Chapter 2

Fourier Series and Integrals

> Having contested the various results [Biot and Poisson] now recognise that they are exact but they protest that they have invented another method of expounding them and that this method is excellent and the true one. If they had illuminated this branch of physics by important and general views and had greatly perfected the analysis of partial differential equations, if they had established a principal element of the theory of heat by fine experiments ... they would have the right to judge my work and to correct it. I would submit with much pleasure ... But one does not extend the bounds of science by presenting, in a form said to be different, results which one has not found oneself and, above all, by forestalling the true author in publication.
>
> Joseph Fourier, from John Herivel, *Joseph Fourier: The Man and the Physicist*, 1992

It is often useful to decompose a given function into components, analyze them, and then reassemble the function again, possibly in a different way. One classical and mathematically very interesting method is to use trigonometric functions. This is the basis for the theory of Fourier analysis.

One can think of a sound (a certain tone played on the violin, say) as consisting of countably many oscillations with different discrete frequencies, which together define the pitch and the specific timbre of the tone. These component frequencies can be identified via Fourier analysis, in particular by computing the Fourier series of a periodic function. Of course, in reality no oscillation is precisely periodic, and a sound will consist of a continuum of frequencies. Mathematically, this is analyzed by taking the continuous Fourier transform. Thus, the Fourier transform arises from Fourier series by taking more and more frequencies into account, a process described by the Poisson summation formula. Finally, the question arises as to whether a function thus analyzed can be reconstructed from its Fourier series. It turns out that even for a continuous function, the Fourier series may not be everywhere convergent. Thus one is led to consider special summation methods, or "kernels."

These are among the topics covered in the next few sections. The interested reader can find more details than we have room to give here in Chapter 8 of Stromberg's book [203], in the classical treatment available in Zygmund [217], or in the more modern books by Katznelson [149] and Butzer-Nessel [75].

One reason for studying Fourier methods is that they are very useful for evaluation of sums and integrals. Once such an object is identified as a Fourier series or integral, that knowledge can be used for the evaluation, or methods such as the Parseval equation or Poisson summation can be brought to bear. We will see numerous examples of this in the present chapter and elsewhere in the book.

2.1 The Development of Fourier Analysis

We start with some historical background here, which we have adapted in part from the MacTutor web site at http://www-gap.dcs.st-and.ac.uk/~history, and also from R. Bhatia's monograph on Fourier series [25].

Joseph Fourier was one of the more colorful figures of mathematical history. Originally intending to be a Catholic priest, Fourier declined to take his vows when he realized that he could not extinguish his interest in mathematics. Shortly afterwards, he became involved in the movement that led to the French Revolution in 1793, but, fortunately for modern mathematics, he was spared the guillotine, and was able to study mathematics at the Ecole Normale in Paris under the tutelage of Lagrange. A few years later, he was appointed as a scientific adviser for Napoleon's expedition to Egypt. When Napoleon's army was defeated by Nelson at the battle of the Nile, Fourier and the other French advisers insisted that they be able to retain some of the artifacts they had found there. The British refused, but at least permitted the French to make a catalog of what they felt were the more important items. Fourier was given this task by the French commanders. The eighth item of his catalog was the Rosetta stone, which had been recognized by the French scientists on the expedition as a possible key to understanding the Egyptian language. Later in Europe, when published copies of the inscriptions were made available, Champollion, a student who had been inspired by Fourier himself to study Egyptology, succeeded in the first translation.

Fourier's principal contributions to mathematics, namely Fourier analysis and Fourier series, paralleled and even stimulated the development of the entire field of real analysis. Fourier analysis had its origin in the 1700s, when d'Alembert derived the wave equation that describes the motion of a vibrating string, starting with an initial "function," which at the time was restricted to an analytic expression. In 1755, Daniel Bernoulli gave another solution for the problem in terms of *standing waves*, namely waves associated with the $n+1$ points $0, 1/n, 2/n, \cdots, (n-1)/n, 1$ on the string that remain fixed. The motion for $n = 1, 2, \cdots$ is the *first harmonic*, the *second harmonic*, and so on. Bernoulli asserted that every solution to the problem of the plucked string is merely a sum of these harmonics.

Beginning in 1804, Joseph Fourier began to analyze the conduction of heat in solids. He not only discovered the basic equations governing heat conduction, but he developed methods to solve them, and, in the process, developed and extended Fourier analysis to a much broader range of scientific problems. He described this work in his book *The Analytical Theory of Heat*, which is regarded as one of the most important books in the history of physics.

Like Bernoulli, Fourier asserted that any continuous function can be written as

$$f(x) \;=\; \sum_{n=-\infty}^{\infty} A_n\, e^{int}. \tag{2.1.1}$$

But Fourier claimed that his representation is valid not only for f given by a single analytical formula, but for f given by any *graph*, which at the time was a more general object, encompassing, for example, a piecewise combination of different analytic expressions. Fourier was not able prove his assertions, at least not to the satisfaction of the mathematical community at the time (and certainly not to the standards required today). But other mathematicians were intrigued, and they pursued these questions with renewed determination.

Dirichlet was the first to find a rigorous proof. He defined a real function as we now understand the term, namely as a general mapping from one set of reals to another, thus decoupling analysis from geometry. He then was able to prove that for every "piecewise smooth" function f, the Fourier series of f converges to $f(x)$ at any point x where f is continuous, and to the average value $(f(x+)+f(x-))/2$ if f has a jump discontinuity at x. This was the first major convergence result for Fourier series.

Mathematicians realized that to handle functions that have infinitely many discontinuities, it was necessary to generalize the notion of an integral beyond the intuitive idea of the area under a curve. Riemann succeeded in developing a theory of integration that could handle such functions, and using this theory, he was able to exhibit an example of a function that did not satisfy Dirichlet's piecewise continuous condition, yet still possessed a pointwise convergent Fourier series. Cantor observed that changing a function f at a few points does not change its Fourier coefficients. In the course of asking how many points can be changed while preserving Fourier coefficients, he was led to the notion of countably infinite and uncountably infinite sets. Ultimately, Lebesgue extended Dirichlet's, Riemann's, and Cantor's results into what we now know as measure theory, where sets of measure zero, almost everywhere equality of functions, and almost everywhere convergence of functions supersede the simple concepts that prevailed in the 1700s.

In summary, it is not an exaggeration to say that all of modern real and complex analysis has its roots in Fourier series and Fourier analysis.

2.2 Basic Theorems of Fourier Analysis

2.2.1 Fourier Series

We will consider 2π-periodic functions $f : \mathrm{R} \to \mathrm{C}$. For $p > 0$, we write $f \in L^p(\mathrm{T})$ if such an f is Lebesgue measurable and satisfies

$$\|f\|_p \;=\; \left[\int_{-\pi}^{\pi} |f(t)|^p \, dt\right]^{1/p} \;<\; \infty$$

(T stands for *Torus*). Note that $\|\cdot\|_p$ is not a norm on L^p, since any function f which is 0 almost everywhere will have $\|f\|_p = 0$. (Later we will identify functions which are equal a.e.) In what follows, we will be mainly interested in the spaces $L^1(\mathrm{T})$ and $L^2(\mathrm{T})$. Note that $\|\cdot\|_1 \le \|\cdot\|_2$ and $L^2(\mathrm{T}) \subsetneq L^1(\mathrm{T})$. If f is 2π-periodic and continuous on R, then we write $f \in C(\mathrm{T})$ and equip this space with the uniform norm.

For a function $f \in L^1(\mathrm{T})$, define the n-th Fourier coefficient ($n \in \mathrm{Z}$) by

$$\widehat{f}_n = \int_{-\pi}^{\pi} f(t)\, e^{-int}\, dt.$$

This is motivated by the insight that if we write

$$s_n(t) = \frac{1}{2\pi} \sum_{k=-n}^{n} c_k\, e^{ikt}, \tag{2.2.2}$$

for some sequence c_k and assume $L^1(\mathrm{T})$-convergence of (s_n) to some $f \in L^1(\mathrm{T})$, then $\widehat{f}_n = c_n$. Thus, for arbitrary $f \in L^1(\mathrm{T})$, we will write (formally and suggestively)

$$f(t) \sim \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}_n\, e^{int},$$

where in general no assertion about convergence of this series is implied. Any convergence statement is to be read in the sense of symmetric limits, i.e., of

$$s_n(f,t) = \frac{1}{2\pi} \sum_{k=-n}^{n} \widehat{f}_k\, e^{ikt}.$$

Fourier coefficients usually are complex numbers, even when f is a real-valued function. Sometimes it is desirable to have a real-valued series for f. Then the Fourier series can be equivalently written as

$$f(t) \sim \frac{1}{2\pi} a_0 + \frac{1}{\pi} \sum_{n=1}^{\infty} (a_n \cos(nt) + b_n \sin(nt)),$$

where

$$\begin{aligned} a_n &= \int_{-\pi}^{\pi} f(t)\, \cos(nt)\, dt \\ b_n &= \int_{-\pi}^{\pi} f(t)\, \sin(nt)\, dt. \end{aligned}$$

Note that

$$s_n(f,t) = \frac{1}{2\pi} a_0 + \frac{1}{\pi} \sum_{k=1}^{n} (a_k \cos(kt) + b_k \sin(kt)).$$

Example 2.2.1 *Fourier series of a symmetric function.*

Define $f \in L^1(\mathrm{T})$ by $f(t) = (\pi - t)/2$ for $t \in [0, 2\pi)$. Then $\widehat{f}_n = -i\pi/n$ for $n \neq 0$ and $\widehat{f}_0 = 0$, and its Fourier series is given by

$$f(t) \sim \frac{-i}{2} \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{1}{n} e^{int}, \tag{2.2.3}$$

or, equivalently,

$$f(t) \sim \sum_{n=1}^{\infty} \frac{\sin(nt)}{n}. \tag{2.2.4}$$

The right-hand side of (2.2.4) equals 0 at $t = 0$, while the left-hand side by definition equals $\pi/2$. Thus, equality cannot hold pointwise here. The situation would improve if we were to define $f(0) = 0$. In fact, it follows from

$$\sum_{n=1}^{\infty} \frac{z^n}{n} = -\mathrm{Ln}(1 - z) \quad \text{for } |z| \leq 1,\ z \neq 1, \tag{2.2.5}$$

by setting $z = e^{it}$ and taking real and imaginary parts, that

$$\sum_{n=1}^{\infty} \frac{\sin(nt)}{n} = \frac{\pi - t}{2} \quad \text{and} \tag{2.2.6}$$

$$\sum_{n=1}^{\infty} \frac{\cos(nt)}{n} = -\ln|2\sin(t/2)|, \tag{2.2.7}$$

for all $t \in (0, 2\pi)$, and even with uniform convergence on every closed subinterval of $(0, 2\pi)$. □

The question now is under what condition the Fourier series of a function converges to that function. The answer depends on the definition of "convergence," and is most interesting in the cases of pointwise, $L^1(\mathrm{T})$- and $L^2(\mathrm{T})$-convergence.

Pointwise and uniform convergence. As we have seen in the above example (and as is clear from the computation of Fourier series), $L^1(\mathrm{T})$-functions which are equal a.e. will have the same Fourier series. By the *uniqueness theorem for Fourier series*, the converse is also true: Functions with the same Fourier series are equal a.e. It is *not* true that the Fourier series of any continuous function is pointwise convergent to that function. An example due to Lebesgue is given in Item 6 at the end this chapter. Such functions must have a complicated structure, because they cannot be of bounded variation. A function $f : \mathrm{R} \to \mathrm{C}$ is of bounded variation on (a, b) if there is an $M > 0$ with

$$\sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)| \ \leq \ M \quad \text{for all } a < t_0 < t_1 < \cdots < t_n < b.$$

The infimum of all such M is the total variation of f, thus

$$V_a^b(f) \ = \ \sup \left\{ \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)| : a < t_0 < t_1 < \cdots < t_n < b, \ n \in \mathrm{N} \right\}.$$

The set of all functions of bounded variation on (a, b) is denoted $BV(a, b)$. An equivalent characterization (due to Lebesgue) is that f can be written as the difference of two bounded increasing functions. Thus, any BV-function f is differentiable a.e., and the one-sided limits $f(t+)$ and $f(t-)$ exist in (a, b).

Theorem 2.2.2 (Jordan test). *For $f \in L^1(\mathrm{T}) \cap BV(a, b)$ we have*

$$\lim_{n\to\infty} s_n(f, t) \ = \ \frac{f(t-) + f(t+)}{2} \quad \text{for every } t \in (a, b)$$

and uniformly on every compact subinterval of (a, b) where f is continuous. If $f \in C(\mathrm{T})$, then the convergence is uniform on R.

Note that this theorem is proved via Cesàro summation, and thus via the Fejér kernel, which we will discuss in Section 2.3.3.

The Jordan test is another explanation of the convergence properties of the Fourier series for the function f from the previous example. If another function $f \in C(\mathrm{T})$ is defined by $f(t) = t^2$ on $[-\pi, \pi]$, then

$$f(t) \ \sim \ \frac{1}{3}\pi^2 + 4 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(nt).$$

Since f is continuous on R, the Fourier series converges uniformly to f, by the Jordan test. However, since $s_n(f, t)$ is uniformly convergent, this also follows directly from the uniqueness theorem for Fourier series.

For $t = 0$, we get from this the identity

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2} = \frac{\pi^2}{12}.$$

By separating even and odd parts, this proves that

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

To conclude the section on pointwise convergence, we note that the Fourier series of an $f \in L^1(\mathrm{T})$ may be divergent everywhere (by an example due to Kolmogorov in 1926). If, however, $f \in L^p(\mathrm{T})$ for $p > 1$, then $s_n(f, t)$ converges to $f(t)$ a.e., a result proved by Carleson and Hunt in 1966 [81, 143].

Convergence in $L^1(\mathrm{T})$*.* From now on it makes sense to identify functions that are equal almost everywhere, so that $\|\cdot\|_1$ and $\|\cdot\|_2$ are now norms in the respective spaces. Thus we will deal, strictly speaking, with equivalence classes of functions instead of pointwise defined functions, although this will not be denoted explicitly. For example, the statement that such a function is continuous will mean that in its equivalence class we can find a continuous function, then denoted by the same symbol. This notation is unusual but convenient, and it corresponds to how one deals with these objects informally. It can lead to dangerous pitfalls, though, as we will see later. In general, the Fourier series of an $f \in L^1(\mathrm{T})$ need not be convergent to f with respect to the $L^1(\mathrm{T})$-norm. Of the many restrictions on f which imply convergence, we mention here only one which is particularly simple (and has a convincing analog in the case of Fourier transforms).

Theorem 2.2.3 *Let* $f \in L^1(\mathrm{T})$*. If* $\sum_{n=-\infty}^{\infty} |\widehat{f}_n| < \infty$*, then* $f \in C(\mathrm{T})$ *and*

$$f(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}_n \, e^{int},$$

with convergence in the $L^1(\mathrm{T})$*-norm as well as uniformly on* T*.*

There are functions in $L^1(\mathrm{T})$ (even continuous ones) for which the Fourier coefficients are not absolutely summable. It is a difficult and in general open problem to characterize those sequences which are Fourier sequences of an $L^1(\mathrm{T})$-function. A weak necessary condition follows from the Riemann-Lebesgue lemma (see next subsection and Exercise 14): The Fourier coefficients of any $f \in L^1(\mathrm{T})$ satisfy $\lim_{|n|\to\infty} \widehat{f}_n = 0$.

Convergence in $L^2(\mathrm{T})$. In contrast to the L^1-case, for $p > 1$ the Fourier series of any $f \in L^p(\mathrm{T})$ converges to f in the $L^p(\mathrm{T})$-norm. For $p = 2$, this follows directly from the usual Hilbert space theory: The trigonometric functions constitute an orthogonal basis for $L^2(\mathrm{T})$. This implies that the Fourier coefficients of an $f \in L^2(\mathrm{T})$ are square-summable and that every square-summable sequence is the sequence of Fourier coefficients of an $f \in L^2(\mathrm{T})$. (That is the Riesz-Fischer theorem.) Another consequence of Hilbert space theory is the Parseval equation.

Theorem 2.2.4 (Parseval's formula). *For any $f, g \in L^2(\mathrm{T})$, the identity*

$$\frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}_n \, \overline{\widehat{g}_n} \;=\; \int_{-\pi}^{\pi} f(t)\, \overline{g(t)}\, dt \tag{2.2.8}$$

holds. In particular, for $f = g$ we get $\frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left|\widehat{f}_n\right|^2 = \int_{-\pi}^{\pi} |f(t)|^2 \, dt$.

Example 2.2.5 *Parseval's formula and the zeta function.*

Applying the Parseval equation to the function $f(t) = (\pi - t)/2$ on $(0, 2\pi)$ from the first example in this section again gives the identity, after simplifying,

$$\zeta(2) \;=\; \sum_{n=1}^{\infty} \frac{1}{n^2} \;=\; \frac{\pi^2}{6}.$$

This is the same result as in the previous example, but the Parseval equation is conceptually simpler than the Jordan test. Applying the Parseval equation to $f(t) = t^2/4 - \pi t/2 + \pi^2/6$ gives

$$\zeta(4) \;=\; \sum_{n=1}^{\infty} \frac{1}{n^4} \;=\; \frac{\pi^4}{90},$$

and this can be continued to give $\zeta(2n)$ as a rational multiple of π^{2n} for any $n \in \mathrm{N}$. The general formula is given in the next chapter, by a different method. Similar formulas for $\zeta(2n+1)$ are unknown (and highly unlikely to exist; see the next chapter for more information). □

Example 2.2.6 *Parseval's formula and Euler sums.*

Multiplying (2.2.6) and (2.2.7), using the Cauchy product, simplifying, and performing a partial fraction decomposition gives

$$-\ln|2\sin(t/2)| \cdot \frac{\pi - t}{2} \;=\; \sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=1}^{n-1} \frac{1}{k} \sin(nt) \quad \text{on } (0, 2\pi). \tag{2.2.9}$$

Now Parseval proves the Euler sum formula that we first mentioned in the first volume:

$$\frac{1}{4}\int_0^{2\pi} (\pi - t)^2 \ln^2(2\sin(t/2))\, dt \;=\; \pi \sum_{n=1}^{\infty} \frac{\left(\sum_{k=1}^{n-1} k^{-1}\right)^2}{n^2}.$$

□

2.2.2 Fourier Transforms

We now consider functions $f : \mathrm{R} \to \mathrm{C}$. For $p > 0$, we write $f \in L^p(\mathrm{R})$ if such an f is Lebesgue measurable and satisfies

$$\|f\|_p \;=\; \left[\int_{-\infty}^{\infty} |f(t)|^p \, dt\right]^{1/p} \;<\; \infty.$$

As before, functions which are equal a.e. will be identified, so that $\|\cdot\|_p$ is a norm. In what follows, we will be mainly interested in the spaces $L^1(\mathrm{R})$ and $L^2(\mathrm{R})$. In contrast to the periodic case, there is now no inclusion relation between these spaces. If $f : \mathrm{R} \to \mathrm{C}$ is continuous, we write $f \in C(\mathrm{R})$ and equip this space with the uniform norm.

The Fourier transform on $L^1(\mathrm{R})$ is now a direct analog of the Fourier coefficients on $L^1(\mathrm{T})$. By further analogy to the previous subsection, a Fourier transform on $L^2(\mathrm{R})$ would also be of interest. However, the definition of such an $L^2(\mathrm{R})$-transform is not as straightforward as before, since these spaces are not contained in each other. There is, however, a meaningful transform on $L^2(\mathrm{R})$, and we will discuss this after giving the properties of the $L^1(\mathrm{R})$-transform.

Fourier Transform on $L^1(\mathrm{R})$. For a function $f \in L^1(\mathrm{R})$, define the Fourier transform of f to be the function $\widehat{f} : \mathrm{R} \to \mathrm{C}$ given by

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

As the example $f = I_{(-\pi,\pi)}$ (characteristic function) with $\widehat{f}(x) = 2\sin(\pi x)/\pi$ shows, the Fourier transform of an $f \in L^1(\mathrm{R})$ need not be in $L^1(\mathrm{R})$. It is not difficult to show, however, that such an $\widehat{f}$ is always continuous, with $\|\widehat{f}\|_\infty \leq \|f\|_1$. The *Riemann-Lebesgue lemma* says that, additionally, $\lim_{|x|\to\infty} \widehat{f}(x) = 0$. It is proved by approximating f by step functions.

Under what conditions can an $f \in L^1(\mathrm{R})$ be reconstructed from its Fourier transform? In principle, this is always possible: By the *uniqueness theorem for Fourier transforms*, functions with the same Fourier transform must be equal a.e. In practice, one would like a simple formula for this inversion. Such a formula is given in the inversion theorem below, whose proof is not easy: It depends on constructing and investigating a suitable summation kernel for the Fourier transform (often the Gauss or the Fejér kernel is used).

Theorem 2.2.7 *If $f \in L^1(\mathrm{R})$ is such that $\widehat{f} \in L^1(\mathrm{R})$, then $f, \widehat{f} \in C(\mathrm{R})$ and*

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(x)\, e^{ixt}\, dx \tag{2.2.10}$$

for all $t \in \mathrm{R}$.

Conditions which are good for $\widehat{f} \in L^1(\mathrm{R})$ are given in Item 10 at the end of this chapter.

Example 2.2.8 *Sinc integrals.*

For $f(t) = \max\{1 - |t|, 0\}$, we compute $\widehat{f}(x) = \operatorname{sinc}^2(x/2) \in L^1(\mathrm{R})$, where $\operatorname{sinc} x = (\sin x)/x$. Thus by Theorem 2.2.7, we immediately get

$$\int_{-\infty}^{\infty} \operatorname{sinc}^2(x/2)\, e^{ixt}\, dx = 2\pi \max\{1 - |t|, 0\} \tag{2.2.11}$$

for all t, and especially $\int_{-\infty}^{\infty} \operatorname{sinc}^2(x/2)\, dx = 2\pi$. The integral transform package of *Maple* knows this integral:

```
inttrans[fourier]((sin(x/2)/(x/2))^2,x,t);
```

returns

```
2 Pi (-2 t Heaviside(t) + t Heaviside(t - 1) - Heaviside(t - 1)
   + t Heaviside(t + 1) + Heaviside(t + 1))
```

This is an inelegant form of the answer we gave. Note that newer versions of *Maple* can evaluate the integral directly as well, while in previous versions this was only possible with the use of the Fourier transform, so that researchers required more knowledge about what they were doing. Thus this is another example of being able to watch technology in progress! □

Fourier Transform on $L^2(\mathrm{R})$. Since $L^2(\mathrm{R})$ is not a subset of $L^1(\mathrm{R})$, in contrast to the periodic case, the definition of the Fourier transform cannot be directly transferred onto this space: The function $f(t)\, e^{-ixt}$ may not be integrable! For this reason, the Fourier transform on $L^2(\mathrm{R})$ is usually defined as the continuation of the Fourier transform on $L^1(\mathrm{R})\cap L^2(\mathrm{R})$ which is dense in $L^2(\mathrm{R})$. An equivalent (and more practical) definition is the following. First of all, note that for $f \in L^2(\mathrm{R})$, $f_A = f \cdot \chi_{(-A,A)} \in L^1(\mathrm{R})$ for any $A > 0$. It can be proved, with some effort, that then $\widehat{f_A} \in L^2(\mathrm{R})$ and that the $L^2(\mathrm{R})$-limit of the functions $\widehat{f_A}$ for $A \to \infty$ exists in $L^2(\mathrm{R})$.

Definition 2.2.9 *Let $f \in L^2(\mathrm{R})$. The $L^2(\mathrm{R})$-limit of the functions $\widehat{f_A}$ for $A \to \infty$ is called the Fourier (or Plancherel) transform of f and is again denoted by $\widehat{f}$.*

By definition, the Fourier transform of an $f \in L^2(\mathrm{R})$ is always in $L^2(\mathrm{R})$. It need not be continuous, nor does the Riemann/Lebesgue lemma hold. It can be proved that $\|f\|_2 = \|\widehat{f}\|_2$ *(Parseval equation)*, and that every function in $L^2(\mathrm{R})$ is the Fourier transform of an $f \in L^2(\mathrm{R})$. An $f \in L^2(\mathrm{R})$ is reconstructible from its Fourier transform by the same process as in Theorem 2.2.7.

Theorem 2.2.10 *For any $f \in L^2(\mathrm{R})$,*

$$f(t) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{f}(x)\, e^{ixt}\, dx \;=\; \lim_{A\to\infty} \frac{1}{2\pi}\int_{-A}^{A} \widehat{f}(x)\, e^{ixt}\, dx \tag{2.2.12}$$

in $L^2(\mathrm{R})$.

Example 2.2.11 *Fourier transform of sine-exponential.*

We have already computed $\widehat{f}(x) = 2\sin(\pi x)/x$ for $f(t) = \chi_{(-\pi,\pi)}(t)$. This $\widehat{f}$ is in $L^2(\mathrm{R})$, but not in $L^1(\mathrm{R})$. Theorem 2.2.10 now says that

$$\int_{-\infty}^{\infty} \frac{\sin(\pi x)}{\pi x}\, e^{ixt}\, dx \;=\; \chi_{(-\pi,\pi)}(t) \quad \text{a.e.} \qquad \square$$

It would be interesting to replace this a.e. statement with a pointwise statement, since a.e. statements are more difficult to handle when the goal is exact evaluation of a series or integral, and so they are less useful for experimental mathematics. In fact, a standard theorem of Fourier analysis (Jordan's theorem, see [75, page 205]) is: *If $f \in L^1(\mathrm{R})$ is of bounded variation in an interval including the point t, then*

$$\lim_{a\to\infty} \frac{1}{2\pi} \int_{-a}^{a} \widehat{f}(x)\, e^{ixt}\, dx \;=\; \frac{1}{2}\left(f(t+) + f(t-)\right).$$

Applied to the above example, this gives

$$\lim_{a\to\infty} \int_{-a}^{a} \frac{\sin(\pi x)}{\pi x}\, e^{ixt}\, dx \;=\; \begin{cases} 0 & \text{for } |t| > \pi, \\ 1 & \text{for } |t| < \pi, \\ \frac{1}{2} & \text{for } |t| = \pi. \end{cases} \qquad (2.2.13)$$

2.3 More Advanced Fourier Analysis

2.3.1 The Poisson Summation Formula

There are obvious similarities between Fourier series and Fourier transforms. Thus one would think that there are connections between the two concepts. That is indeed so, and the link is provided by the Poisson summation formula.

As a first example, take a function $F \in L^1(\mathrm{T})$, and note that $\widehat{F}_n = \widehat{f}(n)$ if $f \in L^1(\mathrm{R})$ is defined to equal F on $(-\pi, \pi)$ and to vanish everywhere else. This already is a simple special case of the Poisson summation formula.

A second approach to Poisson summation can be motivated by the following question. Take a function $g \in C(\mathrm{R})$, and assume that the sequence $(g(\frac{n}{w}))$ is absolutely summable for each $w > 0$. Consider $F_w(t) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{w}\right) e^{int}$ for $w > 0$. Such functions F_w are often investigated in connection with summation procedures for Fourier series, and we will do so later. A first graphical observation is that the restrictions to $(-\pi, \pi)$ of these functions F_w tend to be concentrated more and more around 0 for increasing w. If, however, the functions $(1/w) \cdot F_w(t/w)$ are plotted, then there seems to be convergence to a limit function that depends on g. What is this limit function, and how can we prove convergence? The answer is again given by the Poisson summation formula.

In general, the Poisson formula links the finite and the infinite transforms via the so-called *periodization*, an operation which associates $L^1(\mathrm{R})$-functions in a natural way with 2π-periodic functions: For $f \in L^1(\mathrm{R})$, set $F(t) = \sum_{j\in\mathrm{Z}} f(t+2\pi j)$ for all t for which the limit exists. The next theorem is (one version of) the classical Poisson formula, linking the Fourier series of F with the Fourier transform of f.

Theorem 2.3.1 (Poisson summation formula). *Let $f \in L^1(\mathrm{R})$.*

(a) The periodization F exists for almost every $t \in \mathrm{T}$, and we have $F \in L^1(\mathrm{T})$ and $\|F\|_1 \le \|f\|_1$.

(b) The Fourier series of F is

$$\sum_{j=-\infty}^{\infty} f(t + 2\pi j) \;\sim\; \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}(n)\, e^{int}, \tag{2.3.14}$$

or in other words, we have $\widehat{F}_n = \widehat{f}(n)$ for all $n \in \mathrm{Z}$.

Proof.
(a) Obviously $f(t + 2\pi j)$ is integrable over $[-\pi, \pi]$, and we have

$$\begin{aligned}\sum_{j=-\infty}^{\infty} \int_{-\pi}^{\pi} |f(t+2\pi j)|\, dt &= \sum_{j=-\infty}^{\infty} \int_{-\pi+2\pi j}^{\pi+2\pi j} |f(t)|\, dt \\ &= \int_{-\infty}^{\infty} |f(t)|\, dt = \|f\|_{L^1(\mathrm{R})} < \infty.\end{aligned}$$

By B. Levi's theorem, the series F is absolutely convergent a.e., we have $F \in L^1(\mathrm{T})$, and summation and integration can be exchanged. Now we also get

$\|F\|_{L^1(\mathrm{T})} \leq \|f\|_{L^1(\mathrm{R})}$ by using the triangle inequality, doing the exchange and then using the above computation.

(b) Now applying B. Levi's theorem to the functions $f(t+2\pi j) \cdot e^{-nt}$ for fixed n, and using the fact that the function e^{-int} is 2π-periodic, we get with the same summation trick as before,

$$\widehat{F}_n \;=\; \int_{-\pi}^{\pi} F(t)\, e^{-int}\, dt \;=\; \int_{-\infty}^{\infty} f(t)\, e^{-int}\, dt \;=\; \widehat{f}(n).$$

□

Of course, it is interesting to ask when the identity holds pointwise instead of just in the sense of Fourier series. From the Jordan test, it can be deduced that if $f \in L^1(\mathrm{R}) \cap BV(\mathrm{R})$ and $f(t) = \frac{1}{2}(f(t+) + f(t-))$ everywhere, then equality holds for all t.

Example 2.3.2 *Fourier series of hyperbolic functions.*

Choose $y > 0$ and define

$$f(t) \;=\; \begin{cases} e^{-yt} & \text{for } t > 0, \\ 0 & \text{for } t < 0, \end{cases}$$

and $f(0) = 1/2$. Then $\widehat{f}(x) = (y + ix)^{-1}$, and Poisson says that

$$\sum_{j=-\infty}^{\infty} f(t + 2\pi j) \;=\; \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \frac{1}{y + in}\, e^{int},$$

which is equivalent to

$$\frac{1}{2\pi}\left(\frac{1}{y} + 2\sum_{n=1}^{\infty} \frac{y\cos(nt) + n\sin(nt)}{y^2 + n^2}\right) \;=\; \sum_{j > t/(2\pi)} e^{-y(t+2\pi j)} + \begin{cases} \frac{1}{2}, & t \in 2\pi\mathrm{Z}, \\ 0, & \text{otherwise.} \end{cases}$$

Setting $t = 0$, we get

$$\pi \coth(\pi y) \;=\; \frac{1}{y} + 2y \sum_{n=1}^{\infty} \frac{1}{y^2 + n^2}.$$

Setting $t = 1/2$, we get

$$\pi\mathrm{cosech}(\pi y) = \frac{1}{y} + 2y \sum_{n=1}^{\infty} \frac{(-1)^n}{x^2 + n^2}. \qquad \square$$

Example 2.3.3 *A quadratic exponential identity.*

Now choose $s > 0$ and set $f(t) = e^{-st^2}$. Then $\widehat{f}(x) = \sqrt{\pi/s}\, e^{-x^2/(4s)}$ (see Section 5.6 of the first volume), and Poisson says that

$$\sum_{j=-\infty}^{\infty} e^{-s(t+2\pi j)^2} = \frac{1}{2}\sqrt{\frac{1}{\pi s}} \sum_{n=-\infty}^{\infty} e^{-n^2/(4s)}\, e^{int}. \tag{2.3.15}$$

By analyticity, this can be extended to hold for all $\mathrm{Re}(s) > 0$. We will meet this formula again in Section 3.1. $\square$

Example 2.3.4 *An experimental version of Poisson's formula.*

Take an even real-valued continuous $g \in L^1(\mathrm{R})$ such that $(g(\frac{n}{w}))$ is absolutely summable for each $w > 0$ and assume that $\widehat{g} \in L^1(\mathrm{R})$. Then using Poisson and the inversion theorem 2.2.7 it follows that

$$\sum_{n=-\infty}^{\infty} g\left(\frac{n}{w}\right) e^{int} = w \sum_{j=-\infty}^{\infty} \widehat{g}\left(w(t + 2\pi j)\right), \tag{2.3.16}$$

where equality is meant in $L^1(\mathrm{T})$. This equality does not hold pointwise! In Katznelson's book [149], an example is given of a function $f \in L^1(\mathrm{R})$ with $\widehat{f} \in L^1(\mathrm{R})$ for which equality does not hold pointwise in Poisson's formula (2.3.14), because the periodization does not converge uniformly. This is the "dangerous pitfall" that we mentioned earlier: Even though the left-hand side of (2.3.16) is a continuous function, the right-hand side is not necessarily continuous; it is only equal to a continuous function a.e. To deduce pointwise equality from this would be wrong! This is a noteworthy difference to the situation in Theorem 2.2.7.

In any case, this gives an answer to the question at the beginning of the present subsection: The limit function (in the $L^1(\mathrm{R})$ sense) of the F_w, restricted to $(-\pi,\pi)$ and then suitably rescaled, is the Fourier transform of g. This follows from (2.3.16), since

$$\begin{aligned}\|\tfrac{1}{w}F_w\left(\tfrac{t}{w}\right)\cdot\chi_{(-\pi w,\pi w)}-\widehat{g}(t)\|_1 &= \int_{-\pi w}^{\pi w}\Big|\sum_{\substack{j=-\infty\\ j\neq 0}}^{\infty}\widehat{g}(t+2\pi jw)\Big|\,dt+\int_{|t|>\pi w}|\widehat{g}(t)|\,dt\\ &\leq 2\int_{|t|>w}|\widehat{g}(t)|\,dt \;\to\; 0 \quad (w\to\infty)\end{aligned}$$

if $\widehat{g}\in L^1(\mathrm{R})$. Often enough (but not always) this convergence is even uniform on R.

This is a very visual theorem; it can be discovered and explored on the computer. In this sense, Formula (2.3.16) is the experimental version of Poisson's summation formula! $\square$

2.3.2 Convolution Theorems

As we have seen, the Fourier series even of a continuous function need not converge back to the function, neither in the L^1-sense nor pointwise. Often, convergence properties of such a series can be improved by putting additional factors into the series to "force convergence." For example, it can be proved (and will be in the next subsection) that for $f\in C(\mathrm{T})$, the series

$$\sigma_n(f,t) \;=\; \frac{1}{2\pi}\sum_{k=-n}^{n}\left(1-\frac{|k|}{n+1}\right)\widehat{f}_k\,e^{ikt}$$

converges to f uniformly as well as in the L^1-sense; this is Fejér's famous theorem.

How do these convergence factors work? If we set

$$F_n(t) \;=\; \sum_{k=-n}^{n}\left(1-\frac{|k|}{n+1}\right)e^{ikt},$$

then the Fourier coefficients of $\sigma_n(f,t)$ are the product of those of f and those of F_n. Thus it seems reasonable that convergence properties of σ_n can be deduced from suitable properties of F_n. An important relation between these objects is given by the following theorem.

Theorem 2.3.5 (Convolution theorem for the torus.) *For $f, g \in L^1(\mathrm{T})$, define*

$$h(t) \;=\; (f * g)(t) \;=\; \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t-u)\, g(u)\, du.$$

(a) Then the integral exists for a.e. $t \in \mathrm{T}$, and we have $h \in L^1(\mathrm{T})$ and $\|h\|_1 \leq \frac{1}{2\pi} \|f\|_1 \|g\|_1$.

(b) Moreover, $\widehat{h}_n = \frac{1}{2\pi} \widehat{f}_n \cdot \widehat{g}_n$ holds for the Fourier coefficients.

Applied to the question above, this convolution theorem says that $\sigma_n(f,t) = (F_n * f)(t)$. In the next section, this will lead to a proof of Fejér's theorem. We will in fact be able to treat much more general summation kernels.

In the following sections, we will also need a convolution theorem in $L^1(\mathrm{R})$.

Theorem 2.3.6 (Convolution theorem for the real line.) *For $f, g \in L^1(\mathrm{R})$, define*

$$h(t) \;=\; (f * g)(t) \;=\; \int_{-\infty}^{\infty} f(t-u)\, g(u)\, du.$$

(a) Then the integral exists for a.e. $t \in \mathrm{R}$, and we have $h \in L^1(\mathrm{R})$ and $\|h\|_1 \leq \|f\|_1 \|g\|_1$.

(b) Moreover, $\widehat{h}(x) = \widehat{f}(x) \cdot \widehat{g}(x)$ holds for the Fourier transforms.

The convolution in $L^1(\mathrm{R})$ tends to make functions smoother but less localized. If g is, for example, the characteristic function of an interval, then $h = f * g$ will be absolutely continuous for every L^1-function f. If, on the other hand, both f and g are L^1-functions with bounded supports, equal to, say, $[-a, a]$ and $[-b, b]$, then the support of $h = f * g$ will be equal to $[-(a+b), a+b]$.

Convolution theorems are quite important in computational mathematics! To compute a convolution, one usually has to perform many multiplications, so that it is expensive in terms of time and memory. On the other hand, a single multiplication is often cheap. The convolution theorems (which have many analogs for different types of convolutions) say that convolutions can be transformed into multiplications and thus may be much easier to compute than appears on first glance. Techniques for computing convolutions using the fast Fourier transform (FFT) are presented in Chapter 6 of the first volume. The speed of convolutions computed using FFTs is a principal reason that Fourier theory in general, and the FFT in particular, are so important in computational science.

2.3.3 Summation Kernels

With regard to Fejér's sum σ_n, the convolution theorem says that $\sigma_n(f,t) = (F_n * f)(t)$, and we are interested in the question of whether $F_n * f \to f$ for $n \to \infty$ in a suitable norm (L^1 or uniformly). This question can be generalized: Under what conditions on a family of functions $(K_w) \subseteq L^1(\mathrm{T})$ is there convergence $\|K_w * f - f\|_1 \to 0$ $(w \to \infty)$ for all $f \in L^1(\mathrm{T})$? Under what conditions is there uniform convergence for continuous f? Any family of type (K_w) is called a *kernel*. If there is suitable convergence, then (K_w) is also called an *approximate identity*, and the question is now open to systematic experimentation. Which conditions of a kernel make it an approximate identity?

It is proved in Katznelson's book [149] that the following conditions imply that a kernel $(K_w) \subseteq L^1(\mathrm{T})$ is an approximate identity for $L^p(\mathrm{T})$ $(1 \leq p < \infty)$ and for $C(\mathrm{T})$:

$$\text{(S1)} \qquad \frac{1}{2\pi}\int_{-\pi}^{\pi} K_w(t)\,dt \;=\; 1 \qquad \text{for all } w,$$

$$\text{(S2)} \qquad \int_{-\pi}^{\pi} |K_w(t)|\,dt \;\leq\; M \qquad \text{uniformly in } w,$$

$$\text{(S3)} \qquad \lim_{w\to\infty} \int_{\delta \leq |t| \leq \pi} |K_w(t)|\,dt \;=\; 0 \qquad \text{for all } 0 < \delta < \pi.$$

Of particular interest here are kernels of the form $K_w(t) = \sum_{k=-\infty}^{\infty} g(\frac{k}{w})\, e^{ikt}$ for suitable functions g, since many classical kernels are of this form. Now our direction should be clear: The "experimental" version (2.3.16) of the Poisson formula will be of use.

Theorem 2.3.7 *Let* $K_w(t) = \sum_{k=-\infty}^{\infty} g(\frac{k}{w})\, e^{ikt}$ *and assume that* $g \in C(\mathrm{R}) \cap L^1(\mathrm{R})$*, that* $(g(\frac{n}{w}))$ *is absolutely summable for each* $w > 0$*, and that* $g(0) = 1$ *and* $\widehat{g} \in L^1(\mathrm{R})$*. Then* (K_w) *is an approximate identity.*

Proof. We have to check (S1)–(S3) above. Condition (S1) is a condition on the middle Fourier coefficient of K_w and follows from $g(0) = 1$. Regarding (S2), we have, using (2.3.16) and the theorem of B. Levi,

$$\begin{aligned}
\int_{-\pi}^{\pi} |K_w(t)|\,dt &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left| w\sum_{j=-\infty}^{\infty} \overline{\widehat{g}}(w(t+2\pi j))\right| dt \\
&\leq w\sum_{j=-\infty}^{\infty}\int_{-\pi}^{\pi} |\widehat{g}(w(t+2\pi j))|\,dt \\
&= w\int_{-\infty}^{\infty} |\widehat{g}(wt)|\,dt \;=\; \|\widehat{g}\|_1,
\end{aligned}$$

which is a uniform bound. Finally, we get by similar computations as before that

$$\int_{\delta\leq|t|\leq\pi} |K_w(t)|\,dt \;\leq\; \int_{|t|>w\delta} |\widehat{g}(t)|\,dt,$$

which tends to 0 for $w \to \infty$ since $\widehat{g} \in L^1(\mathrm{R})$. □

All of the conditions in Theorem 2.3.7 are easily checked symbolically. Moreover, the methods employed here allow more detailed investigations of the approximation properties of the kernels by direct computation of the Fourier transform. For example, often the *Lebesgue constants*, defined as $\|K_w\|_1$, determine the rate of convergence of $K_w * f$ to f, or, if K_w is not an approximate identity, the growth rate of $K_w * f$. The Lebesgue constants, as computed in the proof of (S2), satisfy $\|K_w\|_1 \leq \|\widehat{g}\|_1$. This bound is precise, since the same Poisson methods also give

$$\|K_w\|_1 \;\geq\; \|\widehat{g}\|_1 - 2\int_{|t|>\pi w} |\widehat{g}(t)|\,dt.$$

Example 2.3.8 *The Fejér kernel.*

The kernel F_n as defined above comes from the Cesàro summation method applied to Fourier series; thus $\sigma_n(f,t) = (F_n * f)(t)$. By geometric summation, F_n can also be written as

$$F_n(t) \;=\; \frac{\sin^2(\frac{n+1}{2}t)}{(n+1)\sin^2(\frac{t}{2})}.$$

If we set $g(x) = \max\{1 - |x|, 0\}$, then $F_n = K_{n+1}$. Since $\widehat{g}(t) = \mathrm{sinc}^2(t/2) \in L^1(\mathrm{R})$, the conditions of Theorem 2.3.7 are satisfied. Thus we deduce directly that F_n is an approximate identity in $L^p(\mathrm{T})$ and in $C(\mathrm{T})$. Moreover, since $\widehat{g}$ is non-negative, for the Lebesgue constants we obtain $\|F_n\|_1 = \|\widehat{g}\|_1 = 2\pi$ (compare with the example in Section 2.2.2). $\square$

Example 2.3.9 *The Poisson kernel.*

Another important summation method is Abel summation. Applied to Fourier series, it leads to the Poisson kernel, defined as

$$P_r(t) \;=\; \sum_{k=-\infty}^{\infty} r^{|k|}\, e^{ikt} \;=\; \frac{1-r^2}{1-2r\cos(t)+r^2} \quad \text{for } 0 \le r < 1.$$

The question is whether $P_r * f \to f$ in $L^p(\mathrm{T})$ or in $C(\mathrm{T})$ for $r \to 1$. By setting $w = -1/\ln(r)$ and $g(x) = e^{-|x|}$, we have $P_r(t) = K_w(t)$. Since $\widehat{g}(t) = 2/(1+t^2) \in L^1(\mathrm{R})$, Theorem 2.3.7 produces convergence of $P_r * f$ to f. Similarly to the Fejér kernel, P_r as well as g and $\widehat{g}$ are nonnegative, so that we again have $\|P_r\|_1 = \|K_w\|_1 = \|\widehat{g}\|_1 = 2\pi$. $\square$

Example 2.3.10 *The Dirichlet kernel.*

The same methods also explain why the usual summation of Fourier series does not always give convergence. This summation corresponds to the Dirichlet kernel

$$D_n(t) \;=\; \sum_{k=-n}^{n} e^{ikt} \;=\; \frac{\sin\left(\left(n+\frac{1}{2}\right)t\right)}{\sin\left(\frac{t}{2}\right)}$$

via $s_n(f,t) = (D_n * f)(t)$. The Dirichlet kernel is of the form K_n with $g = \chi_{[-1,1]}$. This g is neither continuous, nor does it have a Fourier transform in $L^1(\mathrm{R})$ (its Fourier transform is $\widehat{g}(t) = 2\ \mathrm{sinc}(t)$.) Thus Theorem 2.3.7 is not applicable. But the experimental Poisson formula (2.3.16) can still be used to estimate the Lebesgue constants and gives $\|D_n\|_1 = \int_{-n\pi}^{n\pi} |2\ \mathrm{sinc}(t)|\, dt + O(1) = (8/\pi)\ln n + O(1)$. These are unbounded, and so the Dirichlet kernel can be expected to have worse summation properties than the Fejér and Poisson kernels above. $\square$

In summary, the methods described here are very useful to gauge norm-convergence properties of kernels of the special form $K_w(t) = \sum g(\frac{k}{w})\, e^{ikt}$ in a direct, computational way, and they open the door to further experimentation. Our description has been adapted from [123], but reportedly these methods go back at least to Korovkin.

These methods deal with norm- and uniform convergence. But what about pointwise convergence? As described in [149], instead of (S1)–(S3), the following properties of a kernel can be used to prove pointwise convergence of $K_w * f$ to f : K_w satisfies (S1), is non-negative and even, and satisfies

$$\lim_{w\to\infty} \left(\sup_{\delta\le|t|\le\pi} K_w(t) \right) = 0 \quad \text{for all } 0 < \delta < \pi.$$

This allows one to decide, for given $f \in L^1(\mathrm{T})$ and $t_0 \in \mathrm{R}$, whether $(K_w * f)(t_0)$ converges to $f(t_0)$.

For the Fejér kernel, it leads to Lebesgue's condition: *If there exists a value* $\check{f}(t_0)$ *such that*

$$\lim_{h\to 0} \frac{1}{h} \int_0^h \left| \frac{f(t_0+t)+f(t_0-t)}{2} - \check{f}(t_0) \right| dt = 0,$$

then $\sigma_n(f,t) \to \check{f}(t_0)$ *for* $n \to \infty$*. In particular,* $\sigma_n(f,t) \to f(t)$ *a.e.*

For the Poisson kernel, it leads to Fatou's condition: *If there exists a value* $\check{f}(t_0)$ *such that*

$$\lim_{h\to 0} \frac{1}{h} \int_0^h \left(\frac{f(t_0+t)+f(t_0-t)}{2} - \check{f}(t_0) \right) dt = 0,$$

then $(P_r * f)(t_0) \to \check{f}(t_0)$ *for* $r \to 1-$*. In particular,* $(P_r * f)(t) \to f(t)$ *a.e.*

The convergence is uniform on closed subintervals where f is continuous.

2.4 Examples and Applications

2.4.1 The Gibbs Phenomenon

If a function $f \in L^1(\mathrm{T})$ is of bounded variation, it may have jump discontinuities. The Jordan test says that the Fourier series of f converges to the center of the gap at such a point. Directly to the left and right of the jump, the series converges pointwise, but not uniformly on any interval containing the discontinuity, to the function. The function $f(t) = (\pi - t)/2 = \sum n^{-1} \sin(nt)$ on $[0, 2\pi]$ is a good example for this behavior (see Figure 2.1), where the series for $(\pi - t)/2$ is evaluated to 20 terms.

One notices that the truncated Fourier series "overshoots" the function at the discontinuity. These oscillations do not diminish when more terms are added; they just move closer to the discontinuity. When the experimental physicist A. Michelson (famous for the Michelson-Morley experiment, which led to Einstein's special relativity) had built a machine to calculate Fourier series and fed it a discontinuous function, he noticed this phenomenon. It was unexpected for him, but after hand calculations confirmed this behavior, he wrote a letter to *Nature* in 1898, expressing his doubts that "a real discontinuity can replace a sum of continuous curves" (cited after Bhatia [25]). J. Willard Gibbs, one of the founders of modern thermodynamics, replied to this letter and clarified the matter. Thus here we have another example of a mathematical theorem that was experimentally discovered (by an experimental physicist)!

Now what is the explanation for the Gibbs phenomenon? Inspection of the picture shows that the largest overshoot seems to occur around the point π/N if N terms of the Fourier series are added. Thus we compute

$$s_N\left(f, \frac{\pi}{N}\right) = \sum_{n=1}^{N} \frac{\sin\left(\frac{n\pi}{N}\right)}{n} = \frac{\pi}{N} \sum_{n=1}^{N} \frac{\sin\left(\frac{n\pi}{N}\right)}{\frac{n\pi}{N}},$$

where the last sum is a Riemann sum for the integral

$$I = \int_0^{\pi} \frac{\sin(t)}{t}\, dt.$$

Therefore, $s_N(f, \pi/N) \to I$ for $N \to \infty$. Since $I/f(0+) = I/(\pi/2) \approx 1.178979744$, this explains why the overshoot does not go away for large N. This overshoot of roughly 18% is not dependent on the function f used here as an example, but can be observed (and proved) for any jump discontinuity.

Does the Gibbs phenomenon vanish when we use Fejér's series instead of the Fourier series? Figure 2.2 shows the Fejér approximation to f, again to 20 terms. The oscillation is now replaced by a pronounced "undershoot" to the right of 0 (this can be explained by the positivity of the Fejér kernel); again, the undershoot can be observed to move closer to the discontinuity, but not vanish altogether when more terms are added. In fact, we have to pay for the increased smoothness of the approximation by its reduced willingness to snuggle up to the limit function.

2.4.2 A Function with Given Integer Moments

The k-th moment of a function $f \in L^1(\mathrm{R})$ is defined as

$$\mu_k(f) \;=\; \int_{-\infty}^{\infty} f(t)\, t^k\, dt,$$

provided that $t \mapsto f(t)\, t^k \in L^1(\mathrm{R})$. The *Hamburger moment problem* is the problem to find a function f with a given sequence of moments (μ_k). This problem is underdetermined: There can be nonvanishing functions whose every moment is 0. This is easily seen by the following argument. Assume that f is k times differentiable with every derivative in $L^1(\mathrm{R})$. Then by partial integration,

$$\widehat{f^{(k)}}(x) \;=\; \int_{-\infty}^{\infty} f^{(k)}(t)\, e^{-ixt}\, dt \;=\; (ix)^k\, \widehat{f}(x),$$

and by the inversion theorem, $\int_{-\infty}^{\infty} \widehat{f}(x)\, x^k\, dx = 2\pi f^{(k)}(0)$. Thus if $f \in L^1(\mathrm{R})$ is infinitely differentiable with every derivative in $L^1(\mathrm{R})$ and satisfies $f^{(k)}(0) = 0$ for all k, then all moments of $\widehat{f}$ vanish. Of course, such nontrivial functions f exist, even with compact support.

Obviously, for an even function every odd moment vanishes. To generalize this, it is quite easy to find, for given $n \in \mathrm{N}$, a function whose k-th moment is nonzero precisely when $k \bmod n = 0$. Just choose an infinitely differentiable function $f \in L^1(\mathrm{R})$ with all derivatives in $L^1(\mathrm{R})$, and such that all its derivatives are nonzero at 0. Then set $g(t) = f(t^n)$, and by the chain and product rule of differentiation, $g^{(k)}(0)$ is nonzero precisely for $k \bmod n = 0$. Thus, $\widehat{g}$ satisfies the moment condition. If f is analytic, say $f(t) = \sum_{j=0}^{\infty} a_j\, t^j$, then $g^{(k)}(0) = (qn)!\, a_q$ if $k = qn$ and $g^{(k)}(0) = 0$ otherwise. An example for such a function f is $f(t) = (1+t)\, e^{-t^2/2}$.

When this was first investigated some years ago, numerical fast Fourier transforms were used—as a test—to calculate the moments for $g(t) = f(t^n)$, where $f(t) = (1+t)\, e^{-t^2/2}$ as above. The scheme for doing this is presented in Section

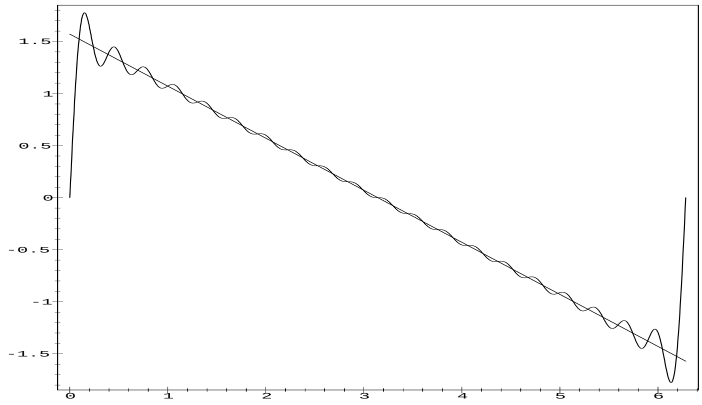

Figure 2.1: The Gibbs phenomenon for Fourier series.

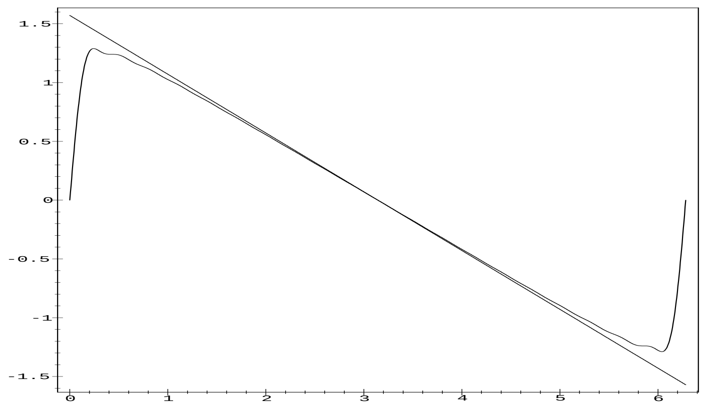

Figure 2.2: The Gibbs phenomenon for Fejér series.

6.1 of the first volume. When this was done, it was noticed that the resulting moment values were extremely accurate, far more than one would expect based on what amounts to a simple step-function approximation to the Fourier integrals. Readers may recall that we also encountered this phenomenon in Section 5.2 of the first volume. This phenomenon, which is rooted in the Euler-Maclaurin summation formula, is the foundation of some extremely efficient and highly accurate numerical quadrature schemes, which we shall see in Sections 7.4.2 and 7.4.3.

2.4.3 Bernoulli Convolutions

Consider the discrete probability density on the real line with measure 1/2 at each of the two points ± 1. The corresponding measure is the so-called Bernoulli measure, denoted $b(X)$. For every $0 < q < 1$, the infinite convolution of measures

$$\mu_q(X) \quad = \quad b(X) * b(X/q) * b(X/q^2) * \cdots \tag{2.4.17}$$

exists as a weak limit of the finite convolutions. The most basic theorem about these infinite Bernoulli convolutions is due to Jessen and Wintner ([145]). They proved that μ_q is always continuous, and that it is either absolutely continuous or purely singular. This statement follows from a more general theorem on infinite convolutions of purely discontinuous measures (Theorem 35 in [145]); however, it is not difficult to prove the statement directly with the use of Kolmogoroff's 0-1-law (which can be found, e.g., in [26]). The question about these measures is to decide for which values of the parameter q they are singular, and for which q they are absolutely continuous.

This question can be recast in a more real-analytic way by defining the distribution function F_q of μ_q as

$$F_q(t) \quad = \quad \mu_q(-\infty, t], \tag{2.4.18}$$

and to ask for which q this continuous, increasing function F_q is singular, and for which it is absolutely continuous. Note that F_q satisfies $F_q(t) = 0$ for $t < -1/(1-q)$ and $F_q(t) = 1$ for $t > 1/(1-q)$.

Another way to define the distribution function F_q is by functional equations: F_q is the only bounded solution of the functional equation

$$F(t) \quad = \quad \frac{1}{2}F\left(\frac{t-1}{q}\right) + \frac{1}{2}F\left(\frac{t+1}{q}\right) \tag{2.4.19}$$

with the above restrictions. Moreover, if F_q is absolutely continuous and thus has a density $f_q \in L^1(\mathrm{R})$, then f_q satisfies the functional equation

$$2q\, f(t) \;=\; f\left(\frac{t-1}{q}\right) + f\left(\frac{t+1}{q}\right) \tag{2.4.20}$$

almost everywhere. This is a special case of a much more general class of equations, namely two-scale difference equations. Those are functional equations of the type

$$f(t) \;=\; \sum_{n=0}^{N} c_n\, f(\alpha t - \beta_n) \qquad (t \in \mathrm{R}), \tag{2.4.21}$$

with $c_n \in \mathrm{C}$, $\beta_n \in \mathrm{R}$ and $\alpha > 1$. They were first discussed by Ingrid Daubechies and Jeffrey C. Lagarias, who proved existence and uniqueness theorems and derived some properties of L^1-solutions [105, 106]. One of their theorems, which we state here in part for the general equation (2.4.21) and in part for the specific case (2.4.20), is the following:

Theorem 2.4.1 *(a) If $\alpha^{-1}(c_0 + \cdots + c_N) = 1$, then the vector space of $L^1(\mathrm{R})$-solutions of (2.4.21) is at most one-dimensional.*

(b) If, for given $q \in (0,1)$, equation (2.4.20) has a nontrivial L^1-solution f_q, then its Fourier transform satisfies $\widehat{f_q}(0) \neq 0$, and is given by

$$\widehat{f_q}(x) \;=\; \widehat{f_q}(0) \prod_{n=0}^{\infty} \cos(q^n\, x). \tag{2.4.22}$$

In particular, for normalization we can assume $\widehat{f_q}(0) = 1$.

(c) On the other hand, if the right-hand side of (2.4.22) is the Fourier transform of an L^1-function f_q, then f_q is a solution of (2.4.20).

(d) Any nontrivial L^1-solution of (2.4.21) is finitely supported. In the case of (2.4.20), the support of f_q is contained in $[-1/(1-q), 1/(1-q)]$.

This implies in particular that the question of whether the infinite Bernoulli convolution (2.4.17) is absolutely continuous is equivalent to the question of whether (2.4.20) has a nontrivial L^1-solution. Now what is known about these questions?

It is relatively easy to see that in the case $0 < q < 1/2$, the solution of (2.4.19) is singular; it is in fact a Cantor function, meaning that it is constant on a dense set of intervals. This was first proved by R. Kershner and A. Wintner [151]. (An example of a Cantor function is depicted in Figure 6.1 of the first volume.)

It is also easy to see that in the case $q = 1/2$, there is an L^1-solution of (2.4.20), namely $f_{1/2}(t) = \frac{1}{4}\chi_{[-2,2]}(t)$. Moreover, this can be used to construct a solution for every $q = 2^{-1/p}$ where p is an integer, namely

$$f_{2^{-1/p}}(t) = 2^{(p-1)/2} \cdot \left[f_{1/2}(t) * f_{1/2}(2^{1/p}t) * \cdots * f_{1/2}(2^{(p-1)/p}t)\right]. \qquad (2.4.23)$$

This was first noted by Wintner via the Fourier transform [214]. Explicitly, we have

$$\begin{aligned}\widehat{f_{2^{-1/p}}}(x) &= \prod_{n=0}^{\infty} \cos(2^{-n/p}x) = \prod_{m=0}^{\infty}\prod_{k=0}^{p-1} \cos(2^{-(m+k/p)}x) \\ &= \widehat{f_{1/2}}(x) \cdot \widehat{f_{1/2}}(2^{-1/p}x) \cdots \widehat{f_{1/2}}(2^{-(p-1)/p}x),\end{aligned}$$

which is equivalent to (2.4.23) by the convolution theorem.

Note that the regularity of these solutions $f_{2^{-1/p}}$ increases when p and thus $q = 2^{-1/p}$ increases: $f_{2^{-1/p}} \in C^{p-2}(\mathrm{R})$. From the results given so far, one might therefore surmise that (2.4.20) would have a nontrivial L^1-solution for every $q \geq 1/2$ with increasing regularity when q increases. This supposition, however, would be wrong, and it came as a surprise when Erdős proved in 1939 [116] that there are some values of $1/2 < q < 1$ for which (2.4.20) does not have an L^1-solution, namely, the inverses of Pisot numbers. A *Pisot number* (discussed further in Exercise 13 of Chapter 7) is defined to be an algebraic integer greater than 1 all of whose algebraic conjugates lie inside the unit disk. The best known example of a Pisot number is the golden mean $\varphi = (\sqrt{5}+1)/2$. The characteristic property of Pisot numbers is that their powers quickly approach integers: If a is a Pisot number, then there exists a θ, $0 < \theta < 1$, such that

$$\operatorname{dist}(a^n, \mathrm{Z}) \leq \theta^n \quad \text{for all } n \in \mathrm{N}. \qquad (2.4.24)$$

Erdős used this property to prove that if $q = 1/a$ for a Pisot number a, then $\limsup_{x\to\infty} \left|\widehat{f_q}(x)\right| > 0$. Thus in these cases, f_q cannot be in $L^1(\mathrm{R})$, since that would contradict the Riemann-Lebesgue lemma. Erdős's proof uses the Fourier transform $\widehat{f}_q$: Consider, for $N \in \mathrm{N}$,

$$\left|\widehat{f}_q(q^{-N}\pi)\right| \;=\; \prod_{n=1}^{\infty} |\cos(q^n\pi)| \cdot \prod_{n=0}^{N-1} \left|\cos(q^{-n}\pi)\right| \;=:\; C \cdot p_N,$$

where $C > 0$. Moreover, choose $\theta \neq 1/2$ according to (2.4.24) and note that

$$\begin{aligned}
p_N \;&=\; \prod_{\substack{n=0 \\ \theta^n \leq 1/2}}^{N-1} \left|\cos(q^{-n}\pi)\right| \cdot \prod_{\substack{n=0 \\ \theta^n > 1/2}}^{N-1} \left|\cos(q^{-n}\pi)\right| \\
&\geq\; \prod_{\substack{n=0 \\ \theta^n \leq 1/2}}^{N-1} \cos(\theta^n\pi) \cdot \prod_{\substack{n=0 \\ \theta^n > 1/2}}^{N-1} \left|\cos(q^{-n}\pi)\right| \\
&\geq\; \prod_{\substack{n=0 \\ \theta^n \leq 1/2}}^{\infty} \cos(\theta^n\pi) \cdot \prod_{\substack{n=0 \\ \theta^n > 1/2}}^{\infty} \left|\cos(q^{-n}\pi)\right| \;=\; C' \;>\; 0,
\end{aligned}$$

independently of N.

In 1944, Raphaël Salem [191] showed that the reciprocals of Pisot numbers are the only values of q where $\widehat{f}_q(x)$ does not tend to 0 for $x \to \infty$. In fact, no other $q > 1/2$ are known at all where F_q is singular. Moreover, the set of explicitly given q with absolutely continuous F_q is also not very big: The largest such set known to date was found by Adriano Garsia in 1962 [120]. It contains reciprocals of certain algebraic numbers (such as roots of the polynomials $x^{n+p} - x^n - 2$ for $\max\{p, n\} \geq 2$) besides the roots of $1/2$.

Matters remained in this state for more than 30 years; the question remained settled only for countably many $q \in [1/2, 1)$. The most recent significant progress then was made in 1995 by Boris Solomyak [201], who developed exciting new methods in geometric measure theory to prove that F_q is in fact absolutely continuous for almost every $q \in [1/2, 1)$. (See also [177] for a simplified proof and [176] for a survey and some newer results.)

This, however, yields no explicit result; the set of q for which the behavior of F_q is known explicitly is the same as before. Here we suggest an experimental approach to at least identify q-values for which the behavior of F_q can be guessed. In fact, define a map T_q, mapping the set of L^1-functions with support in $[-1/(1-$

$q), 1/(1-q)]$ and with $\widehat{f_q}(0) = 1$ into itself, by

$$(T_q f)(t) \;=\; \frac{1}{2q}\left(f\left(\frac{t-1}{q}\right) + f\left(\frac{t+1}{q}\right)\right) \quad \text{for } t \in \mathrm{R}.$$

Then note that the fixed points of T_q are the solutions of (2.4.20) and that T_q is nonexpansive (see Chapter 6). Therefore, one may have hope that by iterating the operator, it may be possible to approximate the fixed point. In fact, if a sequence of iterates $T_q^n f$ converges in $L^1(\mathrm{R})$ for some initial function f, then the limit will be a fixed point of T_q, since T_q is continuous. It is, however, not easy to prove convergence; no convergence proof is known. It is, on the other hand, possible to prove a weaker result, namely convergence in the mean, provided that a fixed point exists: *If a solution $f_q \in L^1(\mathrm{R})$ with $\widehat{f_q}(0) = 1$ of (2.4.20) exists, then for every initial function $f \in L^1(\mathrm{R})$ with support in $[-1/(1-q), 1/(1-q)]$ and with $\widehat{f}(0) = 1$, we have,*

$$\lim_{n\to\infty} \left\| \frac{1}{n}\sum_{k=0}^{n-1} T_q^k f - f_q \right\|_1 \;=\; 0.$$

This theorem follows from properties of Markov operators [155] and from a result by Mauldin and Simon [164], showing that if an L^1-density f_q exists, then it must be positive a.e. on its support.

In practice, we observe that the iterates usually seem to converge directly, even without the means. Plotting them, we hope to infer existence and regularity of L^1-solutions by visual inspection. The figures on the next pages show the 25th iterate for $f = (1-q)/2\,\chi_{[-1/(1-q),1/(1-q)]}$ as initial function. Figure 2.3 shows convergence to $2^{1/2} \cdot \left[f_{1/2}(t) * f_{1/2}(2^{1/2}t)\right]$ for $q = 1/\sqrt{2}$; Figure 2.4 shows that for $q = (\sqrt{5}-1)/2$, the iterates do not converge to a meaningful function. It is not known if there is a density for any rational $q \in (1/2, 1)$. Figures 2.5 and 2.6 show that there seems to be a continuous limit in both cases shown; moreover, regularity seems to increase when q increases.

2.5 Some Curious Sinc Integrals

Define

$$I_n \;=\; \int_0^\infty \operatorname{sinc} x \cdot \operatorname{sinc}\left(\frac{x}{3}\right) \cdots \operatorname{sinc}\left(\frac{x}{2n+1}\right)\, dx.$$

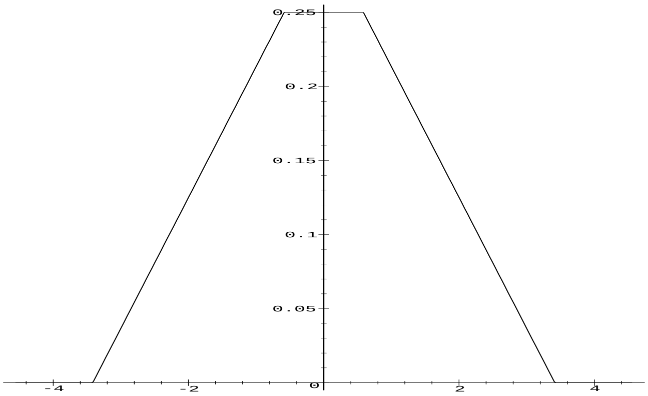

Figure 2.3: 25th iterate for $q = 1/\sqrt{2}$.

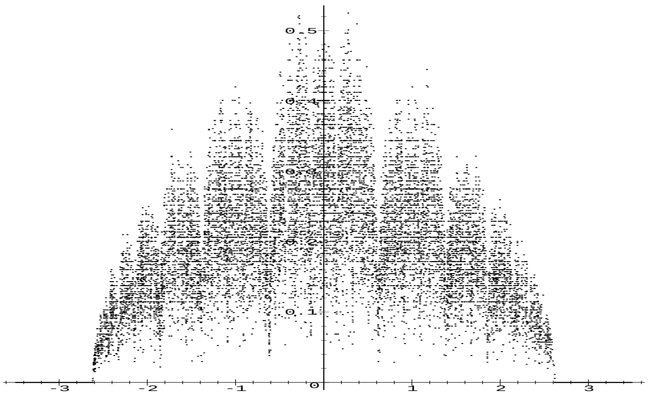

Figure 2.4: 25th iterate for $q = (\sqrt{5} - 1)/2$.

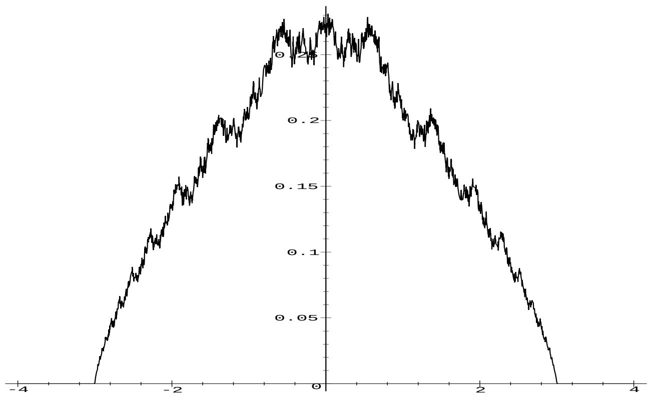

Figure 2.5: 25th iterate for $q = 2/3$.

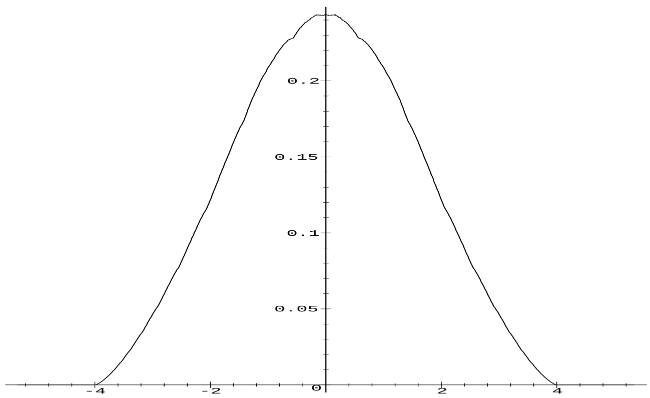

Figure 2.6: 25th iterate for $q = 3/4$.

Then *Maple* and *Mathematica* evaluate

$$\begin{aligned}
I_0 &= \int_0^\infty \operatorname{sinc} x\, dx \;=\; \frac{\pi}{2},\\
I_1 &= \int_0^\infty \operatorname{sinc} x \cdot \operatorname{sinc}\left(\frac{x}{3}\right) dx \;=\; \frac{\pi}{2},\\
&\;\vdots\\
I_6 &= \int_0^\infty \operatorname{sinc} x \cdot \operatorname{sinc}\left(\frac{x}{3}\right) \cdots \operatorname{sinc}\left(\frac{x}{13}\right) dx \;=\; \frac{\pi}{2}, \quad \text{but}\\
I_7 &= \int_0^\infty \operatorname{sinc} x \cdot \operatorname{sinc}\left(\frac{x}{3}\right) \cdots \operatorname{sinc}\left(\frac{x}{15}\right) dx\\
&= \frac{467807924713440738696537864469}{935615849440640907310521750000}\pi,
\end{aligned}$$

where the fraction is approximately $0.499999999992646\ldots$

When this fact was recently verified by a researcher using a computer algebra package, he concluded that there must be a "bug" in the software. This conclusion may be too hasty, but it does raise the question: How far can we trust our computer algebra system? Or as computer scientists often ask, "Is it a bug or a feature?"

In this section, we will derive general formulas for this type of sinc integrals, thereby proving that all of the above evaluations are in fact correct. Thus, this is a somewhat cautionary example for too enthusiastically inferring patterns from symbolic or numerical computations. The material comes from [33]; additional information can be found in [34].

2.5.1 The Basic Sinc Integral

It will turn out that the general multi-fold sinc integral can be reduced to the integral I_0, so that it makes sense to first evaluate this integral. Note that the function $\operatorname{sinc} x$ is not an element of $L^1(\mathrm{R})$! Thus, Lebesgue theory cannot be applied here directly, and in fact the integral has to be interpreted correctly. Here we use the usual interpretation

$$I_0 \;=\; \int_0^\infty \operatorname{sinc} x\, dx \;=\; \lim_{a\to\infty} \int_0^a \operatorname{sinc} x\, dx.$$

Thus we interpret it as an improper Riemann integral, or at best as the limit of Lebesgue integrals.

Of course, *Maple* and *Mathematica* directly evaluate $I_0 = \pi/2$, but this is not helpful for those who demand understanding or a proof. Where does this evaluation come from? Peering behind the covers, we find that *Maple* knows

$$\int_0^a \operatorname{sinc} x \, dx \;=\; \operatorname{Si}(a) \;\to\; \frac{\pi}{2} \quad \text{for } a \to \infty.$$

However, this just shifts the problem to another level, since Si equals the integral by definition. We will now give several proofs of this identity: One will be short (and incomplete), one will be wrong, and one will be constructive!

For the first proof we just remember the Jordan theorem in Section 2.2.2, which directly implies that

$$\lim_{a\to\infty} \int_{-a}^{a} \operatorname{sinc}(\pi x)\, e^{ixt} \, dx \;=\; \frac{1}{2}\left(\chi_{(-\pi,\pi)}(t+) + \chi_{(-\pi,\pi)}(t-)\right),$$

so that $t = 0$ gives the desired evaluation. However, this is only a proof modulo the Jordan theorem. A direct proof would still be preferable.

The second "proof" is not a proof, just an idea: Write the sinc function as an inner integral and then use Fubini. Writing $1/x = \int_0^\infty e^{-tx}\, dt$, we would have to use Fubini on the function $g(t,x) = e^{-tx} \sin x$ on $\mathrm{R} \times \mathrm{R}$. The double integral that results from exchanging the integration order does, in fact, give

$$\int_0^\infty \int_0^\infty e^{-tx} \sin x \, dx \, dt \;=\; \int_0^\infty \frac{1}{1+t^2}\, dt \;=\; \frac{\pi}{2}.$$

However, this exchange is not allowed, since the integrand is not in $L^1(\mathrm{R}^2)$.

But this idea can now be made into a proof which is valid and constructive. If g is not L^1 on R^2, then we just have to restrict the domain of g at first. Now Fubini is applicable in

$$\begin{aligned}
\int_0^a \operatorname{sinc} x \, dx &= \int_0^a \int_0^\infty e^{-xt} \sin x \, dt \, dx \\
&= \int_0^\infty \int_0^a e^{-xt} \sin x \, dx \, dt \\
&= \int_0^\infty \frac{1}{1+t^2}\left[1 - e^{-at}(t \sin a + \cos a)\right] dt, \\
&= \frac{\pi}{2} - \int_0^\infty \frac{e^{-at}}{1+t^2}\,(t \sin a + \cos a)\, dt,
\end{aligned}$$

and the final integral goes to 0 for $a \to \infty$ as follows by elementary estimates (see [26]).

Another constructive method to evaluate the sinc integral is given in Item 26 in the Exercises at the end of this chapter.

2.5.2 Iterated Sinc Integrals

Now let $n \geq 1$ and $a_0, a_1, \cdots, a_n$ be positive reals. Our goal is to find inequalities (explicit formulas will be given in the exercises) for the integral

$$\tau = \int_0^\infty \prod_{k=0}^{n} \operatorname{sinc}(a_k\, x)\, dx,$$

which in particular explain the behavior of the integrals I_n. For simplicity and without loss, we can assume that $a_0 = 1$.

Theorem 2.5.1 *Let $s = \sum_{k=1}^n a_k$. If $s \leq 1$, then $\tau = \pi/2$; if $s > 1$, then $\tau < \pi/2$.*

Proof. Let $\tau(x) = \prod_{k=0}^n \operatorname{sinc}(a_k\, x)$. Note that $\operatorname{sinc}(a_k\, x) = \widehat{f_k}(x)$ with $f_k = (1/(2a_k))\, \chi_{[-a_k, a_k]}$. Thus by the convolution theorem, $\tau(x) = (f_0 * \cdots * f_n)\widehat{\ }(x)$, and by the inversion theorem,

$$\int_{-\infty}^{\infty} \tau(x)\, dx \;=\; 2\pi\, (f_0 * \cdots * f_n)(0) \;=\; 2\pi\, \frac{1}{2} \int_{-1}^{1} (f_1 * \cdots * f_n)(u)\, du. \quad (2.5.25)$$

Now since the support of f_k equals $[-a_k, a_k]$, the support of $(f_1 * \cdots * f_n)$ equals $[-s, s]$. If $s \leq 1$, then

$$\begin{aligned}\int_{-1}^{1} (f_1 * \cdots * f_n)(u)\, du \;&=\; \int_{-\infty}^{\infty} (f_1 * \cdots * f_n)(u)\, du \\ &=\; (f_1 * \cdots * f_n)\widehat{\ }(0) \;=\; \prod_{k=1}^{n} \operatorname{sinc}(a_k\, 0) \;=\; 1,\end{aligned}$$

and $\int_{-\infty}^{\infty} \tau(x)\, dx = \pi$ follows. If, on the other hand, $s > 1$, then the interval $[-1, 1]$ is strictly inside the support of $(f_1 * \cdots * f_n)$. Since $(f_1 * \cdots * f_n)$ is strictly positive in the interior of its support, we get

$$\int_{-1}^{1} (f_1 * \cdots * f_n)(x)\, dx \;<\; \int_{-\infty}^{\infty} (f_1 * \cdots * f_n)(x)\, dx \;=\; 1,$$

and $\int_{-\infty}^{\infty} \tau(x)\, dx < \pi$ follows. $\square$

This theorem explains why the values of I_n suddenly drop below $\pi/2$ at $n = 7$ and not before: We have $1/3 + 1/5 + \cdots + 1/13 < 1$; however, $1/3 + 1/5 + \cdots + 1/13 + 1/15 > 1$.

A geometric interpretation of this behavior can also be given. Consider the polyhedra

$$\begin{aligned} P &= \{(x_1, \cdots, x_n) : -1 \le \sum_{k=1}^{n} x_k \le 1,\ -a_k \le x_k \le a_k \text{ for } k = 1, \cdots, n\}, \\ Q &= \{(x_1, \cdots, x_n) : -1 \le \sum_{k=1}^{n} a_k x_k \le 1,\ -1 \le x_k \le 1 \text{ for } k = 1, \cdots, n\}, \\ H &= \{(x_1, \cdots, x_n) : -1 \le x_k \le 1 \text{ for } k = 1, \cdots, n\}. \end{aligned}$$

Then by formula (2.5.25),

$$\begin{aligned} \tau &= \frac{\pi}{2^n\, a_1 \cdots a_n} \int_0^{\min(1,s)} \left(\chi_{[-a_1,a_1]} * \cdots * \chi_{[-a_n,a_n]}\right)(x)\, dx \\ &= \frac{\pi}{2} \frac{\operatorname{Vol}(P)}{2^n\, a_1 \cdots a_n} = \frac{\pi}{2} \frac{\operatorname{Vol}(Q)}{\operatorname{Vol}(H)}. \end{aligned}$$

Thus the value of τ drops below $\pi/2$ precisely when the constraint $-1 \le \sum a_k x_k \le 1$ becomes active and "bites" into the hypercube H.

Of course, the same methods will also work for infinite products. Consider the function

$$C(x) = \prod_{n=1}^{\infty} \cos\left(\frac{x}{n}\right)$$

which is continuous since the product is absolutely convergent. We are interested in the integral $\mu = \int_0^{\infty} C(x)\, dx$. High precision numerical evaluation of this highly oscillatory integral is possible, but by no means straightforward. We get

$$\int_0^{\infty} C(x)\, dx \approx 0.785380557298632873492583011467332524761,$$

while $\pi/4 \approx 0.785398$ only differs in the fifth significant place. Can this numerical evaluation $\mu < \pi/4$ be confirmed symbolically? Indeed it can, by reduction to a sinc integral of the above type, only this time with an infinite product. Recall the sine product (1.2.11) and note that a corresponding product for the cosine can be derived:

$$\sin(x) \;=\; x\prod_{n=1}^{\infty}\left(1-\frac{x^2}{\pi^2\,n^2}\right), \qquad \cos(x) \;=\; \prod_{k=0}^{\infty}\left(1-\frac{4x^2}{\pi^2\,(2k+1)^2}\right).$$

Using this result, and exchanging the order of multiplication, we obtain

$$C(x) \;=\; \prod_{n=1}^{\infty}\prod_{k=0}^{\infty}\left(1-\frac{4x^2}{\pi^2\,n^2\,(2k+1)^2}\right) \;=\; \prod_{k=0}^{\infty}\operatorname{sinc}\left(\frac{2x}{2k+1}\right).$$

Now apply the theorem to get that

$$\mu \;=\; \int_0^{\infty} C(x)\,dx \;=\; \lim_{N\to\infty}\int_0^{\infty}\prod_{k=0}^{N}\operatorname{sinc}\left(\frac{2x}{2k+1}\right)dx \;<\; \frac{\pi}{4}.$$

This remarkable observation was made by Bernard Mares, then 17, and led to the entire development that we have given of the iterated sinc integrals. More examples are given in the Exercises. There is an interesting connection with random harmonic series in [192].

2.6 Korovkin's Three Function Theorems

In 1953, Pavel Korovkin [153] provided an approach to uniform approximation results that is especially well suited to computational assistance and discovery. While the result can be given much more generally, we limit ourselves to the two most basic cases.

In what follows, by ι we denote the identity function $t \mapsto t$; by $C[0,1]$ we denote continuous functions on the unit interval; and by $\rightrightarrows$ we denote uniform convergence (i.e., in the supremum norm). The interval $[0,1]$ can easily be replaced by any finite interval $[a,b]$.

Recall that an operator between continuous function spaces is positive if it maps nonnegative functions to nonnegative functions (when linear, this is necessarily a monotone and bounded linear operator). The motivating example is

Example 2.6.1 *Bernstein operators.*

For n in N, let $\mathcal{B}_n(f)$ be defined by

$$\mathcal{B}_n(f)(t) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} t^k (1-t)^{n-k}. \tag{2.6.26}$$

It is clear that the Bernstein operators are linear and positive on C$[0,1]$ and indeed take values which are polynomials. □

Theorem 2.6.2 (First Korovkin three function theorem). *Let L_n be a sequence of positive linear operators from $C[0,1]$ to $C[0,1]$. Suppose that*

$$L_n(1) \rightrightarrows_n 1, \quad L_n(\iota) \rightrightarrows_n \iota, \quad L_n(\iota^2) \rightrightarrows_n \iota^2.$$

Then

$$L_n(f) \rightrightarrows_n f$$

as $n \to \infty$ for all f in $C[0,1]$.

Proof. The hypotheses imply that $L_n(q) \rightrightarrows_n q$ for all quadratic q. Fix f in $C[0,1]$, x in $[0,1]$, and $\varepsilon > 0$. We claim that one can find a quadratic q_x^ε with $f \leq q_x^\varepsilon$ and $f(x) + \varepsilon \geq q_x^\varepsilon$. Thus

$$L_n(f) \leq L_n\left(q_x^\varepsilon\right) \rightrightarrows_n q_x^\varepsilon \leq f(x) + \varepsilon.$$

A compactness argument completes the proof. The details are left for the reader as Exercise 33. □

Corollary 2.6.3 (Stone-Weierstrass). *The Bernstein polynomials are uniformly dense in $C[0,1]$.*

Proof. We check by hand or in a computer algebra system that $\mathcal{B}_n(1) = 1$, $\mathcal{B}_n(\iota) = \iota$, and slightly more elaborately, $\mathcal{B}_n(\iota^2) = \iota^2 + \frac{1}{n}\left(\iota - \iota^2\right) \rightrightarrows_n \iota^2$. □

In the periodic case, the role of t and t^2 is taken by sin and cos, as the second Korovkin theorem shows.

Theorem 2.6.4 (Second Korovkin three function theorem). *Let L_n be a sequence of positive linear operators from $C(\mathrm{T})$ to $C(\mathrm{T})$. Suppose that*

$$L_n(1) \rightrightarrows_n 1, \quad L_n(\sin) \rightrightarrows_n \sin, \quad L_n(\cos) \rightrightarrows_n \cos .$$

Then

$$L_n(f) \rightrightarrows_n f$$

as $n \to \infty$ for all f in $C(\mathrm{T})$.

The great virtue of the Korovkin approach is that it provides us with a well-formed program. We illustrate with the second theorem. For any kernel (K_n), we may induce a sequence of linear operators $\mathcal{K}_n(f) = K_n * f$ and must answer two questions: (i) Is each $\mathcal{K}_n$ positive? (ii) Does $\mathcal{K}_n(f) \rightrightarrows_n f$ for the three functions $f = 1, \sin, \cos$? The first is usually easy to answer; the second is frequently a direct computation.

Example 2.6.5 *Dirichlet and Fejér Operators.*

We revisit the uniform convergence properties of the Dirichlet and Fejér kernels from Section 2.3.3.

1. The Dirichlet kernel induces the operator

$$\mathcal{D}_n(f) = D_n * f$$

where $D_n = \sin((n + 1/2)t) / \sin(t/2)$. This is quite easily seen not to be positive. This is a good thing, since we know that $\mathcal{D}_n(f)$ might not converge uniformly to f, for f in $C(\mathrm{T})$.

2. The Fejér kernel induces the operator

$$\mathcal{F}_n(f) = F_n * f$$

where $F_n = \sin^2((n+1)/2)t) / [(n+1)\sin^2(t/2)] \geq 0$. Thus, to recover Fejér's theorem on the uniform convergence of the Cesàro averages, it suffices to compute

$$\mathcal{F}_n(1) = 1, \quad \mathcal{F}_n(\sin) = \frac{n}{n+1} \sin, \quad \mathcal{F}_n(\cos) = \frac{n}{n+1} \cos . \qquad \Box$$

2.7 Commentary and Additional Examples

1. **An error function evaluation.** (*Monthly* Problem 11000, Mar. 2003) [154].

   (a) Work out the ordinary generating function of $\binom{2n}{n}$, and so evaluate

   $$\sum_{n+m=N} \binom{2n}{n} \binom{2m}{m}.$$

   (b) Evaluate

   $$\int_0^{\frac{\pi}{2}} \cos^{2n}(x) \sin^{2m}(x)\, dx.$$

   (c) Recall the *error function*, $\operatorname{erf}(x) = (2/\sqrt{\pi}) \int_0^x \exp(-t^2)\, dt$, and show for $a > 0$ that

   $$a \int_0^{\frac{\pi}{2}} \operatorname{erf}\left(\sqrt{a}\cos x\right) \operatorname{erf}\left(\sqrt{a}\sin x\right) \sin(2x)\, dx = e^{-a} + a - 1.$$

   (d) The previous evaluation can be viewed as an inner product of the functions $\operatorname{erf}(\sqrt{a}\sin x)\sin x$ and $\operatorname{erf}(\sqrt{a}\cos x)\cos x$. Determine that

$$\begin{aligned}
&\int_0^{\pi/2} \operatorname{erf}^2\left(\sqrt{a}\cos x\right)\cos^2(x)\; dx \\
&= \sum_{N=0}^{\infty} \frac{\left(\frac{-a}{4}\right)^{N+1}(8N+12)\binom{2N}{N}}{(N+2)!} \mathrm{F}\left(\frac{1}{2}, -N, -N-\frac{1}{2}; \frac{3}{2}, -N+\frac{1}{2}; -1\right) \\
&= \frac{1}{2\pi} - 2\int_0^1 \frac{e^{-1/2\,a\left(1+x^2\right)}}{1+x^2}\left\{\mathrm{I}_0\left(\frac{1}{2}a\left(1+x^2\right)\right) - \mathrm{I}_1\left(\frac{1}{2}a\left(1+x^2\right)\right)\right\} dx.
\end{aligned}$$

2. **Failure of Fubini.** Evaluate these integrals:

   (a)

   $$\int_0^1 \int_0^1 \frac{x^2 - y^2}{\left(x^2+y^2\right)^2} dx\, dy = -\frac{\pi}{4}$$

   and

   $$\int_0^1 \int_0^1 \frac{x^2 - y^2}{\left(x^2+y^2\right)^2} dy\, dx = \frac{\pi}{4}$$

(b)

$$\int_0^1 \int_1^\infty \left(e^{-xy} - 2\,e^{-2\,xy}\right) dy\,dx - \int_1^\infty \int_0^1 \left(e^{-xy} - 2\,e^{-2\,xy}\right) dx\,dy = \ln 2$$

(c)

$$\int_0^\infty \int_0^\infty \frac{4\,xy - x^2 - y^2}{(x+y)^4} dy\,dx = \int_0^\infty \int_0^\infty \frac{4\,xy - x^2 - y^2}{(x+y)^4} dx\,dy = 0$$

but, for all $m, c > 0$

$$\int_0^{mc} \int_0^c \frac{4\,xy - x^2 - y^2}{(x+y)^4} dx\,dy = \frac{m}{(1+m)^2} \not\to 0,$$

as $c \to \infty$.

In each case, explain why they differ without violating any known theorem.

3. **Failure of l'Hôpital's rule.** Evaluate these limits:

   Let $f(x) = x + \cos(x)\sin(x)$ and $g(x) = e^{\sin(x)}\,(x + \cos(x)\sin(x))$.

   Then $\lim_{x\to\infty} f(x)/g(x)$ does not exist although $\lim_{x\to\infty} f'(x)/g'(x) = 0$. This is a caution against carelessly dividing by zero!

4. **Various Fourier series evaluations.**

   (a) Compute the Fourier series of $t/2$, $|t|$, t^2, and $(t^3 - \pi^2\,t)/3$ on $[-\pi, \pi]$.

   (b) Plot the 6th and 12th Fourier polynomials against the function in each case.

   (c) Compute enough Fourier coefficients of $\sin(x^3)$ on $[-\pi, \pi]$ to be convinced of Parseval's equation.

   (d) Compute the Fourier series of t^2 and $(t^3 - \pi^2\,t)/3$ on $[0, 2\,\pi]$.

   (e) Use Parseval's equation with $(t^3 - \pi^2\,t)/3$ to evaluate $\zeta(6)$. Then apply Parseval to $t^4/4$.

   (f) Show that

   $$\int_0^{\pi/2} \log\left(2\,\sin(t/2)\right)\,dt = -G,$$

   where G is Catalan's constant.

(g) Show that for $a > 0$,

$$\cos(ax) = \frac{\sin(\pi a)}{\pi a} - 2\,\frac{\sin(\pi a)\,a\cos(x)}{(a^2-1)\,\pi} + 2\,\frac{\sin(\pi a)\,a\cos(2x)}{(a^2-4)\,\pi} - 2\,\frac{\sin(\pi a)\,a\cos(3x)}{(a^2-9)\,\pi} + \cdots. \tag{2.7.27}$$

Similarly evaluate the Fourier series for $\exp(ax)$.

(h) Substitute $x = \pi$ in (2.7.27) to obtain the partial fraction expansion for cot (compare the first example in Section 2.3.1) and integrate to recover the product formula for sin (justifying all steps).

(i) Evaluate $\sum_{n\geq 0} 1/(4n^2-1)$.

5. **Two applications of Parseval's equation.** Use Parseval's equation in $L^2(\mathrm{R})$ to evaluate

$$\int_{-\infty}^{\infty} \frac{\sin^2(t)}{t^2}\,dt = \pi \qquad \text{and} \qquad \int_{-\infty}^{\infty} \frac{\sin^4(t)}{t^4}\,dt = \frac{2\pi}{3}.$$

See also Exercise 28.

6. **Lebesgue's function.** An example of a continuous function with divergent Fourier series.

Construction. We follow Stromberg, page 557, and let $a_k = 2^{\sum_{j=1}^{k} j!}$ for $k \geq 0$ and define

$$f_n(x) = \sum_{k=1}^{n} \frac{\sin(a_k|x|)}{k}\,\chi_k(|x|)$$

on $[-\pi, \pi]$, where χ_k is the characteristic function of $[\pi/a_k, \pi/a_{k-1}]$, and extend f by 2π-periodicity onto R. Then $f(x) = \lim_{n\to\infty} f_n(x)$ is continuous and the Fourier series is uniformly convergent on $[\delta, 2\pi-\delta]$ for $\delta > 0$, but $s_{a_k}(f, 0) \to \infty$ for $k \to \infty$.

Convergence on $[\delta, 2\pi-\delta]$. This is easy since $f = f_n$ on this interval for large n, so that the "Riemann localization principle" for Fourier series can be used: *If $f_1(t) = f_2(t)$ for every t in some nonvoid open interval I, then $|s_n(f_1,t) - s_n(f_2,t)| \to 0$ for every $t \in I$.*

Divergence at 0. The divergence estimate comes as follows:

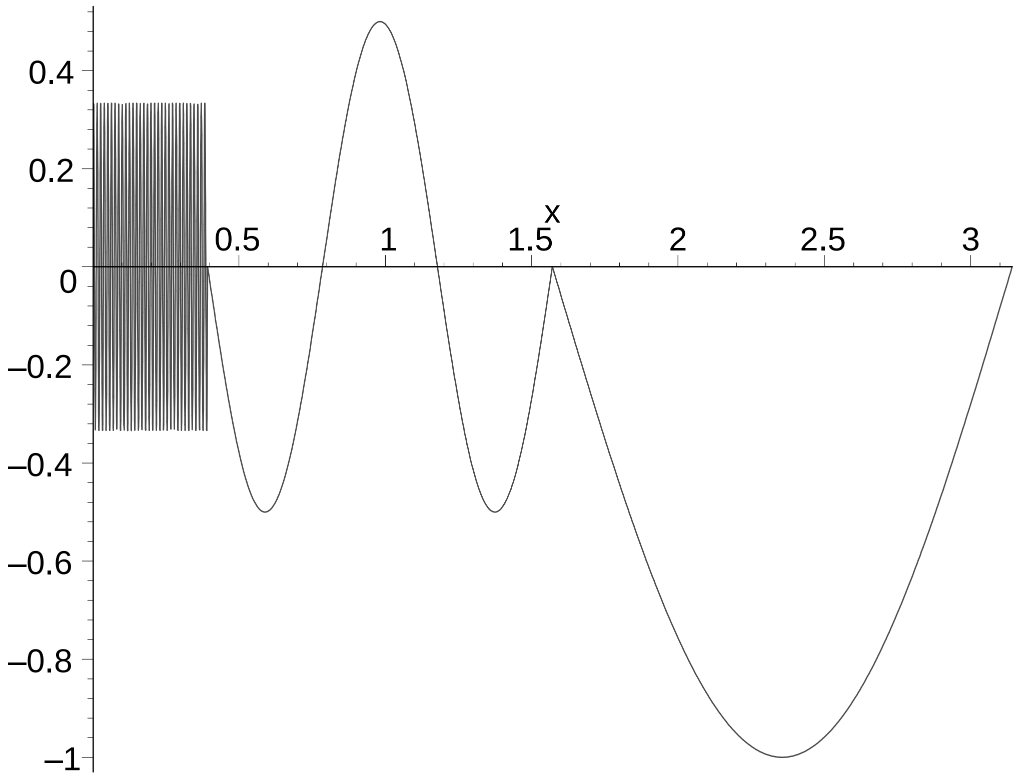

Figure 2.7: Approximation f_3 to Lebesgue's function.

(a) We start with Dirichlet's kernel: It can be proved that the n-th partial sum behaves like

$$s_n(f,0) \;=\; \frac{2}{\pi}\int_0^{\pi} f(t)\,\frac{\sin(nt)}{t}\,dt + \varepsilon_n$$

where $\varepsilon_n \to 0$.

(b) We estimate the first part of the integral as

$$\left|\int_0^{\pi/a_k} f(t)\,\frac{\sin(a_k t)}{t}\,dt\right| \;\leq\; \frac{\pi}{k+1},$$

since $|(\sin x)/x| \leq 1$ and $|f(t)| \leq 1/(k+1)$.

(c) We estimate the second part of the integral via

$$2\int_{\pi/a_k}^{\pi/a_{k-1}} f(t)\frac{\sin(a_k t)}{t}\,dt + \frac{1}{k}\int_{\pi/a_k}^{\pi/a_{k-1}} \frac{\cos(2\,a_k t)}{t}\,dt = k!\,\frac{\ln 2}{k}, \quad (2.7.28)$$

and the second term on the left, say I_k, is no bigger than $\frac{1}{2\pi k}$ on using the Bonnet second mean value theorem to write

$$|I_k| \;=\; \left|\frac{a_k}{k\pi}\int_{\pi/a_k}^{\psi}\cos(2\,a_k t)\,dt\right| \;\leq\; \frac{1}{2k\pi},$$

for some $\frac{a_{k-1}}{\pi} \leq \psi \leq \frac{a_k}{\pi}$.

(d) We estimate the third part of the integral as

$$\begin{aligned}\left|\int_{\pi/a_{k-1}}^{\pi} f(t)\,\frac{\sin(a_k t)}{t}\,dt\right| \;&\leq\; \frac{a_{k-1}}{\pi}\left|\int_{\pi/a_{k-1}}^{\psi} f(t)\sin(a_k t)\,dt\right| \\ &\leq\; \frac{a_{k-1}}{\pi}\left(\left|f(t)\,\frac{\cos(a_k t)}{a_k}\right|_{\pi/a_{k-1}}^{\psi} + \left|\int_{\psi}^{\pi} f'(t)\,\frac{\cos(a_k t)}{a_k}\,dt\right|\right) \\ &\leq\; \frac{a_{k-1}}{a_k}\left(\frac{1}{\pi} + a_{k-1}\right) \to 0,\end{aligned}$$

on using the mean value theorem again, and then applying integration by parts with the estimates that $|f(t)| \leq 1$, $|f'(t)| \leq a_{k-1}$. Thus, the dominant term is the first integral in (2.7.28) and $s_{a_k}(f,0) \to \infty$.

7. **Nonuniqueness of Fourier series.** Can a trigonometric series converge a.e. on R to a function $\varphi \in L^1(T)$ and yet not be the Fourier series of φ? This question was first answered in the affirmative with $\varphi = 0$ in 1916 by the Russian analyst D. E. Menshow. His counterexample involves the Cantor set. For more details, see [203], from which the above text was cited.

8. **Nowhere differentiable continuous functions.** The first and most famous example of a continuous, nowhere differentiable function was constructed by K. Weierstrass in 1872. (Tradition has it that Bolzano and Riemann constructed such examples before Weierstrass, but their examples did not become widely known.) Weierstrass's example was given in

the form of a trigonometric series. We state it here as a series on $[0,1]$, not on $[-\pi,\pi]$ or $[0,2\pi]$ as before, because this will simplify matters later; we analogously write $f \in L^1(0,1)$, and formulas and theorems on Fourier series are easily converted to this case. Weierstrass's example was

$$C_{a,b}(t) \;=\; \sum_{n=0}^{\infty} a^n \cos(b^n\, 2\pi t),$$

with $|a| < 1$ and integral $b > 1$. Weierstrass proved that $C_{a,b}$ is nowhere differentiable when $b \in 2\mathrm{N}+1$ and $ab > 1 + 3\pi/2$. This result settled, once and for all, the question of whether such functions could exist (at the beginning of the 19th century, Ampére "proved" that every continuous function must be differentiable at some point). Some questions were left open, however: It is clear that $C_{a,b}$ is differentiable when $|a|\, b < 1$, since the series is then termwise differentiable. But what happens for $|a|\, b$ between 1 and $1 + 3\pi/2$? This question gave several mathematicians a headache, until in 1916, G. H. Hardy proved that both $C_{a,b}$ and the corresponding sine series

$$S_{a,b}(t) \;=\; \sum_{n=0}^{\infty} a^n \sin(b^n\, 2\pi t)$$

are nowhere differentiable whenever b is a real greater than 1, and $ab > 1$. In his paper, Hardy first treated the case when $b \in \mathrm{N}$, i.e., when the functions are given by their Fourier series, and only afterwards treated the general case of arbitrary real b. Hardy's methods were not easy, not even in the Fourier case (where he used the Poisson kernel, among other things). In the ensuing years, several other, simpler proofs have been published. In the middle of the 20th century, G. Freud and J.-P. Kahane gave conditions for the differentiability of lacunary Fourier series (where nonzero Fourier coefficients are spaced far apart), from which the nondifferentiability of Weierstrass's function follows. Another approach to the Weierstrass functions uses functional equations.

Prove:

(a) The Weierstrass sine series $f = S_{a,2}$ with $b = 2$ satisfies the system of two functional equations

$$f\left(\frac{t}{2}\right) \;=\; a\, f(t) + \sin(t), \quad f\left(\frac{t+1}{2}\right) \;=\; a\, f(t) - \sin(t) \quad (2.7.29)$$

for every $t \in [0,1]$. The cosine series $S_{a,2}$ satisfies an analogous system.

(b) The Weierstrass function is the only bounded solution of the respective system on $[0,1]$.

Hint for (b): Use Banach's fixed point theorem.

9. **Replicative functions.** Let D be an interval containing $(0,1)$. A function $f : D \to \mathrm{C}$ is called *replicative* (on D) if it satisfies the functional equation

$$\frac{1}{p}\sum_{k=0}^{p-1} f\left(\frac{t+k}{p}\right) \;=\; u(p)\, f(t) \quad \text{for all } t \in D \text{ and } p \in \mathrm{N}, \tag{2.7.30}$$

with a $u : \mathrm{N} \to \mathrm{C}$ (which turns out to be unique if $f \not\equiv 0$). This notion was introduced (with more generality) by D. E. Knuth in [152]. Examples are the cotangent ($\cot(\pi t)$ is replicative on $(0,1)$ with $u(p) = 1$), the Bernoulli polynomials ($B_m(t)$ is replicative on R with $u(p) = 1/p^m$), and derivatives of the Psi function (the m-th derivative of $\Psi = \Gamma'/\Gamma$ is replicative on R_+ with $u(p) = p^m$). Functions that are replicative and 1-periodic have multiplicative Fourier coefficients.

Theorem 2.7.1 *(a) Let $f : D \to \mathrm{C}$, $f \not\equiv 0$, be replicative on D with $u(p)$. Then u is necessarily multiplicative, i.e., $u(mn) = u(m)\cdot u(n)$ for all $m, n \in \mathrm{N}$.*

(b) Let $f \in L^1(0,1)$ be replicative on $(0,1)$ with a sequence $u(p)$. Then $\widehat{f}_{mn} = u(n)\widehat{f}_m$.
This implies: If $u \equiv 1$, then $f(t) = \widehat{f}_0$. If $u \not\equiv 1$, then

$$f(t) \;\sim\; \widehat{f}_{-1}\sum_{n=-\infty}^{-1} u(-n)\, e^{2\pi i n t} + \widehat{f}_1 \sum_{n=1}^{\infty} u(n)\, e^{2\pi i n t}.$$

(c) Let u be multiplicative and assume that $f(t) = {}^{L}\sum_{n=1}^{\infty} u(n)\, e^{2\pi i n t}$ is pointwise convergent on $[0,1]$, where L is a linear summation method. Then f is replicative on $[0,1]$.

Part (c) of this theorem makes it easy to construct many different examples of replicative functions on $[0,1]$. Verify the following Fourier series:

(a) $\sum n^{-2}\sin(2\pi nt) = -2\pi\int_0^t \ln(2\sin(\pi x))\,dx$ on $[0,1]$,
$\sum n^{-2}\cos(2\pi nt) = B_2(t)$ on $[0,1]$.

(b) $\sum n^{-1}\sin(2\pi nt) = B_1(t)$ on $(0,1)$ and $=0$ on $0,1$,
$\sum n^{-1}\cos(2\pi nt) = -\ln(2\sin(\pi t))$ on $(0,1)$ and $=\infty$ on $0,1$.

(c) ${}^{C_1}\sum \sin(2\pi nt) = \frac{1}{2}\cot(\pi t)$ on $(0,1)$ and $=0$ on $0,1$,
${}^{C_1}\sum \cos(2\pi nt) = -\frac{1}{2}$ on $(0,1)$ and $=\infty$ on $0,1$,
where C_1 stands for Cesàro summation.

(d) ${}^{A}\sum n\sin(2\pi nt) = 0$ on $[0,1]$,
${}^{A}\sum n\cos(2\pi nt) = -1/(4\sin^2(\pi t))$ on $(0,1)$ and $=\infty$ on $0,1$,
where A stands for Abel summation.

(e) $\sum_{n=0}^{\infty} a^n \sin(p^n\, 2\pi t) = S_{a,p}(t)$ on $[0,1]$,
$\sum_{n=0}^{\infty} a^n \cos(p^n\, 2\pi t) = C_{a,p}(t)$ on $[0,1]$,
for p a prime, i.e., the nowhere differentiable Weierstrass functions can be replicative.

10. **Conditions for integrability.**

(a) As in Theorem 2.3.7, it is often useful to decide whether $\widehat{f} \in L^1(\mathrm{R})$ for a given $f \in L^1(\mathrm{R})$, without having to explicitly compute $\widehat{f}$. The usual conditions assume differentiability properties of f, since smoothness of f translates into shrinkage of $\widehat{f}$. Thus, $f \in C^2(\mathrm{R})$ is sufficient for $\widehat{f} \in L^1(\mathrm{R})$. However, this condition does not cover the Fejér kernel via $g(x) = \max\{1-|x|, 0\}$, for example. A stronger condition, which is good for functions with bounded support, is given in the next theorem.

Theorem 2.7.2 *Let f be an absolutely continuous function on the real line with compact support and let f' be of bounded total variation on* R, *i.e., $V(f') < \infty$. Then $\widehat{f} \in L^1(\mathrm{R})$ and*

$$\|\widehat{f}\|_1 \;\leq\; 4\sqrt{V(f')\,\|f\|_1}. \tag{2.7.31}$$

This theorem presents another experimental challenge: Is the constant "4" appearing there best possible? The answer is not known. Nonsystematic experimentation has found no value for the constant greater than π, which is achieved for the Fejér kernel. It is also not known if the "compact support" condition in the theorem is really needed.

Perform a systematic experiment on Theorem 2.7.2, in analogy to the experimentation for the uncertainty principle described in Section 5.2 of the first volume.

(b) A quite different condition is due to Chandrasekharan: *If* $f \in L^1(\mathrm{R})$, *continuous at* 0, *and satisfies* $\widehat{f} \geq 0$ *on* R, *then* $\widehat{f} \in L^1(\mathrm{R})$. The disadvantage of this condition is that it uses $\widehat{f}$ explicitly. It is applicable, however, to both the Fejér and the Poisson kernel.

11. **More kernels.** For each of the following kernels, decide whether (respectively, for which parameters) it is an approximate identity in $L^1(\mathrm{R})$. Note that sometimes a version of Theorem 2.3.7 with weakened assumptions (allowing more variety in the kernels) is needed.

(a) The *de la Vallée-Poussin kernel* V_m^n, depending on two integer parameters m, n with $n > m$, is defined by $V_m^n(t) = \sum_{k=-(n+m)}^{n+m} a_{m,k}^n \, e^{ikt}$, where

$$a_{m,k}^n = \begin{cases} 1, & \text{if } |k| \leq n-m, \\ \frac{n+m+1-|k|}{2m+1}, & \text{if } n-m \leq |k| \leq n+m, \\ 0, & \text{otherwise.} \end{cases}$$

Hint: Let m, n tend to infinity such that $n/m \to \lambda$.

(b) For $\alpha > 0$, the (C, α)-*kernel* $F_n^{(\alpha)}$ is defined as $F_n^{(\alpha)}(t) = \sum_{k=-n}^{n} a_{n,k}^{(\alpha)} \, e^{ikt}$ where

$$a_{n,k}^{(\alpha)} = \begin{cases} \frac{\Gamma(n-|k|+\alpha+1)\,\Gamma(n+1)}{\Gamma(n-|k|+1)\,\Gamma(n+\alpha+1)}, & \text{if } |k| \leq n+1, \\ 0, & \text{otherwise.} \end{cases}$$

(c) For parameters $\alpha_0, \cdots, \alpha_p \in \mathrm{R}$ with $\alpha_0 + \cdots + \alpha_p = 1$, the *Blackman kernel* $H_n^{(\alpha_0,\cdots,\alpha_p)}$ is defined by $H_n^{(\alpha_0,\cdots,\alpha_p)}(t) = \sum_{k=-n}^{n} h_{n,k}^{(\alpha_0,\cdots,\alpha_p)} e^{ikt}$, where

$$h_{n,k}^{(\alpha_0,\cdots,\alpha_p)} = \sum_{j=0}^{p} \alpha_j \cos(jkt_n),$$

with $t_n = 2\pi/(2n+1)$.

(d) The *Fejér-Korovkin kernel* FK_n is defined as

$$FK_n(t) = \begin{cases} \dfrac{2\sin^2(\pi/(n+2))}{n+2} \left[\dfrac{\cos((n+2)t/2)}{\cos(\pi/(n+2)) - \cos t}\right]^2, \\ (n+2)/2, \end{cases}$$

depending on whether $t \neq \pm\pi/(n+2) + 2j\pi$ or $t = \pm\pi/(n+2) + 2j\pi$, respectively. It can be written in the form $FK_n(t) = \sum_{k=-n}^{n} a_{n,k}\, e^{ikt}$ where

$$a_{n,k} = \frac{(n-|k|+3)\sin\frac{|k|+1}{n+2}\pi - (n-|k|+1)\sin\frac{|k|-1}{n+2}\pi}{2(n+2)\sin(\pi/(n+2))}.$$

(e) Finally, the *Jackson kernel* J_n is a rescaled version of the square of the Fejér kernel, namely

$$J_n(t) = \frac{3}{n(2n^2+1)} \left[\frac{\sin(nt/2)}{\sin(t/2)}\right]^4.$$

12. **The Haar basis.** As we mentioned in the text, the trigonometric functions e^{int} constitute an orthogonal basis for the space $L^2(\mathrm{T})$, so that L^2-statements follow from general Hilbert space theory. Bases other than the trigonometric are, of course, conceivable and in fact are used for the analysis of L^2-functions. Since about 15 years ago, certain bases of $L^2(\mathrm{R})$, called *wavelet bases*, have found widespread use in signal analysis. Such bases are constructed as follows. Take a $\psi \in L^2(\mathrm{R})$ and define $\psi_{j,n}(t) = 2^{n/2}\,\psi(2^n t - j)$. Then ψ is called an *orthogonal wavelet* if $\{\psi_{j,n} : j, n \in \mathrm{Z}\}$ is an orthonormal basis of $L^2(\mathrm{R})$.

    Show: $\psi = \chi_{[0,1/2)} - \chi_{[1/2,1)}$ is an orthogonal wavelet. The associated basis $\{\psi_{j,n}\}$ is called the *Haar basis* of $L^2(\mathrm{R})$.

13. **The Schauder basis.** The foundation of the theory of bases in Banach spaces was laid by J. Schauder in the 1930s. A sequence (x_n) in a Banach space B is called a *basis of* B if for every $x \in B$, there exists a unique sequence of scalars (α_n) with

$$x = \sum_{n=1}^{\infty} \alpha_n\, x_n \quad \text{in } B.$$

The trigonometric functions are not a basis for $L^1(\mathrm{T})$ or for $C(\mathrm{T})$, although their span is dense in these spaces. The standard example of a basis for the space $C[0,1]$ is also due to Schauder (although G. Faber had used the same basis before Schauder, in a different analytical guise). This Faber-Schauder basis is the system of continuous functions $\{\sigma_{0,0}, \sigma_{1,0}\} \cup \{\sigma_{i,n} : n \in \mathrm{N},\ i = 0, \cdots, 2^{n-1} - 1\}$, where $\sigma_{0,0}(t) = 1 - t$, $\sigma_{1,0}(t) = t$, and the function $\sigma_{i,n}$ is the linear interpolation of the points

$$(0,0), \quad \left(\frac{i}{2^{n-1}}, 0\right), \quad \left(\frac{2i+1}{2^n}, 1\right), \quad \left(\frac{i+1}{2^{n-1}}, 0\right), \quad (1,0).$$

This system is a basis of the space $C[0,1]$, more precisely: *Every continuous function $f : [0,1] \to \mathrm{R}$ has a unique, uniformly convergent expansion of the form*

$$f(x) \;=\; \gamma_{0,0}(f)\,\sigma_{0,0}(x) + \gamma_{1,0}(f)\,\sigma_{1,0}(x) + \sum_{n=1}^{\infty} \sum_{i=0}^{2^{n-1}-1} \gamma_{i,n}(f)\,\sigma_{i,n}(x),$$

where the coefficients $\gamma_{i,n}(f)$ are given by $\gamma_{0,0}(f) = f(0)$, $\gamma_{1,0}(f) = f(1)$, and

$$\gamma_{i,n}(f) \;=\; f\left(\frac{2i+1}{2^n}\right) - \frac{1}{2} f\left(\frac{i}{2^{n-1}}\right) - \frac{1}{2} f\left(\frac{i+1}{2^{n-1}}\right).$$

Knowing the Schauder basis expansion of a continuous function f can be useful in the analysis of f. For example, Faber proved in 1910 a criterion for differentiability of f in terms of its Schauder coefficients: *If $f'(x_0) \in \mathrm{R}$ exists for some $x_0 \in [0,1]$, then*

$$\lim_{n\to\infty} 2^n \cdot \min\{|\gamma_{i,n}(f)| : i = 0, \cdots, 2^{n-1} - 1\} \;=\; 0. \tag{2.7.32}$$

Interestingly, this condition can be used to prove nondifferentiability of the Weierstrass functions in an elementary way. Prove:

(a) The Schauder coefficients of the Weierstrass sine series $f = S_{a,2}$ satisfy the recursion
$\gamma_{0,1}(f) = 0$,
$\gamma_{i,n+1}(f) = a\gamma_{i,n}(f) + \gamma_{i,n}(\sin)$ for $n \in \mathrm{N}$, $i = 0, \cdots, 2^{n-1} - 1$,
$\gamma_{i,n+1}(f) = a\gamma_{i-2^{n-1},n}(f) - \gamma_{i-2^{n-1},n}(\sin)$ for $n \in \mathrm{N}$, $i = 2^{n-1}, \cdots, 2^n - 1$.
Hint: Use the functional equations (2.7.29).

(b) Faber's condition (2.7.32) is not satisfied for $f = S_{a,2}$. Thus this function is nowhere differentiable.

It is instructive to experiment with the recursion in (a): to plot the Schauder coefficients and Schauder approximations for the Weierstrass and for other functions that satisfy similar functional equations. Details of this method can be found in [121] and [122].

14. **Riemann-Lebesgue lemma.** Deduce the following from the Riemann-Lebesgue lemma for every Lebesgue integrable function f.

(a) For any real $\sigma(t)$

$$\lim_{t\to\infty} \int_{\mathrm{R}} f(x)\cos^2\left(tx+\sigma(t)\right)\, dx = \frac{1}{2}\int_{\mathrm{R}} f(x)\, dx.$$

(b) The coefficients $\hat{f}(n) \to 0$ as $n \to \infty$.

Conclude that the trigonometric series $\sum_{n>1} \sin(nt)\,/\log(n)$ is not the Fourier series of any integrable function.

When a_n is convex, decreasing with limit zero and with $\sum_{n>0} a_n/n = \infty$, it is in fact the case that $\sum_{n>0} a_n \cos(nt)$ is a Fourier series of an integrable function, but $\sum_{n>0} a_n \sin(nt)$ is not [203, Chapter 8].

15. **A few Fourier transforms.** We have already seen many examples of Fourier transforms and their Laplace transform variants. The specialization to the Mellin transform is explored in the next chapter.

(a) Show that the L^2-Fourier transform of $(\sin t^2)/t$ is the expression $i\pi\left(S(x/\sqrt{2\pi}) - C(x/\sqrt{2\pi})\right)$, where the functions C and S are the Fresnel integrals,

$$C(x) \;=\; \int_0^x \cos(\frac{\pi}{2}t^2)\, dt, \qquad S(x) \;=\; \int_0^x \sin(\frac{\pi}{2}t^2)\, dt.$$

Determine the Fourier transform of $\cos(t^2)/t$. Find "sensible" Fourier transforms for $\sin(t^2)$ and $\cos(t^2)$, even though these functions are neither in $L^1(\mathrm{R})$ nor in $L^2(\mathrm{R})$.

(b) Show that for $a > 0$, the Fourier transform of $|t|\exp(-a|t|)$ is the function $2\,(a^2 - x^2)\,/\,(a^2 + x^2)^2$.

(c) Find the transform of $1/\,(a^2 + t^2)$ and of $1/t^\eta$ (for suitable η, and in a suitable sense).

(d) Find all square-integrable solutions to $\hat{f}/\sqrt{2\pi} = f$ (the fixed points of the normalized Fourier transform). Then experiment with the orbit of $f \mapsto \hat{f}/\sqrt{2\pi}$ for various choices f_0.

16. **The isoperimetric inequality.** The ancient Greek geometers knew already that a circle with given perimeter encloses a larger area than any polygon with the same perimeter. In 1841 Steiner extended this result to simple closed plane curves. Here we will sketch a Fourier series proof (due to Hurwitz), for simplicity restricted to (piecewise) C^1-curves. Thus, assume that we have a simple, closed C^1-curve $(x(t), y(t))$ in R^2 of length $\int_{-\pi}^{\pi}(x'(s)^2 + y'(s)^2)^{1/2}\,ds = 2\pi$. Without loss of generality, we can assume that $x'(s)^2 + y'(s)^2 = 1$ for all s. We wish to minimize the area inside the curve, given by

$$A \;=\; \int_{-\pi}^{\pi} x(s)\,y'(s)\,ds.$$

Show: $A \geq \pi$ with equality if and only if $(x(t)-\widehat{x}_0)^2+(y(t)-\widehat{y}_0)^2 = 1$. This is the *isoperimetric inequality. Hint:* Substitute Fourier series, transfer the formulas for derivatives from Section 2.4.2 to the $L^1(\mathrm{T})$-case, and use Parseval's equation.

17. **The maximum principle.** In like fashion, employ Poisson's kernel to heuristically deduce that the maximum principle discussed briefly in Section 6.5 applies.

18. **The heat equation.** The one-dimensional heat equation

$$\frac{\partial\phi}{\partial t}(x,t) = \frac{\pi}{4i}\,\frac{\partial^2\phi}{\partial x^2}(x,t),$$

is solved by the general theta function $\sum_{n\in\mathrm{Z}} x^n \exp(-\pi i\,tn^2)$. More usefully, when $G : \mathrm{R} \mapsto \mathrm{C}$ is continuous and bounded we may solve the equation

$$\frac{\partial\phi}{\partial t}(x,t) = K\frac{\partial^2\phi}{\partial x^2}(x,t),$$

with boundary condition $\phi(x,t) \to G(x)$ as $t \to 0^+$, by the infinitely differentiable function

$$G * E_{1/\sqrt{2Kt}}(x) = \frac{1}{2\sqrt{\pi Kt}} \int_{\mathrm{R}} G(x-y)\exp(-y^2/2Kt)\,dy,$$

for x in R and $t > 0$.

19. **The easiest three-dimensional Watson integral.** We start with the easiest integral to evaluate. Let

$$W_3(w) = \int_0^\pi \int_0^\pi \int_0^\pi \frac{1}{1 - w\,\cos(x)\cos(y)\cos(z)}\,dx\,dy\,dz,$$

for suitable $w > 0$.

(a) Prove that

$$\begin{aligned} W_3(1) &= \int_0^\pi \int_0^\pi \int_0^\pi \frac{1}{1-\cos(x)\cos(y)\cos(z)}\,dx\,dy\,dz \\ &= \frac{1}{4}\Gamma^4\left(\frac{1}{4}\right) = 4\,\pi\,\mathrm{K}\left(\frac{1}{\sqrt{2}}\right) \end{aligned}$$

via the binomial expansion and [44, Exercise 14, page 188].

(b) More generally,

$$W_3((2kk')^2) = \pi^3\,\mathrm{F}\left(1/2,1/2,1/2;1,1;4\,k^2\left(1-k^2\right)\right) = 4\pi\,\mathrm{K}^2(k)\,.$$

20. **The harder three-dimensional Watson integrals.** We now describe results largely in Joyce and Zucker [146, 147], where more background can also be found. The following integral arises in Gaussian and spherical models of ferromagnetism and in the theory of random walks.

(a) One of the most impressive closed-form evaluations of a multiple integral is Watson's

$$\begin{aligned} W_1 &= \int_{-\pi}^\pi \int_{-\pi}^\pi \int_{-\pi}^\pi \frac{1}{3-\cos(x)-\cos(y)-\cos(z)}\,dx\,dy\,dz \\ &= \frac{1}{96}(\sqrt{3}-1)\,\Gamma^2\left(\frac{1}{24}\right)\Gamma^2\left(\frac{11}{24}\right) \qquad (2.7.33) \\ &= 4\,\pi\left(18+12\sqrt{2}-10\sqrt{3}-7\sqrt{6}\right)\mathrm{K}^2(k_6)\,, \end{aligned}$$

where $k_6 = \left(2-\sqrt{3}\right)\left(\sqrt{3}-\sqrt{2}\right)$ is the sixth singular value of Section 4.2. Note that $W_1 = \pi^3 \int_0^\infty \exp(-3t)\, I_0^3(t)\, dt$ allows for efficient computation [146] where the Bessel function $I_0(t)$ has been written as $(1/\pi)\int_0^\pi \exp(t\,\cos(\theta))\, d\theta$. The evaluation (2.7.33), in its original form, is due to Watson and is really a tour de force. In the next exercise we describe a refined and simplified evaluation due to Joyce and Zucker [147].

(b) Similarly, the integral

$$\begin{aligned} W_2 &= \int_0^\pi\int_0^\pi\int_0^\pi \frac{dx\,dy\,dz}{3-\cos(x)\cos(y)-\cos(y)\cos(z)-\cos(z)\cos(x)} \\ &= \sqrt{3}\,\pi\,\mathrm{K}^2\left(\sin\left(\frac{\pi}{12}\right)\right) = \frac{2^{1/3}}{8\,\pi}\,\beta^2\left(\frac{1}{3},\frac{1}{3}\right), \end{aligned} \tag{2.7.34}$$

where $\sin(\pi/12) = k_3$ is the third singular value, again as in Section 4.2. Indeed, as we shall see in Exercise 21, (2.7.34) is easier and can be derived on the way to (2.7.33).

(c) The evaluation (2.7.34) then implies that

$$\frac{1}{\pi}\,W_2 = \int_0^\pi\int_0^\pi \frac{dy\,dz}{\sqrt{9-8\,\cos(y)\cos(z)-\cos^2(y)-\cos^2(z)+\cos^2(y)\cos^2(z)}}$$

on performing the innermost integration carefully.

(d) The expression inside the square root factors as $(\cos x\cos y+\cos x+\cos y-3)(\cos x\cos y-\cos x-\cos y-3)$. Upon substituting $s = x/2$, and $t = y/2$, one obtains

$$\begin{aligned} &\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} \frac{dy\,dx}{\sqrt{\left(1-\sin^2(x)\sin^2(y)\right)\left(1-\cos^2(x)\cos^2(y)\right)}} \\ &= \frac{1}{4\pi}\sum_{m=0}^\infty\sum_{n=0}^\infty \beta^3\left(n+\frac{1}{2},m+\frac{1}{2}\right)\binom{m+n}{n} = \sqrt{3}\,\mathrm{K}^2\left(\sin\left(\frac{\pi}{12}\right)\right). \end{aligned}$$

21. **More about the Watson integrals.**

(a) For $a > 1, b > 1$, show that

$$\frac{1}{2}\int_0^{\pi} \frac{1}{\sqrt{(a+\cos(y))\,(b-\cos(y))}}\, dy = \frac{\mathrm{K}\left(\sqrt{\frac{2(b+a)}{(1+b)(1+a)}}\right)}{\sqrt{(1+b)\,(1+a)}}$$

$$= \int_0^{1/2\,\pi} \frac{1}{\sqrt{(1+b)\,(1+a)\cos^2(t) + (1-a)\,(1-b)\sin^2(t)}}\, dt.$$

(b) A beautiful, but harder to establish, identity is that

$$\int_0^{\frac{\pi}{2}} \mathrm{K}\left(\sqrt{c^2\cos^2(s)+\sin^2(s)}\right)\, ds = \mathrm{K}\left(\sqrt{\frac{1-c}{2}}\right)\mathrm{K}\left(\sqrt{\frac{1+c}{2}}\right), \tag{2.7.35}$$

or equivalently that

$$\int_0^{\frac{\pi}{2}} \mathrm{K}\left(\sqrt{1-(2kk')^2\,\cos^2(\theta)}\right)\, d\theta = \mathrm{K}\,(k)\,\mathrm{K}\,(k')$$

with $k' = \sqrt{1-k^2}$. Hence,

$$\int_0^{\frac{\pi}{2}} \mathrm{K}\left(\sqrt{1-(2k_Nk_N')^2\,\cos^2(\theta)}\right)\, d\theta = \sqrt{N}\,\mathrm{K}^2\,(k_N)\,,$$

where k_N is the N-th singular value. This is especially pretty for $N = 1, 3, 7$ so that $2\,k_N k_N' = 1, 1/2, 1/8$, respectively.

(c) Deduce that the face centered cubic (FCC) lattice for the Green's function evaluates as

$$\frac{1}{\pi}\,W_2 = \int_0^{\frac{\pi}{2}} \mathrm{K}\left(\sqrt{\frac{3}{4}\cos^2(s)+\sin^2(s)}\right)\, ds = \sqrt{3}\,\mathrm{K}^2\,(k_3)\,.$$

(d) Correspondingly, Watson's evaluation for the simple cubic (SC) lattice relied on deriving

$$W_1 = \sqrt{2}\,\pi\int_0^{\pi} \mathrm{K}\left(\frac{\cos(x)-5}{2}\right)\, dx,$$

and the following extension of (2.7.35):

$$\int_0^{\frac{\pi}{2}} \mathrm{K}\left(\sqrt{c^2\cos^2(s)+d^2\sin^2(s)}\right)ds = \mathrm{K}\left(\sqrt{\frac{1-cd-\sqrt{(d^2-1)(c^2-1)}}{2}}\right) \times \mathrm{K}\left(\sqrt{\frac{1+cd-\sqrt{(d^2-1)(c^2-1)}}{2}}\right).$$

(e) *The generalized Watson integrals.* Let

$$W_1(w_1) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\frac{1}{3-w_1\,(\cos(x)-\cos(y)-\cos(z))}\,dx\,dy\,dz$$

$$W_2(w_2) = \int_0^{\pi}\int_0^{\pi}\int_0^{\pi}\frac{dx\,dy\,dz}{3-w_2\,(\cos(x)\cos(y)-\cos(y)\cos(z)-\cos(z)\cos(x))}.$$

In a beautiful study, Joyce and Zucker [147], using the sort of elliptic and hypergeometric transformations we have explored, show fairly directly that

$$W_2(-w_1(3+zw_1)/(1-w_1)) = (1-w_1)^{1/2}\,W_1(w_1).$$

Verify that, with $w_1=-1$, this leads to a quite direct evaluation of (2.7.33) from (2.7.34).

(f) It is also true that

$$W_1 = \sqrt{2}\,\pi\int_0^{\frac{\pi}{2}}\mathrm{K}\left(\frac{1}{2}+\frac{1}{2}\sin^2(t)\right)dt.$$

(g) *A more symmetric form.* Show that

$$\int_0^{\frac{\pi}{2}}\mathrm{K}\left(\sqrt{1-4\,k^2\,(1-k^2)\cos^2(x)}\right)dx = \int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\frac{dt\,dx}{\sqrt{\cos^2(t)+4\,k^2(1-k^2)\cos^2(x)\sin^2(t)}}$$

for $0<k<1$.

Hint: For (21d) consider $N = 3$ ($c^2 = 3/4$) in (21c), and let a and b be defined as $a = (3 - \cos x) / (1 + \cos x)$ and $b = (3 + \cos x) / (1 - \cos x)$ in (21a).

22. **Watson integral and Burg entropy.** Consider the perturbed *Burg entropy* maximization problem

$$\begin{aligned} v(\alpha) = \sup_{p \geq 0} \{ \log(p(x_1, x_2, x_3)) \ | \int_0^1 \int_0^1 \int_0^1 p(x_1, x_2, x_3)\, dx_1 dx_2 dx_3 &= 1, \\ \text{and for } k = 1, 2, 3, \int_0^1 \int_0^1 \int_0^1 p(x_1, x_2, x_3) \cos(2\,\pi x_k)\ dx_1 dx_2 dx_3 &= \alpha \}, \end{aligned}$$

maximizing the log of a density p with given mean, and with the first three cosine moments fixed at a parameter value $0 \leq \alpha < 1$. It transpires that there is a parameter value $\overline{\alpha}$ such that below and at that value $v(\alpha)$ is attained, while above it is finite but unattained. This is interesting, because:

(a) The general method—maximizing $\int_T \log(p(t))\, dt$ subject to a finite number of trigonometric moments—is frequently used. In one or two dimensions, such spectral problems are always attained when feasible.

(b) There is no easy way to see that this problem qualitatively changes at $\overline{\alpha}$, but we can get an idea by considering

$$\overline{p}(x_1, x_2, x_3) = \frac{1/W_1}{3 - \sum_1^3 \cos(2\pi x_i)},$$

and checking that this is feasible for

$$\overline{\alpha} = 1 - 1/(3W_1) \approx 0.3405373295509991 42833$$

in terms of the first Watson integral, W_1.

(c) By using Fenchel duality [61] one can show that this $\overline{p}$ is optimal.

(d) Indeed, for all $\alpha \geq 0$ the only possible optimal solution is of the form

$$\overline{p}_\alpha(x_1, x_2, x_3) = \frac{1}{\lambda_\alpha^0 - \sum_1^3 \lambda_\alpha^i \cos(2\pi x_i)},$$

for some real numbers λ_α^i. Note that we have four coefficients to determine; using the four constraints we can solve for them. For $0 \le \alpha \le \overline{\alpha}$, the precise form is parameterized by the generalized Watson integral:

$$\overline{p}_\alpha\,(x_1, x_2, x_3) = \frac{1/W_1(w)}{3 - \sum_1^3 w \cos\,(2\pi x_i)},$$

and $\alpha = 1 - 1/(3W_1(w))$, as w ranges from zero to one. Note also that $W_1(w) = \pi^3 \int_0^\infty I_0^3(w\,t)\,e^{-3t}\,dt$ allows one to quickly obtain w from α numerically. For $\alpha > \overline{\alpha}$, no feasible reciprocal polynomial can stay positive. Full details are given in [60].

23. **A "momentary" recursion.** Choose $p \in \mathrm{N}$ and define polynomials $q_k(x)$ recursively by $q_0(x) = -1$ and $q_{n+1}(x) = q_n'(x) - x^{p-1}\,q_n(x)$. Give an explicit formula for $q_n(0)$ (and for $q_n(x)$).

24. **The limit of certain Fourier transforms.** For $p \in 2\mathrm{N}$, let $f_p(t) = e^{-t^p/p}$. In Figure 2.8, the functions $\widehat{f_p}(x)$ are shown for $p = 2, 8, 16$. The figure suggests that there may be a limit function as $p \to \infty$. Identify this limit function!

25. **The Schilling equation.** The Schilling equation is the functional equation

$$4q\,f(qt) \;=\; f(t+1) + 2f(t) + f(t-1) \quad \text{for } t \in \mathrm{R}$$

with a parameter $q \in (0, 1)$. It has its origin in physics, and although it has been studied intensively in recent years, there are still many open questions connected with it. The main question is to find values of q for which the Schilling equation has a nontrivial L^1-solution. Discuss this question! *Hint:* If an L^1-function f satisfies (2.4.20), then a rescaled version of $f * f$ satisfies the Schilling equation.

26. **Another way to evaluate the sinc integral.** The evaluation of the integral $\int_0^\infty \sin y/y\,dy = \pi/2$ also follows on taking the limit, via Binet's mean value theorem [203, page 328], of the absolutely convergent integral

$$\int_0^\infty \frac{\sin y}{y^{1+\varepsilon}}\,dy = \frac{\pi}{2}\,\frac{\sec(\frac{\pi}{2}\varepsilon)}{\Gamma(1+\varepsilon)}.$$

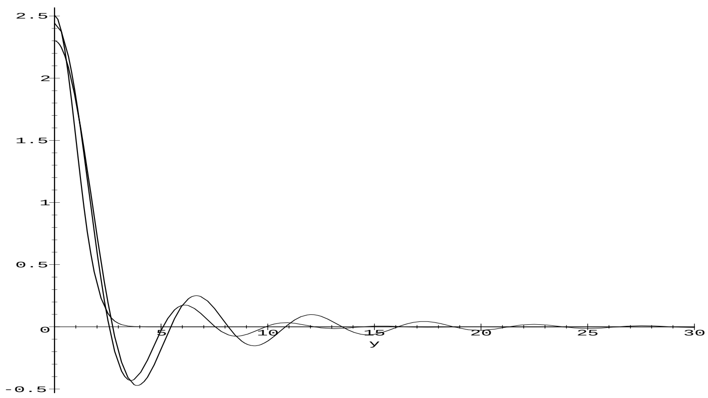

Figure 2.8: The oscillatory Fourier transforms $\widehat{f_2}, \widehat{f_8}, \widehat{f_{16}}$.

Maple happily provides the second integral in a form which simplifies to that we have given. A proof based on the conventional Mellin transform follows.

(a) For $0 < \varepsilon < 1$, use the Γ-function to write

$$\int_0^\infty \frac{\sin y}{y^{1+\varepsilon}}\, dy = \frac{1}{\Gamma(\varepsilon+1)} \int_0^\infty dx \int_0^\infty \sin(x)\exp(-xt)t^\varepsilon\, dt.$$

(b) Interchange variables and evaluate the inner integral to $t^\varepsilon/(t^2+1)$.

(c) Then use the β-function to prove

$$\int_0^\infty \frac{t^\varepsilon}{t^2+1}\, dt = \beta\left(\frac{1}{2} - \frac{1}{\varepsilon}, \frac{1}{2} - \frac{1}{\varepsilon}\right) = \frac{\pi}{2}\sec\left(\frac{\pi}{2}\varepsilon\right).$$

Note:

$$\int_0^\infty \frac{\log^{2n}(s)}{s^2+1}\, ds = (-1)^n \left(\frac{\pi}{2}\right)^{2n+1} \mathrm{E}_{2n}.$$

27. **An explicit formula for the sinc integrals.** Assume that $n \geq 1$ and $a_0, a_1, \cdots, a_n > 0$. For $\gamma = (\gamma_1, \cdots, \gamma_n) \in \{-1, 1\}^n$ define

$$b_\gamma \;=\; a_0 + \sum_{k=1}^{n} \gamma_k \, a_k \quad \text{and} \quad \epsilon_\gamma \;=\; \prod_{k=1}^{n} \gamma_k.$$

Show:

(a) $$\sum_{\gamma \in \{-1,1\}^n} \epsilon_\gamma \, b_\gamma^r \;=\; \begin{cases} 0, & \text{for } r = 0, 1, \cdots, n-1, \\ 2^n \, n! \prod_{k=1}^{n} a_k, & \text{for } r = n, \end{cases}$$

where $b_\gamma^0 = 1$ even if $b_\gamma = 0$. *Hint:* Expand both sides of $e^{a_0 t} \prod_{k=1}^{n} (e^{a_k t} - e^{-a_k t}) = \sum_{\gamma \in \{-1,1\}^n} \epsilon_\gamma \, e^{b_\gamma t}$ into a power series in t and compare coefficients.

(b) $$\prod_{k=0}^{n} \sin(a_k x) \;=\; \frac{1}{2^n} \sum_{\gamma \in \{-1,1\}^n} \epsilon_\gamma \cos(b_\gamma x - \tfrac{\pi}{2} \, (n+1)).$$

(c) $$\int_0^\infty \prod_{k=0}^{n} \frac{\sin(a_k \, x)}{x} \, dx \;=\; \frac{\pi}{2} \frac{1}{2^n \, n!} \sum_{\gamma \in \{-1,1\}^n} \epsilon_\gamma \, b_\gamma^n \operatorname{sign}(b_\gamma).$$

(d) $$\int_0^\infty \prod_{k=0}^{n} \operatorname{sinc}(a_k \, x) \, dx \;=\; \frac{\pi}{2} \frac{1}{a_0} \left(1 - \frac{1}{2^{n-1} \, n! \, a_1 \cdots a_n} \sum_{b_\gamma < 0} \epsilon_\gamma b_\gamma^n \right).$$

(e) The first "bite." If $\sum_{k=1}^{n-1} a_k \leq a_0 < \sum_{k=1}^{n} a_k$, then

$$\int_0^\infty \prod_{k=0}^{n} \operatorname{sinc}(a_k \, x) \, dx \;=\; \frac{\pi}{2} \frac{1}{a_0} \left(1 - \frac{(a_1 + \cdots + a_n - a_0)^n}{2^{n-1} \, n! \, a_1 \cdots a_n} \right).$$

28. **A special sinc integral.** Evaluate

$$\begin{aligned} \int_0^\infty \operatorname{sinc}^n(x) \, dx \;&=\; \frac{\pi}{2} \left(1 + \frac{1}{2^{n-2}} \sum_{1 \leq r \leq \frac{n}{2}} \frac{(-1)^r}{(r-1)!} \frac{(n-2r)^{n-1}}{(n-r)!} \right) \\ &=\; \frac{\pi}{2^n \, (n-1)!} \sum_{0 \leq r \leq \frac{n}{2}} (-1)^r \binom{n}{r} (n-2r)^{n-1}. \end{aligned}$$

In this way, confirm the results of Exercise 5.

29. **A strange cosine integral.** Let $C^*(x) = \cos(2x)\prod_{n=1}^{\infty}\cos(x/n)$. Show symbolically that $\int_0^{\infty} C^*(x)\,dx < \pi/8$, and show numerically that

$$0 \;<\; \frac{\pi}{8} - \int_0^{\infty} C^*(x)\,dx \;<\; 10^{-41}.$$

This is hard to distinguish numerically from $\pi/8$; compare Exercise 39.

30. **Multivariable sinc integrals.** For $x, y \in \mathrm{R}^m$ we write $x \cdot y$ to denote the dot product. Define the *sinc space* $\mathcal{S}^{m,n}$ to be the set of $m \times (m+n)$ matrices $S = (s_1\ s_2\ \cdots\ s_{m+n})$ of column vectors in R^m such that

$$\int_{\mathrm{R}^m} \left| \prod_{k=1}^{m+n} \mathrm{sinc}(s_k \cdot y) \right| dy \;<\; \infty,$$

and a function $\sigma : \mathcal{S}^{m,n} \to \mathrm{R}$ by

$$\sigma(S) \;=\; \int_{\mathrm{R}^m} \prod_{k=1}^{m+n} \mathrm{sinc}(s_k \cdot y)\,dy.$$

Correspondingly, define the *polyhedron space* $\mathcal{P}^{m,n}$ to be the complete set of $m \times (m+n)$ matrices $P = (p_1\, p_2\ \cdots p_{m+n})$ and a function $\nu : \mathcal{P}^{m,n} \to \mathrm{R}$ by

$$\nu(P) \;=\; \mathrm{Vol}\{x \in \mathrm{R}^n : |p_k \cdot x| \le 1 \text{ for } k = 1, 2, \cdots, m+n\}.$$

(a) Note that by change of basis, for $S \in \mathcal{S}^{m,n}$ and $P \in \mathcal{P}^{m,n}$, we have

$$\sigma(S) \;=\; |\det(M)|\,\sigma(MS) \quad \text{and} \quad \nu(P) \;=\; |\det(N)|\,\nu(NP)$$

for nonsingular matrices M $(m \times m)$ and N $(n \times n)$.

(b) The following correspondence between multidimensional sinc integrals and volumes of polyhedra can be proved with some effort (see [34]): *If $n \ge m$, if A is a nonsingular $(m \times m)$-Matrix, and if B is any $(m \times n)$-matrix having m of its columns linearly independent, then*

$$\sigma(A|B) \;=\; \frac{\sigma(I^m|A^{-1}B)}{|\det(A)|} \;=\; \frac{\pi^m}{2^n}\frac{\nu(I^n|(A^{-1}B)^T)}{|\det(A)|}.$$

Similarly, if $n \geq m$, if C is a nonsingular $(n \times n)$-matrix, and if D is any $(n \times m)$-matrix such that $C^{-1}D$ has m linearly independent rows, then

$$\nu(C|D) = \frac{\nu(I^n|C^{-1}D)}{|\det(C)|} = \frac{2^n}{\pi^m}\frac{\sigma(I^m|(C^{-1}D)^T)}{|\det(C)|}.$$

(c) Use the theorem from (b) to determine (with the use of symbolic integration) the volume of $\{x \in \mathrm{R}^6 : |p_k \cdot x| \leq 1,\ k = 1, \cdots, 11\}$, where p_i is the i-th column of the matrix

$$P = \begin{pmatrix} 10 & 0 & 0 & 0 & 0 & 0 & 9 & 10 & -1 & -3 & 7 \\ 0 & 10 & 0 & 0 & 0 & 0 & -2 & -1 & -8 & 2 & -6 \\ 0 & 0 & 10 & 0 & 0 & 0 & -9 & 7 & -5 & 5 & 1 \\ 0 & 0 & 0 & 10 & 0 & 0 & 5 & -2 & -9 & -8 & -9 \\ 0 & 0 & 0 & 0 & 10 & 0 & -10 & -2 & -3 & 6 & -4 \\ 0 & 0 & 0 & 0 & 0 & 10 & -8 & 9 & 2 & 7 & -10 \end{pmatrix}.$$

Hint: $\nu(P) = (32/(5\pi^5)) \int_{\mathrm{R}^5} \prod_{k=1}^{11} \operatorname{sinc}(s_i \cdot y)\, dy$, where

$$S = \begin{pmatrix} 10 & 0 & 0 & 0 & 0 & 9 & -2 & -9 & 5 & -10 & -8 \\ 0 & 10 & 0 & 0 & 0 & 10 & -1 & 7 & -2 & -2 & 9 \\ 0 & 0 & 10 & 0 & 0 & -1 & -8 & -5 & -9 & -3 & 2 \\ 0 & 0 & 0 & 10 & 0 & -3 & 2 & 5 & -8 & 6 & 7 \\ 0 & 0 & 0 & 0 & 10 & 7 & -6 & 1 & -9 & -4 & -10 \end{pmatrix}.$$

31. **Another iterated sinc integral.** For positive constants (a_i), evaluate

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\sin(a_1 x_1)}{x_1} \cdots \frac{\sin(a_n x_n)}{x_n} \frac{\sin(a_1 x_1 + \cdots + a_n x_n)}{x_1 + \cdots + x_n}\, dx_1 \cdots dx_n.$$

Answer: $\pi^n \min(a_1, \ldots, a_n)$.

32. **Infinite series and Clausen's product.** For x and t appropriately restricted:

(a) Use Clausen's product to obtain

$$\sum_{n=0}^{\infty} \frac{(t)_n\,(-t)_n}{(2\,n)!}\,(2\,x)^{2\,n} = \cos\left(2\,t \arcsin\left(x\right)\right)$$

and

$$-\frac{1}{2}\sum_{n=1}^{\infty}\frac{(t)_n\,(-t)_n}{(2\,n)!}\left(4\,\sin^2 x\right)^n=\sin^2(t\,x)\,.$$

(b) Obtain the Taylor series

$$\arcsin^2(x)=\frac{1}{2}\sum_{n\geq 1}\frac{(2\,x)^{2n}}{n^2\binom{2n}{n}}$$

on taking an appropriate limit as $t\to 0$ (see also Exercise 16 of Chapter 1). Hence, show

$$\sum_{n\geq 1}\frac{1}{n^2\binom{2n}{n}}=\frac{\pi^2}{18}\qquad\text{and}\qquad\sum_{n\geq 1}\frac{(-1)^n}{n^2\binom{2n}{n}}=-2\,\log^2\left(\frac{1+\sqrt{5}}{2}\right).$$

Evaluate $\sum_{n\geq 1}3^n/\binom{2n}{n}$ and both of $\sum_{n\geq 1}(\pm 1)^n/\binom{2n}{n}$.

33. **Proof of the Korovkin theorems.** Prove Theorems 2.6.2 and 2.6.4.

34. **Korovkin by inequalities.** An interesting recent approach to the Korovkin theorems is given in [210]. Recall that a subset of a continuous function space is a *subalgebra* if it is closed under pointwise multiplication. Therein, the following elegant lemma is proven:

Lemma 2.7.3 *Suppose that $\mathcal{A}$ is a norm-closed subalgebra of $C[a,b]$ that contains* 1. *Let T be a positive linear operator on $\mathcal{A}$ such that $T(1)\leq 1$. Then*

(a) $\mathcal{E}(h)=T(h^2)-T(h)^2\geq 0$,
(b) $|T(fg)-T(f)T(g)|^2\leq\mathcal{E}(f)\,\mathcal{E}(g)$,
(c) $\|T(fg)-T(f)T(g)\|^2\leq\|\mathcal{E}(f)\|\,\|\mathcal{E}(g)\|$,
(d) $\|T(fg)-T(f)T(g)\|^2\leq\|\mathcal{E}(f)\|\,\|\mathcal{E}(g)+\mathcal{E}(k)\|$,

for all elements f,g,h and k in the algebra.

Proof. (a) is established by observing that $T\left((h+t\,1)^2\right)\geq 0$ for all real t. Then (b) follows with h replaced by $f+tg$, and (c) and (d) are easy consequences. □

It is now a nice problem to show that the first and second Korovkin theorems follow—if one knows that the polynomials are dense in $C[a,b]$. Moreover, the same approach will yield:

Theorem 2.7.4 (Complex Korovkin theorem). *Let $D = \{z \in \mathrm{C} : |z| \leq 1\}$. Let T_n be positive linear operators on $C(D)$ such that $T_n(h) \rightrightarrows h$ for $h = 1, z$ and $|z|^2$. Then this holds for all h in $C(D)$.*

To prove this, it helps to observe that positive operators preserve conjugates: $T\left(\overline{h}\right) = \overline{T\left(h\right)}$ for all h in $C(D)$.

35. **Bézier curves.** The *Bézier curve* of degree n defined by $n+1$ points $b_0, b_1, \ldots, b_n$ is exactly the Bernstein polynomial interpolating the values at k/n

$$\sum_{k=1}^{n} b_k \binom{n}{k} t^k \left(1-t\right)^{n-k}. \tag{2.7.36}$$

Typically, parametric cubic Bézier curves in the plane such as

$$\begin{aligned} x(t) &= -\left(1-t\right)^3 - t\left(1-t\right)^2 + \frac{3}{2}t^2\left(1-t\right) + t^3 \qquad (2.7.37) \\ y(t) &= \frac{1}{2}\left(1-t\right)^3 + t\left(1-t\right)^2 + \frac{3}{4}t^2\left(1-t\right) + \frac{1}{2}t^3 \end{aligned}$$

are fitted together for smoothing purposes. To compute the values, it is useful to observe Castlejau's algorithm that the basis functions $B_{n,k} = t \mapsto \binom{n}{k} t^k \left(1-t\right)^{n-k}$ satisfy the recursion $B_{n,-1} = B_{n-1,n} = 0$ and

$$B_{n,k}(t) = \left(1-t\right) B_{n-1,k}(t) + t\, B_{n-1,k-1}(t),$$

for $0 \leq k \leq n$ and all real t.

36. **Bernstein polynomials.** Determine the appropriate Bernstein polynomials on $[-1,1]$.

37. **Rate of approximation.** As we have seen, the rate of approximation is tied to the smoothness of the underlying function. In Lebesgue's proof of the Stone-Weierstrass Theorem, the main work is in showing that $|\cdot|$

can be uniformly approximated by polynomials on $[-1,1]$. Plot the first few Bernstein polynomials and observe that the approximation is worst at zero, where $|t|$ is not differentiable.

38. **Korovkin kernels.** Apply the Korovkin theorems to the Poisson, Fejér-Korovkin, and Jackson kernels, respectively.

39. **Contriving coincidences.**

(a) A consequence of the theta transform, (2.3.15), in the form $s\,\theta_3^2\left(e^{-\pi s}\right) = \theta_3^2\left(e^{-\pi/s}\right)$, is that

$$\sum_{n\geq 1} e^{-(n/10)^2} \approx 5\,\Gamma\left(\frac{1}{2}\right) - \frac{1}{2}$$

and they agree through 427 digits, with similar more baroque estimates for higher powers of ten.

(b) The fact that $\alpha = \exp(\pi\sqrt{163}/3) \approx 640320$ lies deeper and relates to the fact that the only imaginary quadratic fields with unique factorization are $\mathrm{Q}\left(\sqrt{-d}\right)$, with $d = 1, 2, 4, 7, 11, 19, 43, 67$, and 163.

(c) This leads to a spectacular "billion-digit" fraud

$$\sum_{n=1}^{\infty} \frac{[n\alpha]}{2^n} \approx 1280640.$$

As we saw this is explained by Theorem 1.4.2 and the fact that as a continued fraction,

$$\alpha = [640320, 1653264929, 30, 1, 321, 2, 1, 1, 1, 4, 3, 4, 2, \ldots].$$

(d) Determine the integers N_d such that

$$\sum_{n=1}^{\infty} \frac{[n\alpha_d]}{2^n} \approx N_d,$$

for $\alpha_d = \exp(\pi\sqrt{d}/3)$ with $d = 19, 43, 67, 163$, and determine the error in each case.

These examples signal the danger of inferring a symbolic identity from tools like PSLQ without knowing the context. That said, we know of nearly no cases where such spectacular deception has occurred without contrivance.