C^\infty(\mathbb{R})$ as a function of t.

2.52. An $N \times N$ matrix $A = (a_{jk})$ is *positive semidefinite* if $\bar{c}^T Ac \geq 0$ for every $N \times 1$ matrix $c \in \mathbb{C}^N$, where $\bar{c}$ designates conjugation of each component and T designates transposition. (Unfortunately, "semidefinite" is like "maybe for sure".) Let $P : \widehat{\mathbb{R}} \to \mathbb{C}$ be a function and let $A_P(\gamma_1, \ldots, \gamma_N)$ be the $N \times N$ matrix $\big(P(\gamma_j - \gamma_k)\big)$.

a. Prove that $P >> 0$ if and only if $A_P(\gamma_1, \ldots, \gamma_N)$ is positive semidefinite for each $N \geq 1$ and each set $\{\gamma_1, \ldots, \gamma_N\} \subseteq \widehat{\mathbb{R}}$.

b. Prove the four properties of positive definite functions listed in *Definition 2.7.8b*. [Hint. For part *iii*, let $\gamma_1 = 0$, $\gamma_2 = \lambda$, $\gamma_3 = \gamma$, assume $P(\lambda) \neq P(\gamma)$, and let $c_1 = 1$, $c_2 = -c_3$, and

$$c_2 = \frac{\big|P(\lambda) - P(\gamma)\big|}{P(\lambda) - P(\gamma)} x, \quad x \in \mathbb{R}.$$

Consider the quadratic form (2.7.4) as a polynomial in x, and analyze its discriminant.]

c. Prove that a positive semidefinite matrix $A = (a_{jk})$ is Hermitian, i.e., $a_{jk} = \overline{a}_{kj}$ for all j, k, cf., [Don69, pages 181–182], [Str88] for further properties.

2.53. Suppose $f : \mathbb{R} \to \mathbb{R}$ has the properties that

$$\exists M \text{ such that } \forall a < b, \quad \left| \int_a^b f(t)\, dt \right| \leq M,$$

f is differentiable, and $\|f'\|_{L^\infty(\mathbb{R})} = m$. Prove that $f, \mathcal{H}f \in L^\infty(\mathbb{R})$. This result was alluded to in *Theorem 2.5.12* and is due to Logan [Log83]. [Hint. First show that

$$\lim_{\substack{\varepsilon \to 0 \\ R \to \infty}} \left(\int_{-R}^{-\varepsilon} \frac{f(t)}{t}\, dt + \int_{\varepsilon}^{R} \frac{f(t)}{t}\, dt \right)$$

$$(E2.3) \qquad = \int_{-T}^{T} \left\{ \log \left| \frac{T}{t} \right| - C(T - |t|) \right\} f'(t)\, dt$$

$$+ \int_{|t| \geq T} \frac{F(t)}{t^2}\, dt + (C - \frac{1}{T})\{F(T) + F(-T)\},$$

where $F(t) = \int_0^t f(x)dx$, $T > 0$, and $C \in \mathbb{R}$. We assume everywhere differentiability of f to ensure $f \in AC_{\text{loc}}(\mathbb{R})$, since we use FTC to verify (E2.3). Now estimate the right side of (E2.3), and correctly choose T and C.]

2.54. Apropros *Example 2.5.2*, prove that $\mathcal{S}'(\mathbb{R}) * \mathcal{S}(\mathbb{R})$ is not contained in $L^1(\mathbb{R})$. [Hint. Let $T(u) = u^2$ and let $g(u) = e^{-u^2}$.] It turns out that $\mathcal{S}'(\mathbb{R}) * \mathcal{S}(\mathbb{R}) \subseteq \mathcal{O}_C$, where $f \in \mathcal{O}_C$ means that $f \in C^\infty(\mathbb{R})$ and

$$\exists k = k(f) \in \mathbb{Z} \text{ such that } \forall n \in \mathbb{N},$$

$$\lim_{|t| \to \infty} \left(1 + |t|^2\right)^k \left| f^{(n)}(t) \right| = 0,$$

e.g., [Hor66, pages 420-423].

2.55. We defined the even and odd parts of functions in *Exercise 1.7.* Verify (formally) that $\mathcal{H}f_e = (\mathcal{H}f)_e$ and $(\mathcal{H}f_o) = (\mathcal{H}f)_o$, where $\mathcal{H}$ is the Hilbert transform. Thus, if X is a space of functions, then, formally, $X = X_e \oplus X_o$ and $\mathcal{H}X = (\mathcal{H}X)_e \oplus (\mathcal{H}X)_o$.

2.56. Let f be a causal function, i.e., $\operatorname{supp} f \subseteq [0,\infty)$. Formally define the *unilateral Stieltjes transform* of f as $\mathcal{S}(f)(t) = \int_0^\infty \frac{f(u)}{t+u}\,du$ and the *(unilateral) Laplace transform* of f as $\mathcal{L}(f)(t) = \int_0^\infty f(u)e^{-tu}\,du$, e.g., [Wid41]. Assume that $\mathcal{L}f$ and $\mathcal{L}\mathcal{L}f$ exist on $(0,\infty)$. Show that

$$\mathcal{L}\mathcal{L}f(u) = \mathcal{S}f(u) = -\pi\mathcal{H}f(-u).$$

2.57. a. Verify (formally) the Fourier transform pairings $2fH \leftrightarrow \hat{f} - i\mathcal{H}\hat{f}$ and $f + i\mathcal{H}f \leftrightarrow 2\hat{f}H$, where H is the Heaviside function and $\mathcal{H}$ is the Hilbert transform operator.

b. Verify (formally) that if f is causal, then $\operatorname{Re}\hat{f} = \mathcal{H}\operatorname{Im}\hat{f}$ and $\operatorname{Im}\hat{f} = -\mathcal{H}\operatorname{Re}\hat{f}$.

Parts *a* and *b* give elementary relations between the Fourier and Hilbert transforms. The situation can become more complex (sic). Recall the Paley-Wiener Logarithmic Integral Theorem [PW34] from *Example 1.6.5.* This result asserts the existence of a causal function $f \in L^2(\mathbb{R})$ for which $|\hat{f}| = \phi$ a.e. in the case

$$\int \frac{|\log\phi(\gamma)|}{1+\gamma^2}\,d\gamma < \infty$$

for a given non-negative function $\phi \in L^2(\mathbb{R})$. Formally

$$\hat{f} = \phi e^{-i\mathcal{H}\phi}$$

is a candidate for such a function f, cf., [BT93] for relevant calculations and a wavelet application to speech compression, and consider $\varphi(\gamma) = 1/(1+\gamma^2)$ for a dose of mathematical reality.

2.58. a. Prove that (2.6.7) defines a translation-invariant continuous linear operator $L : L^1(\mathbb{R}) \to L^1(\mathbb{R})$.

b. Let $L : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ be a translation-invariant continuous linear operator, and let $g \in L^\infty(\mathbb{R})$. Prove that there is an $h \in L^\infty(\mathbb{R})$, depending on L and g, such that

$$\forall f \in L^1(\mathbb{R}), \quad \int (Lf)(t)g(t)\,dt = \int f(t)h(t)\,dt.$$

c. Use part *b* to prove that

$$\forall f, k \in L^1(\mathbb{R}), \quad (Lf) * k = L(f * k) = f * L(k).$$

[Hint. Calculate

$$\int (Lf) * k(t)g(t)\ dt = \int k(s) \int (\tau_s f)(t)h(t)\,dt\,ds$$
$$= \int h(t) f * k(t)\ dt = \int L(f * k)(t)g(t)\ dt.]$$

2.59. a. Prove that $\mathcal{H}(\cos 2\pi t\gamma_o) = \sin 2\pi t\gamma_o$ and $\mathcal{H}(\sin 2\pi t\gamma_o) = -\cos 2\pi t\gamma_o$, $\gamma_o > 0$.

b. Let $f \in L^2(\mathbb{R})$ be real-valued. Prove that $\int \mathcal{H}f(t)\overline{f(t)}\ dt = 0$.

c. Prove the symmetry condition (2.6.7) of *Example 2.6.7a*. [Hint. The straightforward calculation depends on the fact that h, described in *Example 2.6.7a*, is real-valued.]

2.60. With regard to the problem of finding $F \in C_c(\widehat{\mathbb{R}})$ for which $F \notin B(\widehat{\mathbb{R}})$, prove that

$$\forall \mu \in M_b(\mathbb{R}), \quad \exists \lim_{\substack{\varepsilon \to 0 \\ R \to \infty}} \int_{\varepsilon \le |\gamma| \le R} \frac{\widehat{\mu}(\gamma)}{\gamma}\,d\gamma,$$

cf., *Exercise 2.19.*

2.61. Prove *Theorem 2.9.4.*

2.62. a. Prove Jordan's Inequality,

$$\forall r > 0, \quad \int_0^{\pi/2} e^{-r\sin\theta}\, d\theta < \frac{\pi}{2r}.$$

b. In light of part *a*, prove that

$$\forall \theta \in (0, \pi/4), \quad (\sin\theta)^{\sin\theta} < (\cos\theta)^{\cos\theta},$$

e.g., Amer. Math. Monthly Problems, 101 (1994), 690.

2.63. The following inequalities are van der Corput's Lemma.

a. Let ϕ be a real, differentiable function on $[a, b]$ for which ϕ' is monotone and for which $|\phi'| \geq r > 0$. Prove that

$$\left| \int_a^b e^{i\phi(t)}\, dt \right| \leq \frac{4}{r}.$$

[Hint. Write the integral in terms of its real and imaginary parts. For each case multiply and divide by ϕ', and use the first mean value theorem for integrals.]

b. Let ϕ be a real, twice differentiable function on $[a, b]$ for which $|\phi^{(2)}| \geq r > 0$. Prove that

$$\left| \int_a^b e^{i\phi(t)}\, dt \right| \leq \frac{8}{\sqrt{r}}.$$

[Hint. ϕ' vanishes at most once in (a, b). If this point is c, write $\int_a^b$ as $\int_a^{c-\epsilon} + \int_{c-\epsilon}^{c+\epsilon} + \int_{c+\epsilon}^b$. The proper choice of ϵ and an application of part *a* on the first and third integrals yield the result.]

It is often important to compute or estimate exponential sums, e.g., the Gauss sums of *Section 3.8*. Finite trigonometric sums $\sum e^{2\pi i\varphi(n)}$ arise in analytic number theory in the process of estimating the growth of the Riemann zeta function. A classical method due to van der Corput is to write

$$(E2.4) \quad \sum_{a<n\leq b} e^{2\pi i\varphi(n)} \approx \sum_{\varphi'(b)-\eta<m\leq\varphi'(a)+\eta} \int_a^b e^{2\pi i(\varphi(t)-mt)}\, dt,$$

and to estimate the right side using refinements of parts *a* and *b*, e.g., [Ivi85, Chapter 2], [Tit51, Chapter 4]. Then it is possible to show that

$$\exists C > 0 \text{ such that } |\zeta(1/2 + iy)| \leq C|y|^{1/6}$$

for large $|y|$. Bombieri and Iwaniec proved a slightly better result using modular forms. In any case, the right side of (E2.4) is an approximate form of the Poisson Summation Formula, cf., *Section 3.10.*

2.64. Complete the details in the proof of *Theorem 2.10.3.*

2.65. The function $s(t) = \frac{1}{\sqrt{\pi i}} e^{it^2}$ of *Section 2.10* is in $L^\infty(\mathbb{R}) \cap A'(\mathbb{R}) \subseteq M(\mathbb{R})$, whereas $s \notin L^1(\mathbb{R})$. Verify whether or not s is a bounded Radon measure.

2.66. Compute the *deterministic cross-correlation*,

$$R_{\gamma,\lambda}(t) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} s_\gamma(t+u)\overline{s_\lambda(u)}\,du, \quad \gamma, \lambda > 0,$$

where s_γ is the dilation of the function s defined in *Exercise 2.65*, cf., *Example 2.10.6* for the case $\gamma = \lambda$.

2.67. Let $X(\mathbb{R}) = L^\infty(\mathbb{R}) \cap A'(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R})$. Note that $X(\mathbb{R})^\wedge = X(\widehat{\mathbb{R}})$, and that $X(\mathbb{R})$ is a Banach space, where the norm of $f \in X(\mathbb{R})$ is defined as $\|f\|_{X(\mathbb{R})} = \|f\|_{L^\infty(\mathbb{R})} + \|\hat{f}\|_{L^\infty(\widehat{\mathbb{R}})}$, cf., *Exercise 2.47.* Prove that $\sin \pi t^2$ and $\cos \pi t^2$ are eigenfunctions of the Fourier transform mapping $\mathcal{F} : X(\mathbb{R}) \to X(\widehat{\mathbb{R}})$.

2.68. a. Verify that if $c \in \mathbb{R}\backslash\{0\}$ then

$$\sup_{a<b} \left| \int_a^b e^{ict^2}\,dt \right| \leq 2 + \frac{4}{|c|},$$

e.g., *Exercise 2.63.* The bound can be refined.

b. Let $f \in AC_{\text{loc}}([0,\infty))$ be positive and decreasing. (Recall that f is not necessarily in $AC[a,b]$ if it is decreasing on $[a,b]$.) Prove that if $c > 0$ then

$$\left| \int_0^\infty f(t) e^{\pi i (ct)^2} e^{-2\pi i t \gamma} \, dt \right| \leq \frac{1}{c}(2 + \frac{4}{\pi}) f(0).$$

Such estimates are used in the *Littlewood Flatness Problem* discussed in *Remark 3.8.11.*