Chapter 3

Fourier Series

3.1 Fourier series – definitions and convergence

3.1.1 Definition. Fourier Series

a. Let $\Omega > 0$ and let $F : \mathbb{R} \rightarrow \mathbb{C}$ be a function. $F$ is $2\Omega$-periodic with period $2\Omega$ if $F(\gamma + 2\Omega) = F(\gamma)$ for all $\gamma \in \mathbb{R}$. For example, $F(\gamma) = \sin \gamma$ is $2\pi$-periodic. If $F$ is defined a.e. then $F$ is $2\Omega$-periodic if $F(\gamma + 2\Omega) = F(\gamma)$ a.e.

b. Let $F \in L^1_{loc} (\mathbb{R})$ be $2\Omega$-periodic. The Fourier series of $F$ is the series,

$$S(F)(\gamma) = \sum f[n] e^{-i\pi n \gamma / \Omega},$$

where

$$\forall n \in \mathbb{Z}, \quad f[n] = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\gamma) e^{i\pi n \gamma / \Omega} d\gamma.$$ 

The numbers $f[n]$ are the Fourier coefficients of $F$. The symbol "$\Sigma$" denotes summation over all of $\mathbb{Z}$, i.e., "$\sum_{n=-\infty}^{\infty}$".

c. Formally, the right side of (3.1.1) can be thought of as defining the Fourier transform $\hat{f}$ or $F$ of the sequence $f = \{f[n]\}$, cf., Remark 3.1.3d. In fact, a sequence $f = \{f[n]\}$ is a function

$$f : \mathbb{Z} \rightarrow \mathbb{C}$$

$$n \mapsto f[n].$$
Letting $\ell^1(\mathbb{Z})$ be the space of all sequences $f = \{f[n]\}$ for which $\|f\|_{\ell^1(\mathbb{Z})} = \sum |f[n]| < \infty$, the right side of (3.1.1) is well-defined for $f \in \ell^1(\mathbb{Z})$. In this context, we shall write

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

just as we did in Definition 1.1.2 for the case of Fourier transforms. Thus, in the case of sequences we write $F[n] = f[n]$.

The notation (3.1.3) is based on the presumption that $S(F)$ should equal $F$, e.g., Remark 3.1.2 and Theorem 3.1.6; and that if the right side of (3.1.1) defines the Fourier transform of the sequence $f$ then (3.1.2) is the Fourier inversion formula on $\mathbb{Z}$ corresponding to the Fourier inversion formula (1.1.1) on $\mathbb{R}$.

In fact, $S(F)$ often does equal $F$ in the sense that the partial sums of the series $S(F)$ will converge in some way to $F$. With this in mind, if $F \in L^1_{\text{loc}}(\mathbb{R})$ is $2\Omega$-periodic we shall write

$$S_{M,N}(F)(\gamma) = \sum_{n=-M}^{N} f[n] e^{-i\pi n \gamma / \Omega},$$

where $f$ is defined by (3.1.2). $S_N(F) \equiv S_{N,N}(F)$ is the $N^{th}$ partial sum of $S(F)$.

d. The venerable subject of Fourier series has its share of venerable treatises, which include [Bary64], [Car30], [Edw67], [HR56], [Kah70], [KS63], [Kat76], [Kör88], [Rog59], and [Zyg59] (Zygmund’s first edition is great, too).

e. We are using the notation “$f[n]$” for a Fourier coefficient of $F$ to distinguish it from the notation “$f(n)$”, which usually designates the value at $n \in \mathbb{Z}$ of a function $f$ defined on $\mathbb{R}$. Similarly, we have chosen “$f[n]$” instead of “$f_n$”, since “$f_n$” often indicates an element of a sequence of functions defined on $\mathbb{R}$. Also, we could use “$c_n$” instead of “$f[n]$”, but then we lose contact with the letter “$F$”.

We can and shall consider Fourier series of periodic functions on $\mathbb{R}$ instead of $\hat{\mathbb{R}}$. Our choice of (3.1.1) to define Fourier series is based on the first part of $c$, the typical setting of spectral frequency information
3.1. FOURIER SERIES – DEFINITIONS AND CONVERGENCE

(in terms of Greek letters such as "\( \gamma \)")) associated with digital signals (sequences), and whim!

Finally, we could have defined Fourier series for 1-periodic or 2\( \pi \)-periodic functions; and then have developed the theory of Fourier series, unburdened by lots of \( \Omega \)s. We have chosen the setting of 2\( \Omega \)-periodic functions to give us the flexibility of dealing with different values of \( \Omega \) which might arise in specific problems or applications. In more theoretical developments, we shall usually let 2\( \Omega = 1 \).

3.1.2 Remark. Formal Calculation and Elementary Examples

a. The reason we deal with (3.1.1) and (3.1.2) as a pair is that the decomposition of \( F \) into its fundamental parts, viz., the formula \( S(F) = F \), is only effective if there is quantitative knowledge of the coefficients \( f[n] \) in (3.1.3). In the case \( S(F) = F \) the following formal calculation allows us to obtain (3.1.2) from (3.1.1):

\[
\int_{-\Omega}^{\Omega} F(\gamma)e^{\pi i n \gamma / \Omega} d\gamma = \int_{-\Omega}^{\Omega} \left( \sum_{n} f[n]e^{-\pi i n \gamma / \Omega} \right) e^{\pi i n \gamma / \Omega} d\gamma
\]

\[
= \sum_{n=-\infty}^{\infty} f[n] \int_{-\Omega}^{\Omega} e^{\pi i (m-n) \gamma / \Omega} d\gamma = 2\Omega f[n].
\]

b. Let \( \Omega > 0 \) and let \( \alpha \in (0, \Omega) \). Define \( F \) as \( F = 1_{[-\alpha, \alpha]} \) on \([-\Omega, \Omega]\), extended 2\( \Omega \)-periodically on \( \mathbb{R} \), i.e., \( F(\gamma + 2n\Omega) = F(\gamma) \) for all \( \gamma \in \mathbb{R} \) and all \( n \in \mathbb{Z} \). The Fourier series of \( F \) is

\[
S(F)(\gamma) = \sum_{d(n)} d[\alpha][n]e^{-\pi i n \gamma / \Omega}
\]

where \( d[\alpha][0] = \frac{\alpha}{\Omega} \) and

\[
\forall n \in \mathbb{Z} \setminus \{0\}, \quad d[\alpha][n] = \frac{\alpha \sin \left( \frac{\pi n \alpha}{\Omega} \right)}{\Omega} \left( \frac{\pi n \alpha}{\Omega} \right),
\]

cf., the Dirichlet function in Example 1.3.1.

Next define \( F \) by \( F(\gamma) = \max(1 - \frac{\left| \gamma \right|}{\alpha}, 0) \) on \([-\Omega, \Omega]\), extended 2\( \Omega \)-periodically on \( \mathbb{R} \). A straightforward calculation, similar to that in Example 1.3.4, shows that the Fourier series of \( F \) is

\[
S(F)(\gamma) = \sum_{w[\alpha][n]} e^{-\pi i n \gamma / \Omega},
\]
where \( w(\alpha)[0] = \frac{\alpha}{2\Omega} \) and

\[
\forall n \in \mathbb{Z}\setminus\{0\}, \quad w(\alpha)[n] = \frac{\alpha}{2\Omega} \left( \frac{\sin \left( \frac{\pi n \alpha}{2\Omega} \right)}{\frac{\pi n \alpha}{2\Omega}} \right)^2,
\]

cf., the Fejér function in Example 1.3.4.

The Fourier coefficients in this example define the Dirichlet and Fejér kernel on \( \mathbb{Z} \), cf., Example 3.4.5.

3.1.3 Remark. Notation and Setting

a. If \( \Omega > 0 \) and \( F \in L^1_{\text{loc}}(\mathbb{R}) \) is \( 2\Omega \)-periodic, then we write \( F \in L^1(\mathbb{T}_{2\Omega}) \).

Mathematically, \( \mathbb{T}_{2\Omega} = \mathbb{R}/(2\Omega \mathbb{Z}) \) is a special quotient group referred to as the circle group depending on \( \Omega \). An engineering student need not be concerned with this terminology for the time being. The point is that, because of the periodicity of \( F \), \( F \in L^1(\mathbb{T}_{2\Omega}) \) can be thought of as being defined on any fixed interval \( I \subseteq \mathbb{R} \) of length \( 2\Omega \); and that this periodicity, combined with knowledge of \( F \) on any such interval, completely determines \( F \) on \( \mathbb{R} \).

Similarly, \( \lambda_0 \in \mathbb{T}_{2\Omega} \) indicates any element of the set \( \{ \lambda + 2n\Omega : n \in \mathbb{Z} \} \subseteq \mathbb{R} \) for some fixed \( \lambda \in \mathbb{R} \). Further, \( J \subseteq \mathbb{T}_{2\Omega} \) indicates a subset \( J = \{ \gamma + 2n\Omega : \gamma \in J \} \), \( n \in \mathbb{Z} \).

If \( \Omega = 1/2 \), we write \( \mathbb{T} = \mathbb{T}_1 \).

This possibly cryptic exposition might be unraveled at this time by performing some of the calculations in Exercise 3.1, cf., part c.

In any case, if \( F \in L^1(\mathbb{T}_{2\Omega}) \), then \( \int_I F(\gamma) \, d\gamma = \int_{\mathbb{T}_{2\Omega}} F(\gamma) \, d\gamma \) for any interval \( I \subseteq \mathbb{R} \) of length \( 2\Omega \). As such we introduce the notation,

\[
\int_{\mathbb{T}_{2\Omega}} F(\gamma) \, d\gamma = \frac{1}{2\Omega} \int_I F(\gamma) \, d\gamma.
\]

The factor \( \frac{1}{2\Omega} \) is a normalization factor in the sense that if \( F = 1 \) on \( \mathbb{R} \) then

\[
F = 1 \in L^1(\mathbb{T}_{2\Omega}) \quad \text{and} \quad \int_{\mathbb{T}_{2\Omega}} 1 \, d\gamma = 1.
\]
3.1. Fourier Series – Definitions and Convergence

The norm of \( F \in L^1(\mathbb{T}_2\Omega) \) is

\[
\|F\|_{L^1(\mathbb{T}_2\Omega)} = \int_{\mathbb{T}_2\Omega} |F(\gamma)| \, d\gamma = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} |F(\gamma)| \, d\gamma.
\]

b. If \( \Omega > 0 \), \( F \) is \( 2\Omega \)-periodic, and \( F^2 \in L^1_{\text{loc}}(\mathbb{R}) \) is \( 2\Omega \)-periodic, then we write \( F \in L^2(\mathbb{T}_2\Omega) \). The norm of \( F \in L^2(\mathbb{T}_2\Omega) \) is

\[
\|F\|_{L^2(\mathbb{T}_2\Omega)} = \left( \int_{\mathbb{T}_2\Omega} |F(\gamma)|^2 \, d\gamma \right)^{1/2} = \left( \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} |F(\gamma)|^2 \, d\gamma \right)^{1/2}.
\]

By Hölder’s Inequality,

\[
\int_{\alpha}^{\beta} |F(\gamma)G(\gamma)| \, d\gamma \\
\leq \left( \int_{\alpha}^{\beta} |F(\gamma)|^2 \, d\gamma \right)^{1/2} \left( \int_{\alpha}^{\beta} |G(\gamma)|^2 \, d\gamma \right)^{1/2},
\]

(3.1.4)

e.g., Theorem A.15; and so we have the inclusion \( L^2(\mathbb{T}_2\Omega) \subseteq L^1(\mathbb{T}_2\Omega) \) and the inequality

\[
\forall F \in L^2(\mathbb{T}_2\Omega), \quad \|F\|_{L^1(\mathbb{T}_2\Omega)} \leq \|F\|_{L^2(\mathbb{T}_2\Omega)}.
\]

(3.1.5)

Recall the analogous (sic) situation on \( \mathbb{R} \), i.e., Exercise 1.35. The inclusion and inequality (3.1.5) follow since, by taking \( G = 1 \) in (3.1.4), we obtain

\[
\|F\|_{L^1(\mathbb{T}_2\Omega)} = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} |F(\gamma)| \, d\gamma
\]

\[
\leq \frac{1}{2\Omega} \left( \int_{-\Omega}^{\Omega} |F(\gamma)|^2 \, d\gamma \right)^{1/2} \left( 2\Omega \right)^{1/2} = \|F\|_{L^2(\mathbb{T}_2\Omega)}.
\]

The inclusion is also proper, e.g., Exercise 3.6.

c. Let \( \Omega = \pi \) and let \( F(\gamma) = \sin \gamma + \cos 2\gamma \) on \( \mathbb{R} \). Then

\[
\int_{\mathbb{T}_2\Omega} F(\gamma) \, d\gamma = \frac{1}{2\pi} \int_{0}^{2\pi} F(\gamma) \, d\gamma
\]

\[
= \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\gamma) \, d\gamma = \frac{1}{2\pi} \int_{\alpha}^{\alpha+2\pi} F(\gamma) \, d\gamma = 0,
\]
for any fixed $\alpha \in \hat{\mathbb{R}}$. Further, if $J_I = \left( \frac{\pi}{2}, \frac{3\pi}{2} \right) \subseteq I = (0, 2\pi]$ then

$$\forall n \in \mathbb{Z}, \quad \frac{1}{2\pi} \int_{J_I} F(\gamma) \, d\gamma = \frac{1}{2\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} F(\gamma) \, d\gamma = \frac{1}{2\pi} \int_{J_I + 2\pi n} F(\gamma) \, d\gamma.$$

Thus, if we let $J$ be any one of the sets $J_I + 2\pi n \subseteq \hat{\mathbb{R}}$, then $J \subseteq T_{2\Omega}$ and

$$\int_{J} F(\gamma) \, d\gamma = \frac{1}{2\pi} \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} F(\gamma) \, d\gamma.$$

Theorem 3.1.1c we defined the Fourier transform of a sequence $f = \{f[n]\} \in \ell^1(\mathbb{Z})$. We now define the Fourier transform of $F \in L^1(T_{2\Omega})$ as the sequence $f = \{f[n]\},$ where

$$\forall n \in \mathbb{Z}, \quad f[n] = \int_{T_{2\Omega}} F(\gamma) e^{-i\pi n \gamma / \Omega} \, d\gamma,$$

and where "$f[n]$" is defined in part a. In this case, the formal inversion formula is

$$F(\gamma) = \sum_{n} f[n] e^{i\pi n \gamma / \Omega}.$$

c. In Chapter 1, for $f \in L^1(\mathbb{R})$, the Fourier transform of $f$ was defined on $\hat{\mathbb{R}}(= \mathbb{R})$. In this chapter, we have two "dual" settings. First, for $f \in \ell^1(\mathbb{Z})$, the Fourier transform of $f$ is defined on $T_{2\Omega}$; and, second, for $F \in L^1(T_{2\Omega})$, the Fourier transform of $F$ is defined on $\mathbb{Z}$. Mathematically, $\mathbb{R}$ and $\hat{\mathbb{R}}$ are locally compact abelian groups (LCAGs) which are not compact and which are dual, in a technical sense, to each other; similarly, the discrete LCAG $\mathbb{Z}$ is the dual group of the compact LCAG $T_{2\Omega}$, and vice-versa, e.g., [Rud62], [Edw67], [Ben75].

3.1.4 Example. Fourier Series of Real-Valued Functions

Let $F \in L^1(T_{2\Omega})$ be real-valued, and let $f = \{f[n]\}$ be the sequence of Fourier coefficients of $F$. Thus,

$$f[n] = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\gamma) \cos \left( \frac{\pi n \gamma}{\Omega} \right) \, d\gamma + \frac{i}{2\Omega} \int_{-\Omega}^{\Omega} F(\gamma) \sin \left( \frac{\pi n \gamma}{\Omega} \right) \, d\gamma.$$
3.1. **FOURIER SERIES – DEFINITIONS AND CONVERGENCE**

Formally, since \( S(F) \) should be equal to \( F \) and therefore be real-valued, we have

\[
S(F)(\gamma) = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \, d\lambda \\
+ \sum' \left( \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \cos \left( \frac{\pi n \lambda}{\Omega} \right) \, d\lambda + \frac{i}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \sin \left( \frac{\pi n \lambda}{\Omega} \right) \, d\lambda \right) \times \\
\left( \cos \left( \frac{\pi n \gamma}{\Omega} \right) - i \sin \left( \frac{\pi n \gamma}{\Omega} \right) \right) \\
= \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \, d\lambda + \sum' \left( \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \cos \left( \frac{\pi n \lambda}{\Omega} \right) \, d\lambda \right) \cos \left( \frac{\pi n \gamma}{\Omega} \right) \\
+ \sum' \left( \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \sin \left( \frac{\pi n \lambda}{\Omega} \right) \, d\lambda \right) \sin \left( \frac{\pi n \gamma}{\Omega} \right) \\
= a_0 + \sum_{n=1}^{\infty} a_n \cos \left( \frac{\pi n \gamma}{\Omega} \right) + \sum_{n=1}^{\infty} b_n \sin \left( \frac{\pi n \gamma}{\Omega} \right),
\]

where

\[
a_0 = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \, d\lambda,
\]

and

\[
a_n = \frac{1}{\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \cos \left( \frac{\pi n \lambda}{\Omega} \right) \, d\lambda,
\]

and

\[
b_n = \frac{1}{\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \sin \left( \frac{\pi n \lambda}{\Omega} \right) \, d\lambda,
\]

for \( n > 0 \). The coefficients \( a_n, b_n \) are obtained since, for example, if \( n > 0 \) then

\[
\left( \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \sin \left( -\frac{\pi n \lambda}{\Omega} \right) \, d\lambda \right) \sin \left( -\frac{\pi n \gamma}{\Omega} \right) \\
+ \left( \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\lambda) \sin \left( \frac{\pi n \lambda}{\Omega} \right) \, d\lambda \right) \sin \left( \frac{\pi n \gamma}{\Omega} \right) \\
= b_n \sin \left( \frac{\pi n \gamma}{\Omega} \right).
\]
CHAPTER 3. FOURIER SERIES

The Riemann-Lebesgue Lemma for $L^1(\mathbb{R})$ (Theorem 1.4.1c) has an
analogue for $L^1(\mathbb{T}_{2\Omega})$.

3.1.5 Theorem. Riemann-Lebesgue Lemma

If $F \in L^1(\mathbb{T}_{2\Omega})$ then $\lim_{|n| \to \infty} f[n] = 0$, where $f = \{f[n]\}$ is the
sequence of Fourier coefficients of $F$, i.e., $\hat{f} = F$.

Proof. a. Assume $F \in C^1(\mathbb{R})$. Then, $G = F' \in L^1(\mathbb{T}_{2\Omega})$ has the
properties that $\int_{-\Omega}^{\Omega} G(\gamma) \, d\gamma = 0$ and

$$\forall \gamma \in [-\Omega, \Omega), \quad F'(\gamma) = \int_{-\Omega}^{\gamma} G(\lambda) \, d\lambda + F(-\Omega),$$

cf., the approach in Theorem 1.4.1c.

We compute (for $n \neq 0$)

$$f[n] = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\gamma) e^{\pi i n \gamma / \Omega} \, d\gamma$$

$$= \frac{1}{2\Omega} \left[ \frac{\Omega}{\pi i n} e^{\pi i n \gamma / \Omega} F(\gamma) \right]_{-\Omega}^{\Omega} - \frac{\Omega}{\pi i n} \int_{-\Omega}^{\Omega} G(\gamma) e^{\pi i n \gamma / \Omega} \, d\gamma$$

$$= \frac{-1}{2\pi i n} \int_{-\Omega}^{\Omega} G(\gamma) e^{\pi i n \gamma / \Omega} \, d\gamma;$$

and, hence,

$$|f[n]| \leq \frac{\Omega}{\pi |n|} \|G\|_{L^1(\mathbb{T}_{2\Omega})}.$$

Consequently, $\lim_{|n| \to \infty} f[n] = 0$.

b. Let $F \in L^1(\mathbb{T}_{2\Omega})$ and $\epsilon > 0$. There is $F_\epsilon \in C^1(\mathbb{R})$ which is
2$\Omega$-periodic and for which $\|F - F_\epsilon\|_{L^1(\mathbb{T}_{2\Omega})} < \epsilon$. Then (3.1.5) is valid
with $F_\epsilon$ and $G_\epsilon = F'_\epsilon \in L^1(\mathbb{T}_{2\Omega})$ instead of $F$ and $G$. (The existence
of $F_\epsilon$ can be proven in many ways, including the convolution of $F$ with
an approximate identity, which we shall discuss in Section 3.4.)

We compute (for $n \neq 0$)

$$|f[n]| \leq |f[n] - f_\epsilon[n]| + |f_\epsilon[n]|$$

$$\leq \|F - F_\epsilon\|_{L^1(\mathbb{T}_{2\Omega})} + \frac{\Omega}{\pi |n|} \|G_\epsilon\|_{L^1(\mathbb{T}_{2\Omega})},$$
3.1. FOURIER SERIES – DEFINITIONS AND CONVERGENCE

where \( f_\varepsilon = \{f_\varepsilon[n]\} \) is the sequence of Fourier coefficients of \( F_\varepsilon \) and where we have invoked part a in the second inequality.

We know that

\[
\lim_{|n| \to \infty} a_n \leq \lim_{|n| \to \infty} b_n + \lim_{|n| \to \infty} c_n
\]

in case \( a_n \leq b_n + c_n \) and \( a_n, b_n, c_n \geq 0 \), e.g., Definition A.1. Consequently,

\[
\lim_{|n| \to \infty} |f[n]| \leq \lim_{|n| \to \infty} \|F - F_\varepsilon\|_{L^1(T_{2\Omega})} < \varepsilon.
\]

Since the left side is non-negative and independent of \( \varepsilon \) we conclude that \( \lim_{|n| \to \infty} |f[n]| = 0 \). \( \square \)

We shall use the Riemann-Lebesgue Lemma to verify Dirichlet’s fundamental theorem which provides sufficient conditions on a function \( F \in L^1(T_{2\Omega}) \) so that \( S(F)(\gamma_0) = F(\gamma_0) \) for a given point \( \gamma_0 \). The following ingenious proof is due to P. Chernoff [Che80], cf., [Lio86] and the classical proof as found in [Zyg59]. Dirichlet’s theorem for Fourier series naturally preceded the analogous inversion theorem for Fourier transforms, e.g., Sections 1.7 and 3.2.

3.1.6 Theorem. DIRICHLET THEOREM

If \( F \in L^1(T_{2\Omega}) \) and \( F \) is differentiable at \( \gamma_0 \) then \( S(F)(\gamma_0) = F(\gamma_0) \) in the sense that

\[
\lim_{M,N \to \infty} \sum_{n=-M}^{N} f[n]e^{-\pi \text{in} \gamma_0 / \Omega} = F(\gamma_0),
\]

where \( f = \{f[n]\} \) is the sequence of Fourier coefficients of \( F \), i.e., \( \hat{f} = F \).

Proof. a. Without loss of generality, assume \( \gamma_0 = 0 \) and \( F(\gamma_0) = 0 \). In fact, if \( F(\gamma_0) \neq 0 \) then consider the function \( F - F(\gamma_0) \) (instead of \( F \)), which is also an element of \( L^1(T_{2\Omega}) \), and then translate this function to the origin.

b. Since \( F(0) = 0 \) and \( F'(0) \) exists we can verify that \( G(\gamma) = F(\gamma)/(e^{-\pi \gamma / \Omega} - 1) \) is bounded in some interval centered at the origin.
To see this, note that

\[(3.1.7) \quad G(\gamma) = \frac{F(\gamma)}{\gamma} \left( \frac{1}{\frac{2i}{\Omega}} + \left(\frac{-\frac{2i}{\Omega}}{\frac{2i}{\Omega}}\right)^2 \frac{1}{2i} + \left(\frac{-\frac{2i}{\Omega}}{\frac{3i}{\Omega}}\right)^3 \frac{1}{3i} \gamma^2 + \cdots \right),\]

and, hence, \(G(\gamma)\) is close to \(\Omega F'(0)/(\pi i)\) in a neighborhood of the origin, e.g., Exercise 3.4.

This boundedness near the origin, coupled with integrability of \(F\) on \(T_{2\Omega}\), yields the integrability of \(G\) on \(T_{2\Omega}\). Therefore, since \(F(\gamma) = G(\gamma)(e^{-\pi i \gamma/\Omega} - 1)\) we compute \(f[n] = g[n-1] - g[n]\), where \(g = \{g[n]\}\) is the sequence of Fourier coefficients of \(G\). Thus, the partial sum 
\[S_{M,N}(F)(0) = \sum_{n=-M}^{N} (g[n-1] - g[n]) = g[-M-1] - g[N].\]

Consequently, we can apply the Riemann-Lebesgue Lemma to obtain
\[\lim_{M,N \to \infty} S_{M,N}(F)(0) = 0. \quad \square\]

3.1.7 Remark. **Fundamental Spaces and Elementary Convergence Results**

a. Let \(\Omega > 0\), and let \(C^m(\mathbb{R})\), \(0 \leq m \leq \infty\), be the space of \(m\)-times continuously differentiable functions on \(\mathbb{R}\). It is convenient to define the following spaces:

\[C^m(T_{2\Omega}) = \{F \in C^m(\mathbb{R}) : F \text{ is } 2\Omega\text{-periodic on } \mathbb{R}\}, \quad 0 \leq m \leq \infty,\]

\[AC(T_{2\Omega}) = \{F \in AC_{loc}(\mathbb{R}) : F \text{ is } 2\Omega\text{-periodic on } \mathbb{R}\},\]

\[BV(T_{2\Omega}) = \{F \in BV_{loc}(\mathbb{R}) : F \text{ is } 2\Omega\text{-periodic on } \mathbb{R}\},\]

\[C(T_{2\Omega}) = C^0(T_{2\Omega}) = \{F \in C^0(\mathbb{R}) : F \text{ is } 2\Omega\text{-periodic on } \mathbb{R}\}.\]

Clearly,
\[C^\infty(T_{2\Omega}) \subseteq \ldots \subseteq C^{m+1}(T_{2\Omega}) \subseteq C^m(T_{2\Omega}) \subseteq \ldots \subseteq C^1(T_{2\Omega}) \subseteq AC(T_{2\Omega}) \subseteq C(T_{2\Omega}) \cap BV(T_{2\Omega}),\]
3.1. FOURIER SERIES – DEFINITIONS AND CONVERGENCE

\text{cf., Remark 1.4.2.}\]

b. The hypothesis in Theorem 3.1.6 that \( F \) is differentiable at \( \gamma_0 \) is strong; and the proof of Theorem 3.1.6 is still valid if the hypothesis on \( F \in L^1(\mathbb{T}_{2\Omega}) \) is weakened to the condition that \( \frac{F(\gamma)-F(\gamma_0)}{\gamma-\gamma_0} \) be integrable on some interval centered at \( \gamma_0 \), cf., Exercise 3.42. In particular, if \( F \in BV(\mathbb{T}_{2\Omega}) \) then

\[
\forall \gamma \in \mathbb{T}_{2\Omega}, \quad \lim_{N \to \infty} S_N(F)(\gamma) = \frac{F(\gamma^+) + F(\gamma^-)}{2}.
\]

Further, Exercises 3.26 and 3.49 deal with properties of Fourier coefficients and rates of convergence of \( \{S_N(F)\} \) for functions \( F \) belonging to the spaces defined in part a.

c. With regard to Theorem 3.1.6, we can further assert that if \( F \in BV(\mathbb{T}_{2\Omega}) \) and if \( F \) is also continuous on a closed subinterval \( I \subseteq \mathbb{T}_{2\Omega} \), then \( \{S_N(F)\} \) converges uniformly to \( F \) on \( I \), cf., [Zyg59, Volume I, pages 57-58]. The Dirichlet Theorem and this version of it for intervals of continuity are often referred to as the Dirichlet-Jordan Test, cf., Section 3.2.3.

3.1.8 Definition. RELATIONS BETWEEN FUNCTIONS DEFINED ON \( \mathbb{Z} \) AND \( \mathbb{T}_{2\Omega} \)

a. If \( f \in \ell^1(\mathbb{Z}) \) and \( \Omega > 0 \) then \( F = \hat{f} \) is an absolutely convergent Fourier series, and the space of such series is denoted by \( A(\mathbb{T}_{2\Omega}) \). By definition, the norm of \( F = \hat{f} \in A(\mathbb{T}_{2\Omega}) \) is

\[
\|F\|_{A(\mathbb{T}_{2\Omega})} = \|f\|_{\ell^1(\mathbb{Z})} = \sum |f[n]|,
\]

cf., Definition 1.1.2 and Example 2.4.6f. We define

\[
L^\infty(\mathbb{T}_{2\Omega}) = \{F \in L^\infty(\mathbb{R}) : F \text{ is } 2\Omega\text{-periodic on } \mathbb{R}\},
\]

and the norm of \( F \in L^\infty(\mathbb{T}_{2\Omega}) \) is \( \|F\|_{L^\infty(\mathbb{T}_{2\Omega})} = \|F\|_{L^\infty(\mathbb{R})} \), e.g., Definition A.10. \( C(\mathbb{T}_{2\Omega}) \) is a closed subspace of \( L^\infty(\mathbb{T}_{2\Omega}) \) if \( C(\mathbb{T}_{2\Omega}) \) is taken with the \( \|\ldots\|_{L^\infty(\mathbb{T}_{2\Omega})} \) norm.

b. We have the proper inclusions,

\[
(3.1.8) \quad \mathcal{A}(\mathbb{T}_{2\Omega}) \subseteq C(\mathbb{T}_{2\Omega}) \subseteq L^\infty(\mathbb{T}_{2\Omega}) \subseteq L^2(\mathbb{T}_{2\Omega}) \subseteq L^1(\mathbb{T}_{2\Omega});
\]
and the identity map corresponding to any of these inclusions is continuous, e.g., Exercise 3.44. In fact,

\[(3.1.9) \quad \|F\|_{L^1(T_{2\pi})} \leq \|F\|_{L^2(T_{2\pi})} \leq \|F\|_{L^\infty(T_{2\pi})} \leq \|F\|_{A(T_{2\pi})}.
\]

We shall show in Example 3.5.3 that \(C^1(T_{2\pi}) \subseteq A(T_{2\pi}).\)

c. Let \(\ell^2(Z) = \{f : Z \rightarrow \mathbb{C} : \|f\|_{A(Z)} = (\sum |f[n]|^2)^{1/2} < \infty\}.
\]

Because of part b, we have the proper inclusions

\[(3.1.10) \quad \ell^1(Z) \subseteq X(Z) \subseteq A'(Z) \subseteq \ell^2(Z) \subseteq A(Z),
\]

where the notation \(X(Z), A'(Z), A(Z)\) is defined as follows:

\[
\begin{align*}
X(Z) &= \{f : Z \rightarrow \mathbb{C} : \hat{f} \in C(T_{2\pi})\}, \\
A'(Z) &= \{f : Z \rightarrow \mathbb{C} : \hat{f} \in L^\infty(T_{2\pi})\}, \\
A(Z) &= \{f : Z \rightarrow \mathbb{C} : \hat{f} \in L^1(T_{2\pi})\},
\end{align*}
\]

e.g., Exercise 3.44. If we define \(\|f\|_{X(Z)} = \|\hat{f}\|_{L^\infty(T_{2\pi})}\) for \(f \in X(Z)\), resp., \(\|f\|_{A'(Z)} = \|\hat{f}\|_{L^\infty(T_{2\pi})}\) for \(f \in A'(Z)\) and \(\|f\|_{A(Z)} = \|\hat{f}\|_{L^1(T_{2\pi})}\) for \(f \in A(Z)\), then the identity map corresponding to any of the inclusions in (3.1.10) is continuous. In fact,

\[(3.1.11) \quad \|f\|_{A(Z)} \leq \|f\|_{A'(Z)} \leq \|f\|_{A(Z)} \leq \|f\|_{X(Z)} \leq \|f\|_{\ell^1(Z)}.
\]

From Theorem 3.1.5, we know that

\[A(Z) \subseteq c_0(Z),\]

where \(c_0(Z) = \{f : Z \rightarrow \mathbb{C} : \lim_{n \to \infty} f[n] = 0\}\), cf., Example 3.3.4a.

d. \(A'(Z)\) is the space of pseudo-measures on \(Z\), cf., Example 2.4.6f.

In light of RRT, it is natural to define the "bounded Radon measures \(M_b(Z)\)" on \(Z\) as the dual space \(c_0(Z)'\) of \(c_0(Z)\), where the norm of \(f \in c_0(Z)\) is defined as \(\|f\|_{c_0(Z)}\), e.g., Appendix B. This handwaving in terms of plausible analogy fails in this case since \(c_0(Z)' = \ell^1(Z)\), i.e., \(M_b(Z) = \ell^1(Z)!\)
3.1. FOURIER SERIES – DEFINITIONS AND CONVERGENCE

3.1.9 Definition. Measures on $T$

a. A linear function $T : C(T) \to \mathbb{C}$ is an element of the dual space $C(T)'$ (of the vector space $C(T)$) if $\lim_{n \to \infty} T(F_n) = 0$ for every sequence $\{F_n\} \subseteq C(T)$ for which $\lim_{n \to \infty} ||F_n||_{L^\infty(T)} = 0$.

b. We denote $C(T)'$ by $M(T)$. $M(T)$ is the space of Radon measures on $T$. The functionals $T \in M(T)$ are often denoted by $\mu, \nu, \text{etc.}$, and in this case we have the usual notation,

$$T(F) = \mu(F) = \int_T F(\gamma) \, d\mu(\gamma),$$

cf., Section 2.7. We also define $M_+(T) = \{\mu \in M(T) : \mu(F) \geq 0 \text{ if } F \geq 0\}$.

c. By the definition of $M(T)$, if $\mu \in M(T)$ then

$$\exists C_\mu > 0 \text{ such that } \forall F \in C(T), \quad |\mu(F)| \leq C_\mu ||F||_{L^\infty(T)}.$$  

$||\mu||_1$ denotes the infimum over all such constants $C_\mu$. As in the case of measures on $\mathbb{R}$, $L^1(T)$ is naturally embedded in $M(T)$, and if the correspondence is denoted by $F \mapsto \mu_F$, then it is easy to show that $||F||_{L^1(T)} = ||\mu_F||_1$.

d. The relationship between $M(T)$ and $M(\mathbb{R})$ is established by the fact that $M(T)$ can be identified with

$$\{\mu \in M(\mathbb{R}) : \tau_1 \mu = \mu\},$$
i.e., $M(T)$ can be considered as the subspace of 1-periodic elements of $M(\mathbb{R})$. Of course, $\mu \in M(T)$ is bounded in the sense of the norm inequality in part c; but if $\mu \in M(T) \setminus \{0\}$, then the 1-periodic measure on $\mathbb{R}$ corresponding to $\mu$ is not in $M_0(\mathbb{R})$. For example, if $F \in L^1(T) \setminus \{0\}$ then the 1-periodic function $\tau_1 F$ on $\mathbb{R}$, which equals $F$ on $[0,1)$, is not in $L^1(\mathbb{R})$. As another example, define the Dirac measure $\delta_\gamma$ at $\gamma \in T$ by the formula $\delta_\gamma(F) = F(\gamma)$, where $F \in C(T)$. Then $\mu = \tau_\gamma(\sum \delta_n)$ is the 1-periodic measure on $\mathbb{R}$ corresponding to $\delta_\gamma$, and $\mu \notin M_0(\mathbb{R})$. The verification of these assertions is left to Exercise 3.50.

3.1.10 Example. Periodicity: Potpourri and Titillation
a. **Periodicity.** We defined $2\Omega$-periodicity on $\mathring{\mathbb{R}}$ in Definition 3.1.1. Let $p \in \mathbb{C}$ and let $F : \mathbb{C} \to \mathbb{C}$ be a function. $F$ is $p$-periodic with period $p$ if $F(z + p) = F(z)$ for all $z \in D$. If $D \subseteq \mathbb{C}$ is a domain then $F$ is doubly periodic with periods $p_1, p_2 \in \mathbb{C}$ if $\text{Im}(p_2/p_1) > 0$ and

$$\forall z \in D, \quad F(z + p_1) = F(z) \text{ and } F(z + p_2) = F(z).$$

$F$ is quasi-periodic if

$$\forall (x, \omega) \in \mathbb{R} \times \mathring{\mathbb{R}},$$

$$F(x + 1, \omega) = e^{-2\pi i \omega} F(x, \omega) \text{ and } F(x, \omega + 1) = F(x, \omega).$$

b. **Jacobi theta function.** As examples, we first note that entire doubly periodic functions are constants. Also $F(z) = e^z$ is $2\pi i$-periodic on $\mathbb{C}$.

The Jacobi theta function $\vartheta_3$ is a $1$-periodic entire function, depending on a parameter $t \in \mathbb{C}$, and is defined as

$$\vartheta(z; t) = \vartheta(z) = \sum e^{-\pi n^2 t + 2\pi i n z},$$

where $\text{Re} \ t > 0$. $\vartheta(z)$ is $1$-periodic on $\mathbb{C}$.

c. **Elliptic functions.** An elliptic function is a meromorphic function in the plane which is doubly periodic in its domain of definition. If $p_1, p_3 \in \mathbb{C}$, $\text{Im}(p_3/p_1) > 0$, and $p_{m,n} \equiv 2mp_1 + 2np_3$ for $m, n \in \mathbb{Z}$, then the Weierstrass $\mathcal{P}$-function is defined as

$$\mathcal{P}(z) = \frac{1}{z^2} + \sum \left\{ \frac{1}{(z - p_{m,n})^2} - \frac{1}{p_{m,n}^2} \right\},$$

where summation is over all $(m, n) \in \mathbb{Z} \times \mathbb{Z} \setminus \{(0, 0)\}$. $\mathcal{P}$ is an elliptic function.

d. **Zak transform.** The Zak transform $Zf$ of $f : \mathbb{R} \to \mathbb{C}$ is formally defined as

$$\forall (x, \omega) \in \mathbb{R} \times \mathring{\mathbb{R}},$$

$$Zf(x, \omega) = a^{1/2} \sum_k f(xa + ka)e^{2\pi ik\omega}, \quad (3.1.12)$$
where $a > 0$. $Zf$ is a quasi-periodic function.

e. **Elliptic integrals.** Seventeenth and eighteenth century problems from astronomy for computing arc length of orbits, or from mechanics for computing the period of a simple pendulum, led to the problem of evaluating integrals of the form

$$s(y) = \int_0^y \left( \frac{1 - k^2 x^2}{1 - x^2} \right)^{1/2} \, dx$$

or

$$s(y) = \int_0^\varphi \frac{d\varphi}{(1 - k^2 \sin^2 \varphi)^{1/2}},$$

respectively. These are examples of elliptic integrals; and Liouville (1833) proved that such integrals can not be evaluated in terms of algebraic, trigonometric, logarithmic, or exponential functions. The study of elliptic integrals was an important part of eighteenth and nineteenth century mathematics Featuring the likes of Fagnano, Euler, and Legendre, and leading to the analyses of Gauss, Abel, and Jacobi. General elliptic integrals are of the form $\int R(x, \sqrt{P(x)}) \, dx$, where $P(x)$ is a third or fourth degree polynomial with distinct roots and $R(x, \omega)$ is a rational function of $x$ and $\omega$.

f. **Quintics.** Instead of studying the function $s(y)$ in (3.1.13) and (3.1.14), Abel (1802-1829) in 1826 and Jacobi in 1827 analyzed the inverse of elliptic integrals; and these are, in fact, the elliptic functions. The analogue in trigonometry is to study the sine function instead of the multiple-valued arcsine,

$$s(y) = \int_0^y \frac{dx}{\sqrt{1 - x^2}}.$$  

Jacobi's theta functions are relatively elementary functions from which elliptic functions can be constructed. Later, in the early 1860s, Weierstrass introduced $\mathcal{P}(z)$ as the elliptic function inverting a specific elliptic integral; and then he proved that every elliptic function can be expressed in terms of $\mathcal{P}(z)$ and $\mathcal{P}'(z)$ [Hil74, page 141].
Ruffini and Abel proved that quintic polynomial equations can not necessarily be solved by algebraic operations. In 1858, Hermite used elliptic functions to obtain solutions of such equations, cf., [Kle56].

**g. Shape of the sun.** We can integrate the Newtonian equations of motion of a secondary body in the equatorial plane of a rotationally symmetric central body; the solution is in terms of the Weierstrass $\mathcal{P}$-function [SB65]. There are important applications of this technique. In particular, the motion of equatorial artificial earth satellites is characterized, and orbital apsidal line shifts of the secondary body can be computed. This latter point is interesting because of the apsidal line shift of mercury's orbit about the sun. This shift can be accounted for by Newtonian methods if the sun is sufficiently "flat", as an oblate spheroid. Robert Dicke and others, e.g., Hill and Stebbins in 1975, have provided an experimental tour de force, and shown that the sun is too spherical, by an order of magnitude, to account for even ten percent of mercury's apsidal line shift by Newtonian methods. Einstein's theory of general relativity does explain the shift.

**h. Coherent states.** A coherent state is a family of functions of the form

\[
\varphi(t) = g(t - a)e^{2\pi itb}e^{2\pi icx\omega},
\]

parameterized by \((a, b, c) \in \mathbb{R}^3\). In the quantum physics literature, \(g\) is often taken to be the Gaussian. There is a natural relation between coherent states and the Heisenberg group, e.g., our *Gabor representations and wavelets*, AMS Contemporary Math., 91(1989), 9–27. Closure problems for coherent states have a history going back to von Neumann's classic from the early 1930s [vN55, page 407]. Zak's role in the evolution of the Zak transform and its use in quantum mechanics has been documented in [Jan88]. The Zak transform and knowledge of its zero set are relevant for solving a variety of closure problems for coherent states, e.g., [BF94, Chapter 3].

As we have seen, \(\vartheta(z)\) is a 1-periodic entire (and therefore non-elliptic) function used in the construction of meromorphic doubly periodic elliptic functions. It turns out that \(\vartheta(z)\) also plays a role in the non-analytic quasi-periodic Zak transform of the Gaussian. In fact, it
is easy to check that the Zak transform of a Gaussian is a product of \( \vartheta(z; t) \), for certain \( t \), and a Gaussian. Thus, the zeros of the Zak transforms of the Gaussian are determined by the zeros of \( \vartheta \), cf., [BF94, computation after Theorem 7.8].

3.2 History of Fourier series

George Sarton, who founded the journal *Isis* in 1912, wrote that the “main duty of the historian of mathematics... is to explain the humanity of mathematics, to illustrate its greatness, beauty and dignity...”. Alas, we can neither achieve such a noble goal with its accompanying deep scholarship, nor even present a lapidary exposition of the history of Fourier series. Fortunately such expositions abound, e.g., the historical commentaries of Riemann [Rie1873], Gibson [Gib1893], Carslaw [Car30], Hobson [Hob26], Plancherel [Pla25], cf., [Zyg59, Preface], the masterful entries on Fourier series in the Encyclopedia Britannicas of the last fifty years, and the relevant biographical entries in the *Dictionary of Scientific Biography*. There are also important historical contributions by Burkhardt, Plessner, and Tonelli referenced in these works.

Our treatment in this section is selective and perhaps idiosyncratic. We shall not discuss the history of Fourier series vis a vis its major applications to heat and light and celestial mechanics by Fourier and Fresnel and Hill, respectively. (Of course, there are brilliant, but perhaps curmudgeonly, thermodynamicists who assert that Fourier’s theory of heat [Fou1822] did not really treat heat.) We shall mostly deal with the relation of Fourier series with real analysis, and to some extent with number theory, cf., [Ben76] and [Mon94], respectively. We shall not discuss its relation with functional analysis, that begins with the profound work of Beurling [Beu89], or with complex analysis, that begins with the work of F. and M. Riesz, Lusin and Privalov, and Hardy and Littlewood.

3.2.1 d’Alembert (1717-1783), Euler (1707-1783), D. Bernoulli (1700-1782), and Lagrange (1736-1813).
In Section 1.8 we discussed some partial differential equations from mathematical physics, and it turns out that Fourier series originated in dealing with such equations.

In 1747, d'Alembert solved the *vibrating string problem*. This problem is to solve the equation,

\[
\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2},
\]

for a function \( u(x, t) \), where \( x \in [0, L] \), \( c \in \mathbb{R} \), \( t \geq 0 \), \( u(0, t) = u(L, t) = 0 \) for all \( t \geq 0 \), and \( u(x, 0) = f(x) \) is given on \([0, L]\). d'Alembert's solution \( u \) is in terms of \( f \), and so \( f \) must be twice differentiable in this case. On the other hand, equation (3.2.1) was derived, after *significant* assumptions, to represent the motion of a taut string, such as a violin string, after it is released from a given initial position \( f \) on \([0, L]\). One can imagine an initial position \( f \) to have corners, so that \( f' \) need not exist everywhere.

In 1748, Euler made an important observation about the vibrating string problem. He noted that the motion of the string is completely determined for \( x \in [0, L] \) and time \( t \geq 0 \) if the form of the string and its velocity at \( t = 0 \) are given. In particular, Euler was able to find the solution of (3.2.1) for a given initial position \( f \) and initial velocity \( g \) of the string. Euler's solution allowed for the initial positions \( f \) to have discontinuous derivatives. This led to a disagreement with d'Alembert on an issue which ultimately comes down to defining the notion of *function* e.g., [Bir73, pages 16ff.] which is taken from [Rie1873], cf., [G-V92]. In any case, Euler felt he had solved the vibrating string problem for very general initial positions \( f \).

Daniel Bernoulli entered the discussion in 1753 in the midst of the d'Alembert-Euler disagreement. Daniel Bernoulli had developed hydrodynamics from the principle of conservation of energy, and was a professor of anatomy and botany at Basel, before becoming a professor of physics. He approached the vibrating string problem with Brook Taylor's observation (1715) that if

\[
(3.2.2) \quad u_n(x, t) = \sin \left( \frac{\pi n x}{L} \right) \cos \left( \frac{\pi n c t}{L} \right), \quad n \in \mathbb{Z},
\]
3.2. HISTORY OF FOURIER SERIES

then (3.2.1) is satisfied for \( u = u_n \), and \( u_n(0, t) = u_n(L, t) = 0 \) for all \( t \geq 0 \).

Using infinite sums of terms of the form (3.2.2), Bernoulli wrote down expressions which he asserted were the most general solutions of the vibrating string problem (3.2.1), i.e., solutions for the most general initial position \( f \). His argument was both formal and in terms of the physics of sound, cf., [BS82], [Pie83] for beautiful treatments of the fundamentals and harmonics used by Bernoulli. Later in 1753, Euler noted that Bernoulli’s claim of general solutions could only be correct if “arbitrary curves” \( f \) defined on \([0, L]\) could be written as, what were later called, Fourier series. Further, because of the periodicity of the individual terms in Bernoulli’s series, Euler judged that Bernoulli was incorrect as far as generality of solution, cf., [Bra86, pages 462-464]. Once again, ill-defined terms such as “arbitrary curves”, instead of a precise definition of function, were the root cause of these different opinions.

To add to the intellectual melee, Lagrange, at age 23, wrote in 1759 in support of Euler’s solution being the most general. Amazingly, his “proof” used trigonometric series similar to Bernoulli’s. For the case \( L = 1 \), Lagrange’s solution was essentially of the form

\[
\begin{align*}
 u(x, t) &= \int_0^1 \sum_{n=1}^{\infty} (\sin \pi nx \cos \pi nct) f(y) \sin \pi ny \, dy \\
&\quad + \frac{2}{c\pi} \int_0^1 \sum_{n=1}^{\infty} \frac{1}{n} (\sin \pi nx \sin \pi nct) g(y) \sin \pi ny \, dy.
\end{align*}
\]

(3.2.3)

Note that if \( t = 0 \) and if there is an interchange of summation and integration, then (3.2.3) gives rise to the Fourier series expansion of \( f \). Lagrange series were almost history! Lagrange seemed intent on verifying Euler’s claims versus d’Alembert. In another bizarre twist, Euler did the formal calculation of Remark 3.1.2 to compute Fourier coefficients in 1777 when Fourier was a 9-year old, cf., [Car30, page 4] for a similar contribution by Clairaut.

3.2.2 Fourier (1768-1830).
Fourier submitted his 234 page manuscript, "Sur la propagation de la chaleur", to the Institut de France in 1807. At that time, Fourier was almost 40, and had only three unrelated published papers. His work as Prefect of the Department of Isère in Grenoble dealt with drainage of marshlands, consultation on the achievements of Napoleon's Institut d'Egypte, modeled after the Institut de France, and with planning the first road between Grenoble and Torino, Lagrange's hometown.

The turbulent story of the evolution of this paper includes its critique by Lagrange, et al., a prize competition which Fourier won in 1812 along with Lagrange's reservations, the publication of the book [Fou1822] in 1822, and the disappearance of the original paper, e.g., [Grat72] along with some of the other references listed at the beginning of Section 3.2. Darboux rediscovered the paper at the École Nationale des Ponts et Chaussées in 1890.

As we saw in Section 3.2.1, the technology was already in place for Fourier series long before Fourier came on the scene. What did Fourier do? He never claimed discovery of the Fourier coefficients (3.1.2) that he used. However, he had a point of view which introduced a "new epoch", to use Riemann's phrase. In the 18th century, Fourier coefficients were an integral part (sic) of trigonometric series which had already been derived by other means. Fourier asserted that an arbitrary function could be expanded in a trigonometric series whose coefficients could be computed as in Remark 3.1.2. Such an assertion led to questions of convergence of series and integration of arbitrary functions (in the definition of Fourier coefficients) and, of course, to questions about the meaning of function. Fourier's examples and applications in [Fou1822] are extraordinary; and were influential in establishing the field of Fourier's series.

There are also related subsequent contributions by Cauchy and Poisson; but we shall go directly to Dirichlet.

3.2.3 Dirichlet (1805-1859).

Who gave the first correct definition of function? Scholars of good will and excellent credentials disagree, cf., [Monn72, especially pages 57-65]. As mentioned in Remark 1.7.7b, it seems to us that Dirichlet has a valid claim.
3.2. HISTORY OF FOURIER SERIES

In 1829 [Dir1829, page 121], he wrote that a continuous function $f$ on $[0, h]$ is defined by the property that it "has a finite and well determined value for each value of $\beta$ between 0 and $h$, and moreover such that the difference $f(\beta + \epsilon) - f(\beta)$ decreases without limit (to 0) when $\epsilon$ becomes smaller and smaller".

In 1837 [Dir1837], he wrote the following, in which the parenthetical remark shows that continuity was not an intrinsic part of his definition. "Imagine a and $b$ to be two fixed values and $x$ a variable, which is supposed to assume one after the other all values between $a$ and $b$. If to each $x$ there corresponds a unique finite $y$ in such a manner that while $x$ runs continuously through the interval from $a$ to $b$, $y = f(x)$ varies gradually also, then $y$ is a continuous function of $x$ for this interval. (Since in what follows, we shall only discuss continuous functions, this attribute can be omitted without loss.)" The lack of rigor in defining continuity in terms of the word "gradually" is compensated by his precision in 1829. "It is not at all necessary that $y$ depends on $x$ in this whole interval by the same law, and it is not even necessary to imagine a dependency expressible by mathematical operations. ...This definition does not prescribe a common law to the different parts of the curve; it can be thought of as being composed of parts of the most different kinds or completely without law." This last part addresses the confusion from the 18th century analysis, when a formula such as $f(x) = x^2$ on $[a, b]$ was often thought to characterize a function, instead of characterizations such as $f = 2$ on $[0, \frac{a+b}{2}]$ and $f(x) = x^2$ on $\left(\frac{a+b}{2}, b\right]$.

Of course, it was precisely in the papers [Dir1829] and [Dir1837], where Dirichlet defined the notion of function, that he also proved the fundamental Theorem 3.1.6. Dirichlet's theorem was generalized by Lipschitz in 1864, supposing so-called Lipschitz conditions; and generalized still further by Dini, who wrote an important book on Fourier series in 1880. In the spirit of Dini's analysis, there is the following Dini-Lipschitz-Lebesgue test for uniform convergence. If $F \in L^2(\mathbb{T}_2\mathbb{N})$ and

$$\lim_{\lambda \to 0} |F(\gamma + \lambda) - F(\gamma)| \log |\lambda| = 0$$

uniformly in an open interval $I$, then $S(F) = F$ on any closed subinterval $J \subseteq I$, and the convergence of the Fourier series $S(F)$ to $F$ is
uniform on \( J \), e.g., [HR56, Theorem 59], cf., [Zyg59, page 52] for the original Dini test.

For perspective, recall that the Jordan Theorem, Theorem 1.7.6, is the analogue for Fourier transforms of the Dirichlet Theorem, Theorem 3.1.6. The second mean value theorem (Lemma 1.7.3) was used to prove Theorem 1.7.6, and, in fact, Bonnet (Memoires des Savant Étrangers of the Belgian Academy, 23 (1948-1850) used Lemma 1.7.3 directly to prove Theorem 3.1.6. Dirichlet’s original proof in 1829 used an argument similar to that required to prove Lemma 1.7.3.

Dirichlet made major contributions to number theory. It is not difficult to prove that the sequence \( \{4n - 1 : n \in \mathbb{N}\} \) contains infinitely many primes. Dirichlet proved the general fact that if \( a \in \mathbb{N}, b \in \mathbb{Z}, \) and \( a \) and \( b \) are relatively prime, then \( \{an + b : n \in \mathbb{N}\} \) contains infinitely many primes, cf., Remark 3.8.11a.

In this book we shall refer to two other number theoretic issues where Dirichlet had seminal ideas. The first concerns the Dirichlet Box Principle, related to rational approximation of irrationals, and leading to the Kronecker Theorem which can be formulated in terms of trigonometric sums, e.g., Exercises 3.40 and 3.41. The second concerns a proof of Gauss’ Law of Quadratic Reciprocity, e.g., Remark 3.8.11. The material of Section 2.10 plays a role, as well as subtle issues concerning Gauss sums and the so-called Littlewood Flatness Problem, e.g., Remark 3.8.11.

3.2.4 Riemann (1826-1866).

Bernhard Riemann’s life was tragic in its briefness and transcendental in its brilliance. The excerpts in [Kli72] about Riemann’s ideas barely scratch the surface on the depth and breadth and lustre of his creativity, cf., [Edwa74] for implications of just one of his gems.

Riemann’s Habilitationsschrift [Rie1873] was presented in 1854 but was only published in 1867 after his death. It is the first part of this work which has provided us with some of the material in Sections 3.2.1–3.2.3. Next, Riemann developed the Riemann integral. His theory of integral was created to define Fourier coefficients and Fourier series expansions for a large class of functions. Finally, he developed the Riemann Localization Principle and several other important tools for
3.2. HISTORY OF FOURIER SERIES

dealing with trigonometric series, e.g., [Zyg59], cf., [Ben71] for the relation between the Riemann Localization Principle and the notion of support.

The Riemann Localization Principle is a key technique in the study of sets of uniqueness (U-sets). A set \( E \subseteq [0, 1) \) is a U-set if

\[
\lim_{N \to \infty} \sum_{|n| \leq N} c_n e^{-2\pi i n y} = 0 \text{ off of } E \implies c_n = 0 \text{ for all } n \in \mathbb{Z}.
\]

Using Riemann's theory, Cantor proved that the empty set \( \emptyset \) is a U-set, cf., Section 3.2.5. Cantor's theorem was apparently known to Riemann, e.g., [Leb06, page 110]. At the other extreme, if \( |E| > 0 \) then \( E \) is not a U-set, e.g., Exercise 3.37.

As is well-known, a bounded function \( f : [a, b] \to \mathbb{C} \) is Riemann integrable if and only if \( f \) is continuous a.e., e.g., [Ben76, pages 94-96]. The concept of measure 0 and the notation "a.e." (Definition A.4) are now part of the Lebesgue theory (1902). Leading to Lebesgue, Vito Volterra (1881), at that time a student of Dini at the Scuola Normale Superiore in Pisa, constructed functions \( f \) whose derivative exists everywhere but for which \( f' \) is not Riemann integrable, e.g., [Ben76, pages 20-21], cf., the interesting examples in [Rie873, Section 13]. Actually, H. J. Smith had solved the same problem in 1875; but Lebesgue was unaware of Smith's result in his thesis [Leb02], where he gives prominent mention of Volterra's example.

The point is that measure 0 was emerging in the late 19th century as an important idea. Norbert Wiener (1938) has made a case for formulating the notion of measure zero based on justifying Maxwell's and Gibbs' theory of statistical mechanics. He wrote that "the ideas of statistical randomness and phenomena of zero probability were current among the physicists and mathematicians in Paris around 1900, and it was in a medium, heavily ionized by these ideas that Borel and Lebesgue solved the mathematical problem of measure" [Wie81, Volume II, pages 794-806], cf., [Carl80].

Besides non-(Riemann) integrability, the issue of non-differentiability was prominent in this part of the 19th century. Weierstrass wrote, at least in a possibly edited version of a lecture he gave at the Royal Academy of Sciences on July 18, 1872, that "Riemann, as I learned
from some of his students, stated decisively (in 1861, or perhaps even earlier)\(^a\) that

\[
f(x) = \sum_{n=1}^{\infty} \frac{\sin n^2 x}{n^2}
\]

is an everywhere continuous nowhere differentiable function, e.g., Weierstrass' Mathematische Werke II, pages 71-74, cf., [BS\textsuperscript{86}] for a profound analysis of this area. Although (3.2.4) does have some points of differentiability, there are now many continuous nowhere differentiable functions including Weierstrass' lacunary Fourier series (1872),

\[
f(x) = \sum_{n=1}^{\infty} b^n \cos(\pi a^n x),
\]

where \(a > 1\) is an odd integer, \(b \in (0, 1)\) and \(ab > 1 + 3\pi/2\), cf., [Ben\textsuperscript{76}, pages 28-29] for other examples and [Dui\textsuperscript{91}] for the relation to selfsimilarity.

Riemann convalesced and toured in Italy during the winter of 1862, arriving in Pisa in 1863. He became friendly with Betti and Beltrami. Betti, of "Betti number" fame, was Director of the Scuola Normale Superior, and there is an interesting Betti-Riemann correspondence at the Scuola. (The Scuola Normale was started by Napoleon, and is modeled after the École Normale Supérieure in Paris.) Dini was a student at the Scuola at the time of Riemann's visit. He graduated in 1864, spent a year studying with Bertrand in Paris, and returned to the Scuola Normale where he spent the next 52 years. Besides Volterra, he counts Vitali as one of his students, e.g., [Ben\textsuperscript{76}] for historical remarks and mathematical contributions of Vitali.

Riemann returned to Germany for the winter of 1864-1865, but then came back to Pisa. He died and was buried at Biganzolo in the northern part of Verbania (the Italian resort town on the western banks of Lago Maggiore just 15 miles south of the Swiss border). A mourner at the local cemetery will surely point you to the marker of "Il Tedesco".

\textbf{3.2.5} Cantor (1845-1918).
3.2. HISTORY OF FOURIER SERIES

Georg Cantor received his Ph.D. in 1867 at Berlin. His dissertation on quadratic Diophantine equations, related to some issue from Gauss' monumental Disquisitiones Arithmeticae [Gau66], was written under the direction of Kummer, cf., [Ben77] for remarks about Kummer, Fermat's Last Theorem, and ideals. Kronecker, who later became an intellectual adversary of Cantor's, was also a professor in Berlin at the time.

Cantor wrote several important papers on U-sets in the early 1870s, including his theorem quoted in Remark 3.2.4. In order to prove this result, that the empty set is a U-set, he first proved what is now known as the Cantor-Lebesgue Lemma: if \( X \subseteq [0,1) \) is a Lebesgue measurable set of positive measure, and if

\[
\forall \gamma \in X, \quad \lim_{N \to \infty} \sum_{|n| \leq N} c_n e^{-2\pi i n \gamma} \in \mathbb{C},
\]

then \( \lim_{|n| \to \infty} c_n = 0 \). Cantor actually proved the result for the case that \( X \) is a nondegenerate interval. Fatou first investigated the converse. In this regard, Lusin found a trigonometric series which was a.e. divergent and for which \( \lim_{|n| \to \infty} c_n = 0 \). Steinhaus clinched the converse by constructing such a series which was everywhere divergent and for which \( \lim_{|n| \to \infty} c_n = 0 \), e.g., [Bary64, Volume I, pages 176-177].

After proving that the empty set was a U-set, Cantor showed that finite sets and certain countably infinite sets are also U-sets. This work certainly influenced his later research on set theory and the infinite.

It was in 1874 that he gave his famous, correct, and controversial proof that of the fact that there are only countably many algebraic numbers. Recall that an algebraic number is a zero of a polynomial with integer coefficients.

In any case, Cantor tried to prove that all countable sets \( E \subseteq [0,1) \) were U-sets; and this was finally achieved by F. Bernstein (1908) and W. H. Young (1909), cf., [But95]. Actually, Bernstein proved somewhat more, cf., Sections 3.2.6 and 3.2.7.

The remainder of Cantor's life, from the mid-1870s, was devoted to the study of the infinite, not only in mathematics as in [Can55], but often delving into various philosophical notions of infinity due to
the Greeks, the Scholastic philosophers, and his contemporaries, e.g., [Daub79]. Cantor certainly did not dote on philosophers. In a letter to Bertrand Russell, who was then at Trinity College, Cambridge, Cantor wrote (1911): "... and I am quite an adversary of Old Kant, who in my eyes has done much harm and mischief to philosophy, even to mankind; as you easily see by the perverted development of metaphysics in Germany in all that followed him, as in Fichte, Schelling, Hegel, Herbart, Schopenhauer, Hartman, Nietzsche, etc. etc. on to this very day. I never could understand why ... reasonable ... peoples ... could follow yonder sophistical Philistine, who was so bad a mathematician."

3.2.6 Mensov (1892-1988).

Dmitrii Mensov proved a key result on $U$-sets in 1916 by finding a non-$U$-set $X$ with Lebesgue measure $|X| = 0$. He did this just after graduating from Moscow University, where he wrote his thesis under N. Lusin. Mensov's example stimulated a great deal of study about sets of measure zero. Actually, on the basis of Mensov's example, Lusin and Bary defined the notion of $U$-sets as such. Earlier, de la Vallée-Poussin had proved that if a trigonometric series converged to $F \in L^1(\mathbb{T})$ off a countable set $E$, then the series is the Fourier series of $F$. It was generally felt that the same would be true for sets $E$ with $|E| = 0$. Mensov changed that perception. Mensov showed that there exists a nontrivial trigonometric series which converges to $F = 0$ off a set of measure zero. Since $F \in L^1(\mathbb{T})$ is the 0-function a.e., and since the series has some nonzero coefficients, the series is not the Fourier series of $F$. Needless to say, Mensov's example had a certain amount of shock value, cf., [Men68], [Ben76, page 115] for other major results by Mensov.

Nina Bary (1923) asked for conditions on the coefficients $\{c_n\}$ of trigonometric series to ensure that $c_n = 0$ for all $n$ whenever

\begin{equation}
\lim_{N \to \infty} \sum_{|n| \leq N} c_n e^{-2\pi iny} = 0 \text{ a.e. on } \mathbb{T}.
\end{equation}

In light of de la Vallée-Poussin's and Mensov's results of the previous paragraph, it is interesting to observe that if $\sum |c_n|^2 < \infty$ and (3.2.5) is true then $c_n = 0$ for all $n$, e.g., Exercise 3.36. There have been
deep results in this problem area, in the case $\Sigma |c_n|^2$ diverges, by Littlewood (1936), Wiener and Wintner (1938), A. C. Schaeffer (1939), Salem (1942), Ivašev-Musatov (1957), Brown and Hewitt (1980), and Körner [Kör87]. For example, Salem proved that for each $\epsilon > 0$ there is a (bounded measure) $\mu \in M(\mathbb{T})\setminus\{0\}$ for which $|\text{supp} \mu| = 0$ and

\begin{equation}
\exists C, N \text{ such that } \forall |n| \geq N, \left| \hat{\mu}[n] \right| \leq C \frac{1}{|n|^{1/2-\epsilon}},
\end{equation}

\[ \text{e.g., [KS63, pages 106–112], cf., [Ben75, pages 96–97]. If the right side of (3.2.6) were } C(1/|n|^{1/2+\epsilon}) \text{ then } \mu \in L^2(\mathbb{T}), \text{ e.g., Theorem 3.4.13; this coupled with the condition } |\text{supp} \mu| = 0 \text{ implies } \mu \text{ is the 0-function.} \]

With regard to (3.2.6) and the Lusin Conjecture of Section 3.2.8, deLeeuw, Kahane, and Katznelson proved the following result:

\[ \forall f \in L^2(\mathbb{T}), \exists g \in C(\mathbb{T}) \text{ such that } \forall n \in \mathbb{Z}, |f[n]| \leq |g[n]| \text{ and } \|G\|_{L^\infty(\mathbb{T})} \leq 9\|f\|_{L^2(\mathbb{T})} \]


3.2.7 Bary (1901-1961) and Rajchman (-1940).

What with Mensov’s example, Alexander Rajchman (who died at Dachau in 1940) “seems to have been the first to realize that for sets of measure zero that occur in the theory of trigonometric series it is not so much the metric as the arithmetic properties that matter” (from Zygmund’s biography of Salem in [Sal67]. Rajchman (1922) proved the existence of some uncountable, closed $U$-sets including the 1/3-Cantor set. He was motivated by some work of Hardy and Littlewood (Acta Math., 37(1914)), and Steinhua (1920), on diophantine approximation to introduce “$H$-sets”; and proved that such sets are $U$-sets. In fact, the 1/3-Cantor set is an $H$-set. In a letter to Lusin, he also expressed his considered opinion, that any $U$-set is contained in a countable union of $H$-sets. Although this particular conjecture was proved false by Pyatetskii-Shapiro(1952), it was such questions that focused the direction of the subject.

Actually, Nina Bary had proved the existence of some uncountable, closed $U$-sets in 1921, and presented her results at Lusin’s seminar at
the University of Moscow. They were unpublished at the time of Rajchman's paper. This does not undermine the importance of Rajchman's theorems, since Rajchman's approach illustrated the need for number theoretic (diophantine) properties in the construction of such sets.

Bary proved her first result on $U$-sets as an undergraduate, and made outstanding contributions to the subject throughout her life. One of her major results is that the countable union of closed $U$-sets is a $U$-set. The problem is open for the finite union of arbitrary $U$-sets. Another one of her theorems, which was proven in 1936–1937, asserts that if $\alpha$ is rational and $E(\alpha)$ is the Cantor set with ratio of dissection $\alpha$, then $E(\alpha)$ is a $U$-set if and only if $\beta = 1/\alpha$ is an integer. This generalizes Rajchman's result about the $1/3$-Cantor set $C$ since $C = E(1/3)$. In general, $E(\alpha)$ is constructed by "throwing away" centered open intervals of length $\alpha(b - a)$, where $[a, b]$ is any remaining closed interval at a given step, and where the first step begins with $[a, b] = [0, 1]$, e.g., [KS63, Chapitre I].

A Pisot-Vigayraghavan (P-V) number is a real algebraic integer $\beta > 1$ with the property that all the other roots of its minimal polynomial have modulus less than 1. Bary's theorem on Cantor sets of uniqueness has the following spectacular sequel announced by Salem (1943): if $\alpha \in (0, 1/2)$ then $E(\alpha)$ is a $U$-set if and only if $\beta = 1/\alpha$ is a P-V number, e.g., [Bary64], [Ben76, pages 116–117], [Mey72], [Sal63] for the proof, a history of the proof, and recent developments.

3.2.8 The Lusin Conjecture.

In his dissertation of 1915 (actually he published a Comptes Rendus Acad. Sci., Paris note on the relevant material), Lusin conjectured that the Fourier series of every $F \in L^2(\mathbb{T})$ is convergent a.e. This is the Lusin Conjecture.

As background for the Lusin Conjecture, du Bois-Reymond (1872) constructed functions $F \in C(\mathbb{T})$ whose Fourier series diverge at some points, cf., [Rog59, pages 75–77], [Zyg59, Volume I, Chapter VIII] and Exercise 3.45. Further, just prior to Lusin, there were contributions in this general area by Fatou(1906), Jerosch and Weyl (1908), Weyl (1909), Fejér (1911), W. H. Young (1912), Hobson(1913), Plancherel (1913), and Hardy (1913).
In 1926, Kolmogorov constructed functions $F \in L^1(\mathbb{T}) \setminus L^2(\mathbb{T})$ whose Fourier series diverge everywhere! His proof used Kronecker’s Theorem, which we shall discuss in Section 3.2.10. There were subsequent relevant “log estimates” by Kolmogorov and Seliverstov (1925), Plessner (1926), and Littlewood and Paley (1931). Finally, Lennart Carleson (1966) proved that if $F \in L^2(\mathbb{T})$ then $S(F) = F$ a.e. [Carl66], cf., [Fel73] for a conceptually different proof and [Moz71] for a superb exposition. R. A. Hunt (1968) used the method of Carleson’s proof and the theory of interpolation of operators to extend Carleson’s result to $L^p(\mathbb{T}), p > 1$, cf., [Ash76, pages 20–37] for an elegant presentation by Hunt, and [Ben76, pages 208–210] for a connection between techniques used by Carleson and the FTC. Also, for perspective vis à vis du Bois-Reymond’s example and Carleson’s Theorem, we have Kahane and Katznelson’s Theorem that if $E \subseteq \mathbb{T}$ is a set of measure zero then there is $F \in C(\mathbb{T})$ such that $S(F)(\gamma)$ diverges for all $\gamma \in E$, e.g., [Kat76, Chapter 2].

We close this section with remarks by Carleson on the occasion of receiving the 1984 Steele Prize. They concern his proof and a remark about the FFT, cf., Section 3.9.

“When I was a student at Harvard in 1950–1951, A. Zygmund and R. Salem were also in Cambridge and I learned very much from them. They also encouraged me to try to use Blaschke products as examples of a Fourier series which diverges a.e. I worked hard at that then, and all through the years I tried different ideas. Then finally, in 1964 or so, I realized the basic reason why there should exist an example. Very briefly we can describe the main feature of the trigonometric system $\cos nx, n \leq 2^m$, by writing down a matrix of $\pm 1$ giving the sequence of sign$(\cos nx)$ which can occur. This matrix is essentially $2^m \times 2^m$, i.e., very few sequences of signs occur which, of course, is very favorable for examples of divergence. (This is also the basic idea behind the fast Fourier transform.) To my great astonishment, it now turned out that for a random $2^m \times 2^m$ matrix there is no example and then a proof of the convergence theorem
came naturally."

3.2.9 The Dirichlet Box Principle.

a. The Dirichlet Box Principle asserts that if $Q$ boxes contain $Q + 1$ objects, then at least one of the boxes contains more than one object. This fact may not seem to be at the usual Dirichlet level of brilliance, but it has been a staple in the method of proof of many results since he first made use of it.

An adaptation of the Dirichlet Box Principle is even used in Wiles' proof of Fermat's Last Theorem. In this case the objects are Hecke rings and an infinite sequence of sets of boxes is created. The assertion, in the part of the proof due to Taylor and Wiles, is that there are Hecke rings in every set of boxes.

b. Originally, Dirichlet used the box principle to give a new proof of the fact that if $x \in (0, 1)$ is irrational then

$$\forall \epsilon > 0, \exists p, q \in \mathbb{N} \text{ such that}$$

$$\left| x - \frac{p}{q} \right| < \epsilon \text{ and } \frac{1}{q^2}. \quad (3.2.7)$$

The pairs $p, q$ can be chosen to be relatively prime. (3.2.7) is an elementary result in Diophantine approximation, and Dirichlet's proof (in part c) has the advantage of being applicable to $d$-dimensional problems, e.g., [HW65, Theorem 201]. The first inequality of (3.2.7) follows from basic properties of $\mathbb{R}$; and the second inequality gives insight into the rapidity of rational approximation to irrationals.

c. To prove (3.2.7), let $Q > 1$, and consider the $Q$ boxes $[\frac{n}{Q}, \frac{n+1}{Q}], n = 0, \ldots, Q - 1$, and the $Q + 1$ numbers, $0, x - [x], 2x - [2x], \ldots, Qx - [Qx]$. By the Dirichlet Box Principle there are integers $0 \leq q_1 < q_2 \leq Q$ and $p \in \mathbb{N}$ for which $|q_2 x - p| < 1/Q$, where $q = q_2 - q_1 \in (0, Q] \cap \mathbb{N}$ and $p = [q_2 x] - [q_1 x]$. Thus the second inequality of (3.2.7) is valid. This part of the proof also works for rational $x$.

Since $\epsilon > 0$ is given in (3.2.7), we choose $Q = Q(\epsilon) = \left[ \frac{1}{\epsilon} \right] + 1$ for the above argument. In particular, $\epsilon \geq 1/Q$ and so $|x - \frac{p}{q}| < \epsilon/q < \epsilon$.

3.2.10 Kronecker Sets and $U$-Sets.
3.3. INTEGRATION AND DIFFERENTIATION OF FOURIER SERIES

A significant refinement of (3.2.7), which is deeper than d-dimensional versions of (3.2.7), is Kronecker's Theorem. Kronecker (1884) proved that if \{1, γ_1, \ldots, γ_d\} ⊆ \mathbb{R} is linearly independent over the rationals, if \{λ_1, \ldots, λ_d\} ⊆ \mathbb{R}, and if ε, N > 0, then there are integers q > N and \(p_1, \ldots, p_d\) such that

(3.2.8) \quad \forall j = 1, \ldots, d, \quad |qγ_j - p_j - λ_j| < ε.

Dirichlet's analysis was for the case λ_j = 0. With the same hypotheses, the conclusion (3.2.8) can be reworded to assert the existence of q for which

(3.2.9) \quad \forall j = 1, \ldots, d, \quad |e^{2πiqt_j} - e^{2πiλ_j}| < ε.

There are several different proofs, e.g., [Ben75, Theorem 3.2.7], [HW65, Chapter 23], [KK64], [Kat76, pages 181-183].

Because of (3.2.9) we say that a closed set \(E \subseteq \mathbb{T}\) is a Kronecker set if for each ε > 0 and continuous function \(F : E \to \mathbb{C}\), for which \(|F| = 1\) on \(E\), there is \(q \in \mathbb{Z}\) such that

\[
\sup_{\gamma \in E} |e^{2\pi i q \gamma} - F(\gamma)| < \epsilon.
\]

In 1962, Paul Malliavin proved that if \(E \subseteq \mathbb{T}\) is closed and if every closed subset of \(E\) is a set of spectral synthesis, then \(E\) is a \(U\)-set. In 1965, Nicholas Varopoulos proved that Kronecker sets satisfy the hypothesis of Malliavin's result; and, hence, Kronecker sets are \(U\)-sets. This is the "tip of the iceberg," cf., [Rud62], [KS63], [Kah70], [Ben71], [LP71], [Mey72].

3.3 Integration and differentiation of Fourier series

3.3.1 Example. Integration of Series

If \(\sum_{n=1}^{\infty} F_n\) is a uniformly convergent series of continuous functions \(F_n\) on \([0, 1]\), then

(3.3.1) \quad \int_0^1 \left( \sum_{n=1}^{\infty} F_n(\gamma) \right) \, d\gamma = \sum_{n=1}^{\infty} \int_0^1 F_n(\gamma) \, d\gamma,
CHAPTER 3. FOURIER SERIES

e.g., [Apo57]. On the other hand, it is well known that the hypotheses, 
\{F_n : n = 1, \ldots \} \subseteq C[0,1] and \sum_{n=1}^{\infty} F_n = F \in C[0,1], where the 
convergence is pointwise on [0,1], are not sufficient to ensure (3.3.1), 
e.g., [Har49], [Ben76, Section 3.3]. For example, if we let 
F_1(\gamma) = 
\gamma(1 - \gamma) 
and 
F_n(\gamma) = n^2 \gamma(1 - \gamma)^n - (n - 1)^2 \gamma(1 - \gamma)^{n-1} for \ n \geq 1, 
then it is easy to see that each \ F_n \in C[0,1] (and each \ F_n \in C(T) since 
F_n(0) = F_n(1) = 0), \sum_{n=1}^{\infty} F_n \equiv F = 0 on [0,1], \ 
\int_0^1 \left( \sum_{n=1}^{\infty} F_n(\gamma) \right) d\gamma = 0, and

\[ \sum_{n=1}^{\infty} \int_0^1 F_n(\gamma) d\gamma = \lim_{N \to \infty} \sum_{n=1}^{N} \int_0^1 F_n(\gamma) d\gamma = \lim_{N \to \infty} \int_0^1 N^2 \gamma(1 - \gamma)^N d\gamma = \lim_{N \to \infty} \frac{N^2}{(N + 1)(N + 2)} = 1, \]

cf., Section 3.4 where we discuss approximate identities for \ L^1(T).

The following theorem is a remarkable feature of Fourier series. It asserts that (3.3.1) is valid when the series \ \sum_{n=1}^{\infty} F_n is replaced by the 
Fourier series of any function in \ L^1(\mathbb{R}). In particular, the Fourier series 
to be integrated can diverge everywhere, as in Kolmogorov's result 
mentioned in Section 3.2.8, cf., Exercise 3.45.

3.3.2 Theorem. Integration of Fourier Series

Let \( F \in L^1(T_{2\Omega}) \). The Fourier series \( S(F) \) of \( F \), with Fourier 
coefficients \( f = \{f[n]\} \), can be integrated term by term, i.e.,

\[ \forall \alpha, \beta \in \mathbb{R}, \quad \int_{\alpha}^{\beta} S(F)(\gamma) d\gamma = \int_{\alpha}^{\beta} F(\gamma) d\gamma, \tag{3.3.2} \]

where the left side of (3.3.2) denotes

\[ \sum f[n] \int_{\alpha}^{\beta} e^{-i n \gamma / \Omega} d\gamma. \]

Proof. Define 

\[ G(\gamma) = \frac{1}{2\Omega} \int_{0}^{\gamma} (F(\lambda) - f[0]) d\lambda \]
3.3. INTEGRATION AND DIFFERENTIATION OF FOURIER SERIES

for $\gamma \in [0, 2\Omega)$. Consequently, $G(0) = G(2\Omega) = 0$, and $G$ can be extended $2\Omega$-periodically to $\mathbb{R}$ with the property that $G \in AC_{loc}(\mathbb{R})$. Let $g = \{g[n]\}$ be the sequence of Fourier coefficients of $G$. As such, we can apply the Dirichlet Theorem, properly modified as in Remark 3.1.7b, to assert that $S(G) = G$ on $\mathbb{T}_{2\Omega}$, i.e.,

$$
(3.3.3) \quad \forall \gamma \in \mathbb{T}_{2\Omega}, \quad G(\gamma) = g[0] + \sum' g[n] e^{-\pi in\gamma/\Omega},
$$

where $\sum'$ denotes summation over $\mathbb{Z}\setminus\{0\}$. For $n \neq 0$, we compute

$$
g[n] = \frac{1}{2\Omega} \int_0^{2\Omega} \left[ \frac{1}{2\Omega} \int_0^\gamma \left( F(\lambda) - f[0] \right) d\lambda \right] e^{\pi in\gamma/\Omega} d\lambda
$$

$$
= \frac{1}{(2\Omega)^2} \frac{\Omega}{\pi in} e^{\pi in\gamma/\Omega} \bigg|_0^\gamma \left( F(\lambda) - f[0] \right) d\lambda - \frac{1}{(2\Omega)^2} \frac{\Omega}{\pi in} \int_0^{2\Omega} \left( F(\gamma) - f[0] \right) e^{\pi in\gamma/\Omega} d\gamma
$$

$$
= -\frac{f[n]}{2\pi in}.
$$

The last equation follows since $\frac{1}{2\Omega} \int_0^{2\Omega} F(\lambda) d\lambda = f[0]$ and $\int_0^{2\Omega} e^{\pi in\gamma/\Omega} d\gamma = 0$. Integration by parts is allowable by the (local) absolute continuity. Combining the above computation for $g[n]$ with (3.3.3), we obtain

$$
(3.3.4) \quad \forall \gamma \in \mathbb{T}_{2\Omega}, \quad G(\gamma) = g[0] - \sum' \frac{f[n]}{2\pi in} e^{-\pi in\gamma/\Omega},
$$

so that, since $G(0) = 0$,

$$
(3.3.5) \quad g[0] = \sum' \frac{f[n]}{2\pi in}.
$$

In particular, the series $\sum' f[n]/n$ converges, cf., Remark 3.1.7a about symmetric convergence.

By definition of $G$, (3.3.4) becomes

$$
(3.3.6) \quad \frac{1}{2\Omega} \int_0^{\gamma} F(\lambda) d\lambda = \frac{\gamma}{2\Omega} f[0] + g[0] - \sum' \frac{f[n]}{2\pi in} e^{-\pi in\gamma/\Omega}.
$$
By definition of \( f^\beta S(F)(\lambda) \, d\lambda \) we have

\[
\frac{1}{2\Omega} \int_0^\gamma S(F)(\lambda) \, d\lambda = \frac{1}{2\Omega} \sum f[n] \int_0^\gamma e^{-i\pi n\lambda/\Omega} \, d\lambda
\]

\[
= \frac{\gamma}{2\Omega} f[0] - \frac{1}{2\Omega} \sum' f[n] \frac{\Omega}{\pi i n} (e^{-i\pi n\gamma/\Omega} - 1)
\]

\[
= \frac{\gamma}{2\Omega} f[0] - \frac{1}{2\Omega} \sum' f[n] \frac{\Omega}{\pi i n} e^{-i\pi n\gamma/\Omega} + g[0].
\]

Consequently, (3.3.2) is obtained for the interval \([0, \gamma] \subseteq [0, 2\Omega]\) by combining (3.3.6) with this last calculation.

The case for the interval \([\alpha, \beta] \subseteq [0, 2\Omega]\) is obtained by writing \( f^\beta = f^\alpha + f^\beta = f^\beta - f^\alpha \); and then making the natural adjustments for other values of \( \alpha \) and \( \beta \). \( \square \)

The following result is a consequence of (3.3.5).

3.3.3 Corollary. A Property of Fourier Coefficients \( f \in A(Z) \)

Let \( F \in L^1(T_{2\Omega}) \). The series,

\[
\sum' f[n] \frac{n}{n}
\]

converges, where \( f = \{f[n]\} \) is the sequence of Fourier coefficients of \( F \).

3.3.4 Example. Necessary Conditions for Integrability of Trigonometric Series

a. The Riemann-Lebesgue Lemma asserts that if \( F \in L^1(T_{2\Omega}) \) then \( \lim_{|n| \to \infty} f[n] = 0 \), where \( f = \{f[n]\} \) is the sequence of Fourier coefficients of \( F \). On the other hand, suppose we are given a trigonometric series \( \sum c_n e^{-2i\pi n\gamma} \) for which \( \lim_{|n| \to \infty} c_n = 0 \). Is there any way we can determine if this series is or is not the Fourier series of some function \( F \in L^1(T_{2\Omega}) \), cf., Section 3.2.6?

We can assert that the trigonometric series,

\[
\sum_{n=2}^{\infty} \frac{\sin(\pi n\gamma/\Omega)}{\log n},
\]

(3.3.7)
3.3. INTEGRATION AND DIFFERENTIATION OF FOURIER SERIES

is not the Fourier series of an element \( F \in L^1(\mathbb{T}_{2\Omega}) \) even though the coefficients tend to 0 at \( \pm \infty \), à la the necessary conditions given by the Riemann-Lebesgue Lemma. To verify this claim we argue as follows. First, we write (3.3.7) as

\[
-\frac{1}{2i} \sum_{|n| \geq 2} \frac{\text{sgn } n}{|n| \log |n|} e^{-\pi in\gamma/\Omega},
\]

where \( \text{sgn } n = n/|n| \). Then we apply Corollary 3.3.3 to observe that if (3.3.7) were a Fourier series (of an element \( F \in L^1(\mathbb{T}_{2\Omega}) \)) then

\[
\sum_{|n| \geq 2} \frac{1}{|n| \log |n|} < \infty,
\]

which is false by the integral test, e.g., Exercise 3.10. Thus, we have constructed a sequence \( f \in c_0(\mathbb{Z}) \setminus A(\mathbb{Z}) \), cf., Example 1.4.4 where we constructed the analogue of (3.3.7) for the case \( C_0(\mathbb{R}) \setminus A(\mathbb{R}) \). An example \( F \in C(\mathbb{T}_{2\Omega}) \setminus A(\mathbb{T}_{2\Omega}) \) is constructed in Exercise 3.4.3, cf., Exercise 3.4.5 and du Bois-Reymond's example mentioned in Section 3.2.8.

b. It does turn out, however, that the series (3.3.7) converges pointwise for each \( \gamma \in \mathbb{R} \), e.g., Exercise 3.2.9b.

3.3.5 Theorem. Differentiation of Fourier Series

Let \( F \in AC(\mathbb{T}_{2\Omega}) \). Then \( F' \in L^1(\mathbb{T}_{2\Omega}) \) (ordinary differentiation) and

\[ S'(F) = S(F'), \]

where \( S'(F) \) denotes the term by term differentiated series

\[
-\sum' \frac{\pi in}{\Omega} f[n] e^{-\pi in\gamma/\Omega},
\]

and where \( f = \{f[n]\} \) is the sequence of Fourier coefficients of \( F \).

Proof. Clearly, \( F' \in L^1(\mathbb{T}_{2\Omega}) \), e.g., Remark A.21. By the absolute continuity and 2\( \Omega \)-periodicity of \( F \), and by the FTC, we compute

\[
(F')^\wedge [n] = -\frac{1}{2\Omega} \frac{\pi in}{\Omega} \int_{-\Omega}^{\Omega} F(\gamma) e^{\pi in\gamma/\Omega} d\gamma = -\frac{\pi in}{\Omega} \hat{F} [n]
\]
for $n \neq 0$. If $n = 0$, then

$$
(F'')''[0] = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F'(\gamma) d\gamma = 0
$$

since the absolute continuity again allows us to use FTC. The result is obtained. \hfill \Box

3.3.6 Example. $\zeta(2)$
a. Let $F(\gamma) = \frac{\pi - \gamma}{2}$ on $[0, 2\pi)$, and consider $F$ as an element of $L^1(T_{2\pi})$.

---

**Figure 3.1**

We shall compute $S(F)$. $F$ is odd on $(-\pi, \pi)$, and we have

$$
S(F)(\gamma) = \sum_{n=1}^{\infty} b_n \sin n\gamma,
$$

where

$$
\forall n \geq 1, \quad b_n = \frac{1}{\pi} \int_{0}^{2\pi} \frac{1}{2} (\pi - \gamma) \sin n\gamma d\gamma.
$$

We calculate

$$
b_n = -\frac{1}{2\pi} \left[ -\frac{\gamma}{n} \cos n\gamma \right]_{0}^{2\pi} + \frac{1}{n} \int_{0}^{2\pi} \cos n\gamma d\gamma = \frac{1}{n}.
$$

Thus,

$$
S(F)(\gamma) = \sum_{n=1}^{\infty} \frac{\sin n\gamma}{n}.
$$
3.3. INTEGRATION AND DIFFERENTIATION OF FOURIER SERIES

By the Dirichlet Theorem, $S(F)(\gamma) = F(\gamma)$ on $\mathbb{R} \setminus \{2\pi n\}$ and $S(F)(2\pi n) = 0$.

b. Since $F \in L^1(T_{2\pi})$, we have

$$\forall \gamma \in [0, 2\pi], \quad \int_0^\gamma S(F)(\lambda) \, d\lambda = \int_0^\gamma F(\lambda) \, d\lambda,$$

by Theorem 3.3.2. This becomes

$$\forall \gamma \in [0, 2\pi], \quad \sum_{n=1}^\infty \frac{1}{n} \int_0^\gamma \sin n\lambda \, d\lambda = \frac{\gamma\pi}{2} - \frac{\gamma^2}{4}$$

and the left side is $-\sum_{n=1}^\infty \frac{1}{n^2} (\cos n\gamma - 1)$. Thus

$$\forall \gamma \in [0, 2\pi], \quad \frac{\gamma}{4}(2\pi - \gamma) = -\sum_{n=1}^\infty \frac{1}{n^2} (\cos n\gamma - 1).$$

c. Integrating both sides of the last expression in part b we obtain

$$\int_0^{2\pi} \frac{\gamma}{4}(2\pi - \gamma) \, d\gamma$$

$$= 2\pi \sum_{n=1}^\infty \frac{1}{n^2} - \sum_{n=1}^\infty \frac{1}{n^2} \int_0^{2\pi} \cos n\gamma \, d\gamma = 2\pi \sum_{n=1}^\infty \frac{1}{n^2}.$$

Consequently,

$$\sum_{n=1}^\infty \frac{1}{n^2} = \frac{\pi^2}{6}.$$  \hspace{1cm} (3.3.8)

d. The Riemann $\zeta$-function $\zeta(s) = \sum_{n=1}^\infty 1/n^s$, discussed in Example 2.4.6g, was defined by Pietro Mengoli (1625-1686) in 1650. He showed that the harmonic series $\zeta(1)$ diverges. The problem of evaluating $\zeta(k)$, for integers $k \geq 2$, attracted the attention of British mathematicians, including James Gregory (1638-1675). Henry Oldenburg (c.1615-1677), first Secretary of the Royal Society in London, wrote to Gottfried Leibniz (1646-1716) asking him to evaluate $\zeta(2)$. This occurred during Leibniz's visit to London in 1673. In 1696, Leibniz admitted his inability to solve this problem. Earlier, in 1689, Jakob
Bernoulli (1654-1705), Daniel's uncle, had also tried and apparently given-up summing \( \zeta(2) \), cf., [P6154, pages 17–22], [Kli72, pages 448–449], [Ebe83] for tantalizing historical remarks and incisive analysis.

Around 1736, Euler was able to state (3.3.8) by an ingenious argument outlined in Exercise 3.28, cf., Exercise 3.38. His *Introductio in Analysin Infinitorum* (1748) contains (3.3.8) and many similar results. In light of our discussion in Sections 3.2.1-3.2.3, it should also be pointed out that the *Introductio* also "defines" a function of a "variable quantity" as "any analytic expression whatsoever made up from that variable quantity and from numbers or constant quantities."

The following result can be used to prove the Classical Sampling Theorem (Theorem 3.10.10), e.g., [Ben92b, pages 447–449]. The proof of Theorem 3.3.7 is similar to that of Theorem 3.3.2, e.g., Exercise 3.15.

**3.3.7 Theorem. Integration of Fourier Series – A Refinement**

Let \( F \in L^1(\mathbb{T}_{2\Omega}) \) and \( G \in BV(\mathbb{T}_{2\Omega}) \). Then

\[
\forall \alpha, \beta \in \mathbb{R}, \quad \int_{\alpha}^{\beta} S(F)(\gamma)G(\gamma) \, d\gamma = \int_{\alpha}^{\beta} F(\gamma)G(\gamma) \, d\gamma,
\]

where the left side of (3.3.9) denotes

\[
\sum_{n} f[n] \int_{\alpha}^{\beta} G(\gamma)e^{-\pi in\gamma/\Omega} \, d\gamma,
\]

and where \( f = \{f[n]\} \) is the sequence of Fourier coefficients of \( F \).

**3.3.8 Remark. Integration of Fourier Series – Evocations**

a. If \( \beta - \alpha = 2\Omega \) in Theorem 3.3.7, then (3.3.9) becomes Parseval's formula,

\[
\lim_{N \to \infty} \sum_{|n| \leq N} f[n]g[-n] = \frac{1}{2\Omega} \int_{\alpha}^{\alpha+2\Omega} F(\gamma)G(\gamma) \, d\gamma,
\]

where \( g = G' \). With the adjustment as in Proposition 1.10.4, we obtain Parseval’s formula in the form,

\[
\lim_{N \to \infty} \sum_{|n| \leq N} f[n]g[n] = \int_{\mathbb{T}_{2\Omega}} F(\gamma)G(\gamma) \, d\gamma,
\]
3.3. INTEGRATION AND DIFFERENTIATION OF FOURIER SERIES

where \( F \in L^1(\mathbb{T}_{2\pi}) \), \( G \in BV(\mathbb{T}_{2\pi}) \), and \( \{f[n]\} \) and \( \{g[n]\} \) are the Fourier coefficients of \( F \) and \( G \), respectively, cf., Theorem 3.4.12b and Exercise 3.4.7 for other statements of Parseval's formula.

b. With respect to Corollary 3.3.3, we recall the following Hardy and Littlewood Theorem: if \( F \in L^1(\mathbb{T}_{2\pi}) \) and \( f[n] = 0 \) for all \( n < 0 \), where \( f = \{f[n]\} \) is the sequence of Fourier coefficients of \( F \), then

\[
\sum_{n=0}^{\infty} \frac{|f[n]|}{n + 1} \leq \pi \|F\|_{L^1(\mathbb{T}_{2\pi})} < \infty.
\]

This result is difficult to prove. One proof involves a fundamental factorization theorem for the so-called Hardy space \( H^1(\mathbb{T}_{2\pi}) \), as well as the following double series theorem due to Hilbert: if \( f, g \in L^2(\mathbb{N} \cup \{0\}) \) then

\[
\sum_{i,k=0}^{\infty} \frac{|f[i]|g[k]|}{j + k + 1} \leq \pi \|f\|_{L^2(\mathbb{N} \cup \{0\})} \|g\|_{L^2(\mathbb{N} \cup \{0\})},
\]

where \( \pi \) is the best possible constant, as proved by Schur, e.g., [Hel83, pages 94-99], cf., Exercise 3.3.4 for related material on Hilbert transforms of sequences.

c. It is well known that the desirable statement,

\[
(3.3.11) \quad \lim_{N \to \infty} \|S_N(F) - F\|_{L^1(\mathbb{T}_{2\pi})} = 0,
\]

is not true for all \( F \in L^1(\mathbb{T}_{2\pi}) \), e.g., Example 3.4.9.

On the other hand, a sequence \( \{F_N\} \subseteq L^1(\mathbb{T}_{2\pi}) \) converges to \( F \in L^1(\mathbb{T}_{2\pi}) \) weakly, i.e.,

\[
\forall G \in L^\infty(\mathbb{T}_{2\pi}), \quad \lim_{N \to \infty} \int_{\mathbb{T}_{2\pi}} (F_N(\gamma) - F(\gamma))G(\gamma) \, d\gamma = 0,
\]

if and only if

\[
(3.3.12) \quad \lim_{N \to \infty} \int_A (F_N(\gamma) - F(\gamma)) \, d\gamma = 0
\]

for every Lebesgue measurable set \( A \subseteq \mathbb{T}_{2\pi} \), e.g., [RN55, page 89]. If we have weak convergence, or, equivalently, (3.3.12), then (3.3.11) is true for \( S_N(F) = F_N \) if \( \{F_N\} \) converges to \( F \) in measure, e.g., [Ben76,
page 226]. (Of course, norm convergence always implies convergence in measure.) Further, Dicudonné and Grothendieck proved that (3.3.12) is true for all Lebesgue measurable sets \( A \subseteq \mathbb{T}_{2\Omega} \) if and only if it is true for all open sets \( A \subseteq \mathbb{T}_{2\Omega} \), e.g., [Ben76, page 225].

Because of this general relationship between weak and norm convergence, and because we would like to have (3.3.11) (or at least know how close we are to it), we note the following reformulation of Theorem 3.3.2: Theorem 3.3.2 is (3.3.12) for all intervals \( A \) in the case \( F_N = S_N(F) \). Further, Theorem 3.3.7 gives (3.3.12) in this case for all finite unions \( A \) of intervals.

### 3.4 The \( L^1(\mathbb{T}) \) and \( L^2(\mathbb{T}) \) theories

We showed in Remark 3.1.3 that \( L^2(\mathbb{T}) \subseteq L^1(\mathbb{T}) \), and have already noted in Section 3.2.8 that Fourier series of \( L^2(\mathbb{T}) \) functions converge a.e., whereas there are \( L^1(\mathbb{T}) \) functions whose Fourier series diverge at every point. Conceptually there are deeper differences between \( L^1(\mathbb{T}) \) and \( L^2(\mathbb{T}) \) than the fact that larger spaces may allow more instances of unusual behavior; briefly, \( L^1(\mathbb{T}) \) has algebraic properties and \( L^2(\mathbb{T}) \) has geometric properties which characterize their Fourier analysis.

#### 3.4.1 Definition. Convolution

a. Let \( F, G \in L^1(\mathbb{T}_{2\Omega}) \). The convolution of \( F \) and \( G \), denoted by \( F * G \), is

\[
F * G(\gamma) = \int_{\mathbb{T}_{2\Omega}} F(\gamma - \lambda)G(\lambda) \, d\lambda = \int_{\mathbb{T}_{2\Omega}} F(\lambda)G(\gamma - \lambda) \, d\lambda.
\]

(Recall that \( f_{\mathbb{T}_{2\Omega}} \) designates \( \frac{1}{2\Omega} \int_{\alpha}^{\alpha+2\Omega} \) for any fixed \( \alpha \in \mathbb{R} \).) As with \( L^1(\mathbb{R}) \) and Exercise 1.31, it is not difficult to prove that \( F * G \in L^1(\mathbb{T}_{2\Omega}) \) and

\[
\forall F, G \in L^1(\mathbb{T}_{2\Omega}),
\]

\[
\|F * G\|_{L^1(\mathbb{T}_{2\Omega})} \leq \|F\|_{L^1(\mathbb{T}_{2\Omega})}\|G\|_{L^1(\mathbb{T}_{2\Omega})}.
\]

b. \( L^1(\mathbb{T}_{2\Omega}) \) is a commutative algebra taken with the operations of addition and convolution, i.e., \( L^1(\mathbb{T}_{2\Omega}) \) is a vector complex vector space
under addition, and convolution is distributive with respect to addition, as well as being associative, and commutative.

3.4.2 Proposition.

Let \( F, G \in L^1(\mathbb{T}_{2\pi}) \), with corresponding sequences \( f = \{f[n]\}, g = \{g[n]\} \in A(\mathbb{Z}) \) of Fourier coefficients, i.e., \( \hat{f} = F \) and \( \hat{g} = G \). Then \( fg = \{f[n]g[n]\} \in A(\mathbb{Z}) \) is the sequence of Fourier coefficients of \( F \ast G \in L^1(\mathbb{T}_{2\pi}) \), i.e.,

\[
\forall n \in \mathbb{Z}, \quad f[n]g[n] = \int_{\mathbb{T}_{2\pi}} F \ast G(\gamma) e^{inx/\Omega} d\gamma.
\]

The proof of Proposition 3.4.2 is the same as that of Proposition 1.5.2.

3.4.3 Definition. Approximate Identity

An approximate identity is a family \( \{K(\lambda) : \lambda > 0\} \subseteq L^1(\mathbb{T}_{2\pi}) \) of functions with the properties

a. \( \forall \lambda > 0, \int_{\mathbb{T}_{2\pi}} K(\lambda)(\gamma) d\gamma = 1 \),

b. \( \exists C \geq 1 \) such that \( \forall \lambda > 0, \|K(\lambda)\|_{L^1(\mathbb{T}_{2\pi})} \leq C \),

c. \( \forall \eta \in (0, \Omega], \lim_{\lambda \to \infty} \frac{1}{2\pi} \int_{|\gamma| \leq \eta} |K(\lambda)(\gamma)| d\gamma = 0 \).

3.4.4 Theorem. Approximate Identity Theorem

a. Let \( F \in C(\mathbb{T}_{2\pi}) \) and let \( \{K(\lambda)\} \subseteq L^1(\mathbb{T}_{2\pi}) \) be an approximate identity. Then

\[
(3.4.2) \quad \lim_{\lambda \to \infty} \|F - F \ast K(\lambda)\|_{L^\infty(\mathbb{T}_{2\pi})} = 0.
\]

b. Let \( F \in L^1(\mathbb{T}_{2\pi}) \) and let \( \{K(\lambda)\} \subseteq L^1(\mathbb{T}_{2\pi}) \) be an approximate identity. Then

\[
(3.4.3) \quad \lim_{\lambda \to \infty} \|F - F \ast K(\lambda)\|_{L^1(\mathbb{T}_{2\pi})} = 0,
\]

cf., Theorem 1.6.9a.
The proof of Theorem 3.4.4a follows the proof of Theorem 1.6.9a, but for the case of \( \| \cdots \|_{L^\infty(T_{2\pi})} \) instead of \( \| \cdots \|_{L^1(\mathbb{R})} \). In fact, by the uniform continuity of \( F \), (1.6.6) can be replaced by the statement that for each \( \varepsilon > 0 \) there is \( \eta > 0 \) for which

\[
\forall |\lambda| < \eta, \quad \| F - \tau_{\lambda} F \|_{L^\infty(T_{2\pi})} < \varepsilon/C.
\]

This allows us to prove

\[
\lim_{\lambda \to \infty} \| F - F * K_{(\lambda)} \|_{L^\infty(T_{2\pi})} \leq \varepsilon
\]

analogous to the proof of Theorem 1.6.9a.

The proof of Theorem 3.4.4b follows from part a, e.g., Exercise 3.16.

3.4.5 Example. The Dirichlet and Fejér Kernels

a. The Dirichlet function \( D_N \) on \( T_{2\pi} \) is defined by

\[
(3.4.4) \quad \forall \gamma \in \mathbb{R}, \quad D_N(\gamma) = \sum_{|n| \leq N} e^{-\pi i n \gamma / \Omega}.
\]

It is not difficult to show that the trigonometric polynomial \( D_N \) can be written as

\[
(3.4.5) \quad D_N(\gamma) = \frac{\sin(N + \frac{1}{2}) \pi \gamma / \Omega}{\sin \frac{\pi \gamma}{2\Omega}},
\]

where \( D_N(2k\Omega) = 2N + 1 \) for \( k \in \mathbb{Z} \), e.g., Exercise 3.11a. The family \( \{D_N : N \in \mathbb{N} \cup \{0\}\} \) is the Dirichlet kernel on \( T_{2\pi} \), cf., the Dirichlet kernels on \( \mathbb{R} \) and \( \mathbb{Z} \) in Remark 1.6.4 and Remark 3.1.2, respectively. Using the notation of Definition 3.4.3, we see that \( K_{(\lambda)} \equiv D_N \) is not an approximate identity since \( D_N \notin L^1(T_{2\pi}) \).

b. The Fejér function \( W_N \) on \( T_{2\pi} \) is defined by

\[
(3.4.6) \quad \forall \gamma \in \mathbb{R}, \quad W_N(\gamma) = \sum_{|n| \leq N} \left( 1 - \frac{|n|}{N + 1} \right)e^{-\pi i n \gamma / \Omega}.
\]

It is not difficult to show that the trigonometric polynomial \( W_N \) can be written as

\[
(3.4.7) \quad W_N(\gamma) = \frac{D_0(\gamma) + \ldots + D_N(\gamma)}{N + 1}
\]

\[
\quad + \frac{1}{N + 1} \left( \frac{\sin(N + 1) \frac{\pi \gamma}{2\Omega}}{\sin \frac{\pi \gamma}{2\Omega}} \right)^2,
\]
where \( W_N(2k\Omega) \) is \( N + 1 \) for \( k \in \mathbb{Z} \), e.g., Exercise 3.11b. The family \( \{W_N : N \in \mathbb{N} \cup \{0\}\} \) is the Fejér kernel on \( T_{2\Omega} \), cf., the Fejér kernel on \( \mathbb{R} \) and \( \mathbb{Z} \) in Remark 1.6.4 and Remark 3.1.2, respectively.

c. The Fejér kernel is an approximate identity. To see this, first note that \( W_N \geq 0 \) by (3.4.7), and

\[
\int_{T_{2\Omega}} W_N(\gamma) \, d\gamma = 1
\]

by the definition (3.4.6) and the fact that \( \int_{T_{2\Omega}} e^{-\pi i n \gamma / \Omega} \, d\gamma = 0 \) for \( n \neq 0 \). Thus, parts \( a \) and \( b \) of Definition 3.4.3 are valid. Finally, we obtain part \( c \) of Definition 3.4.3 by (3.4.7) and the estimate

\[
\frac{1}{2\Omega} \int_{\eta \leq |\gamma| \leq \Omega} |W_N(\gamma)| \, d\gamma \leq \frac{2(\Omega - \eta)}{2\Omega(N + 1)} \sup_{\eta \leq |\gamma| \leq \Omega} \frac{1}{\sin^2 \frac{\pi \gamma}{2\Omega}} \leq \frac{\Omega - \eta}{\Omega(N + 1)} \frac{1}{\sin^2 \frac{\pi \gamma}{2\Omega}}
\]

In Section 3.2.8, we mentioned du Bois-Reymond's example (1872) of a function \( F \in C(T) \) for which \( \{S_N(F)(0)\} \) diverges, cf., Exercise 3.45. At the risk of being an alarmist over 100 years after the fact, it is not an exaggeration to say that this example dampened some of the optimism for a comprehensive theory of the representation of functions by trigonometric series. Fejér's result (1904), stated below in Theorem 3.4.6a, came none too soon, e.g., [Bir73, pages 150-156] for a translation of the relevant parts of Fejér's original paper.

3.4.6 Theorem. FEJÉR THEOREM

a. Let \( F \in C(T_{2\Omega}) \), and let \( f = \{f[n]\} \in A'(\mathbb{Z}) \subseteq \ell^2(\mathbb{Z}) \) be its sequence of Fourier coefficients. Then

\[
(3.4.8) \quad F \ast W_N(\gamma) = \sum_{|n| \leq N} \left(1 - \frac{|n|}{N + 1}\right)f[n]e^{-\pi i n \gamma / \Omega}
\]

and

\[
(3.4.9) \quad \lim_{N \to \infty} \|F - F \ast W_N\|_{L^\infty(T_{2\Omega})} = 0.
\]
b. Let \( F \in L^1(T_{2\pi}) \). Then

\[
\lim_{N \to \infty} \int_{T_{2\pi}} \left| F(\gamma) - \sum_{|n| \leq N} \left(1 - \frac{|n|}{N+1}\right)f[n]e^{-\pi i n \gamma / \Omega}\right| \, d\gamma = 0,
\]

cf., Theorem 1.6.9b.

Proof. a. (3.4.8) follows from a direct computation of the left side, cf., Exercise 3.12a. (3.4.9) follows by combining Example 3.4.5c with Theorem 3.4.4a. Part b follows from part a, e.g., Exercise 3.16. \( \square \)

3.4.7 Corollary. **Uniqueness**

Let \( F \in L^1(T_{2\Omega}) \) and assume \( f[n] = 0 \) for each \( n \in \mathbb{Z} \), where \( f = \{f[n]\} \) is the sequence of Fourier coefficients of \( F \). Then \( F \) is the 0-function, cf., Theorem 1.6.9c.

3.4.8 Remark. **Weierstrass Approximation Theorem**

a. The Weierstrass Approximation Theorem (1885) asserts that if \( F \in C[\alpha, \beta] \) then there is a sequence \( \{P_N\} \) of polynomials for which

\[
(3.4.10) \quad \lim_{N \to \infty} \|F - P_N\|_{L^\infty[\alpha, \beta]} = 0.
\]

(3.4.10) can be derived from (3.4.9) in the following way. By translation we can take \( F \in C[-\Omega, \Omega] \) without loss of generality. Next choose \( c \) such that \( G(-\Omega) = G(\Omega) \) where \( G(\gamma) \equiv F(\gamma) - c\gamma \) for \( \gamma \in [-\Omega, \Omega] \). In fact, let

\[
c = \frac{F(\Omega) - F(-\Omega)}{2\Omega}.
\]

Apply Theorem 3.4.6 to \( G \) considered as an element of \( C(T_{2\Omega}) \). Finally, uniformly approximate the trigonometric polynomials \( G*W_N \) on \([-\Omega, \Omega]\) by polynomial approximants of their Taylor series expansions.

b. There are extensive developments of the Weierstrass Theorem, many of which have evolved from Stone's celebrated Stone-Weierstrass Theorem (1937). We refer to [BD81] and [Bur84], replete with ingenuity and scholarship, for recent contributions to the Stone-Weierstrass Theorem in a functional analytic uniform algebra setting.
3.4. **THE \( L^1(\mathbb{T}) \) AND \( L^2(\mathbb{T}) \) THEORIES**

In Remark 3.3.8c we asserted that the partial sums \( S_N(F) \) do not necessarily converge to \( F \) in \( L^1(\mathbb{T}_{2\pi}) \). We shall now prove this assertion.

### 3.4.9 Example. Lebesgue Constants and an \( L^1 \)-Convergence Problem

a. Analogous to (3.4.8), a direct computation shows that

\[
\forall F \in L^1(\mathbb{T}_{2\pi}), \quad S_N(F)(\gamma) = F \ast D_N(\gamma),
\]

cf., Exercise 3.12.

b. The Lebesgue constants are defined as \( \|D_N\|_{L^1(\mathbb{T}_{2\pi})} \) for each \( N \). We shall prove that there is a bounded sequence \( \{C(N) : N \geq 2\} \) such that

\[
(3.4.11) \quad \forall N \geq 2, \quad \|D_N\|_{L^1(\mathbb{T}_{2\pi})} = \frac{4}{\pi^2} \log N + C(N),
\]

cf., Remark 3.3.8b.

To see this we use (3.4.5) as follows. First,

\[
\|D_N\|_{L^1(\mathbb{T}_{2\pi})} = \frac{2}{\pi} \int_0^{\pi/2} \left| \sin(2N + 1)x \right| \left| \frac{1}{\sin x} - \frac{1}{x} + \frac{1}{x} \right| dx,
\]

so that

\[
(3.4.12) \quad \left| \|D_N\|_{L^1(\mathbb{T}_{2\pi})} - \frac{2}{\pi} \int_0^{\pi/2} \frac{|\sin(2N + 1)x|}{x} dx \right|
\]

\[
\leq \frac{2}{\pi} \int_0^{\pi/2} \left| \frac{1}{\sin x} - \frac{1}{x} \right| dx = C_1,
\]

by the triangle inequality, where \( C_1 < \infty \). In fact, it is easy to check that \( \lim_{x \to 0} \left( \frac{1}{\sin x} - \frac{1}{x} \right) = 0 \).
Next, letting \( t = (2N + 1)x \) and dividing the domain of integration according to the sign of the sine, we have

\[
\frac{2}{\pi} \int_0^{\pi} \frac{\sin(2N + 1)x}{x} \, dx = \frac{2}{\pi} \sum_{k=0}^{N-1} \int_{k\pi}^{(k+1)\pi} \frac{\sin t}{t} \, dt \\
+ \frac{2}{\pi} \int_{N\pi}^{N\pi + \frac{\pi}{2}} \frac{\sin t}{t} \, dt \\
= \frac{2}{\pi} \int_0^\pi \sin u \left( \sum_{k=1}^{N-1} \frac{1}{u + k\pi} \right) \, du \\
+ \frac{2}{\pi} \int_0^\pi \sin u \, du + \frac{2}{\pi} \int_{N\pi}^{N\pi + \frac{\pi}{2}} \frac{\sin u}{u} \, du.
\]

(3.4.13)

Since both \( \sin u \geq 0 \) and

\[
\sum_{k=2}^{N} \frac{1}{k\pi} \leq \sum_{k=1}^{N-1} \frac{1}{u + k\pi} \leq \sum_{k=1}^{N-1} \frac{1}{k\pi}
\]

on \([0, \pi]\), we can use the integral test to compute

\[
\frac{4}{\pi^2} \left( -1 + \log(N + 1) \right) \leq \frac{2}{\pi} \int_0^\pi \sin u \left( \sum_{k=1}^{N-1} \frac{1}{u + k\pi} \right) \, du \\
\leq \frac{4}{\pi^2} \left( 1 + \log(N - 1) \right).
\]

(3.4.14)

We obtain (3.4.11) by combining (3.4.12), (3.4.13), and (3.4.14).

c. Consider the linear mappings

\[
L_N : L^1(\mathbb{T}_{2\Omega}) \rightarrow L^1(\mathbb{T}_{2\Omega}) \\
F \rightarrow S_N(F).
\]

The norm of \( L_N \), defined in Definition B.6 is

\[
\|L_N\| = \sup_{\|F\|_{L^1(\mathbb{T}_{2\Omega})} \leq 1} \|L_N(F)\|_{L^1(\mathbb{T}_{2\Omega})} = \sup_{\|F\|_{L^1(\mathbb{T}_{2\Omega})} \leq 1} \|F \ast D_N\|_{L^1(\mathbb{T}_{2\Omega})}.
\]
and so \( \|L_N\| \leq \|D_N\|_{L^1(\mathbb{T}_{2\pi})} \) by (3.4.1) and part a. To prove the opposite inequality we first note that

\[
\|L_N\| \geq \|L_N(W_n)\|_{L^1(\mathbb{T}_{2\pi})} = \|D_N * W_n\|_{L^1(\mathbb{T}_{2\pi})}.
\]

Then, by Theorem 3.4.4b and Example 3.4.5c, we have

\[
\lim_{n \to \infty} \|D_N * W_n\|_{L^1(\mathbb{T}_{2\pi})} = \|D_N\|_{L^1(\mathbb{T}_{2\pi})}.
\]

Consequently,

(3.4.15) \( \forall N \geq 1, \ \|L_N\| = \|D_N\|_{L^1(\mathbb{T}_{2\pi})} \).

Combining (3.4.15) and (3.4.11) with the Uniform Boundedness Principle (Theorem B.8), we can assert that there are functions \( F \in L^1(\mathbb{T}_{2\pi}) \) for which \( \sup_N \|S_N(F)\|_{L^1(\mathbb{T}_{2\pi})} = \infty \). In particular, for such functions we do not have \( \lim_{N \to \infty} \|S_N(F) - F\|_{L^1(\mathbb{T}_{2\pi})} = 0 \).

We begin our discussion of the \( L^2(\mathbb{T}) \) theory with the following definition (Definition 3.4.10) and background (Remark 3.4.11).

3.4.10 Definition. Orthonormal Basis

a. A sequence \( \{E_n\} \subseteq L^2(\mathbb{T}_{2\pi}) \) is orthonormal in \( L^2(\mathbb{T}_{2\pi}) \) if

\[
\forall m, n \in \mathbb{Z}, \quad \int_{\mathbb{T}_{2\pi}} E_m(\gamma) \overline{E_n(\gamma)} \, d\gamma = \delta(m, n),
\]

where

\[
\delta(m, n) = \begin{cases} 1, & \text{if } m = n, \\ 0, & \text{if } m \neq n. \end{cases}
\]

An orthonormal sequence \( \{E_n\} \subseteq L^2(\mathbb{T}_{2\pi}) \) is an orthonormal basis (ONB) for \( L^2(\mathbb{T}_{2\pi}) \) if

(3.4.16) \( \forall F \in L^2(\mathbb{T}_{2\pi}), \ \exists \{c_n\} \subseteq \mathbb{C} \) such that

\[
F = \sum c_n E_n \quad \text{in} \quad L^2(\mathbb{T}_{2\pi}),
\]

cf., (3.4.18).
b. Using Hölder’s Inequality, it is an elementary calculation to show that

(3.4.17) \[ \lim_{n \to \infty} \int_{T_{2\pi}} F_n(\gamma) \overline{G_n(\gamma)} \, d\gamma = \int_{T_{2\pi}} F(\gamma) \overline{G(\gamma)} \, d\gamma \]

when \( F, G, F_n, G_n \in L^2(T_{2\pi}) \) and

\[ \lim_{n \to \infty} \| F - F_n \|_{L^2(T_{2\pi})} = 0 \quad \text{and} \quad \lim_{n \to \infty} \| G - G_n \|_{L^2(T_{2\pi})} = 0, \]

e.g., Exercise 3.9.

c. In the case of an ONB \( \{E_n\} \), the coefficients \( c_n \) in (3.4.16) are of the form

(3.4.18) \[ \forall n \in \mathbb{Z}, \quad c_n = \int_{T_{2\pi}} F(\gamma) \overline{E_n(\gamma)} \, d\gamma. \]

This follows from part \( b \).

3.4.11 Remark. INTEGRAL EQUATIONS AND THE RIESZ-FISCHER THEOREM

a. In Example 3.1.10f we mentioned Abel in conjunction with elliptic functions. In fact, this work was in the realm of integral equations, and Abel (1823) solved the “tautochrone” equation

\[ \int_0^y \frac{f(x)}{\sqrt{y - x}} \, dx = g(y), \]

for a given forcing function \( g \), by computing

\[ f(x) = \frac{1}{\pi} \int_0^x \frac{g'(y)}{\sqrt{x - y}} \, dy. \]

In the late 19th century it was realized that many problems in mathematical physics could be transformed into solving integral equations of the form

(3.4.19) \[ \int_T F(\gamma) K(\gamma, \lambda) \, d\gamma = G(\lambda) \]

e.g., [CH53], [Die81]. The Dirichlet problem in potential theory was solved in particular cases by Neumann. Vito Volterra (1896) used the
3.4. THE $L^1(\mathbb{T})$ AND $L^2(\mathbb{T})$ THEORIES

Neumann method to solve a certain type of integral equation, and this led to Ivar Fredholm's (still) eminently readable and fundamental paper on integral equations in Acta Mathematica (27(1903), 365-390).

b. Hilbert and (Erhard) Schmidt, in the periods 1904-1912 and 1905-1908, respectively, made great strides in solving (3.4.19), and, in the process, established some of the fundamental ideas of functional analysis. They also set the stage for the Riesz-Fischer Theorem in the following way.

Suppose $\{E_n\} \subseteq L^2(\mathbb{T})$ is orthonormal. Assume we can write $K, F,$ and $G$ of (3.4.19) as $K = \sum_{m,n} k(m,n)E_mE_n$, $F = \sum f(n)E_n$, $G = \sum g(n)E_n$. For a given kernel $K$ and forcing function $G$, the goal is to find $F$. Formally,

$$\int_T F(\gamma)K(\gamma, \lambda) d\gamma = \sum_{m} \sum_{p,q} f(m)k(p,q)E_q(\lambda) \int_T E_m(\gamma)E_P(\gamma) d\gamma$$

$$= \sum_q \left( \sum_m f(m)k(m,q) \right) E_q(\lambda),$$

and so (3.4.19) leads to the infinite system of linear equations

(3.4.20) \quad \forall q \in \mathbb{Z}, \quad \sum_m f(m)k(m,q) = g(q).

Suppose $g \in \ell^2(\mathbb{Z})$, and $k: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{C}$ has the property that

$$\sum_{m,n}|k(m,n) - \delta(m,n)|^2 < \infty.$$  

Then classical methods yield the construction of a unique solution $f \in \ell^2(\mathbb{Z})$ of (3.4.20), e.g., [GG81, pages 70-74].

c. Once the sequence $f \in \ell^2(\mathbb{Z})$ of part b is found, then the major problem in solving (3.4.19) in terms of $F = \sum f(n)E_n$ is accomplished by means of F. Riesz' Theorem: if $\{E_n\} \subseteq L^2(\mathbb{T})$ is orthonormal and $f \in \ell^2(\mathbb{Z})$, then there is $F \in L^2(\mathbb{T})$ for which

$$\forall n \in \mathbb{Z}, \quad f(n) = \int_T F(\gamma)E_n(\gamma) d\gamma.$$
This is Riesz' formulation of the *Riesz-Fischer Theorem* (1906-1907). Fischer's formulation is that $L^2(\mathbb{T})$ is complete, i.e., if $\{F_n\} \subseteq L^2(\mathbb{T})$ is a Cauchy sequence in the $L^2$-norm then there is $F \in L^2(\mathbb{T})$ for which $\lim \|F - F_n\|_{L^2(\mathbb{T})} = 0$. Zygmund refers to the Riesz-Fischer Theorem as "a great achievement of the Lebesgue theory".

Fischer's formulation is a special case of *Theorem A.18*, which itself is a staple in a basic real variables course, e.g., [Ben76, pages 232–233], [HS65, pages 192–194], [Rud66, pages 66–67]. We shall use *Theorem A.18* in the $L^2(\mathbb{T})$ theory which follows.

### 3.4.12 Theorem. ONB, Parseval Formula, and Convergence

a. **ONB.** $\{e^{-i\pi n/\Omega}\}$ is an ONB for $L^2(\mathbb{T}_{2\Omega})$.

b. **Parseval Formula.** Let $F, G \in L^2(\mathbb{T}_{2\Omega})$, and consider the pairings $f \mapsto F$, $g \mapsto G$. Then

$$
\int_{\mathbb{T}_{2\Omega}} F(\gamma) \overline{G(\gamma)} \, d\gamma = \sum f[n]g[n],
$$

and, in particular,

$$
\|F\|_{L^2(\mathbb{T}_{2\Omega})} = \left( \int_{\mathbb{T}_{2\Omega}} |F(\gamma)|^2 \, d\gamma \right)^{1/2} = \left( \sum |f[n]|^2 \right)^{1/2} = \|f\|_{L^2(\Omega)}.
$$

c. **Convergence.** For all $F \in L^2(\mathbb{T}_{2\Omega})$,

$$(3.4.21) \quad \lim_{N \to \infty} \|F - S_N(F)\|_{L^2(\mathbb{T}_{2\Omega})} = 0.
$$

**Proof.** i. $\{e^{-i\pi n/\Omega}\}$ is orthonormal in $L^2(\mathbb{T}_{2\Omega})$ by direct calculation, and $\text{span}\{e^{-i\pi n/\Omega}\} = L^2(\mathbb{T}_{2\Omega})$ by the Fejér Theorem (*Theorem 3.4.6a*) and the method of *Exercise 3.16*.

ii. For any $F \in L^2(\mathbb{T}_{2\Omega})$ and any $N$

$$
0 \leq \|F - S_N(F)\|_{L^2(\mathbb{T}_{2\Omega})}^2 = \|F\|_{L^2(\mathbb{T}_{2\Omega})}^2 - \sum_{|n| \leq N} |f[n]|^2,
$$
and so

(3.4.22) \[ \sum_{|n| \leq N} |f[n]|^2 \leq \|F\|_{L^2(T_{2\Omega})}^2. \]

This is Bessel's Inequality (1828). Further, if \( N > M \) then

\[ \|S_N(F) - S_M(F)\|_{L^2(T_{2\Omega})}^2 = \sum_{M < |n| \leq N} |f[n]|^2, \]

and so, by (3.4.22), \( \{S_N(F)\} \) is a Cauchy sequence in \( L^2(T_{2\Omega}) \). Thus, \( \sum f[n] e^{-\pi in\gamma/\Omega} \) converges to some \( K \in L^2(T_{2\Omega}) \) since \( L^2(T_{2\Omega}) \) is complete.

iii. By i, if \( G \in L^2(T_{2\Omega}) \) and

\[ \forall n \in \mathbb{Z}, \quad \int_{T_{2\Omega}} G(\gamma)e^{\pi in\gamma/\Omega} d\gamma = 0, \]

then (3.4.17) allows us to conclude that \( \|G\|_{L^2(T_{2\Omega})} = 0 \), and so \( G \) is the 0-function.

iv. Now, for any \( F \in L^2(T_{2\Omega}) \) and corresponding \( K \) as in ii, we have (by (3.4.17) again)

\[
\int_{T_{2\Omega}} (F(\gamma) - K(\gamma))e^{\pi in\gamma/\Omega} d\gamma \\
= f[n] - \lim_{N \to \infty} \sum_{|m| \leq N} f[m] \int_{T_{2\Omega}} e^{-\pi i(m-n)\gamma/\Omega} d\gamma = 0.
\]

Therefore, by iii, \( F = K \) a.e.; and so (3.4.21) is obtained, giving part c as well as part a.

v. Part b follows from (3.4.17), (3.4.21), and the calculation

\[
\int_{T_{2\Omega}} F(\gamma)\overline{G(\gamma)} d\gamma = \lim_{N \to \infty} \int_{T_{2\Omega}} S_N(F)(\gamma)\overline{S_N(G)(\gamma)} d\gamma \\
= \lim_{N \to \infty} \sum_{|m|, |h| \leq N} f[m]g[n] \int_{T_{2\Omega}} e^{-\pi i(m-n)\gamma/\Omega} d\gamma \\
= \sum f[n]\overline{g[n]}.
\]

\( \square \)
In light of Bessel's Inequality (3.4.22), we know that if $F \in L^2(\mathbb{T})$ for the pairing $f \leftrightarrow F$, then $f \in \ell^2(\mathbb{Z})$. Riesz's formulation of the Riesz-Fischer Theorem completes the picture as follows:

**3.4.13 Theorem.** $\ell^2(\mathbb{Z}) \longrightarrow L^2(\mathbb{T})$

There is a unique linear bijection $\mathcal{F} : \ell^2(\mathbb{Z}) \rightarrow L^2(\mathbb{T})$ with the properties

i. $\forall f \in \ell^2(\mathbb{Z}), \quad \|f\|_{\ell^2(\mathbb{Z})} = \|\mathcal{F}f\|_{L^2(\mathbb{T})}$,

ii. $\forall F \in L^2(\mathbb{T})$, $f \equiv \mathcal{F}^{-1}F$ is the sequence of Fourier coefficients of $F$.

**Proof.** In light of Corollary 3.4.7 and Theorem 3.4.12, it is sufficient to prove that for any sequence $\{c_n\} \in \ell^2(\mathbb{Z})$ there is $F \in L^2(\mathbb{T})$, uniquely determined a.e., such that $\{c_n\}$ is the sequence of Fourier coefficients of $F$.

If we define $S_N(\gamma) = \sum_{|n| \leq N} c_n e^{-i\omega n\gamma}$, then $\{S_N\} \subseteq L^2(\mathbb{T})$ is a Cauchy sequence since

$$\|S_N - S_M\|_{L^2(\mathbb{T})} = \sum_{M < |n| \leq N} |c_n|^2$$

when $N > M$. By the completeness of $L^2(\mathbb{T})$ there is a unique $F \in L^2(\mathbb{T})$ for which

$$\lim_{N \to \infty} \|F - S_N\|_{L^2(\mathbb{T})} = 0.$$ 

Further, for each $n$ and for $N \geq |n|$, we have

$$|f[n] - c_n| = \left| \int_{\mathbb{T}} (F(\gamma) - S_N(\gamma)) e^{i\omega n \gamma\gamma} d\gamma \right|$$

$$\leq \|F - S_N\|_{L^2(\mathbb{T})}$$

by Hölder's Inequality. Thus, $c_n = f[n]$.

\[ \square \]

**3.5 $A(\mathbb{T})$ and the Wiener Inversion Theorem**

Besides the formidable task of finding an effective intrinsic characterization of $A(\mathbb{T})$, e.g., *Examples 1.4.4 and 3.3.4*, the space $A(\mathbb{T})$
3.5. *A*(T) AND THE WIENER INVERSION THEOREM

is worthy of careful attention because of its algebraic properties and the ramifications of those properties, cf., the introductory paragraph of *Section 3.4*. Fortunately, there is an accessible masterpiece on the subject by Kahane [Kah70]. Our modest goal in this section will be to state these algebraic properties, and to prove Wiener’s theorem on the inversion of absolutely convergent Fourier series. We shall refer to this result as Wiener’s Inversion Theorem.

3.5.1 **Definition. Convolution**

a. Let \( f, g \in \ell^1(\mathbb{Z}) \). The convolution of \( f \) and \( g \), denoted by \( f \ast g \), is

\[
f \ast g[n] = \sum_{k=-\infty}^{\infty} f[n-k]g[k] = \sum_{k=-\infty}^{\infty} f[k]g[n-k].
\]

More simply than the cases of \( L^1(\mathbb{R}) \) and \( L^1(\mathbb{T}_{2\pi}) \), we see that \( f \ast g \in \ell^1(\mathbb{Z}) \) since

\[
\sum_{n=-\infty}^{\infty} |f \ast g[n]| \leq \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |f[n-k]g[k]| = \sum |f[n]| \sum |g[k]|.
\]

Rewriting this expression, we have

\[
(3.5.1) \quad \forall f, g \in \ell^1(\mathbb{Z}), \quad \|f \ast g\|_{\ell^1(\mathbb{Z})} \leq \|f\|_{\ell^1(\mathbb{Z})} \|g\|_{\ell^1(\mathbb{Z})}.
\]

b. \( \ell^1(\mathbb{Z}) \) is a **commutative algebra** taken with the operations of addition and convolution, i.e., \( \ell^1(\mathbb{Z}) \) is a complex vector space under addition, and convolution is distributive with respect to addition, as well as being associative, and commutative.

Further, \( \ell^1(\mathbb{Z}) \) has a **unit** \( u \) under convolution. \( u \) is defined by \( u[n] = \delta(0,n) \), so that

\[
\forall f \in \ell^1(\mathbb{Z}), \quad f \ast u = u \ast f = f
\]

since

\[
\forall n \in \mathbb{Z}, \quad \sum_{k=-\infty}^{\infty} f[n-k] \delta(0,k) = f[n].
\]

A straightforward calculation yields the following result.
3.5.2 Proposition.

Let \( \Omega > 0 \) and consider \( A(\mathbb{T}_{2\Omega}) \) (Definition 3.1.8a).

a. Let \( F, G \in A(\mathbb{T}_{2\Omega}) \) and let \( f = \{f[n]\} \), \( g = \{g[n]\} \in \ell^1(\mathbb{Z}) \) be the sequences of Fourier coefficients of \( F \) and \( G \), i.e., \( \hat{f} = F \) and \( \hat{g} = G \). Then \( f \ast g \in \ell^1(\mathbb{Z}) \) is the sequence of Fourier coefficients of \( FG \in A(\mathbb{T}_{2\Omega}) \), i.e.,

\[
\forall n \in \mathbb{Z}, \quad (f \ast g)[n] = \int_{\mathbb{T}_{2\Omega}} F(\gamma)G(\gamma) e^{2\pi i n \gamma / \Omega} \, d\gamma.
\]

(Recall that "\( f_{\alpha} \)" designates "\( \frac{1}{2\pi} \int_{-\alpha}^{\alpha} f \)" for any fixed \( \alpha \in \mathbb{R} \).)

b. \( A(\mathbb{T}_{2\Omega}) \) is a commutative algebra under the operations of addition and (ordinary pointwise) multiplication of functions. The function \( U \equiv 1 \in A(\mathbb{T}_{2\Omega}) \) is the multiplicative unit of \( A(\mathbb{T}_{2\Omega}) \), and

\[(3.5.1)' \quad \forall F, G \in A(\mathbb{T}_{2\Omega}), \quad \|FG\|_{A(\mathbb{T}_{2\Omega})} \leq \|F\|_{A(\mathbb{T}_{2\Omega})}\|G\|_{A(\mathbb{T}_{2\Omega})}.\]

3.5.3 Example. The \( A(\mathbb{T}) \) Norm

a. Let \( F \in A(\mathbb{T}_{2\Omega}) \) and let \( f = \{f[n]\} \in \ell^1(\mathbb{Z}) \) be the sequence of Fourier coefficients of \( F \). Then

\[(3.5.2) \quad \|F\|_{A(\mathbb{T}_{2\Omega})} = |f[0]| + \sum' \frac{1}{|n|}|nf[n]|.\]

Using (3.5.2) and Exercise 3.26 we can conclude that if \( F \in A(\mathbb{T}_{2\Omega}) \) and \( F' \in L^2(\mathbb{T}_{2\Omega}) \) then

\[(3.5.3) \quad \|F\|_{A(\mathbb{T}_{2\Omega})} \leq |f[0]| + \frac{\Omega}{\sqrt{3}}\|F'\|_{L^2(\mathbb{T}_{2\Omega})}.\]

b. Let \( F_\varepsilon \) be the \( 2\Omega \)-periodic triangle function, \( \max(1 - |\gamma|/\varepsilon, 0) \), defined in Remark 3.1.2b. As we stated there, its sequence \( \{w(\varepsilon)\} \) of Fourier coefficients is an element of \( \ell^1(\mathbb{Z}) \), and so \( F_\varepsilon \in A(\mathbb{T}_{2\Omega}) \). Now define the \( 2\Omega \)-periodic trapezoid function \( V_\varepsilon = 2F_\varepsilon - F_\varepsilon \). Note that

\[(3.5.4) \quad \|V_\varepsilon\|_{A(\mathbb{T}_{2\Omega})} \leq 3.\]

In fact, since \( w(\varepsilon) \geq 0 \) in \( \mathbb{Z} \), we have

\[
2 \sum w(2\varepsilon)[n] + \sum w(\varepsilon)[n] = 2F_\varepsilon(0) + F_\varepsilon(0) = 3.
\]
3.5. $A(T)$ AND THE WIENER INVERSION THEOREM

We shall show that if $F \in A(T_{2\Omega})$ and $F(0) = 0$ (and so $F(2\Omega n) = 0$ for all $n \in \mathbb{Z}$) then

$$\lim_{\epsilon \to 0} \|FV_\epsilon\|_{A(T_{2\Omega})} = 0. \tag{3.5.5}$$

First, $FV_\epsilon \in A(T_{2\Omega})$ by Proposition 3.5.2. If $F \in C^1(T_{2\Omega})$ then $F \in A(T_{2\Omega})$ by (3.5.3). Also, the pointwise a.e. derivative $(FV_\epsilon)'$ is not only supported by $[-2\epsilon, 2\epsilon]$ on $[-\Omega, \Omega)$, but is uniformly bounded independent of $\epsilon$. In fact,

$$\|(FV_\epsilon)'\|_{L^\infty(T_{2\Omega})} \leq \|F'\|_{L^\infty(T_{2\Omega})} + \sup_{\gamma \in [-2\epsilon, -\epsilon] \cup [\epsilon, 2\epsilon]} \frac{1}{\epsilon} |F(\gamma)|,$$

and the second term on the right side is bounded independent of $\epsilon$ since $F'(\gamma) = F(\gamma) - F(0)$ and the mean value theorem applies. Further, it is clear that

$$\lim_{\epsilon \to 0} \int_{-2\epsilon}^{2\epsilon} F(\gamma)V_\epsilon(\gamma) \, d\gamma = 0.$$

Thus, remembering the support of $(FV_\epsilon)'$, we can apply (3.5.3) again to obtain (3.5.5) for $F \in C^1(T_{2\Omega})$.

For arbitrary $F \in A(T_{2\Omega})$, we define the trigonometric polynomials

$$F_N(\gamma) = \sum_{1 \leq |n| \leq N} F^\nu[n]\epsilon^{-\pi \nu \gamma/\Omega} + a_N[0],$$

where $a_N[0] \equiv -\sum_{1 \leq |n| \leq N} F^\nu[n]$. Thus, $F_N \in C^1(T_{2\Omega})$ and $F_N(0) = 0$ so that the result of the previous paragraph can be used to obtain

$$\lim_{\epsilon \to 0} \|F_NV_\epsilon\|_{A(T_{2\Omega})} = 0.$$

Thus, since

$$\|FV_\epsilon\|_{A(T_{2\Omega})} \leq \|(F - F_N)V_\epsilon\|_{A(T_{2\Omega})} + \|F_NV_\epsilon\|_{A(T_{2\Omega})},$$

we can further apply Proposition 3.5.2 and (3.5.4) to compute

$$\lim_{\epsilon \to 0} \|FV_\epsilon\|_{A(T_{2\Omega})} \leq 3\|F - F_N\|_{A(T_{2\Omega})}.$$
CHAPTER 3. FOURIER SERIES

The left side is independent of $N$ and the right side is

$$3 \left\| F^\gamma[0] - \sum_{|n|>N} F^\gamma[n] e^{-\pi i n y/\Omega} + \sum_{1 \leq |n| \leq N} F^\gamma[n] \right\|_{A(T_{2\Omega})} \leq 3 \left( \sum_{|n| \leq N} |F^\gamma[n]| \right) + 3 \sum_{|n|>N} |F^\gamma[n]|.$$

As $N \to \infty$, the first term tends to 0 since $F(0) = 0$, and the second term tends to 0 since $F \in A(T_{2\Omega})$.

The proof of (3.5.5) is complete.

There are far reaching generalizations of (3.5.5) related to spectral synthesis and the ideal structure of $L^1(T)$, e.g., [Ben75, Section 1.2].

3.5.4 Proposition.

$A(T_{2\Omega}) = L^2(T_{2\Omega}) \ast L^2(T_{2\Omega})$.

Proof. The inclusion $L^2(T_{2\Omega}) \ast L^2(T_{2\Omega}) \subseteq A(T_{2\Omega})$ is a consequence of Parseval's Formula and Hölder's Inequality. In fact,

$$\sum |f[n]g[n]| \leq \left( \sum |f[n]|^2 \right)^{1/2} \left( \sum |g[n]|^2 \right)^{1/2} < \infty,$$

where $F, G \in L^2(T_{2\Omega})$ and $\hat{f} = F, \hat{g} = G$; and so $f g \in \ell^1(\mathbb{Z})$.

For the inclusion, $A(T_{2\Omega}) \subseteq L^2(T_{2\Omega}) \ast L^2(T_{2\Omega})$, let $F \in A(T_{2\Omega})$, where $\hat{f} = F$. For each $n$, we can write $f[n] = (g[n])^2$ for some $g[n] \in \mathbb{C}$; and we define the sequence $g = \{g[n]\}$. $g \in \ell^2(\mathbb{Z})$ since $f \in \ell^1(\mathbb{Z})$, and, hence, $F = \hat{g} \ast \hat{g} \in L^2(T_{2\Omega}) \ast L^2(T_{2\Omega})$.

3.5.5 Remark. Factorization

a. The factorization, $A(T) = L^2(T) \ast L^2(T)$, or, equivalently, $\ell^1(\mathbb{Z}) = \ell^2(\mathbb{Z}) \ell^2(\mathbb{Z})$, of Proposition 3.5.4 is elementary. A consequence of Proposition 3.5.2 is the inclusion $A(T)A(T) \subseteq A(T)$. In this context, we observe that $A(T) = A(T)A(T)$, or, equivalently, $\ell^1(\mathbb{Z}) = \ell^1(\mathbb{Z}) \ast \ell^1(\mathbb{Z})$, is also valid since $U \equiv 1 \in A(T)$.

Further, $A(\mathbb{Z}) = A(\mathbb{Z})A(\mathbb{Z})$ and $A(\mathbb{R}) = A(\mathbb{R})A(\mathbb{R})$. However, these two results are far from trivial and are due to Salem and Rudin, respectively. Paul Cohen (1959) proved that $A(T) = A(T)A(T)$ for any
3.5. $A(\mathbb{T})$ AND THE WIENER INVERSION THEOREM

locally compact abelian group $\Gamma$, cf., [Koo64], [Ptá72] for elegant proofs of the Cohen Factorization Theorem. Using Salem's Theorem, or an argument with convex functions, we also have $A(\mathbb{T}) = L^1_+(\mathbb{T}) \ast A(\mathbb{T})$, where $L^1_+(\mathbb{T}) = \{ F \in L^1(\mathbb{T}) : F \geq 0 \}$.

b. Although the proof that $A(\mathbb{T}) = A(\mathbb{T})A(\mathbb{T})$ is elementary, we have the following more intricate relationship: if $F \in A(\mathbb{T})$ never vanishes then

$$\forall H \in A(\mathbb{T}), \exists G \in A(\mathbb{T}) \text{ such that } H = FG.$$ 

In particular, if $F \in A(\mathbb{T})$ never vanishes then $1/F \in A(\mathbb{T})$. This last fact is Wiener's Inversion Theorem. There are Banach algebra proofs [Ben75, pages 22–23], a "spectral radius" proof [Ben75, pages 23–24], extensive generalizations which are documented and compared in [Ben75], and classical proofs going back to Wiener's original techniques [Wie81, Volume II, pages 519–623, esp., page 532], [Wie33, page 91]. We shall proceed in this last direction in Sections 3.5.6–3.5.9.

c. Before proving Wiener's Inversion Theorem, let us point out that a modified version of it is a natural component in the proof of Wiener's Tauberian Theorem, Theorem 2.9.12 and Remark 2.9.13, cf., the historical remark in [Ben75, pages 142–143] and the proofs in [Ben75, Sections 1.1–1.4]. Also, as indicated earlier, these results are fundamental in spectral synthesis, e.g., [Ben75, Section 2.5]. We shall apply Wiener's Inversion Theorem in Section 3.6.

3.5.6 Definition. Local Membership

Let $I \subseteq A(T_{2\Omega})$ be an ideal in the algebra $A(T_{2\Omega})$, i.e., $I$ is a subalgebra of $A(T_{2\Omega})$ and $FG \in I$ whenever $F \in A(T_{2\Omega})$ and $G \in I$.

A function $F : T_{2\Omega} \to \mathbb{C}$ belongs to $I$ locally at $\gamma \in T_{2\Omega}$ if

$$\exists G_\gamma \in I \text{ and } \exists N_\gamma = (\alpha, \beta) \subseteq T_{2\Omega} \text{ such that } \gamma \in (\alpha, \beta) \text{ and } \forall \lambda \in N_\gamma, \ G_\gamma(\lambda) = F(\lambda).$$

In this case we write $F \in I_{loc}(\gamma)$.

3.5.7 Theorem. Local Membership Theorem
Let $I \subseteq A(T_{2\Omega})$ be an ideal and let $F : T_{2\Omega} \to \mathbb{C}$ be a function. If $F \in I_{\operatorname{loc}}(\gamma)$ for each $\gamma \in T_{2\Omega}$ then $F \in I$.

**Proof.** For each $\gamma \in T_{2\Omega}$ choose $G_\gamma$ and $N_\gamma$ (as in Definition 3.5.6) for which $G_\gamma = F$ on $N_\gamma$. Clearly we can take $N_\gamma$ centered at $\gamma$; and for each $\gamma$ we choose a closed interval $C_\gamma \subseteq N_\gamma$, also centered at $\gamma$ and whose length $|C_\gamma|$ equals $\frac{1}{2}|N_\gamma|$. Since $T_{2\Omega}$ is a compact set we can find $\gamma_1, \ldots, \gamma_n$ so that

\begin{equation}
T_{2\Omega} = \bigcup_{j=1}^{n} C_j,
\end{equation}

\[\text{e.g., Definition B.1.}\]

Next, choose $V_j \in A(T_{2\Omega})$, $j = 1, \ldots, n$, where $V_j = 1$ on $C_{\gamma_j}$ and $V_j = 0$ off $N_{\gamma_j}$, e.g., Example 3.5.3, cf., Exercise 1.5\theta and the general construction in [Ben75, Proposition 1.1.4]. Since $I$ is an ideal, we have $V_j G_j \in I$ for each $j = 1, \ldots, n$. It is also clear that

\begin{equation}
\forall \gamma \in T_{2\Omega} \text{ and } \forall j = 1, \ldots, n, \quad V_j(\gamma) G(\gamma) = V_j(\gamma) F(\gamma).
\end{equation}

Defining

\begin{equation}
F_0 = F(1 - (1 - V_1)(1 - V_2) \cdots (1 - V_n)),
\end{equation}

we see that $F_0 \in I$ because of (3.5.7) and the fact that the "1s" cancel when we compute the right side of (3.5.8). Finally, $F_0 = F$ on $T_{2\Omega}$. In fact, if $\gamma \in T_{2\Omega}$, then there is $k$ for which $1 - V_k(\gamma) = 0$ because of (3.5.6).

3.5.8 Example. 2\text{\[m\]} > \|F\|_{A(T_{2\Omega})} \text{ IMPLIES } 1/F \in A(T_{2\Omega})

Let $F \in A(T_{2\Omega}) \setminus \{0\}$ and let $f = \{f[n]\}$ be the sequence of Fourier coefficients of $F$. Assume

\begin{equation}
2|f[0]| > \|F\|_{A(T_{2\Omega})}.
\end{equation}

We shall show that $1/F \in A(T_{2\Omega})$. To this end we combine the series expansion, $1/(1 + G) = 1 - G + G^2 - G^3 + \cdots$, and (3.5.9) to compute for $G(\gamma) \equiv \sum_j f[j] e^{-\pi in\gamma/\Omega}$ that

\[
\frac{1}{F(\gamma)} = \frac{1}{f[0]} (1 - G(\gamma) + G(\gamma)^2 - G(\gamma)^3 + \cdots).
\]
3.5. $A(T)$ AND THE WIENER INVERSION THEOREM

Thus, by (3.5.9) and Proposition 3.5.2b,

$$
\|\frac{1}{F}\|_{A(T_{2\Omega})} \leq \frac{1}{|f[0]|} \sum_{n=0}^{\infty} \|G\|_{A(T_{2\Omega})}^{n} = \frac{1}{|f[0]|} \frac{1}{1 - \|G\|_{A(T_{2\Omega})}} = \frac{1}{2|f[0]| - \|F\|_{A(T_{2\Omega})}} < \infty.
$$

3.5.9 Theorem. Wiener Inversion Theorem

Let $F \in A(T_{2\Omega})$.

a. If $F(\gamma_{0}) \neq 0$ then there is $G \in A(T_{2\Omega})$ such that $F = G$ on some open interval $N$ centered at $\gamma_{0}$ and $1/G \in A(T_{2\Omega})$.

b. If $F$ never vanishes then $1/F \in A(T_{2\Omega})$.

Proof. a. Without loss of generality, let $\gamma_{0} = 0$, and define

(3.5.10) $\forall \gamma \in T_{2\Omega}, \ G_{\varepsilon}(\gamma) = F(0) + V_{\varepsilon}(\gamma)(F(\gamma)) - F(0),$

where $V_{\varepsilon}$ was defined in Example 3.5.3. Choose $\eta \equiv |F(0)|/3 > 0$, and apply (3.5.5) to find $\varepsilon > 0$ for which $\|V_{\varepsilon}(F - F(0))\|_{A(T_{2\Omega})} < \eta$. For this $\varepsilon$, set $G = G_{\varepsilon}$.

Since

$$
\left| \int_{T_{2\Omega}} V_{\varepsilon}(\gamma)(F(\gamma) - F(0)) \, d\gamma \right| \leq \|V_{\varepsilon}(F - F(0))\|_{A(T_{2\Omega})},
$$

we have

$$
\left| \int_{T_{2\Omega}} G(\gamma) \, d\gamma \right| = \left| F(0) + \int_{T_{2\Omega}} V_{\varepsilon}(\gamma)(F(\gamma) - F(0)) \, d\gamma \right|
$$

(3.5.11)

$$
\geq |F(0)| - \|V_{\varepsilon}(F - F(0))\|_{A(T_{2\Omega})} > |F(0)| - \eta.
$$

On the other hand, it is immediate from (3.5.10) that

(3.5.12) $\|G\|_{A(T_{2\Omega})} < |F(0)| + \eta.$

Combining (3.5.11) and (3.5.12) with the definition of $\eta$, we obtain

(3.5.13) $2 \left| \int_{T_{2\Omega}} G(\gamma) \, d\gamma \right| > \frac{4}{3} |F(0)| > \|G\|_{A(T_{2\Omega})}.$
From (3.5.10) we see that \( G = F \) on \( N_0 = (-\epsilon, \epsilon) \), and, because of (3.5.13) and Example 3.5.8, \( 1/G \in A(T_{2\Omega}) \). This completes the proof of part a.

b. For each \( \gamma \in T_{2\Omega} \), we use part a to choose \( G_\gamma \in A(T_{2\Omega}) \) such that \( G_\gamma = F \) on some open interval \( N_\gamma \) centered at \( \gamma \) and \( 1/G_\gamma \in A(T_{2\Omega}) \).

Thus, \( 1/F \in A(T_{2\Omega})_{\text{loc}}(\gamma) \) for each \( \gamma \in T_{2\Omega} \), and so \( 1/F \in A(T_{2\Omega}) \) by Theorem 3.5.7.

\[ \square \]

3.5.10 Remark. \( A(T) \) and \( A(\hat{\mathbb{R}}) \)

a. Wiener’s concept of local membership leads one to investigate the relationship between \( A(T) \) and \( A(\hat{\mathbb{R}}) \). For example, it is natural to ask the question: supposing \( F \in A(T) \), \( F(\pm \frac{1}{2}) = 0 \), and \( G \) is defined on \( \hat{\mathbb{R}} \) as

\[ G(\gamma) = \begin{cases} 
F(\gamma), & \text{if } |\gamma| \leq \frac{1}{2}, \\
0, & \text{otherwise,}
\end{cases} \tag{3.5.14} \]

is \( G \in A(\hat{\mathbb{R}}) \)? Wiener proved the following theorem. Let \( F : T \to \mathbb{C} \) be a function vanishing on \([\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon]\) and define \( G : \hat{\mathbb{R}} \to \mathbb{C} \) by (3.5.14); then \( F \in A(T) \) if and only if \( G \in A(\hat{\mathbb{R}}) \), and

\[ \exists C_1(\epsilon), C_2(\epsilon) > 0 \text{ such that } C_1(\epsilon)\|G\|_{A(\hat{\mathbb{R}})} \leq \|F\|_{A(T)} \leq C_2(\epsilon)\|G\|_{A(\hat{\mathbb{R}})}. \]

The proof is not difficult, e.g., [Wie33, pages 80–82].

b. The following extension of Wiener’s result from part a was proved by Wik [Wik65]. Let \( F \in L^\infty(T) \) vanish on \([\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon]\), define \( G : \hat{\mathbb{R}} \to \mathbb{C} \) by (3.5.14), and let \( w : \mathbb{R} \to \mathbb{R} \) be an even positive function, which is increasing on \((0, \infty)\) and which satisfies the condition

\[ \exists C > 0 \text{ such that } \forall t \in \mathbb{R} \setminus \{0\}, w(2t) \leq Cw(t); \tag{3.5.15} \]

then

\[ \sum |F^\gamma[n]|w(n) < \infty \]

if and only if

\[ \int G^\gamma(t)w(t) \, dt < \infty. \]
3.5. \textit{A}(\mathbb{T}) \text{ AND THE WIENER INVERSION THEOREM}  

In this statement, we use "w" to denote a so-called \textit{weight}, not the Fejér function.

c. Using the result of part \textit{b}, Wik [Wik65] proved that if \( F \in A(\mathbb{T}) \), 
\(-1/2 < \alpha < \beta < 1/2\), \( F(\alpha) = F(\beta) = 0 \), and
\[
\sum' |F^\gamma[n]| \log |n| < \infty,
\]
then \( F_{\alpha,\beta} \in A(\mathbb{T}) \) where
\[
\forall \gamma \in [-\frac{1}{2}, \frac{1}{2}), \quad F_{\alpha,\beta}(\gamma) = F(\gamma) 1_{[\alpha,\beta]}(\gamma)
\]
and \( F_{\alpha,\beta} \) is defined \textit{1}-periodically on \( \mathbb{R} \).

d. Condition (3.5.15) is a \textit{doubling condition} for weights. Such conditions play a conceptually important role in an interesting and unresolved set of problems categorized as \textit{weighted norm inequality problems}, e.g., [G-CRdeF85]. An example of a weighted norm inequality is
\begin{equation}
(3.5.16) \quad \left( \int |\hat{f}(\gamma)|^2 \, d\mu(\gamma) \right)^{1/2} \leq C \left( \int |f(t)|^2 \, w(t) \, dt \right)^{1/2},
\end{equation}
where \( w > 0 \) and \( \mu \) is a positive measure, cf., \textit{Definition 2.6.5} and \textit{Theorem 3.7.2}, which are applicable raisons d'etre for dealing with such inequalities. The problem is to characterize the relationship between \( w \) and \( \mu \) so that (3.5.16) is valid for a large class of functions, e.g., [BH92] for measure weights \( \mu \).

e. Wik's Theorem from part \textit{c} can be thought of in terms of local membership or weighted norm inequalities. In the context of \textit{local membership}, we can obtain \( F_{\alpha,\beta} \in A(\mathbb{T}) \) by \textit{Theorem 3.5.7} if we can show \( F_{\alpha,\beta} \in A(\mathbb{T})_{\text{loc}}(\gamma) \) for \( \gamma = \alpha, \beta \), since local membership is obvious for other values of \( \gamma \). In the context of \textit{weighted norm inequalities}, define \( w(n) = \log |n| \) and \( Fv = F_{\alpha,\beta}, \text{i.e., } v = 1_{[\alpha,\beta]} \) on \([-\frac{1}{2}, \frac{1}{2})\); and consider the weighted norm inequality,
\begin{equation}
(3.5.17) \quad \|Fv\|_{A(\mathbb{T})} \leq C \sum' |F^v[n]|w(n),
\end{equation}
for all \( F \in A(\mathbb{T}) \) for which \( F(\alpha) = F(\beta) = 0 \). Then Wik's Theorem can be restated by saying that if the right side of (3.5.17) is finite then \( Fv = F_{\alpha,\beta} \in A(\mathbb{T}) \).
f. Geometrical considerations play a significant role in a class of weighted norm inequalities referred to as restriction theory, e.g., [Ash76, pages 107–117 by E. M. Stein], [Ste93]. Also, extensions of the classical uncertainty principle inequality (2.8.5) are critical in quantifying the implications of inequalities such as (3.5.16), e.g., [BF94, Chapter 7], [BL94].

3.6 Maximum entropy and spectral estimation

We shall discuss the Maximum Entropy Theorem and prove a spectral estimation theorem. In so doing we shall prove the Fejér-Riesz Theorem and indicate the role of \( A(T) \) in such matters.

3.6.1 Remark. The Spectral Estimation and Extension Problem

a. We gave a qualitative statement of the spectral estimation problem in Definition 2.8.6b. We shall now aim to quantify that statement for both the stochastic setting of Section 2.8 and the deterministic setting of Section 2.9. As a first step, we say that the spectral estimation problem is to estimate the power spectrum in terms of given autocorrelation data on a finite interval.

In the context of Fourier series, we are given \( N > 0 \) and data \( X_N \equiv \{ r_n : n = 0, \pm 1, \ldots, \pm N \} \subseteq \mathbb{C} \), and the extension problem associated with spectral estimation is to find nonnegative functions \( S \in L^1(T) \) for which \( S^n[n] = r_n \) for \( n = 0, \pm 1, \ldots, \pm N \). Because of Herglotz's Theorem on \( \mathbb{Z} \) (Theorem 2.7.10), \( X_N \) must satisfy some positive definiteness condition, e.g., Definition 3.6.2. With this stipulation on \( X_N \) there are generally many nonnegative solutions \( S \in L^1(T) \) as Krein (1940) first showed in the setting of \( \mathbb{R} \). After Krein, contributions to the extension problem were made by Chover, Doss, Dym, Gohberg, R. R. Goldberg, Rudin, et al.; and [Rud63] also analyzed the difficulties in extending Krein's Theorem to higher dimensions. The fact that there can be many solutions to the extension problem leads to Fourier uniqueness problems in the spirit of Example 1.10.6, e.g., [Pric85, pages 149–170].
3.6. **MAXIMUM ENTROPY AND SPECTRAL ESTIMATION** 247

b. From the point of view of spectral estimation there are various ways of choosing a specific solution \( S \) from part \( a \), depending on the type of application. One procedure, the *Maximum Entropy Method* (MEM), involves choosing the function \( S = S_{MEM} \) which maximizes a certain logarithmic integral associated with entropy, e.g., Theorem 3.6.3.

Mathematically, we shall see that this choice restricts us to \( A(T) \), instead of \( L^1(T) \). Physically, since entropy is a measure of disorder in a system, \( S_{MEM} \) represents maximum uncertainty with regard to what we do not know about the system, whereas it depends on all the known autocorrelation data \( X_N \). Thus, the choice of \( S_{MEM} \) is a mathematical guarantee that the least number of assumptions has been made regarding the information content of the unmeasured data at \( |n| > N \), e.g., [Chi78], [IEEE82] for expert physical rationales and expositions.

John Parker Burg invented MEM in 1967 and van den Bos (1971) [Chi78, pages 92-93] showed that the Maximum Entropy Method of choosing a *spectral estimator* \( S \) is equivalent to least squares linear prediction, used in speech processing, and autoregression, used in statistics, e.g., Section 3.7. MEM is also related to the maximum likelihood method, e.g., [Chi78, page 3 and pages 132-133]. There is a deep study of MEM and moment problems by Landau [Lan87], *Definition 2.7.8c*, as well as an important new mathematical contribution by Gabardo [Gab93], cf., our extension of MEM to \( \mathbb{R} \) in *A quantitative maximum entropy theorem for the real line*, Integral Equations and Operator Theory, 10(1987), 761-779.

### 3.6.2 Definition. Positive Definite Matrices

a. An \( (N+1) \times (N+1) \) matrix \( R = (r_{jk}) \), where \( r_{jk} \in \mathbb{C} \) and \( 0 \leq j, k \leq N \), is *Hermitian* if \( r_{jk} = \overline{r_{kj}} \). An \( (N+1) \times (N+1) \) matrix \( R = (r_{jk}) \) is *positive semidefinite* if

\[
\forall c = (c_0, c_1, \ldots, c_N) \in \mathbb{C}^{N+1}, \langle Rc, c \rangle = \sum_{j,k} r_{jk}c_k\overline{c}_j \geq 0.
\]

Positive semidefinite matrices are Hermitian, e.g., *Exercise 2.52*. 
b. An \((N + 1) \times (N + 1)\) positive semidefinite matrix \(R\) is positive definite written \(R \gg 0\), if \(\langle Rc, c \rangle = 0\) implies \(c = 0\). Clearly, if \(R \gg 0\) then \(R\) is nonsingular and \(R^{-1}\) exists. In fact, if \(Rc = 0\) then \(\langle Rc, c \rangle = 0\), and so \(c = 0\) by hypothesis; thus, \(R : \mathbb{C}^{N+1} \to \mathbb{C}^{N+1}\) is a linear injection and we have the result.

c. Let \(R\) be an \((N + 1) \times (N + 1)\) matrix with eigenvalues \(\{\lambda_0, \ldots, \lambda_N\}\). By definition, the trace of \(R\) is \(\sum_{j=0}^{N} r_{jj}\). It can be shown that the trace of \(R\) equals \(\sum_{j=0}^{N} \lambda_j\) and that the determinant of \(R\) is \(\prod_{j=0}^{N} \lambda_j\). Also, \(R\) is Hermitian if and only if

\[
\forall c, d \in \mathbb{C}^{N+1}, \quad \langle Rc, d \rangle = \langle c, Rd \rangle;
\]

and the eigenvalues of an Hermitian matrix are real.

Finally, if \(R\) is Hermitian, then \(R \gg 0\) if and only if each eigenvalue \(\lambda_j > 0\), e.g., [Str88], cf., part a.

d. Let \(\{r_j : j = 0, \pm 1, \pm 2, \ldots, \pm N\} \subseteq \mathbb{C}\) satisfy the condition \(r_j = \bar{r}_{-j}\) for each \(j\), and define the \((N + 1) \times (N + 1)\) matrix \(R = (r_{jk})\), where \(j, k \geq 0\) and \(r_{jk} = r_{j-k}\). \(R\) is a Toeplitz matrix, i.e., it takes constant values on "diagonals of negative slope". \(R\) is Hermitian since \(r_j = \bar{r}_{-j}\). From the previous discussion, \(\sum_{j=0}^{N} \lambda_j = (N + 1)r_0\); and if \(R \gg 0\) then the determinant of \(R\) is positive, \(r_0 > 0\), and \(R^{-1} = (c_{jk})\) exists, cf., Exercise 3.52.

e. Let \(S \in L^1(\mathbb{T})\setminus\{0\}\) be nonnegative, and let \(s = \{s[n]\}\) be the sequence of Fourier coefficients of \(S\). Then, for each \(N\), the \((N + 1) \times (N + 1)\) matrix \(R = (s[j-k])\), where \(0 \leq j, k \leq N\), is positive definite. In fact,

\[
\sum_{0 \leq j, k \leq N} s[j-k]c_k \bar{c}_j = \int_{\mathbb{T}} S(\gamma) \left| \sum_{j=0}^{N} c_j e^{-2\pi i j\gamma} \right|^2 d\gamma > 0,
\]

for all \((c_0, \ldots, c_N) \in \mathbb{C}^{N+1}\setminus\{0\}\), since \(\sum_{j=0}^{N} c_j e^{-2\pi i j\gamma} = 0\) for at most finitely many points.

3.6.3 Theorem. The Maximum Entropy Theorem

Let \(\{r_j : j = 0, \pm 1, \pm 2, \ldots, \pm N\} \subseteq \mathbb{C}\) satisfy the condition \(r_j = \bar{r}_{-j}\) for each \(j\), and assume the \((N + 1) \times (N + 1)\) matrix \(R = (r_{jk}) \gg 0\).
3.6. MAXIMUM ENTROPY AND SPECTRAL ESTIMATION

0, where \( j, k \geq 0 \) and \( r_{jk} \equiv r_{j-k} \). There is a unique function \( S \in A(\mathbb{T}) \), with Fourier coefficients \( \{s[n]\} \in l^1(\mathbb{Z}) \), satisfying the following properties:

a. \( \forall |n| \leq N, \quad s[n] = r_n \);

b. \( S > 0 \) on \( \mathbb{T} \), and, hence (by Theorem 3.5.9) \( S^{-1} \in A(\mathbb{T}) \);

c. \( S = |S_+|^2 \) where \( S_+ \in A(\mathbb{T}) \) has the form

\[
S_+(\gamma) = \sum_{n=0}^{\infty} s_+[n] e^{-2\pi in\gamma};
\]

d. \( \forall |n| > N, \quad (S^{-1})'[n] = 0 \) and

\[
S(\gamma) = 1/ \sum_{n=-N}^{N} \left( \frac{1}{c_0} \sum_{m=0}^{N} c_{0m} c_{0, m-n} \right) e^{-2\pi in\gamma},
\]

where \( R^{-1} = (c_{mn}) \);

e. For all \( F \in A(\mathbb{T}) \), for which \( F > 0 \) on \( \mathbb{T} \) and \( F'[n] = r_n \) when \( |n| \leq N \), we have

\[
\int_{\mathbb{T}} \log F(\gamma) \, d\gamma \leq \int_{\mathbb{T}} \log S(\gamma) \, d\gamma,
\]

and equality is obtained if and only if \( F = S \).

Our proof of Theorem 3.6.3 in [Pric85, pages 95–97] depends on the fact that if the matrix \( R \) of Theorem 3.6.3 is positive definite and if \((a_0, a_1, \ldots, a_N)^T = R^{-1}(1, 0, 0, \ldots, 0)^T\), then

\[
\forall \gamma \in \mathbb{T}, \quad \sum_{n=0}^{N} a_n e^{-2\pi in\gamma} \neq 0,
\]

e.g., [GS58], [GL94], cf., [DG79] for an important extension. We shall not prove Theorem 3.6.3 since we shall not prove this fact. Instead we shall prove Theorem 3.6.6 below, which is essentially Theorem 3.6.3 and which depends on the Fejér-Riesz Theorem.

3.6.4 Theorem. FEJÉR-RIESZ THEOREM
Let

\[ A(\gamma) = \sum_{n=-N}^{N} a_n e^{-\pi in\gamma/\Omega} \geq 0 \quad \text{on} \quad T_{2\Omega} \]

and define \( A_*(z) = \sum_{n=-N}^{N} a_n z^n \) so that \( A_*(e^{-\pi i\gamma/\Omega}) = A(\gamma) \). There is a unique trigonometric polynomial \( B(\gamma) = \sum_{n=0}^{N} b_n e^{-\pi in\gamma/\Omega} \) with the properties that

\[ A = |B|^2 \quad \text{on} \quad T_{2\Omega}, \]

and if \( B_*(z) = 0 \) then \(|z| \leq 1\).

**Proof.** Since \( A \) is real, the Fourier coefficients \( \{a_n\} \) have the property that \( a_n = a_{-n} \) for \( n = -N, \ldots, 0, 1, \ldots, N \). In particular, \( a_0 \in \mathbb{R} \). Without loss of generality, assume \( a_{-N} \neq 0 \).

Now define the polynomial \( P_* \) by

\[ P_*(z) = z^N A_*(z) = a_{-N} + a_{-N+1} z + \cdots + a_0 z^N + a_1 z^{N+1} + \cdots + a_N z^{2N}, \quad z \in \mathbb{C}. \]

Clearly,

\[ \forall z \in \mathbb{C} \setminus \{0\}, \quad z^{2N} P_*(1/z) = P_*(z). \]

Let \( P_*(z_0) = 0 \). If \( z_0 \neq 0 \) then \( 1/z_0 \) is also a zero of \( P_* \).
3.6. MAXIMUM ENTROPY AND SPECTRAL ESTIMATION  251

If $0 < |z_0| < 1$ then, by differentiation, we see that $z_0$ and $1/z_0$ have the same multiplicity. If $|z_0| = 1$ then $z_0$ has even multiplicity since $A \geq 0$. We shall say that a polynomial has $0$ as a zero of order $m$ if its first $m$ coefficients, that is, those of $1, z, \ldots, z^{m-1}$, are 0; and it has $\infty$ as a zero of order $m$ if its last $m$ coefficients are 0. Thus, if $z_0 = 0$ is a zero of order $m$ for $P_\gamma$, then $P_\gamma$ has $\infty$ as a zero of order $m$.

The previous discussion allows us to write

$$P_\gamma(z) = C \prod_{j=1}^{p} (z - z_j) \left( z - \frac{1}{z_j} \right) \prod_{j=1}^{q} (z - u_j)^2,$$

where $0 < |z_j| < 1$, $|u_j| = 1$, and $2p + 2q = 2N$, i.e., $p + q = N$, cf., Exercise 3.28i.

Since $A \geq 0$, (3.6.2) allows us to write

$$A(\gamma) = |A_\gamma(z)| = |z^N P_\gamma(z)| = |P_\gamma(z)|$$

(3.6.3)

$$= |C| \prod_{j=1}^{p} |z - z_j||z - \frac{1}{z_j}| \prod_{j=1}^{q} |z - u_j|^2$$

for $z = e^{-\pi i \gamma /\Omega}$, where $p + q = N$. Because $|z - \frac{1}{z_j}| = |\frac{z - z_j}{z_j}|$ for such $z$ on the unit circle, (3.6.3) becomes

$$A(\gamma) = \left| C^{1/2} \prod_{j=1}^{p} (e^{-\pi i \gamma /\Omega} - z_j) z_j^{-1/2} \prod_{j=1}^{q} (e^{-\pi i \gamma /\Omega} - u_j) \right|^2.$$

(3.6.1) is obtained by setting

$$B(\gamma) = C^{1/2} \prod_{j=1}^{p} (e^{-\pi i \gamma /\Omega} - z_j) z_j^{-1/2} \prod_{j=1}^{q} (e^{-\pi i \gamma /\Omega} - u_j).$$

The claim about the zeros of $B_\gamma$ is immediate from (3.6.4).

3.6.5 Remark. FEJÉR-RIESZ THEOREM: POTPOURRI AND TITILLATION

a. The Fejér-Riesz Theorem was proved by Fejér and F. Riesz, and published by Fejér (Jour. für Reine und Angew. Math., 146(1915),

b. The Fejér-Riesz Theorem is a critical component in the classical proof of the Spectral Theorem for Unitary Operators in a Hilbert space, e.g., [RN55, pages 280–284], cf., the historical note on operator-theoretic applications in [Bur79].

c. Herglotz's Theorem (1911) on $\mathbb{Z}$ (Theorem 2.7.10) can be proved as a consequence of the Fejér-Riesz Theorem, e.g., [RN55, pages 115–118], where the context is in terms of the moment problems mentioned in Definition 2.7.8c.

d. Using the Fejér-Riesz Theorem, it is elementary to prove that if $F \in C(T_{2\Omega})$ is nonnegative, then there is a sequence $\{B_N\}$ of trigonometric polynomials on $T_{2\Omega}$ for which

$$\lim_{N \to \infty} \|F - |B_N|^2\|_{L^\infty(T_{2\Omega})} = 0,$$

(cf., Exercise 3.48).

e. Krein Theorem. Using the Fejér-Riesz Theorem, Krein proved that if $f \in PW_\Omega$ is nonnegative then there is $b \in PW_{\Omega/2}$ for which $f = |b|^2$ on $\mathbb{R}$ and for which the zeros of the entire function,

$$b(z) = \int_{-\Omega/2}^{\Omega/2} b(\gamma)e^{2\pi i z \gamma} d\gamma,$$

are in the half-plane $\text{Im} z \geq 0$, e.g., [Ach56, page 154]. ($PW_\Omega$ is the Paley-Wiener space defined in Remark 1.10.8.) In fact, this result is true for a larger space than $PW_\Omega$, e.g., [Ach56, pages 137–152].

f. If $A(\gamma) = \sum_{n=-N}^{N} a_n e^{-2\pi in\gamma} \geq 0$ on $\mathbb{T}$ and $a_0 = 1$ then $A(0) \leq 2N + 1$. This is an immediate consequence of the fact that $\{a_n\} >> 0$:

$$0 \leq P(0) = \sum_{n=-N}^{N} a_n \leq \sum_{n=-N}^{N} |a_n| \leq (2N + 1)a_0.$$

In his paper referenced in part a, Fejér proved that $A(0) \leq N + 1$.

g. The fact that the zeros $z_0$ of the polynomial $B_\ast$, of Theorem 3.6.4, satisfy $|z_0| < 1$ for $A > 0$ on $T_{2\Omega}$ is useful in filter design. A rational
function $H$, all of whose zeros and poles $z_0$ satisfy $|z_0| < 1$, is a minimum phase filter, e.g., [OS75, pages 345–353], [Dau92, pages 194 ff.].

h. **Daubechies Theorem.** One of the early stunning successes of wavelet theory was Daubechies' Theorem (1987): for any $r \geq 0$, there is a constructible function $\psi \in C_r^0(\mathbb{R})$ for which $\{\psi_{m,n} : m,n \in \mathbb{Z}\}$ is an ONB for $L^2(\mathbb{R})$, where

$$\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n).$$

Her proof requires the Fejér-Riesz Theorem, e.g., [Dau92, Chapter 6], especially pages 167–174 for the role of Theorem 3.6.4.

i. The functions $A_\ast$ and $B_\ast$ of Theorem 3.6.4 are called $z$-transforms of $\{a_n\}$ and $\{b_n\}$, respectively, e.g., Exercise 3.22.

**3.6.6 Theorem. A Spectral Estimation Theorem**

Let $\{r_j : j = 0, \pm 1, \pm 2, \cdots \pm N\} \subseteq \mathbb{C}$ satisfy the condition $r_j = r_{-j}$ for each $j$, and assume the $(N + 1) \times (N + 1)$ matrix $R = (r_{jk}) > 0$, where $j,k \geq 0$ and $r_{jk} = r_{j-k}$. Let $\Omega > 0$. There is a positive function $S \in A(\mathbb{T}_{2\Omega})$, with Fourier coefficients $\{s[n]\} \in \ell^1(\mathbb{Z})$, satisfying the following properties:

$$\forall |n| \leq N, \quad s[n] = r_n,$$

and

$$\forall \gamma \in \mathbb{T}_{2\Omega}, \quad S(\gamma) = 1/A(\gamma) = 1/\sum_{n=-N}^{N} a_n e^{-\pi i n \gamma/\Omega},$$

where $A$ designates the sum in the denominator on the right side of (3.6.6).

**Proof.** i. In order to prove this result we shall proceed in the following devious way.

Given the hypotheses on $R$ we shall momentarily assume both (3.6.5) and (3.6.6) for some nonnegative function $S \in L^1(\mathbb{T}_{2\Omega})$. Using these hypotheses and conclusions, we shall show in parts ii–v how to obtain the coefficients $\{a_n : |n| \leq N\}$ of (3.6.6) from the given data $\{r_n : |n| \leq N\}$. 

In order to give an honest proof of the theorem, we work backwards. In particular, we take the hypotheses on \( R \) (without the conclusions of the theorem!), and solve the system of equations in part \( v \) to obtain \( \{a_n\} \). Then we define \( S \) in terms of these \( a_n \) by means of (3.6.6). The calculations in parts \( u-v \) allow us to conclude that \( S \) is nonnegative, \( S \in A(T_{\Delta}) \), and \( \hat{S}[n] = r_n \) for all \( |n| \leq N \).

ii. We shall prove that \( S \) is not only in \( L^1(T_{\Delta}) \), but \( S \in A(T_{\Delta}) \). In fact, \( A \) is a trigonometric polynomial, and so it is a continuous function on \( T_{\Delta} \) with at most finitely many zeros. However, \( S \) is not integrable over any small interval centered at such a zero, e.g., Exercise 3.27, and thus \( S \) and \( A \) are really positive on \( T_{\Delta} \). Thus, \( S = 1/A \in A(T_{\Delta}) \) by Wiener’s Theorem (Theorem 3.5.9).

iii. Since \( A > 0 \) on \( T_{\Delta} \), we can apply the Fejér-Riesz Theorem (Theorem 3.6.4). Thus, there is a trigonometric polynomial \( B(\gamma) = \sum_{n=0}^{N} b_n e^{-\pi i n / \Delta} \) with the properties that \( A = |B|^2 \) on \( T_{\Delta} \), \( b_0 \neq 0 \), and for which
\[
B_*(z) = \sum_{n=0}^{N} b_n z^n = 0 \quad \text{implies} \quad |z| < 1,
\]
where \( B_*(e^{-\pi i n / \Delta}) \equiv B(\gamma) \). In fact, the proof of Theorem 3.6.4 allows us to write
\[
B_*(z) = C^{1/2} \prod_{j=1}^{N} (z - z_j) z_j^{-1/2},
\]
where \( \{z_j\} \) is the set of zeros of \( B_* \) and each \( z_j \) satisfies \( 0 < |z_j| < 1 \).

Therefore,
\[
\overline{B_*\left(\frac{1}{z}\right)} = C^{1/2} \prod_{j=1}^{N} \left(\frac{1}{z} - z_j\right) \overline{z_j}^{-1/2} = \sum_{n=0}^{N} b_n z^{-n},
\]
and its zeros \( z = 1/\overline{z_j} \) are outside the unit circle. Hence,
\[
(3.6.7) \quad \frac{1}{B_*\left(\frac{1}{z}\right)} = \sum_{n=0}^{\infty} c_n z^{-n}, \quad |z| \leq 1.
\]
Further, if \( z = e^{-\pi i n / \Delta} \) then
\[
(3.6.8) \quad \overline{B_*\left(\frac{1}{z}\right)} = B(\gamma).
\]
3.6. MAXIMUM ENTROPY AND SPECTRAL ESTIMATION

iv. Combining the factorization \( A = |B|^2 \) with (3.6.6), we obtain

\[
S(\gamma)B(\gamma) = \frac{1}{B(\gamma)} \in \mathcal{A}(\mathbb{T}_m).
\]

Because of (3.6.7) and (3.6.8) the Fourier coefficients \((SB)^\vee[n]\) vanish if \( n > 0 \).

Since \((SB)^\vee[n] = \hat{s} * \hat{b}[n]\), we have from (3.6.9) that

\[
\forall n > 0, \quad \sum_{k=0}^{N} s[n - k]b_k = 0.
\]

v. Because of property (3.6.5), the \( N \) cases of (3.6.10) for \( n = 1, \ldots, N \) give rise to the \( N \) equations

\[
\begin{align*}
r_1 b_0 + r_0 b_1 + r_{-1} b_2 + \cdots + r_{1-N} b_N &= 0, \\
r_2 b_0 + r_1 b_1 + r_0 b_2 + \cdots + r_{2-N} b_N &= 0, \\
& \quad \ldots \\
r_N b_0 + r_{N-1} b_1 + r_{N-2} b_2 + \cdots + r_0 b_N &= 0.
\end{align*}
\]

We rewrite (3.6.11) as the matrix equation,

\[
Rb^T = -b_0(r_1, \ldots, r_N)^T,
\]

where \( b = (b_1, b_2, \ldots, b_N) \). \( R \) is invertible since \( R >> 0 \), e.g., Definition 3.6.2c,d. Thus, we compute \( b \), cf., Example 3.7.9; and then we compute \( \{a_n : -N \leq n \leq N\} \) in terms of \( b \) since \( A = |B|^2 \).

3.6.7 Example. Spectral Estimation and Maximum Entropy

a. Equation (3.6.6) gives spectral (frequency) information about a digital signal in the case that relatively little data \( X \) is known about the signal. In fact, a small value of \( A(\gamma) > 0 \) in (3.6.6) allows one to guess that this value of \( \gamma \) is a frequency component of the signal which generates given autocorrelation data \( X \). Of course, we very rarely get something for nothing, and so this method of spectral estimation can only be used effectively when, as indicated in Remark 3.6.1b, certain physical parameters make sense.
This method of spectral estimation is a form of the MEM, and should be compared with Fourier transform methods, e.g., Proposition 2.8.8 and Example 2.9.7, which really give accurate spectral information but which usually require large data sets $X$.

b. Let $\{r_n : |n| \leq N\} \subseteq \mathbb{C}$ satisfy the hypotheses of either the MEM Theorem (Theorem 3.6.3) or Theorem 3.6.6. The relationship between these theorems was alluded to in part a, and is quantified by the following suggestive calculation. The calculation itself was made early on in the development of MEM, e.g., [Chi78, page 55], [IEEE82, page 944]. It is a rationale (not a proof) for supposing that $S$ has the form (3.6.6) in the case that $\int_T \log S(\gamma) \, d\gamma$ is the largest (or smallest) value of \( \int_T \log F(\gamma) \, d\gamma \), when $F$ ranges over all positive functions $F$ in $A(T)$ for which $\hat{F}[n] = r_n$, $|n| \leq N$.

Let $|n| > N$ be fixed, and consider the continuous (complex) variable $r \equiv r_n$. Assuming $\int_T \log S(\gamma) \, d\gamma$ is an optimum we have

$$0 = \frac{\partial}{\partial r} \int_T \log S(\gamma) \, d\gamma = \int_T \frac{1}{S(\gamma)} \frac{\partial S(\gamma)}{\partial r} \, d\gamma,$$

where $\gamma$ is fixed in the expression $\frac{\partial S(\gamma)}{\partial r}$, cf., the proof of Theorem 3.7.7. Writing $S(\gamma) = \sum r_j e^{-2\pi i j \gamma}$ we have $\frac{\partial S(\gamma)}{\partial r} = e^{-2\pi i \gamma}$. Thus, $(1/S)^\gamma[n] = 0$ for all $|n| > N$, and so $S$ has the form (3.6.6).

### 3.7 Prediction and spectral estimation

An extension of the factorization given by the Fejér-Riesz Theorem (Theorem 3.6.4) is the Szegö Factorization Theorem: let $A \in L^1(T)$ be nonnegative; then $\log A \in L^1(T)$ if and only if $A = |B|^2$ for some $B \in H^2(T)$, where

$$H^2(T) = \{F \in L^2(T) : \forall n < 0, \quad \hat{F}[n] = 0\}.$$  

Proofs can be found in [GS58, Section 1.14], [Hof62, pages 48–54]. We shall not prove Szegö's Factorization Theorem, but mention it because Szegö's original proof (Math. Ann. 84(1921), 232–244) depended on
3.7. PREDICTION AND SPECTRAL ESTIMATION

the following result which he proved in 1920 (Math. Zeit. 6(1920), 167-202). (The proof of the Szegö Factorization Theorem in [GS58, Chapter 1.14] is joint work of F. Riesz and Szegö which actually appeared before Szegö’s original proof, cf., [MW57, pages 115 ff.].)

3.7.1 Theorem. Szegö Alternative

Let \( W \in L^1(\mathbb{T}) \) be nonnegative, and define the “geometric mean” of \( W \) as

\[
g(W) = \begin{cases}  
\exp \int_{\mathbb{T}} \log W(\gamma) \, d\gamma, & \text{if } \log W \in L^1(\mathbb{T}), \\
0, & \text{if } \log W \notin L^1(\mathbb{T}).
\end{cases}
\]

Then

\[
\inf_{P} \int_{\mathbb{T}} |1 - P(\gamma)|^2 W(\gamma) \, d\gamma = g(W),
\]

where the infimum is taken over all trigonometric polynomials \( P \) on \( \mathbb{T} \) of the form

\[
P(\gamma) = \sum_{n=1}^{N} a_n e^{-2\pi i n \gamma},
\]

and where \( N \geq 1 \) and \( a_1, \ldots, a_N \) vary.

The Szegö Alternative has been generalized in several directions, e.g., [Ach56, pages 256 ff.], [DM76], [Hel64, pages 19–24], [Hof62, pages 48–50]. It can also be reformulated as follows.

3.7.2 Theorem. Kolmogorov Theorem (1940)

Let \( W \in L^1(\mathbb{T}) \) be nonnegative, and define the space,

\[
L^2_W(\mathbb{T}) = \left\{ F : \|F\|_{L^2_W(\mathbb{T})} = \left( \int_{\mathbb{T}} |F(\gamma)|^2 W(\gamma) \, d\gamma \right)^{1/2} < \infty \right\}.
\]

Then

\[
(3.7.1) \quad \mathop{\text{span}}\{ e^{-2\pi i n \gamma} : n \leq 0 \} = L^2_W(\mathbb{T})
\]

if and only if \( \log W \notin L^1(\mathbb{T}) \), where \( \mathop{\text{span}}X \) is defined as the closure in \( L^2_W(\mathbb{T}) \), taken with the \( \| \cdot \|_{L^2_W(\mathbb{T})} \)-norm, of the linear span of elements from \( X \).
In the spirit of Section 2.8, we state the following definition.

3.7.3 Definition. **Stationary Sequences and Power Spectra**

a. Let \( f \in L^2(\mathbb{Z}) \). The \( L^2 \)-autocorrelation of \( f \) is the sequence \( p : \mathbb{Z} \to \mathbb{C} \) defined as

\[
\forall n \in \mathbb{Z}, \quad p[n] = \sum_m f[n + m] \overline{f[m]},
\]

cf., Example 2.7.9 for \( L^2 \)-autocorrelation. By the Parseval Formula, \( \hat{p} = |F|^2 \in L^1(\mathbb{T}) \). \( S \equiv |F|^2 \) is the **power spectrum** of \( f \).

b. A sequence \( x = \{x[n] : n \in \mathbb{Z}\} \) in a complex Hilbert space \( H \) is **stationary** if the inner product

\[
r[n] = \langle x[n + k], x[k] \rangle, \quad n \in \mathbb{Z},
\]
is independent of \( k \), e.g., Definition B.4. The sequence \( r \) is the autocorrelation of \( x \), and it is elementary to check that \( r \) is a positive definite function on \( \mathbb{Z} \). By Herglotz's Theorem, there is \( S \in M_+(\mathbb{T}) \) for which \( \hat{r} = S \). \( S \) is the power spectrum of \( x \), e.g., Definition 3.10.3.

3.7.4 Remark. **Prediction and Kolmogorov's Theorem**

a. Theorem 3.7.2 is valid in more general contexts, including replacing the weight \( W \) by any \( \mu \in M_+(\mathbb{T}) \), e.g., [Kol41], cf., [Ach56, pages 261–263].

b. Let \( H \) be a Hilbert space, and let \( x \in H \) be a stationary sequence with power spectrum \( S \). Define \( H(x) \) as \( \overline{\text{span}}\{x[n]\} \subseteq H \). Kolmogorov noted that there is a unique linear mapping,

\[
Z : L^2_S(\mathbb{T}) \longrightarrow H(x),
\]
defined on the exponentials as \( Z(e^{inx}) = x[n] \), which is an isometric isomorphism, i.e., \( Z \) is a bijection and \( \|F\|_{L^2_S(\mathbb{T})} = \|Z(F)\| \) for all \( F \in L^2_S(\mathbb{T}) \), where \( \|\ldots\| \) is the norm on \( H \).

c. Using Theorem 3.7.2 and the result of part b, Kolmogorov solved the problem of "predicting the future from the whole past" [Kol41], cf., [DM76], [Ben92a]. We shall not go into the details of defining deterministic sequences which are required for a clear statement of the
prediction problem which Kolmogorov solved. However, one has the intuition of prediction from (3.7.1) in the sense that the past information \( \{e^{2\pi i n\gamma} : n \leq 0\} \) is sufficient to approximate or predict any \( F \in L^2_W(\mathbb{T}) \), including exponentials \( F(\gamma) = e^{2\pi i n\gamma} \) for times \( n > 0 \).

We shall now attempt to quantify Remark 3.7.4c for the practical matter of addressing prediction problems that arise in analyzing bioelectric traces, speech data, economic and weather trends, and a host of time series from a variety of subjects. We begin with the following example.

3.7.5 Example. Prediction Estimates

a. Let \( \Omega > 0 \) and let \( \tilde{f} = F \), where \( F \in L^2(\mathbb{T}_{2\Omega}) \) has Fourier coefficients \( f = \{f[n]\} \in \ell^2(\mathbb{Z}) \). Suppose that for a given value of \( n \), \( f[n] \) is not explicitly known, whereas \( H_n \equiv \{f[n - k] : k \geq 1\} \) is known. When is it possible to predict, i.e., approximate or evaluate, \( f[n] \) in terms of \( H_n \)? One way of addressing this question is to write

\[
(3.7.2) \quad \epsilon_N[n] = f[n] - \sum_{k=1}^{N} a_k f[n - k],
\]

so that

\[
\epsilon_N[n] = \int_{\mathbb{T}_{2\Omega}} F(\gamma) e^{\pi i n\gamma/\Omega} d\gamma - \sum_{k=1}^{N} a_k \int_{\mathbb{T}_{2\Omega}} F(\gamma) e^{\pi i (n-k)\gamma/\Omega} d\gamma,
\]

from which we have the estimate

\[
|\epsilon_N[n]| \leq \int_{\mathbb{T}_{2\Omega}} |F(\gamma)| \left| 1 - \sum_{k=1}^{N} a_k e^{-\pi i k\gamma/\Omega} \right| d\gamma
\]

\[
(3.7.3) \quad \leq \|F\|_{L^2(\mathbb{T}_{2\Omega})} \left( \int_{\mathbb{T}_{2\Omega}} \left| 1 - \sum_{k=1}^{N} a_k e^{-\pi i k\gamma/\Omega} \right|^2 d\gamma \right)^{1/2}.
\]

In the transition from (3.7.2) to the right side of (3.7.3) we have squandered our information \( H_n \). Further, \( 1 \notin \text{span}\{e^{-\pi i k\gamma/\Omega} : k \geq 1\} \) in \( L^2(\mathbb{T}_{2\Omega}) \) by Theorem 3.4.12. This fact is corroborated by the Szegö
Alternative since the weight $W \equiv 1$ on the right side of (3.7.3) has the property that $\log W \in L^1(T_{2\Omega})$.

b. We now adjust the calculation of part $a$ by implementing the Parseval formula to compute

$$
\|\epsilon_N\|_{L^2(\Omega)} = \left( \int_{T_{2\Omega}} |F(\gamma)|^2 \left| 1 - \sum_{k=1}^{N} a_k e^{-\pi i k / \Omega} \right|^2 d\gamma \right)^{1/2}.
$$

Of course, $n$ is no longer fixed as it is in (3.7.2). On the other hand, if $\log W \notin L^1(T_{2\Omega})$, where $W \equiv |F|^2$, then the Szegő Alternative (Theorem 3.7.1) or Kolmogorov's Theorem (Theorem 3.7.2) can be used to glean prediction theoretic information in the following way, cf., Exercise 3.54. For each fixed $n$, $|\epsilon_N[n]| \leq \|\epsilon_N\|_{L^2(\Omega)}$, and so $\inf |\epsilon_N[n]| = \inf \|\epsilon_N\|_{L^2(\Omega)} = 0$ in the case $\log |F| \notin L^1(T_{2\Omega})$, where the infima are taken over all $N \geq 1$ and $a_1, \ldots, a_N \in \mathbb{C}$.

c. Our next adjustment of the calculation in part $a$ deals with analogue signals $f \in PW_{\Omega}$, instead of discrete signals in $L^2(\mathbb{Z})$. We write

$$
\epsilon_N(t) = f(t) - \sum_{k=1}^{N} a_k f(t - k \Omega)
$$

for a fixed $\alpha > 0$ and for a fixed $t \in \mathbb{R}$. Then we compute

$$
\epsilon_N(t) = \int \hat{f}(\gamma) e^{2\pi i t \gamma} d\gamma - \sum_{k=1}^{N} a_k \int \hat{f}(\gamma) e^{2\pi i (t - k\Omega) \gamma} d\gamma,
$$

cf., [Ben92b, Proposition 9]. Thus,

$$
(3.7.4) \ |\epsilon_N(t)| \leq \left( \frac{\Omega}{\alpha} \right)^{1/2} \|f\|_{L^2(\mathbb{R})} \left( \int_{-\alpha}^{\alpha} \left| 1 - \sum_{k=1}^{N} a_k e^{-2\pi i k \gamma} \right|^2 d\gamma \right)^{1/2}.
$$

As in (3.7.3), we have squandered information about $t$ on the right side of (3.7.4). However, it can be shown that if $\alpha < 1/2$ then

$$
(3.7.5) \ \text{span} \{e^{-2\pi i k \gamma} : k \geq 1\} = L^2[-\alpha, \alpha].
$$

Thus,

$$
\inf |\epsilon_N(t)| = 0,
$$
3.7. PREDICTION AND SPECTRAL ESTIMATION

where the infimum is taken over all \( N \geq 1 \) and \( a_1, \ldots, a_N \in \mathbb{C} \).

The density result (3.7.5) is due to Carleman, e.g., [You80, pages 114-116], cf., Remark 3.7.11.

d. Suppose we are given a fixed discrete "time" \( n \), resp., a fixed continuous "time" \( t \). In part b, resp., part c, we have shown in theory how to predict the value \( f[n] \), resp., \( f(t) \), in terms of its known values at previous times. This prediction requires \( f \) to satisfy certain conditions; and the prediction itself is made in terms of a given error bound \( \epsilon \). In fact, in both cases it can be proved that coefficients \( a_j, j = 1, \ldots, N \), exist with the property that \( |\epsilon_n[n]| < \epsilon \), resp., \( |\epsilon_N(t)| < \epsilon \).

We wish to design an applicable tool based on the idea of Example 3.7.5. We begin with the definition of linear prediction.

3.7.6 Definition. LINEAR PREDICTION MODELS

a. Let \( M, N \geq 1 \). A pole-zero linear prediction model of a sequence \( f \in \ell^2(\mathbb{Z}) \) is an equation of the form

\[
(3.7.6) \quad f[n] = \sum_{k=1}^{N} a_k f[n-k] + \sum_{j=0}^{M} b_j u[n-j],
\]

for each \( n \), where \( u \in \ell^2(\mathbb{Z}) \), \( \{a_k\}, \{b_k\} \subseteq \mathbb{R} \), and \( b_0 \neq 0 \).

\( f[n] \) can be thought of as the output of a linear translation-invariant system with some unknown input \( u \), including its past and present values, as well as input consisting of past values of \( f \), viz., \( f[n-1], \ldots, f[n-N] \). In this point of view, the goal is to estimate the system parameters \( \{a_k\}, \{b_k\} \subseteq \mathbb{R} \), so that (3.7.6) is a meaningful predictor of \( f \).

b. If \( b_1 = b_2 = \cdots = b_M = 0 \) then (3.7.6) is an all-pole model, or, equivalently, an autoregressive (AR) model. If \( a_1 = a_2 = \cdots = a_N = 0 \) then (3.7.6) is an all-zero model, or, equivalently, a moving average (MA) model. The full pole-zero model (3.7.6) is also called an autoregressive moving average (ARMA) model.

c. In the following, as in Example 3.7.5, we shall deal with the all-pole model. This type of linear prediction goes back to Yule's work (1927) on sun spot analysis, cf., the beginning of Section 2.9 for another
remark on sun spots. The mathematics underlying the effectiveness of linear prediction is Gauss' technique of linear least squares estimation from 1795, cf., Proposition 1.10.9. Major contributions were made by Kolmogorov and Wiener, independently, in the early 1940s, e.g., [Wei49] including Appendix C by Norman Levinson. Linear prediction is a staple in time series analysis, e.g., [BSS88], [JN84], [Mak75], [Pri81]. We mention, for example, the introduction of linear prediction into speech analysis and data compression by Fant, Jury, Atal and Schroeder, and Itakura and Saito in the 1960s, e.g., [MG82].

3.7.7 Theorem. Least Squares Method for All-Pole Model

Let \( N \geq 1 \), assume \( f \in \ell^2(\mathbb{Z}) \setminus \{0\} \) is real-valued with \( \ell^2 \)-autocorrelation \( r \), and consider the sequence of equations

\[
(3.7.7) \quad \forall n \in \mathbb{Z}, \quad f[n] = \sum_{k=1}^{N} a_k f[n - k] + u[n].
\]

There are unique coefficients \( a_1^*, \ldots, a_N^* \in \mathbb{R} \) and a sequence \( \epsilon_N \in \ell^2(\mathbb{Z}) \) defined by

\[
\epsilon_N[n] = f[n] - \sum_{k=1}^{N} a_k^* f[n - k]
\]

for each \( n \), such that

\[
\forall a = (a_1, \ldots, a_N) \in \mathbb{R}^N,
\]

\[
(3.7.8) \quad 0 \leq r[0] - \sum_{k=1}^{N} a_k^* r[k] = \| \epsilon_N \|^2_{\ell^2(\mathbb{Z})} \leq \| u \|^2_{\ell^2(\mathbb{Z})},
\]

where \( u \) is defined by (3.7.7). (Thus, \( u \) depends on \( a \in \mathbb{R}^N \) as opposed to the interpretation of Definition 3.7.6).

Proof. a. For a given \( n \in \mathbb{Z} \) and \( a = (a_1, \ldots, a_N) \in \mathbb{R}^N \), we define the approximation \( f_a[n] \) of \( f[n] \) as

\[
f_a[n] = \sum_{k=1}^{N} a_k f[n - k].
\]
Clearly, \((f - f_s) \in \ell^2(\mathbb{Z})\), and we set
\[
E_N(a) = \sum_n \left( f[n] - \sum_{k=1}^N a_k f[n - k] \right)^2 < \infty.
\]

b. For each \(k\), we have the formal calculation
\[
(3.7.9) \quad \frac{\partial E_N(a)}{\partial a_k} = -2 \sum_n \left( f[n] - \sum_{j=1}^N a_j f[n - j] \right) f[n - k].
\]

By Hölder's Inequality, the right side converges uniformly on any bounded interval of the \(a_k\)-axis. Thus, the right side of (3.7.9) is in fact \(\frac{\partial E_N(a)}{\partial a_k}\).

c. A necessary condition in order that \(E_N\) have a local minimum (or maximum) at \(a^* \in \mathbb{R}^N\) is that
\[
(3.7.10) \quad \forall k = 1, \ldots, N, \quad \frac{\partial E_N(a^*)}{\partial a_k} = 0.
\]

A sufficient condition in order that \(E_N\) (which satisfies (3.7.10)) have a local minimum at \(a^* \in \mathbb{R}^N\) is that \(D_{N-k}(a^*) > 0\) for each \(k = 0, 1, \ldots, N\), where \(D_0(a) = 1\) and \(D_{N-k}(a)\) is the determinant obtained from the \(N \times N\) matrix,

\[
\begin{pmatrix}
\frac{\partial^2 E_N(a)}{\partial a_m \partial a_n}
\end{pmatrix}, \quad m, n = 1, \ldots, N,
\]

by deleting the last \(k\) rows and columns, and taking the determinant of the resulting \((N - k) \times (N - k)\) matrix, e.g., [Apo57].

d. In our case, by (3.7.9), we have
\[
(3.7.11) \quad \forall a \in \mathbb{R}^N, \quad \frac{\partial^2 E_N(a)}{\partial a_m \partial a_n}
\]

\[
= 2 \sum_j f[j - m] f[j - n] = 2r[m - n] = 2r[n - m].
\]

Since \(f \in \ell^2(\mathbb{Z})\), we compute \(\hat{f} = |F|^2 \in L^1(\mathbb{T})\) by the Parseval Formula; and the \(N \times N\) matrix \(R = (r[m - n])\), \(m, n = 1, \ldots, N\), is positive
definite, as we showed in Definition 3.6.2e. An elementary characterization of positive definite matrices is that \( D_0(a), D_1(a), \ldots, D_N(a) \) be positive [Per52, Theorem 9-26].

Thus, our goal is to find values \( a^* \in \mathbb{R}^N \) for which (3.7.10) is valid; and, by the previous paragraph and the sufficient conditions for minima in part c, these values will in fact be minimizers of \( E_N \).

e. In order to obtain a candidate \( a^* \) for a minimizer, we rewrite (3.7.9) and (3.7.10) as the system of \( N \) linear equations in \( N \) unknowns \( a_1, \ldots, a_N \):

\[
\forall 1 \leq k \leq N,
\sum_{j=1}^{N} a_j \sum_n f[n-j]f[n-k] = \sum_n f[n]f[n-k],
\]

(3.7.12)

or, equivalently,

\[
\forall 1 \leq k \leq N, \quad \sum_{j=1}^{N} a_j r[k-j] = r[k].
\]

(3.7.13)

Since \( D_N(a) > 0 \) by (3.7.11) and the discussion in part d, the system (3.7.13) has a unique solution \( a^* = (a_1^*, \ldots, a_N^*) \in \mathbb{R}^N \).

Thus, the inequality in (3.7.8) is obtained.
f. Expanding the square in the definition of $E_N$, and substituting
the minimizer $a^*$ into (3.7.12), we obtain

$$0 \leq E_N(a^*) = \sum_n f[n]^2 - 2 \sum_{k=1}^N a_k^* \sum_n f[n]f[n-k]$$

$$+ \sum_n \left( \sum_{k=1}^N a_k^* f[n-k] \right)^2$$

$$= \sum_n f[n]^2 - 2 \sum_{k=1}^N a_k^* \sum_n f[n]f[n-k] + \sum_{k=1}^N a_k^* \sum_n f[n]f[n-k]$$

$$= \sum_n f[n]^2 - \sum_{k=1}^N a_k^* \sum_n f[n]f[n-k]$$

$$= r[0] - \sum_{k=1}^N a_k^* r[k].$$

This completes (3.7.8). \qed

3.7.8 Example. Spectral Estimation and the All-Pole Model

a. A typical and important issue in many problems and fields is to
find the spectral peaks or fundamental frequencies in a given sig-
nal $f \in \ell^2(\mathcal{Z})$. In theory, the graph of the Fourier series $F = \hat{f}$ will
provide this spectral information. In reality, there are potential prob-
lems. For example, trigonometric polynomial approximations of $\hat{f}$, or
approximations such as Proposition 2.8.8, may be inadequate because
they are either too good or too bad! In the former case, a very good ap-
proximation of $\hat{f}$ may have so much spectral information from "noise"
embedded in $f$ that desirable information about pure tones (in $f$) is
obscured when observing $\hat{f}$. In the latter case, the approximation may
not be developed enough to specify relevant frequencies in $\hat{f}$.

b. In cases, such as those hypothesized in part a, where the Fourier
transform can not be directly used to observe some fundamental fre-
quencies in a signal, there are other methods which sometimes provide
spectral information. The all-pole model is one of these methods. The
prediction estimates in Example 3.7.5 show that the prediction error
$\epsilon_N$ in Theorem 3.7.7 tends to 0 as $N \to \infty$ in many cases. The all-pole
model in (3.7.7) of Theorem 3.7.7 shows that such a model and its corresponding prediction error can be used to specify spectral peaks by the following process and rationale.

By taking the Fourier transforms of the sequences $f$ and $u$, (3.7.7) becomes

\[(3.7.14)\]
\[F(\gamma) \left(1 - \sum_{k=1}^{N} a_k e^{-2\pi i k \gamma}\right) = U(\gamma),\]

where $F$ and $U$ are Fourier series with Fourier coefficients $\{f[n]\}$ and $\{u[n]\}$, respectively. With the minimization effected by Theorem 3.7.7, equation (3.7.14) becomes

\[F(\gamma) = \frac{\hat{e}_N(\gamma)}{1 - \sum_{k=1}^{N} a_k^* e^{-2\pi i k \gamma}}.\]

Assume $0 < A \leq |\hat{e}_N(\gamma)| \leq B$, which is reasonable for many applications, e.g., [Chi78]. Suppose $F$ is continuous and $|F(\gamma_0)|$ is large in comparison to $|F(\gamma)|$ for values of $\gamma$ near $\gamma_0$, i.e., suppose $F$ has a spectral peak at $\gamma_0$. Define the polynomial

\[P_*(z) = 1 - \sum_{k=1}^{N} a_k^* z^k,\quad z \in \mathbb{C}.\]

We write $P(\gamma) = P_*(e^{-2\pi i \gamma})$. Then $|P(\gamma_0)|$ is small. This simple observation is the basis of the all-pole model method of spectral estimation.

In order to describe this method, we consider the following procedure for a given $N \geq 1$ and a given sequence $f$. First, choose a threshold $\epsilon > 0$ and consider the annular region $A_\epsilon = \{z : 1 - \epsilon \leq |z| \leq 1 + \epsilon\}$. The choice of $\epsilon$ can be adjusted according to the amount of spectral information desired. Next, compute $\{a_k^* : k = 1, \ldots, N\}$, cf., Example 3.7.9. Compute the zeros $z_0$ of $P_*$, e.g., in MATLAB use the roots command. If $z_0 \in A_\epsilon$, compute $\gamma_0$ by taking the projection $e^{-2\pi i \gamma_0}$ of $z_0$ to the unit circle.

$F(\gamma_0)$ is a candidate for a spectral peak.
c. As a caveat for our presentation in part b, we note that we have been precise about certain matters, e.g., Theorem 3.7.7, but quite cavalier about others. For example, the structure of \( u \) is important for the type of application at hand; and the proper behavior of \( \hat{u} \) is important for the success of the all-pole model method of spectral estimation in that application, e.g., [JN84, Section 2.4], [Pri81].

We should also mention that we are not using the all-pole method to predict values of \( f \) so much as to determine its spectral behavior. In fact, in order to compute \( a^* \) in Theorem 3.7.7 we assume knowledge of each \( f[n] \).

3.7.9 Example. Levinson Recursion Algorithm

a. Norman Levinson (1947) was the first to use the structure of a Toeplitz matrix (Definition 3.6.2d) to solve the system of linear equations (3.7.13) recursively, e.g., [Wie49, Appendix B]. This system not only plays a role in linear prediction (Theorem 3.7.7), but was also essential in the proof of Theorem 3.6.6, e.g., (3.6.11), which is associated with MEM. (Recall the equivalence of these methods noted in Remark 3.6.1b.) In statistics, equations (3.6.11) and (3.7.13) are called the Yule-Walker Equations. In making use of the Toeplitz structure, Levinson's algorithm has led to numerically realistic computations of the prediction coefficients \( a_1^*, \ldots, a_N^* \). For example, the classical Gauss elimination method requires \( N^3 + KN^2 \) multiplications or divisions, whereas even Levinson's original method only required \( N^2 + KN \) such operations.

b. If \( \{c_k : k \geq 0\} \subseteq \mathbb{C} \) and \( \exp(\sum_{k=0}^{\infty} c_k z^k) = \sum_{k=0}^{\infty} b_k z^k \), then one example of the Lebedev-Milin Inequalities is the inequality

\[
|b_n|^2 \leq \exp \left( \sum_{k=1}^{n} k |c_k|^2 - \sum_{k=1}^{n} \frac{1}{k} \right), \quad b_0 = 1.
\]

These inequalities have applications in univalent function theory, as well as in number theory and spectral synthesis, e.g., our construction of idelic pseudo-measures in Zeta functions for idelic pseudo-measures, Ann. Scuola Norm. Sup., 6(1979), 367–377.
It turns out that

$$b_{k+1} = \sum_{j=0}^{k} \left(1 - \frac{j}{k+1}\right) c_{k+1-j} b_j$$

[Pou84]; and that this recursion formula can be implemented numerically to deal with linear prediction problems. For example, in dealing with the Fejér-Riesz Theorem (Theorem 3.6.4), \(A = |B|^2\), we let \(\{c_k\}\) be the sequence of Fourier coefficients of \(\log A\), where \(A \geq 0\) is the given nonnegative polynomial. Then, by way of a standard argument in complex analysis, we obtain \(b_0 = e^{\alpha/2}\) and (3.7.15) for \(k = 0, \ldots, N - 1\), where \(b_0, \ldots, b_N\) are the Fourier coefficients of \(B\), e.g., [Pou84].

### 3.7.10 Remark. Ramifications of Szegö Factorization on \(\mathbb{R}\)

a. We first stated and discussed the Paley-Wiener Logarithmic Integral Theorem in Example 1.6.5c. It asserts that if \(\varphi \in L^2(\mathbb{R}) \setminus \{0\}\) is nonnegative, then there is \(f \in L^2(\mathbb{R})\), for which \(\text{supp } f \subseteq [0, \infty)\) and \(|\hat{f}| = \varphi\) a.e., if and only if

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

Although it is elementary to prove the Szegö Factorization Theorem on \(\mathbb{R}\) from this theorem of Paley and Wiener, e.g., [Pric85, pages 156-157], they were, in fact, motivated in their research by a result of Carleman on quasi-analytic functions [PW33], [PW34, Theorem XII], cf., [Koo88], [Rud66, Chapter 19] for the theory of quasi-analytic functions.

b. By an approximate identity argument, the Paley-Wiener Logarithmic Integral Theorem can be used to prove the following result. If \(\mu \in M(\mathbb{R})\), then \(\text{supp } \mu \subseteq [T, \infty)\) if and only if

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

c. The result for compactly supported measures analogous to that of part b is due to Beurling and Malliavin: let \(W\) be a continuous
3.7. PREDICTION AND SPECTRAL ESTIMATION

function for which \(|W| > 1\) and \(\log |W|\) is uniformly continuous; then the condition,

\[
\int \frac{\log |W(\gamma)|}{1 + \gamma^2} \, d\gamma < \infty,
\]

is necessary and sufficient that for all \(\varepsilon > 0\) there exists \(\mu \in M(\mathbb{R})\) such that \(\text{supp} \mu \subseteq [-\varepsilon, \varepsilon]\) and \(\bar{\mu} K \in L^\infty(\mathbb{R})\) [BM62], [Mal79].

3.7.11 Remark. Closure Theorems for Sets of Exponentials

In Example 3.7.5 we saw a role for the closure theorems of Szegö, Kolmogorov, and Carleman. There are other landmark contributions by Paley and Wiener [PW34], Levinson [Lev40], and Beurling and Malliavin [BM67], as well as deep results by others, cf., the superb expositions of [Red77], [You80], [Koo88]. We shall close this section with some perspective on such theorems.

a. Equation (3.7.5) was one of the first substantial results of an area which culminated in the work of Beurling and Malliavin [BM67]. Beurling and Malliavin solved the following closure problem for a given discrete subset \(D \subseteq \mathbb{R}\): find the upper bound \(\Omega \geq 0\) of the set of \(\alpha \geq 0\) for which

\[
\text{span}\{e^{-2\pi i t \gamma} : t \in D\} = L^2[-\alpha, \alpha].
\]

Their solution includes writing \(\Omega\) in terms of a density condition on \(D\).

b. Density results such as (3.7.6) are a weak form of sampling theorems—weak but not necessarily elementary. In fact, if (3.7.6) is valid and \(f \in PW_\alpha\), then there is a sequence of trigonometric polynomials \(P_n\), where

\[
P_n(\gamma) = \sum_{t \in D_n} a_{t,n} e^{-2\pi i t \gamma} \quad \text{and} \quad D_n \subseteq D,
\]

for which

\[
\lim_{n \to \infty} \|\hat{f} - P_n\|_{L^2[-\alpha, \alpha]} = 0.
\]

Distributionally, each \(P_n = \tilde{p}_n\), where \(p_n = \sum_{t \in D_n} a_{t,n} \delta_t\). Thus,

\[
\|\hat{f} - P_n\|_{L^2[-\alpha, \alpha]} = \|\hat{f} - P_n 1_{[-\alpha, \alpha]}\|_{L^2(\mathbb{R})} = \|f - p_n * d_{2\pi \alpha}\|_{L^2(\mathbb{R})},
\]
and so
\[ \lim_{n \to \infty} \sum_{i \in D_n} a_{i,n} \tau_i d_{2\pi a} = f \text{ in } L^2(\mathbb{R}), \]
cf., the sampling theorems in [Ben92b], [BF94], [BSS88], [Hig85].

c. Let $\mu \in M_+(\mathbb{R})$ and define
\[ L^1_\mu(\mathbb{R}) = \{ f : \|f\|_{L^1_\mu(\mathbb{R})} = \int |f(t)| d\mu(t) < \infty \}, \]
cf., Remarks 2.7.4a and 3.5.10d.

By Theorem 2.7.6, $\mu = w + \mu_s$ where $w \in L^1(\mathbb{R})$ is nonnegative and $\mu_s$ is the sum of the discrete and continuous singular parts of $\mu$. Krein proved the following $L^1$-version of Kolmogorov’s Theorem (Theorem 3.7.2) on $\mathbb{R}$:
\[ \text{span}\{e^{-2\pi it\gamma} : \gamma \leq 0\} = L^1_\mu(\mathbb{R}) \]
if and only if
\[ \int \frac{|\log w(t)|}{1 + t^2} \, dt = \infty. \]

d. Suppose $w \geq 1$ on $\mathbb{R}$. In particular, $w \notin L^1(\mathbb{R})$, whereas $w \in L^1(\mathbb{R})$ in part c. Assume $w$ is even, $1 = w(0) \leq w(t)$, and $w(u + t) \leq w(u)w(t)$ for $u, t \in \mathbb{R}$. $L^1_\mu(\mathbb{R}) \subseteq L^1(\mathbb{R})$ is an algebra under convolution, and Beurling (1938) posed the spectral analysis question: is every proper closed ideal $I \subseteq L^1_\mu(\mathbb{R})$ contained in a regular (i.e., $L^1_\mu(\mathbb{R})/I$ has a unit under convolution) maximal ideal? For an equivalent analytic means of posing this question, consider the following property of $I \equiv L^1_\mu(\mathbb{R})$:

\[ \forall \gamma \in \mathbb{R}, \exists f \in I \text{ such that } \hat{f}(\gamma) \neq 0. \tag{3.7.17} \]

Then the spectral analysis question is equivalent to finding conditions on $w$ so that, whenever a closed ideal $I \subseteq L^1_\mu(\mathbb{R})$ satisfies (3.7.17), we can conclude that $I = L^1_\mu(\mathbb{R})$.

We have posed the spectral analysis question in this section since there is an equivalent dual formulation in terms of sets of exponentials in $L^\infty_{1/\omega}(\mathbb{R})$, e.g., [Ben75].
3.8. DISCRETE FOURIER TRANSFORM

e. Beurling (1938) proved that if

\[ \int \frac{\log w(t)}{1 + t^2} \, dt < \infty \]  

(3.7.18)

then the spectral analysis question has an affirmative answer [Beu89].

Let \( w = 1 \) on \( \mathbb{R} \). Then (3.7.18) is satisfied and Beurling’s Theorem reduces to the Wiener Tauberian Theorem: \( f \in L^1(\mathbb{R}) \) has a nonvanishing Fourier transform if and only if the closed principal ideal \( I \subseteq L^1(\mathbb{R}) \) generated by \( f \) is all of \( L^1(\mathbb{R}) \), cf., Remark 2.9.13 and the formulation of Wiener’s Tauberian Theorem in Theorem 2.9.12.

3.8 Discrete Fourier Transform

In Section 1.1 we defined the Fourier transform \( \hat{f} \) of \( f : \mathbb{R} \rightarrow \mathbb{C} \); in this case \( \hat{f} \) is defined on \( \mathbb{R} \). In Section 3.1 we defined the Fourier transform \( \hat{f} \) of \( f : \mathbb{Z} \rightarrow \mathbb{C} \); in this case, and for a given \( \Omega > 0 \), \( \hat{f} \) is defined on \( \mathbb{T}_{2\Omega} \equiv \mathbb{R}/(2\Omega \mathbb{Z}) \), i.e., the Fourier series \( \hat{f} \) is a \( 2\Omega \)-periodic function on \( \mathbb{R} \) with Fourier coefficients \( f = \{f[n]\} \). The next step is the following definition.

3.8.1 Definition. DISCRETE FOURIER TRANSFORM

a. Let \( N \) be a positive integer, and let \( \mathbb{Z}_N \) be the set of integers \( 0, 1, \ldots, N - 1 \) under addition modulo \( N \). This means that if \( m, n \in \mathbb{Z}_N \) and the ordinary sum \( m + n \leq N - 1 \), then the addition modulo \( N \) of \( m \) and \( n \) has the value \( m + n \). However, if \( m, n \in \mathbb{Z}_N \) and the ordinary sum \( m + n > N - 1 \), then the addition modulo \( N \) of \( m \) and \( n \) has the value \( m + n - N \). For example, the addition table for \( \mathbb{Z}_6 \) is given in Figure 3.3.

When dealing with \( \mathbb{Z}_N \) we shall denote the addition modulo \( N \) of \( m, n \in \mathbb{Z}_N \) by \( m + n \in \mathbb{Z}_N \).

In order to define functions \( f \) on \( \mathbb{Z}_N \), we assign values \( f[n] \) for \( n = 0, 1, \ldots, N - 1 \), and then we extend \( f \) as an \( N \)-periodic function on \( \mathbb{Z} \). Thus, \( f[m + nN] = f[m] \) for any \( m \in \{0, 1, \ldots, N - 1\} \) and for all \( n \in \mathbb{Z} \). In this case, we write \( f : \mathbb{Z}_N \rightarrow \mathbb{C} \), cf., part e.
b. The Fourier transform of $f : \mathbb{Z}_N \rightarrow \mathbb{C}$ is the function $F : \mathbb{Z}_N \rightarrow \mathbb{C}$ defined as

\[(3.8.1) \quad \forall n \in \{0, 1, \ldots, N - 1\}, \quad F[n] = \sum_{m=0}^{N-1} f[m] e^{-2\pi i mn/N}.\]

<table>
<thead>
<tr>
<th>+</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

*Figure 3.3*

Because of the setting $\mathbb{Z}_N$, $F$ is called the *Discrete Fourier Transform* (DFT) of $f$. In this context we shall write

$$f \leftrightarrow F, \quad \widehat{f} = F, \quad f = \overline{F},$$

just as we did in *Definitions 1.1.2 and 3.1.1* for the cases of Fourier transforms (on $\mathbb{R}$) and Fourier series.

c. Letting $f : \mathbb{Z}_N \rightarrow \mathbb{C}$, and introducing the (standard) notation,

$$W_N = e^{-2\pi i/N},$$

the DFT $F$ of $f$ is defined as

\[(3.8.2) \quad \forall n \in \mathbb{Z}_N, \quad F[n] = \sum_{m \in \mathbb{Z}_N} f[m] W_N^{mn}.\]

Clearly, (3.8.1) and (3.8.2) are equivalent.

We could have defined (3.8.1) for each $n \in \mathbb{Z}$ by our definition in part a; in fact,

$$\forall n \in \mathbb{Z}, \quad F[n] = \sum_{m=0}^{N-1} f[m] e^{-2\pi i mn/N}$$

$$= \sum_{m=0}^{N-1} f[m] e^{-2\pi i (m+N)/N} = F[n+N].$$
3.8. DISCRETE FOURIER TRANSFORM

In any case, \( \mathbb{Z}_N \) is the natural domain for \( F \).

d. \( \mathbb{Z}_N \) is a commutative group under the operation of addition modulo \( N \).

A character \( \gamma \) of \( \mathbb{Z}_N \) is a homomorphism from \( \mathbb{Z}_N \) into the multiplicative group \( \{ z \in \mathbb{C} : |z| = 1 \} \), i.e.,

\[ \forall m \in \mathbb{Z}_N, \quad |\gamma(m)| = 1 \]

and

\[ \forall m, n \in \mathbb{Z}_N, \quad \gamma(m + n) = \gamma(m)\gamma(n). \]

The set of characters of \( \mathbb{Z}_N \) is denoted by \( \hat{\mathbb{Z}}_N \), and \( \hat{\mathbb{Z}}_N \) becomes a commutative group by defining the addition \( \gamma_1 + \gamma_2 \) of characters \( \gamma_1, \gamma_2 \) by means of the formula

\[ \forall m \in \mathbb{Z}_N, \quad (\gamma_1 + \gamma_2)(m) = \gamma_1(m)\gamma_2(m). \]

In this setting it can be proved that \( \hat{\mathbb{Z}}_N = \mathbb{Z}_N \), e.g., [Rud62], cf., Remark 3.1.3e.

Algebraic considerations are fundamental in many aspects of harmonic analysis; but, in this book, I am coming closest to cheating the reader by their omission in my treatment of the DFT and FFT.

e. Let \( L(\mathbb{Z}_N) \) denote the vector space of complex sequences on \( \mathbb{Z}_N \), and define the DFT mapping

\[ \mathcal{F}_N : L(\mathbb{Z}_N) \rightarrow L(\hat{\mathbb{Z}}_N), \]

\[ f \mapsto \hat{f}. \]

Since \( \mathbb{Z}_N \) is a finite set, \( L(\mathbb{Z}_N) \) can be considered any one of the \( L^p \)-spaces on \( \mathbb{Z}_N \). In fact, \( L(\mathbb{Z}_N) \) is the \( N \)-dimensional space of all functions on \( \mathbb{Z}_N \).

When one thinks of \( L(\mathbb{Z}_N) \) as \( L^2(\mathbb{Z}_N) \), we define the inner product

\[ \forall f, g \in L(\mathbb{Z}_N), \quad \langle f, g \rangle = \frac{1}{N} \sum_{m=0}^{N-1} f[m]g[m]. \]

f. In order to make the analogy between the DFT \( F \) of \( f \) on \( \mathbb{Z}_N \) (\( N \) even) and Fourier series on \( \mathbb{T} \) with Fourier coefficients on \( \mathbb{Z} \), we could
consider the finite sets,
\[ \{0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}\} \text{ and } \{-\frac{N}{2}, \frac{N}{2} + 1, \ldots, \frac{N}{2} - 1\}, \]
corresponding to approximations of T and Z, respectively.

The inversion theorem for the DFT is elementary.

### 3.8.2 Theorem. Inversion Formula for the DFT

Let \( N > 1 \) and let \( f : \mathbb{Z}_N \to \mathbb{C} \) have DFT \( F \). Then

\[
(3.8.3) \quad \forall m = 0, 1, \ldots, N - 1, \quad f[m] = \frac{1}{N} \sum_{n=0}^{N-1} F[n] e^{2\pi i mn/N}.
\]

**Proof.** Note that \( W_N^N = 1 \), and if \( N > 1 \) then \( W_N \neq 1 \). Thus, since

\[
(3.8.4) \quad 1 + r + r^2 + \cdots + r^{N-1} = \frac{1 - r^N}{1 - r}, \quad r \neq 0,
\]

we see that

\[
(3.8.5) \quad 1 + W_N + W_N^2 + \cdots + W_N^{N-1} = 0.
\]

For fixed \( m \), the right side of (3.8.3) is

\[
\frac{1}{N} \sum_{n=0}^{N-1} F[n] e^{2\pi i mn/N} = \frac{1}{N} \sum_{n=0}^{N-1} F[n] W_N^{-mn}
\]

\[
= \frac{1}{N} \sum_{n=0}^{N-1} \left( \sum_{j=0}^{N-1} f[j] W_N^{jn} \right) W_N^{-mn} = \frac{1}{N} \sum_{j=0}^{N-1} f[j] \left( \sum_{n=0}^{N-1} W_N^{(j-m)n} \right).
\]

If \( j = m \) then \( \sum_{n=0}^{N-1} W_N^{(j-m)n} = N \). If \( j \neq m \) and \( r \equiv W_N^{j-m} \), then \( r \neq 1 \) and (3.8.4) gives

\[
(3.8.7) \quad \sum_{n=0}^{N-1} W_N^{(j-m)n} = \frac{1}{1 - r} \left( 1 - e^{2\pi i (m-j)n} \right) = 0.
\]
3.8. DISCRETE FOURIER TRANSFORM

Substituting this information into the right side of (3.8.6) gives \( f[m] \).

\[ \square \]

The simplicity of (3.8.5) or (3.8.7) in Theorem 3.8.2 evaporates if one considers the Gauss sum,

\[ (3.8.8) \quad G_N^\pm = \sum_{n=0}^{N-1} e^{\pm 2\pi i n^2/N}, \]

e.g., Example 3.8.6, Theorem 3.8.7, Theorem 3.8.9, and Theorem 3.8.10.

3.8.3 Corollary.

\( \mathcal{F}_N : L(\mathbb{Z}_N) \rightarrow L(\mathbb{Z}_N) \) is a linear bijection.

3.8.4 Theorem. ONB AND PARSEVAL FORMULA

a. ONB. \( \{ \frac{1}{\sqrt{N}} W_N^n : n = 0, \ldots, N - 1 \} \) is an ONB for \( L(\mathbb{Z}_N) \) taken with the inner product defined in Definition 3.8.1e.

b. Parseval's Formula. Let \( f, g \in L(\mathbb{Z}_N) \) and consider the pairings \( f \leftrightarrow F, g \leftrightarrow G \). Then

\[ (3.8.9) \quad \sum_{m=0}^{N-1} f[m] \overline{g[m]} = \frac{1}{N} \sum_{n=0}^{N-1} F[n] \overline{G[n]}, \]

and, in particular,

\[ \left( \sum_{m=0}^{N-1} |f[m]|^2 \right)^{1/2} = \left( \frac{1}{N} \sum_{n=0}^{N-1} |F[n]|^2 \right)^{1/2}. \]

Proof. a. For a fixed \( k, n \in \mathbb{Z}_N \), we consider \( W_N^k \) and \( W_N^n \) as functions on \( \mathbb{Z}_N \) defined by \( W_N^k[m] \equiv W_N^{mn} \). Then

\[ (3.8.10) \quad \langle \frac{1}{\sqrt{N}} W_N^k, \frac{1}{\sqrt{N}} W_N^n \rangle = \frac{1}{N} \sum_{m=0}^{N-1} W_N^{mk} \overline{W_N^{mn}} = \frac{1}{N} \sum_{m=0}^{N-1} \left( e^{-2\pi i (k-n)/N} \right)^m. \]

The right side of (3.8.10) is 1 if \( k = n \). If \( k \neq n \) then \( r \equiv e^{-2\pi i (k-n)/N} \neq 1 \) since \( (k-n)/N \notin \mathbb{Z} \). Thus the right side of (3.8.10) is

\[ (1 - r^N)/(1 - r) = 0 \]
since $r^N = 1$. Therefore, \( \left\{ \frac{1}{N} W_N^n : n \in \mathbb{Z}_N \right\} \) is orthonormal.

Linear independence follows from orthonormality. In fact, if $k \in \mathbb{Z}_N$ is fixed and $\sum_{n=0}^{N-1} a_n W_N^n W_N^{-mk} = 0$, then

$$\forall m \in \mathbb{Z}, \quad \frac{1}{N} \sum_{n=0}^{N-1} a_n W_N^{mn} W_N^{-mk} = 0,$$

and so

$$\sum_{n=0}^{N-1} a_n \left( \frac{1}{N} \sum_{m=0}^{N-1} W_N^{mn} W_N^{-mk} \right) = 0.$$

By the orthonormality the left side is $a_k$.

The result follows since $L(\mathbb{Z}_N)$ is $N$-dimensional.

b. It is sufficient to prove (3.8.9); and (3.8.9) follows from Theorem 3.8.2, part a, and the fact that

$$\sum_{m=0}^{N-1} f[m] g[m] = \frac{1}{N} \sum_{n=0}^{N-1} F[n] \sum_{k=0}^{N-1} G[k] \left( \frac{1}{N} \sum_{m=0}^{N-1} W_N^{mn} W_N^{-mk} \right). \quad \square$$

3.8.5 Example. The DFT Matrix

a. The DFT $N \times N$-matrix $\mathcal{D}_N$ is defined as $\left( \frac{1}{\sqrt{N}} W_N^{mn} \right)$, $m, n = 0, \ldots, N - 1$, i.e.,

$$\mathcal{D}_N = \frac{1}{\sqrt{N}} \begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & e^{-2\pi i / N} & e^{-2\pi i 2/N} & \cdots & e^{-2\pi i (N-1)/N} \\
1 & e^{-2\pi i 2/N} & e^{-2\pi i 4/N} & \cdots & e^{-2\pi i 2(N-1)/N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & e^{-2\pi i (N-1)/N} & e^{-2\pi i 2(N-1)/N} & \cdots & e^{-2\pi i (N-1)(N-1)/N} 
\end{pmatrix}.$$

Thus,

$$\mathcal{D}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

If $f \in L(\mathbb{Z}_N)$ is considered as a $1 \times N$ column vector then the DFT $F$ of $f$ is the $1 \times N$ column vector

$$F = \sqrt{N} \mathcal{D}_N f.$$
3.8. DISCRETE FOURIER TRANSFORM

By Theorem 3.8.2 we have

\[ \mathcal{D}_N^{-1} = \frac{1}{\sqrt{N}} \overline{\mathcal{D}_N}, \]

where \( \overline{\mathcal{D}_N} \) denotes complex conjugation of the entries of \( \mathcal{D}_N \).

b. Note that the trace (Definition 3.6.2) of \( \mathcal{D}_N \) is

\[ \text{trace}(\mathcal{D}_N) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{-2\pi i n^2/N} = \frac{1}{\sqrt{N}} \mathcal{G}_N^2, \]

which we shall evaluate in Theorem 3.8.9 and Theorem 3.8.10.

c. Let \( U_N^n : \mathbb{Z}_N \rightarrow \mathbb{C} \), for fixed \( n \in \mathbb{Z}_N \), be defined by \( U_N^n[m] = \delta_{m,n} \)
for \( m \in \mathbb{Z}_N \). Clearly, \( \{ \sqrt{N} U_N^n \} \) is an ONB for \( L(\mathbb{Z}_N) \), with the inner product defined in Definition 3.8.1c.

It is easy to check that

\[(3.8.11) \quad \mathcal{D}_N U_N^n = \frac{1}{\sqrt{N}} W_N^n, \]

where \( U_N^n \) and \( W_N^n \) are considered as \( 1 \times N \) column vectors and the left side is matrix multiplication. Similarly, a direct calculation shows that

\[(3.8.12) \quad \mathcal{D}_N W_N^{-n} = \sqrt{N} U_N^n. \]

For example,

\[(3.8.13) \quad \mathcal{D}_5 W_5^{-2} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & e^{2\pi i 2/5} & e^{2\pi i 4/5} & e^{2\pi i 6/5} & e^{2\pi i 8/5} \\ 1 + e^{2\pi i 5} & 1 + e^{2\pi i 2/5} & 1 + e^{2\pi i 3/5} & 1 + e^{2\pi i 4/5} \\ 1 + e^{2\pi i 5} & 1 + e^{2\pi i 2/5} & 1 + e^{2\pi i 3/5} & 1 + e^{2\pi i 4/5} \\ 1 + e^{-2\pi i 5} & 1 + e^{-2\pi i 2/5} & 1 + e^{-2\pi i 3/5} & 1 + e^{-2\pi i 4/5} \\ 1 + e^{-2\pi i 5} & 1 + e^{-2\pi i 2/5} & 1 + e^{-2\pi i 3/5} & 1 + e^{-2\pi i 4/5} \end{pmatrix}. \]

The \( 1 \times 5 \) column matrix on the right side of (3.8.13) is \( \sqrt{5} U_5^2 \). In fact, \( \sum_{m=0}^{N-1} e^{-2\pi i m n/N} = 0 \) for a fixed \( n \in \mathbb{Z}_N \) by (3.8.4).
Combining (3.8.11) and (3.8.12) we see that

\[ \forall n \in \mathbb{Z}_N, \quad \mathcal{D}_N^n U_n^N = U_n^N. \]

Since \( \{\sqrt{N} U_n^N\} \) is an ONB for \( L(\mathbb{Z}_N) \) we conclude that

\[ (3.8.14) \quad \mathcal{D}_N^4 = I, \]

where \( I \) is the identity matrix, cf., Example 1.10.12 where we did the analogous calculation for the Fourier transform on \( \mathbb{R} \).

d. (3.8.14) leads naturally to investigating the eigenvalue problem for the DFT. In fact, because of (3.8.14), the eigenvalues of \( \mathcal{D}_N \) are \( \pm 1, \pm i \). The more difficult aspect of the eigenvalue problem is the multiplicity problem, viz., finding the eigenvectors of \( \mathcal{D}_N \) and the dimension of the space of eigenvectors for each of the eigenvalues.

In any case the complete eigenvalue problem was essentially solved by Gauss, fundamental related calculations were expounded in E. Landau's classic book, Vorlesungen über Zahlentheorie (Volume 1, 1927, pages 164–165), and the explicit solution was recorded in [Goo62, page 261]. Recent comprehensive contributions are due to McClellan and Parks [MP72] and Auslander and Tolimieri, e.g., [AT79]. Using a technique due to Schur, e.g., [BSh66, pages 349–353], it is shown in [AT79, page 856] that the solution of the multiplicity problem is equivalent to evaluating trace (\( \mathcal{D}_N \)) for all \( N \).

In the following material, from Example 3.8.6 to Theorem 3.8.10, we shall deal with the evaluation of the Gauss sum \( \mathcal{G}_N^\pm \). (\( \mathcal{G}_N^\pm \) was defined in (3.8.8)).

3.8.6 Example. \( \mathcal{G}_N^\pm = 0 \) if 4 divides \( N - 2 \)

a. If \( N - 2 = 4k, \ k \geq 0 \), then \( N = 2(2k + 1) \equiv 2M, \) where \( M \geq 1 \) is odd. Then

\[ (3.8.15) \quad \text{trace}(\mathcal{D}_N) = \sum_{m=0}^{M-1} e^{-2\pi im^2/(2M)} + \sum_{m=0}^{M-1} e^{-2\pi i(m+M)^2/(2M)}. \]
Expanding \((m + M)^2\) and using the fact that \(e^{-2\pi i M/2} = -1\) for \(M\) odd, we see that

\[
\sum_{m=0}^{M-1} e^{-2\pi i (m+M)^2/(2M)} = - \sum_{m=0}^{M-1} e^{-2\pi i m^2/(2M)} e^{-2\pi i m},
\]

and, hence, the right side of (3.8.15) is 0.

b. Note that \(G_3^+ = i\sqrt{3}\) and \(G_3^- = -i\sqrt{3}\). In fact,

\[
\sum_{m=0}^{2} e^{2\pi i m^2/3} = 1 + e^{2\pi i/3} + e^{2\pi i4/3} = 1 + 2e^{2\pi i/3}
\]

\[
= 1 + 2\cos \frac{2\pi}{3} + 2i \sin \frac{2\pi}{3} = i\sqrt{3}
\]

and

\[
\sum_{m=0}^{2} e^{-2\pi i m^2/3} = 1 + 2\cos \frac{2\pi}{3} - 2i \sin \frac{2\pi}{3} = -i\sqrt{3}.
\]

3.8.7 Theorem. \(|\sum_{m=0}^{N-1} e^{2\pi i m^2/N}| = \sqrt{N}, \ N \text{ odd}\)

Let \(N \geq 1\) be an odd integer. Then

\[
(3.8.16) \quad \left| \sum_{m=0}^{N-1} e^{\pm 2\pi i m^2/N} \right| = \sqrt{N}.
\]

Proof. Let \(g[m] = e^{-2\pi i m^2/N}, m = 0, 1, \ldots, N - 1,\) and let \(G = \hat{g}\). We shall prove (3.8.16) with the minus sign in the exponent. The plus sign case is a consequence of the same argument with the signs in the DFT properly adjusted. For \(n = 0, \ldots, N - 1,\) we compute

\[
G[2n] = \sum_{m=0}^{N-1} e^{-2\pi i m^2/N} e^{-2\pi i m(2n)/N}
\]

\[
= e^{2\pi i n^2/N} \sum_{m=0}^{N-1} e^{-2\pi i (m+n)^2/N}
\]

\[
= e^{2\pi i n^2/N} \sum_{m=0}^{N-1} e^{-2\pi i m^2/N} = e^{2\pi i n^2/N} G[0].
\]
CHAPTER 3. FOURIER SERIES

The second step follows since \((m+n)^2 - n^2 = m^2 + 2mn\); and the third step follows since \(W_N^{(m+N)^2} = W_N^{m^2}\) and by noting that the Gauss sum for either “\(m+n\)” or “\(m\)” has a domain of \(N\) consecutive integers. By (3.8.17) we have

\[
\sum_{n=0}^{N-1} |G[2n]|^2 = N|G[0]|^2.
\]

Since \(N\) is odd,

\[
\sum_{n=0}^{N-1} |G[n]|^2 = \sum_{n=0}^{N-1} |G[2n]|^2.
\]

For example, let \(N = 5\) so that \(G[2n], n = 0,1,2\), consists of \(G[0], G[2], G[4]\); and \(G[2\cdot 3] = G[1]\) and \(G[2\cdot 4] = G[3]\) since 5 divides \(6 - 1\) and 5 divides \(8 - 3\), respectively. The same phenomenon occurs for any odd \(N \geq 3\) by properties of \(\mathbb{Z}_N\).

Combining (3.8.18) and (3.8.19) with the Parseval Formula we obtain

\[
N|G[0]|^2 = N \sum_{m=0}^{N-1} |g[m]|^2 = N^2.
\]

(3.8.16) is obtained.

It is more difficult to evaluate \(G_N^\pm\) than \(|G_N^\pm|\). We shall first compute \(G_N^\pm\) for \(N\) odd in Theorem 3.8.9. Gauss gave the first proof of Theorem 3.8.9 in 1805 after working “with all efforts” for four years, e.g., [BE81, pages 109–110]. There have been many other derivations of Theorem 3.8.9, e.g., [BE81]. We shall give Daniel Shanks’ proof [Sha58].

Shanks originally devised ingenious finite term identities to prove deep infinite term identities of Euler and Gauss. The original idea for his proof’s goes back to his PhD thesis at the University of Maryland in 1954. The identity by which he obtained Gauss’ infinite term identity is Lemma 3.8.8, and in [Sha58] he used it to obtain Gauss’ Theorem (Theorem 3.8.9).
3.8. DISSERTE FOURIER TRANSFORM

3.8.8 Lemma. Shanks Finite Identity

If \( x > 0 \), \( P_0(x) = 1 \), and

\[
\forall N \in \mathbb{N}, \quad P_N(x) = \prod_{n=1}^{N} \left( \frac{1-x^{2n}}{1-x^{2n-1}} \right),
\]

then

\[
(3.8.20) \quad \sum_{n=1}^{2N} x^{n(n-1)/2} = \sum_{n=0}^{N-1} \frac{P_N(x)}{P_n(x)} x^n(2N+1).
\]

The proof of Lemma 3.8.8 begins with the identity

\[
(1-x^{2N})x^{n(2N+1)} = (1-x^{2N-1})x^{n(2N-1)}
\]

\[
+(1-x^{2n+1})x^{n+1(2N-1)} - (1-x^{2n})x^{n(2N-1)},
\]

which we multiply by \( P_{N-1}(x)/P_n(x)(1-x^{2N-1}) \).

3.8.9 Theorem. Gauss Computation of \( G_N^+ \), N odd

Let \( N \geq 1 \) be an integer. Then

\[
\frac{N-1}{4} \in \mathbb{N} \cup \{0\} \quad \text{implies} \quad G_N^+ = \sqrt{N}
\]

and

\[
\frac{N-3}{4} \in \mathbb{N} \cup \{0\} \quad \text{implies} \quad G_N^+ = i\sqrt{N}.
\]

(These two cases include all odd integers \( N \).)

Proof. We shall consider the case \( (N-1)/4 \in \mathbb{N} \cup \{0\} \). The other case is similar.

Let \( x = v^2 \) and \( v = e^{i\theta} \). Clearly,

\[
(3.8.21) \quad P_M(x) = v^M \prod_{n=1}^{M} \left( \frac{v^{2n} - v^{-2n}}{v^{2n-1} - v^{1-2n}} \right).
\]

Further, if \( Q_0 \equiv 1 \) and

\[
Q_M = \prod_{n=1}^{M} \left( \frac{\sin 2n\theta}{\sin(2n-1)\theta} \right),
\]
then

\[(3.8.22) \quad \sum_{n=1}^{2M} v^{n(n-1)} = \sum_{n=0}^{M-1} \frac{Q_M v^{M+n(4M+1)}}{Q_n}\]

and

\[(3.8.23) \quad \sum_{n=1}^{2M+1} v^{n(n-1)} = \sum_{n=0}^{M} \frac{Q_M v^{M+n(4M+1)}}{Q_n}\]

by Lemma 3.8.8. For example, in the case of (3.8.22), the left side is precisely the left side of (3.8.20); and so

\[(3.8.24) \quad \sum_{n=1}^{2M} v^{n(n-1)} = \sum_{n=0}^{M-1} \frac{P_M(x)}{P_n(x)} x^{n(2M+1)},\]

from which (3.8.22) is obtained by substituting (3.8.21) into the right side of (3.8.24).

Letting \(N = 2K + 1\) and \(\theta = 2\pi/N\), we have

\[v^{2K} = \exp i \left(2K \left(\frac{2\pi}{N}\right)\right) = \exp i \left(\left(2K + 1\right) \left(\frac{2\pi}{N}\right) - \left(\frac{2\pi}{N}\right)\right) = v^{-1}.\]

Thus,

\[(3.8.25) \quad G_N^+ = \sum_{n=0}^{N-1} e^{i2\pi n^2/N} = \sum_{n=0}^{N-1} e^{i\theta n^2} = \sum_{n=0}^{N-1} v^{n^2} = \sum_{n=0}^{N-1} \sum_{n=0}^{K} v^{n(K+n)^2} = v^{K^2} \sum_{n=0}^{K} v^{2nK+n^2} = v^{K^2} \sum_{n=0}^{K} v^{n(n-1)} = v^{K^2} \left[ \sum_{n=1}^{K} v^{n(n-1)} + \sum_{n=1}^{K+1} v^{n(n-1)} \right].\]
Finally, let \( N = 4M + 1 \) and \( K = 2M \). Combining (3.8.25) with (3.8.22) and (3.8.23), we obtain

\[
G_N^\pm = v^{K^2} \left[ \sum_{n=1}^{K} v^n(n-1) + \sum_{n=1}^{K+1} v^n(n-1) \right]
\]

\[
= v^{4M^2} \left[ \sum_{n=0}^{M-1} \frac{Q_M}{Q_n} v^{M+n(4M+1)} + \sum_{n=0}^{M} \frac{Q_M}{Q_n} v^{M+n(4M+1)} \right]
\]

\[
= \sum_{n=0}^{M-1} \frac{Q_M}{Q_n} v^{(M+n)(4M+1)} + \sum_{n=0}^{M} \frac{Q_M}{Q_n} v^{(M+n)(4M+1)}.
\]

Note that \( v^N = 1 \) since \( v^{2K} = v^{-1} \). Hence,

\[
(3.8.26) \quad G_N^\pm = 1 + 2 \sum_{n=0}^{M-1} \frac{Q_M}{Q_n}.
\]

Also, \( Q_0 = 1 \) and \( Q_N > 0 \) for \( n = 1, \ldots, M \) since \( \theta = 2\pi/N \). In fact, for such \( n, 0 < 2n\theta \leq 4M\pi/N < \pi \). Thus, \( G_N^\pm > 0 \) if \( \frac{N-1}{4} \in \mathbb{N} \cup \{0\} \). The result follows by (3.8.16). (\( 3.8.26 \) gives a cryptic way to write \( \sqrt{N} \) as a sum of products of quotients of sines!)

In 1835, Dirichlet used Fourier series to evaluate \( G_N^\pm \) for all \( N \). In this paper, Dirichlet also gave Gauss’ proof of the Law of Quadratic Reciprocity (Remark 3.8.11a) once he (Gauss) had Theorem 3.8.9, e.g., Dirichlet’s Werke, Chelsea Publishing Company, New York, pages 257–270. The following is Dirichlet’s computation of \( G_N^\pm \). It not only shows Dirichlet’s brilliance, but also the power of elementary harmonic analysis.

**3.8.10 Theorem. Dirichlet Computation of \( G_N^\pm \)**

\[
G_N^\pm =
\]

\[
\begin{cases}
(1 + i)\sqrt{N}, & \text{resp.,} \ (1 - i)\sqrt{N}, \quad \text{if} \ N/4 \in \mathbb{N}, \\
\sqrt{N}, & \quad \text{if} \ (N - 1)/4 \in \mathbb{N} \cup \{0\}, \\
0, & \quad \text{if} \ (N - 2)/4 \in \mathbb{N} \cup \{0\}, \\
i\sqrt{N}, & \quad \text{resp.,} \ -i\sqrt{N}, \quad \text{if} \ (N - 3)/4 \in \mathbb{N} \cup \{0\},
\end{cases}
\]
where the two values for two cases on the right side of (3.8.27) indicate \( G_N^+ \) and \( G_N^- \), respectively.

**Proof.** We shall evaluate \( G_N^- \). The calculation of \( G_N^+ \) is similar.

i. Let \( G(\gamma) = e^{-2\pi i \gamma^2/N} \) on \( \hat{\mathbb{R}} \). Then

\[
\sum_{k=0}^{N-1} (\tau_{-k} G(0) + \tau_{-k} G(1)) = (1 + e^{-2\pi i/N}) \\
+ (e^{-2\pi i/N} + e^{-2\pi i^2/N}) + (e^{-2\pi i^2/N} + e^{-2\pi i^3/N}) \\
+ \ldots + (e^{-2\pi i(N-1)^2/N} + e^{-2\pi iN^2/N}) = 2G_N^-.
\]

Let

\[
F(\gamma) = \tau_{-0} G(\gamma) + \tau_{-1} G(\gamma) + \ldots + \tau_{-(N-1)} G(\gamma)
\]

for \( \gamma \in \hat{\mathbb{R}} \). Then

\[
F(\gamma + 1) = F(\gamma) + e^{-2\pi i \gamma^2/N} (e^{-4\pi i \gamma} - 1)
\]

and so \( F(1) = F(0) \). At this point, we shall only consider \( F \) defined on \([0,1]\), and extend \( F \) as a 1-periodic function on \( \hat{\mathbb{R}} \).

By the smoothness of \( G \), and the fact that \( F \in C(\mathbb{T}) \), we can invoke Dirichlet's Theorem (Theorem 3.1.6) and Remark 3.1.1b to write

\[
G_N = \frac{F(1) + F(0)}{2} = F(0) = \sum f[m],
\]

where the sequence \( f = \{f[m]\} \) of Fourier coefficients of \( F \) is evaluated by

\[
f[m] = \int_0^1 \sum_{k=0}^{N-1} \tau_{-k} G(\gamma) e^{2\pi im \gamma} d\gamma
\]

\[
= \sum_{k=0}^{N-1} \int_k^{k+1} G(\lambda) e^{2\pi im \lambda} d\lambda = \int_0^N e^{-2\pi i \lambda^2/N} e^{2\pi im \lambda} d\lambda
\]

for each \( m \in \mathbb{Z} \). Completing the square, we obtain

\[
f[m] = e^{\pi im^2 N/2} \int_0^N e^{-2\pi i (\gamma - m N/2)^2/N} d\gamma,
\]
3.8. **DISCRETE FOURIER TRANSFORM**

and, hence,

\[(3.8.28) \quad G_N = \sum e^{\pi i m^2 N/2} \int_0^N e^{-2\pi i (\gamma - m N/2)^2/N} d\gamma.\]

If \(m\) is even then \(e^{\pi i m^2 N/2} = 1\) and if \(m\) is odd then \(e^{\pi i m^2 N/2} = e^{\pi i N/2} = i^N\). Separating the sum on the right side of (3.8.28) into even \((m = 2k)\) and odd \((m = 2k + 1)\) parts, and making the corresponding changes of variable, (3.8.28) becomes

\[G_N = I_e + I_o,\]

where

\[I_e = \sum \int_{-kN}^{(1-k)N} e^{-2\pi i \lambda^2/N} d\lambda = \int e^{-2\pi i \lambda^2/N} d\lambda\]

and

\[I_o = i^N \sum \int_{(-k-\frac{1}{2})N}^{(1-k-\frac{1}{2})N} e^{-2\pi i \lambda^2/N} d\lambda = i^N \int e^{-2\pi i \lambda^2/N} d\lambda.\]

Using Theorem 2.10.1, we see that

\[\int e^{-2\pi i \lambda^2/N} d\lambda = \frac{\sqrt{N}}{2} (1 - i).\]

Combining this information we have the formula,

\[(3.8.29) \quad G_N = \frac{\sqrt{N}}{2} (1 - i)(1 + i^N).\]

ii. (3.8.27) follows by letting \(N = 4k, 4k+1, 4k+2, 4k+3\), respectively, cf., Example 3.8.6a for the case \(N = 4k+2\) and Example 3.8.6b for the case \(N = 4k+3\). \(\square\)

3.8.11 **Remark. Gauss Sums: Potpourri and Titillation**

a. **Law of Quadratic Reciprocity.** The Law of Quadratic Reciprocity asserts that if \(p\) and \(q\) are distinct odd primes then

\[(3.8.30) \quad \left(\frac{p}{q}\right) \left(\frac{q}{p}\right) = e^{2\pi i (p-1)(q-1)/8},\]
where \((\frac{p}{q})\) is the Legendre Symbol. It is defined as follows. If \(r\) and \(m\) are relatively prime then \(r\) is the quadratic residue of \(m\) if

\[
\exists n \in \mathbb{N} \text{ such that } \frac{n^2 - r}{m} \in \mathbb{N} \cup \{0\},
\]

and it is a quadratic nonresidue if there is no such \(n\). Then \((\frac{p}{q}) = 1\) if \(p\) is a quadratic residue of \(q\), and \((\frac{p}{q}) = -1\) if it is a quadratic nonresidue of \(q\).

Equation (3.8.30) is important in Diophantine polynomial equations, e.g., [BE81], [HW65], as well as more advanced (and just as difficult!) number theoretic topics relates to the Riemann \(\zeta\)-function, more general Dirichlet \(L\)-functions, etc., e.g., [BSh66], [Cha68], [Cha70].

Equation (3.8.30) was first stated by Euler (1783). Legendre had an incomplete proof since he used a property, only later proved by Dirichlet, about primes in arithmetic progressions, e.g., Section 3.2.2. Independently, Gauss discovered (3.8.30), which he called the \(\text{"Theorema aureum"}\), in March 1795 before his 18th birthday on April 30, 1777; and gave the rigorous proof by April 1796. He went on to give seven other proofs. His third proof is in [HW65, Chapter VI]. His fourth proof, published in 1809, used Theorem 3.8.9. This means of obtaining the Law of Quadratic Reciprocity by means of Gauss sums has been refined by magisterial lineage: Schaar (1848), Kronecker (1880), Hecke (1919), C. L. Siegel (1966), e.g., [Cha68, pages 34–42].

b. Littlewood Flatness Problem. Let \(\mathcal{U}_N\) denote the class of unimodular trigonometric polynomials \(U(\gamma) = \sum_{n=0}^{N} u_n e^{-2\pi i n \gamma}\), i.e., \(|u_n| = 1\) for \(n = 0, 1, \ldots, N\). In light of a question asked in [Lit66] and expanded upon in [Lit68, Problem 19], we state the following Littlewood Flatness Problem: determine whether or not there are unimodular polynomials \(U_N \in \mathcal{U}_N\) and positive constants \(\epsilon_N > 0\) tending to zero, as \(N \to \infty\), such that

\[
(3.8.31) \quad (1 - \epsilon_N)\sqrt{N+1} \leq |U_N(\gamma)| \leq (1 + \epsilon_N)\sqrt{N+1}
\]

for all large \(N\). It turns out that Gauss sums and their variants play a natural role in dealing with (3.8.31).
3.8. DISCRETE FOURIER TRANSFORM

The inequalities (3.8.31) assert that there are trigonometric polynomials \( \frac{1}{\sqrt{N+1}} U_N \), whose modulus is almost identically 1 on \( T \) and whose coefficients have moduli which are all identically 1. Also, using the Parseval formula, they imply that

\[
\forall N, \exists U_N \in \mathcal{U}_N \text{ such that } \lim_{N \to \infty} \frac{\|U_N\|_{L^\infty(T)}}{\|U_N\|_{L^2(T)}} = 1.
\]

There have been herculean attempts to prove (3.8.31), sometimes in concert with subtle failures. Finally, in 1980, Kahane [Kah80] proved (3.8.31) by showing that such polynomials \( U_N \) exist, cf., [QS96]. His proof has a fundamental probabilistic component, and it remains to construct the \( U_N \). There are constructions in the more general case where it is only assumed that the moduli of the coefficients are bounded by 1, e.g., [BNe74], [Benk92].

c. Antenna Theory. The ratio \( \|F\|_{L^\infty(T)}/\|F\|_{L^2(T)} \) is called the crest factor of \( F \). This relationship between \( L^\infty \) ("maximum and minimum values") and \( L^2 \) ("energy") norms plays a role in a number of applications. Further, under certain constraints, trigonometric polynomials provide a natural model for arrays of energy transmitters and receptors in fields such as acoustics, electromagnetism, and seismology, e.g., [Sche60].

In particular, the space \( \mathcal{U}_N \) combined with (3.8.32) can be used in antenna array signal processing, e.g., [Steil76], [Hay85]. The space \( \mathcal{U}_N \) gives way to other models for other aspects of antenna theory. For example, transmitters and receptors can be placed at points \( x_1, \ldots, x_N \in \mathbb{R}^3 \); and space factors for the corresponding array outputs can be modeled by trigonometric polynomials \( P(\gamma) = \sum_{n=1}^N \exp 2\pi ix_n \cdot \gamma \). Since each of the coefficients is 1, it becomes relevant to check how close \( P \) is to the \( \delta \) measure, and to analyze the impact of the sidelobes of \( P \) outside of a neighborhood of the origin, e.g., [BHe93], cf., equation (UP) in Remark 1.1.4 and Example 2.4.8b, as well as the case of absolute values for the examples considered in part b and Theorem 2.10.9a.

d. Lagrange's Theorem and Surgery. The quadratic forms \( \sum r_{ij} e_i e_j \) we have dealt with have had complex entries; and we have established central results in harmonic analysis in the case these forms are nonnegative. It turns out that quadratic forms with integer or rational number
entries have been central to the development of number theory and algebra. The illustrious Dedekind, Frobenius, Minkowski, E. Noether, E. Artin, C. L. Siegel, Eichler, and Hasse are all major contributors.

Lagrange (1770) proved that every positive integer is the sum of four squares of integers. More generally, if \( \{r_{jk} : j, k = 1, \ldots, N\} \subseteq \mathbb{N} \), it is natural to ask which integers \( n \in \mathbb{N} \) are of the form \( n = \sum r_{jk} c_k c_j \) for some sequence \( c = \{c_j : 1, \ldots, N\} \subseteq \mathbb{N} \), and how many such "solutions" \( c \) there are, e.g., [HW65, Chapter 20]. The problem in this generality is unsolved. Many of the special cases that are known, as well as Minkowski’s classification of quadratic forms over \( \mathbb{Q} \), involve Gauss sums, e.g., [Scha85, Chapter 5]. These Gauss sums are of the form \( \sum e^{2\pi i q(x)} \) where \( q \) is a quadratic form associated with some algebraic structure. For example, if \( b \) is a bilinear function \( b : V \times V \to K \), where \( V \) is a vector space over a field \( K \), then \( q(x) = b(x, x) \) defines a quadratic form.

Surgery invariants can also be determined by evaluating Gauss sums of this type. The purpose of surgery in differential topology is to characterize simply connected manifolds in higher dimensions, and a basic Gauss sum was computed in this setting by J W. Morgan and D. P. Sullivan (Annals of Mathematics, 1974). The first use of Gauss sums in surgery problems was made by Edgar Brown.

### 3.9 Fast Fourier Transform

Let \( N > 1 \) and let \( f : \mathbb{Z}_N \to \mathbb{C} \) have DFT \( F \). As is apparent from Example 3.8.5a, \( \{F[n] : n \in \mathbb{Z}_N\} \) can be computed with \( N^2 \) operations, where an operation is defined to mean a complex multiplication followed by a complex addition. The Fast Fourier Transform (FFT) is an algorithm to compute a DFT by \( N \log_2 N \) operations in the case \( N = 2^r \), e.g., Example 3.9.2. The fundamental paper on the FFT is due to Cooley and Tukey [CT65], and it involves an idea which they refer to as a two step algorithm.

It turns out that two step Fourier analysis algorithms have been used in various applications since early in the 19th century. The first published results are due to Francesco Carlini (1828) in his research
on hourly barometric variations. This historical fact, as well as the fact that Gauss had a general form of the FFT as early as 1805, is found in the fascinating article by M. T. Heideman, D. H. Johnson, and C. S. Burrus [HJB84] on the history of the FFT, cf., [DV90], [IEEE96], [OS75, Chapter 6], [RR72], [Walk91] for other developments. Gauss’ results were published posthumously in 1866.

3.9.1 Theorem. Two Step FFT Algorithm

Let $N_1, N_2 > 1$ and let $f : \mathbb{Z}_N \to \mathbb{C}$ have DFT $F$, where $N = N_1 N_2$.

a. Each $n = 0, 1, \ldots, N - 1$ has the unique representation $n = n_2 N_1 + n_1$, for some $n_1 = 0, 1, \ldots, N_1 - 1$ and some $n_2 = 0, 1, \ldots, N_2 - 1$; and each $m = 0, 1, \ldots, N - 1$ has the unique representation $m = m_1 N_2 + m_2$, for some $m_1 = 0, 1, \ldots, N_1 - 1$ and some $m_2 = 0, 1, \ldots, N_2 - 1$.

b. In the format of part a, $F$ can be written as

\begin{equation}
F[n_2 N_1 + n_1] = \sum_{m_2 = 0}^{N_2 - 1} \left( \sum_{m_1 = 0}^{N_1 - 1} f[m_1 N_2 + m_2] W_N^m W_N^{m_1 n_1} W_N^{m_2 n_2} \right) W_N^{m_2 n_2}
\end{equation}

for each $n = n_2 N_1 + n_1$.

c. The total number of operations required to compute \{ $F[n] : n = n_2 N_1 + n_1$, $n_1 \in \mathbb{Z}_{N_1}$, and $n_2 \in \mathbb{Z}_{N_2}$ \} is

\[ N(N_1 + N_2). \]

Proof. Part a is immediate when, for the case of $n$, we think of the $N_2$ equispaced elements $0, N_1, 2 N_1, \ldots, (N_2 - 1) N_1$ of $\mathbb{Z}_N$. In fact, for each fixed $n_2 N_1$ we obtain all the elements of $\mathbb{Z}_N$ which are greater than or equal to $n_2 N_1$ and less than $(n_2 + 1) N_1$ by adding $n_1 = 0, 1, \ldots, N_1 - 1$ to $n_2 N_1$. (If $n_2 = N_2 - 1$ then $(n_2 + 1) N_1 = 0$.)

b. For a given $n$, choose $n_1, n_2$ for which $n = n_2 N_1 + n_1$. Then

\begin{equation}
F[n_2 N_1 + n_1] = \sum_{m_1, m_2} f[m_1 N_2 + m_2] W_N^{m_1 N_2 n_1} W_N^{m_2 n_2}.
\end{equation}

We calculate

\[ e^{-2\pi i m_1 N_2 n/N} = e^{-2\pi i m_1 n/N_1} = e^{-2\pi i m_1 (n_2 N_1 + n_1)/N_1} \]
and
\[ e^{-2\pi i m_2 (n_2 N_1 + n_1) / N} = e^{-2\pi i m_2 n_2 / N_2} e^{-2\pi i m_1 n_1 / N}, \]
and so the right side of (3.9.2) is
\[ \sum_{m_1, m_2} f[m_1 N_2 + m_2] W_{N_1}^{m_1 n_1} W_N^{m_2 n_1} W_{N_2}^{m_3 n_2}. \]
Writing out the domains of \( m_1 \) and \( m_2 \), this is precisely (3.9.1), which we shall also write as
\[ F[n] = \sum_{m_2=0}^{N_2-1} F[m_2, m_1] W_{N_2}^{m_2 n_2}, \]
where
\[ F[m_2, m_1] = \sum_{m_1=0}^{N_1-1} f[m_1 N_2 + m_2] W_{N_1}^{m_1 n_1} W_N^{m_2 n_1}. \]

(c) For fixed \( n_1, n_2, \) and \( m_2 \), we can compute
\[ (3.9.3) \sum_{m_1=0}^{N_1-1} f[m_1 N_2 + m_2] W_{N_1}^{m_1 n_1} \]
by \( N_1 - 1 \) operations, where the first of these operations yields
\[ f[0 \cdot N_2 + m_2] + f[1 \cdot N_2 + m_2] W_{N_1}^{1 n_1}, \]
and the \((N_1 - 1)st\) operation yields the complete sum (3.9.3). Multiplying this sum by \( W_N^{m_2 n_1} \) we obtain \( F[m_2, n_1] \) in \((N_1 - 1) + 1 = N_1\) operations.

Now, using the data \( \{F[m_2, n_1]\} \) which required \( N_1 \) operations to compute, we can compute each
\[ F[n] = \sum_{m_2=0}^{N_2-1} F[m_2, n_1] W_{N_2}^{m_2 n_2} \]
by an additional \( N_2 \) operations. Thus, for each pair \((n_1, n_2)\), \( F[n] = F[n_2 N_1 + n_1] \) can be computed by means of (3.9.1) in \( N_1 + N_2 \) operations. Since there are \( N = N_1 N_2 \) pairs \((n_1, n_2)\), all the data \( \{F[n] : n = 0, \ldots , N - 1\} \) can be computed by means of (3.9.1) in \( N(N_1 + N_2) \) operations. □
3.9. Fast Fourier Transform

3.9.2 Example. Order $N \log_2 N$ Algorithm

a. Let $N = 2N_1$. Then each $n = 0, 1, \ldots, N - 1$ can be written as $n = n_2N_1 + n_1$ for $n_1 = 0, 1, \ldots, N_1 - 1$ and $n_2 = 0, 1$. Hence, if $F: \mathbb{Z}_N \rightarrow \mathbb{C}$, defined as

$$F[n] = \sum_{m=0}^{N-1} f[m]W_N^{mn},$$

is the DFT of $f: \mathbb{Z}_N \rightarrow \mathbb{C}$, then

$$F[n] = F[n_2N_1 + n_1] = \sum_{m_1, m_2} f[2m_1 + m_2]W_{2N_1}^{(2m_1+m_2)n},$$

where $m_1 = 0, 1, \ldots, N_1 - 1$ and $m_2 = 0, 1$. We compute

$$F[n_2N_1 + n_1] = \sum_{m_2=0}^{N_1-1} \sum_{m_1=0}^{N_1-1} f[2m_1 + m_2]W_{2N_1}^{(2m_1+m_2)n}$$

$$= \sum_{m_2=0}^{N_1-1} \sum_{m_1=0}^{N_1-1} f[2m_1 + m_2]W_{N_1}^{m_1n_1}W_{N_1}^{m_2n_2} e^{-\pi i m_2 n_1},$$

since

$$e^{-2\pi i 2m_1 n/(2N_1)} = e^{-2\pi i m_1 (n_2N_1 + n_1)/N_1} = W_{N_1}^{m_1n_1}$$

and

$$e^{-2\pi i m_2 n/(2N_1)} = e^{-2\pi i m_2 (n_2N_1 + n_1)/ (2N_1)}$$

$$= e^{-\pi i m_2 n_2} e^{-2\pi i m_2 n_1/(2N_1)} = e^{-\pi i m_2 n_2} W_{N_1}^{m_2n_1}.$$

Therefore,

$$F[n_2N_1 + n_1] = \sum_{m_1=0}^{N_1-1} f[2m_1]W_{N_1}^{m_1n_1} + e^{-\pi i n_2}W_{N_1}^{n_1} \sum_{m_1=1}^{N_1-1} f[2m_1 + 1]W_{N_1}^{m_1n_1}.$$

Consequently, because

$$e^{-\pi i n_2}W_{N_1}^{n_1} = e^{-2\pi i n_2 N_1/(2N_1)}W_{N_1}^{n_1} = W_{N_1}^{n_1},$$

we obtain the FFT algorithm,

(3.9.4)

$$F[n_2N_1 + n_1] = \sum_{m_1=0}^{N_1-1} f[2m_1]W_{N_1}^{m_1n_1} + W_{N_1}^{n_1} \sum_{m_1=1}^{N_1-1} f[2m_1 + 1]W_{N_1}^{m_1n_1}.$$
for each $n = n_2 N_1 + n_1$.

b. Let $\#(K)$ be the number of multiplications required to compute the DFT of $f : \mathbb{Z}_K \to \mathbb{C}$. Clearly, $\#(K) \leq (K - 1)^2$.

In dealing with the right side of (3.9.4), we see that $2\#(N_1)$ multiplications are required to compute

$$s_e(n_1) = \sum_{m_1=0}^{N_1-1} f[2m_1] W_{N_1}^{m_1 n_1} \quad \text{and} \quad s_0(n_1) = \sum_{m_1=0}^{N_1-1} f[2m_1 + 1] W_{N_1}^{m_1 n_1}$$

for all $n_1 \in \mathbb{Z}_{N_1}$.

Now note that $W_N^n = \pm W_{N_1}^n$ depending on whether $n_2 = 0, 1$ in the representation $n = n_2 N_1 + n_1$. Thus, $N_1$ multiplications are required to compute $\{W_{N_1}^n, s_0(n_1) : n_1 \in \mathbb{Z}_{N_1}\}$ for given data $\{W_{N_1}^n, s_0(n_1) : n \in \mathbb{Z}_{N_1}\}$. Consequently, equation (3.9.4) allows us to write

(3.9.5) \hspace{1cm} \#(2N_1) = 2\#(N_1) + N_1.

(c. Because of (3.9.5) and the fact that $\#(2) = 1$ (since the DFT of $\{f[0], f[1]\}$ is $\{f[0] + f[1], f[0] + f[1]W_2^1\}$), we have

$$\#(2) = 1$$
$$\#(4) = 2 + 2 = 4 = \frac{1}{2}4 \log_2 4$$
$$\#(8) = 8 + 4 = 12 = \frac{1}{2}8 \log_2 8$$

\[\cdots\]

In fact, if $N = 2^r$ then the number of multiplications required to compute the DFT on $\mathbb{Z}_N$ is

(3.9.6) \hspace{1cm} \frac{N}{2} \log_2 N.

To prove (3.9.6) we proceed by induction. We just checked the result for $N = 2, 4, 8$. Now assume (3.9.6) is true for $N = N_1 = 2^r$, i.e., make the induction hypothesis that $\#(N_1) = \frac{N_1}{2} \log_2 N_1$. Letting $N = 2N_1$, (3.9.5) and the induction hypothesis imply

$$\#(N) = 2\#(N_1) + N_1 = N_1 \log_2 N_1 + N_1$$

$$= \frac{N}{2} (\log_2 N_1 + 1) = \frac{N}{2} \log_2 N,$$
and the proof is complete.

d. The FFT algorithm, which achieves (3.9.6), is defined by means of (3.9.4).

Equation (3.9.4), or the more general (3.9.1), can be viewed in terms of replacing the DFT $N \times N$-matrix $D_N$, having no 0-entries, in terms of a collection of matrices with many 0-entries, cf., Example 3.9.3 and Carleson’s remark at the end of Section 3.2.8.

In this spirit, C. Rader (1968) introduced an important idea, based on the fact that $\mathbb{Z}_p$ is a field, which ultimately led to Winograd’s celebrated algorithms (1978) for computing the DFT on $\mathbb{Z}_p$, e.g., [RR72], [DV90].

3.9.3 Example. Sparse FFT Matrices

We shall quantify the remark in Example 3.9.2d about matrices with 0 entries.

The right side of (3.9.4) splits the domain of $f : \mathbb{Z}_N \rightarrow \mathbb{C}$, $N = 2N_1$, into its even and odd parts. This can be accomplished by a matrix operation. For example, if $N = 8$ and

$$C_8 = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}$$

then

(3.9.7) \[ C_8 f^T = (f[0], f[2], f[4], f[6], f[1], f[3], f[5], f[7])^T. \]

The rule for constructing $C_N$ is obvious.

The next operation in (3.9.4) tells us that $\sqrt{N}D_N$ is related to two copies of $\sqrt{N_1}D_{N_1}$ embedded into an $N \times N$ array. These two copies address the even and odd parts of the domain of $f$, resp. Thus, we
introduce the \( N \times N \) matrix

\[
B_N = \begin{pmatrix}
\sqrt{N_1} D_{N_1} \\
\sqrt{N} D_N
\end{pmatrix}
\]

where the first and third quadrants are each \( N_1 \times N_1 \) 0-matrices. For example, if \( N = 8 \) then

\[
B_8 C_8 f^T = (s_e(0), s_e(1), s_e(2), s_e(3), s_o(0), s_o(1), s_o(2), s_o(3))^T,
\]

where the "even" and "odd" sums \( s_e \) and \( s_o \) are the DFTs defined in Example 3.9.2b for the case \( N_1 = 4 \).

Finally, we have to introduce a matrix which incorporates the factor \( W_N^n \) of \( s_e(n_1) \) (when \( n = n_2 N_1 + n_1 \)) and adds it to \( s_o(n_1) \). Recall from Example 3.9.2b that \( W_N^n = W_N^{n_1} \) if \( n_2 = 0 \) and \( W_N^n = -W_N^{n_1} \) if \( n_2 = 1 \). Therefore, if \( N = 8 \) we set

\[
A_8 = \begin{pmatrix}
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & W_8^1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & W_8^2 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & W_8^3 \\
1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & -W_8^1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & -W_8^2 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & -W_8^3
\end{pmatrix}
\]

and obtain

\[
A_8 B_8 C_8 f^T = F^T.
\]

The rule for constructing \( A_N \) is obvious, and so

\[
A_N B_N C_N = \sqrt{N} D_N
\]

(as easy as ...).
3.9. FAST FOURIER TRANSFORM

3.9.4 Remark. Butterflies and Bit Reversal

a. Let \( N = 2^r \) and let \( f : \mathbb{Z}_N \rightarrow \mathbb{C} \) have DFT \( F \). Besides the matrix formulation of the FFT algorithm (3.9.4) in Example 3.9.3, we can also formulate it in terms of diagrams whose basic components are called butterflies.

To define a butterfly, we begin by rewriting (3.9.4) as

\[
F[k] = F_0[k] + W_N^k F_1[k], \quad k = 0, 1, \ldots, (1/2)N - 1
\]

and

\[
F[k + (1/2)N] = F_0[k] - W_N^k F_1[k], \quad k = 0, 1, \ldots, (1/2)N - 1,
\]

where

\[
F_0[k] = \sum_{m=0}^{N/2 - 1} f[2m] W_N^{mk} \quad \text{and} \quad F_1[k] = \sum_{m=0}^{N/2 - 1} f[2m + 1] W_N^{mk},
\]

and where we have again used the fact that \( W_N^{r(N/2)+k} = \pm W_N^k \) depending of whether \( r_2 = 0, 1 \).

The DFT \( F[k], k = 0, 1, \ldots, N - 1 \), written in terms of the calculations (3.9.8)–(3.9.10), can be visualized as

\[
\begin{align*}
F_0[k] &\rightarrow F_0[k] + W_N^k F_1[k] \\
F_1[k] &\rightarrow F_0[k] - W_N^k F_1[k],
\end{align*}
\]

where \( k = 0, 1, \ldots, (1/2)N - 1 \); and this diagram is called a butterfly.

b. The butterfly (3.9.11) can be viewed as a construction of the DFT \( F : \mathbb{Z}_N \rightarrow \mathbb{C} \) in terms of the two DFTs, \( F_0 \) and \( F_1 \), on \( \mathbb{Z}_{N/2} \).

In the same way, \( F_0 \) and \( F_1 \) can each be constructed in terms of pairs of DFTs on \( \mathbb{Z}_{N/4} \). For example,

\[
F_0[k] = F_{00}[k] + W_{N/2}^k F_{01}[k]
\]

and

\[
F_0[k + (1/4)N] = F_{00}[k] - W_{N/2}^k F_{01}[k],
\]
for \( k = 0, 1, \ldots, (1/4)N - 1 \), where \( F_{00} \) is the DFT of
\[
\{ f[0], f[4], f[8], \ldots, f[N - 4] \}
\]
and \( F_{01} \) is the DFT of \( \{ f[2], f[6], f[10], \ldots, f[N - 2] \} \).

Since \( N = 2^r \) this procedure can be reduced to the consideration of
\( 2^{r-1} \) DFTs on \( \mathbb{Z}_2 \). Each of these is the DFT of a pair \( \{ f[j], f[k] \} \), and
the pairs are mutually disjoint. Clearly, the stepwise evolution of \( F \)
from these \( 2^{r-1} \) DFTs on \( \mathbb{Z}_2 \) can be pictured and understood in terms
of butterflies, e.g., [OS75, Chapter 6], [Walk91]. Computationally, it is
convenient to compute the 2-point DFTs first, then the 4-point DFTs,
etc.

c. Suppose \( f : \mathbb{Z}_N \to \mathbb{C} \) is given and the computation of its DFT
\( F \) is desired in the natural ordering \( \{ F[0], F[1], \ldots, F[N - 1] \} \), i.e., for
a given \( f \), a computational device will compute an ordered \( N \)-tuple
\( (F_0, \ldots, F_{N-1}) \), and we want to be sure that \( F_k = F[k] \) for each \( k \).

From (3.9.10) it is clear that if we begin with the DFTs of the pairs
\( \{ f[0], f[1] \}, \{ f[2], f[3] \}, \) etc., we shall not obtain \( F \) in its natural
ordering. It turns out that \( f \) can be ordered in such a way so that the
DFTs of consecutive pairs \( \{ f[j], f[k] \} \) in this ordering yield the desired
natural ordering of \( F \). The procedure is called bit reversal. For example, if \( N = 8 \) then the DFTs of \( \{ f[0], f[4] \}, \{ f[2], f[6] \}, \{ f[1], f[5] \},
\{ f[3], f[7] \} \) will yield the ordered \( N \)-tuple \( \{ F[0], F[1], \ldots, F[N - 1] \} \)
when the halving procedure of part b is implemented.

Bit reversal is defined as follows for \( N = 2^r \). At level \( r = 1 \) the
bit reversal ordering of the set \{0, 1\} is the ordered 2\(^1\)-tuple \( (0, 1) \).
At level \( r = 2 \) the bit reversal ordering of the set \{0, 1, 2, 3\} is the
ordered 2\(^2\)-tuple \( (0, 2, 1, 3) \). Inductively, at level \( m \) suppose the set
\{0, 1, \ldots, 2^m - 1\} has as its bit reversal ordering the ordered \( M = 2^m \)-
tuple,
\[
(b_0, \ldots, b_{M-1}).
\]
Then, by definition, at level \( m + 1 \), the bit reversal ordering of the set
\{0, 1, \ldots, 2M - 1\} is
\[
(2b_0, 2b_1, \ldots, 2b_{M-1}, 2b_0 + 1, 2b_1 + 1, \ldots, 2b_{M-1} + 1).
\]
3.9. FAST FOURIER TRANSFORM

For example the bit reversal orderings at levels 3 and 4 are

\[(0, 4, 2, 6, 1, 5, 3, 7)\]

and

\[(0, 8, 4, 12, 2, 10, 6, 14, 1, 9, 5, 13, 3, 11, 7, 15),\]

respectively. The term “bit reversal” is appropriate since the coefficients of the binary expansions (part d) of integers are reversed at the critical step in the above process. This is the reason for the subscripts 0 and 1 in part b.

d. The binary expansion of \(n \in \{0, 1, \ldots, 2^r\}\) is

\[n = \sum_{j=1}^{r} \epsilon_j 2^{j-1},\]

where \(\epsilon_j \in \{0, 1\}\). For example, if \(\epsilon_0 = \epsilon_1 = \ldots = \epsilon_r = 0\) then \(n = 0\), and if \(\epsilon_0 = \epsilon_1 = \ldots = \epsilon_{r-1} = 1\) then \(n = \sum_{j=1}^{r} 2^j = (2^r - 1)/(2-1) = 2^r - 1\). Thus, each such \(n\) is well-defined by an \(r\)-array \((\epsilon_1, \ldots, \epsilon_r)\) of 0s and 1s.

e. Suppose \(\{X^n_r\}\) is a “tree of spaces” where \(r\) designates the level, and where, for each fixed \(r \geq 0\) there are \(N = 2^r\) elements \(X^n_r\), indexed by \(n\). Using the binary expansion of part d we write

\[X^n_r = X^n_{(\epsilon_1, \ldots, \epsilon_r)};\]

and, using the bit reversal ordering, the tree \(\{X^n_r\}\) has the form

\[
\begin{array}{ccc}
X^0_0 & & \\
X^1_0 & X^1_1 & \\
X^2_{(0,0)} & X^2_{(0,2)} & X^2_{(1,0)} & X^2_{(1,1)} & \\
\end{array}
\]

etc. At level \(r - 1\) the space \(X^{r-1}_{(\epsilon_1, \ldots, \epsilon_{m-1}, 0)}\) is the (single) parent of

\[X^r_{(\epsilon_1, \ldots, \epsilon_{m-1}, 0)} \quad \text{and} \quad X^r_{(\epsilon_1, \ldots, \epsilon_{m-1}, 1)}.\]
CHAPTER 3. FOURIER SERIES

Tree models, including some where bit reversal ordering is essential, abound in mathematics and engineering. There are applications in image compression, channel crosstalk reduction, Walsh functions and the waveletpackets of Coifman, Meyer, and Wickerhauser, subband coding and the theory of nonlinear waveletpackets, frequency localization, multirate systems and bit allocation, C. Fefferman's proof of Carleson's Theorem proving the Lusin Conjecture, etc., e.g., [BA83], [BF94, Chapter 10], [BSa94], [Dau92], [Fef73], [Mey90].

f. With regard to the algorithmic butterflies of part a, recall the locally lacunary butterfly from Example 1.4.4 in Chapter 1. Even Chapter 2 has a lepidopteral connection: Laurent Schwartz is not only a world class mathematician but has seven butterflies named after him!

The extension of Carleson's Theorem (Section 3.2.8) to 2-dimensions is true for partial sums taken over squares \(\{(m, n): -N \leq m, n \leq N\}\). C. Fefferman's proof uses butterfly-shaped subsets of \(\mathbb{Z} \times \mathbb{Z}\) in a fundamental way [Fef71], cf., the proofs by Sjölin and Tevzadze at about the same time.

3.9.5 Example. Computation of \(W_N, N = 2^r\)

Let \(C_r = \cos(2\pi/2^r)\) and \(S_r = \sin(2\pi/2^r)\). If \(r = 1\) then \(C_1 = -1\) and \(S_1 = 0\). Using the half-angle formulas,

\[
\cos \frac{x}{2} = \sqrt{\frac{1 + \cos x}{2}} \quad \text{and} \quad \sin \frac{x}{2} = \sqrt{\frac{1 - \cos x}{2}},
\]

we can compute \(C_r\) and \(S_r\) for large values of \(r\) by the recurrence equations,

\[
(3.9.12) \quad C_{r+1} = \sqrt{\frac{1 + C_r}{2}} \quad \text{and} \quad S_{r+1} = \sqrt{\frac{1 - C_r}{2}}.
\]

The problem with (3.9.12) from a computational point of view is the possibility that \(\lim_{r \to \infty} (1 - C_r) = 0\) might lead to computational instability. On the other hand, since \(\sin^2 x + \cos^2 x = 1\), the double angle formulas can be written as

\[
\cos x = \sqrt{\frac{1 + \cos 2x}{2}} \quad \text{and} \quad \sin x = \frac{\sin 2x}{2 \cos x} = \frac{\sin 2x}{2 \sqrt{1 + \cos 2x}};
\]
d. $F(\gamma) = \gamma$, $\gamma \in [-\Omega, \Omega)$.

e. $F(\gamma) = \gamma$, $\gamma \in [0, 2\Omega)$.

f. $F(\gamma) = \gamma^2$, $\gamma \in [-\Omega, \Omega)$.

g. $F(\gamma) = \gamma^2$, $\gamma \in [-2\Omega, 0)$.

h. $F(\gamma) = \begin{cases} c \neq 0, & \gamma \in (0, \Omega), \\ 0, & \gamma \in [\Omega, 2\Omega] \end{cases}$.

i. $F(\gamma) = \gamma \cos \left( \frac{\pi \gamma}{\Omega} \right)$, $\gamma \in [-\Omega, \Omega)$.

j. $F(\gamma) = \gamma \sin \left( \frac{\pi \gamma}{\Omega} \right)$, $\gamma \in [-\Omega, \Omega)$.

k. $F(\gamma) = |\sin \left( \frac{\pi \gamma}{\Omega} \right)|$, $\gamma \in [0, 2\Omega)$.

l. $F'(\gamma) = \begin{cases} \cos \left( \frac{\pi \gamma}{\Omega} \right), & \gamma \in (0, \Omega), \\ 0, & \gamma \in [\Omega, 2\Omega] \end{cases}$.

m. $F'(\gamma) = \begin{cases} \sin \left( \frac{\pi \gamma}{\Omega} \right), & \gamma \in (0, \Omega), \\ 0, & \gamma \in [\Omega, 2\Omega] \end{cases}$.

3.2. Using MATLAB, graph the $N$th partial sums, for $N = 1, 2, 4, 8$, of the Fourier series of the functions defined in parts a, b, c, f, j, k of Exercise 3.1a.

3.3. Designate $F$ in Exercise 3.1a by $F_a$, and similarly for parts b, c, $\cdots$, m.

a. Show that, formally,

$$S(F_b)' = S(F_a) \quad \text{and} \quad S(F_f)' = 2S(F_d),$$

where $S(F)'$ is the term by term differentiation of $S(F)$.
(This calculation is legitimate when we consider $S(F)'$ as the distributional derivative of the function $F$ defined 2$\Omega$-periodically on $\mathbb{R}$, cf., Definition 3.1.9.)

b. Evaluate the Fourier coefficients of $F_m - \frac{1}{2}F_k$. 
3.4. Using (3.1.7), prove that the function $G$ defined in part $b$ of the proof of Theorem 3.1.6 is bounded in some interval centered at the origin.

3.5. Compute the Fourier series of

$$F(\gamma) = \frac{\pi - \gamma}{2}, \quad \gamma \in [-\pi, \pi),$$

defined $2\pi$-periodically on $\hat{\mathbb{R}}$, cf., Example 3.3.6a and Exercise 3.29.

3.6. Prove that the inclusion, $L^2(\mathbb{T}_{2\pi}) \subseteq L^1(\mathbb{T}_{2\pi})$, is proper.

3.7. a. Compute the Fourier series of the following functions defined $2\pi$-periodically.

   i. $F(\gamma) = \sin \gamma, \quad \gamma \in [0, 2\pi)$.

   ii. $F(\gamma) = \sin \gamma, \quad \gamma \in [0, \pi)$, and extended evenly to $(-\pi, \pi)$.

   b. Compute the Fourier series of $F(\gamma) = \sin \gamma, \quad \gamma \in [0, \pi)$, considered as a $\pi$-periodic function. Compare this result with part $a$.

3.8. Compute

$$\frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \left| \sum_{n=-100}^{100} e^{i n \gamma / \Omega} \right|^2 d\gamma.$$

3.9. Prove (3.4.17), which states the continuity of the inner product.

3.10. Prove the divergence of the series

$$\sum_{|n| \geq 2} \frac{1}{|n| \log |n|}.$$  

This fact played a role in Example 3.3.4.

3.11. a. Prove that

$$\sum_{|n| \leq N} e^{inx} = \frac{\sin(N + \frac{1}{2})x}{\sin \frac{x}{2}}$$

for $x \notin 2\pi \mathbb{Z}$. For $x = -\pi \gamma / \Omega$, this establishes (3.4.5).
b. Prove that
\[ \sum_{|n| \leq N} \left( 1 - \frac{|n|}{N + 1} \right) e^{inx} = \frac{1}{N + 1} \left( \frac{\sin(N + 1) \pi}{\sin \frac{\pi}{2}} \right)^2 \]
for \( x \notin 2\pi \mathbb{Z} \), and complete the proof of (3.4.7).

3.12. a. Let \( P(\gamma) = \sum_{|n| \leq N} c_n e^{-\pi in\gamma/\Omega} \) be a trigonometric polynomial and let \( F \in L^1(\mathbb{T}_2) \). Compute \( F \ast P \). Is \( F \ast P \) a trigonometric polynomial?

b. Let \( F \in L^1(\mathbb{T}_2) \) and \( G \in L^\infty(\mathbb{T}_2) \). Prove that
\[ \|F \ast G\|_{L^\infty(\mathbb{T}_2)} \leq \|F\|_{L^1(\mathbb{T}_2)} \|G\|_{L^\infty(\mathbb{T}_2)}. \]

c. Let \( F \in L^2(\mathbb{T}_2) \) and \( G \in L^1(\mathbb{T}_2) \). Prove that
\[ \|F \ast G\|_{L^2(\mathbb{T}_2)} \leq \|F\|_{L^2(\mathbb{T}_2)} \|G\|_{L^1(\mathbb{T}_2)}. \]

3.13. a. Abel's partial summation formula is
\[
(E3.1) \quad \sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p,
\]
where \( \{a_n, b_n : n = 0, 1, \ldots \} \subseteq \mathbb{C} \), \( 0 \leq p \leq q \), \( A_{-1} \equiv 0 \), and \( A_n \equiv \sum_{j=0}^{n} a_j \). Prove (E3.1).

b. Prove that if \( p, q \in \mathbb{Z}, p < q \), and \( x \neq 2\pi k \) for any \( k \in \mathbb{Z} \), then
\[ \left| \sum_{n=p}^{q} e^{inx} \right| \leq \frac{1}{|\sin \frac{\pi}{2}|}; \]
and so
\[ \forall \gamma \neq 2\Omega k, \quad \left| \sum_{n=p}^{q} e^{-\pi in\gamma/\Omega} \right| \leq \frac{1}{|\sin \frac{\pi}{2}|}. \]

c. Let \( f[p] \geq f[p + 1] \geq \ldots \geq f[q] \geq 0 \). Prove that
\[ \forall \gamma \neq 2\Omega k, \quad \left| \sum_{n=p}^{q} f[n] e^{-i\pi n\gamma/\Omega} \right| \leq \frac{f[p]}{|\sin \frac{\pi}{2}|}. \]

[Hint. Use the partial summation formula (E3.1).]
3.14. Let \( \Omega > 0 \) and let \( \alpha \in (0,\Omega) \). Let \( F = 1 \) on \( T_{2\Omega} \) and define \( F_\alpha \) on \( T_{2\Omega} \) by setting \( F_\alpha = 1_{[-\alpha,\alpha]} \) on \([-\Omega,\Omega)\).

a. Compute the Fourier coefficients of \( F \) and of \( G = F - cF_\alpha \) for \( c \in \mathbb{C} \).

b. Does there exist \( c \neq 0 \) such that for all \( \alpha \in (0,\Omega) \),
\[
\|F\|_{L^2(T_{2\Omega})} = \|G\|_{L^2(T_{2\Omega})}?
\]

c. Let \( k \in \mathbb{N} \) and let \( \alpha = \Omega/k \). For which values of \( n \) can we assert that \( \hat{G}[n] = \hat{F}[n] = 0 \).

The point of part a is that even a small frequency perturbation of a signal causes almost all of the "temporal" data on \( \mathbb{Z} \) to change. This is precisely the uncertainty principle phenomenon discussed in Remark 1.1.4. This lack of time and frequency localization in Fourier analysis can be circumvented to some extent in wavelet theory, e.g., Daubechies' book [Dau92].

3.15. Prove Theorem 3.3.7. The right side of (3.3.9) in Theorem 3.3.7 is defined for \( G \in L^\infty(T_{2\Omega}) \). Why does Theorem 3.3.7 generally fail in this case?

3.16. Prove Theorem 3.4.4b. [Hint. Let \( \epsilon > 0 \) and let \( G \in C(T_{2\Omega}) \) have the property that \( \|F - G\|_{L^1(T_{2\Omega})} < \epsilon/(2C) \), where \( \|K_\lambda\|_{L^1(T_{2\Omega})} \leq C \), \( C \geq 1 \). Then, by (3.4.1), you can show that
\[
\|F - F * K_\lambda\|_{L^1(T_{2\Omega})} < \frac{\epsilon}{2} + \|G - G * K_\lambda\|_{L^\infty(T_{2\Omega})} + C\|G - F\|_{L^1(T_{2\Omega})}.
\]

The result follows from Theorem 3.4.4a and by taking a \( \lim \).]

3.17. In this excercise we shall approximate Fourier transforms on \( \mathbb{R} \) by DFTs (Definition 3.8.1). Consider intervals \([a,b] \subseteq \mathbb{R} \) and \([\alpha,\beta] \subseteq \mathbb{R} \), and denote their lengths by \( T = b - a \) and \( 2\Omega = \beta - \alpha \). Suppose \( 2T\Omega = N \) is a positive integer, and define
\[
t_m = a + m\Delta t \quad \text{and} \quad \gamma_n = \alpha + n\Delta \gamma,
\]
where $\Delta t \equiv 1/(2\Omega)$, $\Delta \gamma \equiv 1/T$, and $m, n = 0, 1, \ldots, N - 1$. Let $f$ be a continuous function on $[a, b]$, set $g[m] \equiv f(t_m)e^{-2\pi im\alpha/(2\Omega)}$, and let $G$ be the DFT of $g$. Show that

$$\forall n = 0, 1, \ldots, N - 1, \quad \int_a^b f(t)e^{-2\pi i\gamma t} \, dt \approx \frac{1}{2\Omega} e^{-2\pi i\gamma n} G[n].$$

[Hint. Use a Riemann sum approximation of the left side.]

We can effectively use this exercise to compute Fourier transforms, e.g., Exercises 3.18 and 3.19. The symbol "≈" means that the right side approximates the left side. This imprecise but meaningful statement can be quantified in terms of the Poisson Summation Formula (Theorem 3.10.8), e.g., [AG89], [BSS88], [BrHe95].

3.18. We computed the Fourier transform of the Gaussian $g(t) = \frac{1}{\sqrt{\pi}} e^{-t^2}$ in Example 1.3.3.

a. Using the MATLAB $\text{fft}$ function and the approach in Exercise 3.17, verify numerically that $\hat{g}(\gamma) = e^{-(\pi\gamma)^2}$. [Hint. To begin, let $[a, b] = [-32, 32], [\alpha, \beta] = [-32, 32], \Delta t = 1/64$, and $\Delta \gamma = 1/64$, and consider the vector $t = -32 : (1/64): 32 - (1/64)$.]

b. With regard to PSF (Theorem 3.10.8), and in MATLAB terminology, compare sum $(f.*f)$ and sum $(fhat.*fhat)$, where $f = \exp(-t.*t)$ and $fhat = sqrt(pi)\exp(-(pi*t).^2)$.

3.19. a. Using the MATLAB $\text{fft}$ function and the approach in Exercise 3.17, graph the following Fourier transform pairs:

$$d_\lambda \leftrightarrow 1_{[-\frac{\lambda}{2\pi}, \frac{\lambda}{2\pi}]}.$$

$$w_\lambda \leftrightarrow \max(1 - \frac{|2\pi\lambda|}{\lambda}, 0),$$

$$p_\lambda \leftrightarrow e^{-2|\lambda|/\lambda},$$

$$g_\lambda \leftrightarrow e^{-(\pi\gamma/\lambda)^2},$$
where \( \{d_\lambda\}, \{w_\lambda\}, \{p_\lambda\}, \) and \( \{g_\lambda\} \) are the Dirichlet, Fejér, Poisson, and Gauss kernels, respectively.

b. Graph the de la Vallée-Poussin kernel \( \{v_\lambda\} \) which was defined in Exercise 1.43, cf., Exercise 1.50 and Example 3.5.3b.

c. For each of the pairs in part \( a, \) as well as the de la Vallée-
Poussin kernel, compare the behavior of the function with its transform as \( \lambda > 0 \) increases. Compare the manner and speed with which each transform converges to 1 as \( \lambda > 0 \) increases.

3.20. The following is a recursive MATLAB program (m-file) which calculates the two step FFT algorithm of Theorem 3.9.1:

```matlab
function y = myfft(x)
    n = length(x);
    if n == 1
        y = 1;
    else
        xo = x(2:2:n); xe = x(1:2:n-1);
        w = exp((-2*pi*i/n)*(0:n-1));
        xohat = myfft(xo);
        y = [xe xehat] + w*[xohat xohat];
    end
```

a. Make an ascii file called `myfft.m` in some directory, say `c:\mymfiles`. Give the MATLAB command

```matlab
path('c:\mymfiles',path);
```

b. Make a vector \( f \) of random numbers, e.g., set \( f \equiv \text{rand}(1,256) \).

Compute the DFT of \( f \) by means of `myfft` as well as by the MATLAB `fft` function. Compare the results.

3.21. Let \( f, g \in l^1(\mathbb{Z}) \) have "compact support", i.e., there are integers \( A, B, C, D \) for which \( f[n] = 0 \) except for \( A \leq n \leq B \) and \( g[n] = 0 \) except for \( C \leq n \leq D \). Recall the definition of \( f * g \) in Definition 3.5.1. Part \( a \) is a fast convolution algorithm.
EXERCISES

3.22. The z-transform $F$ of $f : \mathbb{Z} \rightarrow \mathbb{C}$ is the function $F(f) : \mathbb{C} \rightarrow \mathbb{C}$ defined by

$$F(f)(z) = \sum f[n]z^{-n}.$$ 

Of course, $F(f)$ may not exist for some $z \in \mathbb{C}$. Note that there is an intrinsic relationship between Fourier series and z-transforms; in fact, we have $\hat{f}(\gamma) = F(f)(e^{2\pi i\gamma})$, $\gamma \in \mathbb{R}$.

Let $a, \omega \in \mathbb{C}$ and define $z_k = a\omega^k$ for $C \leq k \leq D$. Consider the sequences $f$, $g$, and $h$, where $f[n] = 0$ except for $A \leq n \leq B$, $h[n] = 0$ except for $C - B \leq n \leq D - A$ where it is defined as $h[n] = \omega^{n^2/2}$, and $g[n] = 0$ except for $A \leq n \leq B$ where it is defined as $g[n] = f[n]a^{-n}h[n]$. Prove that the z-transform $F$ of $f$ can be written as

$$F(f)(z_k) = \overline{h[k]}(g * h)[k].$$

This is the chirp z-algorithm.

3.23. a. Using MATLAB, implement the chirp z-algorithm of Exercise 3.22 to approximate $\int e^{-it}e^{-2\pi i\gamma} dt$, cf., Example 2.10.5. [Hint. Approximate the Fourier transform at the points $\gamma \equiv \gamma_n \equiv n\Delta \gamma$, where $\Delta \gamma \equiv .01$ and $-100 \leq n \leq 100$. Use the values $f[m] \equiv e^{-(tm)^2}$ of the Gaussian, where $t_m \equiv m\Delta t$, $\Delta t \equiv .01$, and $-300 \leq m \leq 300$. Show that the desired integral is approximately

$$F(f)(z_k)\Delta t,$$
where \( f[m] = 0 \) except for \(-300 \leq m \leq 300\), and where \( z_k \equiv (e^{-2\pi i \Delta t \Delta \gamma})^k \). Finally, invoke the fast convolution algorithm (with 1024 point FFTs) from Exercise 3.21 to perform the convolution used in the chirp \( z \)-algorithm of Exercise 3.22.

b. Compare the number of multiplications required for the chirp \( z \)-transform in part a with the number of multiplications required in the computation of the same integral in Exercise 3.18. Note that the chirp \( z \)-transform calculation provides points \( \gamma_n \) on a finer mesh than the FFT approach of Exercise 3.18, and requires only 65% of the number of multiplications.

3.24. Evaluate
\[
\sum_{n \geq 1, \text{odd}} \frac{1}{n^2}.
\]

[Hint. Consider Exercise 3.16.]

3.25. a. Let \( F \in L^1(\mathbb{T}) \) and \( G \in L^\infty(\mathbb{T}) \), and let \( f = \{f[n]\} \) and \( g = \{g[n]\} \) be their sequences of Fourier coefficients. Prove that
\[
\lim_{n \to \infty} \int_{\mathbb{T}} F(\gamma)G(n\gamma) \, d\gamma = f[0]g[0].
\]

The Riemann-Lebesgue Lemma (Theorem 3.1.5) is the special case of this result for \( G(\gamma) = e^{2\pi i \gamma} \).

b. Let \( F \in L^1(\mathbb{T}) \) be real valued. Using the notation from Example 3.1.4 and the result from part a, prove that
\[
\lim_{k \to \infty} \sum_{n=1}^{\infty} \frac{b_k n}{n} = \lim_{k \to \infty} \sum_{n=1}^{\infty} (-1)^{n} \frac{a_k(2n+1)}{2n + 1} = 0,
\]
cf., [Lux62].

3.26. a. Let \( F \in AC(T_{2n}) \), and let \( f = \{f[n]\} \) be the sequence of Fourier coefficients of \( F \). Prove that \( F' \in L^1(T_{2n}) \), \((F')'[n] = -\frac{2\pi i}{n} f[n]\) for each \( n \in \mathbb{Z} \), and \( \lim_{|n| \to \infty} nf[n] = 0 \).
EXERCISES

Thus, if $F \in C(\mathbb{T}_{2\Omega}) \cap BV(\mathbb{T}_{2\Omega})$ has Fourier coefficients $f[n], n \in \mathbb{Z}$, and if

$$\lim_{|n| \to \infty} |nf[n]| > 0$$

then $F \not\in AC(\mathbb{T}_{2\Omega})$.

b. Prove (3.5.3). [Hint. For $F \in AC(\mathbb{T}_{2\Omega})$, let $G = F' \in L^1(\mathbb{T}_{2\Omega})$ so that $G'[0] = 0$ and $F(\gamma) = \int_0^\gamma G(\lambda) d\lambda + F(0)$ on $[0, 2\Omega]$. Apply part a, Hölder’s Inequality, and Parseval’s formula to the right side of (3.5.2).]

c. Let $G \in L^1(\mathbb{T}_{2\Omega})$ and suppose $G'[0] = 0$. Prove

$$\int_0^{2\Omega} \int_0^\gamma G(\lambda) d\lambda d\gamma = -\int_0^{2\Omega} \gamma G(\gamma) d\gamma.$$ 

d. Let $F \in BV(\mathbb{T}_{2\Omega})$, and let $f = \{f[n]\}$ be the sequence of Fourier coefficients of $F$. Prove that

$$\exists M > 0 \text{ such that } \forall n \in \mathbb{Z}, \ |nf[n]| \leq M.$$ 

$M$ can be taken as the variation of $F$, i.e., $M = \inf\{\sum |F(\gamma_j) - (\gamma_{j-1})|\}$, where the infimum is taken over all finite partitions $\{\gamma_j\}$ of $[0, 2\Omega]$, cf., the definition of bounded variation in Definition 1.1.5.

Part d can be proved by a calculation for step functions and an approximation, but we also recommend an ingenious calculation due to M. Taibleson (Fourier coefficients of functions of bounded variation, Proc. Amer. Math. Soc. 18(1967), 766), e.g., [Ben76, page 120].

3.27. Let $S(\gamma) = 1/(1 + \sin 2\pi \gamma)$. Then $1/S \geq 0$ on $\mathbb{T}$ has a zero at $\gamma = -\frac{1}{4} + n$. Prove that $S \not\in L^1(\mathbb{T})$.

3.28. Let $P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a polynomial of degree $n \geq 1$ with complex coefficients. By the Fundamental Theorem of Algebra there is $\alpha \in \mathbb{C}$ such that $P(\alpha) = 0$. The following facts are well-known.
CHAPTER 3. FOURIER SERIES

i. \( P(x) \) has the unique representation

\[ P(x) = (x - \alpha_1)(x - \alpha_2)\ldots(x - \alpha_n), \]

where \( \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{C} \) is the set of zeros of \( P \);

ii. the zeros and coefficients of \( P \) are related by the formulas,

\[
\begin{align*}
  a_{n-1} &= -\sum_{j=1}^{n} \alpha_j \\
  a_{n-2} &= \sum_{i<j} \alpha_i \alpha_j \\
  a_{n-3} &= -\sum_{i<j<k} \alpha_i \alpha_j \alpha_k \\
  &\vdots \\
  a_0 &= \pm \alpha_1 \alpha_2 \ldots \alpha_n.
\end{align*}
\]

Fill in the details of the following brilliant persuasive rationale by Euler to establish that \( \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \). The zeros of

\[ P(x) = \frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \ldots \]

are \( \{n\pi : n \in \mathbb{Z}\setminus\{0\}\} \). Arguing by analogy, Euler obtains

\[ \frac{\sin x}{x} = \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{(n\pi)^2}\right), \]

cf., Exercise 1.36; and, hence,

\[ \frac{1}{3!} = \sum_{n=1}^{\infty} \frac{1}{(n\pi)^2}, \]

by parts i and ii.

3.29. a. Prove that

\[ \lim_{N \to \infty} \sum_{0 < |n| \leq N} \frac{e^{-in\gamma/\Omega}}{n} \]

exists for each \( \gamma \in (0, 2\Omega) \). [Hint. See Example 3.3.6a.]
EXERCISES

b. Let \( f[n] \geq f[n+1] \), for \( n \geq 1 \), and let \( \lim_{n \to \infty} f[n] = 0 \). Prove that

\[
\sum_{n=1}^{\infty} f[k] e^{-\pi in \gamma / \Omega}
\]

exists for each \( \gamma \in (0, 2\Omega) \), and that the convergence is uniform on any closed subset of \((0, 2\Omega)\).

c. Let \( f[n] \geq f[n+1] \), for \( n \geq 1 \), and assume \( nf[n] \leq C \). Prove that

\[
\sup \left| \sum_{n=1}^{N} f[n] \sin \frac{\pi n \gamma}{\Omega} \right| < \infty,
\]

and that the series converges to a continuous function on \((0, 2\Omega)\), cf., Exercise 3.43.

The relation between this exercise and pseudo-measures supported by arithmetic progressions is found in [Ben71, Section 4.3].

3.30. a. Prove Abel’s original “Abelian theorem”: if \( \sum_{n=1}^{\infty} a_n x^n \) converges on \([0, 1)\) and \( \sum_{n=1}^{\infty} a_n = S \), then

\[
(E3.2) \quad \lim_{x \to 1^-} \sum_{n=1}^{\infty} a_n x^n = S.
\]

[Hint. Use (E3.1).]

The first Tauberian Theorem, by Tauber in 1897, came as a response to the problem of finding some sort of converse to Abel’s result. Tauber proved that if \( f(x) = \sum_{n=1}^{\infty} a_n x^n \) converges on \([0, 1)\), \( f(x-) \equiv S \), and \( \lim_{n \to \infty} na_n = 0 \), then \( \sum_{n=1}^{\infty} a_n = S \). The boundedness condition \( \lim_{n \to \infty} na_n = 0 \) is the “Tauberian condition” required to effect the converse. This is a special case of the Tauberian condition \( \varphi \in L^\infty(\mathbb{R}) \) used in Wiener’s Tauberian Theorem (Theorem 2.9.12), e.g., [Ben75, Section 2.3].

3.31. Prove that

\[
(E3.3) \quad \forall \gamma \in (0, 2\pi), \quad -\log\left(\frac{2}{\sin \frac{\gamma}{2}}\right) = \sum_{n=1}^{\infty} \frac{\cos \frac{\pi n \gamma}{n}}{n}.
\]
[Hint. By a power series expansion we have]

\[ \text{Re} \log \left( \frac{1}{1 - z} \right) = \sum_{n=1}^{\infty} \frac{r^n \cos n\gamma}{n}, \]

where \( z = re^{i\gamma} \) and \( r \in [0, 1) \). Compute

\[ \lim_{r \to 1} \sum_{n=1}^{\infty} \frac{r^n \cos n\gamma}{n} = -\log(2 \sin \frac{\gamma}{2}) \]

on \((0, 2\pi)\). Since the right side of (E3.3) exists on \((0, 2\pi)\) by Exercise 3.29b, we can apply Exercise 3.30 to obtain (E3.3).

3.32. Let \( \Delta(ABC) \) be a triangle with vertices \( A, B, C \); and let \( a \) and \( c \) be points on the segments \( BC \) and \( BA \), respectively, so that

\[ <cCB> = <cCA> \quad \text{and} \quad <aAB> = <aAC>. \]

Assume \(|Aa| = |Cc|\). Prove that \( \Delta(ABC) \) is an isosceles triangle. [Hint. Use the law of sines.] This is a difficult exercise.

3.33. Prove that \( \sum_{n=1}^{\infty} \cos(2\pi n\gamma) \) diverges for all \( \gamma \in [0, 1) \), and that \( \sum_{n=1}^{\infty} \sin(2\pi n\gamma) \) diverges for all \( \gamma \in (0, 1) \).

3.34. Analogous to the definition of the Hilbert transform in Definition 2.5.11, we define the Hilbert transform of the sequence \( f = \{f[n]\} \) as the sequence \( \mathcal{H}f \), where

\[ (\mathcal{H}f)[n] = \frac{1}{\pi} \sum_{m \neq n} \frac{f[m]}{n - m}. \]

Prove that if \( f \in \ell^2(\mathbb{Z}) \setminus \{0\} \) is real valued then

\[ ||\mathcal{H}f||_{\ell^2(\mathbb{Z})} < ||f||_{\ell^2(\mathbb{Z})}. \]

This is the analogue on \( \mathbb{Z} \) of Theorem 2.5.12a.

[Hint. First show that if \( m, n \) are fixed and unequal and if \( j \neq m, n \), then

\[ \sum_{j} \frac{1}{(j - n)(j - m)} = \frac{2}{(m - n)^2}. \]
Use this fact to compute
\[(E3.4)\]
\[
\|Hf\|_2^2(z) = \sum_n f[n]^2 \sum_{j \neq n} \frac{1}{(j-n)^2} + \sum_n \sum_{m \neq n} f[n]f[m] \frac{2}{(m-n)^2}
\]
for \(f\) vanishing off some finite set. Since \(2f[n]f[m] \leq f[n]^2 + f[m]^2\), we can invoke (3.3.8) to bound the right side of (E3.4) by \(\pi^2 \sum f[n]^2\). This clever, elementary proof is due to [Graf94], cf., [HLP52, pages 206ff., 212, 226, 235] for other proofs, [OS75, Chapter 7] for signal processing applications, and [IVi85, pages 129 ff.] for applications in analytic number theory.

3.35. Let \(\{z_1, \ldots, z_n\} \subseteq \mathbb{C}\). Prove there is \(S \subseteq \{1, \ldots, n\}\) such that
\[(E3.5)\]
\[
\sum_{j=1}^n |z_j| \leq 4\sqrt{2} \sum_{j \in S} |z_j|.
\]
[Hint. Begin by dividing \(\mathbb{C}\) into four “diagonal” quadrants, e.g., [Ben76, page 217]. In particular, if \(F(\gamma) = \sum f[n]e^{-\pi i n \gamma / N}\), where \(f = \{f[n]\}\) is a finite sequence and \(\{n_s\} \subseteq \mathbb{R}\), then
\[
\|F\|_{L^\infty(\mathbb{R})} \leq \|f\|_{l^2(\mathbb{Z})} \leq 4\sqrt{2} \inf_{\gamma \in \mathbb{R}} |\sum f[n]e^{-\pi i n \gamma / N}|,
\]
for some finite sequence \(S \subseteq \mathbb{Z}\), cf., Definition 3.1.8.

Inequalities such as (E3.5) are used in measure theory and to prove versions of Schur’s Lemma, e.g., [Ben76, Section 6.2]. The constant \(4\sqrt{2}\) can be replaced by \(\pi\), which is best possible, e.g., [Ben76, page 172] for references to more advanced material.

3.36. Assume \(\sum |c_n|^2 < \infty\) and
\[
\lim_{N \to \infty} \sum_{|n| \leq N} c_n e^{-2\pi i n \gamma} = 0 \text{ a.e.}
\]
Prove that \(c_n = 0\) for all \(n\), cf., the discussion in Section 3.2.6.

3.37. Let \(0 \leq a < b < 1\). Prove that \([a,b]\) is not a \(U\)-set. In fact, if \(E \subseteq \mathbb{T}\) is measurable and \(|E| > 0\) then \(E\) is not a \(U\)-set, cf., the discussion in Section 3.2.4.
3.38.  

a. Prove that

\[
\left( \int_0^1 \frac{dt}{1 + t^2} \right)^2
\]

\(= \int_0^{\pi/4} \frac{\log(1 + \cos 2\theta)}{\cos 2\theta} d\theta = \frac{3}{4} \int_0^1 \frac{\log(1 + t)}{t} dt,\)

e.g., [Ebe83].

b. We know the value of the left side of (E3.6). Expanding \(t^{-1} \log(1 + t)\) in a Maclaurin series, evaluate

\[
\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2}.
\]

c. Using part b, evaluate \(\zeta(2)\) using the fact that

\[
\zeta(2) - \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2} = \frac{1}{2} \zeta(2).
\]

3.39. Prove that \(\pi\) is irrational, cf., [Ben77] for history and perspective on this material. [Hint. Assume \(\pi^2 = a/b\), where \(a, b \in \mathbb{N}\). Define \(f_n(x) = \frac{\pi^{n-1}}{n!} x^n\) on \([0, 1]\) for each \(n \geq 1\); and set

\[
F_n(x) = b^n \{ \pi^{2n} f_n(x) - \pi^{2n-2} f_n^{(2)}(x) + \pi^{2n-4} f_n^{(4)}(x) - \ldots + (-1)^j \pi^{2n-2j} f_n^{(2j)}(x) \ldots (-1)^n f_n^{(2n)}(x) \}.
\]

Compute

\[
\frac{d}{dx} (F_n(x) \sin \pi x - \pi F_n(x) \cos \pi x) = \pi^2 a^n f_n(x) \sin \pi x,
\]

and deduce that

\[
\forall n \geq 1, \quad \pi \int_0^1 f_n(x) \sin \pi x dx \in \mathbb{Z}.
\]

Obtain a contradiction for large \(n\) since \(0 < f_n(x) < 1/n!\) and \(\sin \pi x > 0\) on \((0, 1)\).]
EXERCISES

One can also show, more simply than for the case of $\pi$, that $e \notin \mathbb{Q}$. In fact, $e$ and $\pi$ are not only irrational, but are transcendental, i.e., they are not zeros of any polynomial $P$ having rational coefficients. It is not known whether or not $\pi + e$ or $\pi^e$ is irrational. On the other hand, $e^\pi$ is transcendental.

3.40. A sequence \( \{r_n : n \in \mathbb{N}\} \subseteq \mathbb{R} \) is equidistributed modulo 1 or, equivalently, uniformly distributed modulo 1 if the sequence \( \{(r_n)\} \subseteq [0,1) \) of fractional parts of the $r_n$ are uniformly distributed in the sense that
\[
\lim_{N \to \infty} \frac{\nu(I,N)}{N} = |I|,
\]
for every interval $I \subseteq (0,1)$, where $\nu(I,N)$ is the number of elements from $\{(r_1), \ldots, (r_N)\}$ contained in $I$. (\((r)\) is defined as $r - [r]$, where $[r]$ is the largest integer less than or equal to $r$.)

The Weyl Equidistribution Theorem (Remark 1.9.2) asserts that \( \{r_n\} \) is equidistributed modulo 1 if and only if
\[
\forall n \in \mathbb{Z}\setminus\{0\}, \quad \lim_{N \to \infty} \frac{e^{2\pi inr_1} + e^{2\pi inr_2} + \cdots + e^{2\pi inr_N}}{N} = 0,
\]
e.g., [KK64], [Cha68, pages 84–90]. As a corollary, prove Kronecker’s Theorem in one-dimension: if $r \in \mathbb{R}\setminus\mathbb{Q}$ then \( \{(nr) : n \in \mathbb{N}\} \) is dense in $[0,1)$. [Hint. Use Exercise 3.13b.]

3.41. The $d$-dimensional version of Kronecker’s Theorem was stated in Section 3.2.10. Prove that it is equivalent to the following assertion. Let \( \gamma_1, \ldots, \gamma_d \subseteq \mathbb{R} \) be linearly independent over the rationals, let $\gamma_0 = 0$, and let $c_0, c_1, \ldots, c_d \in \mathbb{C}$; then
\[
\sup_{t \in \mathbb{R}} |\sum_{j=0}^{d} c_j e^{-2\pi i r_j}| = \sum_{j=0}^{d} |c_j|.
\]

3.42. Let $F \in L^1(\mathbb{T}_{2\pi})$. Prove the following results.

a. If $\lim_{\lambda \to 0} \frac{1}{|\lambda|} \|\tau_{\lambda} F - F\|_{L^1(\mathbb{T}_{2\pi})} = 0$ then $F = 0$ a.e.
CHAPTER 3. FOURIER SERIES

b. If \( \lim_{\lambda \to 0} \frac{1}{|\lambda|} \| \tau_{\lambda} F - F \|_{L^1(\mathbb{T}_2^\infty)} = 0 \) then \( F \) is a constant a.e.

c. If there exist \( c, C > 0 \) such that

\[
\forall |\lambda| \leq c, \quad \| \tau_{\lambda} F - F \|_{L^1(\mathbb{T}_2^\infty)} \leq C |\lambda|
\]

then \( F \in BV(\mathbb{T}_2^\infty) \).

d. \( \lim_{\lambda \to 0} \| \tau_{\lambda} F - F \|_{L^1(\mathbb{T}_2^\infty)} = 0. \)

Part a is a consequence of part b; part b follows from a refinement of FTC, e.g., [Ben76, pages 141-142] and Remark 1.7.9 on the Lebesgue set; part c is due to Hardy and Littlewood (1928), e.g., [Ben76, pages 124-126]; and part d is an elementary and fundamental fact from Appendix A. The converse of part c results from a straightforward calculation using the classical form of the Jordan Decomposition Theorem (Remark 1.7.4).

3.43. a. Let \( \Omega > 0 \) and let \( f \in c_0(\mathbb{Z}) \) have the property that \( \{nf[n] : n \in \mathbb{N}\} \) decreases monotonically to 0 after a certain point. Prove that \( \sum_{n=1}^{\infty} f[n] \sin \frac{\pi n x}{\Omega} \) converges uniformly on \( \hat{\mathbb{R}} \), cf., Example 3.3.4, Exercise 3.29, and [Zyg59, Volume I, pages 182-183].

b. Let \( f[n] = 1/(n \log n) \), \( n \geq 3. \) By part a,

\[
F(\gamma) = \sum_{n=3}^{\infty} \frac{1}{n \log n} \sin \left( \frac{\pi n \gamma}{\Omega} \right)
\]

converges uniformly on \( \hat{\mathbb{R}} \). Prove that \( F \in C(\mathbb{T}_2^\infty) \setminus A(\mathbb{T}_2^\infty) \), cf., Example 1.4.4 and Example 3.3.4a.

3.44. Verify the inclusions and inequalities (3.1.8)-(3.1.11). Prove that the inclusions are proper, e.g., Exercise 3.43.

3.45. Use the Uniform Boundedness Principle (Theorem B.8), in a manner similar to Example 3.4.9c, to prove that there are continuous functions whose Fourier series diverge at a point, cf., the discussion in Section 3.2.8. This proof, due to Lebesgue (1905), is short
and nonconstructive. [Hint. Define the linear functionals
\[ L_N : C(\mathbb{T}) \rightarrow \mathbb{C} \]
\[ F \mapsto S_N(F)(0). \]
Let \( F_N \in C(\mathbb{T}) \) equal \( \text{sgn} \, D_N \) except on small intervals about the discontinuities of \( \text{sgn} \, D_N \); further, construct \( F_N \) so that \( |F_N| \leq 1 \). Then
\[ \|L_N\| \geq |L_N(F_N)| = \left| \int_{\mathbb{T}} D_N(\gamma) F_N(\gamma) \, d\gamma \right|, \]
and the right side is close to \( \|D_N\|_{L^1(\mathbb{T})} \). Thus, \( \{\|L_N\|\} \) is unbounded, and the result is obtained analogous to Example 3.4.9c.]

Fejér’s construction in 1911 of continuous functions whose Fourier series diverge at a point makes implicit use of the Uniform Boundedness Principle. His proof provides divergence at much larger sets of measure 0 than single points, e.g., [Rog59, pages 75-77], [Zyg59, Volume I, Chapter VIII.1], cf., the theorem of Kahane and Katznelson quoted in Section 3.2.8. On the other hand, Fejér’s type of example has unbounded partial sums. Using Riesz products, Zygmund (1948) constructed \( F \in L^\infty(\mathbb{T}) \) for which \( \{S_N(F)\} \) is uniformly bounded and \( S(F) \) diverges on uncountable dense sets of measure 0, e.g., [Zyg59, Volume I, page 302].

3.46. If \( F \in L^1(\mathbb{T}) \), define
\[ M(F)(\gamma) = \sup_N \{|S_N(F)(\gamma)|\}. \]
The main part of Carleson’s proof of the Lusin Conjecture (Section 3.2.8) is his theorem that
\[ (E3.7) \exists C \text{ such that } \forall F \in L^2(\mathbb{T}), \|M(F)\|_{L^2(\mathbb{T})} \leq C\|F\|_{L^2(\mathbb{T})}. \]
Using (E3.7) prove the Lusin Conjecture:
\[ \forall F \in L^2(\mathbb{T}), \lim_{N \to \infty} S_N(F) = F \text{ a.e.} \]

With regard to (E3.7), compare Zygmund’s result quoted in Exercise 3.45.
3.47. Let $F \in L^1(\mathbb{T})$ and $G \in L^\infty(\mathbb{T})$ have Fourier coefficients $f = \{f[n]\} \in A(\mathbb{Z})$ and $g = \{g[n]\} \in A'(\mathbb{Z})$, respectively. Compute $(FG)^\vee$ and $(F \ast G)^\vee$ in terms of $f$ and $g$.

3.48. a. Let $F \in A(\mathbb{T})$ be nonnegative. With regard to Remark 3.6.5d on the Fejér-Riesz Theorem, verify whether or not there is a sequence $\{B_n\}$ of trigonometric polynomials on $\mathbb{T}$ for which

$$\lim_{n \to \infty} \|F - |B_n|^2\|_{A(\mathbb{T})} = 0,$$

cf., the Szegö Factorization Theorem stated in Section 3.7.

b. Let $F \in C(\mathbb{T})$, and suppose the sequence $f = \{f[n]\}$ of Fourier coefficients of $F$ is nonnegative, i.e., $f[n] \geq 0$ for each $n$. Prove that $F \in A(\mathbb{T})$.

[Hint. $f$ is a positive distribution of $\mathbb{Z}$ so that $F$ is continuous and positive definite on $\mathbb{T}$.]

There are also several classical ways of proving part b.

3.49. A fundamental problem in harmonic analysis is to quantify properties such as support, smoothness, convergence, and decay of $f$ and its approximants in terms of the behavior of $\hat{f}$; and many of our results can be put in this context, e.g., the Riemann-Lebesgue Lemma, Exercise 1.18, Exercise 3.26, and Example 3.5.3 where we proved that $C^1(\mathbb{T}) \subseteq A(\mathbb{T})$. In this regard, prove that if $1 \leq m < \infty$ then

$$\forall F \in C^m(\mathbb{T}), \quad \lim_{N \to \infty} \|F - S_N(F)\|_{L^\infty(\mathbb{T})} = 0.$$

Estimate $\|F - S_N(F)\|_{L^\infty(\mathbb{T})}$ in terms of $N$ and $m$.

3.50. Prove the assertions in Definition 3.1.9d.

3.51. a. Let $L \in \mathcal{L}(L^2(\mathbb{T}))$, i.e., $L : L^2(\mathbb{T}) \to L^2(\mathbb{T})$ is a continuous linear function (Definition B.6). If $F \in L^2(\mathbb{T})$, prove that

$$L(F)(\gamma) = \sum (L_d F^\vee)[n]e^{-2\pi in \gamma}$$
where
\[(L_d F^\gamma)[n] = \sum_m \ell_{mn} F^\gamma[m]\]
and
\[\ell_{mn} = \langle L e_{-m}, e_{-n} \rangle.\]

Thus, the action (operation) of \(L\) on \(L^2(\mathbb{T})\) is equivalent to the action of the matrix \(L_d \equiv \{\ell_{mn} : m, n \in \mathbb{Z}\}\) on \(\ell^2(\mathbb{Z})\).

b. Now, further suppose that \(L\) is translation invariant, i.e., \(L(\tau_\gamma F) = \tau_\gamma(LF)\) for all \(\gamma \in \mathbb{T}\) and all \(F \in L^2(\mathbb{T})\), Exercise 2.58. Prove that \(L_d\) is a diagonal matrix.

Parts a and b allow us to conclude that \(\{e_{-m}\}\) diagonalizes all continuous translation invariant linear functions on \(L^2(\mathbb{T})\); and it is also clear that each \(e_{-m}\) is an eigenfunction of the derivative operator. This fact accounts for the classical success of Fourier analysis in dealing with linear partial differential equations, since it provides exactly the advantage of doing matrix calculations in a basis that diagonalizes the matrix. In more recent times, parts a and b are the starting point for the harmonic analysis of singular integral operators, e.g., the Calderón-Zygmund Theory [Ste70], as well as LTI systems in engineering, e.g., Definition 2.6.5.

3.52. Let \(\{r_j : j = 0, \pm 1, \ldots, \pm N\} \subseteq \mathbb{C}\) satisfy the condition \(r_j = \bar{r}_{-j}\) for each \(j\), and define the \((N + 1) \times (N + 1)\) matrix \(R = (r_{jk})\), where \(j, k \geq 0\) and \(r_{jk} \equiv r_{j-k}\). Also, define the functional \(L\) on the trigonometric polynomials \(F(\gamma) = \sum_{n=-N}^{N} f[n] e^{-2\pi im\gamma}\) by the rule
\[L(F) = \sum_{n=-N}^{N} f[n] \bar{r}_n.\]

a. Prove that if \(R >> 0\) and \(1 \leq j \leq N\) then \(r_0 > |r_j|\).

b. Prove that for each polynomial \(G(\gamma) = \sum_{n=0}^{N} g[n] e^{-2\pi im\gamma}\),
\[L(|G|^2) = \sum r_{j-k} g[k] \bar{g}[j].\]
c. We shall say that the functional $L$ is positive, written $L \geq 0$, if $L(F) \geq 0$ and

$$L(F) = 0 \implies F = 0$$

for all trigonometric polynomials $F(\gamma) = \sum_{n=-N}^{N} f[n] e^{-2\pi i n \gamma}$. Prove that $R >> 1$ if and only if $L \geq 0$ cf., Exercise 2.52a for the analogous situation in terms of positive semidefinite matrices. [Hint. Part b is used in both directions, and the direction to prove $L \geq 0$ requires the Fejér-Riesz Theorem.]

3.53. Prove that $(\sin x)^{-2} < x^{-2} + 1$ on $(0, \pi/2]$. The proof is elementary, e.g., [Mon71, page 155–156]. This inequality should be compared with the inequality, $\frac{2}{\pi} \leq \sin x$ on $[0, \frac{\pi}{2}]$, which is usually used in the proof of Jordan’s Inequality (Exercise 2.62).

This inequality can be used to prove the following number theoretic “sieve theorem” due to Bombieri, e.g., [Mon71, Chapters 2 and 3]. Let $\delta > 0$ and $R \subseteq \mathbb{R}$ have the property that the distance of $r - s$ to any integer is greater than or equal to $\delta$ for all unequal $r, s \in R$; then

$$\delta \sum_{\gamma \in R} |F(\gamma)|^2 \leq (\delta N + \frac{2}{\sqrt{3}} + 3\delta) \sum |c_n|^2$$

for all trigonometric polynomials

$$F(\gamma) = \sum_{n=M+1}^{M+N} c_n e^{-2\pi i n \gamma}.$$ 

Because of the Parseval formula, it is interesting to note that if the elements of $R$ are (relatively) equally spaced modulo 1 then $\delta \sum_{\gamma \in R} |F(\gamma)|^2$ is a Riemann sum approximating $\|F\|_{L^2(T)}^2$. The relation of such approximations to sampling theory is the subject of [Ben92b, pages 492–494].

3.54. a. Prove that if $W \in L^1(T)$ is positive on $T$, then $\log W \in L^1(T)$. 
b. Prove that if \( W \in L^1(\mathbb{T}) \) is nonnegative, then \( \log W \in L^1(\mathbb{T}) \) if and only if
\[
\int_{\mathbb{T}} \log W(\gamma) \, d\gamma > -\infty.
\]

3.55. We discussed the Littlewood Flatness Problem in Remark 3.8.11b. In this regard, prove that if \( F(\gamma) = \sum_{n=0}^{N-1} f[n] e^{-2\pi i n \gamma} \), then \(|F| \) and \(|f| \) cannot simultaneously take constant values on \( \mathbb{T} \) and \( \{0, 1, \ldots, N-1\} \), respectively.

3.56. Consider the following formal calculation for \( f : \mathbb{R} \times \mathbb{R} \to \mathbb{C} \), where \( f(t, u) \equiv f(t-u) \).
\[
\iint f(t, u) e^{-2\pi i (\lambda + u \gamma)} \, dt \, du
\]
\[
= \int \left( \int f(u) e^{-2\pi i (u+u) \lambda} e^{-2\pi i u \gamma} \, du \right) \, du
\]
\[
= \hat{f}(\lambda) \int e^{-2\pi i u (\lambda + \gamma)} \, du = \hat{f}(\lambda) \delta(\lambda + \gamma)
\]
\[
= \begin{cases} 
\hat{f}(\lambda), & \text{if } \lambda = -\gamma, \\
0, & \text{if } \lambda \neq -\gamma.
\end{cases}
\]

(a) Recalling the definition of a Toeplitz matrix \( R = (r_{jk}) \), viz., \( r_{jk} = r_{j-k} \), prove that the two dimensional DFT of a Toeplitz matrix is a diagonal matrix, cf., Exercise 3.51.

(b) Provide the hypotheses and details to make (E3.8) into a theorem.

3.57. Let \( \mu \in M(\mathbb{T}) \) and assume \( \lim_{n \to +\infty} \mu^\vee[n] = c \in \mathbb{C} \). Prove that \( \lim_{n \to -\infty} \mu^\vee[n] = c \). [Hint. By subtracting \( c \delta \) from \( \mu \) we can assume, without loss of generality, that \( \lim_{n \to +\infty} \mu^\vee[n] = 0 \). By the Radon-Nikodym Theorem, let \( \mu = F|\mu \), where \( F \) is \( \mu \)-measurable and \( |F| \equiv 1 \), e.g., [Ben76, Chapter 5]. If \( G \equiv F/F \) then
\[
\mu^\vee[-n] = \int_{\mathbb{T}} e^{2\pi i n \gamma} G(\gamma) \, d\mu(\gamma).
\]
CHAPTER 3. FOURIER SERIES

If $G$ is a trigonometric polynomial the right side tends to 0 as $n \rightarrow \infty$, and so the result is obtained by (E3.9). The result for $G = \overline{F}/F$ follows by the Weierstrass Approximation Theorem. (Theorem 3.4.6 and Remark 3.4.8) and approximation properties of bounded $\mu$-measurable functions (such as $G$) by continuous functions, cf., [Ben76, pages 201].

3.58. Complete the proof of the Classical Sampling Theorem which uses the PSF, i.e., provide the mathematical justification for the formal steps in Proof 1 of Theorem 3.10.10.

3.59. Let $s(t) = d_{2\pi \Omega}(t) = \frac{\sin 2\pi \Omega t}{\pi t}$ and assume equation (3.10.21),

$$f = T \sum f(nT) \tau_{nT}s,$$

is true for all $f \in PW_\Omega$. Prove that $2T\Omega \leq 1$, e.g., [Ben92b, page 450].

3.60. Let $F \in L^2(\mathbb{T})$, and let $f = \{f[n]\}$ be the sequence of Fourier coefficients of $F$. Assume

$$\sum_{n=1}^{\infty} \frac{1}{n^{1/2}} \left( \sum_{|k|=n} |f[k]|^2 \right)^{1/2} < \infty.$$

Prove that $F \in A(\mathbb{T})$.

We have commented often on the difficulty of characterizing $A(\mathbb{T})$. The result of this exercise can be considered the first step in a fairly sophisticated classical line of thinking, e.g., [Ben75, pages 150–154].

3.61. We shall say that $L \in \mathcal{L}(L^2(\mathbb{T}))$ is translation invariant if

$$\forall F \in L^2(\mathbb{T}) \text{ and } \forall \gamma \in \mathbb{T}, \quad L(\tau_\gamma F) = \tau_\gamma (LF),$$

cf., Definition 2.6.5. Prove that the translation invariant continuous linear operators $L : L^2(\mathbb{T}) \rightarrow L^2(\mathbb{T})$ are precisely of the form

(E3.10)

$$\forall F \in L^2(\mathbb{T}), \quad LF = G * F,$$
where \( G \in A'(\mathbb{T}) \), i.e., \( g \equiv G^\vee \in \ell^\infty(\mathbb{Z}) \); and also show that \( \|g\|_{\ell^\infty(\mathbb{Z})} = \|L\| \), where \( \|L\| \) is defined in Definition B.6. [Hint. If \( L \) is defined by (E3.10) then it is easy to see that \( L \in \mathcal{L}(L^2(\mathbb{T})) \) and that \( L \) is translation invariant; and it is elementary to prove the norm equality. The converse is a consequence of Exercise 3.51.]

In fact, if \( L \in \mathcal{L}(L^2(\mathbb{T})) \) is translation invariant, then \( (LF)^\vee = L_d f \), where \( L_d \) is the diagonal matrix of Exercise 3.51 which has entries \( g[n] \equiv \langle L e_n, e_n \rangle \) on the diagonal. If \( g \equiv \{g[n]\} \) then \( G^\vee \equiv g \in \ell^\infty(\mathbb{Z}) \) since \( |g[n]| \leq \|L\| \), and \( LF = G * F \) for this \( G \).

In Example 2.6.6 we stated the \( L^1 \) analogue of the above result. In the \( L^1 \) setting, \( A'(\mathbb{T}) \) is replaced by \( M(\mathbb{T}) \). In \( L^1(\mathbb{T}) \), if \( L \) is also a projection, i.e., \( L \circ L = L \) on \( L^1(\mathbb{T}) \), then the corresponding measure \( \mu \in M(\mathbb{T}) \) is an idempotent measure. Idempotent measures were discussed in Remark 3.10.13.

3.62. Let \( T > 0 \) and define \( \mu = \sum e^{2\pi int/T} \) on \( \mathcal{S}(\mathbb{R}) \) by the rule

\[
\forall f \in \mathcal{S}(\mathbb{R}), \quad \mu(f) = \lim_{N \to \infty} \left( \sum_{\lvert n \rvert \leq N} e^{2\pi int/T} \right)(f(t)).
\]

Recall from Chapter 2 that \( g(f) \) denotes \( \int g(t)f(t) \, dt \).

a. Prove that \( \mu \in \mathcal{S}'(\mathbb{R}) \).

b. Prove that \( \mu \in M(\mathbb{T}) \) and

\[
T \sum \delta_n T = \sum e^{2\pi int/T}.
\]